# INTERACTIONS BETWEEN DEMOGRAPHIC RATES, POPULATION DENSITY AND THE ENVIRONMENT - THE SPATIAL STRUCTURE OF THE RANGE OF THE HOLLY LEAF-MINER

**by Andrew M. Brewer**

Thesis submitted to the University of Sheffield for the degree of Doctor of Philosophy in the Department of Animal and Plant Sciences.

NOVEMBER 2000

# Summary

Generalisations about the abundance structure of species' geographic ranges may have implications within a number of areas of applied ecology. However, empirical data is largely restricted to a single taxonomic group within one geographic region. One approach to the study of range structure and its ecological consequences is to examine the relationship between demographic rates and environmental conditions. However, most studies of population processes are at small spatial scales and it is not known to what extent patterns found at these scales may be extrapolated.

This thesis addresses these issues using surveying techniques to measure spatial structure in both the densities and demographic rates of the holly leaf-miner (*Phytomyza ilicis* Curt.) at a wide variety of spatial scales. Geostatistical tools are used to analyse the data collected.

At a regional scale, spatial structure in leaf-miner densities is apparent despite considerable variation between hosts within individual habitat patches. This structure can in part be accounted for by variation in habitat and altitude. Over the entire range, broad scale trends in population density can be detected which can also, be correlated with environmental variation.

Many demographic rates components exhibit spatial structure. However, their relationships both with population density and environmental variation are more complex. *Per capita* mortality rates did not correlate well with population density at any spatial scale. The difficulties inherent in relating population density with demographic rates and regulation by natural enemies are discussed.

The demographic approach has also been used to explain the positive interspecific abundance-distribution relationship. Computer simulation techniques are employed to explore this model. The results indicate that the positive relationship should be extremely robust under a high degree of variation between the demographic rates.

However, data from the holly leaf-miner surveys suggest that current models of range structure make unrealistic assumptions about environmental variation. The key to understanding range structure and its consequences may lie in our ability to make generalisations about environmental structure.

# Acknowledgements

This project was financially supported by the Alfred Denny Studentship at the University of Sheffield. The European holly leaf-miner survey was funded by the Natural Environment Research Council.

I am indebted to a number of people for their assistance during this project. Thanks to Melody McGeoch for help in the fieldwork and for useful discussion during the Sheffield Surveys. Many thanks to Rob and Kerry Briers, Sian Gaston, Trond Hofsvang, Jens Rolff, Phil Warren and most especially John Marçal for help and advice during the European survey. Thanks also to Bob Holt for an extremely helpful day of discussion about the vital rates model.

I am particularly grateful for the support of my supervisor Kevin Gaston, who has been a continual source of inspiration throughout the project.

*This thesis is dedicated to my parents.*

# Declaration

Two chapters in this thesis were jointly authored. The contributions I have made to these chapters are detailed below.

---

Chapter 4:

Brewer A.M. & Gaston K.J. (manuscript) The population density structure of the geographic range of the holly leaf-miner.

Chapter 5:

Brewer A.M. & Gaston K.J. (manuscript) Spatial structure in the demographic rates of the holly leaf-miner across its geographic range.

These manuscripts have been prepared for submission to *Oikos* and are based on an original idea by KJG. Data throughout Europe were collected primarily by AMB with the help of a research assistant. AMB also created the database, conducted all data analyses and drafted the manuscripts.

---

# Table of Contents

# Chapter 1

# Introduction

> *"As we cannot know the reason for a fact before we know the fact. We cannot know what a thing is before knowing that it is."*
> Aristotle

The aims of population ecology are to describe patterns in the distribution and abundance of organisms through space and time, and to explain these patterns in terms of their underlying mechanisms. Much of our understanding of population processes is based upon research conducted at small spatial scales (Brown *et al.* 1995), typically at the level of individual populations confined to a single habitat patch. It has not been established whether the information collected from these studies can be applied to larger spatial scales. This is an important issue, since many of the environmental problems we face today are related to processes operating at these broad scales. The effects of global climate change, shifts in land-use and biological invasions on the current distribution and abundance of species is likely to have important consequences for agriculture, the conservation of biodiversity and the spread of disease.

Macroecology (Brown & Maurer, 1989) is a relatively recent attempt to understand large scale ecological phenomena using the substantial database that has already been assembled on many characteristics of plant and animal populations. Where traditional ecological studies have largely focused upon the roles of individual organisms in community and population processes, macroecology aims to identify and explain patterns that become apparent when comparing species characteristics within assemblages. Typical units of study are the size and shape of the geographic range, mean population density and the mean body size of species. The approach has met with some success. A number of statistical relationships between such characteristics have been identified and some historically well-known patterns have received renewed attention and attempts at explanation (e.g. Gaston, 1994; Brown, 1995). However, it can be difficult to distinguish between competing hypotheses. Experimental work is seldom practical at these taxonomic and spatial scales and collection of the additional data required may be difficult and expensive.

For example, there has recently been a renewed attempt to make statistical generalisations about the internal abundance structure of the geographic ranges of species. If such generalisations could be made there would be obvious consequences within several areas of applied ecology such as the identification of high risk areas for pest outbreaks and the design and location of nature reserves. A number of intuitively-appealing assertions have been made about range structure and its causes (e.g. Brown, 1984; Hengeveld, 1989; Maurer & Brown, 1989) but these are derived from just a few empirical studies which have been based on narrow taxonomic and geographic domains.

One potentially useful approach to the investigation of range structure is given by Maurer & Brown (1989). They propose that it may best be understood in terms of changes across space in the four demographic rates that underlie population regulation; births, deaths, immigration and emigration. Birth and immigration sum to a net rate of population gain and death and emigration sum to a net rate of population loss. Where population gain and population loss intersect there is a stable equilibrium point. The general spatial pattern of population abundance may therefore be characterised as the set of equilibrium points of local populations across space. If the demographic rates are dependent upon local environmental conditions, then range structure should ultimately reflect the spatial structure of the environment. Maurer & Brown (1989) make a number of assumptions about the nature of environmental structure in order to create a theoretical model of the structure of the geographic range. To date, the model has not been tested empirically since there are inherent difficulties in estimating the demographic rates of populations. However, this thesis is an attempt to apply a similar approach to the structure of the range of the holly leaf-miner *Phytomyza ilicis* Curt., a species which may be considered particularly amenable to studies of this nature.

The interactions between the spatial structure of the environment, demography and population density are the main focus of this thesis. In order to address these issues use is made of a number of statistical techniques that have been developed within the geological sciences for the analysis of spatially explicit data. It is not always clear what spatial scale is appropriate for documenting and understanding ecological phenomena (Maurer, 1999 and references therein). This issue will be addressed directly by the work described in the thesis. Spatial structure in leaf-miner densities and demographic rates will be described within a single habitat patch, across regional environmental variation and finally over the whole of the geographic range.

Patterns in the variation in local population densities may be described in a number of ways incorporating varying degrees of spatial information. Below is a brief review of the techniques that have been commonly used:

### *Frequency distributions*

Frequency distributions of local densities give a graphical summary of the number of sites at which a species occurs at different levels of abundance. They incorporate no explicit spatial information but nevertheless may reveal that individuals of a species are not uniformly distributed through space. There are often differences in densities between sample sites of several orders of magnitude. Moreover the plots often show a characteristic shape. A species distribution can be characterised by a few sites of high density ("hot spots" Brown, 1995) and many sites of comparatively low density ("cool spots"). In addition there will often be a large number of sites where the species does not occur at all ("cold spots"). Brown *et al.* (1995) analysed data sets from the American Breeding Bird Survey (BBS) and found that there are similar frequency distributions for many bird species. Brown (1995) uses unpublished data to show that this pattern holds for a variety of taxonomic groups and spatial scales.

Attempts have been made to fit relatively simple statistical distributions to these plots. If Poisson distributions were adequate to describe them, it would indicate that individuals are randomly distributed between sample sites. However, at all but the smallest of spatial scales, individuals tend to deviate from the Poisson distribution showing some degree of aggregation (Taylor *et al.*, 1978). The negative binomial is commonly used to describe this clumped distribution for parasites and for pest insects among hosts, although there is some debate on the best statistical distribution to use and several other alternatives have been offered (e.g. Perry & Taylor, 1986). There is an empirical relationship between frequency of occurrence and population density for the negative binomial distribution. This has allowed estimates of overall density to be assessed by counting the number of samples where at least one individual is present. In agriculture this method has been used for rapid estimation of crop pest infestation levels to determine when control methods should be used. Its application in pest control has meant that most work has been done at fine spatial scales (e.g. individual leaves or plants within a crop). Little work has been done at coarser resolutions where it could be useful in estimating the density of larger organisms for conservation purposes (Gaston, 1994). Frequency distributions for *P. ilicis* are presented in this thesis at two spatial scales, both regional and across the entire geographic range.

### *Spatial autocorrelation analysis*

A more spatially-explicit method of describing variation in local abundance is to incorporate distances between sample sites into the analysis. The separation distance between sites is often referred to as the lag. Spatial autocorrelation analysis (Cliff & Ord, 1973) can be used to identify correlations between densities within different distance intervals and these correlations can then be plotted as an autocorrelation profile or correlogram (figure 1.1).

Patterns of spatial autocorrelation have been investigated at several scales. For example, Thomson *et al.* (1996), in a relatively small scale study of the factors that affect the distribution of glacier lily (*Erythronium grandiflorum*), show how correlograms produced from spatial autocorrelation analysis can reveal the dimensions of patches of mature plants. They also use this technique to show that abundances and a variety of other physical and biotic factors are significantly correlated.

Correlograms have also been used to reveal patterns of abundance over entire geographic ranges. Brown *et al.* (1995) show that there are distinctive patterns of spatial variation within the ranges of American bird species (figure 1.1). There are often two peaks of positive values. At short lag distances there tends to be quite high positive correlation between abundances. This is despite the large amount of variation shown at small spatial scales revealed by frequency abundance plots. This pattern shows that sites that are close together tend to support more similar densities than those that are widely separated, regardless of their position within the geographic range. A second, high correlation is found at the maximum lag distance revealing that sites at opposite edges of a range tend to have similar densities.

Correlograms are used widely throughout the research reported in this thesis and are the main statistical tool used to detect spatial structure in local population densities and demographic rates. The correlograms for *P. ilicis* densities across the whole geographic range, the first data set at such a scale that is not for North American Birds, will be compared directly to the findings of Brown *et al.* (1995).

### *Environmental Gradients*

Relating changes in abundance to environmental variables can be used to understand some of the spatial variation in abundance of a species and it can provide the first step in understanding the reasons for the size and shape of its geographic range (Hoffmann & Blows, 1994). Gradients do not necessarily have a physical reality as continua in space, they may be expressed as a complex mosaic of different habitat types.

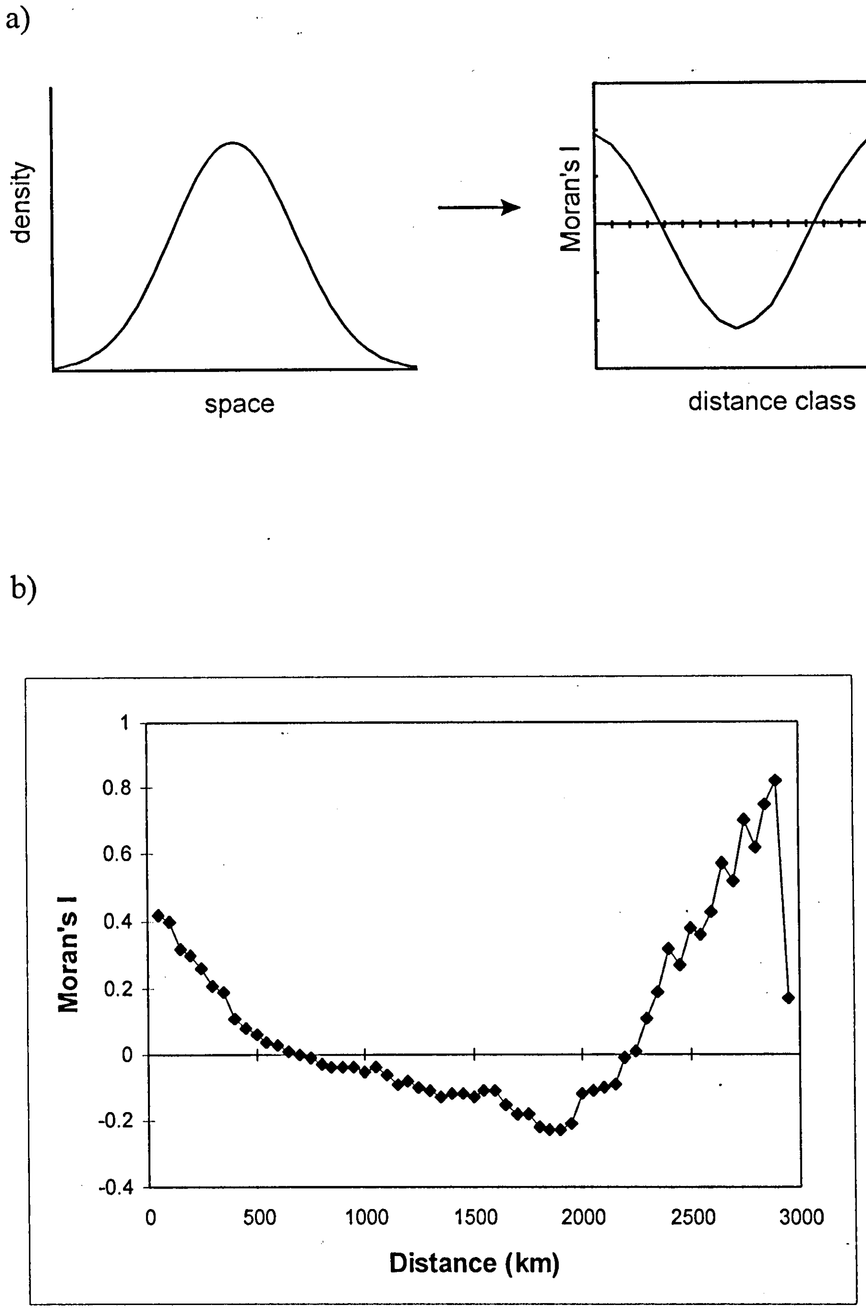


Figure 1.1 The use of spatial autocorrelation analysis to detect structure in local population densities. (a) A hypothetical species with a Gaussian distribution of densities along a transect and the resulting correlogram. (b) Correlogram of carolina wren (*Thryothorus ludovicianus)* abundance across its geographic range (Brown *et al*,. 1995).

However, considering them as a continuous gradient for the purpose of analysis is a useful abstraction (Austin, 1985).

There has been a great deal of work done in the field of gradient analysis which attempts to identify changes in community composition along environmental gradients (Gauch, 1982 and Ter Braak & Prentice, 1988 provide references). Direct gradient analysis uses regression techniques to describe each species abundance as a function of measured environmental variables. Abundance tends to be highest around a particular environmental optimum and decrease from this point. Over a short interval of an environmental gradient there may be an approximately linear response of a species' abundance. However if a large range of environmental variation is measured, the species' response may be non-linear. Often a Gaussian curve will be more appropriate (Gauch & Whittaker, 1972) where the logarithm of species abundance is a quadratic in the environmental variable (figure 1.2). This can be extended to the multivariate form if several environmental variables are measured. Ter Braak & Prentice (1988) review the statistical techniques behind gradient analysis.

### *Mapping and interpolation*

A maximum amount of spatial information is incorporated when the local abundance of a species is mapped across all or part of its geographic range. Unfortunately, there are only a limited number of studies, mostly of birds, which are of sufficient geographic scope to be used to ascertain the generality of any patterns (see Gaston 1994 for references). Root (1988) has mapped winter abundances of over 200 species of North American birds. In addition the abundances of many species of aphids in Britain have been monitored by a network of suction traps since the early 1970s as part of the Rothamsted Insect Survey. Maps of some of the more common species have been published (e.g. Taylor & Taylor, 1979).

Not only is there rather a paucity of data available but also there are currently limits in the methodology for interpreting any patterns that maps may reveal. Until recently, most interpretation has been by visual means which, while useful, limits the extent to which generalisations about patterns can be made between species (Gaston, 1994). In addition, data collected for maps tend to be from discrete locations and there are a number of different methods of interpolating between these points to produce the final map (Isaaks & Srivistava, 1989). Each technique has its own strengths and weaknesses. However, use of geostatistical statistical methods in conjunction with maps may offer a

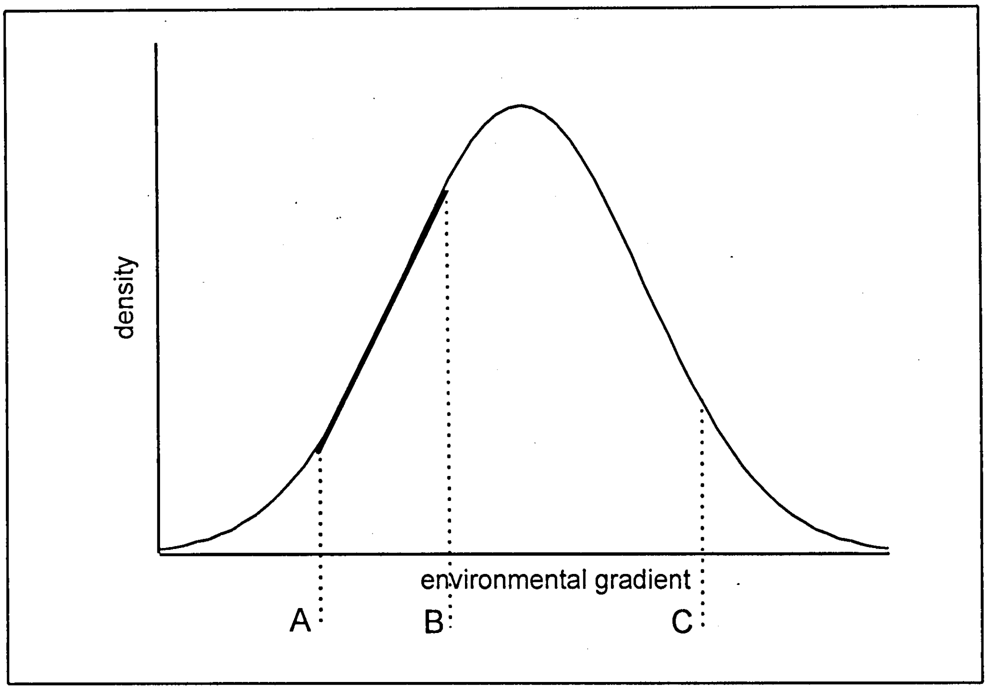


Figure 1.2 Across relatively short intervals, modelling population density as a linear response to environmental variation may be a good approximation (e.g. between A and B). However over a larger range of environmental variation, a Gaussian model may be more appropriate (e.g. between A and C).

more objective method of interpretation. This approach is used within the thesis.

Because of these limitations only rather crude generalisations about spatial patterns in abundance have been made to date. A number of researchers maintain that at a large spatial scale, abundance tends to be greatest near the centre of a species' geographic range and tends to gradually decline toward the boundaries (Chapter 4). Brown (1984) identified two ways that this can occur. First, a species may occupy a smaller proportion of local patches towards the edge of the range, that is its area of occupancy declines (Gaston, 1994). Secondly, average population density within a patch may be less at the edge of the range.

It has been further suggested (e.g. Brown, 1984; Maurer, 1999) that the distribution of abundances over an entire range can often be regarded as a surface approximating a bivariate normal distribution, with a single, central peak in abundance and declining monotonically towards the range limits. However, the fit of data to this model is seldom tested (Gaston, 1994). Moreover, there are obviously many exceptions to such a generalisation (Maurer, 1999). In practice, spatial 'textures of abundance' (Lawton, 1993) may be far more complex. Multi-modal patterns are very common over a geographic range as the maps provided by Root (1988), for example, show.

At a smaller spatial scale, local maps of abundance can reveal interesting patterns, not least the enormous amount of spatial heterogeneity in the abundance of a species in adjacent sites. The figures provided in Thomson *et al.* (1996) of the abundance of glacier lily adults and immature plants show obvious aggregation of plants which is reflected by the underlying substrate and the mode of seed dispersal. Spatial autocorrelation analysis was used to provide a measure of patch size (see above) but from the maps patch shape can also be appreciated.

Patterns of abundance mapped at coarse resolutions such as entire geographic ranges are necessarily a very simplified picture of patterns that occur at regional and local spatial scales. A consequence of this is that patterns detected at one spatial scale may not apply to others. This is an important theme throughout the thesis.

There are a number of studies that have investigated variation in components of demographic rates between populations and many of these will be referred to in the subsequent chapters. Very few have attempted to use this information to provide an explicit picture of how demographic rates, population density and environmental structure might interact to create the patterns in population densities that would be observed if they were plotted on a map. However, there has been some interest in comparing population dynamics between central and marginal populations.

For example, Whittkaker (1971), in response to what he saw as the polarised views of theorists on the mechanisms that regulated animal populations, investigated the population dynamics of a cercopid bug (*Neophilaenus lineatus*). At a lowland site, where environmental conditions were considered favourable for the bug, populations were relatively stable from year to year and largely governed by weather conditions at the end of the season which determined the number of egg batches that could be laid. Whittaker found that the population suffered density dependent mortality, largely as a result of a dipteran parasitoid (*Verallia aucta*). He concluded that at this site, well within its range, *N. lineatus* populations were closely regulated. In contrast, harsh environmental conditions at an upland site caused large oscillations in populations between years, occasionally resulting in temporary local extinctions when weather conditions did not permit any eggs to be laid. He found no evidence of density dependent mortality and concluded that fringe populations were not well regulated.

Similarly, Randall (1982) showed that mortality and birth rates of a case moth, *Coleophora alticolella*, feeding on the seed capsules of *Juncus squarrosus*, changed systematically with altitude. At high altitudes, populations suffered density independent mortality from harsh climatic conditions which limited the amount of food available for the larvae to feed upon. Occasionally local extinctions could occur in these populations when conditions were bad enough to prevent seed production. At low altitudes where there was an excess of food available for the moth larvae, parasitism was an important mortality factor and thought to hold populations well below the carrying capacity. Five species of hymenopteran parasitoids were recorded attacking the larvae. The intensity of parasitism and the number of species attacking the larvae decreased markedly with altitude. At intermediate altitudes, where populations were at their highest, populations were regulated by density dependent overwintering mortality and reduction in birth rates. This was largely as a result of intraspecific competition for food. The factors regulating *C. alticolella* are summarised in figure 1.3. This figure highlights the changing intensity of the two external mortality factors with altitude, leaving a middle region where their combined effect is low.

Caughley *et al.* (1988) suggested a theoretical framework for identifying a variety of causative, environmental factors limiting distribution by comparing three attributes of central and marginal populations. These are density, the intrinsic rate of increase ($r_m$)

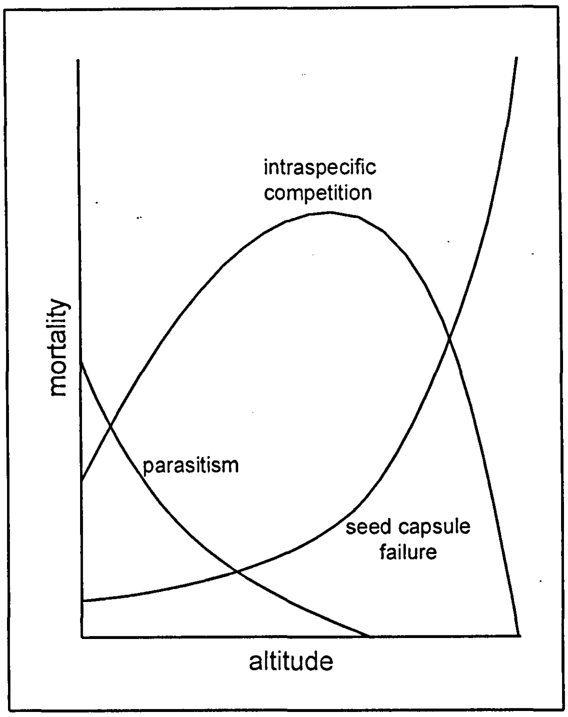


Figure 1.3 Diagrammatic representation of the changing intensity of mortality factors of *C. alticolella* along an altitudinal gradient (After Randall, 1982).

| **Factor affecting range boundary** | **$r_m$** | **Density** | **Condition** |
|---|---|---|---|
| Climate | ramp | ramp | ramp |
| Substrate | step | step | step |
| Unmodifiable resource | ramp | ramp | ramp |
| Resource affected by organism | step | ramp | step |
| Facultative predator, parasite or pathogen | ramp | ramp | ramp |
| Obligate predator, parasite or pathogen | - | - | - |

Table 1.1. Changes in population characteristics from central to marginal populations that may allow the identification of any single factor that limits distribution. (Adapted from Caughley *et al.*, 1988).

and the condition of individuals. The population characteristics can either change suddenly (a step) or more progressively (a ramp), reflecting the change in the influence of a limiting factor. Their approach is summarised in Table 1.1. Two species of kangaroo (*Macropus* spp.) were used to illustrate the method. Due to the difficulty of measuring $r_m$, only density and condition were compared. *Macropus fuliginosus* showed a progressive decrease (ramp) in both attributes toward the margin suggesting a climatic factor or unmodifiable resource was involved in limiting distribution, since there was no evidence of significant predation, parasitism or disease. However, *Macropus giganteus* populations showed a ramp in density and a step in the condition of individuals suggesting a renewable resource was involved.

In a study to understand the distributions of three species of tsetse flies (*Glossina* spp.) in Africa, Rogers and Randolph (1986) used local data on the factors influencing abundance to extrapolate to predictions about regional distributions. Mortality rates for each species were calculated from long-term data sets obtained from fly-round catches (a method of assessing population abundance by sampling flies along set routes). Monthly mortality figures were plotted against temperature and a measure of relative humidity to produce mortality climograms. A reproductive plane was added to these to identify meteorological conditions where reproduction exceeded mortality. This defined the bioclimatic limits within which any species' geographic distribution was expected to lie. In addition, the bioclimatic optimum was identified, where fertility minus mortality was at a maximum. These analyses were used to predict the abundance of tsetse within its distributional limits and so identify areas where disease transmission was high. The three species showed different responses to environmental conditions and this information was used to describe the differences in their distributions with some success. Several areas of Africa with known problems of trypanosomiasis corresponded well with the bioclimatic optimum of one species, *G. morsitans*.

To summarise, there appear to be consistent patterns in density variation at a number of spatial scales for a variety of species. In addition, it appears that these patterns may be usefully interpreted in terms of underlying demographic rates. However, there is an obvious need to apply this approach systematically to species within a spatially explicit framework across a broad range of spatial scales. This thesis aims specifically to answer the following broad questions about the distribution of *P. ilicis*:

1. What is the nature of the variation in local population densities of *P. ilicis* and to what extent does this change as spatial scale increases?
2. Can variation in local population densities be understood in terms of measured components of demographic rates?
3. To what extent can variation in local population densities and demographic rates be explained in terms of environmental variation?
4. What are the implications of the results for other areas of macroecology?

### *Thesis outline*

### *Chapter 2 The holly leaf-miner system*

This chapter assembles the relevant background information for the research into the distribution *P. ilicis*. It serves as an introduction to the holly leaf-miner system, describing both the natural history of *P. ilicis* and its host, European holly. In addition, a general review of the literature on leaf-miner population dynamics, with an emphasis on the holly leaf-miner, is provided. Included, are data and amendments to published work on *P. ilicis* phenology, derived from my own observations in the field. The holly leaf-miner is particularly amenable to studies on the spatial structure of populations. The reasons for this are discussed at the end of the chapter.

### *Chapter 3 The spatial structure of holly leaf-miner populations and variation in mortality rates at a regional scale.*

In this chapter, variation in densities of *P. ilicis* is investigated between trees separated by distances of a few metres to over 25km. Previous studies have found substantial variation in leaf-miner densities within individual habitat patches. One aim of the surveys described here was to detect spatial structure in leaf-miner population densities at a regional spatial scale despite the documented local variation. In addition, data are presented on differences between local populations in components of demographic rates. An attempt is made to correlate leaf-miner densities both with the demographic rates components and environmental conditions.

### *Chapter 4 The population density structure of the geographic range of the holly leaf-miner*

This chapter describes part of an extensive survey on the distribution of *P. ilicis* across Western Europe. The aims of the survey were to detect and characterise spatial structure in population densities across the entire geographic range and to relate this structure to broad scale changes in environmental conditions. The results of this research are compared to existing data on range structure and the generality of these previous studies is critically assessed.

### *Chapter 5 Spatial structure in the demographic rates of the holly leaf-miner across its geographic range*

The second part of the European survey focused on differences in components of the underlying demographic rates of *P. ilicis* populations across the geographic range. Broad scale spatial structure in these rates is described and the relationships between this structure, environmental variation and population densities are investigated.

### *Chapter 6 The interspecific abundance-distribution relationship and the vital rates model*

The positive correlation found between the mean local abundance and the distribution (expressed either as the extent of occurrence or area of occupancy) within species assemblages is a common pattern found at a broad range of spatial scales (Gaston, 1996) and has potentially important consequences for a variety of ecological applications. Recently Holt *et al.* (1997) have proposed a model to explain this relationship in terms of differences in demographic rates between species across environmental gradients. This chapter explores the model in some detail using computer simulation techniques. The simulations are used to investigate the robustness of the abundance-distribution relationship by altering the amount of variability in demographic rates between species. The utility of simulation techniques as an intermediate step between theoretical and empirical work is discussed.

### *Chapter 7 Discussion*

The results obtained during the course of the investigations are considered together here in order to address a number of issues related to understanding spatial patterns in density. The results of all of the *P. ilicis* surveys are compared and the effects of spatial scale on the interpretation of ecological pattern is discussed. In addition, the assumptions behind theoretical models of range structure and the abundance-distribution relationship model are assessed in the light of the data collected for *P. ilicis*. The adequacy of the statistical techniques employed during the thesis are also discussed and some alternative approaches are briefly explored. Recommendations for future work on the holly leaf-miner system in particular and into range structure as a whole are also presented.

# Chapter 2

# The holly leaf-miner system

One of the major obstacles to our understanding of spatial patterns of abundance at any scale is the difficulty in obtaining accurate estimates of local population densities and key demographic rates at a sufficient number of sites. Population estimation can be a time consuming process for many species (see, for example, the discussion in Wiens, 1989) and birth and death events are often effectively hidden from direct measurement. However, some species may provide suitable model systems that could lead to the establishment of more general principles concerning the spatial structure of local populations.

This chapter introduces the holly leaf-miner system, discussing the natural history of both the leaf-miner and its host plant. Some new information about the system which was collected in addtion to the studies described in the subsequent chapters is also provided. Next, an overview of the population studies that have already been conducted on holly leaf-miners is given and these are set in the context of what is already known about leaf-miner population dynamics in general. Finally, the reasons why the holly leaf-miner system may be considered a suitable candidate for studies on the spatial structure of local populations are summarised.

Leaf-mining insects feed on live foliage while simultaneously living within it. They often leave the leaf epidermis intact, consuming only the internal tissues. This results in characteristic blotches or serpentine mines visible on the leaf surface. Most species are present in mines as larvae but may leave them to begin different feeding habits in later stages. In comparison to external feeders, they are relatively small insects (Gaston *et al.*, 1991), physically constrained by the thickness and area of the leaves they occupy. There are approximately 10,000 species of leaf-miners but all are within the holometabolous orders Coleoptera, Diptera, Hymenoptera and Lepidoptera. Leaf-mining has evolved independently several times in each of these orders, derived from other forms of herbivory (Connor & Taverner, 1997). There have been a number of possible reasons given for the evolution of leaf-mining, including protection from natural predators, environmental stress and plant defences by selective feeding. The relative importance of each of these factors is likely to vary between different species.

## The holly leaf-miner system

The holly leaf-miner, *Phytomyza ilicis* Curtis, feeds on European holly (*Ilex aquifolium* L.). This is a dioecious, evergreen tree from a large, cosmopolitan genus of approximately 400 species world wide (Galle, 1997).

European holly is relatively small, typically reaching a maximum height of 10-15m although it has been recorded at heights of up to 23m (Peterken & Lloyd, 1967). Its appearance varies considerably from a densely branched, pyramidal habit in open habitats to a pendulous or bushy form in woodlands. Historically, many trees were also cut to ground level or pollarded at 2-3m for timber and forage for livestock. As a result, they are now often found with multiple main stems. The foliage is dark green, glossy and usually spiny with a relatively thick cuticle. However, the size and shape of the leaves may also vary a great deal dependent upon a number of factors. Leaves on mature trees are often large with entire margins, particularly on the upper branches, whereas on younger trees the leaves are smaller and more spiny. Habitat type also appears to play an important role in leaf morphology. Leaves on trees in open upland areas have a thicker cuticle, are more spiny and have shorter internodal distances than those on trees in lowland, shaded areas (personal observation). In addition, shoots that are subject to browsing by deer and livestock often support small leaves with proportionately longer spines (Peterken & Lloyd, 1967; Obeso, 1997). Dormer & Hucker (1957) propose a number of physiological mechanisms that may influence spinescence in holly. In addition to the influence of biotic and environmental factors on the appearance of *I. aquifolium*, it is a species noted for the large number of varieties that may appear in wild populations. This has made it a popular plant with horticulturists and many cultivars have been developed with different forms of leaf-variegation, bark, habit and fruit. Galle (1997) lists 400 varieties that are known from various collections.

*I. aquifolium* has one flush of new leaves per year. In the UK this typically occurs in May, followed by a two month period of rapid shoot elongation and leaf development. After this, leaf growth is slow (Peterken & Lloyd, 1967). Most leaves are retained for up to five years before they are shed at a low rate throughout the seasons. European holly has small white flowers, borne on the leaf axils of old growth. Flower buds first appear in late summer on the current year's growth but do not open until the following May or June. Determining the sex of a flowering tree is quite straightforward. Female flowers have reduced stamens and the four stigmas form a disc-like plate in the centre. Male flowers have large, conspicuous stamens and reduced ovaries.

Occasionally, trees may be found bearing flowers of both sexes (Peterken & Lloyd, 1967). The fruit on female trees, which are typically red and approximately 5-10mm in diameter, are generally found only on individuals above 1.5m in height and only commonly on those above 3m. They usually mature by November or December.

*Ilex aquifolium* has an oceanic southern distribution (Peterken & Lloyd, 1967). Its natural range extends throughout north-western, central and southern Europe (see figure 2.1; Peterken & Lloyd, 1967; Hultén & Fries, 1986). It can also be found less commonly in scattered localities in parts of North Africa and has been reported as having a narrow band of distribution extending into Asia Minor although there is some doubt as to whether this is indeed so. European holly has also been introduced into North America. In the 1850s it was brought to Oregon as seed and small rooted plants from England and France. Since the early 1900s it has been extensively cultivated in the Pacific Northwest for decoration at Christmas (Cameron, 1939; Galle, 1997)

Iversen (1944) found that the northern and eastern limit of the natural range of holly is most probably determined by the amount of winter cold. From a study of frost damage on holly in Denmark, he showed that the mean temperature of the coldest month is a good indicator of the ability of holly to survive in a particular region. Holly is absent from regions where this value is below -0.5°C. In addition, he also suggested that holly is generally absent from areas where the mean warmest month is below 12°C. Perring & Walters (1962) found that in the UK holly is not found naturally in regions where July temperatures fail to reach 12.8°C. Some circumstantial evidence (see Peterken & Lloyd, 1967) suggests that this is the minimum summer temperature requirement for successful fruit production. Although individuals are able to survive in regions below this limit, they are invariably sterile. The reasons for the southern and south-eastern limits to the geographic range have not received similar empirical treatment. However, Sykes *et al.* (1996) have used a bioclimatic model to simulate distributions of north European tree species and have found that many species ranges correlate well with an index of water availability in the south and south-east. This index is calculated from the ratio of actual transpiration to equilibrium evapotranspiration to give a measure of drought stress in a region. For *Ilex aquifolium*, the actual and simulated range limits are remarkably similar so it appears that drought stress is likely to be an important factor determining its distribution in the south.

In addition to browsing of the young leaves by deer and livestock, there are several phytophagous insect groups that feed on holly within its natural geographic range, including aphids and some Lepidoptera (Peterken & Lloyd, 1967; Blackman & Eastop, 1994). Galle (1997) also list a number of species of weevils and scale insects that feed on holly in North America. During the surveys conducted for this thesis, characteristic leaf-chewing damage was found at a number of localities in the UK and the rest of Europe which seems likely also to be due to weevils. In addition, scale insects were also uncommonly observed on some leaves. However, *Phytomyza ilicis*, the holly leaf-miner, is by far the most common insect herbivore of European holly.

*P. ilicis* belongs to an economically important genus of small, leaf-mining agromyzid flies. It is strictly monophagous, feeding exclusively on *Ilex aquifolium*. A number of closely related species within the genus are known to feed on other holly species and most are also host-specific (Kulp, 1968; Griffiths & Piercey-Normore, 1995; Galle, 1997).

The life history of *P. ilicis* in the UK is given by Cameron (1939) and Lewis & Taylor (1967). The holly leaf-miner exhibits a univoltine life-cycle. Adult flies emerge from their mines in late May. Females puncture holly leaves with their ovipositor to produce wounds from which both they and male flies feed. This results in numerous characteristic scars on the foliage which can be used to measure feeding activity (Valladares & Lawton, 1991). After fertilisation, the female lays eggs on the new flush of leaves in June. They are inserted into the base of the midrib on the underside of the leaf. The female bores a very narrow tunnel into the midrib and lays a small (0.383 by 0.160 mm; Cameron, 1939), white egg at the end of it. This process also scars the leaf and can be used to measure the amount of oviposition (Valladares & Lawton, 1991; see also this thesis) that occurs on a tree.

After 3 or 4 days (Galle, 1997) the egg hatches and the larva eats through the midrib. It continues to do this until autumn when it will emerge from the midrib into the leaf lamina itself. It is at this stage that the characteristic blotch mine begins to form. The larva gradually eats away the parenchyma below the epidermis until a large irregular mine is formed. This mine reaches its maximum size in March and may occupy a considerable area beneath the epidermis. Between July and March the larva moults twice and passes through three stages, the first lasting from July to December, the second from December to January, and the third from February until the formation of the pupa at about the end of March (Cameron, 1939).

Just before pupation, the mature larva prepares a thin triangular area on the cuticle of the leaf, the larva presses its ventral surface against the epidermis and then begins pupation within the last instar skin which becomes the pupa. The anterior spiracles of the pupa are projected through the attenuated area of cuticle. The imago emerges from the leaf by pressing with the ptilinum against a hinged emergence plate on the puparium, which opens and breaks through the thin leaf epidermis above it. This leaves a triangular raised flap or hole on the surface of the leaf which is an indication that an adult fly has completed its development successfully.

### *Natural Enemies*

Auerbach *et al.* (1995) point out that in nearly every study examining the mortality of leaf-miners, larval parasitism and occasionally pupal parasitism is an important source of mortality. Moreover, the parasitoid communities associated with leaf-mining insects are more species rich than those associated with any other guild of phytophagous insects. *P. ilicis* is no exception to this general trend.

Nine species of parasitoids have been recorded attacking *P. ilicis* in the UK (Cameron, 1939). These are listed in Table 2.1. Of these species, there are four that are most important (Lewis & Taylor, 1967). *Chrysocharis gemma* is the only recorded endoparasitic larval parasite. The adult parasitoid inserts a single egg through the leaf cuticle into the body cavity of the larva in the early spring. Occasionally superparasitism can occur when two or more eggs are laid in the same larva. Adult *C. gemma* usually begin to emerge from their parasitised mines at the beginning of June. The imago chews a small, round hole in the leaf. This, together with the absence of a *P. ilicis* pupa and its own conspicuous black pupal skin within the mine allows simple identification of many of the mines that have been parasitised by *C. gemma*. However, during the studies described in subsequent chapters of this thesis, some localities were found to have a significant proportion of parasitised mines without the round emergence hole. Dissection of the mines revealed that *C. gemma* was able to undergo full development but then fail to escape from the holly leaf itself. Many examples of larval parasitism were found with dead adults inside the mines and an emergence hole absent or only partially completed. The cause of the deaths of the adult parasitoids is unknown but may be related to the size of the adult parasitoid and the toughness of the leaf cuticle from which it tries to escape.

*Sphegigaster pallicornis* and *Chrysocharis pubicornis* are the two most common pupal parasitoids. Pupal parasitism may be identified at the end of the life cycle by the presence of the leaf-miner pupal scar on the leaf-surface, a round emergence hole (as

| **Superfamily / family** | **Species** |
|---|---|
| *Chalcidoidea* | |
| *Eulophidae* | *Chrysocharis gemma* Walker |
| | *Chrysocharis pubicornis* Zetterstedt *(C. syma)* |
| | *Pedobius acantha* Walker *(Pleurotropis amyntas)* |
| | *Closterocerus trifasciatus* Westwood |
| | *Tetracambe sp.* |
| *Pteromalidae* | *Sphegigaster pallicornis* Spinola *(S. flavicornis)* |
| | *Cyrtogaster vulgaris* Walker |
| | *Eutelus* sp. |
| *Ichneumonoidea* | |
| *Braconidae* | *Opius ilicis* Nixon |

Table 2.1. Parasitoids of *Phytomyza ilicis* in the UK. Taxonomy according to Fitton *et al.* (1978). Synonyms used by Cameron (1939) and Lewis & Taylor (1967) are given in brackets.

opposed to the triangular flap denoting successful emergence of the adult leaf-miner) and the parasitoid's own black pupal skin. Often the leaf-miner pupal case will also be present in the mine although this is not always the case. *Pedobius acantha* is less common than the other parasitoids but is unique among the leaf-miners' parasites because it may be either a primary parasite on the pupa of the leaf-miner or hyperparasitic on the leaf-miner's pupal parasites. This species may have two generations in a year, the first in spring on the leaf-miner pupae and then another in August on the persisting pupae of *P.ilicis* which now only contain pupae of the other parasitoid species, all successful flies having emerged by this time. Figure 2.1 summarises the phenology of *P. ilicis* and these hymenopterous parasitoids.

Another important source of mortality for *P. ilicis* in some areas is larval (and more rarely pupal) predation by birds. Evidence of bird predation is easily detected since the mines are torn open leaving a characteristic, triangular flap.

## Population studies on *Phytomyza ilicis* and *P. ilicicola*

There are two species of holly leaf-miner which have received significant attention in population studies. Only *P. ilicis* occurs in Britain but North America has its own indigenous species, *Phytomyza ilicicola* Loew which lives on American holly (*Ilex opaca* Aiton), as well as *P. ilicis* which was accidentally introduced when European holly was imported for cultivation (Cameron, 1939). Their life-histories are very similar and for the purpose of this review the two species will be considered together. This is because the emphasis of research on *P. ilicicola* has been on tree responses to leaf-miner infestation and intraspecific competition, whereas for *P. ilicis* it has been on parasitism and predation.

### *Oviposition rate and its consequences for mine distribution*

Oviposition rates of *P. ilicis* between trees were estimated by Valladares & Lawton (1991) by counting the number of oviposition scars visible on the midribs of holly leaves. Over a three year study period consistent differences were found between trees, reflecting preferential oviposition by adult female *P. ilicis*. Surprisingly, no physical or chemical characteristics of the bushes that they measured appeared to reflect this host selection by females and larval performance was not significantly higher in preferentially selected trees. The proportion of eggs that successfully hatched and subsequent larval performance were not significantly higher in preferentially selected trees.

Since numbers of egg scars and numbers of mines are strongly correlated, the distribution of mines reflects female choice in oviposition sites. In a study of a holly hedge, Heads & Lawton (1983) observed that mines are overdispersed among leaves (the number of leaves with more than one mine is less than that expected by chance) but that mine distribution is spatially clumped in patches of high leaf density. The mean number of mines per leaf remained constant at a number of spatial scales so leaf-miner aggregation simply reflected the number of available leaves in a patch.

However, this result is inconsistent with a short study that I conducted on the distribution of mines within trees. Two holly trees were selected in a mixed woodland and from each, 100 shoots were haphazardly selected from all around the tree. On each shoot, the number of leaves and the number of mines were counted. To test whether leaf-miners were randomly distributed, a short simulation program was written that randomly allocated mines to each leaf and then counted the number of shoots that would be occupied. The simulation was repeated 16000 times for each tree to create a distribution of the expected number of occupied shoots if mines were randomly distributed. For both trees, the observed number of shoots which had mines on them was significantly less than the number of shoots expected (tree 1 leaves = 609, mines = 88, p=0.006; tree 2 leaves = 641, mines = 61, p=0.000). This indicates that females are preferentially laying eggs on some shoots and not others.

Potter (1985) also noted the highly aggregated distribution of oviposition sites in *P. ilicicola* and speculated that host suitability may be important in causing these patterns. He had observed that adult females can only lay eggs on leaves less than a few days old. One consequence of this would be that only a small proportion of leaves are suitable for egg-laying during the peak activity of the adults. Differences in host phenology may therefore be reflected in leaf-miner infestation. Auerbach *et al.* (1995) point out that the importance of host phenology in leaf-miner population dynamics is a common theme in many studies and may be a critical determinant of population density.

### *Miscellaneous larval mortality*

Both *P. ilicis* and *P. ilicicola* are subject to very high mortality rates both from their natural enemies and a variety of other sources. An important component of this, usually referred to, in this context, as miscellaneous larval mortality (Heads & Lawton, 1983; Valladares & Lawton, 1991), is when larvae die within the leaf with no obvious signs of attack from other organisms. Miscellaneous larval mortality in *P. ilicis* is density-independent (Heads & Lawton, 1983) and it is possible that some of this mortality can be attributed to climatic factors mediated through the tree during the winter months, although a leaf-miner's immediate environment may provide some protection from the weather. Overwintering mortality has been shown to be important for other leaf-mining species on deciduous hosts (Connor, 1984; Connor *et al.*, 1994) but there is no direct evidence for such an effect on leaf-miners on evergreen hosts.

During the winter of 1997-1998, I set up a small experiment to test whether leaf-mines afforded protection to their occupants from fluctuations in temperature. Three relatively large mines were selected in October 1997. Into each of these, a small thermocouple was inserted and the mine was resealed using candle wax. An additional thermocouple was attached to the underside of each mined leaf, as close as possible to the leaf-surface without touching it. Temperatures from all six thermocouples were logged at hourly intervals to an accuracy of 0.1°C until March 1998. During this period no differences were detected between the temperatures inside and outside the mines. This suggests that mines do not help to insulate the leaf-miners during the winter months. However, the mine does prevent physical displacement of the leaf-miner from the host by the wind and the rain and it is also possible that some protection against wind chill is provided.

During the survey work carried out for this thesis, two potential types of miscellaneous larval mortality were identified. In addition to the typical form which is found on medium-sized mines, in some places there is a significant amount of mortality that occurs when the leaf-mines are at a very early stage of development. The mines are so small that they are often not immediately recognisable using the usual method of counting mines in the leaf laminae alone. They are most easily spotted when simultaneously counting oviposition scars at the base of the leaf midrib. The size of the mine suggests that this mortality is occurring in late autumn or winter and may be evidence that climatic factors are an important source of leaf-miner mortality.

For *Phytomyza ilicis*, then, it is certainly possible that overwintering mortality may be important since there is no evidence to suggest that the mines help to protect the leaf-miner from fluctuations in temperature and, at some localities, many larvae die at a point in their life-cycle that coincides with the coldest months.

Two other causes of miscellaneous larval mortality have been suggested. One is that it is the result of attacks from natural enemies, either stinging by parasitoids or predation. The other is that it may be the result of a physiological response by the tree to infestation. Holly is not considered to be a very attractive host to herbivores (Potter & Kimmerer, 1986, and references therein). It has tough, spiny leaves, low leaf nitrogen content for much of the time and produces toxic and digestibility-reducing allelochemicals. Potter & Kimmerer (1986) suggested that the prolonged life-cycle of *P. ilicicola* is a response to the seasonal allocation of defence investment in American holly. Anstead (1996) found some evidence to suggest that levels of phenolic compounds are higher in leaves mined by *P. ilicis* on unshaded parts of holly. This could be due to a plant response to wounding though there is currently no data available on the relationship between miscellaneous larval mortality and levels of phenolics.

Potter (1985) showed that early leaf abscission may also be a response to infestation by *P. ilicicola.* A proportionately higher number of mined leaves were shed than unmined ones and there was high miner mortality in these leaves. However, Kahn & Cornell (1983) have argued that in some areas the effects of early leaf abscission on mortality may be moderated by the fact that shed leaves provide miners with a refuge from pupal parasitism, since parasitoids do not search fallen leaves for hosts. Potter (1985) has examined the effects of intraspecific competition on *P. ilicicola.* He found a positive relationship between the number of mines on leaves and miner mortality. In addition, there was a significant negative relationship between mine number and pupal weight. Heads & Lawton (1983) found that, at their study site, more than one mature *P. ilicis* mine per leaf is a rare event. However, in some areas there can be a high incidence of leaves with multiple mines particularly where overall density is high (personal observation) and in these areas the effects of competition may be important.

### *Natural Enemies*

Holly leaf-miners can suffer high levels of predation by birds. Heads & Lawton (1983) found that bird predation accounted for an average of approximately 40% of total mortality in one study, though there was a great deal of variation between sampling sites. In addition, there was a surprising, negative relationship between mine density and levels of bird predation. However, they point out that the ability to detect density-dependent mortality and the nature of the relationship may depend on the spatial scale over which samples are taken.

The number of species of parasites attacking holly leaf-miners and the amount of mortality imposed by them is subject to considerable spatial variation. At a broad spatial scale, Cameron (1939) noted that some species of *P. ilicis* parasitoids were absent from his sampling sites spread across Southern England, though this could have been due in part to the fact that rarer species may have been too scarce to detect in some samples. Potter (1985) notes that the larval parasite, *Opius striativentris* Gahan (Hymenoptera: Braconidae) is the predominant species that attacks *P. ilicicola* in Kentucky whereas in other areas of North America pupal parasites are also very important (e.g. Kahn & Cornell, 1983). At a local scale there is also considerable spatial variation in parasitism levels. *Chrysocharis gemma*, the most important parasitoid of *P. ilicis* (Cameron, 1939), shows a significant aggregative response to levels of leaf-miner infestation (Heads & Lawton, 1983; Valladares & Lawton, 1991), whereas pupal parasitoids do not. Potter (1985) found that parasitism of *P. ilicicola* by *O. striativentris* was density independent but did show wide variation between trees.

*Dispersal*

The importance of immigration and emigration in governing levels of abundance of holly leaf-miner has not been documented. This is probably because of the practical difficulties in quantifying insect dispersal. However, there must be at least some migration, since small populations do seem to persist on isolated holly trees which would otherwise be prone to local extinction due to stochastic effects.

### Suitability for studies on the spatial structure of local populations.

There are a number of features of the holly leaf-miner system that make it suitable for studies on spatial variation in population density and demographic rates.

First, because most larvae produce large, irregular mines that are clearly visible, this makes estimating the densities on individual trees a relatively simple task. Similarly, oviposition rates can be estimated by counting the number of egg scars. This also helps to identify individuals that have succumbed to miscellaneous larval mortality at an early stage in the life-cycle. Moreover, holly is evergreen and its leaves typically stay on the tree for a number of years. This, in conjunction with the fact that *P. ilicis* is univoltine, means that leaves of known age can be sampled for quite a long period after the life-cycle of the leaf-miner is completed to give a density estimate for a particular generation. Many of the sources of mortality of *P. ilicis* leave characteristic traces in the mines and the number of successful adult emergences can also be identified. For many animal species there is often no way of measuring sources of mortality without direct observation at the time that it occurs.

*P. ilicis* is monophagous and so its distribution is strictly limited by the availability of holly trees. Within its natural range, European holly is a distinctive species and this allows rapid identification and censusing of suitable sites for leaf-miners. In addition, it may occupy a broad range of habitats (Peterken & Lloyd, 1967) and shows considerable morphological variation.

Previous studies, outlined above, have identified a great deal of variation in the densities of *P. ilicis*. A large part of this thesis exploits features of the holly leaf-miner system in order to determine the extent to which this variation in leaf-miner densities is spatially structured and whether this can be attributed to the interaction between variation in demographic rates and environmental conditions.

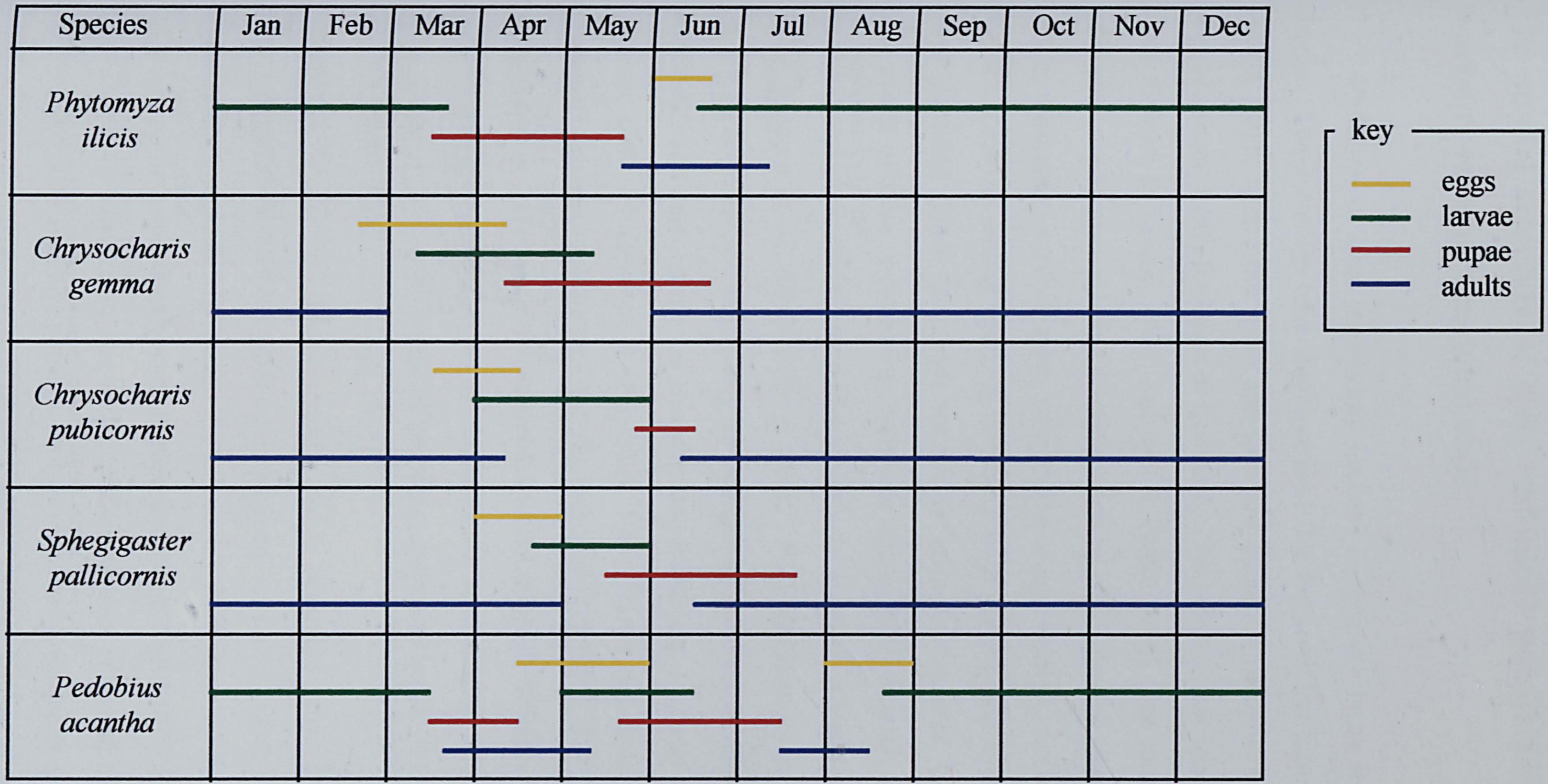


Figure 2.1 Phenology of *Phytomyza ilicis* and its four most important parasitoids in the UK (adapted from Lewis & Taylor, 1967).

# Chapter 3

## The spatial structure of holly leaf-miner populations and variation in mortality rates at a regional scale

Variation in local population densities is typically large for phytophagous insect species. Many studies have demonstrated this between individual host plants within a habitat patch, where the mean and standard deviation of densities may be almost equal (e.g. Harrison, 1997; Redfearn & Cameron, 1993; Harrison & Maron, 1995; Nestel *et al.*, 1994; Bylund & Tenow, 1994; Bultman & Faeth, 1985). The cause of such variation has been attributed to graded host discrimination by adult females when ovipositing (Rausher, 1983), differential survival rates, or patterns of aggregation resulting from movements of individual insects among hosts. A variety of both genetic and environmental factors have been identified which contribute to this variation in densities. Genotypic differences in host plants offering resistance to herbivore attack, both through variation in morphology and the production of defensive allelochemicals, are well known (e.g. Rosenthal & Janzen, 1979; Mopper *et al.*, 1990; Fritz & Sims, 1992). Differences in the quality of the host plant as a food source which may be strongly affected by physical factors such as water availability and host density are also very important (e.g. Mattson & Addy, 1975; Athey & Connor, 1989; Minkenberg & Ottenheim, 1990; Hanks & Denno, 1993). Leaf-abscission rates may vary between individual hosts and this has been shown to affect some herbivore populations (e.g. Potter, 1985; James & Pritchard, 1988; Stiling *et al.*, 1991). In addition, attack rates from predators and parasitoids and competition with other herbivores have been found to influence levels of herbivory between hosts (e.g. Hairston *et al.*, 1960; Karban, 1989; Hanks & Denno, 1993).

Population density variation occurs at a variety of spatial scales, from differences between parts of an individual host plant up to regional variation within the geographic range of the species (Denno & McClure, 1983). It seems likely that the factors that influence such variation will often differ dependent upon the spatial scale under examination. In addition, it is unclear to what extent the variation that has been recorded between individual plants obscures broader regional patterns. Whilst there has been much work on spatial variation in trap catches of insects (e.g. Taylor *et al.*, 1980; Taylor & Woiwod, 1982; Taylor, 1986), there have been remarkably few surveys of regional or geographic density variation on hosts, other than for economically significant species (e.g. Cammell *et al.* 1989; but see McClure & Price, 1976; Eber and Brandl, 1994).

In this chapter, variation in densities of the holly leaf-miner *Phytomyza ilicis* are investigated between trees separated by distances of a few metres to over 25km. Several previous studies have remarked on very high between-tree variation in the densities of holly leaf-miners within a habitat patch, and have endeavoured to understand the determinants of this variation (e.g. James & Pritchard, 1988; Heads & Lawton, 1983; Valladares & Lawton, 1991). The purpose of this chapter is to determine whether any broad scale spatial structure can be discerned despite this documented variation and, if found, whether it can be correlated with environmental variation and differences in demographic processes between local populations.

## Methods

Data collected from three surveys are used in the analyses:

1. Sheffield density survey.

This survey investigated whether spatial structure could be detected in local population densities of *P. ilicis* at a regional scale. It was conducted within a 22×22 km square region centred upon the middle of the city of Sheffield, England (figure 3.1, Ordnance Survey grid coordinates SK425376 to SK447398). This area encloses part of the edge of the Peak District to the west, the urban and suburban areas of Sheffield and Rotherham, and low-lying agricultural land to the east. Altitudes ranged from 10m to over 400m above sea level. The region overlies coal measures and millstone grit from the Upper Carboniferous. Holly is very common throughout this area and can be found in a wide variety of habitat types ranging from sites within the city itself, to woodland, moorland and farmland in the outlying regions. Data were collected between October and December 1996. During this time an attempt was made to sample as many as possible of the 1km grid squares in the region as defined by the British National Grid. On the whole, measurement of leaf-miner density was restricted to just one tree per grid square. Trees were selected haphazardly in areas of high holly density. In areas of low density the sampled tree was often simply the first to be observed. To sample trees in a strictly random manner would have been impractical since this would involve knowledge of the positions of all the trees in each grid square, and ready access to all

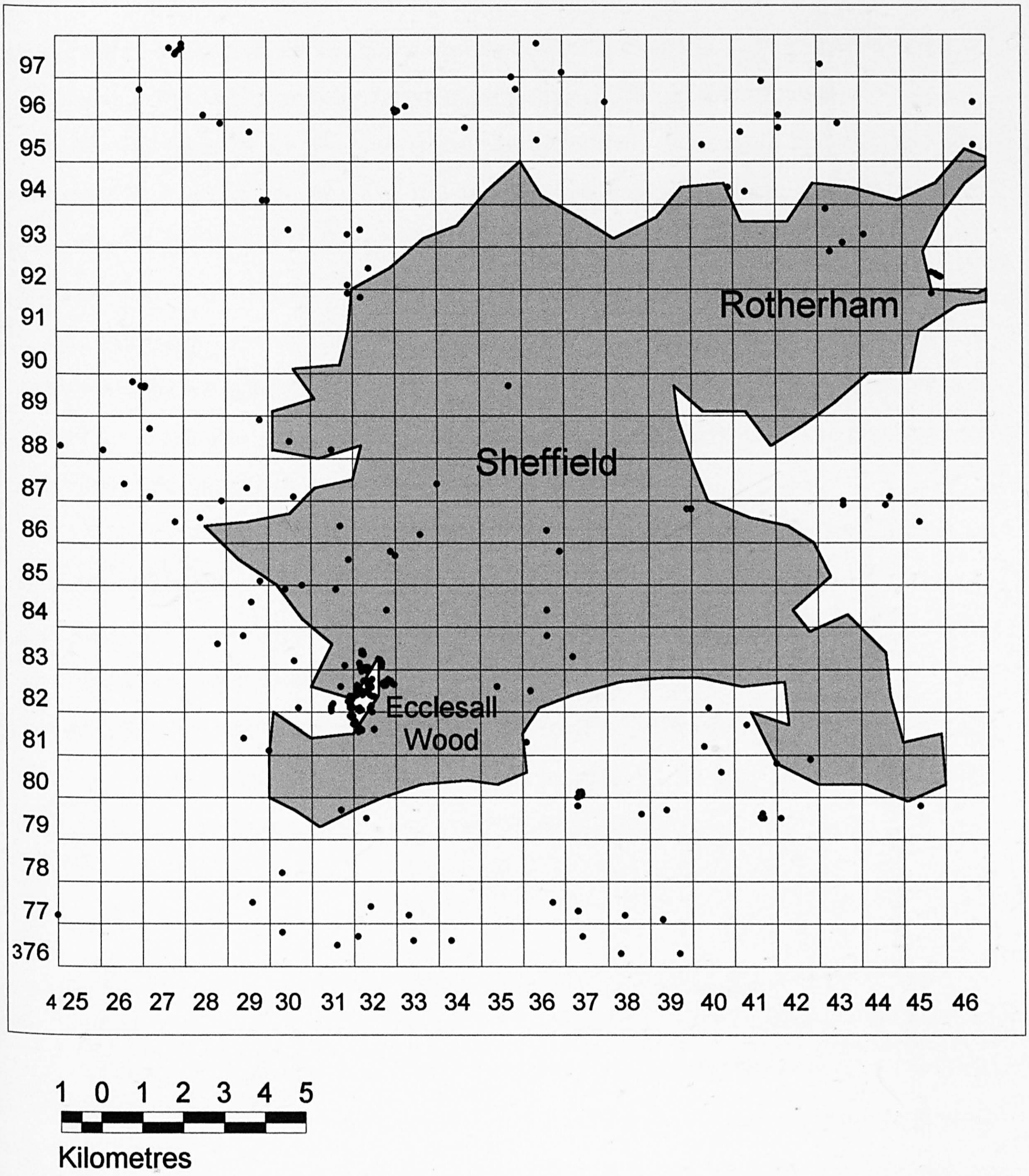


Figure 3.1 Sampling locations during the 1996 Sheffield density survey. The grey region represents the city limits of Sheffield and Rotherham. The1km grid corresponds to the Ordnance Survey National Grid.

areas within the region. In all, holly was found and sampled in 123 grid squares out of the possible 484, which corresponds to approximately 25% of the study area.

*P. ilicis* density for the generation that developed from eggs laid in 1995 was measured by haphazardly sampling 100 one year-old leaves from around a tree, between ground level and a height of approximately 2m. Leaves of a suitable age can readily be distinguished from the others by their position between the annual nodes on the branch. If no mines were found after this sampling, the presence or absence of mines was established by an exhaustive search of the tree and if a mine was found, density was treated as an arbitrarily small number (0.1 mines per 100 leaves). This was necessary for only 5 trees during the course of the survey.

Since evidence of bird predation is easily measured *in situ* (Chapter 2) rates of bird predation on holly leaf-miners, often a significant mortality agent (Heads & Lawton, 1983; Valladares & Lawton, 1991), were measured by inspection of 50 mined leaves on each of a subsample of 104 trees.

Finally, six well-separated patches of woodland were selected from the survey area and from these sites five trees were sampled to provide data in the spatial autocorrelation analyses for trees separated by small distances.

### 2. Ecclesall Wood Survey

The second data set came from another study (McGeoch and Gaston, 2000) of *P. ilicis* densities and mortalities in Ecclesall Wood. This is an ancient woodland of approximately 130 hectares located in a suburban area to the south-west of Sheffield (Ordnance Survey grid coordinates SK323825). Holly comprises the most important part of the middle-storey of this deciduous wood. Holly leaf-miner densities were collected in the same way as in the Sheffield Survey from 100 holly trees between July and November 1996. The data set was used to look for evidence of structure in population densities at this smaller spatial scale.

### 3. Sheffield demographic rates survey

Following the Sheffield density survey, a further investigation was conducted concentrating on the different habitat types within the region. Data were collected between August and November 1997. This allowed full identification of the fates of leaf-miners that developed from eggs laid in June 1996 (Chapter 2). Leaf-miner densities on each tree were again measured by haphazardly sampling 100 leaves but in addition to this, oviposition rate was determined by counting the number of egg-scars on the same leaves. Next a sample of approximately 100 mined leaves (or as many as possible) from the same generation were collected from each tree. These leaves were then dissected under a microscope. Successful emergence or the causes of the mortality of the larvae and pupae inhabiting the mine were identified using the characteristic features described in Chapter 2. During the course of this survey, 106 trees were sampled from a variety of localities that varied in extent from isolated, individual trees in high-altitude, exposed habitats to trees scattered throughout woodland and urban areas.

## Statistical Analyses

### *Spatial autocorrelation*

Spatial dependence in leaf-miner densities was investigated using spatial autocorrelation analysis (Cliff & Ord, 1973; Sokal and Oden, 1978; see Appendix 1 for the program code listing). Distances between all pairs of trees were calculated and then divided into discrete classes. Moran's I, an autocorrelation coefficient, was calculated for each distance class and these values were then plotted as correlograms (Cliff and Ord, 1981), which give an autocorrelation profile for these data. Analyses were performed both for equal and logarithmic distance classes in order to emphasise patterns of autocorrelation at different scales. Equal distance classes give a good impression of the pattern of autocorrelation across the whole study area. Logarithmic distance classes emphasise any spatial structure at shorter distances. The Sheffield survey and Ecclesall Wood data sets were analysed separately to avoid biasing the correlograms with a large number of points (100 out of 253) from the relatively small area of Ecclesall Wood.

## Results

### *Frequency distribution of leaf-miner densities*

The mean density of *P. ilicis* on all trees from both density surveys was 13.7 mines per 100 leaves with a variance of 66.6. A variance/mean ratio of 4.86 indicates that the distribution of leaf-miners between trees is obviously not random. A negative binomial distribution was constructed, estimating *k* (the "clumping" parameter) using the maximum likelihood method given in Krebs (1989). A chi-squared goodness-of-fit test of the data to this distribution demonstrated that it is an adequate description of the aggregation apparent in the data set (Figure 3.2; $\chi^2 = 21.17$, d.f. = 24 (after pooling), $p>0.6$).

*Spatial patterns of leaf-miner density*

The results of the spatial autocorrelation analyses of leaf-miner densities from both the Sheffield and Ecclesall Wood surveys are summarised in figures 3.3 and 3.4. Each value of Moran's I was tested for significance, and the overall significance of each correlogram was calculated using the Bonferroni correction for multiple comparisons. Individual coefficients may not differ significantly from zero for two reasons, the value of Moran's I may be close to zero or there may be an insufficient number of comparisons within a distance class.

For all the correlograms, there is a general trend of decreasing autocorrelation in leaf-miner densities with distance.. That is, holly trees close to one another have more similar densities than those further apart. Figure 3.3a shows that there is positive autocorrelation in densities across the Sheffield area up to a distance of approximately 6km. Moreover, there appears to be a gradient in densities across the study area since Moran's I becomes significantly negative for the largest distance classes. The pattern for Ecclesall Wood (figure 3.4a) is qualitatively similar, but there is positive autocorrelation only to 200-300m. Considering just the significant values of Moran's I, there appears also to be a gradient in densities across the wood, but the pattern is rather more complicated than that for the Sheffield survey as a whole.

The correlograms based on logarithmic distance classes (figures 3.3b, 3.4b) show that the pattern of spatial autocorrelation also holds strongly at smaller scales. The gradients across the entire study areas are not so clear in these correlograms because a large number of data pairs, representing a broad range of distances, are confined to the last distance class.

Figure 3.5 is an interpolated map of leaf-miner densities across the study area. The pattern of leaf-miner densities is certainly not simple but neither is it random. Elements of the spatial structure are apparent. For example there are several regions of high population density (e.g. around SK431386) and densities tend to be generally higher towards the North and East. This broader scale structure is the origin of the gradient in densities identified from the correlogram (figure 3.3a).

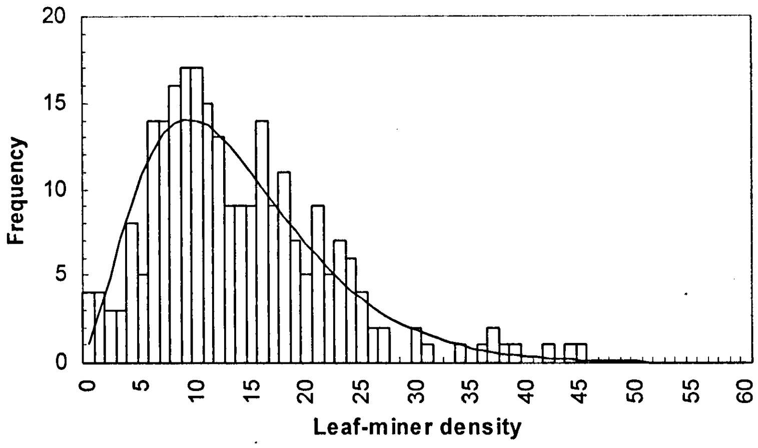
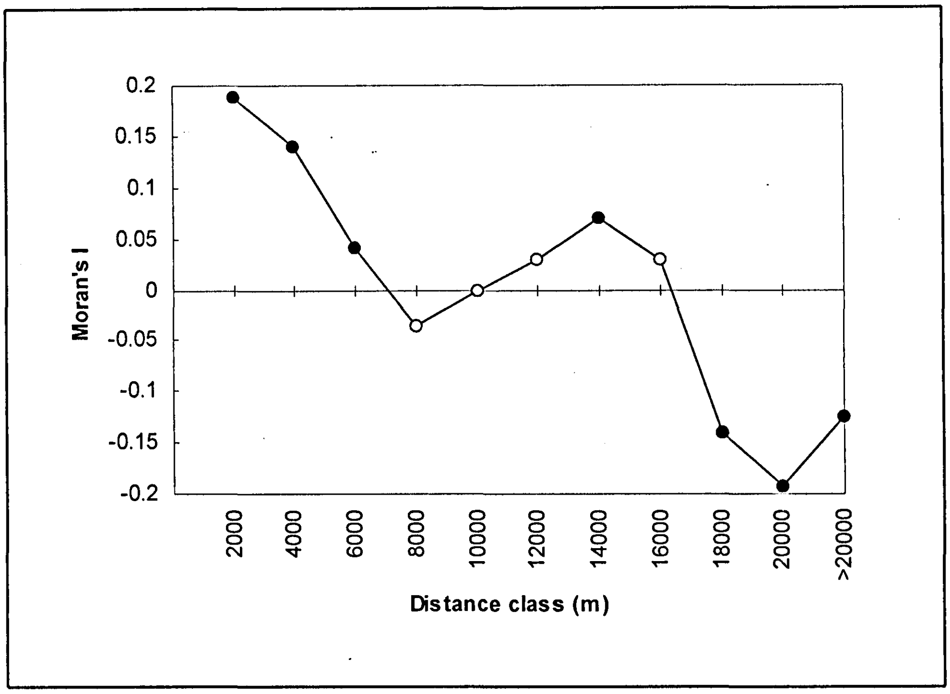


Figure 3.2. Frequency histogram of leaf-miner densities from the Sheffield density survey. Columns are the observed frequencies of trees supporting each of the population densities. The line represents frequencies from a negative binomial distribution (k = 3.48).

a)



b)

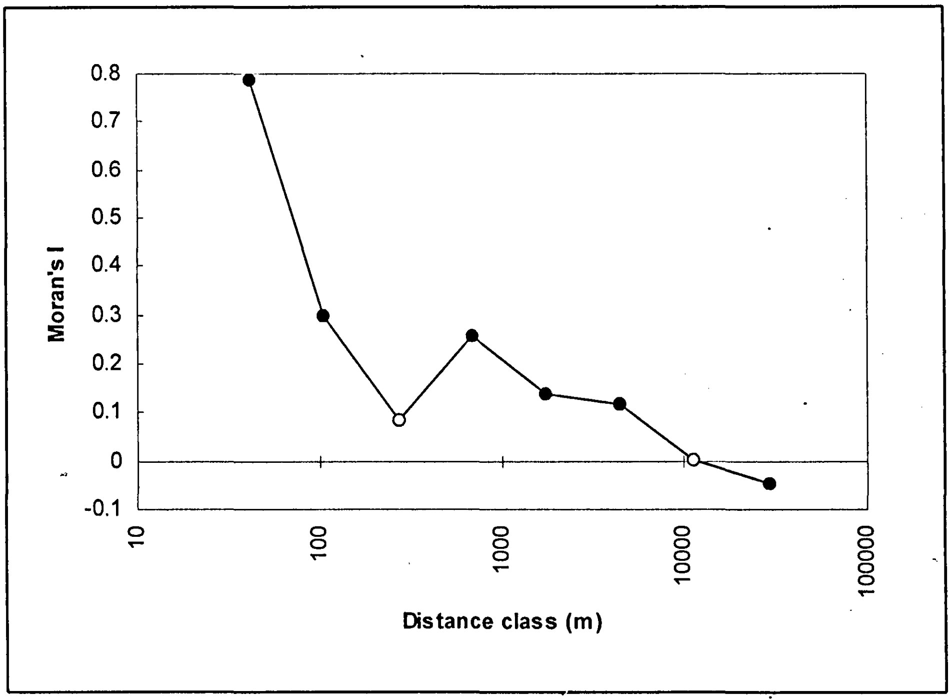
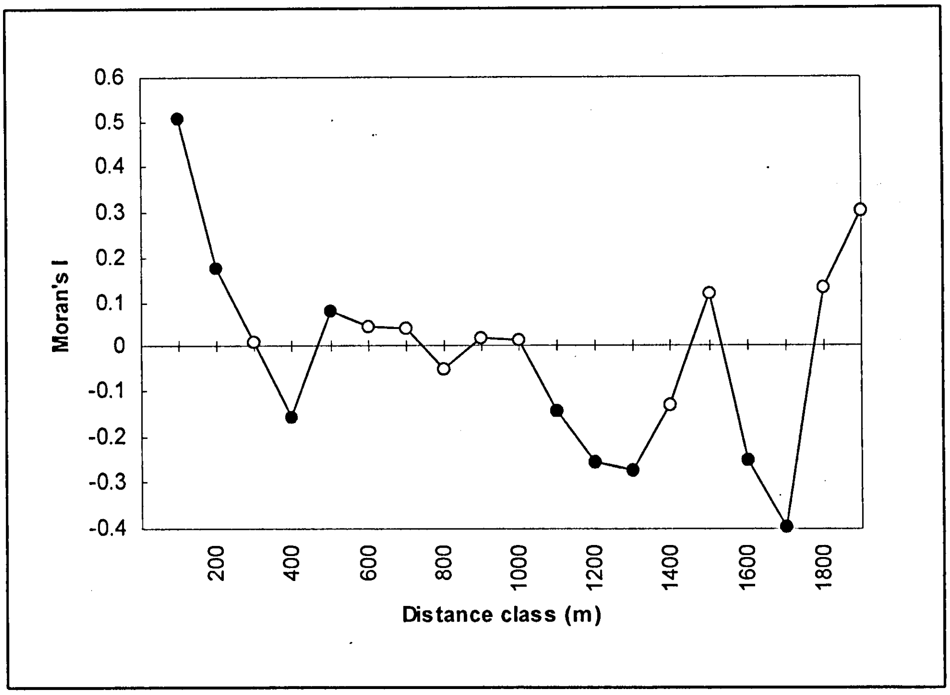


Figure 3.3 Spatial correlograms of *P. ilicis* densities from the Sheffield density survey. Scale on the *x*-axis represents maximum distance in each distance class. Filled circles indicate significant values for Moran's I ($p < 0.05$). Overall correlogram significance was tested using Bonferroni's correction for multiple comparisons. (a) Equal distance classes ($n = 153$; overall significance: $n = 11$, $p < 0.01$) (b) Logarithmic distance classes ($n = 153$; overall significance: $n = 8$, $p < 0.01$.

a)



b)

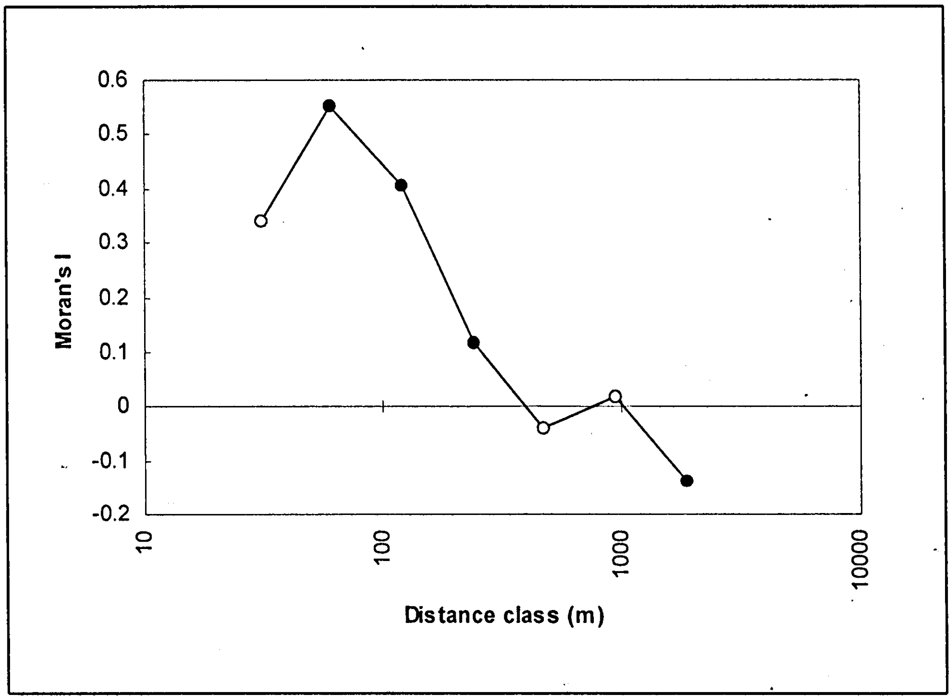


Figure 3.4 Spatial correlograms of *P. ilicis* densities from the Ecclesall Wood survey. Scale on the *x*-axis represents maximum distance in each distance class. Filled circles indicate significant values for Moran's I ($p < 0.05$). Overall correlogram significance was tested using Bonferroni's correction for multiple comparisons. (a) Equal distance classes ($n = 100$; overall significance: $n = 19$, $p < 0.01$) (b) Logarithmic distance classes ($n = 100$; overall significance: $n = 7$, $p < 0.01$.

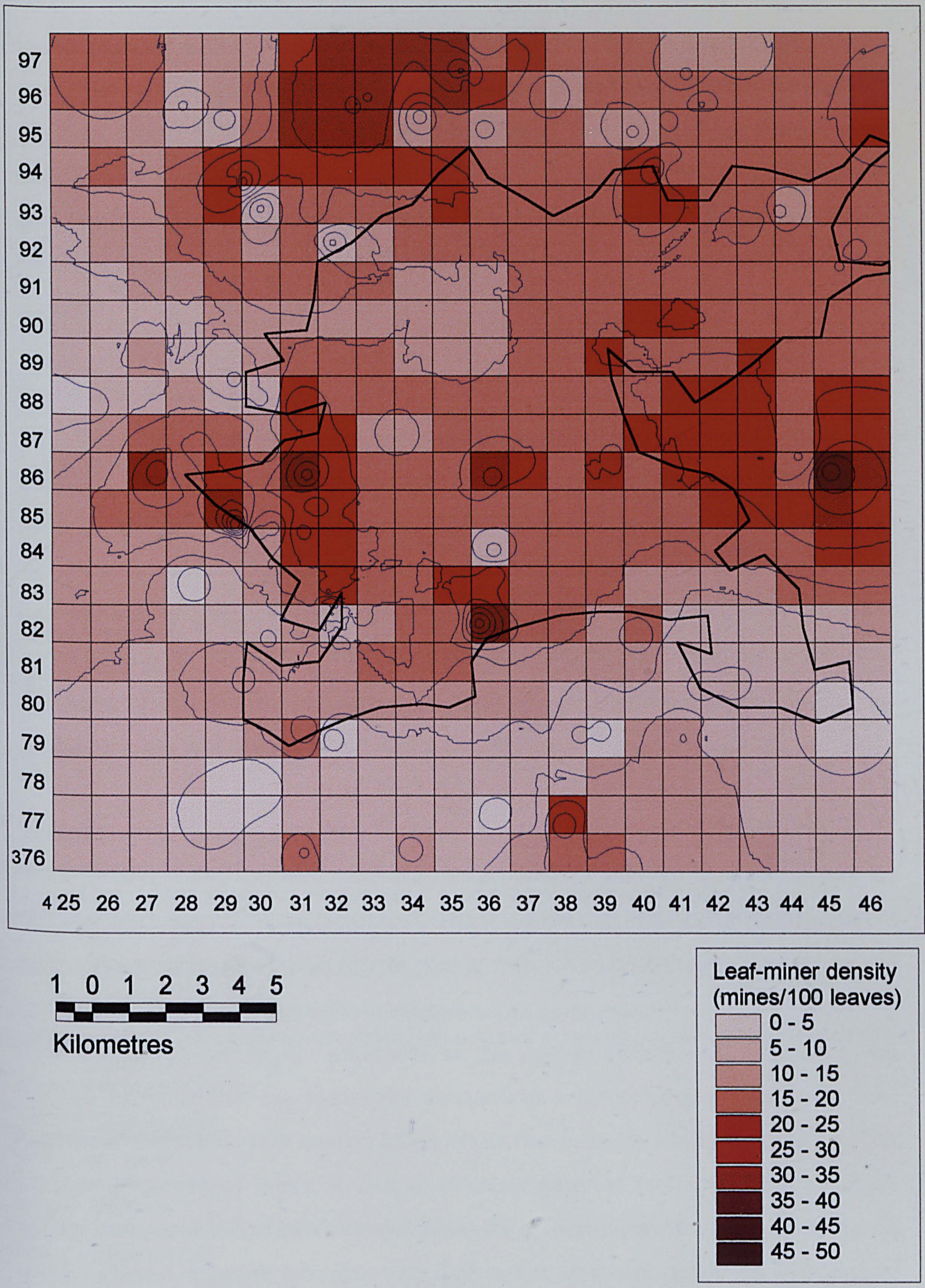


Figure 3.5 Interplated surface of Phytomyza ilicis densities across the Sheffield Region. The 1km grid corresponds to the Ordnance Survey National Grid. The blue contours indicate intervals in leaf-miner density of 5 mines/100 leaves.

### *Spatial patterns in bird predation*

There is little spatial autocorrelation in levels of bird predation across the whole of the study area (Fig. 3.6a). However, within Ecclesall Wood (Fig. 3.7a) there is some indication of a gradient. The logarithmic plots (Fig. 3.6b and 3.7b) reveal significant autocorrelation for both data sets on a scale of hundreds of metres. 'Patchiness' in bird predation appears to occur at a much smaller scale than it does for overall leaf-miner density.

### *Correlation between years*

Where trees were sampled in both the 1996 and 1997 surveys, their densities were compared in order to establish whether trees tend to have consistent densities between years (noted by Valladares & Lawton, 1991, and apparent during the course of the survey work). Figure 3.8 shows that although trees in 1997 tended to have higher densities than in 1996, relative densities between years are reasonably closely correlated ($r = 0.77$, $n = 26$, $p < 0.001$). This correlation suggests that factors found to influence densities in the 1997 vital rates survey are likely to have been important in creating the spatial structure described in the 1996 density survey.

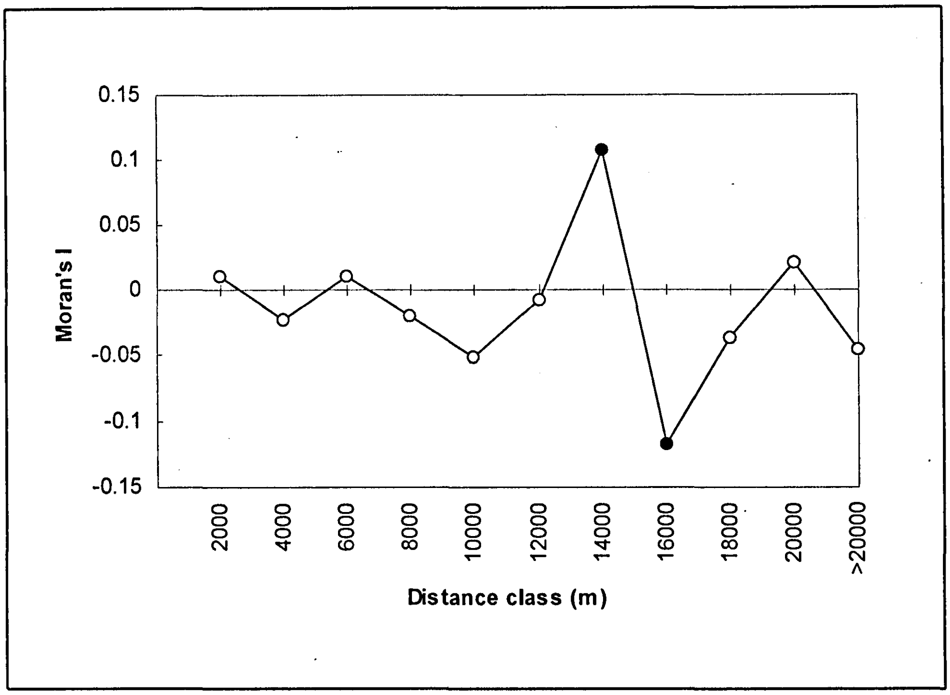
### *Summary of the demographic rates survey*

In total, 8974 leaf-mines were dissected and the fates of their occupants are summarised in table 3.1. Although the selection of sites was not random with respect to habitat type and location, these figures give an impression of the relative importance of the different mortalities across the study area.

Mortalities are largely sequential in this system (Chapter 2). Therefore, the number of mines that can potentially succumb to a particular mortality is the total number of mines minus the number killed before that mortality has a chance to act. This smaller proportion of mines is used to calculate apparent mortality (Bellows *et al.*, 1992) as opposed to real mortality (which is the proportion of all mines). In all of the analyses below, apparent mortalities are used unless otherwise stated. An advantage of using this method of reporting mortalities is that they reflect the real 'attack rate' from any particular mortality source and allow statistical comparisons between mortalities to be made.

Across all sites, larval parasitism by *Chrysocharis gemma* and miscellaneous larval mortality were the most important sources of mortality of *P. ilicis* mines. On

a)



b)

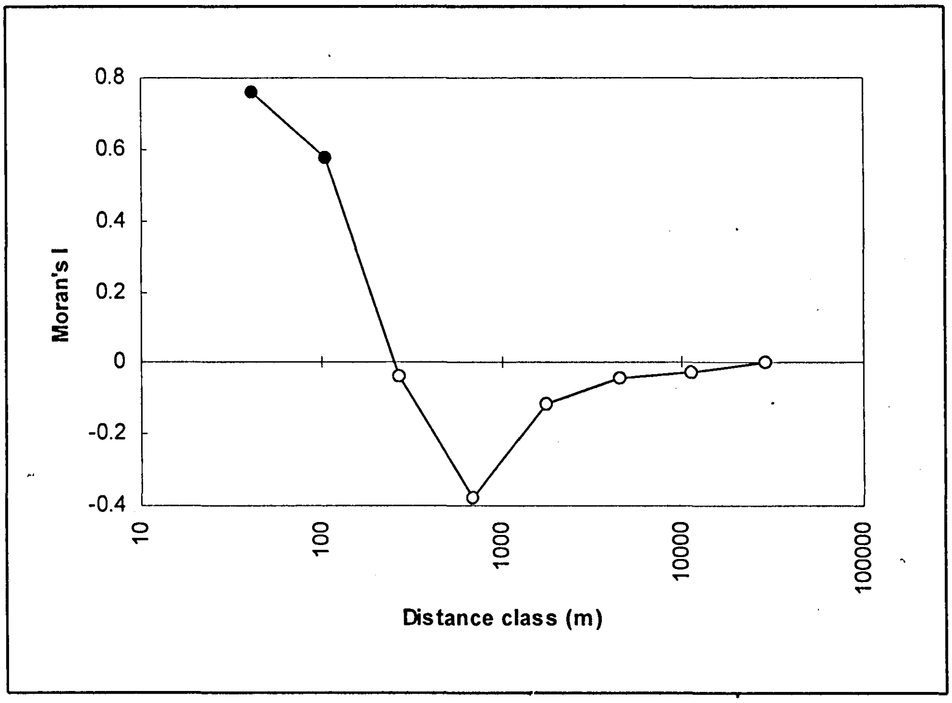
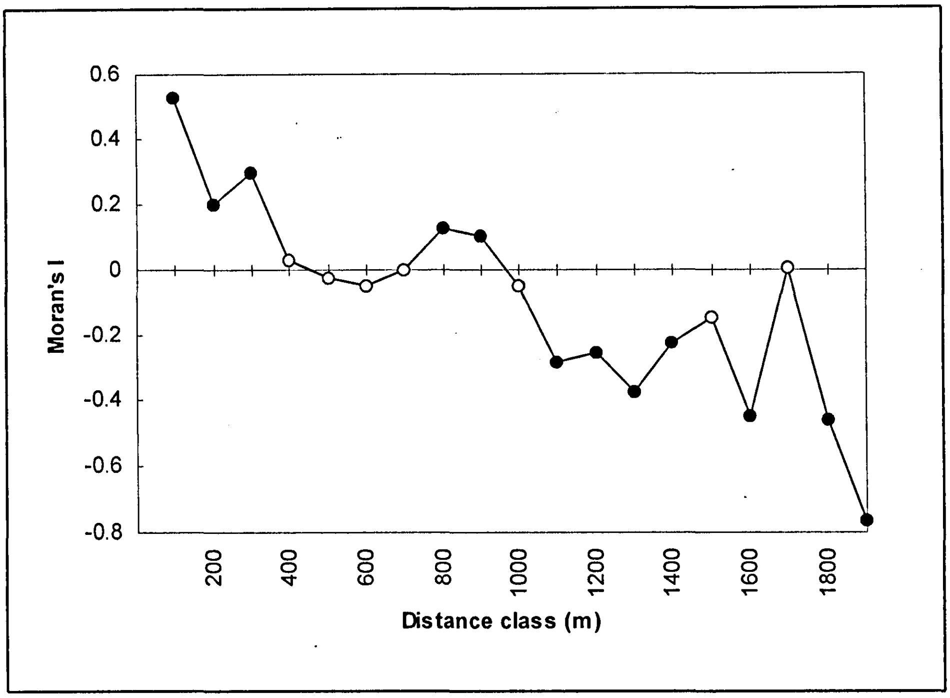


Figure 3.6 Spatial correlograms of bird predation from the Sheffield density survey. Scale on the *x*-axis represents maximum distance in each distance class. Filled circles indicate significant values for Moran's I ($p < 0.05$). Overall correlogram significance was tested using Bonferroni's correction for multiple comparisons. (a) Equal distance classes ($n = 104$; overall significance: $n = 11$, $p < 0.05$) (b) Logarithmic distance classes ($n = 104$; overall significance: $n = 10$, $p < 0.01$.

a)



b)

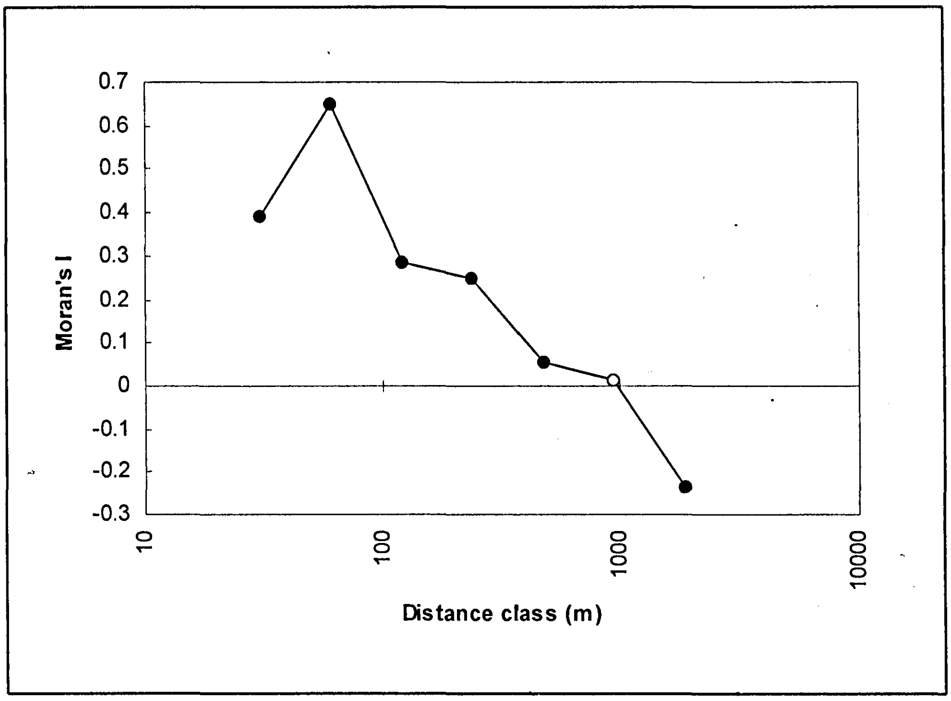


Figure 3.7 Spatial correlograms of bird predation from the Ecclesall Wood survey. Scale on the $x$-axis represents maximum distance in each distance class. Filled circles indicate significant values for Moran's I ($p < 0.05$). Overall correlogram significance was tested using Bonferroni's correction for multiple comparisons. (a) Equal distance classes ($n = 100$; overall significance: $n = 19$, $p < 0.01$) (b) Logarithmic distance classes ($n = 100$; overall significance: $n = 7$, $p < 0.01$.

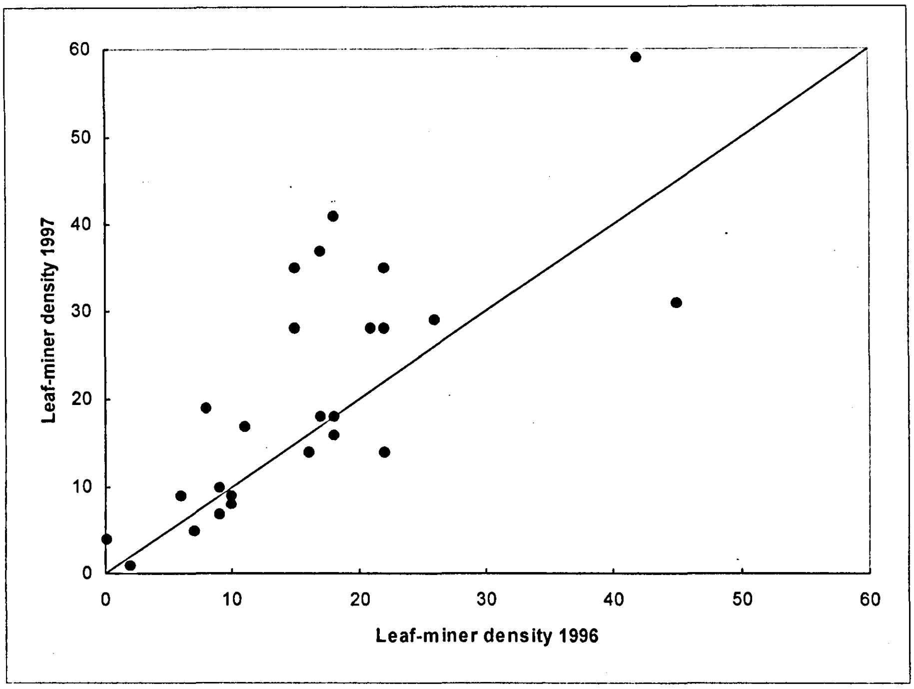


Figure 3.8 Plot of leaf-miner densities in the 1997 survey against leaf-miner densities in the 1996 survey for trees that were sampled in both years ($r = 0.773$, $n = 26$, $p < 0.001$). The line of slope 1indicates that, on average, trees had slightly higher densities in 1997.

| | mean (%) | standard deviation (%) |
|---|---|---|
| miscellaneous larval mortality | 29.5 | 20.1 |
| larval parasitism | 45.1 | 32.2 |
| bird predation | 35.4 | 31.3 |
| pupal parasitism | 47.7 | 27.6 |
| miscellaneous pupal mortality | 28.7 | 31.4 |
| successful emergence | 10.3 | 17.6 |

Table 3.1 Fates of established *Phytomyza ilicis* mines from all 106 trees in the 1997 vital rates survey.

average, only 10.3% of established mines resulted in the successful emergence of an adult fly across all trees. However, levels of mortality and successful emergence were extremely variable between trees and no single mortality was recorded at all of the sites sampled. The population dynamics of *P. ilicis* differs depending upon location.

A principal component analysis (PCA) was performed on the leaf-miner demographic data from all trees to determine the extent to which trees could be characterised in terms of their mortality rates. Sixty percent of the variation in the data was accounted for in the first 2 principal components (table 3.2). The scores and correlations of the original variables with these components are plotted in figure 3.9. The first principal component reflects differences between trees from sites of high successful emergence to those of high miscellaneous larval mortality and larval parasitism.

### *Oviposition and mine density at all sites*

As Valladares & Lawton (1991) point out, there are two ways in which differences in infestation levels between trees can originate. First, female flies may lay different numbers of eggs on different trees and this reflects the final level of infestation. Alternatively, a similar number of eggs are laid on all trees, but egg and/or early larval mortality rates vary before mines become visible. The data from this survey support the first hypothesis. The number of egg scars and mines are strongly correlated (figure 3.10) with approximately 80% of egg-scars giving rise to viable mines. Leaf-miner density levels are therefore the result of selective oviposition by adult flies given a free choice of all trees across the region, a reflection of the capacity of each site to produce gravid females in the previous generation or a combination of these two effects.

### *Oviposition and successful emergence at all sites*

Since there is some evidence that the relative abundance of *P. ilicis* between trees remains the same between years, one might expect that levels of oviposition and successful emergence would be correlated. This would indicate that infestation levels are a reflection of the suitability of the site for *P. ilicis*. However, when all sites are considered together, there is no evidence for this ($r = 0.001$; $n = 106$; $p = 0.99$).

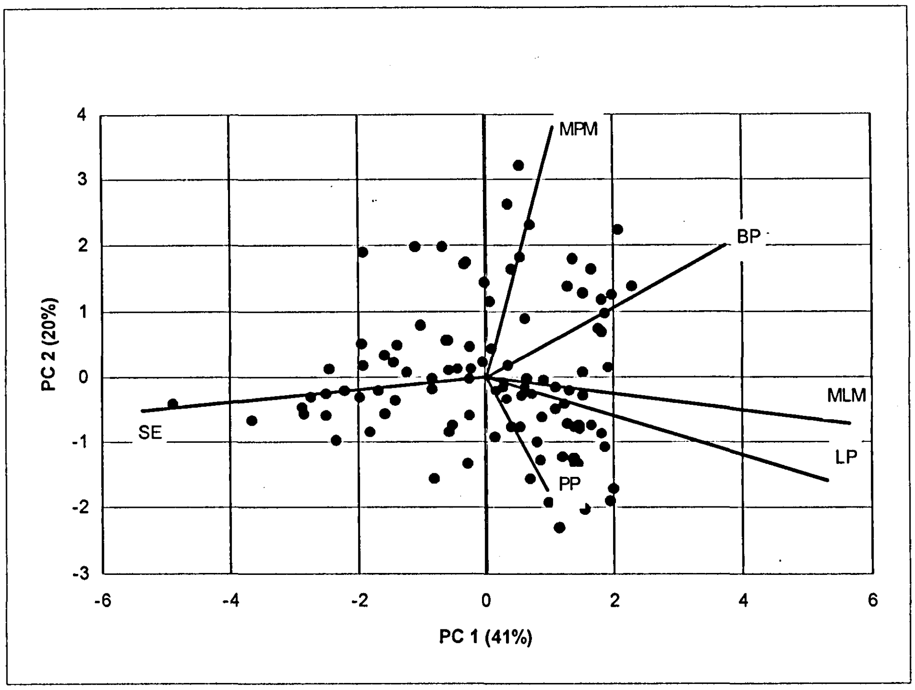


Figure 3.9 Score plot for the first two principal components of variation in the leaf-miner demographic data for all sites, including loadings of the variables against these components. MLM = miscellaneous larval mortality, LP = larval parasitism, BP = bird predation, PP = pupal parasitism, MPM = miscellaneous pupal mortality.

| | PC 1 | PC 2 | PC 3 | PC 4 | PC 5 | PC 6 |
|---|---|---|---|---|---|---|
| **Eigenvalues** | 2.44 | 1.19 | 1.02 | 0.74 | 0.39 | 0.21 |
| Percentage variation | 40.6 | 19.8 | 17.0 | 12.4 | 6.6 | 3.6 |
| Cumulative percentage | 40.6 | 60.4 | 77.4 | 89.8 | 96.4 | 100.0 |
| **Vectors** | | | | | | |
| Miscellaneous larval mortality | 0.552 | -0.146 | -0.256 | -0.050 | 0.429 | -0.650 |
| Larval parasitism | 0.519 | -0.321 | -0.213 | -0.297 | -0.114 | -0.694 |
| Bird predation | 0.363 | 0.400 | 0.049 | 0.789 | 0.156 | 0.241 |
| Miscellaneous pupal mortality | 0.105 | 0.763 | 0.210 | -0.525 | 0.289 | 0.067 |
| Successful emergence | -0.524 | -0.104 | -0.245 | 0.099 | 0.782 | 0.183 |

Table 3.2 Results of the principal component analysis on the demographic components at all sites (see also figure 3.8).

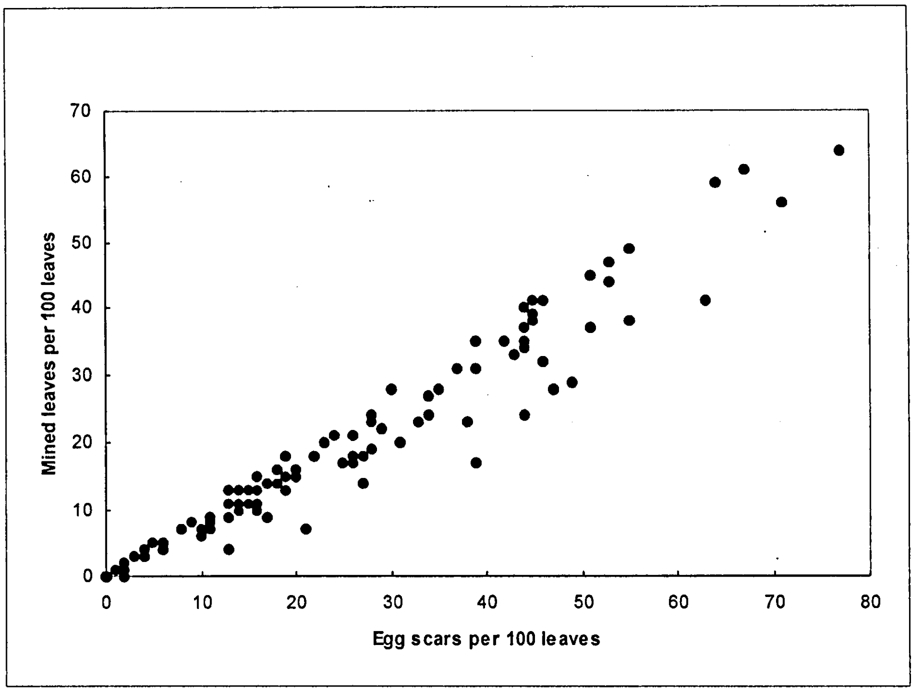


Figure 3.10 Relationship between number of mines per 100 leaves and number of egg-scars per 100 leaves for all trees in the 1997 survey. The equation is $y = 0.80x - 0.61$, $r^2 = 0.94$, $p = 0.00$.

### *Mortality and mine densities*

The lack of concordance between oviposition and leaf-miner performance may be driven by density-dependent reductions in leaf-miner success on more heavily infested trees. This hypothesis was tested by regressing angular transformed mortality data on mine density for all trees in the survey. However, no single mortality factor showed density dependence at this scale.

Tests for density dependence were also conducted within each habitat separately, with similar results for both woodland and open habitat types. However, at urban sites significant relationships were found between mine density and larval parasitism, miscellaneous larval mortality and successful emergence (table 3.3). Surprisingly, both larval parasitism and miscellaneous larval mortality were negatively correlated with mine density.

### *Oviposition and environmental variation*

Potential differences in oviposition (and hence mine) density due to environmental effects were investigated using an Analysis of Covariance (ANCOVA) on angular transformed data for all trees. Habitat type was simplified to 3 levels: woodland habitats, including both coniferous and deciduous woodland types; open habitats, including farmland and moorland; urban habitats, sites within the city limits of the two main towns within the study area, Sheffield and Rotherham. Altitude was treated as a covariate. Significant differences were found between the three habitat types with urban habitats supporting the highest leaf-miner densities and open habitats the lowest. In addition, a significant negative correlation with altitude was detected within these habitats. The analysis accounted for 45% of the variation in oviposition density for all trees ($F_{2,103} = 23.7$, $p = 0.001$). However, as figure 3.11 shows, the variation within habitats around each regression line is extremely high and within the urban habitat the range of altitudes is limited to sites below 250m (urban sites above this altitude did not exist within the study area). Therefore, while there certainly appear to be differences that are affected by environmental conditions the predictive value of this simple model is low and this result should be treated with caution. There are obviously a number of other unmeasured factors that are important in determining oviposition rate. However, one could at least reasonably conclude that the highest leaf-miner densities are to be found in some urban areas and *P. ilicis* may often be absent from high-altitude, open sites.

|  | Regression equation | $r^2$ | significance test |
|---|---|---|---|
| Miscellaneous larval mortality | $y = -0.010x + 0.896$ | 0.522 | $F_{1,11} = 12.036$<br>$p = 0.005$ |
| Larval parasitism | $y = -0.013x + 1.097$ | 0.373 | $F_{1,11} = 6.530$<br>$p = 0.027$ |
| Successful emergence | $y = 0.081x + 0.012$ | 0.610 | $F_{1,11} = 17.175$<br>$p = 0.002$ |

Table 3.3 Significant relationships between angular transformed demographic rates components and mine density for the urban habitat type. These were the only significant relationships found in all of the analyses of density dependence.

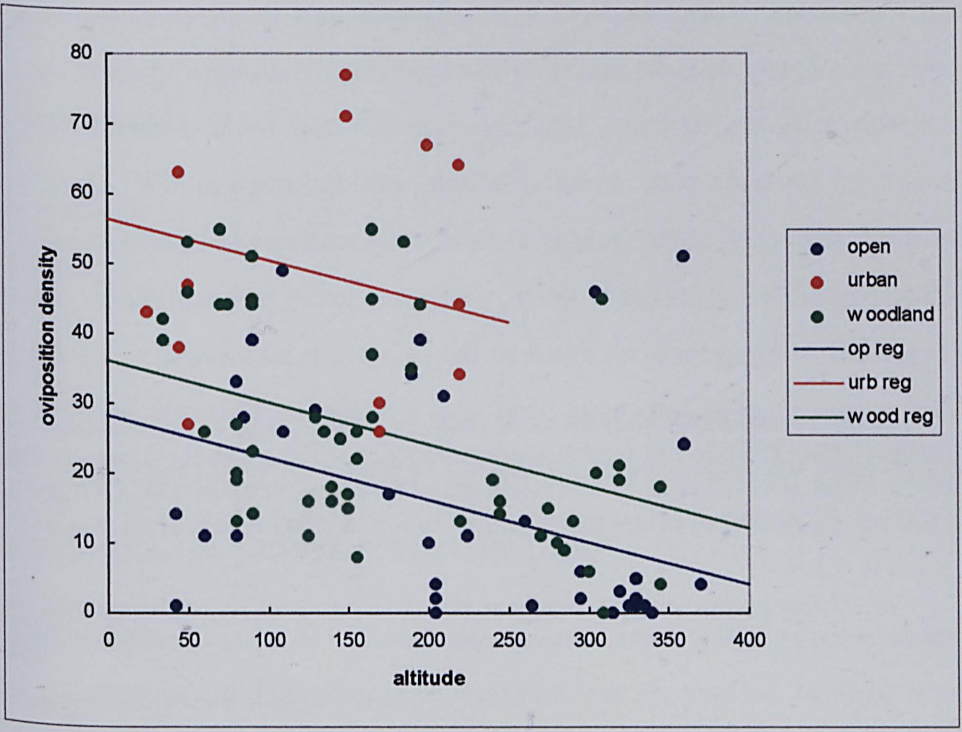


Figure 3.11 Plot of oviposition rate against altitude, incorporating the results of Analysis of Covariance. Refer to text for details

*Differences in demographic components and environmental variation*

As the principal component analysis indicates, the demographic rates components do not vary independently of one another. In particular, a reasonably strong correlation was detected between miscellaneous larval mortality and larval parasitism across all trees ($r = 0.77$, $n = 106$, $p < 0.01$; figure 3.12).

When considering all trees sampled, no obvious consistent relationships between the demographic rates components and altitude were apparent. Therefore direct comparisons between levels of the demographic rates between habitats by ANCOVA could not be carried out. However, regression analyses of angular transformed demographic data on altitude, performed separately for each of the three habitat types, showed that within individual habitats some trends were present (tables 3.4-3.6). Relationships were strongest within the woodland habitat type, where both miscellaneous larval mortality (figure 3.13) and larval parasitism (figure 3.14) were significantly negatively correlated with altitude whereas pupal parasitism (figure 3.15), miscellaneous pupal mortality and successful emergence were positively correlated with altitude. Within open habitats, similar patterns were apparent for miscellaneous larval mortality, larval parasitism and successful emergence. However, both pupal parasitism and miscellaneous pupal mortality were negatively correlated with altitude. No significant relationships were found between the demographic rates and altitude within the urban habitat type but this may have been due to the narrower range of altitudes available and the lower power of the statistical tests since less trees were sampled in this habitat. Bird predation was not correlated with altitude within any of the habitat types.

## Discussion

Despite the high degree of density variation that has been reported for the holly leaf-miner and other phytophagous insect species at a small scale, the results of this study demonstrate that this is not sufficient to completely mask a degree of regional spatial structure which can be identified using spatial autocorrelation techniques and has been illustrated with the interpolated map. The variation between trees is effectively constrained dependent upon where in the region the holly trees are located.

The autocorrelation profiles differed between Ecclesall Wood and the Sheffield Survey. Indeed, one would not expect them to be the same since the correlograms themselves are only derived from the variation in the data taken from the particular area under study and not the total amount of variation possible. This itself reflects spatial

|  | Regression equation | $r^2$ | significance test |
|---|---|---|---|
| Miscellaneous larval mortality | $y = -0.001x + 0.909$ | 0.557 | $F_{1,52} = 65.501$<br>$p < 0.001$ |
| Larval parasitism | $y = -0.002x + 1.298$ | 0.464 | $F_{1,52} = 44.953$<br>$p < 0.001$ |
| Bird predation | $y = -0.000x + 0.228$ | 0.013 | $F_{1,52} = 0.707$<br>$p = 0.404$ |
| Pupal parasitism | $y = 0.002x - 0.008$ | 0.484 | $F_{1,52} = 48.818$<br>$p < 0.001$ |
| Miscellaneous pupal mortality | $y = 0.001x - 0.041$ | 0.405 | $F_{1,52} = 35.334$<br>$p < 0.001$ |
| Successful emergence | $y = 0.001x + 0.019$ | 0.315 | $F_{1,52} = 23.945$<br>$p < 0.001$ |

Table 3.4 Results of regressions of the demographic rate components on altitude within the woodland habitat type.

|  | Regression equation | $r^2$ | significance test |
|---|---|---|---|
| Miscellaneous larval mortality | $y = -0.001x + 0.613$ | 0.177 | $F_{1,37} = 7.964$<br>$p = 0.008$ |
| Larval parasitism | $y = -0.003x + 1.172$ | 0.497 | $F_{1,37} = 36.513$<br>$p < 0.001$ |
| Bird predation | $y = -0.001x + 0.657$ | 0.035 | $F_{1,37} = 1.336$<br>$p = 0.255$ |
| Pupal parasitism | $y = -0.001x + 1.070$ | 0.120 | $F_{1,37} = 5.035$<br>$p = 0.031$ |
| Miscellaneous pupal mortality | $y = -0.001x + 0.630$ | 0.172 | $F_{1,37} = 7.698$<br>$p = 0.009$ |
| Successful emergence | $y = 0.001x + 0.010$ | 0.187 | $F_{1,37} = 8.530$<br>$p = 0.006$ |

Table 3.5 Results of regressions of the demographic rate components on altitude within the open habitat type.

|  | Regression equation | $r^2$ | significance test |
|---|---|---|---|
| Miscellaneous larval mortality | $y = -0.001x + 0.639$ | 0.106 | $F_{1,11} = 1.304$<br>$p = 0.278$ |
| Larval parasitism | $y = -0.001x + 0.814$ | 0.104 | $F_{1,11} = 1.281$<br>$p = 0.282$ |
| Bird predation | $y = -0.003x + 0.924$ | 0.202 | $F_{1,11} = 2.777$<br>$p = 0.124$ |
| Pupal parasitism | $y = 0.000x + 0.686$ | 0.001 | $F_{1,11} = 0.013$<br>$p = 0.912$ |
| Miscellaneous pupal mortality | $y = 0.002x + 0.090$ | 0.290 | $F_{1,11} = 4.497$<br>$p = 0.058$ |
| Successful emergence | $y = 0.001x + 0.173$ | 0.186 | $F_{1,11} = 2.520$<br>$p = 0.141$ |

Table 3.6 Results of regressions of the demographic rate components on altitude within the urban habitat type.

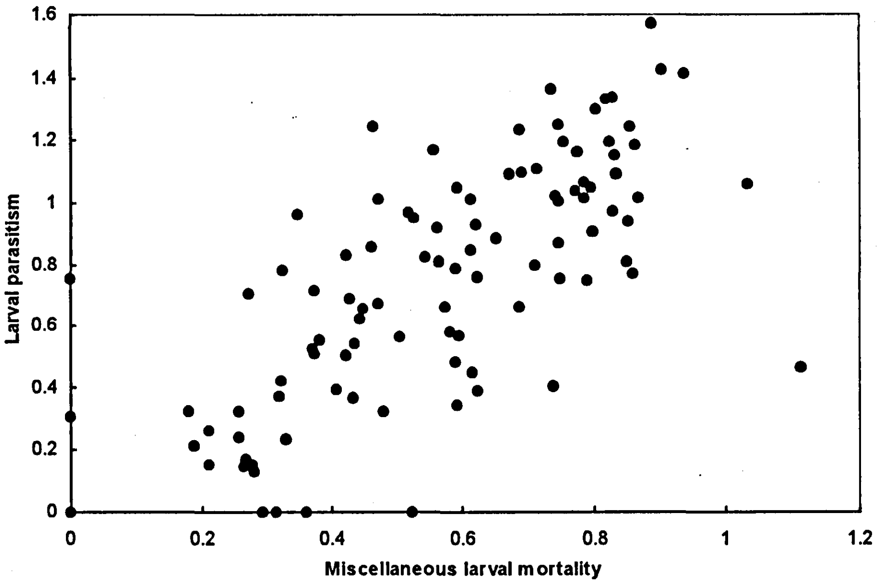


Figure 3.12 Plot of larval parasitism against miscellaneous larval mortality ($r = 0.77$, $p < 0.01$). Data has been angular transformed.

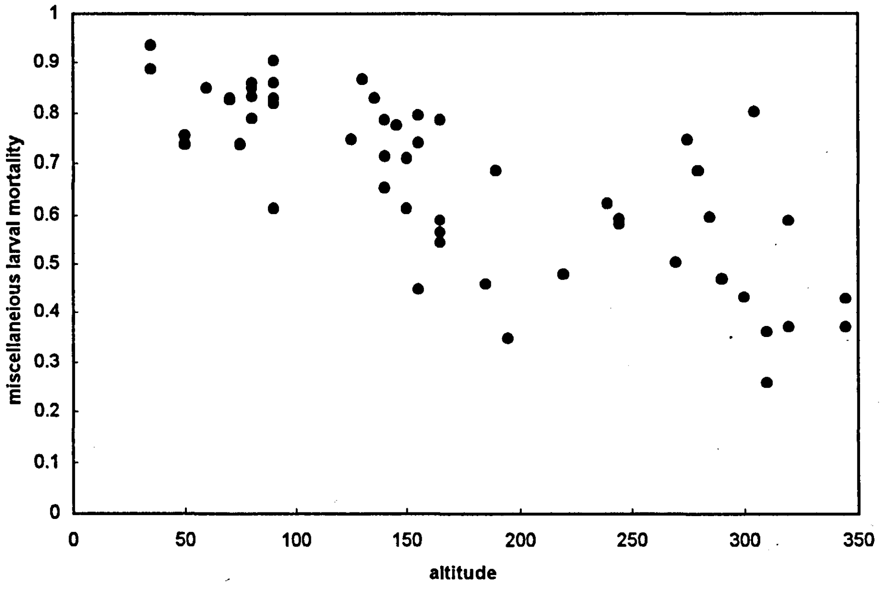


Figure 3.13 Plot of angular transformed miscellaneous larval mortality against altitude within the woodland habitat type.

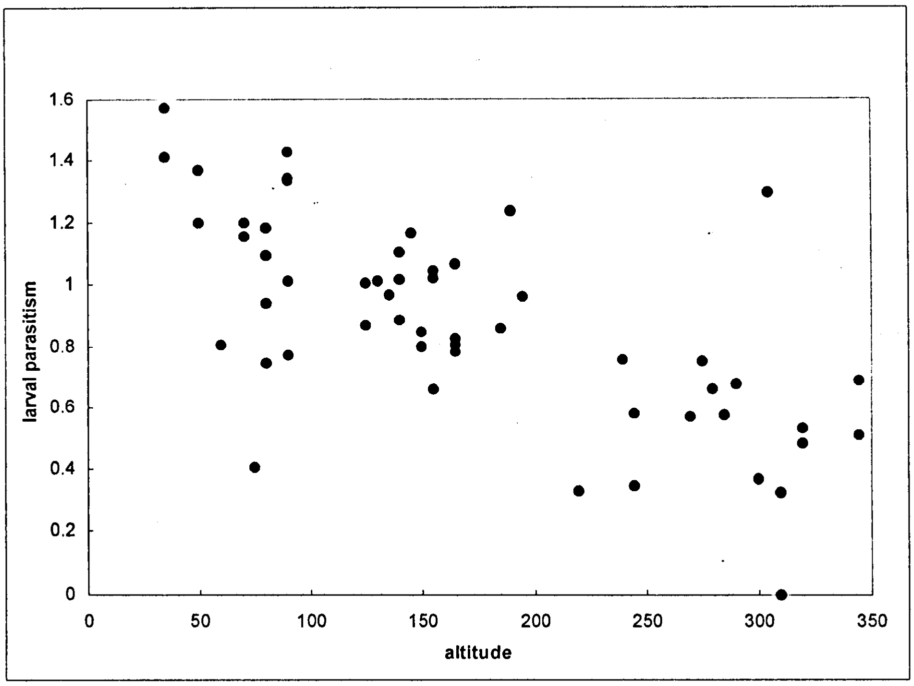


Figure 3.14 Plot of angular transformed larval parasitism against altitude within the woodland habitat type.

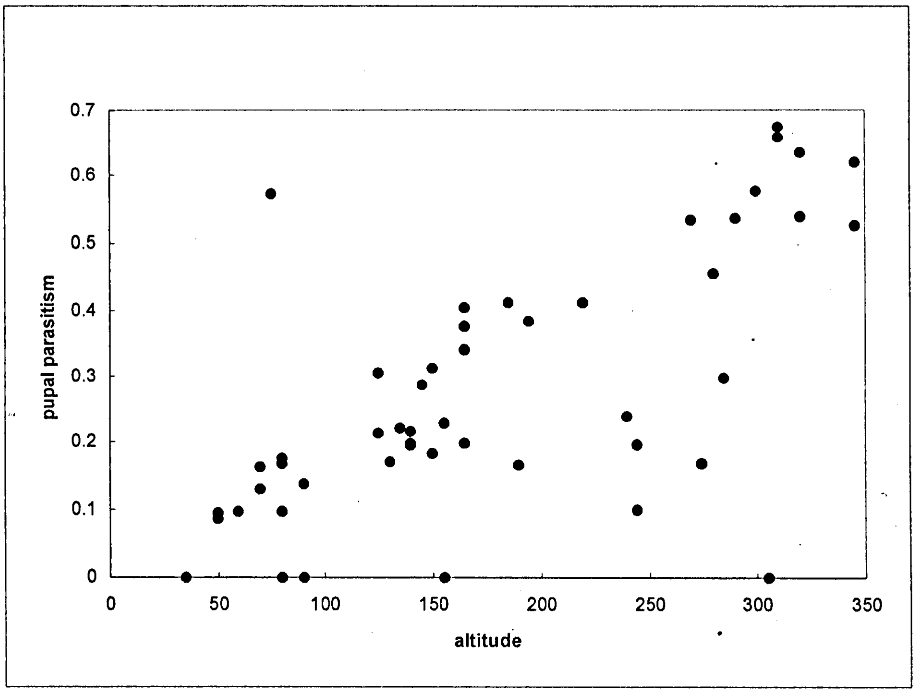


Figure 3.14 Plot of angular transformed pupal parasitism against altitude within the woodland habitat type.

structuring in leaf-miner populations at two different scales. Within Ecclesall Wood significant positive autocorrelation occurs over a scale of a few hundred metres relative to the extent of the sampling, whereas for the Sheffield survey a pattern can be discerned over several kilometres. Patterns of spatial autocorrelation, therefore, may often be dependent upon the areal extent of a survey if spatial structuring processes are operating at a variety of scales. This should be borne in mind when making comparisons between different studies.

No attempt has been made to interpret the distributions of individual points on the correlograms as indicative of more complex spatial patterns than the simple trends discussed. There are two reasons for this. First, such patterns of variation, but not the bold signatures, are rather sensitive to the numbers of distance classes used. Second, one of the assumptions for this method of analysis is second-order stationarity in the data (i.e., the expected mean and variance of the variable over the study area should be constant and finite); the autocorrelation function (Moran's I) must depend only on the length and orientation of the vector between any two points and not on their position in the study area. Over the relatively large area of the Sheffield survey (and most other applications of these techniques in ecology), strictly this assumption is likely to be violated. Thus, whilst Legendre & Fortin (1989) provide examples of correlograms for a number of artificially constructed spatial structures as an aid to interpretation of the common patterns found, it is perhaps wiser not to read any more into those which we present (see Chapters 1 and 4).

The trees sampled during the 1997 Sheffield demographic rates survey were necessarily aggregated, reflecting the availability of some environmental combinations within the study area. Moreover, the site where a tree was sampled did not necessarily reflect the typical habitat type at that location. For example, woodland containing holly was generally common throughout the region but finding holly in open habitats at higher altitudes was difficult. If holly was found in such a habitat then it was sampled whether or not that habitat was typical of the area. For these reasons, spatial autocorrelation analysis of the demographic data would be misleading and so none was attempted. Another related consequence of the aggregation of the sampled trees is that of pseudo-replication due to spatial dependence in the data. This can be an important issue in surveys of this nature, often leading to artificially high estimates of the number of degrees of freedom when conducting statistical tests. Unfortunately, the effect can be difficult to quantify and impossible to avoid completely when there is little choice available for the locations of sites for sampling. However, it is unlikely to create completely erroneous trends in survey data and is generally thought to be more important in experimental studies. One technique to account for spatial dependence, which is applicable under certain sampling regimes, is used in Chapters 4 and 5. The issue is further discussed in Chapter 7, together with a brief overview of another possible solution to the problem.

The ANCOVA revealed that there are significant differences in leaf-miner population densities between the three broad habitat types after accounting for altitude. Trees in urban sites generally experience the largest populations of leaf-miners and trees in open habitats the lowest, although there is considerable variation within these categories as one would expect. Several other studies have also found that trees in urban areas support higher densities of phytophagous insects than in more typical natural habitats (e.g. Hanks & Denno, 1993; Speight *et al.*, 1998). Altitude has often been found to influence phytophagous insect densities. However, its effect seems largely dependent upon the autecology of the species under study. In addition to negative relationships (e.g. Alonso, 1999), positive and unimodal relationships have also been found (e.g. Randall, 1982; Virtanen & Neuvonen, 1999).

As well as variation in leaf-miner local population densities, there is also a great deal of variation in the components of the demographic rates across the region. *P. ilicis* experiences very different population dynamics at different locations. Whilst attempts to categorise trees in terms of different patterns of mortalities has proven difficult in this study, the principal component analysis and correlations between some of the mortalities demonstrate that they do not vary wholly independently of one another at this scale.

The most obvious example of this is the relatively high correlation between miscellaneous larval mortality and larval parasitism. There are two possible reasons for this correlation. First, there may be no causal link between the two mortalities, rather they are affected similarly by environmental variation across the region. Second, it is possible that an important component of miscellaneous larval mortality may be due to attacks by *Chrysocharis gemma*, the larval parasitoid. Cameron (1939) noted that it can be relatively simple to identify parasitised *P. ilicis* larvae even shortly after egg-laying since the leaf-miner larva often becomes 'flaccid and of a pale, dirty hue'. Perhaps many of these larvae die before the parasitoid larva has had sufficient time to become established. While this might seem maladaptive on the part of the parasitoid, it should also be borne in mind that in some localities a high proportion of adult *C. gemma* fail to escape the leaf mine after pupation (Chapter 2). Losses of *C. gemma* during the parasitic phase of its life-cycle may be high. There is an obvious exclusion experiment that could be conducted to test this hypothesis. Mined leaves on a number of trees could be protected from parasitoid attack during the early months of the year (figure 2.1) and levels of miscellaneous larval mortality in this treatment could then be compared to those of unprotected mines on the same trees.

Valladares & Lawton (1991) found a significant negative correlation between miscellaneous larval mortality and larval parasitism for 15 trees at one site. These data are not necessarily inconsistent with the results presented here since the scale of the investigations are very different. A negative relationship between miscellaneous larval mortality and larval parasitism within a site may reflect differences in the suitability of *P. ilicis* larvae as hosts for the parasitoid between trees, if both are a measure of attack by *C. gemma*, whereas a positive relationship at a larger scale may reflect that only sites with significant levels of larval parasitism are likely to have high levels of miscellaneous larval mortality too.

Within woodland habitats and open habitats, levels of larval parasitism decrease with altitude. One reason for this may be that *C. gemma* overwinters as an adult in holly (Cameron, 1939; Hansson, 1985) and that conditions during these months cause significant mortality. Another reason may be that densities of holly are often lower at higher altitudes in the region (personal observation) and that there is not a sufficient number of trees and hence leaf-miners to maintain large larval parasite populations. This hypothesis could be tested by comparing levels of larval parasitism at high altitude sites of different tree densities, although considerable care would be required in order to account for a number of other potentially confounding factors.

Pupal mortality, from parasitoids and other unknown factors, is positively correlated with altitude in woodland habitats and so, to some extent, replaces larval parasitism with increasing elevation. However, in open habitats, the sources of pupal mortality are significantly negatively correlated with altitude, although the relationships are weaker. Not only do the patterns of some demographic rates components differ between habitat types, but also the way in which these components respond to the same continuous environmental variable within these habitats.

None of the components of the demographic rates measured during this study were significantly correlated with local population densities when all trees were included in the analyses or in woodland and open habitats. Perhaps most surprising, is that although levels of successful emergence increase with altitude, oviposition and mine densities generally decrease. Whilst independence between successful emergence and infestation level at small spatial scales is not unusual (Thompson, 1988), this result is rather counter-intuitive. It could arise for a number of possible reasons.

First, even though the *per capita* rate of successful emergence may not be highest at high density sites, the absolute number of adult flies after pupation is higher at these sites. This would be likely to result in similarly high levels of infestation in the following year. *Per capita* rates of successful emergence may not be a good indicator of contemporary habitat suitability. If a local population is near its carrying capacity (equilibrium level) then density dependent effects may serve to reduce birth rates and increase death rates. If this hypothesis is correct, there will be important implications for our ability to understand range structure in terms of differences in demographic rates by surveying methods alone. This issue is further discussed in Chapter 7.

Second, despite the ease with which a large proportion of the life-cycle of the holly leaf-miner may be monitored, there are currently no data for the adult stage of the life-cycle. It is possible that this comparatively short period is a critical stage in determining local population densities, perhaps through a number of mechanisms, including differences in adult mortality, fecundity and dispersal. In addition, the relationship between the phenology of the host plant and the oviposition behaviour of the adult flies may be important, although Valladares & Lawton (1991) note that levels of adult activity correlated well with the presence of leaves suitable for oviposition within the location of their study. Experimental work required to investigate these factors might prove difficult since measuring the mortality and dispersal of small free-living insects is not easy.

In contrast to the general independence of the demographic rate components from infestation levels found during the survey, larval parasitism and miscellaneous larval mortality on urban trees were found to be negatively correlated with mine density and this was accompanied by a concomitant increase in successful emergence. This effect was reasonably strong. Indeed, during the course of the field work it was apparent that where levels of larval parasitism were particularly low in urban environments, mine densities reached very high levels. These trees seemed to be in the most built-up regions and next to the busiest roads, although this aspect of environmental variation was not formally measured. Both Hanks & Denno (1993) and Speight *et al.* (1998) give evidence that elevated densities of phytophagous insects in urban environments can be the result of both a release from natural enemies and environmentally induced stress upon the host plant. To some extent, the Sheffield data are consistent with these results.

It is interesting to note that if the pattern found for this small sample is typical of the system, then regulation by a natural enemy will also appear differently depending upon spatial scale. Both Valladares & Lawton (1991) and Heads & Lawton (1983) found significant positive correlations between larval parasitism and leaf-miner density within a habitat patch. Density dependence at this scale is likely to reflect a behavioural (aggregative) response to leaf-miner populations. However, at larger scales differences in the numbers of *C. gemma* present between sites is more likely to be due to variation in habitat suitability. A negative relationship with leaf-miner density at these scales would be indicative of a regulatory role by the parasitoid. This underlines the issues raised by Heads & Lawton (1983) concerning scale and the detection of density dependence but over a much larger range of scales. However, a more intensive survey of urban holly trees would be required before any firm conclusions could be drawn from this relatively small sample.

To summarise, a significant component of the variation in infestation levels of holly leaf-miners is spatially structured and is correlated with environmental variation at a regional scale. Broad differences between habitat types and trends along altitudinal gradients have been discovered during this study. The population dynamics of *P. ilicis* appears to be qualitatively different between habitats. Larval parasitism may play a regulatory role in some urban habitats but the effects appear to be overshadowed by unmeasured (and presumably abiotic) factors in others. Comparison of the results to earlier within-habitat studies of *P. ilicis* also emphasise the importance of spatial scale in determining the patterns that may be apparent. The fact that population dynamics may differ between habitats is an important consideration when attempting to characterise the spatial structure of local population densities at a large scale. In the next chapters that examine the whole of the geographic range of *P. ilicis*, this information is taken into account and some attempt is made to control for it by choosing only one habitat type that is likely to be well represented throughout the range.

Hanks & Denno (1993) state that the multiple factor approach to explaining herbivore distribution and abundance has proven more effective than examining single factors. That is, the combined influence of host plant suitability and natural enemies often better explains the abundance of some herbivores than do either host plant condition or natural enemies alone. The data presented here are consistent with these ideas if the concept of host plant suitability can be expanded to include its location with respect to regional abiotic factors. However, the Sheffield data may also be consistent with a metapopulation approach in some habitats where consideration of the distance between trees and the size of the habitat patch may be appropriate. Whilst this study has succeeded in identifying important patterns in the face of considerable variation and complexity, ultimately extensive surveys of this nature can only generate hypotheses which then require testing by more focused, intensive surveys and experiments.

# The structure of the geographic range of the holly leaf-miner

> *"The populations near the centre of the species range are usually completely contiguous; they also show a relatively high population density (per unit area), and greater individual variation than is average for populations of the species. Peripheral populations tend to have opposite values for each of the three characteristics (frequent isolation, low population density, low individual variation). This broad generalisation, long accepted by naturalists, is based on general observations but lacks, so far, detailed quantitative support."*
>
> Mayr (1963)

Chapters four and five present the results of a survey of *P. ilicis* local population densities and demographic rates across the whole of its natural geographic range. The chapters are copies of manuscripts prepared for submission to the journal *Oikos*. They have been reformatted to remain consistent with this thesis and the acknowledgements and references have also been moved to the appropriate sections. Since each of these manuscripts is intended to be a stand alone article, there has necessarily been some repetition of the methodology, though this has been kept to a minimum. Further discussion of some of the issues raised in these chapters and their relationships with the other results presented in this thesis can be found in Chapter 7.

# Chapter 4

# The population density structure of the geographic range of the holly leaf-miner

Andrew M. Brewer* and Kevin J. Gaston

Biodiversity and Macroecology Group, Department of Animal & Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

*Contact details: Tel: 0114 2220034, Fax: 0114 2220002
e-mail: bop96amb@sheffield.ac.uk

## Abstract

The local population density structure of a phytophagous insect, the holly leaf-miner *Phytomyza ilicis* Curtis, was examined across its natural geographic range in Europe. The frequency distribution of the number of sample sites at which the leaf-miner attained different densities per tree was strongly right skewed, with the species being absent from a large number of sites at which its host occurred, particularly in southern regions. There was a decline in the spatial autocorrelation of leaf-miner densities with increasing distance between sample sites, with negative autocorrelation at long lags resulting in part from high densities being attained at the north-eastern range limits and low densities at the southern range limits. Partial regression analysis was used to model leaf-miner densities in terms of the spatial position within the geographic range and the broad climate experienced at the sample localities, and accounted for between 40 and 65 % of the variation in densities, dependant upon how the leaf-miner's geographic range was defined. Whilst overall the results are at odds with common and intuitively appealing assertions about the abundance structure of geographic ranges, they can readily be interpreted in terms of a simple modification of a general model of such structures.

## Introduction

It has long been accepted that, at least to a first approximation, in general local population densities of a species can be regarded as tending to peak towards the centre of its geographic range and to decline towards the periphery (Shelford, 1911; Kendeigh, 1974; Hengeveld *et al.*, 1979; Hengeveld, 1989; Lawton, 1993; Safriel *et al.*, 1994; Brown *et al.*, 1996; Maurer, 1999). Indeed, a number of theoretical models have been developed to explain why this should be so (Brown, 1984; Williams, 1988; Hengeveld, 1989; Maurer & Brown, 1989; Hall *et al.*, 1992; Maurer, 1999), and the pattern provides a central presumption of many discussions in biogeography and macroecology (e.g. Hengeveld, 1990; Brown, 1995; Brown & Lomolino, 1998; Maurer, 1999). Perhaps surprisingly, however, it finds support from rather few published empirical studies, even in an amended form that recognises that low abundances may be found across the full breadth of a geographic range but maintains that mean or maximum densities will still tend to decrease towards range limits (see Brown *et al.*, 1995; Enquist *et al.*, 1995; Brown & Lomolino 1998). Mayr (1963) observed that the pattern itself had long been accepted by naturalists because of its intuitive appeal but that it was based on a general observation that lacked detailed quantitative support. Relatively little has changed since then.

Many assertions regarding patterns of local abundance are based on impressions derived from visual inspection of maps of abundance surfaces across all or parts of the geographic range of a species. These are problematic on the grounds that the human visual system is adept at seeing regularities in images which may not be present, particularly when there is some preconception as to what will be found. A variety of more formal analyses have suggested support for a decline in population abundances towards the edges of geographic ranges (e.g. McClure & Price, 1976; Bock, Bock & Lepthien, 1977; Hengeveld & Haeck, 1981, 1982; Randall, 1982; Emlen *et al.*, 1986; Bart & Klosiewski, 1989; Svensson, 1992; Tellería & Santos, 1993; Maurer, 1994; Maurer & Villard, 1994; Whitcomb *et al.*, 1994; Brown *et al.*, 1995; Carey *et al.*, 1995; Curnutt *et al.*, 1996). A number of others have failed to find convincing evidence for the pattern (e.g. Griggs, 1914; Rapoport, 1982; Brussard, 1984; Carter & Prince, 1985; Wiens, 1989; Woods & Davis, 1989; Blackburn *et al.*, 1999).

A paucity of high quality data, the use of very indirect measures of abundance, and complexities in the application of analytical techniques render the results of some of these studies difficult to interpret (see Grice *et al.*, 1986; Wiens, 1989; Gaston, 1994). In addition, a number of the analyses consider rather small portions of the geographic ranges of species, raising concerns as to how well they reflect the pattern of change in local abundances across entire geographic ranges. The principal exceptions, and some of the studies providing the most convincing evidence for a peak in abundances towards the centre of geographic ranges, are those for North American birds, for which interpolated maps have now been published of variation in the abundances of both summering and wintering species across much of the continent, from data generated by the Breeding Bird Survey and Christmas Bird Counts (Root, 1988; Price *et al.*, 1995). Range-wide studies are required for other taxa on other continents.

In this paper we examine the local abundance structure of a phytophagous insect, the holly leaf-miner *Phytomyza ilicis* Curtis (Diptera: Agromyzidae), across the bulk of its natural overall geographic range, in Europe. Previous studies of this species have revealed a great deal of variation in densities at small spatial scales (e.g. within a single host plant or between host plants within a small locality: Heads & Lawton, 1983; Valladares & Lawton, 1991; McGeoch and Gaston, 2000). However, it is not known how this variability may contribute to or obscure any spatial structure that may be present at larger scales.

## The study system

European (or English) holly *Ilex aquifolium* L. is a relatively small, dioecious, evergreen tree. It has distinctive, dark green, glossy leaves that are usually spiny with a relatively thick cuticle. Its natural range extends throughout north-western, central and southern Europe (Figure 1; Peterken & Lloyd, 1967; Hultén & Fries, 1986). It can also be found less commonly in scattered localities in parts of North Africa and has been reported as having a narrow band of distribution extending into Asia Minor although there is some doubt as to whether this is indeed so. European holly was also introduced to North America and Canada in the late nineteenth century where it has been extensively cultivated (Cameron, 1939).

*Phytomyza ilicis*, the holly leaf-miner, is the most common insect herbivore of European holly. It is strictly monophagous, so its geographic range is ultimately limited by the availability of holly trees. The life histories of *P. ilicis* and its natural enemies in Britain have been described in detail by Cameron (1939) and Lewis & Taylor (1967). Put briefly, the holly leaf-miner exhibits a univoltine life-cycle. Eggs are laid in June on new holly leaves (the tree has just one flush of new leaves per year) into the base of the underside of the midrib. The larvae eat through the midrib and enter the outer parenchyma of the leaf lamina during the autumn. They feed throughout the following winter months and pupate in the mine in March, emerging from the leaf as adults in late May or June.

Several factors contribute to making the holly leaf-miner an ideal model system for large scale population surveys. First, the larvae produce large, irregular mines that are clearly visible. Moreover, since holly is evergreen and its leaves typically stay on the tree for up to five years (Peterken & Lloyd, 1967), leaves of a particular age can be sampled for quite a long period after the life-cycle of the leaf-miner is completed to give an estimate of the density for that generation. Therefore local population densities of this species can be relatively easily censused and mapped.

## Methods

### *Data collection*

Between June and December 1998 we conducted a survey of leaf-miner densities throughout most of the natural range of holly. Our aim was to obtain a good coverage, sampling both the centre and the range edges. The survey was organised as a series of routes of varying duration: Spain and Portugal; Norway; Germany, Italy and France; United Kingdom; Ireland; Greece (Figure 4.1). During each survey we sampled leaf-miner densities at as many sites as time allowed.

Previous surveys conducted at a much smaller spatial scale have shown that broad differences in habitat type can be an important determinant of holly leaf-miner density. Thus, for example, holly trees in urban areas typically support higher mine densities than those in deciduous woodland (A. M. Brewer, unpublished data). Such habitat effects could potentially seriously confound the results of a large scale survey. It would have been impractical to attempt to obtain representative samples from all possible habitats in every part of the range, and strictly random sampling could have led to census results from different regions largely reflecting differences between habitat types. Therefore, we kept habitat type broadly constant by sampling deciduous woodland, the predominant habitat for holly, wherever possible.

Leaf-miner density estimates were taken from 10 trees at each sampling site, or as many as possible where less than 10 trees were present. Following Heads & Lawton (1983) and Valladares & Lawton (1991), density was estimated on each tree by haphazardly sampling 200 one year-old leaves (or all the leaves of this age if less than 200 were present on the tree) from all around the canopy, between ground level and a height of approximately 2m. This protocol gives an estimate of density for the previous

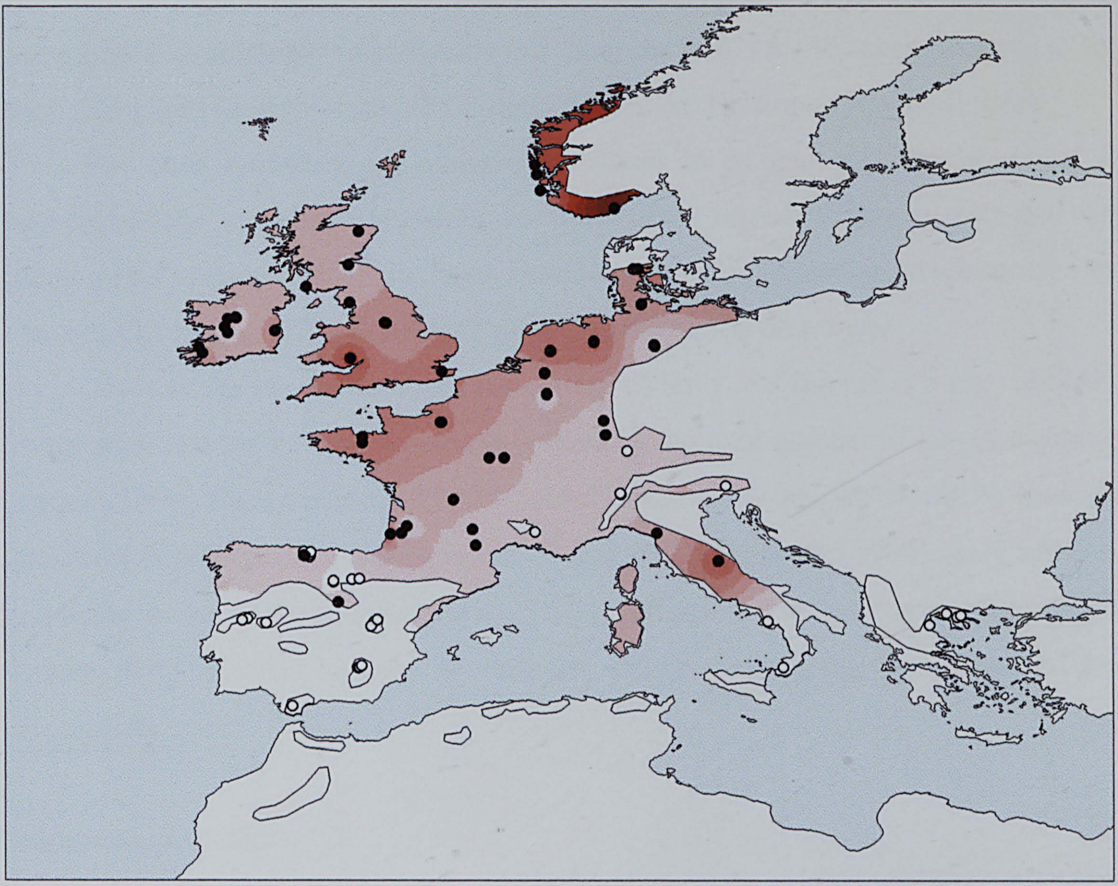


Figure 4.1 Interpolated surface of leaf-miner densities within the range of European holly. Geographic range of holly compiled from Hultén & Fries (1986) and our own observations. Points on map indicate the positions of sites sampled. Filled circles indicate sites where leaf-miners were present. Open circles indicate where leaf-miners were absent.

growing season. Leaves of a suitable age can readily be distinguished by their position between the annual nodes on the branch. For each leaf, the number of mines present were recorded, although typically only one mine per leaf is found except in areas of very high leaf-miner density (Heads & Lawton 1983, A. M. Brewer unpublished data).

For trees on which no mines were found in samples of 200 leaves of the appropriate age, an exhaustive search was made of the rest of the tree. In very rare cases some mines would be found, in which case densities were treated as an arbitrarily small number (0.1 mines per 200 leaves). This occurred for only four trees.

The unit for the analyses was the density of holly leaf-miners at a site, calculated as the total number of mines recorded divided by the total number of leaves examined, across all trees that were sampled. The spatial relations of sites, which can be viewed as one form of potential non-independence of data points (Legendre 1993), were addressed explicitly in the analyses (see below). The geographic location (expressed as decimal degrees longitude and latitude) of each site and its altitude (m) were obtained from appropriate maps.

### *Meteorological data*

Regional climatic data over a 0.5°×0.5° grid were obtained from a coverage of twentieth century terrestrial surface climate (New *et al.*, 2000). From this coverage, mean values for winter and summer temperatures, precipitation and humidity were calculated from data for the 10 year period between 1986 and 1995 inclusive.

### *Data analysis*

A number of geostatistical techniques are available for handling data sets where each sample can be represented as a point estimate in two-dimensional space. We briefly outline those employed here, but for further details see Legendre & Fortin (1989), Rossi *et al.* (1992), Legendre (1993), Maurer (1994) and Legendre & Legendre (1998).

#### Spatial autocorrelation

Spatial autocorrelation analysis (Cliff & Ord, 1973) was used to establish whether local population densities of *P. ilicis* exhibited any simple spatial structure across the geographic range. Moran's I was calculated for 15 equal distance intervals, and spatial correlograms were produced and tested for significant spatial dependence in the data.

Since the survey covered a relatively large geographic area, site coordinates were not treated as Cartesian coordinates when measuring distances between them. Instead, distances along great circles were calculated to take into account the curvature of the earth's surface. Bonferroni's correction for multiple comparisons was used when assessing overall correlogram significance.

### Interpolation

Leaf-miner densities were interpolated across space to estimate the overall local density structure of the geographic range of the species. Densities were entered into a geographic information system (GIS; ArcInfo and ArcView, Environmental Systems Research Institute Inc., 1997) as point coverages, and interpolated using an inverse distance weighting technique (IDW) to create a grid comprising 0.1°×0.1° cells over the study area. Following the recommendations of Isaaks & Srivastava (1989), twelve nearest neighbours were used to estimate the values with weights inversely proportional to the square of the distance from the estimated cell.

### Partial regression analysis

Partial regression analysis is one of several techniques that can be used when attempting to model data showing spatial dependence (see Clifford & Richardson, 1985; Clifford *et al.*, 1989; Getis, 1990; Legendre, 1993; Carroll & Pearson, 1998; Legendre & Legendre, 1998, for applications and reviews of the available techniques). Following the methodology suggested by Legendre (1993), we used the method to estimate how much variation in leaf-miner densities can be attributed to regional climatic variation and altitude once the effect of spatial location has been taken into account. The spatial component of the holly leaf-miner density data was modelled using a third-order polynomial of the form:

$$f(x,y) = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3$$

where $x$ and $y$ represent longitude and latitude respectively. This expression is sufficient to extract any linear gradients from the data as well as more complex features such as patches or gaps (Legendre, 1990; Borcard *et al.*, 1992). Significant terms, determined by conducting a stepwise regression of leaf-miner densities on a matrix of all of the terms in the expression, were retained to construct a new matrix of spatial variables to be used in the subsequent analysis. To ensure that this matrix provided an adequate description of the spatial structure of leaf-miner densities, the residuals from the regression were checked for spatial dependence using autocorrelation analysis. Similarly, leaf-miner densities were regressed onto a matrix containing the climate and altitude data and significant terms determined after stepwise regression by elimination were retained.

The combined effect of both the environmental and spatial variables on leaf-miner densities was calculated by multiple regression of leaf-miner densities onto both sets of predictive variables combined. The explanatory potential of the environmental variables, after correcting for spatial dependence, was calculated by measuring the change in deviance explained by the regression model after the environmental variables were removed. This fraction was then tested to determine whether the change in deviance was statistically significant.

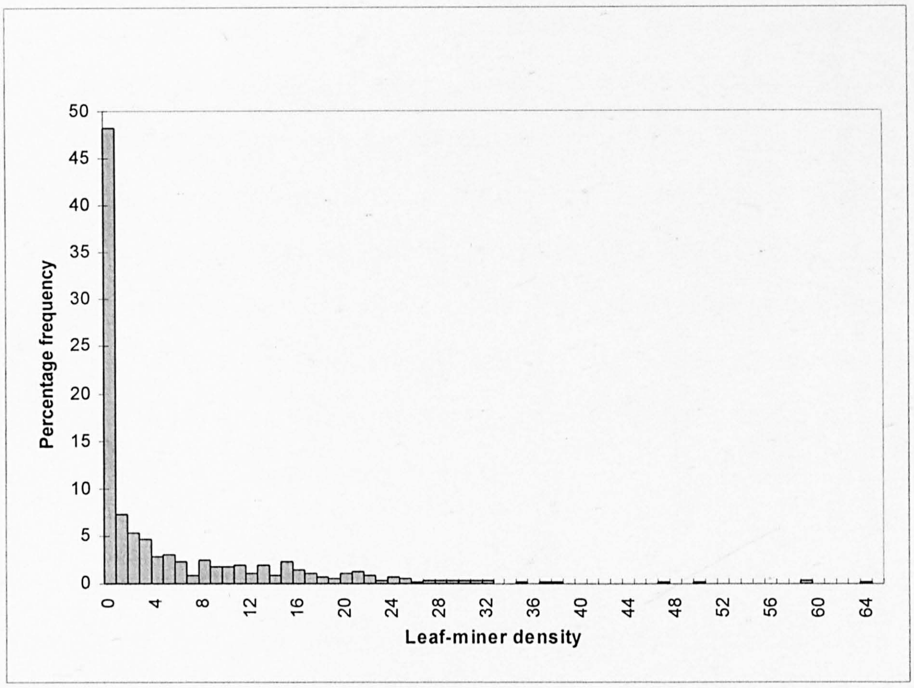
At the end of the partial regression analysis, variation in leaf-miner densities could be divided into four components: (a) non-spatial environmental - the fraction that can be explained by the environmental variables independent of any spatial structure; (b) spatially structured environmental - spatial structuring in leaf-miner densities that is shared with the environmental data; (c) non-environmental spatial - spatial structure in leaf-miner densities that is not shared with the environmental variables; and (d) unexplained (residual) variation. These components and their associated probabilities allow several hypotheses about the causal relationships between environmental variation, spatial position and leaf-miner density to be tested.

The method employed here is essentially the same as Legendre's (1993) third extension to partial regression analysis with some minor modifications. Since our leaf-miner abundance data were in the form of proportions (number of mines per number of leaves sampled), binomial errors were modelled. This results in proportions of deviance explained which are analogous to the partial $r^2$ statistics at the end of the analysis. In addition, during the stepwise regressions, any overdispersion in the residual deviance was corrected for using Williams' procedure where appropriate (Collett 1991, Crawley 1993) before testing for parameter significance. All of the partial regression analyses were conducted using GLIM version 3.77.

## Results

A total of 703 holly trees were sampled during the course of the survey at 96 sites. This resulted in the examination of 133685 leaves. A total of 7284 mines of *P. ilicis* were found, giving a mean infestation rate of approximately 5.5% across all trees. However, no mines were recorded on nearly 50% of the trees sampled (Figure 4.2a).

a)



b)

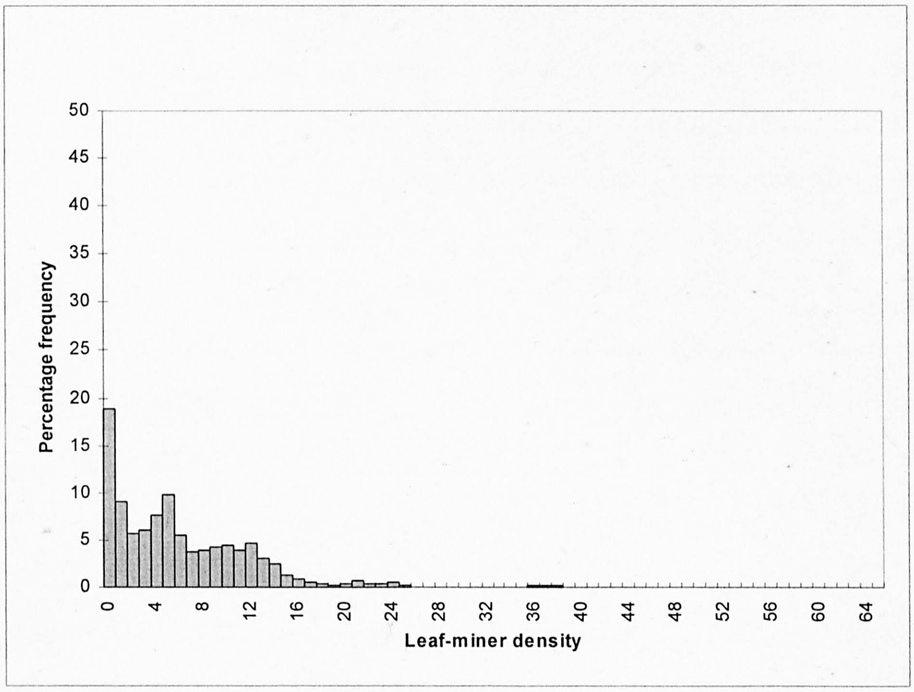


Figure 4.2 Frequency histograms of leaf-miner densities across its natural geographic range (a) Leaf-miner densities on trees sampled (b) Leaf-miner densities taken from the interpolated surface (figure 4.1).

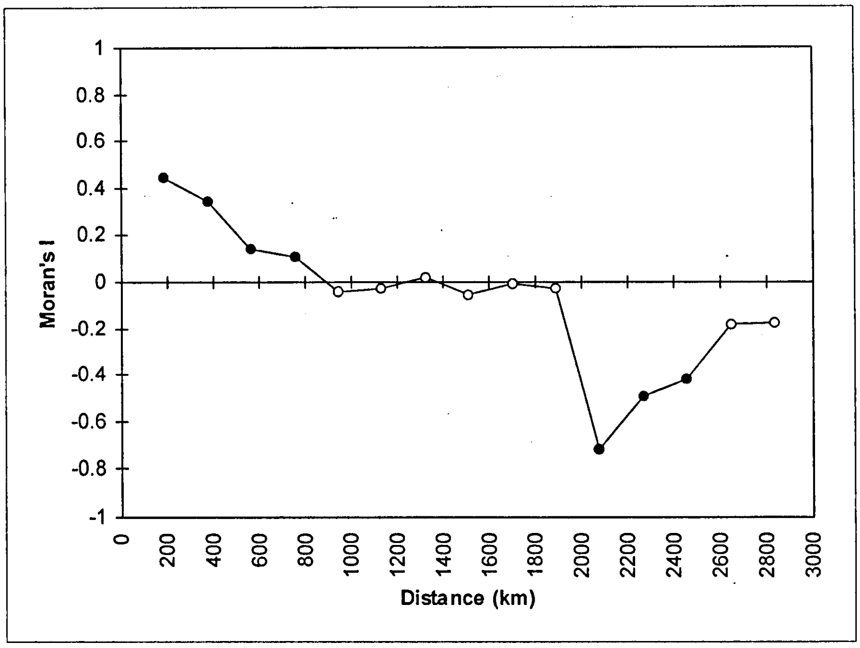
The holly leaf-miner was almost entirely absent from large regions of the geographic range of holly, particularly in Spain, Portugal, Greece and most of Italy. This was true even though in some of these areas (e.g. parts of Spain) holly itself is locally abundant. The overall frequency distributions of local densities at sample sites exhibits a strong right-hand skew, with most sites at which the miner was found being occupied at relatively low densities (Figure 4.2a). Some of the highest densities found were at two sites in Italy, but the peak densities were consistently found in the extreme north-east of the range, along the Norwegian coast. In some areas where holly leaf-miners attained high densities holly itself was scarce (e.g. Norway).

*Spatial autocorrelation*

Since not all sites sampled supported populations of the holly leaf-miner, and it is rather difficult to define the boundaries of the range of a species exactly (Gaston 1994, Blackburn *et al.* 1999), there is some difficulty in deciding which sites to include in a spatial autocorrelation analysis. If all sites where the leaf-miner was absent are omitted from the analysis, we risk losing information about the spatial structure of the local populations since some of these sites lie well within the outermost boundaries of the geographic occurrence of the species. However, incorporating data from all of the sites sampled would result in the inclusion of a large number of sites from which *P. ilicis* was entirely absent, many of which are separated from one another by relatively short distances, and represent regions of the geographic range of holly which lie outside that of the leaf-miner. This could artificially inflate the value of Moran's I obtained for short separation distances (lags). As a pragmatic solution, we present and compare both correlograms, one for all sites sampled and one which omits those sites where leaf-miners were absent.

The correlogram for all sites sampled shows that there is significant and relatively strong spatial dependence in the local densities of the holly leaf-miner across its geographic range (Figure 4.3a). That is, sites relatively close together are more similar to each other than those further apart. Up to approximately 750km, leaf-miner densities are positively autocorrelated. In addition, at lags of over 2000km there is a sharp discontinuity in the autocorrelation profile and densities become negatively autocorrelated. This largely represents the relationship between sampling sites at opposite extremes of the geographic range and suggests a highly asymmetric abundance structure to the range of this species. The correlogram constructed only for those sites

a)



b)

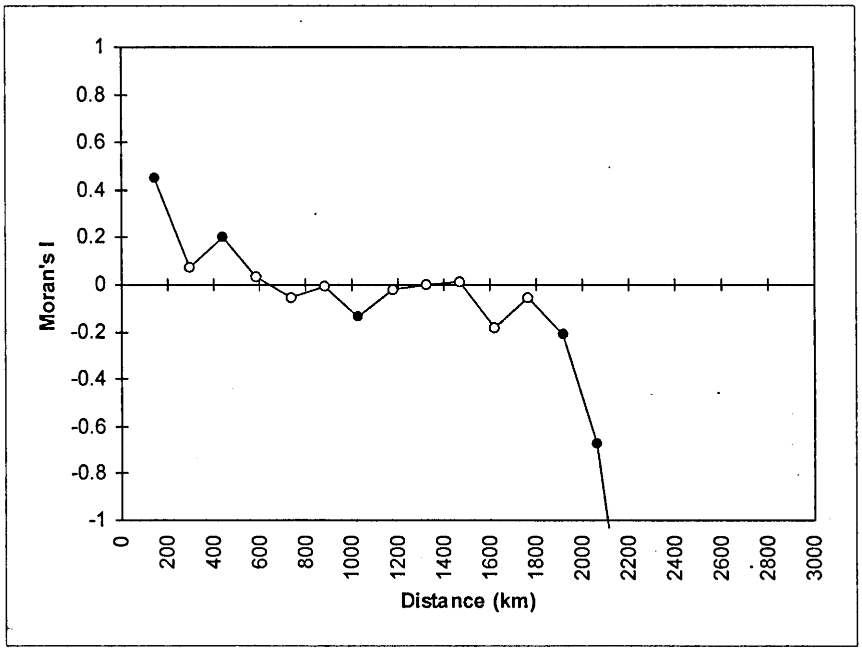


Figure 4.3 Spatial correlograms of *P. ilicis* densities from across its geographic range. Scale on the *x*-axis represents maximum distance in each distance class. Filled circles indicate significant values for Moran's I ($p < 0.05$). Overall correlogram significance was tested using Bonferroni's correction for multiple comparisons. (a) All sites sampled ($p < 0.01$) (b) Sites with leaf-miners present only ($p < 0.01$).

with leaf-miners present shows qualitatively the same spatial structure (Figure 4.3b). However, note that Moran's I falls to zero at a shorter lag of 500km.

### *Interpolation*

The interpolated map of leaf-miner densities clearly shows strong spatial structure in the geographic distribution of local densities of *P. ilicis* (Figure 4.1). High to moderate densities predominantly fall in a band running north-east to south-west, from the south-west coast of Norway, across Denmark, Holland, Belgium, northern France and southern Britain. Densities largely decline away from this band, with the principal exception of a peak of high density in central Italy. The interpolation also shows why at the longest lags (e.g. between Norway and Spain) Moran's I was strongly negative, with the density peak in Norway contrasting with the absence of holly leaf-miners in Spain.

Summing under the interpolated surface also produces a strongly right-skewed frequency distribution of the extent of areas over which *P. ilicis* attains different local densities (Figure 4.2b). This frequency distribution differs principally from that for the raw sample data in that the zero density class is significantly smaller (figure 4.2a).

### *Partial regression analysis*

Deciding which sites to include in the partial regression presents a similar problem to that encountered in the autocorrelation analysis. Again, we conducted analyses for all sites and also only for the subset of sites where leaf-miners were present. In an additional analysis the leaf-miner density data was simplified to a binary response variable, representing presence or absence from a site. This was considered to be another way of determining how well the geographic range of the holly leaf-miner could be explained by environmental variation after accounting for spatial dependence.

#### *(i) Regression over all sites (including zero densities)*

Including all 96 sites in the analysis, after stepwise regression by elimination the environmental variables that were found to contribute significantly in accounting for variation in leaf-miner densities were mean winter and summer temperatures, winter precipitation and humidity. When considered alone, these accounted for 48.0% of the deviance in leaf-miner densities ($F_{4,91} = 19.9$, $p < 0.01$). In addition, five terms from the coordinates polynomial contributed significantly to explaining the spatial structure of leaf-miner densities. These terms were $x$, $y$, $xy$, $y^2$ and $xy^2$ ($F_{5,90} = 18.1$, $p < 0.01$), where $x$ and $y$ represent longitude and latitude respectively. When all of these variables were considered together, 55.2% of the total variation in leaf-miner densities was accounted for by the full model (Table 4.1). Once spatial position was taken into account, only 7.9% could be attributed to environmental variation alone, the rest was shared with the spatial component. However, this reduced fraction was still significant at the 5% level. Therefore, despite the spatial structure shared by both the local population densities of *P. ilicis* and environmental variation, a significant environmental effect on leaf-miner density can be discerned.

*(ii) Regression over sites with leaf-miners present only (no zeros)*

In this analysis 61 sites which had at least one infested tree were included. After stepwise regression, winter and summer temperatures and winter precipitation were found to contribute significantly to the model and three terms from the coordinates polynomial; $x$, $xy$ and $xy^2$ (Table 4.1). When these variables were considered together in the partial regression, 40.0% of the total variation could be accounted for by the full model. However, the environmental variables were not found to contribute significantly after taking into account the spatial component. Moreover, the spatially structured environmental fraction was negative.

*(iii) Binary response model*

Finally, using a simplified binary response variable for all 96 sites, winter humidity was found to be the only environmental variable that contributed significantly to predicting whether leaf-miners would be present at a site ($\chi^2 = 77.23$, $p < 0.01$) (Table 4.1). From the coordinates matrix, the terms $x$, $y$, $xy$ and $x^2$ were significant ($\chi^2 = 77.23$, $p < 0.01$). After the partial regression, the total amount of variation explained was 65.4% with the environment only component accounting for just 4.1% of the variation. However, this component was significant at the 5% level.

## Discussion

The holly leaf-miner occupies much of the geographic range of its host plant (Figure 4.1). However, on the basis of extensive sampling it is clear that it is also absent from large areas, particularly in the south (we predict that it will be absent from the peripheral, but unsampled, occurrences in North Africa). Leaf-miners in general might be expected to occupy a higher proportion of the geographic ranges of their hosts than do externally feeding insects because of their more intimate host associations (Cornell, 1989). Given that many externally feeding herbivorous insects plainly occupy a rather

| Leaf-miner data | Environmental variables | Co-ordinate terms | Deviance accounted for in leaf-miner data (%) | | | | Significance test |
|---|---|---|---|---|---|---|---|
| | | | Total | Env | Env * Space | Space | |
| All sites | *wt, st, wp, wh* | $x\ y\ y^2\ xy\ xy^2$ | 55.2 | 7.8 | 40.7 | 6.7 | $F_{4,86} = 2.49$<br>$p < 0.05$ |
| Occupied sites only | *wt, st, wp* | $x\ xy\ xy^2$ | 40.0 | 8.9 | -2.2 | 33.3 | $F_{3,57} = 0.943$<br>NS |
| Presence/absence (binary response) | *wh* | $x\ y\ xy\ x^2$ | 65.4 | 4.1 | 50.4 | 10.9 | $\chi 2 = 5.15$<br>$p < 0.05$ |

Table 4.1 Results of the partial regressions of *Phytomyza ilicis* densities on environmental data taking into account a spatial component. Environmental variables: *wt* = winter temperature; *st* = summer temperature; *wp* = winter precipitation; *wh* = winter humidity.

| | *Winter Temperature* | *Summer Temperature* | *Winter Precipitation* | *Summer Precipitation* | *Winter Humidity* | *Summer Humidity* |
|---|---|---|---|---|---|---|
| Summer Temperature | 0.71 | | | | | |
| Winter Precipitation | -0.15 | -0.53 | | | | |
| Summer Precipitation | -0.45 | -0.52 | 0.77 | | | |
| Winter Humidity | -0.38 | -0.62 | 0.24 | 0.25 | | |
| Summer Humidity | -0.36 | -0.75 | 0.51 | 0.51 | 0.86 | |
| Altitude | 0.19 | 0.36 | -0.34 | -0.33 | -0.77 | -0.71 |

Table 4.2 Matrix of correlation coefficients between the environmental variables at all sites sampled during the survey.

small proportion of the range of their host (e.g. Strong *et al.*, 1984; Quinn *et al.*, 1997, 1998), the data for the holly leaf-miner are not obviously at odds with this suggestion.

The strong right skew to the frequency distribution of estimated local densities of the holly leaf-miner across the geographic range of its host plant conforms to general expectations about the shapes of intraspecific abundance distributions (e.g. Taylor *et al.*, 1978; Perry & Taylor, 1985, 1986; Gaston, 1994; Strayer, 1991; Brown *et al.*, 1995). However, examples are scarce in which the pattern has been demonstrated for sample sites distributed across the majority of a species' geographic range, and in which the zero class has been retained (Figure 4.2a). Where it occurs, most sites support relatively low densities of the holly leaf-miner, and only a very few support high ones. Of course, to some extent this is a biased estimate of the true population frequency distribution since the number of sites sampled was not equal between different regions. Rather, sampling intensity is more a reflection of the time available to sample in any one region. For example, since there were very few mines found in Spain and Portugal there was more time to visit further sites. Therefore one would expect bias to enter into any global estimate of the frequency-abundance histogram based on raw data alone and for this bias to make the sample histogram more right-skewed. Possibly, a more representative histogram is the one based on the interpolated surface of leaf-miner densities across the range of holly (Figure 4.2b). Here we see a similar pattern to the raw data plot but with a much lower frequency of sites supporting no leaf-miners. However, even this is not necessarily completely unbiased since the density of holly trees, that is the number of suitable habitat patches available for the holly leaf-miner to colonise, also varies in space.

Correlograms are increasingly being used to check for spatial dependence in local abundance data. Koenig (1999) warns that despite the fact that a correlogram may be shown to be significantly non-random (due to the large number of comparisons - and hence degrees of freedom - between pairs of samples), individual values of Moran's I may be so low within each distance class as to be biologically meaningless. However, the autocorrelation profiles for *P. ilicis* across its geographic range indicate very strong spatial structuring of local population densities. As documented here, positive autocorrelation has been found to be a common (albeit not ubiquitous e.g. Koenig, 1998, 1999) feature at short lags (e.g Eber & Brandl, 1994; Brown *et al.*, 1995). As one expects intuitively, many species have similar densities at sites that are close to one another. Across an entire geographic range, Brown *et al.* (1995) suggested that autocorrelograms of densities may for many species show a characteristic bowl-shape, with marked positive autocorrelation at both short and very long lags. They argue that this reflects a symmetric unimodal abundance structure to ranges, with the high autocorrelation at long lags indicating the similarly low levels of density found at opposing range edges. The correlograms for leaf-miner densities do not show this pattern, and the abundance structure of its range plainly does not exhibit a central peak which declines towards the range limits (Figure 4.1). However, it should also be noted that although a symmetric unimodal abundance distribution would have a bowl-shaped autocorrelation profile the converse is not necessarily true. Brown *et al.* (1995) give four examples of bird distributions that do show this autocorrelation pattern using data from the North American Breeding Bird Survey. However, reference to abundance maps for these species generated from the same data source (Price *et al.* 1995) reveal that arguably none of them have a central abundance peak. Most notably, the map for the Carolina wren *Thryothorus ludovicianus* is extremely asymmetric, with peak abundances toward the south-east coast of the United States, yet it's autocorrelation profile is plainly bowl-shaped.

Despite the pitfalls present when interpreting such maps visually, even a cautious interpretation of the interpolated surface of densities of the holly leaf-miner suggests that it too has a strongly asymmetric spatial pattern of local abundance, with peak values being attained toward the north-eastern extremity of its geographic range (figure 4.1). However, the pattern is clearly quite complex. Across several regions, such as Britain, there appear to be gradients in densities that might, falsely, suggest simple trends of increasing density towards the range centre if these were considered in isolation. This raises the spectre that the results of previous analyses of abundance structure based on only parts of geographic ranges may indeed prove misleading (see Introduction).

As Legendre & Legendre (1998) point out, there are two motivations for analysing data using analyses related to the partial regressions we present here. One is that spatial structuring can be a major source of false correlations which do not indicate causal relationships between environmental variation and the response variable of interest. The other is that both the spatial and non-spatial components of the environmental variation may be considered to be causal and the magnitude of the spatially-structured environmental effect or some other component of the variation in the response variable may be of interest in itself.

Taking the first more conservative approach, it is encouraging that despite the high degree of spatial dependence in our data that some effect of environmental variation on holly leaf-miner populations can be discerned when this is taken into account (Table 4.1). However, when considering the subset of the data which included only occupied sites, the amount of variation accounted for by the effect of environment alone was not found to be significant. This does not mean that there is no causal relationship between environmental variation and leaf-miner densities but it does indicate that if such a relationship exists, it cannot be disentangled from the confounding effect of space for this particular analysis. Considering all three analyses together, environmental variation has been shown to influence whether holly leaf-miner populations will be present at particular sites across the range of its host plant, although we have not demonstrated that there is a consistent significant effect on the internal density structure of the leaf-miner's range itself. The partitioning of variation in an ecological response variable between spatial and environmental effects is an appealing approach yet since the components are not strictly additive (see Legendre & Legendre, 1998, for a full account) some caution must be used in interpretation of the results. The small negative value obtained for the spatially-structured environmental component in the second analysis indicates that the environmental and spatial variables together explain leaf-miner densities better than the sum of their individual effects and that the environmental variables themselves have significant spatial structure. However, the two sets of variables are having effects of opposite sign on holly leaf-miner densities. In the other two analyses the amount of spatially-structured environmental variation is by far the largest proportion of the total explained variation. This is typical of the results found by others when analysing ecological data in this and similar ways. At a smaller, regional scale, in a study of the factors that affect the species composition of oribatid mite communities, Borcard *et al.* (1992) found that two-thirds of the variation explained by environmental variables could equally well be predicted by spatial position. Kitron *et al.* (1996) used the method proposed by Getis (1990) in their analysis of tsetse fly distribution using remotely sensed environmental data. They also found that the spatial component of their environmental data contributed more to explaining fly catches than the non-spatial component. The remaining spatial variation may be linked to unknown environmental variables that also contribute to the unexplained variation (Borcard *et al.* 1992). At smaller scales, it may reflect the effects of biotic processes between patches such as dispersal, predation or disturbance (Legendre 1993), though we feel that these effects are unlikely to be detected across an entire geographic range (see below).

The stepwise regressions for each of the data sets yielded different combinations of the environmental and spatial variables. We have not attempted to attribute any biological significance to this since the environmental variables themselves are often closely correlated (Table 4.2) and the results of a stepwise regression can be sensitive to the method used. Our intention here is simply to provide the minimum adequate model for explaining leaf-miner densities in terms of both environmental variation and spatial position. The results of our analyses demonstrate the difficulty in identifying important controlling factors on local population densities in the presence of correlated environmental variables and a high degree of spatial dependence. These are problems that will often be unavoidable in observational surveys of this nature, which ultimately can only help to formulate hypotheses about the regulation of abundance in space and would require subsequent testing by experimental studies.

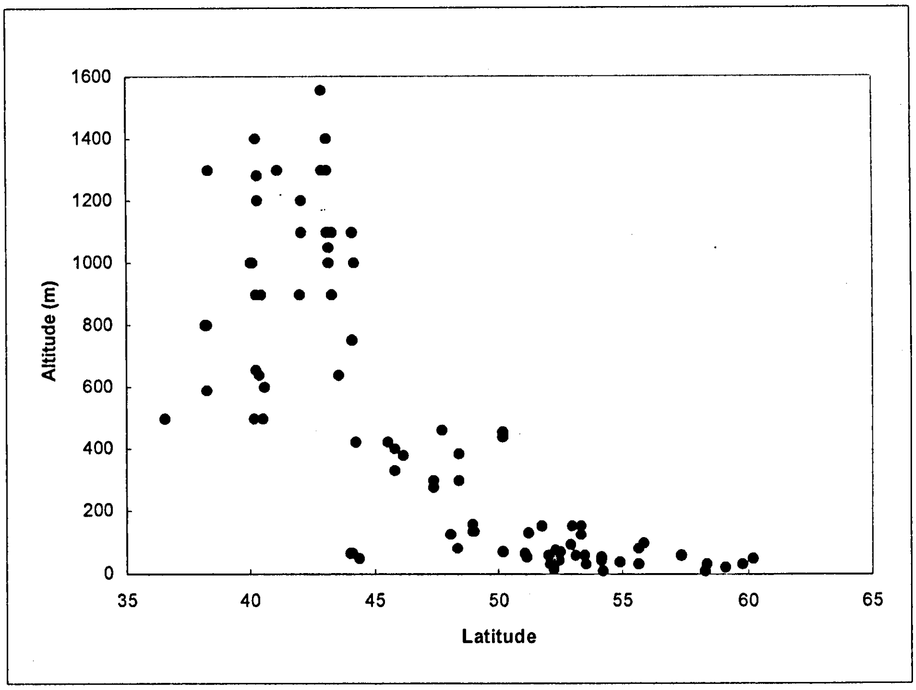
Taking the rather less conservative approach to interpreting the results of the partial regression analysis, which is at least in part justified since some causal influence of environment on leaf-miner densities can be inferred, it is interesting to look at the findings reported here in the light of a popular ecological hypothesis for the structure of the geographic range. In so doing, it should, of course, be emphasised that the pattern of a unimodal symmetric spatial distribution of local abundances has long been emphasised to be a statistical one (Hengeveld & Haeck, 1982; Brown, 1995; Maurer, 1999) with many exceptions to the general rule. That is, for any one species or at any one point in time the pattern may not be apparent, but when looking at many species averaged over longer periods it generally holds true. It would be naïve to dispute the existence or otherwise of the general pattern based on the data collected for one phytophagous insect species in Europe. However, in the case of North American birds we have seen that autocorrelation profiles do not necessarily say what they appear to. Moreover, in a study cited by Wiens (1989), Root (1988c) found that only 4% of 48 species of birds wintering in North America could be said to have such a distribution according to her criteria. Other studies have also brought the pattern into question (see Introduction).

Brown (1984, 1995; see also Brown & Lomolino 1998) has argued that if (i) the abundance and distribution of a species are determined by combinations of many physical and biotic variables, and that spatial variation in population density reflects the probability density distribution of the required combinations of these variables, and (ii) some sets of environmental variables are distributed independently of each other, and environmental variation is spatially autocorrelated, then density should be highest at the centre of the range of a species and should decline towards the boundary.

Examining each of these assumptions in turn, first, the abundance and distribution of the holly leaf-miner certainly seem likely to be determined by combinations of several physical and biotic variables. As a specialist herbivore, its occurrence at a site necessitates that of its host plant, holly. The northern and eastern extent of the distribution of holly are thought to be limited by minimum winter temperatures, where the trees can become prone to frost damage (Iversen, 1944; Perring & Walters, 1962). In addition, summer temperatures have to be sufficiently high to allow fruit production. In the most northerly parts of its range, holly is limited to low altitudes (e.g. <200m in Norway), whereas in southern regions it is confined to the mountains (usually in sheltered ravines with relatively high humidity). In central areas, the tree is present throughout a large altitudinal range (Peterken & Lloyd, 1967). A plot of the altitude of our sampling sites against latitude certainly reflects this trend in the distribution of holly (Figure 4.4a). Within the geographic range of its host the holly leaf-miner does not occur in all regions. The results of the partial regression analysis show that this can be attributed in part to the effect of climatic variables. These effects may be direct, mediated through responses of the host plant, or may act upon agents of mortality across the geographic range of the leaf-miner (Brewer & Gaston, in preparation).

Second, spatial variation in the population density of the holly leaf-miner seems likely also to reflect the probability density distribution of the required combinations of physical and biotic variables (i.e. it is an optimum response surface; Hengeveld, 1990), although the results of our analyses are not entirely consistent in this regard. The most obvious way in which this might not occur would be if the species exhibited some form of sink-source dynamics, such that in some areas immigration maintained densities at levels above those which could otherwise be sustained (Pulliam, 1988). It is difficult to envisage that this could occur on the scale of the entire geographic range of this species. Thus, for example, it seems improbable that the high density of the holly leaf-miner in Norway is a product of immigration. In general, even small isolated populations appear to be self-sustaining, although the adult insects are probably good dispersers and may readily locate any unoccupied but suitable trees.

a)



b)

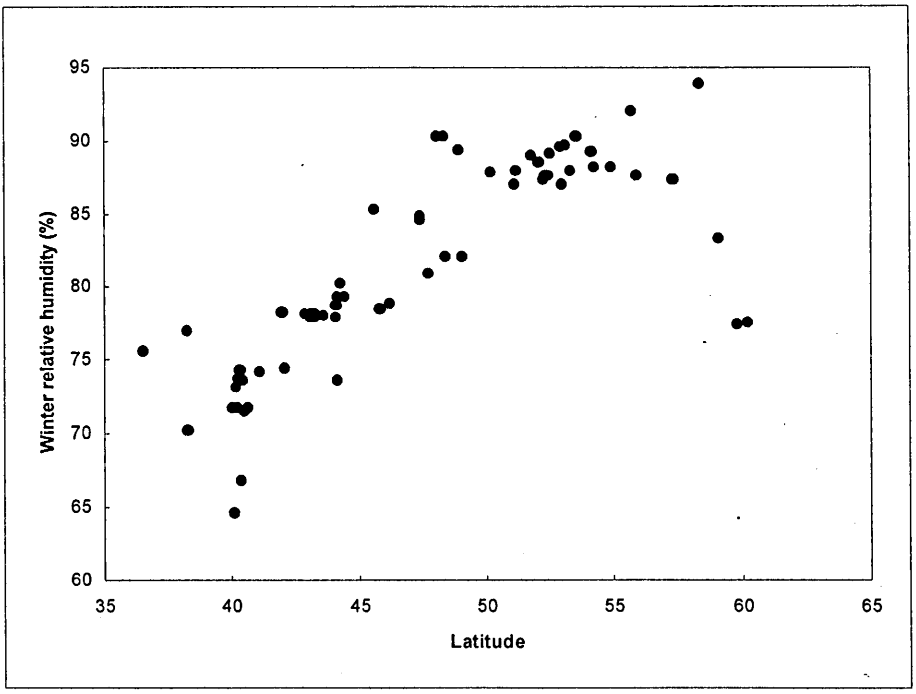


Figure 4.4 Examples of spatial trends in environmental variation. Plots of two of the environmental variables used in the partial regression analyses against latitude. The purpose of fitting the trend surface to the environmental data was to remove such trends before testing for a significant effect on leaf-miner densities (a) Altitude (b) Winter relative humidity.

The third and fourth assumptions of Brown's mechanism are that some sets of the environmental variables are distributed independently of each other and that environmental variation is spatially autocorrelated. Since our data set is comprised mainly of climatic variables, it is unsurprising that most of them show significant correlation with one another (Table 4.2). However, it is possible that some of the environmental variables we did not measure (e.g. aspects of soil type dependent upon underlying geology) may influence holly leaf-miner populations independently of climatic variation. We certainly found significant spatial structure in all of our environmental variables and in two out of the three partial regressions this structure was shared with holly leaf-miner populations. However, this spatial structure does not yield a simple Gaussian template of habitat suitability for *P. ilicis*. Population densities may well reflect local environmental conditions across the geographic range, but the spatial structure of the environment as it is presented to the organism certainly does not appear to be a simple one.

In sum, examination of the local abundance structure of holly leaf-miner populations across the majority of its natural global geographic range suggests a rather different pattern from that which has been asserted to generally hold, namely a rather complex pattern of local density as opposed to a trend of increasing density towards the range centre. This result can readily be explained as a consequence of a violation of the assumptions about environmental structure underlying Brown's (1984) general explanation for why such a peaked pattern should exist. Further investigation into environmental structure as it presents itself to local populations may prove to be a fruitful area of research into understanding the internal structure of the geographic range.

# Chapter 5

## Spatial structure in the demographic rates of the holly leaf-miner across its geographic range

Andrew M. Brewer* and Kevin J. Gaston

Biodiversity and Macroecology Group, Department of Animal & Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

*Contact details: Tel: 0114 2220034, Fax: 0114 2220002
e-mail bop96amb@sheffield.ac.uk

### Abstract

There are few studies that have reported differences in demographic rates in local populations across the geographic range despite the fact that such information might yield insights into the causes of its internal abundance structure. Components of demographic rates are reported here for the holly leaf-miner *Phytomyza ilicis* Curtis across its natural geographic range in Europe. Correlograms reveal that a number of the these components exhibit significant spatial structure across the geographic range and interpolated maps show that some of these patterns are remarkably simple. Most mortalities are largely independent from one another and from leaf-miner population densities. *P. ilicis* experiences very dissimilar patterns of mortality in different parts of its range. While some correlations between broad-scale environmental variation and the demographic rates components were found, it was not possible to separate this effect from shared spatial structure. However, the results from this initial survey are encouraging and suggest that understanding range structure in terms of underlying demographic rates may be tractable for some species.

### Introduction

The local population density of a species is the net outcome of four basic demographic rates: births and immigration minus deaths and emigration. For most species, accurate measurement of any of these rates is extremely difficult, if not impossible. As a result, there is little data on how they may vary between local populations in different parts of a species distribution. However, such research has the potential to go some way towards explaining the internal structure and boundaries of the geographic ranges of species. Indeed, Maurer and Brown (1989) have proposed that the pattern of a single species' local population densities across its geographic range may best be understood by examining the interaction between smoothly varying demographic rates across space

A few studies have shown that there can be systematic changes in particular demographic rates, or components thereof, across ranges. Most notably a number of studies on reproductive investment in birds have demonstrated latitudinal trends in clutch size (Peakall, 1970; Ricklefs, 1972; Koenig, 1984; Koenig, 1986; Bell, 1996; Corbacho et al., 1997; Sanz, 1997; Dunn et al., 2000). Others have considered the differences in the population dynamics of a species between the centre and edge of the geographic range (See Hoffmann & Blows, 1994, for a review). However, to date there have been no studies that have attempted to build a relatively complete picture of changes in population dynamics across the whole range of any species.

In a previous paper (Brewer & Gaston, *manuscript*), we reported the results of a survey of the spatial structure in local population densities of the holly leaf-miner *Phytomyza ilicis* Curtis across its geographic range and showed that a distinct pattern at a broad spatial scale could be detected despite the high degree of variation in leaf-miner densities that is typically found at local scales (Heads & Lawton, 1983; Valladares & Lawton 1991; McGeoch & Gaston, 2000). Several previous studies of range structure have claimed that most species tend to exhibit higher local population densities in the centre of the range which gradually decrease towards the edges (Brown *et al.*, 1995). However, holly leaf-miner exhibits a range structure markedly different from this pattern. Peak densities were discovered on the north-eastern edge of the range, with high to moderate densities predominantly falling in a band running north-east to south-west. Densities largely declined away from this region but with some notable deviations from this general pattern. We also showed that this pattern could, to some extent, be correlated with broad scale meteorological variables across the range of holly and suggested how Brown's (1984) niche-based theory of the spatial structure of the geographic range could be modified to accommodate our findings.

In addition to collecting data on leaf-miner densities during the survey, we also measured oviposition rate and the dominant sources of mortality of holly leaf-miner populations at each sampling site. In this paper we report the results of this work and explore how local population dynamics of the holly leaf-miner vary across its geographic range. We investigate whether there are systematic changes in the components of the demographic rates between local populations and whether these changes can be correlated with differences in population density and environmental variation.

### The study system

*Phytomyza ilicis* Curtis, the holly leaf-miner, is a univoltine agromyzid (Diptera) leaf-miner that feeds exclusively on European holly *Ilex aquifolium*. In the UK, eggs are laid in about June on new holly leaves into the base of the underside of the midrib. When laying an egg, the adult female *P. ilicis* leaves a characteristic scar due to the insertion of her ovipositor into the midrib. These scars mean that oviposition density for a local population may be easily censused. A large proportion of the life-history of each individual is spent inhabiting a single holly leaf. After hatching, the larvae eat through the midrib and enter the outer parenchyma of the leaf lamina during the autumn. They feed throughout the following winter months and pupate in the mine in March, emerging from the leaf as adults in late May or June.

During the period spent within the holly leaves, a leaf-miner population may be subject to a number of potential mortalities which are largely sequential, albeit with some overlap. They have been relatively well documented (Cameron, 1939; Lewis & Taylor, 1967; Heads & Lawton, 1983) and include miscellaneous larval deaths, larval parasitism by the parasitoid *Chrysocharis gemma* Walker (Hymenoptera: Eulophidae), bird predation, pupal parasitism by at least eight species of hymenopteran parasitoids and miscellaneous pupal mortality. These causes of mortality can be identified by dissection of the mine at the end of the life-cycle of the leaf-miner. In addition, larvae that successfully complete their development and emerge as adult flies can also be identified from the characteristic emergence holes on the leaf surface. Therefore, the holly leaf-miner has a life-cycle that is more suitable than the majority for measuring a number of the components of the basic demographic rates of a local population. The level of oviposition (births), a number of components of mortality and the rate of successful emergence for a local population in any one year can all be assessed by rapid censusing techniques and dissection of a representative sample of leaves from the holly trees at a site.

## Methods

### *Data collection*

A full description of the population density survey can be found in Brewer & Gaston (*manuscript*). Leaf-miner densities and demographic rates were recorded throughout the natural range of holly, between June and December 1998. Our aim was to obtain a relatively complete coverage of the whole geographic range, by sampling both the centre and edges. The survey was organised as a series of routes of varying duration in different parts of the range where holly was sampled at as many sites as time allowed.

Leaf-mine and oviposition density estimates were taken from 10 trees at each sampling site, or as many as possible where less than 10 trees were present. 200 one-year-old leaves (or all the leaves of this age present on the tree if there were less than 200) were sampled from each tree. Following Valladares & Lawton (1991) density was estimated by haphazardly sampling from all around the canopy between ground level and a height of approximately 2m. For each leaf, the number of mines and egg scars present were recorded. Leaves of the correct age can be readily distinguished by their position between the annual nodes on the branch. This protocol gives an estimate of density for the previous growing season.

For trees on which no mines or egg scars were found in a sample of 200 leaves, an exhaustive search was made of the rest of the tree. In very rare cases these were found, in which case densities were treated as an arbitrarily small number (0.1 mines or egg scars per 200 leaves). This occurred for only 4 trees of the 703 sampled. When mines were found, a haphazardly chosen sample of 50 mined leaves was removed from each tree. If 50 mines could not be found on a particular tree, all the visible mines of the appropriate age were removed. These mines were dissected to ascertain the number of *P. ilicis* individuals that were subject to the different mortalities or emerged successfully as adults, following the procedures given by Heads & Lawton (1983).

The geographic location (expressed as decimal degrees longitude and latitude) of each site and its altitude (m) were obtained from appropriate maps.

### *Meteorological data*

Regional climatic data over a 0.5°×0.5° grid were obtained from a coverage of twentieth century terrestrial surface climate (New *et al.*, 2000). From this coverage, mean values for winter and summer temperatures, precipitation and humidity were calculated from data for the 10 year period between 1986 and 1995 inclusive.

### *Data analysis*

#### The demographic data sets

Since the mortalities are largely sequential in this system, the number of mines that can potentially succumb to a particular mortality is the total number of mines minus the number killed in previous mortalities. This smaller proportion of mines is used to calculate *apparent* mortality (Bellows *et al.*, 1992) as opposed to *real* mortality (which is the proportion of all mines). In all of the analyses below, apparent mortalities were used unless otherwise stated. These were calculated as the total number of each mortality recorded divided by the number of mines available to that mortality across all trees that were sampled. Oviposition rate was measured as the total number of egg scars divided by the number of leaves examined during density estimation at each site. Successful emergence is measured as the proportion of the total number of leaf-miners collected, since it will always be 100% of the mines that have not succumbed to any previous mortality.

#### Spatial autocorrelation

Spatial autocorrelation analysis (Cliff & Ord, 1973) was used to characterise the spatial structure of the demographic rates of *P. ilicis* that we measured. Moran's I was calculated for 15 equal distance intervals, and spatial correlograms were produced and tested for significant spatial dependence. Since the survey covered a relatively large geographic area, site coordinates (measured as decimal degrees longitude and latitude) were not treated as Cartesian coordinates when measuring distances between them. Instead, distances along great circles were calculated to take into account the curvature of the earth's surface. Bonferroni's correction for multiple comparisons was used when assessing overall correlogram significance.

#### Interpolation

The demographic rates were interpolated across space to give a visual impression of the changes in these rates across the geographic range. Each demographic rate was entered as a point coverage into a geographic information system (GIS), and interpolated using an inverse distance weighting technique (IDW) to create a grid comprising 0.1°×

0.1° cells over the study area. Seven nearest neighbours were used to estimate the values, with weights inversely proportional to the square of the distance from the estimated cell.

### Partial regression analysis

As with the density survey (Brewer & Gaston *manuscript*), partial regression analysis was used to model oviposition, the components of mortality and successful emergence in terms of the environmental variables. Following the methodology suggested by Legendre (1993), we used the method to estimate how much variation in the components of leaf-miner demographic rates can be attributed to regional climatic variation and altitude once the effect of spatial location has been taken into account. The spatial component of the holly leaf-miner demographic data was modelled using a third-order polynomial of the form:

$$f(x,y) = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3$$

where $x$ and $y$ represent longitude and latitude respectively. Significant terms, determined by conducting a stepwise regression of leaf-miner demographic rates on a matrix of all of the terms in the expression, were retained to construct a new matrix of spatial variables to be used in the subsequent analysis. Similarly, the leaf-miner demographic rates were regressed onto a matrix containing the climate and altitude data and significant terms determined after stepwise regression by elimination were retained.

The combined effect of both the environmental and spatial variables on leaf-miner demographic rates was calculated by multiple regression of the rates onto both sets of predictive variables combined. The explanatory potential of the environmental variables, after correcting for spatial dependence, was calculated by measuring the change in deviance explained by the regression model after the environmental variables were removed. This fraction was then tested to determine whether the change in deviance was statistically significant.

At the end of the partial regression analysis, variation in leaf-miner demographic rates could be divided into four components: (a) non-spatial environmental - the fraction that can be explained by the environmental variables independent of any spatial structure; (b) spatially structured environmental - spatial structuring in leaf-miner demographic data that is shared with the environmental data; (c) non-environmental spatial - spatial structure in leaf-miner demographic data that is not shared with the environmental variables; and (d) unexplained (residual) variation. These components and their associated probabilities allow several hypotheses about the causal relationships between environmental variation, spatial position and leaf-miner demographic rates to be tested.

The method employed here is essentially the same as Legendre's (1993) third extension to partial regression analysis with some minor modifications (see also Legendre & Legendre, 1998). Since the demographic data were in the form of proportions (e.g. apparent mortality rate per number of mines available), binomial errors were modelled. This results in proportions of deviance explained which are analogous to the partial $r^2$ statistics at the end of the analysis. In addition, during the stepwise regressions, any overdispersion in the residual deviance was corrected for using William's procedure where appropriate (Collett 1991, Crawley 1993) before testing for parameter significance. All of the partial regression analyses were conducted using GLIM version 3.77.

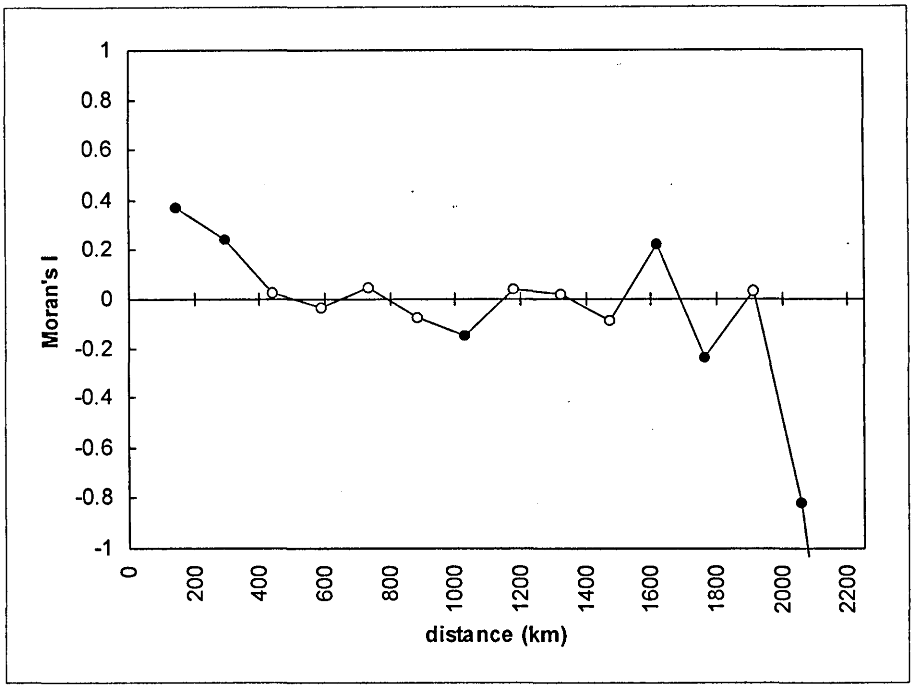
## Results

### *1. Spatial structure in the demographic data*

### Spatial autocorrelation

The correlograms show that significant spatial dependence is apparent in oviposition, successful emergence and a number of the components of leaf-miner mortality that we measured across the geographic range of the holly leaf-miner (figure 5.1). At least some significant positive autocorrelation was found at shorter lags for all of the rates measured, indicating that sites near to each other tend to have more similar levels of natality and mortality compared to those further apart. This correlation falls close to or below zero at intermediate lags followed by a considerable amount of variation between the rates at the longest separation distances between sites. At these long lags it should be noted that Moran's I is often calculated from a relatively low number of pairs of sites (figure 5.1h) and so care should be taken when interpreting these values. It is clear that where some of the mortality rates measured show what might be considered strong systematic spatial structure (e.g. larval parasitism) others showed a lot weaker (e.g. bird predation) or no (pupal parasitism) spatial structure across the whole of the geographic range. Oviposition rates show a similar pattern of autocorrelation to population densities (Brewer & Gaston *manuscript*) with positive autocorrelation up to approximately 400km and marked negative autocorrelation

a)



b)

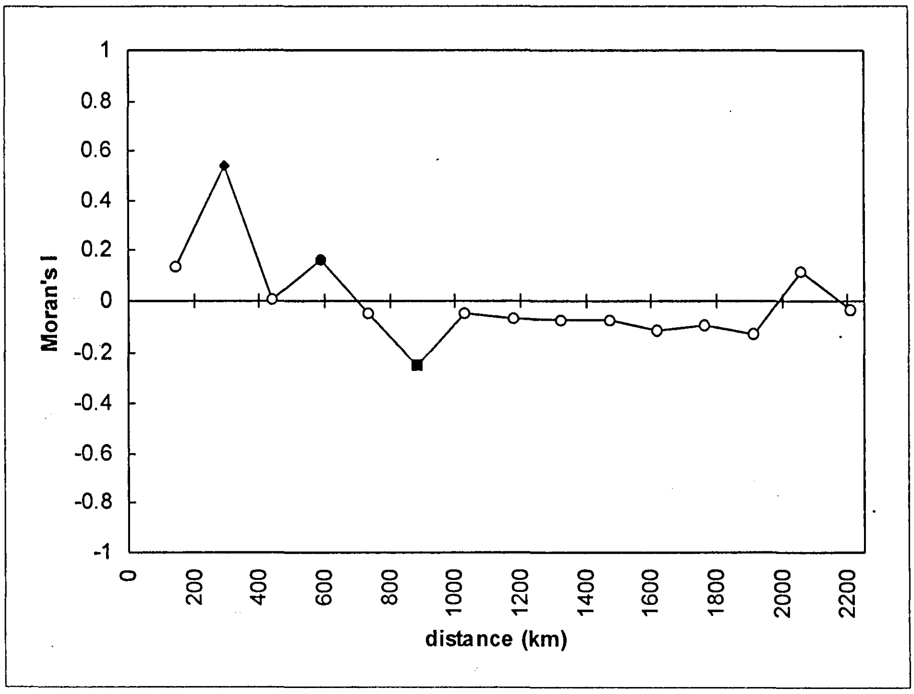
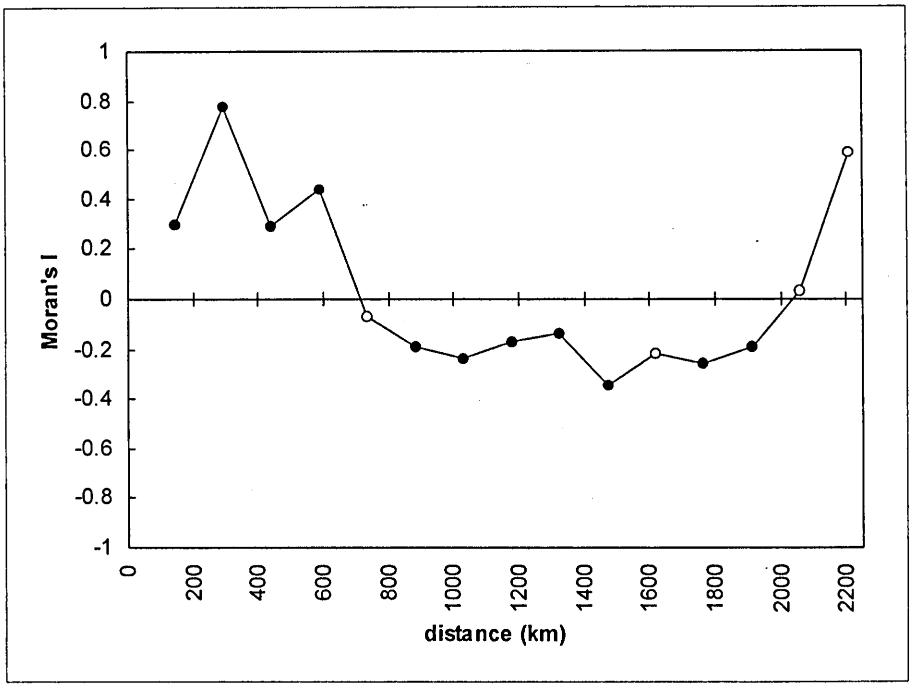


Figure 5.1 Spatial correlograms of *P. ilicis* demographic rate components from across its geographic range. Scale on the *x*-axis represents maximum distance in each distance class. Filled circles indicate significant values for Moran's I ($p < 0.05$). Overall correlogram significance was tested using Bonferroni's correction for multiple comparisons. (a) oviposition density ($p < 0.01$) (b) miscellaneous larval mortaliy ($p < 0.01$).

c)



d)

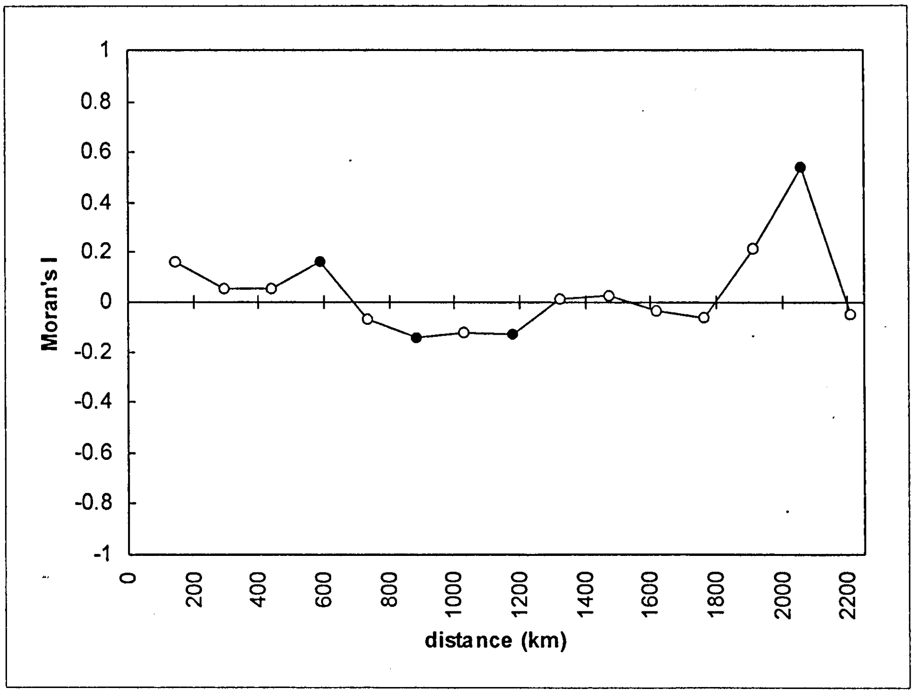
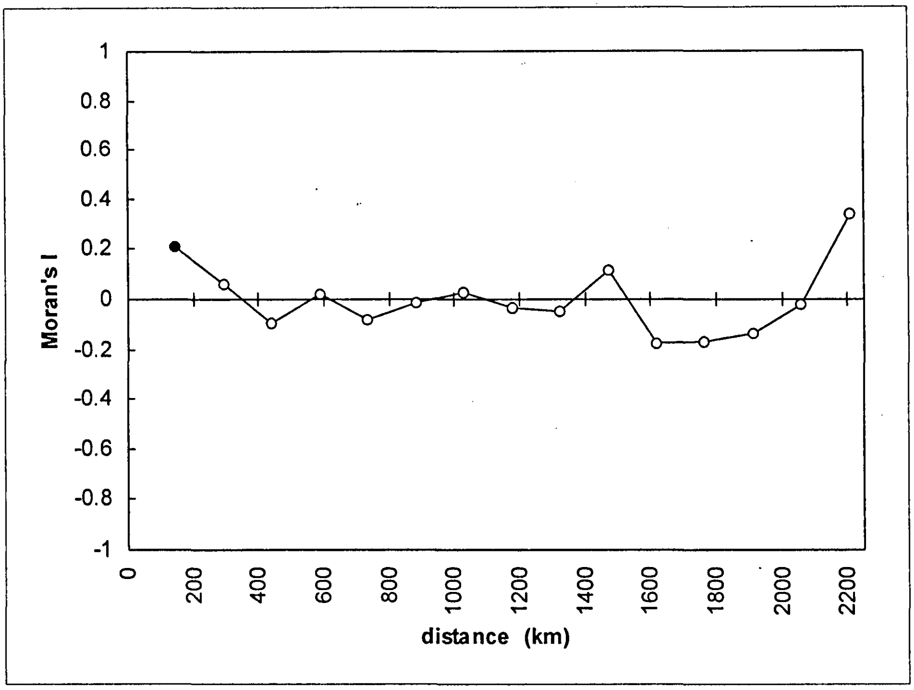


Figure 5.1(*continued*) Spatial correlograms of *P. ilicis* demographic rate components from across its geographic range. (c) larval parasitism ($p < 0.01$) (d) bird predation ($p < 0.05$).

e)



f)

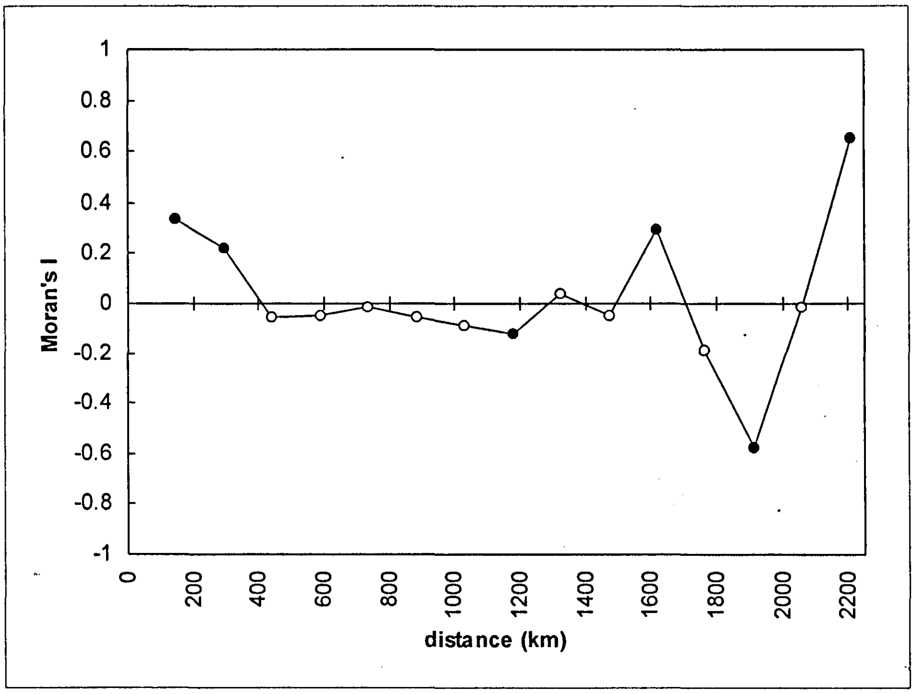
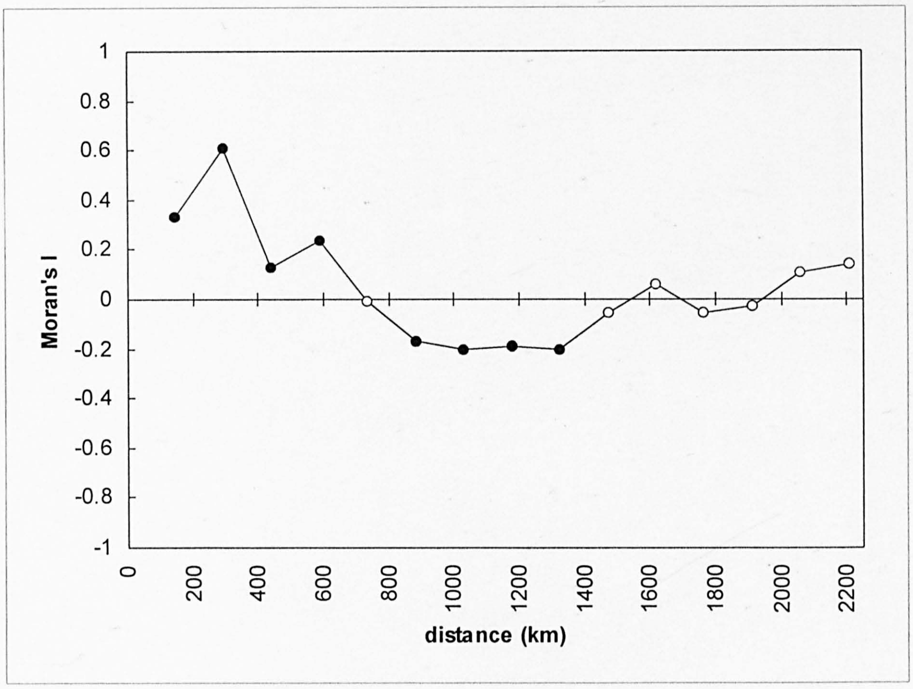


Figure 5.1(*continued*) Spatial correlograms of *P. ilicis* demographic rate components from across its geographic range. (e) pupal parasitism ($p > 0.01$) (f) miscellaneious pupal mortality ($p < 0.05$).

g)



h)

Figure 5.1(*continued*) Spatial correlograms of *P. ilicis* demographic rate components from across its geographic range. (g) successful emargence ($p<0.01$) (h) The number of pairs of comparisons within each distance class for the autocorrelation analysis.

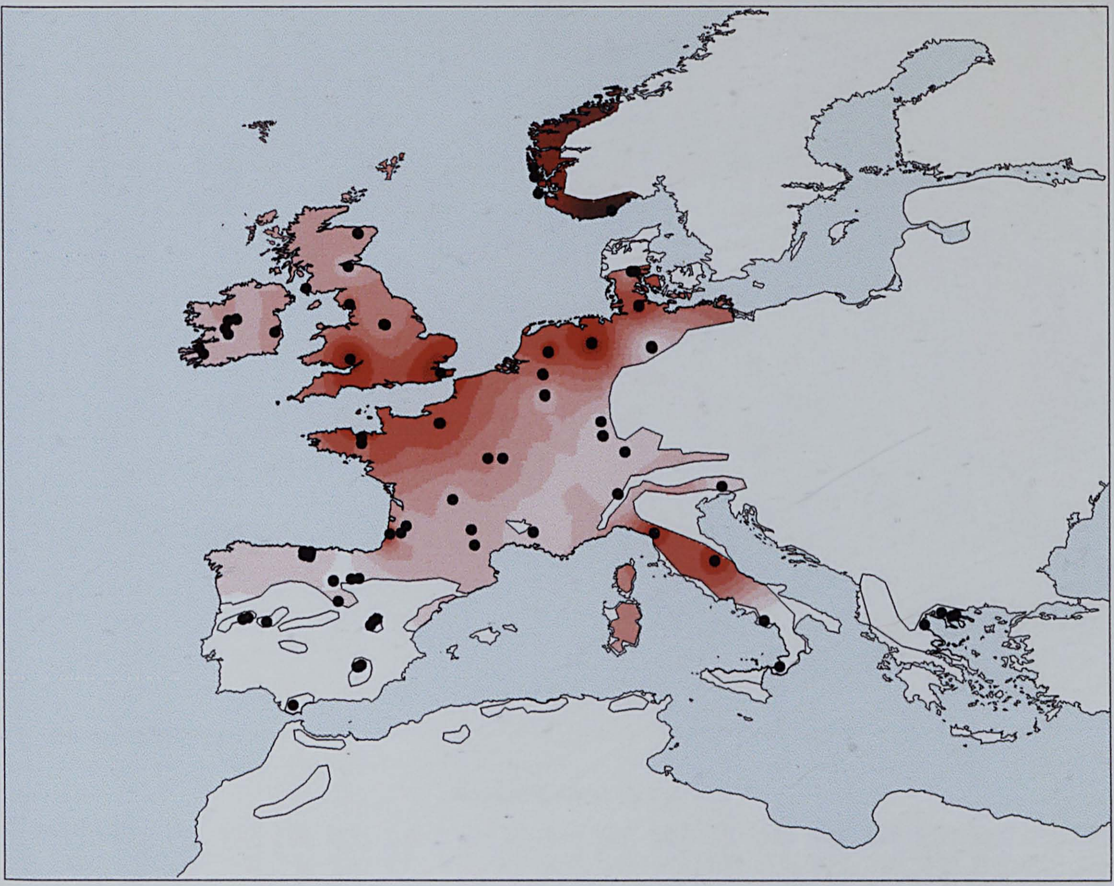


Figure 5.2a Interpolated surface of oviposition density across the range of European holly. Points on the map indicate location of sites where density estimates were made.

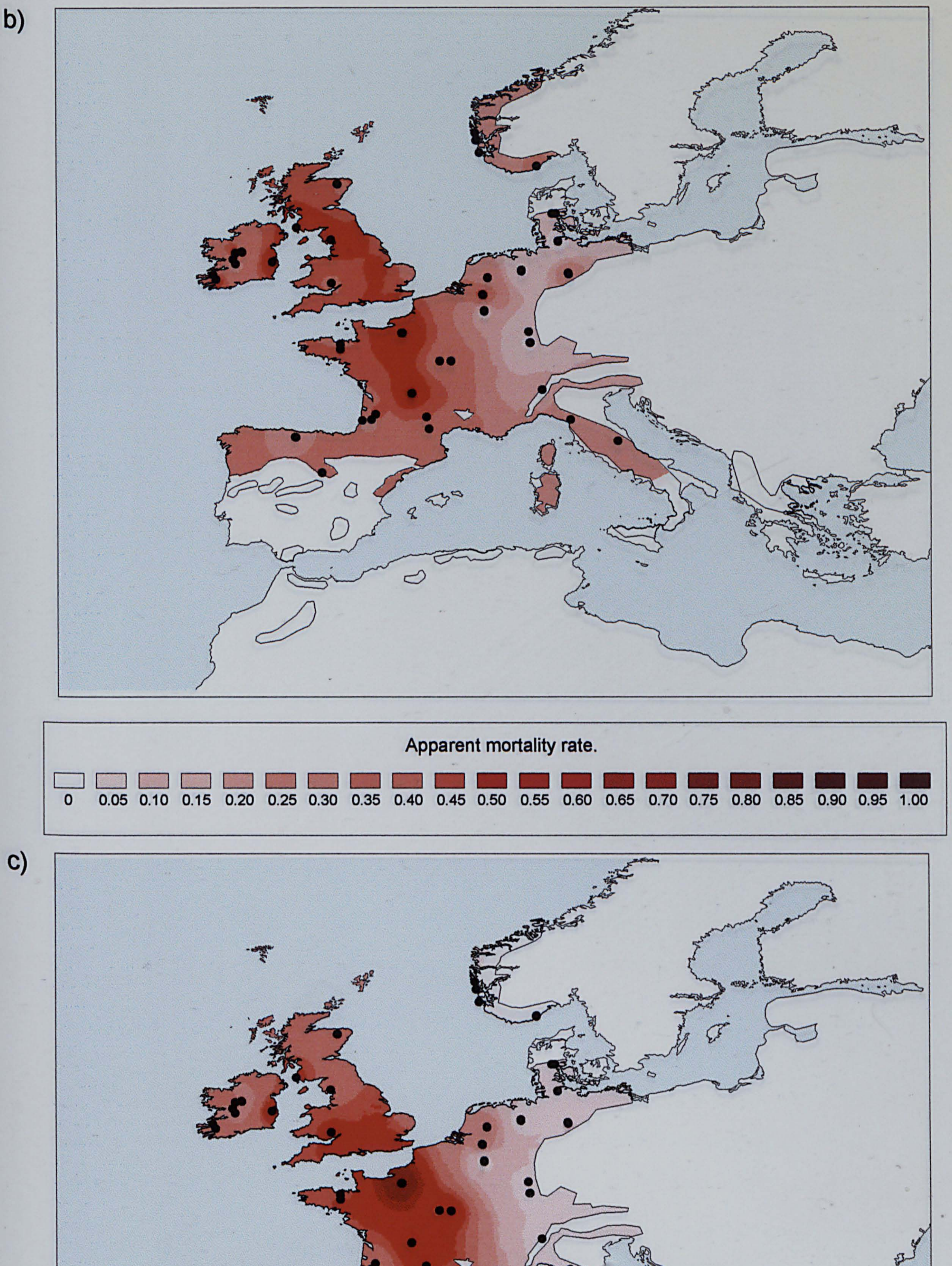


Figure 5.2 Interpolated surfaces of (b) miscellaneous larval mortality (c) larval parasitism rates.

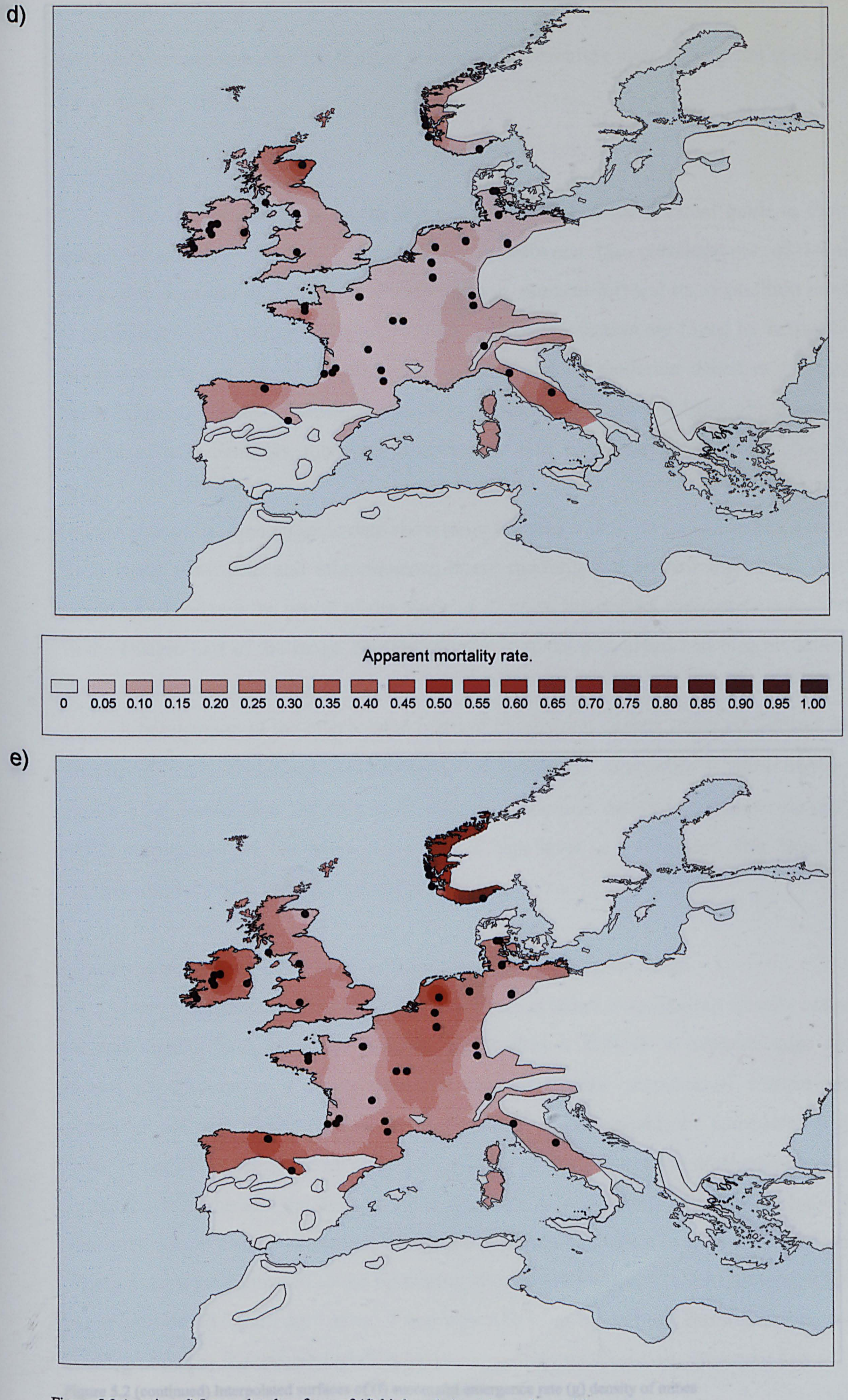


Figure 5.2 (continued) Interpolated surfaces of (d) bird predation (e) pupal parasitism rates.

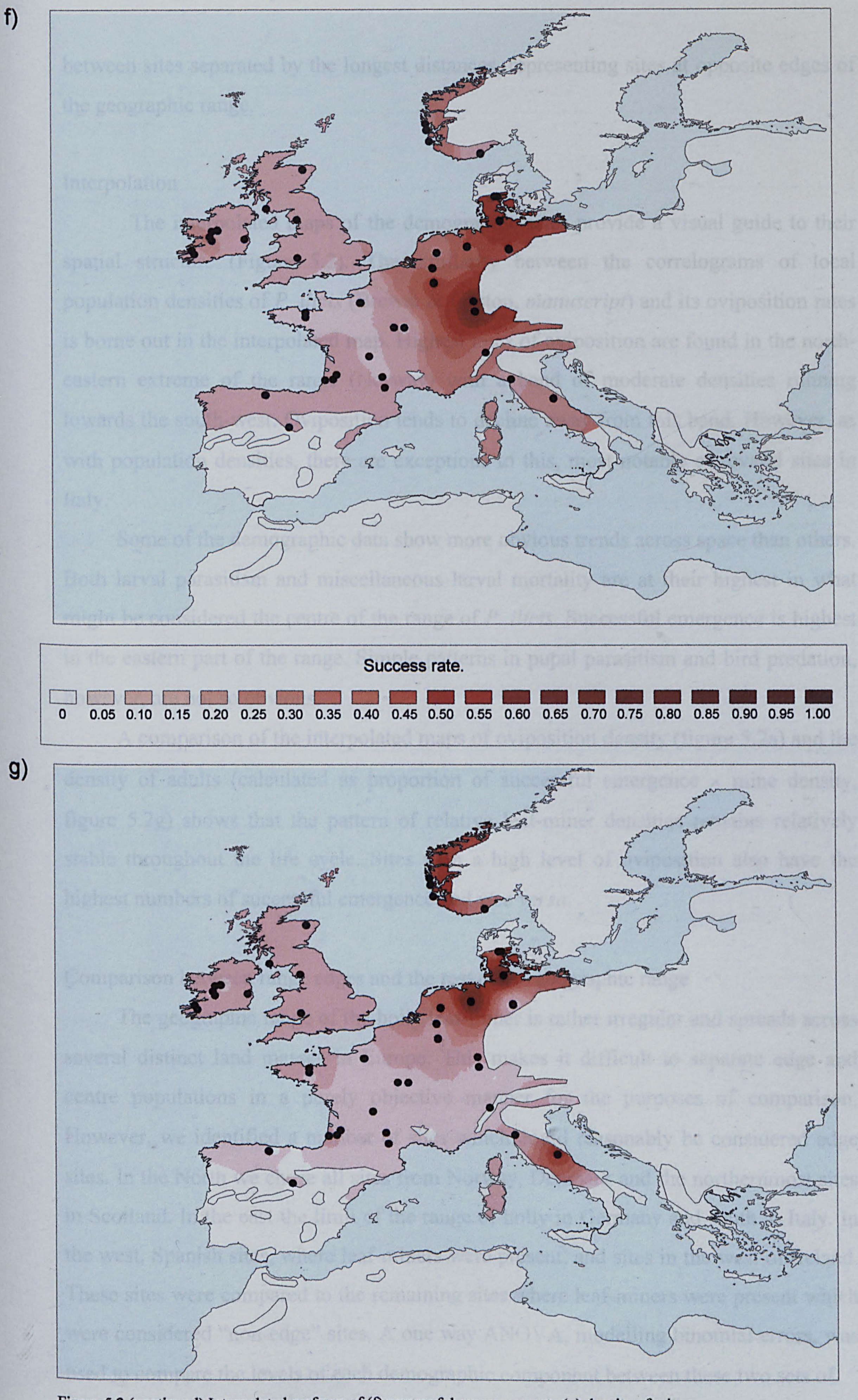


Figure 5.2 (continued) Interpolated surfaces of (f) successful emergence rate (g) density of mines that successfully emerge.

between sites separated by the longest distances, representing sites at opposite edges of the geographic range.

### Interpolation

The interpolated maps of the demographic rates provide a visual guide to their spatial structure (Figure 5.2). The similarity between the correlograms of local population densities of *P. ilicis* (Brewer & Gaston, *manuscript*) and its oviposition rates is borne out in the interpolated map. Highest rates of oviposition are found in the north-eastern extreme of the range (Norway) with a band of moderate densities running towards the south-west. Oviposition tends to decline away from this band. However, as with population densities, there are exceptions to this, most notably at several sites in Italy.

Some of the demographic data show more obvious trends across space than others. Both larval parasitism and miscellaneous larval mortality are at their highest in what might be considered the centre of the range of *P. ilicis*. Successful emergence is highest in the eastern part of the range. Simple patterns in pupal parasitism and bird predation, however, are not so obvious.

A comparison of the interpolated maps of oviposition density (figure 5.2a) and the density of adults (calculated as proportion of successful emergence × mine density, figure 5.2g) shows that the pattern of relative leaf-miner densities remains relatively stable throughout the life cycle. Sites with a high level of oviposition also have the highest numbers of successful emergence and *vice versa*.

### Comparison between range edges and the rest of the geographic range

The geographic range of the holly leaf-miner is rather irregular and spreads across several distinct land masses in Europe. This makes it difficult to separate edge and centre populations in a purely objective manner for the purposes of comparison. However, we identified a number of sites which could reasonably be considered edge sites. In the North we chose all sites from Norway, Denmark and the northernmost sites in Scotland. In the east the limit of the range of holly in Germany and south to Italy. In the west, Spanish sites, where leaf-miners were present, and sites in the west of Ireland. These sites were compared to the remaining sites where leaf-miners were present which were considered "non-edge" sites. A one way ANOVA, modelling binomial errors, was used to compare the levels of each demographic component between these two sets of

| **Demographic rate component** | **Significance test** |
|---|---|
| Mine density | $F_{1,61}= 0.08$<br>No significant difference |
| Oviposition density | $F_{1,61} = 1.14$<br>No significant difference |
| Miscellaneous larval mortality | $F_{1,54} = 7.67$<br>$p < 0.05$<br>Edge sites have lower rates of miscellaneous larval mortality |
| Larval parasitism | $F_{1,54} = 22.30$<br>$p < 0.05$<br>Edge sites have lower rates of larval parasitism |
| Bird predation | $F_{1,54} = 4.80$<br>$p < 0.05$<br>Edge sites have higher rates of bird predation |
| Pupal parasitism | $F_{1,54} = 0.07$<br>$p > 0.05$<br>No significant difference |
| Miscellaneous pupal mortality | $F_{1,54} = 0.38$<br>No significant difference |
| Successful emergence | $F_{1,54} = 6.12$<br>$p < 0.05$<br>Edge sites have higher rates of successful emergence |

Table 5.1 Comparison of levels of the demographic rates components between edge and centre sites using ANOVA, modelling binomial errors.

sites, again correcting for any over-dispersion before conducting the significance tests (Crawley 1993). Edge sites tend to have lower levels of larval parasitism and miscellaneous larval mortality than other sites but higher rates of bird predation and successful emergence (Table 5.1). However, there were no significant differences in miscellaneous larval mortality or pupal parasitism.

### *2. Correlations between demographic rate components and local population densities*

The extent to which the mortality rates covary across the whole of the range was calculated. Significant positive correlation was found between miscellaneous larval mortality and larval parasitism ($r=0.679$, $n=57$, $p<0.001$, figure 5.3), but no other mortality rates were correlated between sites. No significant correlations were found between each of the mortality rates and leaf-miner population densities at this scale. Successful emergence showed a weak negative ($r=-0.294$; $n=57$; $p>0.05$) correlation with mine density but this was not significantly different from zero. However, when considering all trees, figure 5.4 shows that while there is not a simple linear relationship between the two variables there does appear to be an upper limit on the rate of successful emergence which declines with leaf-miner density. At low densities successful emergence is very variable but as the proportion of mined leaves increases, the maximum rate of successful emergence drops consistently.

### *3. Demographic rate components and environmental variation*

#### Partial regression analyses

Using stepwise regressions, all of the demographic rates could to some extent be significantly described by the environmental data when considered alone (table 5.2). However, a large proportion of the environmental variation that contributed to explaining the demographic rates is spatially structured itself. As a result, when spatial position is taken into account using the terms from the coordinates polynomial, a significant environmental effect was detected for oviposition only.

## **Discussion**

Holly leaf-miner densities exhibit a distinct spatial structure across the geographic range. In addition, oviposition density, a number of the components of mortality, and successful emergence are positively spatially autocorrelated at short separation distances and also show a broad scale spatial structure across the whole geographic range.

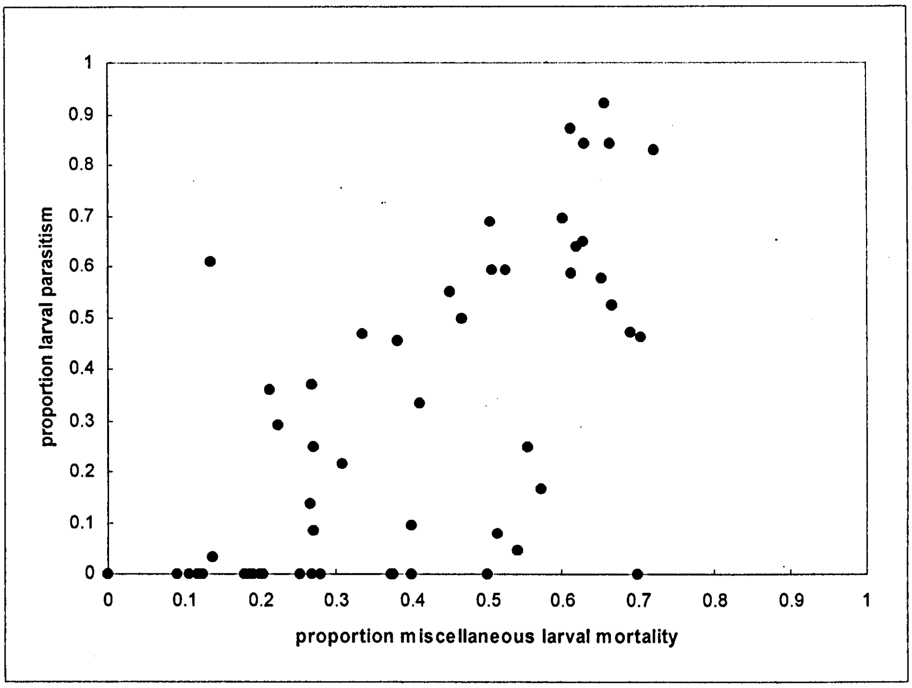


Figure 5.3 Plot of larval parasitism against miscellaneous larval mortality for all sites. (r=0.679, n=57, p<0.01)

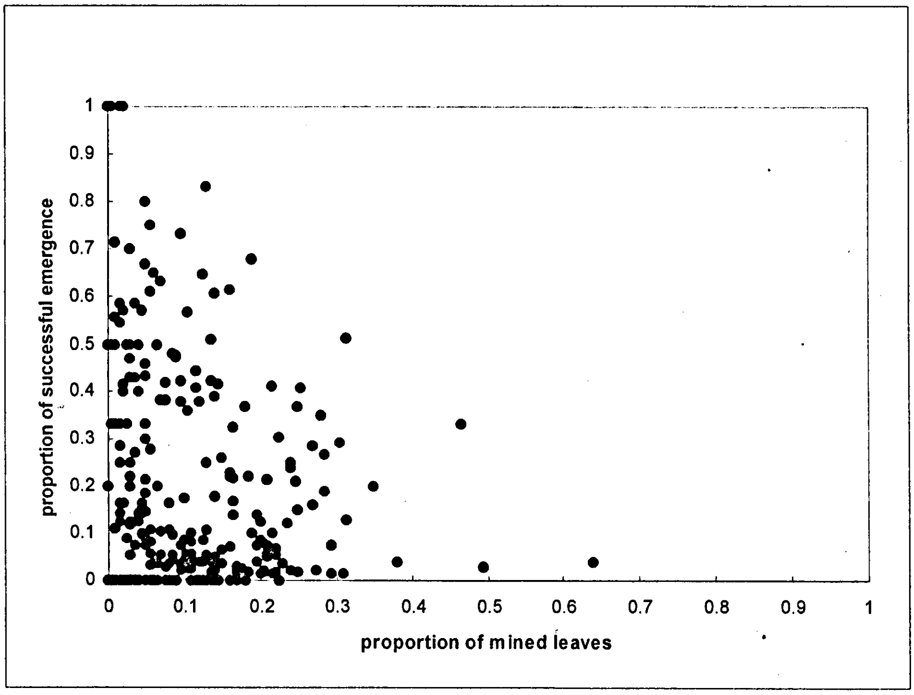


Figure 5.4 Plot of the proportion of successful emergence against the proportion of mined leaves for all trees sampled in the survey

| Leaf-miner data | Environmental variables | Co-ordinate terms | Deviance accounted for in leaf-miner data (%) | | | | Significance test |
|---|---|---|---|---|---|---|---|
| | | | Total | Env | Env * Space | Space | |
| Oviposition | *wt st wp sp sh* | $x\ y\ x^2\ y^2\ xy\ xy^2$ | 69.3 | 14.1 | 35.8 | 19.4 | $F_{5,80}=4.64$<br>$p<0.05$ |
| Miscellaneous larval mortality | *wt* | $x\ xy\ x^2\ x^2y\ xy^2\ x^3\ y^3$ | 64.6 | 1.6 | 12.2 | 50.8 | $F_{1,48}=0.561$<br>NS |
| Larval parasitism | *wt wh* | $x\ y\ y^2$ | 77.9 | 0.3 | 26.9 | 50.7 | $F_{2,49}=0.280$<br>NS |
| Bird predation | *st sp wh* | $x^2\ x^2y\ y^3$ | 47.1 | 13.6 | 32.7 | 0.8 | $F_{3,50}=2.67$<br>NS |
| Pupal parasitism | *sp wh* | $x\ xy\ xy^2$ | 39.4 | 2.4 | 16.9 | 20.1 | $F_{2,51}=0.560$<br>NS |
| Miscellaneous pupal mortality | *alt* | $x\ y\ xy\ xy^2$ | 36.0 | 0 | 18.3 | 17.7 | $F_{1,50}=1.69$<br>NS |
| Successful emergence | *wt st* | $x\ x^2\ xy\ xy^2\ x^3$ | 68.8 | 1.7 | 11.9 | 55.2 | $F_{2,49}=0.937$<br>NS |

Table 5.2 Results of the partial regressions of *Phytomyza ilicis* oviposition, mortality and successful emergence on environmental data taking into account a spatial component. Environmental variables: *wt* = winter temperature; *st* = summer temperature; *wp* = winter precipitation; *sp* = summer precipitation; *wh* = winter humidity; *sh* = summer humidity; *alt* = altitude.

However, with the exception of miscellaneous larval mortality and larval parasitism, each of the mortality components are largely independent of one another. The consequences of this are that sites close together are likely to have similar mortality profiles whereas populations separated by large distances experience very different patterns of mortality. The general picture that emerges then is that, at the broadest spatial scale, the range of *P. ilicis* is set upon a complex yet structured template of mortality with populations experiencing different mortality profiles dependent upon their location.

There does not appear to be a simple relationship between any of the components of mortality and population densities between sites. No single component of mortality that we were able to measure during the survey can be responsible for regulating leaf-miner populations across the whole range and hence creating the spatial abundance structure that we have recorded. This should not be surprising given the geographic variation in the distribution of natural enemies and climate. In addition, previous studies have shown that different factors may be responsible for determining the range limits of a species in different regions (MacArthur, 1972; see Hoffmann & Blows, 1994 for a brief review) and this may also be true for determining variation in local population densities.

### Oviposition

Oviposition levels necessarily set an upper bound to leaf-miner densities since the two were measured together during the survey. Across all sites, 48% of all eggs, as indicated by oviposition scars, successfully hatched to become larvae visible in the leaf lamina. This is consistent with the findings of Valladares & Lawton (1991). However, a plot for all of the trees in the survey shows that the proportion of eggs that hatch successfully decreases at higher leaf-miner densities (figure 5.5).

The results of the partial regression analysis (table 5.2) show that a significant effect of environmental variation on oviposition rate can be separated from the potentially confounding effect of spatial position at this geographic scale. This is in accordance with the results from the density survey (Brewer & Gaston, *manuscript*).

### Miscellaneous larval mortality

Miscellaneous larval mortality is thought to occur relatively early in the life-cycle since mines suffering from this mortality are typically small. It may be due to a number of possible causes, including death during the overwintering phase of the life-cycle, as

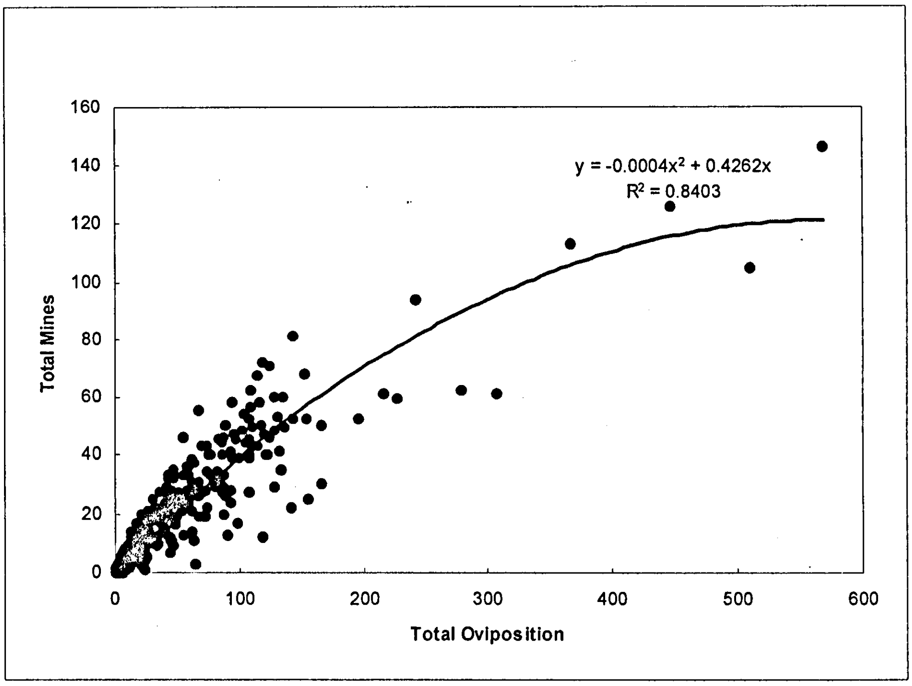


Figure 5.5 Regression of total number of mines on total oviposition for all trees ($p<0.001$).

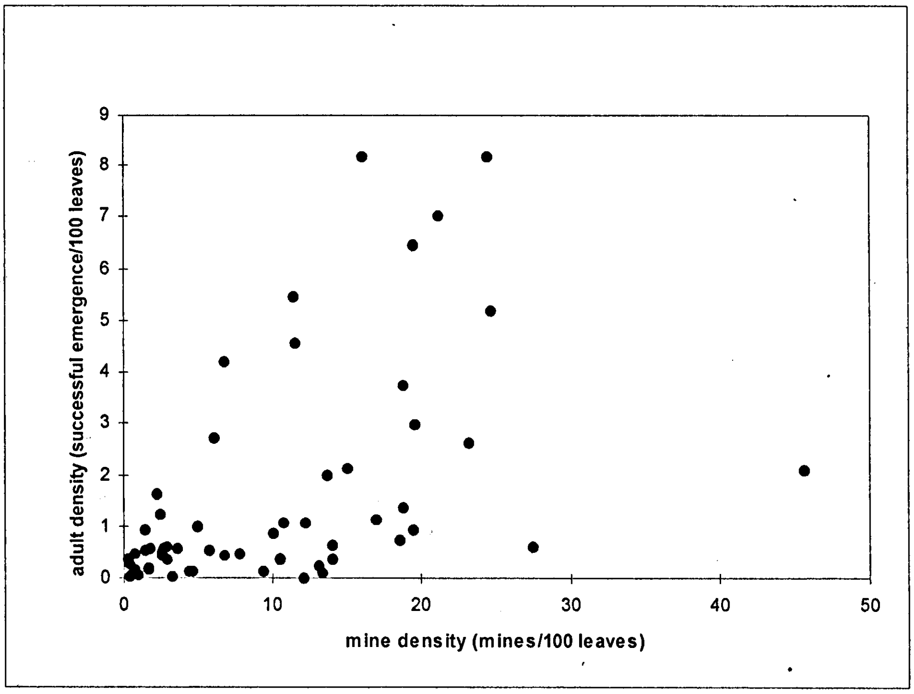


Figure 5.6 Plot of adult density (calculated as the number of leaf-miners that successfully emerge per 100 leaves against leaf-miner density at all sites

the result of a physiological response of the host tree or from attacks by natural enemies. Miscellaneous larval mortality in *P. ilicis* has previously been found to be density-independent at local spatial scales (Heads & Lawton, 1983) so that it seems unlikely that it is a consequence of intraspecific competition. We did not find any evidence for density-dependence across the whole range (r=0.002 , n=57 , p=1.00 )

There is, however, a positive correlation between miscellaneous larval mortality and larval parasitism by *Chrysocharis gemma* (figure 5.3). While some degree of miscellaneous mortality is present at most sites, regions with high rates of larval parasitism tend also to have higher rates of miscellaneous larval mortality. It is therefore possible that a component of miscellaneous larval mortality may be due to attacks by *C. gemma*, when it is ovipositing into the leaf-miner larva but this would be in addition to other factors. The correlation could, of course, have no causative basis but exclusion experiments would help to determine the role (if any) of *C. gemma* in this component of mortality.

### Larval parasitism

*Chrysocharis gemma* is the only recorded larval parasite of the holly leaf-miner and has been introduced into Canada as a biological control agent in commercial holly crops (Hansson, 1985). Auerbach *et al.* (1995) point out that in nearly every study examining the mortality of leaf-miners, larval parasitism is an important source of mortality in leaf-miners and our data support this observation over a considerable proportion of the range of *P. ilicis*. Highest levels of larval parasitism are found in the centre of the range. However, it is almost completely absent from north-eastern regions. Moreover, leaf-miner densities are at their highest in the north-east which suggests that *C. gemma* may contribute towards reducing leaf-miner densities elsewhere. However, across the whole of the range, there is no obvious relationship between leaf-miner density and larval parasitism between sites (r=0.002, n=56, p>0.05). Heads & Lawton (1983) found that *C. gemma* exhibited an aggregative response to leaf-miners within a single habitat patch. We conducted regression analyses of *C. gemma* attack rates on both real and apparent leaf-miner densities between trees separately for 21 sites where *C. gemma* was present, but no significant positive relationships were found. However, the power of each statistical test was not high since we were limited to 10 trees at each site.

The absence mortality caused by *C. gemma* in the north-east is in accordance with its known distribution (Hansson, 1985). *C. gemma* overwinters as an adult in the evergreen host plants of the leaf-miners that it parasitises. It is possible that its absence from northern parts of the leaf-miner's range may be as a result of winter conditions being too extreme for it to survive in these regions. However, despite the fact that winter temperature and humidity did contribute significantly to the incidence of larval parasitism when considered alone, this effect could not be disentangled from the confounding effect of spatial position.

The difficulties we have encountered here in trying to discern the role of *C. gemma* in influencing leaf-miner populations and the factors that influence its own distribution highlight the problems that will often be encountered in such correlative studies. Experimental investigation may be the only way of resolving these two particular issues.

### Bird predation

*P. ilicis* can suffer high levels of bird predation in some UK habitats (Heads & Lawton, 1983; McGeoch & Gaston, 2000). Heads & Lawton (1983) found that bird predation accounted for approximately 40% of total mortality in their local study, although there was a great deal of variation between samples. In addition, they found a surprising negative relationship between mine density and levels of bird predation. However, they point out that the ability to detect density-dependent mortality and the nature of the relationship may depend on the spatial scale over which the samples are taken. We did not find such a relationship over a much broader spatial scale with our data.

Despite the fact that no obvious spatial structure could be discerned from the correlogram and the spatial component of variation in the partial regression analysis was small, bird predation was significantly higher at the edges of the range than elsewhere. This is highlighted well in the interpolated map (figure 5.2d).

### Pupal parasitism

Pupal parasitism can be an important source of mortality for leaf-miners (Auerbach *et al.*, 1995), often with species-rich parasitoid communities associated with a single leaf-miner species. The holly leaf-miner is no exception. In the UK, eight species of parasitoids have been recorded attacking the pupae of *P. ilicis* with considerable variation in the species found between regions (Cameron, 1939; KJG personal observation). We did not find significant spatial autocorrelation in the amount of pupal parasitism across the range. This is perhaps not surprising since if this component of mortality can be subdivided into individual mortalities as a result of each parasitoid species, and if each of these species has a range largely independent of the others, then one would not expect to find spatial structure in pupal parasitism across the whole range.

The map for pupal parasitism (figure 5.2e) does show an elevated incidence of pupal parasitism in the north-east. This may in part be due to the absence of larval parasitism which would leave more mines available to the pupal parasitoids.

### Miscellaneous pupal mortality

The causes of miscellaneous pupal mortality are unknown but may be due to attack by pupal parasitoids, predators or pathogenic fungi. We found significant spatial autocorrelation in levels of this mortality at short lags. However, there does not appear to be a simple pattern at larger scales..

### Successful emergence

Relatively high levels of successful emergence were found to the east of the range of *P. ilicis*. The weak negative relationship between maximum levels of successful emergence and local population densities is somewhat surprising. Intuitively, one might expect sites supporting high population densities to have a higher *per capita* rate of successful emergence, indicative of their suitability as habitats for *P. ilicis*. However, the converse seems to be true. All sites supporting high densities of leaf-miners have relatively high rates of mortality (per mine). However, the density of successful emergences (the number of adults produced at each site) is still generally higher at these sites (figure 5.2g).

Our picture of the *P. ilicis* populations across its range, although better than for most animal species, is, of course, not complete. There are presently no data on the longevity, dispersal and potential agents of mortality of adult *P. ilicis*. This stage necessarily represents a small proportion of the total life cycle since the availability of leaves for oviposition is limited to at most only a few weeks in the year. However, it should not automatically be assumed that this means that the stage is unimportant in determining population densities. In a review of leaf-miner population dynamics, Auerbach *et al.* (1995) note that host-plant phenology and adult fecundity may have important consequences on population sizes. These are difficult parameters to measure, although for *P. ilicis* at least there may be some value in measuring oviposition rates for two successive generations (the one in which leaf-miner densities and mortalities are recorded and the subsequent one). This would at least provide a measure of the mean number of eggs produced from each successful emergence and hence an estimate for the effects of all of these factors considered together.

The population dynamics of *Phytomyza ilicis* across Europe are complex and there is still some way to go before the spatial structure of its geographic range is properly understood. No single source of mortality appears to be responsible for limiting population numbers across the whole range. At any site, the mortality that a population suffers is the sum of largely independent yet spatially structured components. These components may be to some extent environmentally determined but it has not been possible to separate the environmental effects from the spatial structure they share with the rates themselves. This should not be surprising yet it highlights the difficulties that face extensive surveys of this nature.

However, the results of this initial survey are encouraging. The strong spatial structure that *P. ilicis* exhibits in a number of components of its demographic rates suggest that it may indeed be possible to characterise range structure in such terms. More intensive, surveys at a smaller number of selected sites across the geographic range and simple experiments would help to clarify the causes some of the patterns that are suggested here for *P. ilicis*. Testing the generality of these results for most other taxa may be more difficult, since *P. ilicis* is an unusually amenable system for studies of this nature.

# Chapter 6

# The interspecific abundance-distribution relationship and the vital rates model

> *"It is an occupational hazard of all computer simulators, and it is simply their high-tech version of the philosopher's fundamental foible: mistaking a failure of imagination for an insight into necessity. A prosthetically enhanced imagination is still liable to failure, especially if it is not used with sufficient rigor."*
>
> Dennett (1995)

## Introduction

The positive correlation found between the mean local abundance and the distribution (expressed either as the extent of occurrence or area of occupancy) within species assemblages is an extremely common pattern found at a broad range of spatial scales (see Gaston, 1996, for a review). Locally abundant species tend to be widespread and locally rare species tend to have restricted distributions. The relationship has potentially important consequences for a variety of ecological applications such as inventorying biodiversity, population monitoring, harvesting, conservation and the study of invasions (Gaston, 1999).

Several mechanisms have been proposed to explain this abundance-distribution relationship (Gaston *et al.*, 1997). Some suggest that it may not be real but arises as a result of sampling artefact, phylogenetic non-independence or the range position of each species within the region under study. The others propose ecological mechanisms based upon aggregation, resource breadth, resource availability, habitat selection, metapopulation dynamics and vital rates. The general consensus is that the pattern is real and ecologically interesting but it is currently not known which of the proposed mechanisms are making the most important contributions to it. Gaston *et al.* (1997) suggest four reasons why this should be so: the mechanisms have not been adequately tested empirically; it is not possible to discriminate between them, since their predictions may not be explicitly stated; none of the mechanisms may be appropriate; some of them may be complementary.

One of the most recently proposed and largely untested mechanisms for the abundance-distribution relationship is the vital rates model (Holt *et al.*, 1997). It emphasises that explanations for the patterns in the distribution and abundance of species must ultimately be reducible to assumptions about how the vital rates of birth, death and dispersal vary in space and time. This reasoning is similar to that used by Maurer and Brown (1989) who proposed that the pattern of a single species' local population densities across its geographic range may best be understood by examining the interaction between smoothly varying vital rates across space.

In this chapter the vital rates model is explored in some detail. It's robustness in predicting the positive abundance-range size relationship is tested by adding random variation from several sources to its parameters and by extending the model to encompass variation in more of the model parameters than originally presented.

*Description of the vital rates model*

Briefly, the vital rates model is based upon the following argument. The distribution of any species is given by the number of sites at which its intrinsic rate of increase, *r*, is greater than zero, whereas its equilibrium local abundance is often directly proportional to *r*. Any factor which tends to increase *r* across all sites will increase the number of sites potentially occupied and also increase abundance at each occupied site. This means that range size and local abundance are related. Consider a set of species that have had sufficient time to disperse to all sites where they can potentially exist in a biogeographic region and have population dynamics that are driven primarily by local birth and death rates rather than dispersal. If these species differ principally in their responses to density-independent factors influencing birth and death rates, but are similar in density dependence, then the positive abundance-distribution correlation should be observed.

Holt *et al.* (1997) develop the vital rates model more formally, using a logistic model for the population dynamics of each species at a number of sites along an environmental gradient and then considering the availability of these sites within the region under study.

The local population dynamics of a species is modelled with the logistic equation:

$$(1/N_i)(\mathrm{d}N_i/\mathrm{d}t) = (b_0(x) - d_i) - uN_i(x)$$

where $b_0(x)$ is the birth rate of the species at site $x$ at low densities (the basal birth rate), $u$ is the strength of the density dependent reduction in birth rates, and $d_i$ is a constant death rate for species $i$ at all sites. The equilibrium density for species $i$ at site $x$ is therefore equal to

$$N_i^*(x) = \left(b_0(x) - d_i\right)/u$$

Since the intrinsic rate of increase of the local population is

$$r_i(x) = b_0(x) - d_i$$

the equilibrium population density of species $i$ at site $x$ can be restated as

$$N_i^*(x) = r_i(x)/u$$

If the set of species all have the same basal birth rate at each site ($b_0(x)$) and each species can be ranked by its density-independent death rate ($d_i$), any site occupied by a species with the highest death rate can also be occupied by any lower ranked species, leading to a nested distribution with the range of species $i$ encompassing the ranges of all the higher ranked species. In addition, at any site occupied by two species $i$ and $i$+1, species $i$ will have the higher equilibrium abundance $N^*_i(x) > N^*_{i+1}(x)$ if $d_i < d_{i+1}$.

The final requirement for the model to yield the positive abundance-range size relationship is that the potentially habitable sites can be ordered along a gradient such that

$$b_0(x) = max\ b_0 - s|x|$$

where $s$ measures the sharpness (i.e. slope) of the gradient in birth rates (figure 6.1a). As Holt *et al.* (1997) point out, this assumption is equivalent to assuming that different habitat types (as assessed operationally by realised birth rates) are equally available. The range limit for species $i$ is determined by identifying the largest value of the gradient position $|x|$ such that the *per capita* growth rate there is just equal to zero

$$r_{0i} = \max b_0 - s|x| - d_i$$

or

$$|x| = \left(\max b_0 - d_i\right)/s$$

This is the distance from the point on the gradient with the highest density to the point of zero density and so is a measure of range size, $R$. In this simple model, the average density over the entire range, $\langle N_i \rangle$, is half the maximal local density so

$$\langle N_i \rangle = 1/2\left(b_0 - d_i\right)/u = 1/2\left(s/u\right)R$$

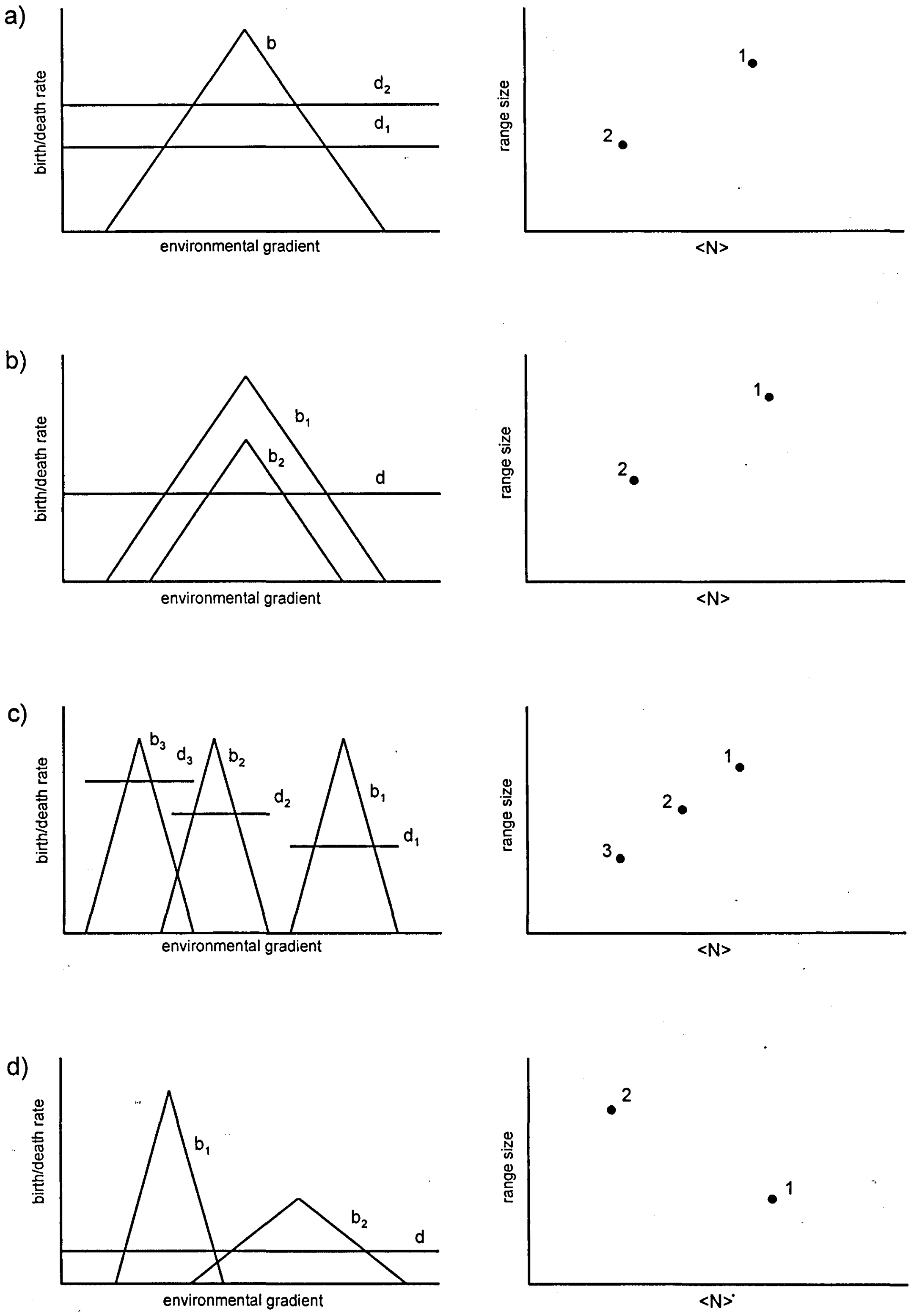


Figure 6.1. The vital rates model illustrated geometrically (a) Two species with identical basal birth rates along an environmental gradient but with different density independent death rates (b) Two species with different basal birth rates and the same death rate (c) The position of the vital rate response along the environmental gradient is not important (d) An example of a negative abundance-distribution relationship generated by the vital rates model.

Hence, given that species differ solely in rates of density-independent mortality, have similar spatial variation in birth rates, and different habitats are equally available, the model predicts that range size and mean local abundance are directly proportional (figure 6.1a).

The vital rates model can be elaborated upon graphically to show that a positive abundance-distribution relationship will result even if the levels of birth rates vary between species (as long as the slopes stay the same, figure 6.1b). In addition, the position of the vital rates responses along the gradient is unimportant as long as all sites are equally available to each species (figure 6.1c). The model can also account for cases where abundance and range size are negatively correlated (figure 6.1d).

## Method

*Modification of the vital rates model to allow for interspecific variation in both birth rates and death rates*

Perhaps the most obvious way to expand the original model is to allow density-independent death rates and basal birth rates to vary between both species and sites. The logistic model can be restated to reflect this

$$(1/N_i)(\mathrm{d}N_i/\mathrm{d}t) = (b_{0i}(x) - d_{0i}(x)) - uN_i(x)$$

Here, $d_{0i}(x)$ is the density-independent death rate for species $i$ at site $x$. Similarly, $b_{0i}(x)$ is the basal birth rate for species $i$ at site $x$. In addition, $u$ now reflects both the density dependent reduction in birth rate and the increase in death rate. This leads to an equilibrium density of

$$N_i^*(x) = (b_{0i}(x) - d_{0i}(x))/u$$

and since the intrinsic rate of increase is

$$r_i(x) = b_{0i}(x) - d_{0i}(x)$$

this can be restated as

$$N_i^*(x) = r_i(x)/u$$

which is equivalent to the original model given by Holt *et al.* (1997).

The ranking (and hence availability) of the sites to produce systematic response curves remains the same as in the original model since, as Holt *et al.* (1997) demonstrate, this is a particularly important requirement for the abundance-range size relationship to remain positive (see the discussion).

*The computer simulation*

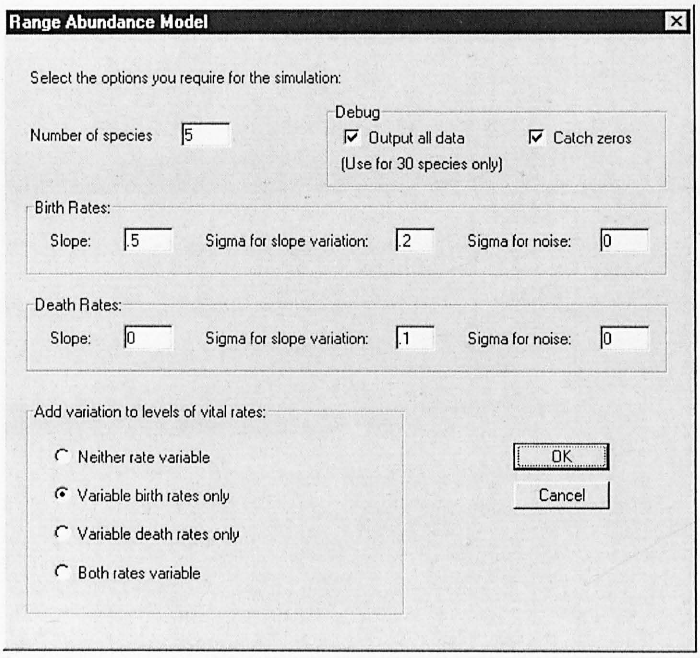
Since there can be a rather broad variety of scenarios with the modified model, depending on the nature of the response in birth rates and death rates, a simulation approach is used to determine the outcomes. The variation between each species is made manifest by adding variation to the slopes of the birth and death rate responses along the environmental gradient. In addition, random noise is added to the birth and death rates at each site to simulate temporal stochastic variation. The computer simulation allows easy adjustment of the model's parameters, reflecting different scenarios, and for the model to be repeated numerous times for a large number of species across a considerable number of potentially habitable sites.

The simulation was written using Visual Basic for Applications (VBA) in Microsoft Excel 95. This language provides relatively easy access to useful resources such as dialog boxes and graphical output as well as the considerable data-handling and statistical capabilities of a spreadsheet. The essential features of the computer implementation of the vital rates model are outlined below. A complete program listing, along with its accompanying internal documentation, may be found in Appendix 2.

The heart of the simulation, like the vital rates model itself, is relatively simple. The environmental gradient is represented as a set of 2001 discrete sites. For each species an array representing birth rates and death rates at every site is filled, incorporating the variation between species selected in the dialog box. After this any random noise is added to these slopes. Then $r$ is calculated, by subtracting the local death rate from the local birth rate. Range size is calculated simply by counting the number of sites where $r$ is greater than zero. Since density dependent losses are treated as a constant for all species in the vital rates model, the population density at each site is proportional to $r$. Therefore a regression of range size on mean $r$ is equivalent to regressing it on mean population density.

Figure 6.2 shows the dialog which appears when the program is run, presenting the various parameters that may be set for the simulation. These include: the number of species in the assemblage; the mean and standard deviation in the slopes of both the birth and death rates curves; and the amount of random noise to be added to each of these curves. The levels (i.e. the maximum height of the curves before the addition of noise) of both the birth and death rates curves can also be set to be variable.

When the parameters have been set in the dialog box and the simulation is started, the birth and death rates for each species at each site are initialised. The mean birth rate



Range Abundance Model

Select the options you require for the simulation:

Number of species 5

Debug

☑ Output all data ☑ Catch zeros

(Use for 30 species only)

Birth Rates:

Slope: .5 Sigma for slope variation: .2 Sigma for noise: 0

Death Rates:

Slope: 0 Sigma for slope variation: .1 Sigma for noise: 0

Add variation to levels of vital rates:

○ Neither rate variable
● Variable birth rates only
○ Variable death rates only
○ Both rates variable

Figure 6.2. The dialog box for the vital rates model simulation, illustrating the parameters that may be set at the beginning of the program. Refer to the main text for a full explanation.

slope for the whole assemblage is taken from the dialog box and for each species this slope is modified by adding a value from the normal distribution with mean zero and standard deviation specified in the dialog box. The slope for the birth rates is constrained to be positive before the central site (site 1001). Next, the maximum, central value for the birth rate of each species is established by taking a random number between a minimum and maximum limit (see below) from the uniform distribution. The birth rate for this species at all of the sites is then calculated using the value for the slope and its value at the central site. For death rates, the procedure is essentially the same with the exception that the central site death rate is not necessarily the maximum value since the simulation allows the death rate slope to be positive or negative before the central site. Finally, if random noise on the curves has been selected in the dialog box, this is added to the birth and/or death rates for each species at each site. This random noise is sampled from a normal distribution with mean zero and standard deviation specified in the dialog box.

Once the values for births and deaths have been initialised, $r_i(x)$ is calculated ($r$ for each species at each site). This is simply the difference between the birth and death rates unless this value is negative in which case it is set to zero. It is then a simple matter to calculate the mean value of $r$ (at all sites where $r>0$) and range size. The range size of each species is measured in two ways, analogous to those highlighted by Gaston (1991): area of occupancy is the number of sites where $r>0$; extent of occurrence is the distance between the two most distal sites where $r>0$. (see discussion about extent of occurrence). When the values for mean $r$ and range size have been calculated for all species, the linear regression functions of Microsoft Excel are used to obtain the value for the regression coefficient, the coefficient of determination ($r^2$) and the probability of the abundance-distribution relationship. The scatter-plot of mean $r$ against range size for all the species is drawn. In addition, the values for birth rates, death rates and $r$ are plotted on separate charts.

### *Constraints on the simulation*

Holt *et al.* (1997) do not discuss constraints on the species' vital rates in their original description of the model since the paper is intended simply to introduce the idea that the abundance-distribution relationship may best be understood in terms of these vital rates. However, when implementing a computer simulation of a mathematical model, some constraints on the amount of variation possible must be imposed in order to prevent both numerical errors and to allow the simulation to finish within a reasonable time. It is extremely important to bear such constraints in mind to avoid over-interpretation of results that may have simply arisen as artefacts of the simulation. Below the constraints of the simulation model are stated explicitly.

1. Number of sites: The maximum number of sites that any species may occupy in the simulation is 2001. This number was chosen as a balance between allowing sufficient variation in the range sizes that the assemblage of species could exhibit and the time taken to run each simulation, since a number of replicates for each combination of parameters were run.

2. Maximum and minimum values for each vital rate: Maximum birth rates were constrained to be integer values lying between 25 and 100. The absolute values are of course arbitrary but this interval was chosen so that a reasonable amount of variation (75 levels) within the assemblage could be simulated. Death rates were extrapolated from a random integer value at the central site between 1 and the maximum birth rate for each species. This of course prevents any species having zero abundance and zero range size i.e. not existing at all on the environmental gradient.

3. Maximum vital rates slope: In the original model the environmental gradient and vital rates slopes are continuous whereas in the simulation there are discrete sites that a population may occupy. If the vital rates slopes are allowed to get too high then numerical errors may result. Figure 6.3a shows three species with variable birth rate slopes and levels and a constant death rate. Note that, despite the fact that mean $r$ would be different between the species and that the ranges on a continuous gradient would also be different, the simulation would report all species having a range size of 4, since all species have positive $r$ at site 2 but negative $r$ at site 1 (similarly for sites 6 and 7). If all of the birth slopes in the simulation are this similar to each other, the resulting abundance-distribution relationship would report mean $r$ and range size to be independent of each other due to the rounding error. In less extreme circumstances the abundance-distribution relationship would at least be weakened if a subset of species had similar slopes. To overcome this potential artefact, the birth slopes were constrained to have a maximum gradient of 1.

4. Minimum vital rates slope: Again considering a simple scenario of variable birth rates and a constant death rate for each species, a potential problem arises if the slopes of the birth rates are so shallow that they do not cross the death rate line within the spatial limits of the simulation. In figure 6.3b, species 1 and 2 have the maximum range size

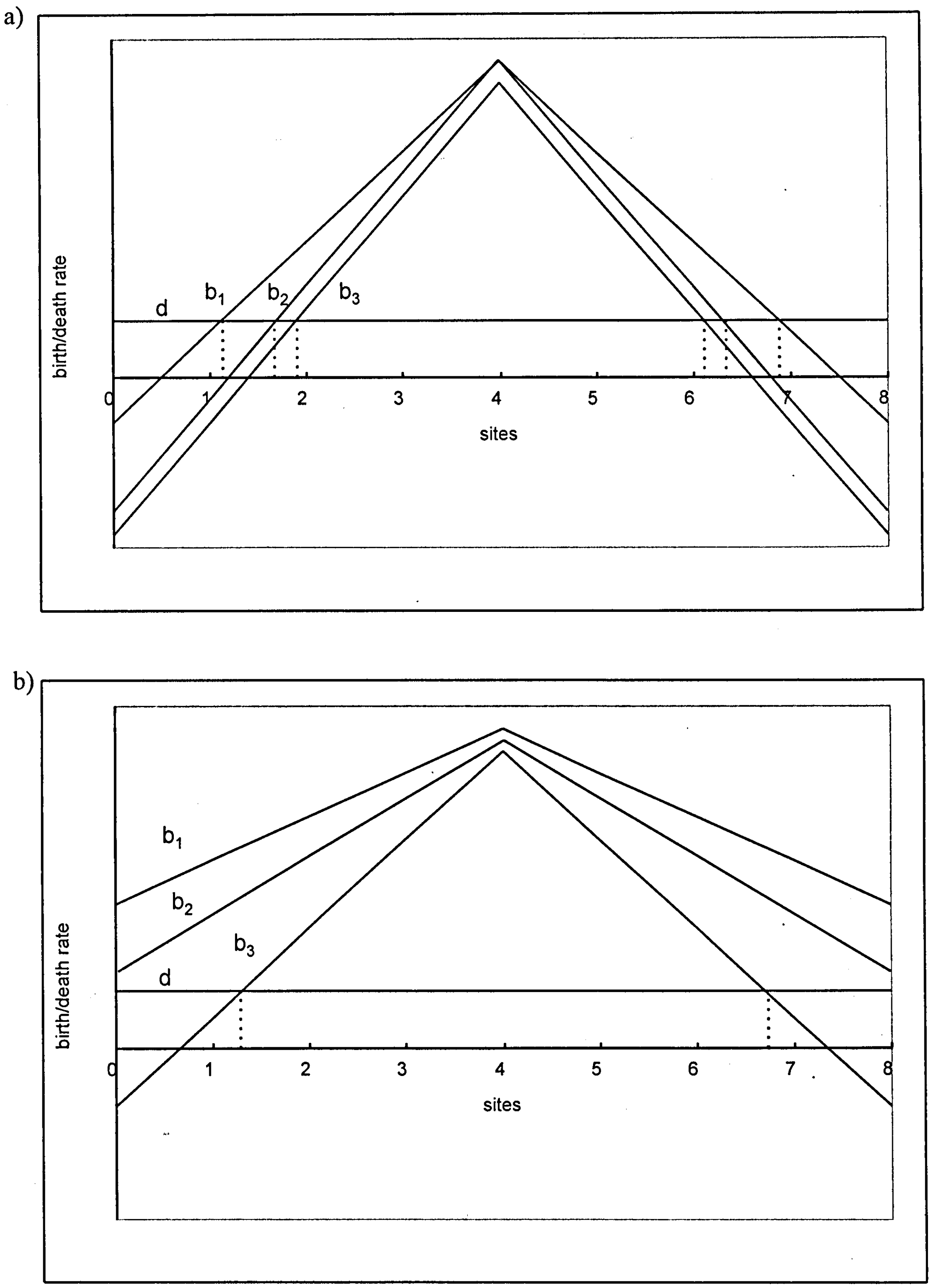


Figure 6.3. Illustration of the need to apply constraints on the birth rates curves. (a) Constraining the maximum slope. All three species have different birth rates, hence mean *r*, but because sites are represented as discrete units, their range sizes are all equal in the simulation. (b) Constraining the minimum slope. Two species have birth slopes that do not cross the death rate line within the spatial limits of the simulation. Therefore, despite having different values of mean *r*, their range sizes would be equal (the maximum possible).

but different values for mean *r*. For a large number of species, this would produce an abundance-distribution relationship which levels out at the maximum range size (see discussion). A constraint on the minimum birth slope dependent upon the number of sites available in the simulation was imposed to prevent this occurring (see internal documentation, Appendix 2). The value for the minimum slope with 2001 sites was 0.1.

## Results

### *The scenarios simulated*

For clarity, each of the six scenarios simulated are illustrated in figure 6.4. Note that, since the position of each species on the gradient is not important as long as each environmental combination has equal representation (figure 6.1c), all species are plotted in the same position.

Each scenario was simulated for n=30 and n=10 species chosen as a statistically 'adequate' sample size and a small sample size, respectively. The mean birth rate slope was set to 0.5 and the mean death rate slope to 0. Variation ($\sigma_b$ and $\sigma_d$) was added to these slopes in increments of 0.05 up to a value of 0.35, which is near the maximum amount of variation this implementation can accommodate given the constraints on the slopes discussed above. Site-specific stochastic variation (hereafter, noise) was added at three levels; 0, 5 and 15. Given that the maximum birth rate can be initialised between 25 and 75, the highest noise level represents a considerable disruption of the shapes of the vital rates curves. For each combination of parameters the simulation was run 10 times initially. Subsequently a subset of these were run 100 times to look in more detail at any interesting patterns that arose.

### *Scenario 1: Variable birth rate slopes and a constant death rate (figure 6.4a)*

Table 6.1a summarises the results of the simulation for 30 species. Since no noise was added to the vital rates slopes, area of occupancy and extent of occurrence are equivalent. As the amount of variation on the slopes of the birth rates increases, the coefficient of the regression of range size on mean *r* (mean local abundance as far as this simulation is concerned) tends to increase and the amount of variation accounted for by this fit tends to decrease. Only at the highest levels of variation in the birth rate slopes are any of the relationships not significantly positive and even at these levels this happens rarely. There were no cases of a significant negative relationship during the course of these simulations.

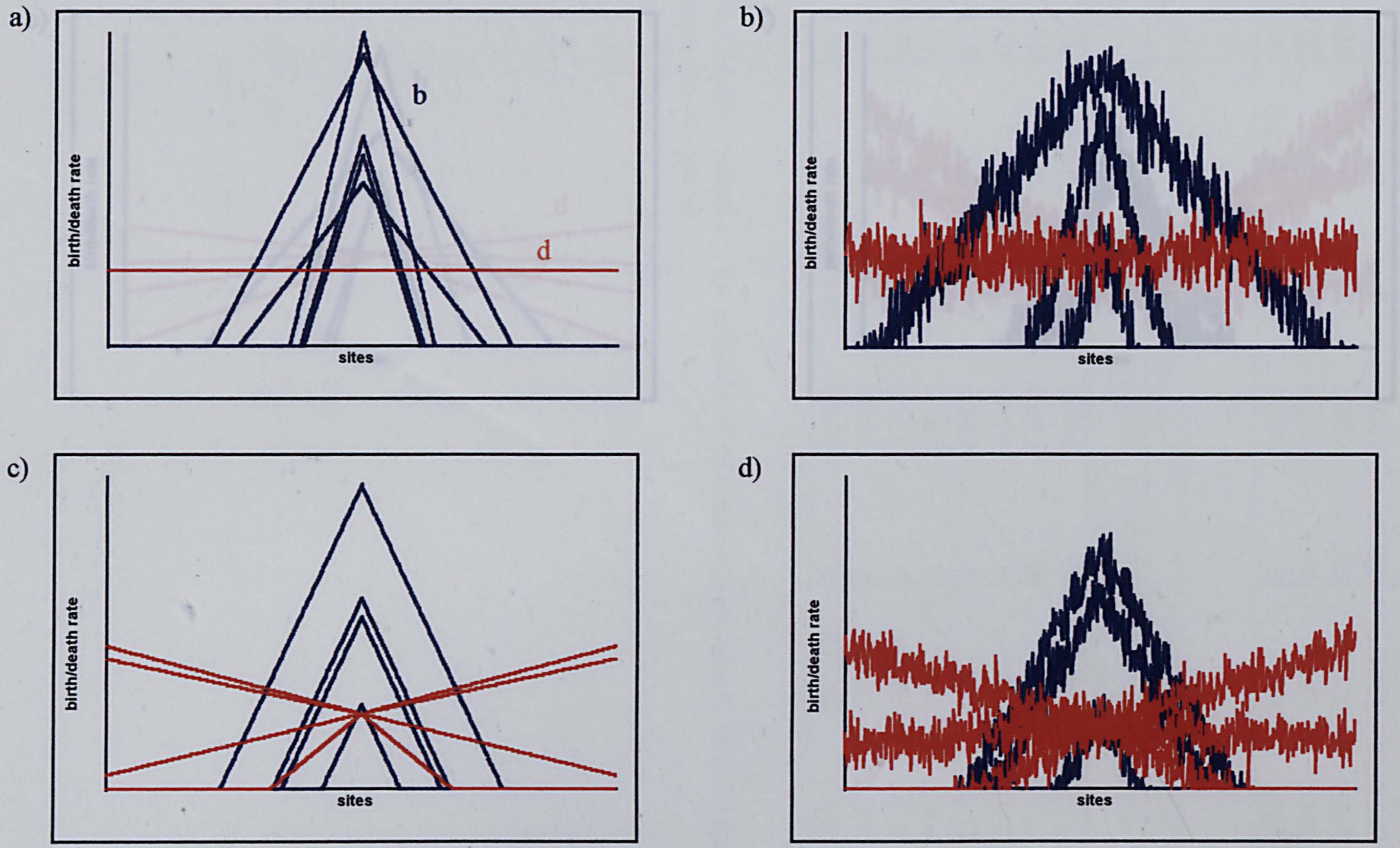


Figure 6.4. The scenarios from the extended vital rates model that are simulated (a) Scenario 1. Variable birth rate slopes between species and constant death rate (b) Scenario 2. Variable birth rate slopes and constant death rate with noise added (c) Scenario 3. Fixed birth rate slopes and variable death rate slopes (d) Scenario 4. Fixed birth rate slopes and variable death rate slopes with noise added.

*Figure continued on the next page.*

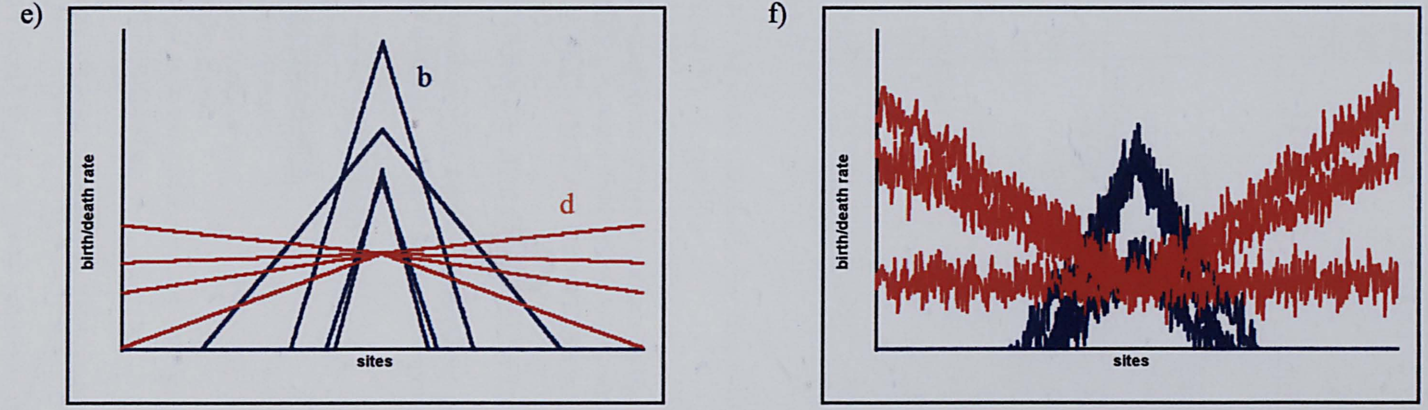


Figure 6.4 *(continued)*. The scenarios from the extended vital rates model that are simulated (e) Scenario 5. Variable birth and death rate slopes (f) Scenario 6. Variable birth and death rate slopes with noise added.

*Scenario 1: Variable birth rate slopes and a constant death rate*

a)

| n=30 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_b$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.002 | 1.000 | 0.000 | 10 | 8.002 | 1.000 | 0.000 | 10 |
| 0.05 | 8.113 | 0.958 | 0.000 | 10 | 8.113 | 0.958 | 0.000 | 10 |
| 0.10 | 8.501 | 0.886 | 0.000 | 10 | 8.501 | 0.886 | 0.000 | 10 |
| 0.15 | 9.383 | 0.632 | 0.000 | 10 | 9.383 | 0.632 | 0.000 | 10 |
| 0.20 | 9.869 | 0.357 | 0.001 | 10 | 9.869 | 0.357 | 0.001 | 10 |
| 0.25 | 11.038 | 0.335 | 0.001 | 10 | 11.038 | 0.335 | 0.001 | 10 |
| 0.30 | 9.982 | 0.369 | 0.000 | 9 | 9.982 | 0.369 | 0.000 | 9 |
| 0.35 | 12.087 | 0.285 | 0.003 | 9 | 12.087 | 0.285 | 0.003 | 9 |

b)

| n=10 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_b$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.001 | 1.000 | 0.000 | 10 | 8.001 | 1.000 | 0.000 | 10 |
| 0.05 | 7.936 | 0.982 | 0.000 | 10 | 7.936 | 0.982 | 0.000 | 10 |
| 0.10 | 8.296 | 0.832 | 0.001 | 10 | 8.296 | 0.832 | 0.001 | 10 |
| 0.15 | 9.008 | 0.712 | 0.002 | 8 | 9.008 | 0.712 | 0.002 | 8 |
| 0.20 | 7.342 | 0.569 | 0.012 | 8 | 7.342 | 0.569 | 0.012 | 8 |
| 0.25 | 16.955 | 0.472 | 0.029 | 8 | 16.955 | 0.472 | 0.029 | 8 |
| 0.30 | 12.971 | 0.367 | 0.064 | 3 | 12.971 | 0.367 | 0.064 | 3 |
| 0.35 | 17.759 | 0.349 | 0.073 | 4 | 17.759 | 0.349 | 0.073 | 4 |

Table 6.1 Median values for the regressions of range-size on mean $r$ after 10 simulations for (a) 30 species (b) 10 species.

For 10 species (table 6.1b), a similar increase in the slope of the final regression and decrease in the amount of variation explained occurs. The number of regressions which are not significantly positive is higher which should not be surprising since the number of degrees of freedom in the regression is lower. Again, there were no cases of a significant negative relationship.

*Scenario 2: Variable birth rate slopes and a constant death rate with noise added (figure 6.4b)*

For area of occupancy, the addition of both the low and high levels of noise to the birth and death rates curves had little effect on the results of the simulations for 30 species (tables 6.2a and c) other than to reduce the amount of variation explained in the final regressions slightly. This is obviously more noticeable at the higher level of noise. For extent of occurrence, however, a larger number of the regressions failed to be significantly positive as the amount of variation in the birth rate slopes increased, again this is more noticeable at the higher levels of noise. With 10 species (tables 6.2b and d) these effects are exaggerated with few regressions significant at the highest levels of $\sigma_b$ and noise.

*Summary of scenarios 1 and 2*

Selected combinations of parameters from scenarios 1 and 2 were repeated with 100 replicates to clarify the trends seen. The box-whisker plots (Tukey, 1977) given in figure 6.5 illustrate the effect of increasing $\sigma_b$ on the regression coefficient of the resulting mean $r$ versus range size relationship.

Figure 6.5a confirms that for 30 species, despite some increase in the variability of the regression coefficient, this is limited and failure to meet significance at the 5% level occurs only rarely. For 10 species (figure 6.5b) much more variation is observed in the final regression coefficient at higher levels of $\sigma_b$ with some instances of a negative relationship between mean $r$ and range size. If noise is added to the vital rates slopes, variation increases slightly in the regression coefficients when considering area of occupancy (figure 6.5c), and this effect is exaggerated when measuring range size as extent of occurrence (figure 6.5d).

*Scenario 2: Variable birth rate slopes and a constant death rate with noise added*

a)

| n=30 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_b$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.503 | 0.990 | 0.000 | 10 | 8.476 | 0.976 | 0.000 | 10 |
| 0.05 | 8.651 | 0.941 | 0.000 | 10 | 8.720 | 0.909 | 0.000 | 10 |
| 0.10 | 8.914 | 0.797 | 0.000 | 10 | 8.973 | 0.713 | 0.000 | 10 |
| 0.15 | 10.574 | 0.609 | 0.000 | 10 | 11.462 | 0.477 | 0.000 | 10 |
| 0.20 | 10.182 | 0.440 | 0.000 | 10 | 9.853 | 0.243 | 0.006 | 9 |
| 0.25 | 13.230 | 0.342 | 0.001 | 10 | 13.460 | 0.182 | 0.019 | 8 |
| 0.30 | 10.365 | 0.249 | 0.005 | 9 | 9.873 | 0.140 | 0.043 | 5 |
| 0.35 | 10.903 | 0.236 | 0.007 | 10 | 11.037 | 0.158 | 0.030 | 6 |

b)

| n=10 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_b$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.273 | 0.988 | 0.000 | 10 | 8.374 | 0.969 | 0.000 | 10 |
| 0.05 | 8.120 | 0.942 | 0.000 | 10 | 8.062 | 0.914 | 0.000 | 10 |
| 0.10 | 9.981 | 0.855 | 0.000 | 10 | 10.117 | 0.777 | 0.001 | 10 |
| 0.15 | 9.459 | 0.705 | 0.003 | 9 | 9.977 | 0.634 | 0.006 | 8 |
| 0.20 | 10.074 | 0.660 | 0.005 | 8 | 10.706 | 0.543 | 0.015 | 6 |
| 0.25 | 11.520 | 0.352 | 0.074 | 4 | 12.086 | 0.300 | 0.104 | 3 |
| 0.30 | 9.984 | 0.373 | 0.065 | 5 | 10.326 | 0.298 | 0.106 | 2 |
| 0.35 | 9.019 | 0.160 | 0.253 | 1 | 7.322 | 0.068 | 0.466 | 0 |

Table 6.2(a-b) Median values for the regressions of range-size on mean *r* after 10 simulations for (a) 30 species, noise level 5 added (b) 10 species, noise level 5 added

*Scenario 2 (continued)*

c)

| n=30 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_b$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 9.454 | 0.957 | 0.000 | 10 | 9.483 | 0.883 | 0.000 | 10 |
| 0.05 | 9.547 | 0.927 | 0.000 | 10 | 10.332 | 0.796 | 0.000 | 10 |
| 0.10 | 9.780 | 0.805 | 0.000 | 10 | 10.466 | 0.505 | 0.000 | 10 |
| 0.15 | 10.311 | 0.617 | 0.000 | 10 | 11.618 | 0.328 | 0.001 | 9 |
| 0.20 | 12.586 | 0.299 | 0.002 | 9 | 13.349 | 0.100 | 0.093 | 4 |
| 0.25 | 16.162 | 0.273 | 0.003 | 10 | 18.869 | 0.145 | 0.056 | 5 |
| 0.30 | 13.129 | 0.213 | 0.010 | 10 | 13.684 | 0.058 | 0.201 | 1 |
| 0.35 | 16.198 | 0.231 | 0.007 | 9 | 16.409 | 0.081 | 0.129 | 4 |

d)

| n=10 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_b$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 9.019 | 0.968 | 0.000 | 10 | 9.874 | 0.905 | 0.000 | 10 |
| 0.05 | 9.489 | 0.953 | 0.000 | 10 | 10.024 | 0.829 | 0.000 | 10 |
| 0.10 | 9.098 | 0.863 | 0.000 | 9 | 8.832 | 0.560 | 0.013 | 8 |
| 0.15 | 9.484 | 0.521 | 0.019 | 7 | 9.910 | 0.148 | 0.273 | 3 |
| 0.20 | 12.258 | 0.423 | 0.042 | 8 | 14.633 | 0.316 | 0.092 | 2 |
| 0.25 | 11.177 | 0.407 | 0.050 | 5 | 7.475 | 0.093 | 0.395 | 3 |
| 0.30 | 18.426 | 0.234 | 0.207 | 4 | 20.207 | 0.151 | 0.338 | 3 |
| 0.35 | 14.830 | 0.282 | 0.115 | 1 | 19.093 | 0.166 | 0.246 | 0 |

Table 6.2(c-d) Median values for the regressions of range-size on mean *r* after 10 simulations for (c) 30 species, noise level 15 added (d) 10 species, noise level 15 added

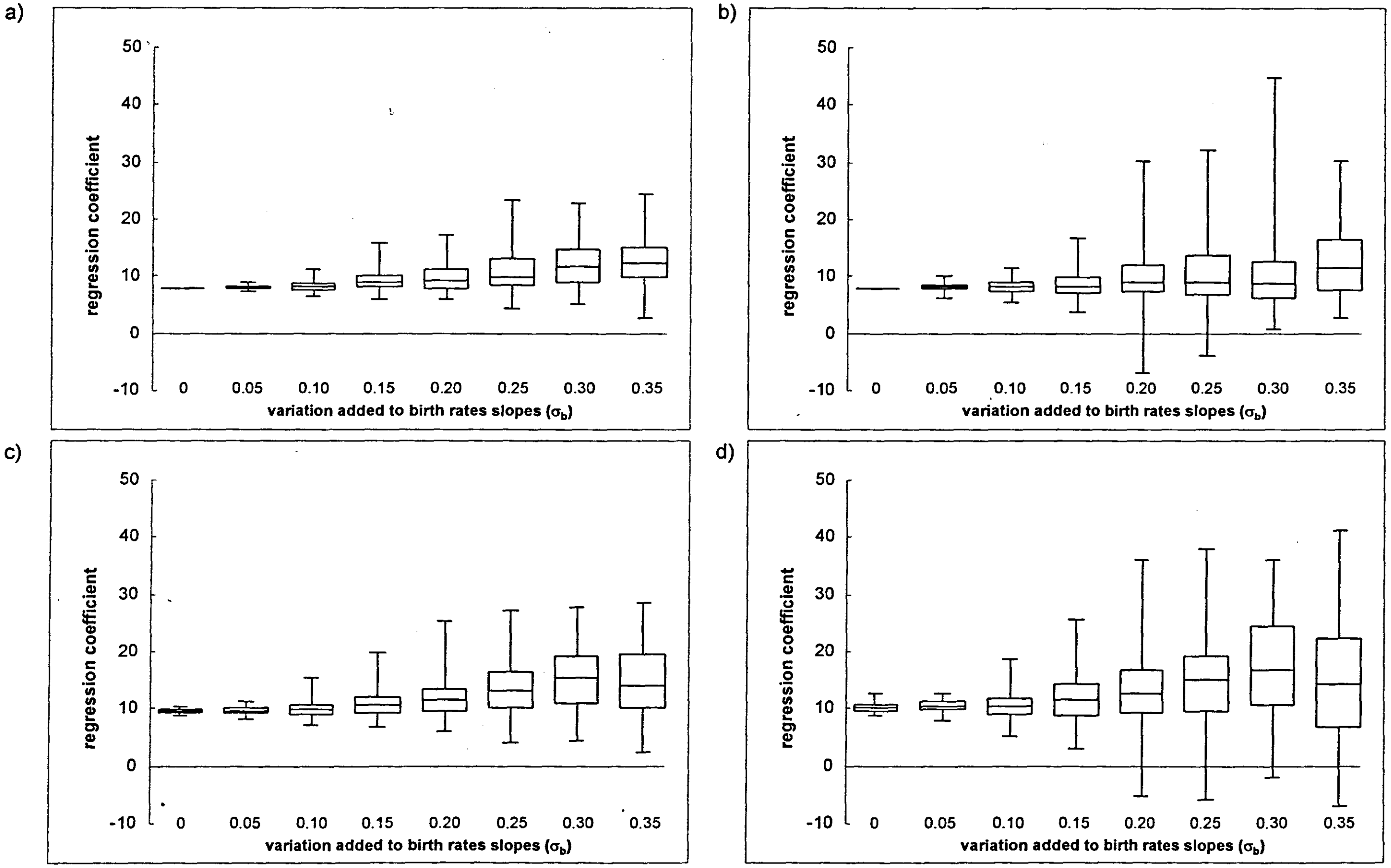


Figure 6.5 Change in the regression coefficient of the abundance-distribution relationship as more variation ($\sigma_b$) is applied to the birth rate slopes. (a) Area of occupancy, 30 species, no noise added (b) Area of occupancy, 10 species, no noise added. (c) Area of occupancy, 30 species, noise = 15 added. (d) Extent of occurrence, 30 species, noise = 15 added. In all cases, 100 simulations were run for each $s_b$ shown and the death rates were held constant.

*Scenario 3: Variable death rate slopes and a constant birth rate slope (figure 6.4c)*

The results from this set of simulations with 30 species (table 6.3a) closely mirror those for scenario 1. Increasing the variability in the slopes of the death rate curves ($\sigma_d$) results in a decrease in the amount of variation explained in the final regression of ranges size on mean *r*. In addition there is a concomitant rise in the median regression coefficient. However, for 10 species there appears to be little trend in runs of 10 replicate simulations (table 6.3b).

*Scenario 4: Variable death rate slopes and a constant birth rate slope with noise added (figure 6.4d)*

Addition of noise to the vital rates curves does not produce the same effect as in scenario 2, neither for 30 species nor 10 species (tables 6.4a-d). The most striking feature of these results is perhaps that very few of the simulations at either level of noise or number of species failed to produce a significantly positive relationship between range-size and mean *r*. In addition, the regression coefficient remains relatively constant as $\sigma_d$ increases. The $r^2$ values, however, do show a gradual decrease as $\sigma_d$ increases as might be expected.

*Summary of scenarios 3 and 4*

Selected combinations of parameters from scenarios 1 and 2 were repeated with 100 replicates to clarify the trends seen. Figure 6.6 gives box-whisker plots for the results of these runs. Figure 6.6a closely resembles the results for scenario 1 (figure 6.5a) with 30 species. However, for 10 species (figure 6.6b), despite a similar overall trend in the median regression coefficients, the variability in the final regressions is much greater than that observed in scenario 1 (compare with figure 6.5b). This increased variability explains why a trend was not evident in the runs involving only 10 replicates of each parameter combination (table 6.3b). Finally, when the highest level of noise is added, the median values for the regression coefficients remain almost constant (perhaps decreasing slightly) as $\sigma_d$ increases. They also show remarkably little variability at all levels of $\sigma_d$.

*Scenario 3: Variable death rate slopes and a constant birth rate slope*

a)

| n=30 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_d$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.002 | 1.000 | 0.000 | 10 | 8.002 | 1.000 | 0.000 | 10 |
| 0.05 | 8.025 | 0.967 | 0.000 | 10 | 8.025 | 0.967 | 0.000 | 10 |
| 0.10 | 7.986 | 0.837 | 0.000 | 10 | 7.986 | 0.837 | 0.000 | 10 |
| 0.15 | 9.091 | 0.619 | 0.000 | 10 | 9.091 | 0.619 | 0.000 | 10 |
| 0.20 | 10.105 | 0.334 | 0.001 | 9 | 10.105 | 0.334 | 0.001 | 9 |
| 0.25 | 12.818 | 0.568 | 0.000 | 8 | 12.818 | 0.568 | 0.000 | 8 |
| 0.30 | 16.802 | 0.449 | 0.000 | 10 | 16.802 | 0.449 | 0.000 | 10 |
| 0.35 | 16.622 | 0.619 | 0.000 | 10 | 16.622 | 0.619 | 0.000 | 10 |

b)

| n=10 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_d$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.002 | 1.000 | 0.000 | 10 | 8.002 | 1.000 | 0.000 | 10 |
| 0.05 | 7.615 | 0.974 | 0.000 | 10 | 7.615 | 0.974 | 0.000 | 10 |
| 0.10 | 8.036 | 0.887 | 0.000 | 10 | 8.036 | 0.887 | 0.000 | 10 |
| 0.15 | 10.298 | 0.706 | 0.003 | 9 | 10.298 | 0.706 | 0.003 | 9 |
| 0.20 | 8.661 | 0.714 | 0.002 | 9 | 8.661 | 0.714 | 0.002 | 9 |
| 0.25 | 16.465 | 0.670 | 0.004 | 8 | 16.465 | 0.670 | 0.004 | 8 |
| 0.30 | 8.019 | 0.280 | 0.128 | 6 | 8.019 | 0.280 | 0.128 | 6 |
| 0.35 | 11.061 | 0.903 | 0.000 | 9 | 11.061 | 0.903 | 0.000 | 9 |

Table 6.3 Median values for the regressions of range-size on mean *r* after 10 simulations for (a) 30 species (b) 10 species.

*Scenario 4: Variable death rate slopes and a constant birth rate slope with noise added*

a)

| n=30 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_d$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.503 | 0.990 | 0.000 | 10 | 8.476 | 0.976 | 0.000 | 10 |
| 0.05 | 8.584 | 0.952 | 0.000 | 10 | 8.749 | 0.919 | 0.000 | 10 |
| 0.10 | 8.461 | 0.906 | 0.000 | 10 | 8.327 | 0.854 | 0.000 | 10 |
| 0.15 | 8.175 | 0.810 | 0.000 | 10 | 7.707 | 0.723 | 0.000 | 10 |
| 0.20 | 8.709 | 0.799 | 0.000 | 10 | 8.367 | 0.735 | 0.000 | 10 |
| 0.25 | 8.138 | 0.799 | 0.000 | 10 | 7.938 | 0.733 | 0.000 | 10 |
| 0.30 | 8.137 | 0.708 | 0.000 | 10 | 7.684 | 0.654 | 0.000 | 10 |
| 0.35 | 7.987 | 0.701 | 0.000 | 10 | 7.564 | 0.618 | 0.000 | 10 |

b)

| n=10 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_d$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 8.563 | 0.995 | 0.000 | 10 | 8.713 | 0.979 | 0.000 | 10 |
| 0.05 | 8.555 | 0.966 | 0.000 | 10 | 8.569 | 0.932 | 0.000 | 10 |
| 0.10 | 8.769 | 0.892 | 0.000 | 10 | 8.603 | 0.873 | 0.000 | 10 |
| 0.15 | 8.278 | 0.846 | 0.000 | 10 | 7.940 | 0.759 | 0.001 | 10 |
| 0.20 | 8.448 | 0.870 | 0.000 | 10 | 8.065 | 0.830 | 0.000 | 10 |
| 0.25 | 7.757 | 0.687 | 0.004 | 10 | 7.054 | 0.612 | 0.011 | 10 |
| 0.30 | 7.633 | 0.735 | 0.002 | 10 | 7.204 | 0.654 | 0.005 | 10 |
| 0.35 | 5.768 | 0.675 | 0.004 | 9 | 5.515 | 0.578 | 0.011 | 7 |

Table 6.4(a-b) Median values for the regressions of range-size on mean $r$ after 10 simulations for (a) 30 species, noise level 5 added (b) 10 species, noise level 5 added

*Scenario 4 (continued)*

c)

| n=30 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_d$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 9.454 | 0.957 | 0.000 | 10 | 9.483 | 0.883 | 0.000 | 10 |
| 0.05 | 9.338 | 0.937 | 0.000 | 10 | 10.430 | 0.868 | 0.000 | 10 |
| 0.10 | 9.370 | 0.869 | 0.000 | 10 | 9.565 | 0.756 | 0.000 | 10 |
| 0.15 | 9.755 | 0.789 | 0.000 | 10 | 9.719 | 0.659 | 0.000 | 10 |
| 0.20 | 9.145 | 0.771 | 0.000 | 10 | 9.090 | 0.625 | 0.000 | 10 |
| 0.25 | 9.367 | 0.697 | 0.000 | 10 | 9.572 | 0.528 | 0.000 | 10 |
| 0.30 | 9.004 | 0.637 | 0.000 | 10 | 8.642 | 0.462 | 0.000 | 10 |
| 0.35 | 9.366 | 0.651 | 0.000 | 10 | 9.306 | 0.518 | 0.000 | 10 |

d)

| n=10 | Area of Occupancy | | | | Extent of occurrence | | | |
|---|---|---|---|---|---|---|---|---|
| $\sigma_d$ | regression coefficient | r-square | probability | number significant | regression coefficient | r-square | probability | number significant |
| 0 | 9.311 | 0.956 | 0.000 | 10 | 9.969 | 0.854 | 0.000 | 10 |
| 0.05 | 9.270 | 0.930 | 0.000 | 10 | 9.452 | 0.830 | 0.000 | 10 |
| 0.10 | 10.169 | 0.873 | 0.000 | 10 | 10.776 | 0.827 | 0.000 | 10 |
| 0.15 | 9.573 | 0.842 | 0.000 | 10 | 10.338 | 0.660 | 0.005 | 8 |
| 0.20 | 8.741 | 0.665 | 0.004 | 10 | 8.361 | 0.516 | 0.021 | 7 |
| 0.25 | 9.513 | 0.726 | 0.002 | 9 | 9.247 | 0.610 | 0.008 | 8 |
| 0.30 | 8.911 | 0.693 | 0.003 | 9 | 8.356 | 0.519 | 0.022 | 7 |
| 0.35 | 9.075 | 0.572 | 0.012 | 9 | 7.963 | 0.405 | 0.056 | 5 |

Table 6.4(c-d) Median values for the regressions of range-size on mean *r* after 10 simulations for (c) 30 species, noise level 15 added (d) 10 species, noise level 15 added

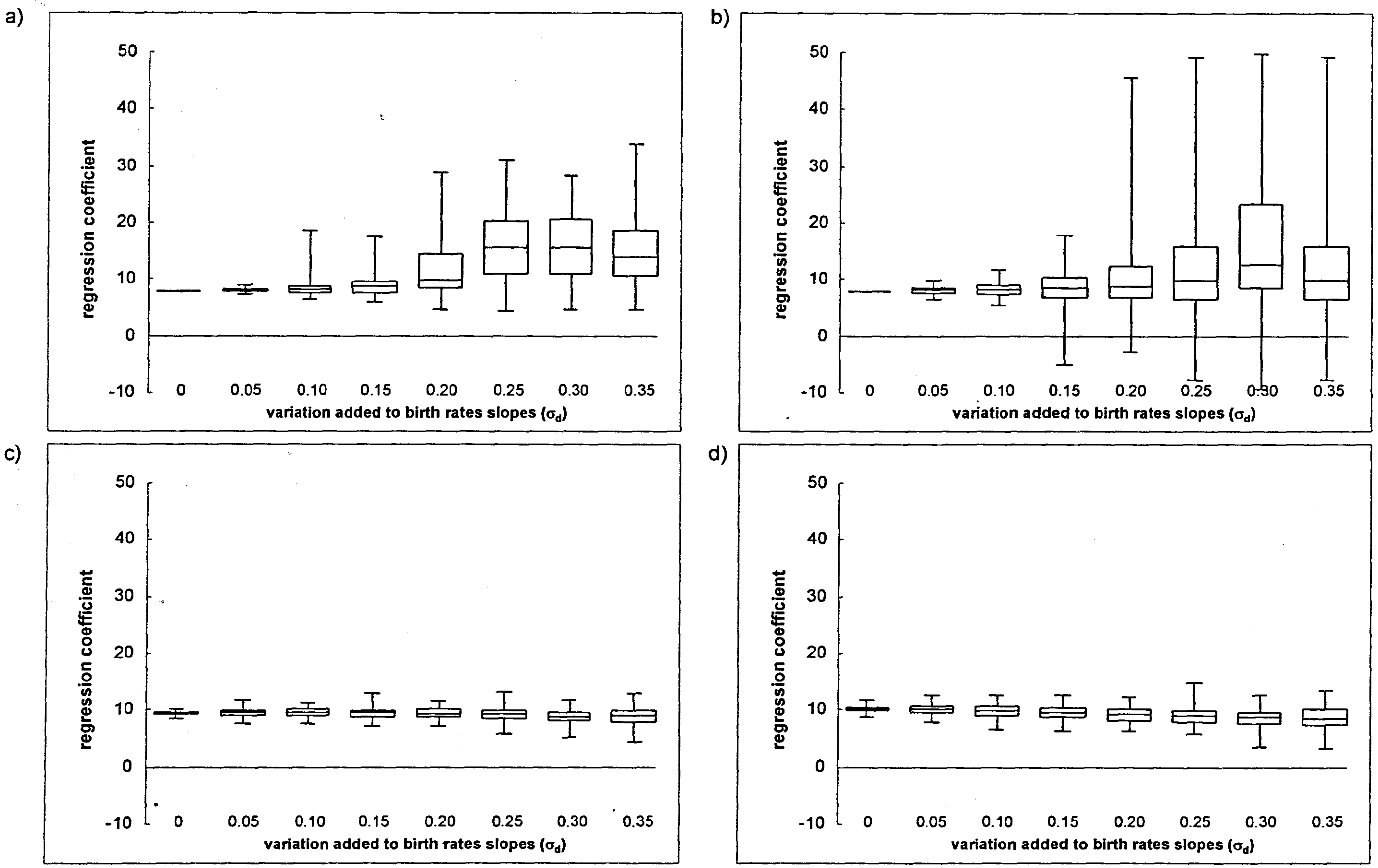


Figure 6.6 Change in the regression coefficient of the abundance-distribution relationship as more variation ($\sigma_d$) is applied to the death rate slopes. (a) Area of occupancy, 30 species, no noise added (b) Area of occupancy, 10 species, no noise added. (c) Area of occupancy, 30 species, noise = 15 added. (d) Extent of occurrence, 30 species, noise = 15 added. In all cases, 100 simulations were run for each $s_d$ shown and the birth rates were held constant.

*Scenario 5: Variable birth and death rate slopes (figure 6.4e)*

Since there are a large number of combinations of parameters for both scenarios 5 and 6, the results tables have been simplified to reflect the number of regressions that were significantly positive, at the $\alpha=0.05$ level, after each set of 10 simulations.

For 30 species, the majority of regressions are significantly positive at all levels of $\sigma_b$ and $\sigma_d$ (table 6.5a). At lower levels of vital rates slope variation, all of the regressions are significantly positive but at the higher levels of $\sigma_d$ and low levels of $\sigma_b$ this drops slightly with typically one or two regressions in ten failing to meet the 5% significance level. For 10 species, the same trend is apparent but rather more regressions fail to meet significance (table 6.5b). The lowest number of significant regressions occur at high $\sigma_d$ and low $\sigma_b$ and for the intermediate values of both.

*Scenario 6: Variable birth and death rate slopes with noise added (figure 6.4f)*

When noise is added to the vital rates curves, area of occupancy and extent of occurrence must be considered separately. Addition of noise at both levels has little effect on the number of significant regressions when considering area of occupancy for 30 species (tables 6.6a and 6.6e), however those that fail to meet significance tend to be at high levels of $\sigma_b$. Extent of occurrence, however, is rather more sensitive to the addition of noise (tables 6.6b and 6.6f). At a noise level of 15 (table 6.6f), few regressions are significantly positive when $\sigma_b$ is high. These patterns are highlighted when considering just 10 species (tables 6.6c, 6.6d, 6.6g and 6.6h) with fewer regressions meeting the significance criteria as would be expected.

*Summary of scenarios 5 and 6*

Selected combinations of parameters were selected from scenarios 5 and 6 and the simulation was run with 100 replicates in order to confirm the patterns seen in the results outlined above (table 6.7). With no noise added to the vital rates curves, the trend that was more obvious for 10 species is again evident (table 6.7a). At low levels of both $\sigma_b$ and $\sigma_d$ all the regressions are significantly positive. The lowest numbers of significant regressions, however, occur at intermediate values of variation in the vital rates curves and not at the highest levels, as might be expected. With the addition of noise, this pattern changes slightly (tables 6.7a-6.7e). The lowest numbers of significant

*Scenario 5: Variable birth and death rate slopes*

a) Area of occupancy and extent of occurrence

| n=30 | | $\sigma_b$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 10 | 9 | 8 | 10 | 10 |
| | 0.05 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| | 0.10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| $\sigma_d$ | 0.15 | 10 | 10 | 9 | 9 | 8 | 10 | 10 | 10 |
| | 0.20 | 10 | 10 | 9 | 8 | 8 | 8 | 10 | 10 |
| | 0.25 | 10 | 8 | 9 | 10 | 10 | 10 | 10 | 9 |
| | 0.30 | 9 | 8 | 9 | 9 | 9 | 10 | 9 | 10 |
| | 0.35 | 9 | 8 | 8 | 7 | 10 | 10 | 9 | 10 |

b) Area of occupancy and extent of occurrence

| n=10 | | $\sigma_b$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 9 | 9 | 8 | 4 | 9 |
| | 0.05 | 10 | 10 | 10 | 9 | 4 | 7 | 9 | 5 |
| | 0.10 | 10 | 10 | 8 | 6 | 10 | 5 | 8 | 4 |
| $\sigma_d$ | 0.15 | 8 | 8 | 7 | 6 | 7 | 3 | 7 | 8 |
| | 0.20 | 8 | 8 | 8 | 4 | 5 | 7 | 7 | 9 |
| | 0.25 | 8 | 2 | 4 | 6 | 7 | 7 | 5 | 7 |
| | 0.30 | 3 | 4 | 7 | 4 | 7 | 7 | 8 | 9 |
| | 0.35 | 3 | 4 | 7 | 7 | 4 | 9 | 9 | 8 |

Table 6.5 The number of significant positive regressions of range-size on mean $r$ at the $\alpha = 0.05$ level after 10 simulations for (a) 30 species (b) 10 species.

*Scenario 6: Variable birth and death rate slopes with noise added (figure 4f)*

a) Area of occupancy

| n=30 | | $\sigma_b$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 9 | 10 |
| | 0.05 | 10 | 10 | 10 | 10 | 10 | 9 | 8 | 7 |
| | 0.10 | 10 | 10 | 10 | 10 | 10 | 10 | 8 | 9 |
| $\sigma_d$ | 0.15 | 10 | 10 | 10 | 10 | 10 | 9 | 7 | 9 |
| | 0.20 | 10 | 10 | 10 | 10 | 10 | 9 | 10 | 9 |
| | 0.25 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 9 |
| | 0.30 | 10 | 10 | 10 | 10 | 10 | 9 | 10 | 10 |
| | 0.35 | 10 | 10 | 10 | 10 | 9 | 8 | 10 | 8 |

b) Extent of occurrence

| n=30 | | $\sigma_b$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 10 | 9 | 8 | 5 | 6 |
| | 0.05 | 10 | 10 | 10 | 10 | 10 | 8 | 8 | 5 |
| | 0.10 | 10 | 10 | 10 | 9 | 10 | 9 | 7 | 6 |
| $\sigma_d$ | 0.15 | 10 | 10 | 9 | 9 | 10 | 8 | 6 | 8 |
| | 0.20 | 10 | 10 | 10 | 10 | 10 | 8 | 9 | 8 |
| | 0.25 | 10 | 10 | 10 | 10 | 9 | 6 | 9 | 8 |
| | 0.30 | 10 | 10 | 10 | 10 | 9 | 8 | 8 | 8 |
| | 0.35 | 10 | 10 | 10 | 9 | 9 | 6 | 9 | 8 |

c) Area of occupancy

| n=10 | | $\sigma_b$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 9 | 8 | 4 | 5 | 1 |
| | 0.05 | 10 | 10 | 10 | 9 | 8 | 8 | 6 | 2 |
| | 0.10 | 10 | 10 | 10 | 9 | 9 | 6 | 5 | 3 |
| $\sigma_d$ | 0.15 | 10 | 10 | 10 | 7 | 7 | 6 | 7 | 6 |
| | 0.20 | 10 | 10 | 9 | 9 | 7 | 5 | 5 | 5 |
| | 0.25 | 10 | 9 | 10 | 7 | 10 | 3 | 8 | 6 |
| | 0.30 | 10 | 9 | 10 | 8 | 6 | 7 | 6 | 5 |
| | 0.35 | 9 | 10 | 8 | 8 | 7 | 3 | 6 | 4 |

d) Extent of occurrence

| n=10 | | $\sigma_b$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 8 | 6 | 3 | 2 | 0 |
| | 0.05 | 10 | 10 | 10 | 9 | 6 | 6 | 6 | 1 |
| | 0.10 | 10 | 10 | 10 | 8 | 5 | 6 | 3 | 2 |
| $\sigma_d$ | 0.15 | 10 | 9 | 10 | 6 | 6 | 4 | 4 | 2 |
| | 0.20 | 10 | 10 | 7 | 6 | 3 | 3 | 4 | 5 |
| | 0.25 | 10 | 9 | 10 | 6 | 9 | 3 | 7 | 5 |
| | 0.30 | 9 | 7 | 8 | 6 | 4 | 5 | 6 | 4 |
| | 0.35 | 6 | 8 | 7 | 7 | 5 | 0 | 6 | 2 |

Table 6.6(a-d) The number of significant positive regressions of range-size on mean $r$ at the $\alpha = 0.05$ level after 10 simulations for (a) Area of occupancy, 30 species, noise level 5 added (b) Extent of occurrence, 30 species, noise level 5 added (c) Area of occupancy, 10 species, noise level 5 added (d) Extent of occurrence, 10 species, noise level 5 added.

*Scenario 6 (continued)*

e) Area of occupancy

| n=30 | | | | | $\sigma_b$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 9 |
| | 0.05 | 10 | 10 | 10 | 10 | 9 | 9 | 8 | 10 |
| | 0.10 | 10 | 10 | 10 | 10 | 10 | 10 | 7 | 8 |
| $\sigma_d$ | 0.15 | 10 | 10 | 10 | 10 | 10 | 8 | 7 | 9 |
| | 0.20 | 10 | 10 | 10 | 10 | 10 | 8 | 10 | 8 |
| | 0.25 | 10 | 10 | 10 | 10 | 10 | 10 | 9 | 7 |
| | 0.30 | 10 | 10 | 10 | 10 | 9 | 9 | 9 | 8 |
| | 0.35 | 10 | 10 | 10 | 10 | 10 | 9 | 7 | 8 |

f) Extent of occurrence

| n=30 | | | | | $\sigma_b$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 9 | 4 | 5 | 1 | 4 |
| | 0.05 | 10 | 10 | 9 | 9 | 6 | 4 | 2 | 5 |
| | 0.10 | 10 | 10 | 10 | 9 | 8 | 8 | 3 | 5 |
| $\sigma_d$ | 0.15 | 10 | 10 | 9 | 8 | 8 | 5 | 3 | 3 |
| | 0.20 | 10 | 10 | 10 | 9 | 6 | 6 | 5 | 3 |
| | 0.25 | 10 | 9 | 10 | 5 | 9 | 5 | 5 | 7 |
| | 0.30 | 10 | 10 | 9 | 8 | 8 | 7 | 5 | 5 |
| | 0.35 | 10 | 10 | 10 | 8 | 6 | 7 | 4 | 5 |

g) Area of occupancy

| n=10 | | | | | $\sigma_b$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 10 | 7 | 8 | 5 | 4 | 1 |
| | 0.05 | 10 | 10 | 10 | 7 | 9 | 3 | 3 | 1 |
| | 0.10 | 10 | 10 | 9 | 8 | 6 | 4 | 4 | 5 |
| $\sigma_d$ | 0.15 | 10 | 10 | 9 | 7 | 6 | 5 | 5 | 4 |
| | 0.20 | 10 | 10 | 8 | 8 | 8 | 3 | 6 | 3 |
| | 0.25 | 9 | 8 | 9 | 6 | 6 | 4 | 5 | 2 |
| | 0.30 | 9 | 9 | 10 | 8 | 4 | 3 | 4 | 3 |
| | 0.35 | 9 | 6 | 9 | 8 | 8 | 5 | 2 | 4 |

h) Extent of occurrence

| n=10 | | | | | $\sigma_b$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 |
| | 0 | 10 | 10 | 8 | 3 | 2 | 3 | 2 | 0 |
| | 0.05 | 10 | 10 | 9 | 2 | 4 | 2 | 1 | 0 |
| | 0.10 | 10 | 10 | 5 | 4 | 3 | 0 | 2 | 1 |
| $\sigma_d$ | 0.15 | 8 | 9 | 6 | 2 | 3 | 2 | 2 | 1 |
| | 0.20 | 7 | 9 | 6 | 2 | 3 | 0 | 4 | 2 |
| | 0.25 | 8 | 5 | 5 | 2 | 2 | 0 | 1 | 1 |
| | 0.30 | 7 | 6 | 5 | 6 | 2 | 1 | 1 | 2 |
| | 0.35 | 5 | 5 | 2 | 4 | 3 | 2 | 0 | 2 |

Table 6.6(e-h) The number of significant positive regressions of range-size on mean $r$ at the $\alpha = 0.05$ level after 10 simulations for (e) Area of occupancy, 30 species, noise level 15 added (f) Extent of occurrence, 30 species, noise level 15 added (g) Area of occupancy, 10 species, noise level 15 added (h) Extent of occurrence, 10 species, noise level 15 added.

*Scenario 6: further simulations with 100 replicates for selected parameters*

a) Area of occupancy and extent of occurrence

| | n=30 | | | $\sigma_b$ | |
|---|---|---|---|---|---|
| | noise = 0 | 0 | 0.10 | 0.20 | 0.30 |
| | 0 | 100 | 100 | 100 | 96 |
| $\sigma_d$ | 0.10 | 100 | 98 | 87 | 82 |
| | 0.20 | 95 | 95 | 88 | 94 |
| | 0.30 | 96 | 95 | 96 | 97 |

b) Area of occupancy

| | n=30 | | | $\sigma_b$ | |
|---|---|---|---|---|---|
| | noise = 5 | 0 | 0.10 | 0.20 | 0.30 |
| | 0 | 100 | 100 | 100 | 94 |
| $\sigma_d$ | 0.10 | 100 | 100 | 97 | 89 |
| | 0.20 | 100 | 100 | 99 | 95 |
| | 0.30 | 100 | 100 | 96 | 95 |

c) Extent of occurrence

| | n=30 | | | $\sigma_b$ | |
|---|---|---|---|---|---|
| | noise = 5 | 0 | 0.10 | 0.20 | 0.30 |
| | 0 | 100 | 100 | 93 | 67 |
| $\sigma_d$ | 0.10 | 100 | 100 | 91 | 68 |
| | 0.20 | 100 | 100 | 96 | 85 |
| | 0.30 | 100 | 100 | 90 | 83 |

d) Area of occupancy

| | n=30 | | | $\sigma_b$ | |
|---|---|---|---|---|---|
| | noise = 15 | 0 | 0.10 | 0.20 | 0.30 |
| | 0 | 100 | 100 | 99 | 87 |
| $\sigma_d$ | 0.10 | 100 | 100 | 98 | 86 |
| | 0.20 | 100 | 100 | 97 | 92 |
| | 0.30 | 100 | 100 | 94 | 83 |

e) Extent of occurrence

| | n=30 | | | $\sigma_b$ | |
|---|---|---|---|---|---|
| | noise = 15 | 0 | 0.10 | 0.20 | 0.30 |
| | 0 | 100 | 99 | 56 | 29 |
| $\sigma_d$ | 0.10 | 100 | 100 | 70 | 41 |
| | 0.20 | 100 | 98 | 69 | 49 |
| | 0.30 | 99 | 99 | 60 | 46 |

Table 6.7 The number of significant positive regressions of range-size on mean *r* at the $\alpha = 0.05$ level after 100 simulations for (a) Area of occupancy and extent of occurrence, 30 species, noise level 0 (b) Area of occupancy, 30 species, noise level 5 added (c) Extent of occurrence, 30 species, noise level 5 added (d) Area of occupancy, 30 species, noise level 15 added (e) Extent of occurrence, 30 species, noise level 15 added.

regressions now occur consistently at high levels of $\sigma_b$ with extent of occurrence showing more sensitivity to this change than area of occupancy.

## Discussion

### *Results of the simulations*

If the results are considered together, a general remark that could be made about the simulations is that as more variability and noise are added to vital rates slopes, the final regressions of range-size upon mean *r* explain less variation with a concomitant decrease in significance. This is hardly a surprising conclusion but it ignores some interesting properties of this implementation of the vital rates model. The main patterns, and some attempt at their explanation, are listed below.

1. Within the bounds of the simulation, the relationship between mean *r* and range size has proved to be rather robust, despite the fact that birth rates may vary by a factor of ten, the death rates slopes can be both positive or negative and a large amount of random noise can be added to the site-specific vital rates. This result is unexpected because there has been no explicit relationship built into the model between the levels of the vital rates of each species and its range size. One might expect any variability in the slopes of the vital rates curves to quickly reduce the resulting abundance-distribution relationship to a random scatter of points. In contrast, the results of the simulation imply that species within an assemblage do not have to show similar responses to an environmental gradient for the positive abundance-distribution relationship to hold.

Considering the simplest of the scenarios, where only birth rates may vary (figure 6.4a), at first glance the vital rates model might be reduced to a comparison between the base (range size) and height (maximum *r*) of random isosceles triangles (which of course must be independent). This would result in points lying randomly within the dashed rectangle on the resulting range size versus mean *r* plot schematically represented in figure 6.7. However, since there are constraints on the slopes of the birth rate curves in the computer simulation, for the reasons outlined above, all species must lie within certain limits in this plot. This can perhaps be best understood by considering the four species illustrated here labelled A-D. Species A has the minimum birth rate slope and maximum birth rate level allowable in the simulation, i.e. it has the highest possible range size and mean *r*. All other species that can be included in the simulation must lie to the left and below this species in a resulting mean *r* versus range size plot. Species D has the maximum slope and minimum level for birth rates, resulting in the

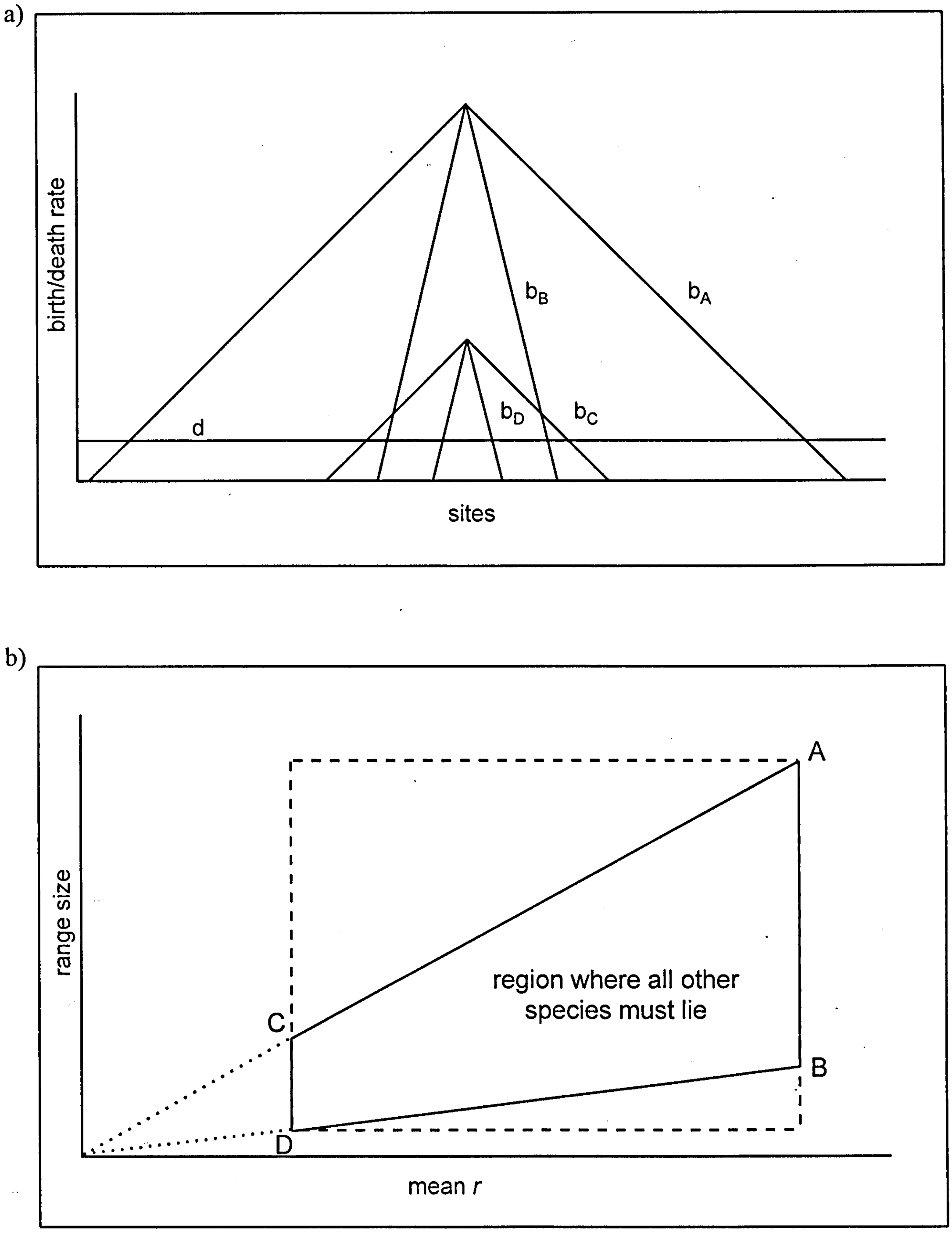


Figure 6.7 Illustration of the effect of constraining birth rates slopes in the vital rates model simulation on the resulting mean *r* versus range size plot (a) Fixed death rates for all species. Species A has the minimum birth rates slope and maximum birth rates level. Species B has maximum slope and birth rates level. Species C has minimum slope and birth rates level. Species D has maximum slope and minimum birth rates level. (b) The rectangular region enclosed by the dashed line represents the region that species might occupy if there were no constraints on the birth rates slopes. However, all species in the simulation are constrained to lie within the region set by species A-D (see main text for a full explanation).

lowest mean $r$ and range size possible in the simulation. Similarly, B and C represent species that also set the limits in the resulting mean $r$ versus range size relationship. Any other species must lie somewhere within the polygon shown. As the constraint upon the minimum slope increases for all species then point C is pulled down towards the ordinate axis. Similarly, if the constraint upon the maximum slope increases then point B is pushed up away from the ordinate axis. The vital rates model, as implemented in the simulation, reduces to a distribution of points within these limits. Intuitively, given a high enough number of degrees of freedom, such a 'relationship' will always be significantly positive. This intuition was tested using a simple program to generate points within similar polygonal limits. It is surprising how small the constraints can be (provided there are enough points generated) yet still produce a significant positive relationship.

A negative relationship and a non-significant relationship can, of course, occur by chance within this polygon and this did occur occasionally in the simulations involving only ten species. A negative relationship can also occur if there is a systematic positive relationship between the slope of the basal birth rates response along the gradient and the maximum value that it can take (e.g. figure 6.1d). This scenario was not investigated using the computer simulation since the purpose of the program was to add random variation onto the slopes, independent of the maximum birth rates. To introduce such a relationship between the maximum value of the vital rates and the slope makes the argument behind the model itself circular.

In the more complicated scenarios, for example when death rates slopes also vary, it is less obvious why the positive relationship should hold despite the high degree of variation in the vital rates curves. However, it seems reasonable that the constraints imposed upon the parameters in the simulation create positively skewed spaces in the resulting mean $r$ versus range size plot in a similar manner to that outlined above. This hypothesis could of course be tested with a more detailed analysis of the model.

Note that, although the simulation yields abundance-distribution data for species that will result in a significant fit when using linear regression, what it is really generating is a distribution of points within a bounded space. Linear regression is not the most appropriate model for this type of data since the simulation output does not conform to its underlying assumptions. However, the technique is used here to illustrate that a significant fit can be obtained when using regression on such distributions of points which could then be interpreted as a linear relationship between species' abundances and distributions. Alternative approaches for dealing with ecological data bounded in these ways are discussed by Brown (1984), Blackburn *et al.* (1992) and Thomson *et al.* (1996).

2. Another property of the simulation is that the regression coefficient often increases as more noise and variation is applied to the vital rates slopes. This suggests that range size increases more quickly with mean *r* as this variability is applied. Considering scenario 1 again, it should be noticed that, even though there is an increase in the overall slope of the regression line with higher levels of $\sigma_b$, less variation in the data is explained by the fit. This is attributable to the fact that as more variation is applied to the vital rates slopes, points on the range abundance plot may occupy more of the region strictly bounded by the constraint lines. This region, however, is asymmetric and the space in which the points are most likely to lie changes shape asymmetrically as $\sigma_b$ increases, with more points falling in the upper region of the constraint space than the lower one. The effect is illustrated in Figure 6.8. Here three simulations are run with 500 species each. Note how, as $\sigma_b$ increases, the upper region where points may lie increases in size more quickly than the lower one, resulting in an increasing regression coefficient but a decrease in the amount of variation explained. This effect is partly a result of the values chosen for the mean slopes of the vital rates and the constraints imposed upon them in this particular study, rather than a necessary property of the vital rates model itself, but is an interesting property of the model under such constraints nonetheless.

3. The addition of noise often had unexpected effects on the resulting regressions. It was initially expected that the addition of random noise to the vital rates curves would simply serve to weaken any relationship between mean *r* and range size. However this was dependent upon the scenario simulated. In scenario 2, where $\sigma_d$ was held constant, this prediction was upheld (e.g. figure 6.5c). Adding noise to the vital rates curves increased the variability in the resulting regression coefficients from the simulations. In contrast, in scenario 4, when $\sigma_b$ was held constant, addition of a noise level of 5 had very little effect upon the resulting regressions and a noise level of 15 served to stabilise the values of the regression coefficients.

There appear to be two main effects operating here. First, addition of the low noise level might be expected to have little influence on the resulting regression since at each site the vital rates change only slightly relative to their original values. If one adopts the operational definition of the environmental gradient given by Holt *et al.* (1997), where sites are ordered to produce the vital rates curve independent of any externally measured

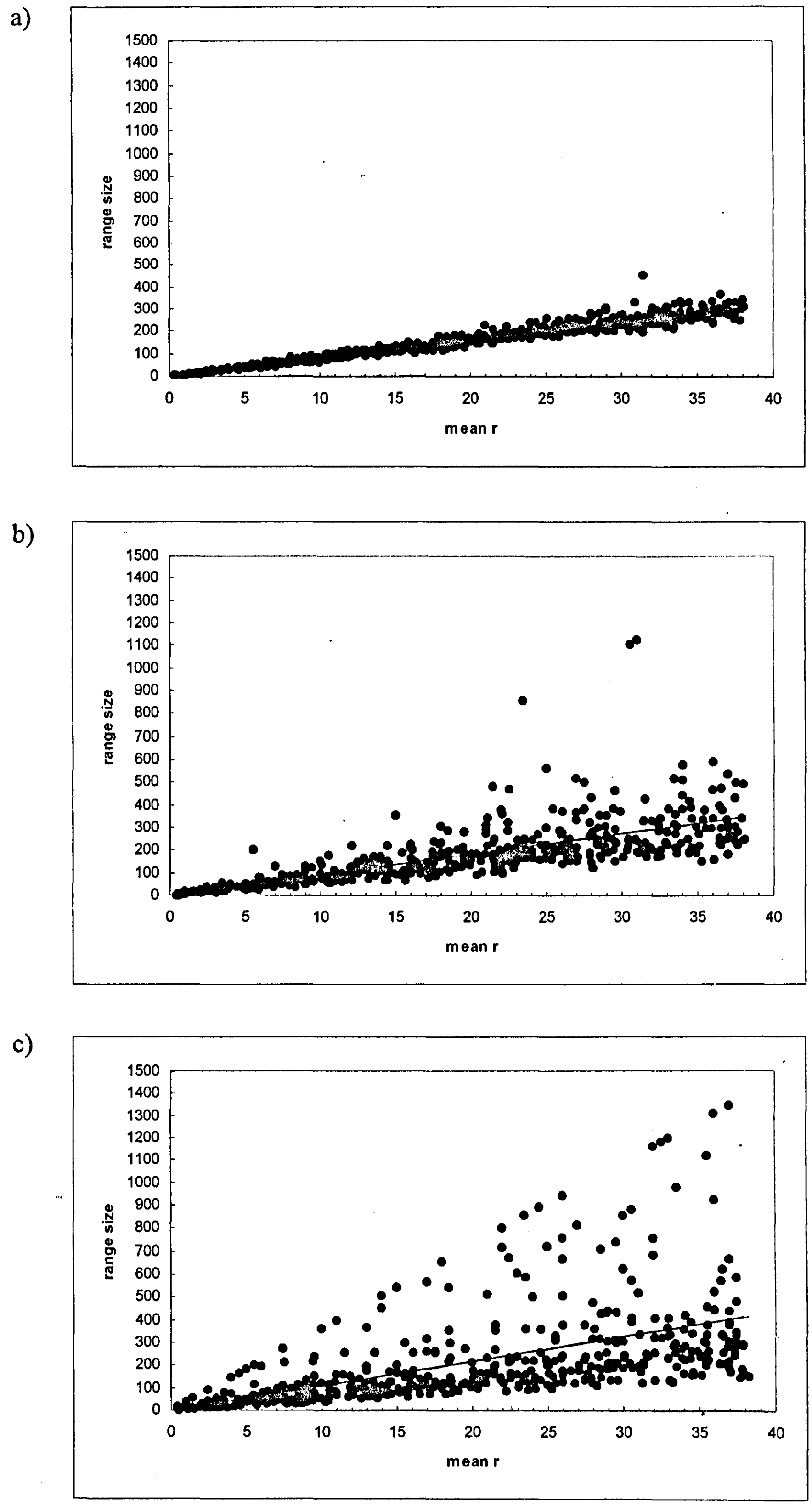


Figure 6.8. The effect of increasing the variation in birth rate slopes on the regression coefficient and fit of the range size versus mean *r* relationship for 500 species:
(a) $\sigma_b = 0.05$, regression coefficient = 8.11, $r^2 = 0.95$
(b) $\sigma_b = 0.15$, regression coefficient = 9.16, $r^2 = 0.54$
(c) $\sigma_b = 0.25$, regression coefficient = 10.96, $r^2 = 0.32$.

set of environmental variables, then the sites could simply be re-ordered along the hypothesised gradient (however see below). This would result in a different set of vital rates curves but would still produce the positive relationship.

At higher levels of noise it appears that, at least when $\sigma_d$ is variable (scenario 4), influence from the variation in the shapes of the death rates curves is almost completely lost amidst the noise. This results in a relatively uniform distribution of regression coefficients at all levels of $\sigma_d$ (figure 6.6c and d). When $\sigma_b$ is variable (scenario 2), the addition of noise has less of a dampening effect. This is to some extent confirmed when considering scenario 6 where both vital rates may vary. In the presence of noise, higher values of $\sigma_b$ have more influence on the number of significantly positive regression lines (tables 6.7 b-e).

It is possible that these effects arise as a result of constraining birth rate slopes to always be positive before their maxima, whereas death rate slopes are not similarly constrained, and that death rates are necessarily lower than birth rates at all of the sites included in the final regressions. One could test these suggestions by modifying the simulation program to calculate the regressions on the same set of species before and after the addition of noise. The present simulation, however, generated species randomly at the beginning of each run.

4. When both $\sigma_b$ and $\sigma_d$ are allowed to vary in the absence of noise (scenario 5), simply increasing variation does not always lead to easily predicted changes in the nature of the final regressions. There are obviously interactions between the parameters occurring here that have resulted in unexpected trends in both the regression coefficients and number of significant fits. The reasons for these results are not immediately clear yet one should bear in mind that if such results were obvious then a simulation approach would have been unnecessary at the outset.

*Comparison between the simulation, the theoretical vital rates model and the real world*

The constraints that were introduced into the simulation were necessary to prevent numerical errors and to allow the simulation to run using limited resources. When considering the results of the simulation it is important to ask whether these constraints are realistic both within the context of the vital rates model as it was originally formulated and in the real world.

Holt *et al.* (1997) point out that the positive abundance-range size relationship can easily be disrupted if there is not an equal representation of each species' birth rate along the environmental gradient. It is implicit in their argument that for the model to work at all it must have the complete response curves of the vital rates curves present. This automatically constrains the slopes of the curves to some minimum (determined by the spatial extent of the survey). The simulation does not differ from the vital rates model in this respect. Therefore, for both the theoretical and simulation models there will be a maximum range that any species may achieve with a given mean *r*. However, it is not obvious that such a constraint should apply in the real world. Moreover, many studies of the abundance-distribution relationship are partial studies covering only a portion of several species ranges. If the environmental variables important in affecting the vital rates of the assemblage exhibit strong spatial autocorrelation then it is possible that a proportion of the vital rates curves will not be fully realised within the extent of the survey. One would expect this to result in a levelling off of the abundance-distribution relationship (representing species with 100% occupancy) at larger values of mean abundance as the full extent of the survey is reached. This does, of course, occur, though not necessarily as a result of unrealised complete birth rates curves! Spatial extent, then, may be important in the strength of the abundance-distribution relationship. Considering the vital rates model, one would expect that relationship to be strongest at larger scales since species vital rates curves would be fully realised. However, as Gaston (1994) points out, increasing the spatial scale also introduces more geographic and hence taxonomic variability into the assemblages under study. This factor at the largest geographic scales, may serve to weaken the abundance-distribution relationship.

A constraint on the maximum slopes of the birth rate curves was necessary in the simulation to prevent numerical error but is neither explicitly nor implicitly suggested in the original model. Relaxing this constraint would result in only the line AC acting as a constraint in figure 6.7. Remarkably, given enough degrees of freedom this would still result in a positive relationship as was confirmed using a simple program to plot points randomly within such a constraint space.

Holt *et al.* (1997) use a triangular response in birth rates along the environmental gradient for the sake of simplicity but note that many rounded, unimodal responses would also yield the positive abundance-distribution relationship. This is also true for the simulation model as long as constraints, similar to those discussed above, apply to these responses. In addition, the responses need not be symmetrical, which one may imagine to be more likely for real species, as long as the constraints still apply.

To summarise, the constraints on the birth rates slopes necessary for the implementation of the vital rates model as a computer simulation do not significantly affect the main conclusion that can be drawn from its results. Even with the constraints relaxed to a considerable extent, the positive relationship will still arise regularly given enough species in the study. In the 82 studies reported by Gaston (1996) on the interspecific abundance-distribution relationship (of which the overwhelming majority reported a significant positive correlation) the mean number of species involved in each study was 71 (median 55). This is considerably more than the maximum of 30 species simulated in this study and so theoretically could be subject to less constraints on their vital rates slopes.

The other results obtained from this simulation study, for example the increasing regression coefficients with $\sigma_b$ and the unpredictable effect of adding noise to the vital rates curves, may be regarded as dependent upon the precise implementation of the vital rates model. Only a small proportion of the possible parameter space is explored in this study yet a variety of interactions and unpredictable effects have been revealed. Further simulation work would help to test the generality of these results but perhaps a more fruitful approach would be to collect ecological data that would help to elucidate the precise nature of any response in the vital rates of local populations to environmental variation.

### *Empirical testing of the model*

In order to test the underlying assumptions of the vital rates model it is necessary to measure birth and death rates within an assemblage along environmental gradients. It would also be necessary to measure spatial variation in density-dependence across a range of sites for the assemblage. There is currently no such data available and as Holt *et al.* (1997) point out these are particularly demanding tasks. The simulation approach has gone some way in revealing several implicit assumptions of the theoretical model and also highlights some difficulties that may arise when undertaking such empirical studies.

Holt *et al.* (1997) state that the environmental gradient in the vital rates model need not indicate spatial structure in birth and death rates. Their environmental gradient could represent an abstract ordering of sites that in reality are arranged in a far more complex spatial mosaic of different qualities (Lawton, 1996). While this is true, they take the argument further by implementing an operational definition of a gradient obtained by ranking sites according to measured vital rates. However, this argument poses some risk of circularity. Consider a transect taken along an unknown but real environmental gradient (e.g. altitude, pH) in a region which yields the measured birth rates according to figure 6.9a. These sites could readily be rearranged to resemble the response curves of the vital rates model (figure 6.9b), yet the environmental gradient would now be meaningless. In fact a bounded, uniform distribution of random values for the vital rates, independent of any environmental gradient, could be rearranged into the desired shape. In the simulation, the limited effect of the addition of random noise at each site serves to highlight this problem. Perhaps if one has to use the operational definition it would be better to label the ordinate axis 'sites' rather than inferring an environmental gradient. Maurer and Brown (1989), who assume spatial autocorrelation in environmental variation in their model of range structure, side step this difficulty by treating the ordinate axis as space. Adjacent sites on this axis are adjacent in reality as well as in the model. Ideally then, the environmental gradient should represent a real, externally measured variable to which the local population dynamics of a species genuinely respond.

At larger scales, responses are likely to be affected by a suite of environmental variables, some of which will be poorly correlated. The definition of the environmental gradient again becomes problematic but could perhaps be represented by the first principal component of environmental variation or increasing the dimensionality of the environmental component in the model.

In the simulation model, range-size has been measured in two different ways; area of occupancy and extent of occurrence (Gaston, 1991). It has been shown that extent of occurrence is considerably more sensitive to variation in the model's parameters than area of occupancy. Since the simulation model makes no necessary assumptions about the spatial structure of the environmental gradient, extent of occurrence as defined here may not be strictly analogous to the way it is used when measuring range size in the real world. Strictly speaking, extent of occurrence in the model is range size measured as the extent of environmental variation in which a species may occur. This would only correlate well with Gaston's (1991) definition if the environment was spatially autocorrelated in the way envisaged by Maurer and Brown (1989).

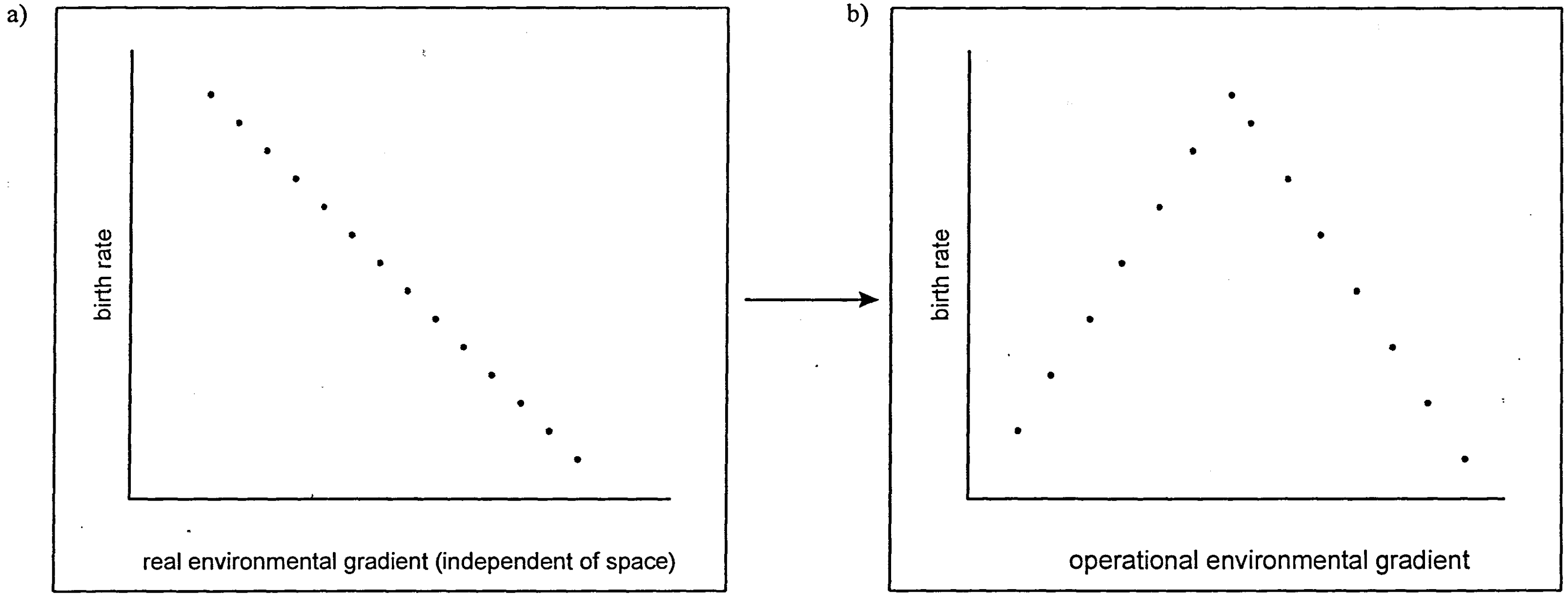


Figure 6.9 The problem inherent in defining an operation environmental gradient for the vital rates model. Many different types of birth rate response along real environmental gradients may be rearranged to resemble a unimodal response along an operational environmental gradient (see main text for further details).

## Conclusion

The results from the simulation reveal that, given the assumptions of the vital rates model, the abundance-distribution relationship may be remarkably robust to random variation in the responses of species to environmental variation and also to noise added to these responses. In addition, it has helped to clarify some of the implicit assumptions behind the theoretical model and highlight some of the problems that may be encountered when trying to test it empirically. The simulation approach can serve as a useful intermediate step between theoretical models and empirical testing.

Gaston (1996) observed that, since there are now a large number of studies that have demonstrated the abundance-distribution relationship empirically, there really is little need to collect more data. A more profitable approach would be research focused on attempting to provide an explanation for such a pattern. Similarly, there are now a variety of hypotheses which attempt to explain the relationship which are not based on data from real assemblages (Gaston *et al.*, 1997). Perhaps the focus of future work should be to collect data that will help to distinguish between these hypotheses. Collecting such comparative data for an assemblage would require co-ordinated research over large geographic areas but this is precisely what is needed if we are ever to understand the patterns that are apparent at these taxonomic and spatial scales.

# Chapter 7

# Discussion

> *"It is important to emphasize at this point that species range boundaries and geographic patterns of diversity ultimately reflect variation in the abiotic environment. Biotic interactions such as competition, predation and mutualism can be proximate limiting factors, but it is the physical template of the earth - its geology, climate and physical oceanography and limnology - that ultimately determines the distribution of life."*
>
> James H. Brown (1995) Macroecology

*Spatial structure in* P. ilicis *local population densities*

It is possible to find the same magnitude of variation in *P. ilicis* densities within certain, environmentally heterogeneous, 1km grid squares of the Sheffield region as across the whole of its geographic range. In the face of such variation, it is perhaps remarkable that any broad scale spatial structure in local population densities can be discerned at all. However, with sufficient sampling, it can be shown that not only do population densities show significant spatial autocorrelation within individual habitat patches but also broad scale trends become apparent. At a regional scale this structure is complex (figure 3.5) and is due, at least in part, to differences in habitat type and altitude. Environmental variation measured at this spatial scale is itself a heterogeneous mosaic across space.

Across the whole of the geographic range, holly leaf-miners do not exhibit the simple pattern in density structure envisaged by Brown (1984). The analyses of holly leaf-miner data reported here (Chapter 4) and the numerous maps created for North American birds, surely indicate that likening range structure statistically to a bivariate normal distribution is not useful and there has never been a serious attempt to assess the goodness of fit of real data to such a distribution. Even in the broadest sense, the internal density structure of the range of the holly leaf-miner is not Gaussian. *P. ilicis* local population density structure is strongly skewed and has at least two distinct peaks (figure 4.1). In addition, the range boundary is very irregular in outline as opposed to a more symmetrical shape that would be expected if it were Gaussian. *P. ilicis* distribution is, of course, determined necessarily by the distribution of holly trees, which itself has an irregular boundary in Western Europe.

However, the internal density structure of the range of *P. ilicis* is far from random and arguably comparatively simple in the light of the variation in population densities apparent at smaller spatial scales, which is present even when habitat type is held constant. A more conservative, and perhaps reasonable, generalisation that could be made about range structure, given our current knowledge, is that species often have a small number of regional centres of consistently high population densities which then obviously decrease with distance away from these centres over portions of the range. There is more structure apparent than this for the holly leaf-miner in Europe, for example the broad gradient of increasing density that runs from the south-west to the north-east. However, it would be rash to attempt to make any further generalisations about range structure as a whole in Europe from this single species.

The results of the partial regression analyses (Chapter 4) demonstrate that a large proportion of the regional spatial structure exhibited by leaf-miner densities is shared with environmental variation. This was also apparent within the regional survey of the Sheffield area. Demonstrating a causal relationship at this spatial scale can be difficult as a direct result of this shared structure (see below). Despite this, the data support the hypothesis that range structure is at least in part determined by broad scale variation in environmental variables. This is consistent with Brown's emphasis on the central importance of abiotic conditions in the creation of spatial patterns apparent in ecological phenomena (see the epigraph to this chapter). However, his assumptions about the nature of the spatial structure of the environment do not hold, at least for Western Europe. Environmental variation is obviously spatially autocorrelated at these scales in the sense that sites close to one another experience similar environmental conditions. However, additional broad scale trends serve to complicate this simple view of environmental variation, including rapid shifts in environmental conditions at coastlines and mountain ranges.

### P. ilicis *demographic rates components*

Many of the components of the demographic rates that were measured for *P. ilicis* also exhibit significant spatial autocorrelation and large scale trends across the range, including some differences between sites at the edge of the range and those closer to the geographic centre. Mortality rates, with the important exceptions of larval parasitism and miscellaneous larval mortality, are distributed largely independently of one another resulting in a complex template of mortality profiles experienced by *P. ilicis* across the range. To summarise, sites close to one another of similar habitat type display similar patterns of mortality. However, leaf-miner populations separated by large distances experience population dynamics that are very different and characteristic of the particular region. Of course, within a region, there may be similar variation in mortalities between local populations in different habitat types as was recorded in the Sheffield area. However, in many parts of its range, holly is restricted to a far narrower range of habitat types than those where it is present in the UK. Controlling for habitat type served to clarify the changes in leaf-miner population dynamics across the range but this simplification should not be mistaken for the complete picture which is also likely to be dependent upon the extent to which particular environmental combinations are represented across space. Unfortunately, there are no other studies with which to compare these results and so the generality of these patterns is yet to be assessed.

At the regional scale, many of the demographic rates components were found to be correlated with environmental variation. However, the responses varied between habitat types, resulting in a complex picture of changes in population dynamics with respect to the environment. An important aim of the European survey was to investigate whether changes in the demographic rates components could be correlated with large scale changes in environmental variables when some attempt was made to control for habitat type. Far less of the variation in mortalities was accounted for by environmental variables than for oviposition and leaf-miner density (which were highly correlated with one another). As a result, after shared spatial structure was taken into account, there was little evidence to support the hypothesis that mortality rates were determined by the environmental variables measured. Given the idiosyncrasies in the responses of mortalities at a regional scale perhaps this shouldn't be too surprising. In addition, the confounding effect of shared spatial structure is likely to make it extremely difficult to attribute variation in leaf-miner demographic rates to environmental variation at these spatial scales (see below).

Miscellaneous larval mortality and larval parasitism are positively correlated at both the regional scale and across the whole range. As was discussed in Chapter 3, this may be due to the two mortalities responding similarly to underlying factors or to miscellaneous larval mortality arising as a result of attacks by larval parasites. However, it seems likely that larval parasitism is no more than a component of miscellaneous larval mortality since miscellaneous larval mortality can occur at sites where larval parasitism is absent. Moreover, in a few areas, a significant proportion of miscellaneous larval mortality occurs when the mine is at too early a stage of development to be attacked by *C. gemma* (Chapter 2). The exclusion experiment outlined in Chapter 3 would serve to clarify this issue.

Few significant correlations between mortality and leaf-miner population densities were found in the regional or European surveys. Larval parasitoids do not show an aggregative response to leaf-miner densities at scales greater than the single habitat patch. As Heads & Lawton (1983) demonstrate, the nature of density dependence between natural enemies and their hosts or prey will often be dependent upon what constitutes a discrete resource patch to the natural enemy in question. This study, which has dealt with a broad range of spatial scales, emphasises that as scale increases beyond the single habitat patch, behavioural aggregative responses of natural enemies become less important as emphasis shifts towards the effects of environmental variation on the presence and absence of the natural enemies in question. The result of this may be that the positive relationship shown between larval parasitism and leaf-miner density within a habitat may switch to a negative relationship across a region. This would be a good indication that larval parasitism is important in the regulation of leaf-miner densities. This pattern was indeed found in urban habitats across Sheffield and Rotherham. It is intriguing to note that in the north-eastern region of the geographic range, where densities are at their highest, larval parasitism is completely absent. This could be considered to indicate regulatory role for *C. gemma* in other areas of the range. However, such a role is not likely to be demonstrated by correlative techniques at this spatial scale since it is likely that different factors are important in limiting leaf-miner densities in different parts of the range. Exclusion experiments, once again, may be the best opportunity for determining the importance of *C. gemma*. Studies employing cages to exclude parasitoids and predators and successful biological control of leaf-miners demonstrate that natural enemies can profoundly affect the population dynamics of leaf-miners. However, Auerbach *et al.* (1985) found that in the field-studies that they reviewed, natural enemies were found to have a regulatory effect on leaf-miner populations in only 36% of the cases, a figure in accord with other similar reviews. It is possible that one reason for this inconsistency is the confounding effect of differences in spatial scale between field studies.

Perhaps most remarkably, the *per capita* rate of successful emergence of adult flies was not a good indicator of local population levels of the most long-lived life stage, the larval one, for most habitat types. While this is not unusual in studies of leaf-miners at smaller spatial scales, it certainly seems counter-intuitive. A number of possible reasons for this are discussed in Chapter 3. It may be that the albeit short adult stage of the life-history has important consequences for local population dynamics. Unfortunately, investigations into the rates of dispersal, fecundity and mortality of adult leaf-miners are likely to meet with the same difficulties that are present with other species which have complete life-histories less suited to the surveying techniques used here.

Perhaps more important is that under the equilibrium model presented by Maurer & Brown (1989) a relationship between *per capita* birth or death rates and differences in density between sites should not be expected once local population densities have reached their equilibrium levels. At equilibrium densities, density-dependent biotic factors would serve to reduce *per capita* birth rates and increase *per capita* death rates. Therefore, comparisons between survival and birth rates between populations in different parts of the range may not give a direct indication of the reasons for differences in density. However, qualitative differences in the components of mortality may certainly provide clues for range structure, for example the potential regulatory effect of larval parasitism in some parts of the range of *P. ilicis*. Experiments would then serve to elucidate these patterns.

### *Macroecological models based on the relationship between demographic rates, population density and the environment.*

Both the vital rates model of Holt *et al.* (1997) and the range structure model of Maurer & Brown (1989), which are based on changes in demographics across the range, would be difficult to test directly by surveys alone because at equilibrium densities birth and death rates would not necessarily reflect population size. Simply collecting data on spatial variation in demographic rates is not sufficient. As Holt *et al.* (1997) admit, collecting data on spatial variation in density-dependence and maximal population growth rates at a large number of sites for a set of species would be a particularly demanding task. However, it is a necessary one if we are to understand broad scale patterns in ecology such as these.

Both the vital rates model and Maurer & Brown's (1989) model of range structure make important assumptions concerning the nature of environmental variation and a species' demographic response to this variation. That is, there is a sufficient range of environmental variation for a unimodal or even Gaussian response in demographic rates and hence densities. This is related to the discussion given in Chapter 1 concerning the selection of appropriate regression models. At a small scale linear regression may be most appropriate whereas at a larger spatial scale, modelling a Gaussian response may be more appropriate (figure 1.2). However, at both the regional and geographic scales, the data from the surveys of *P. ilicis* provided no evidence for a unimodal response in leaf-miner demographic rates and densities to environmental variation. The most likely reason for this is that the assumptions by Maurer & Brown (1989) and Holt *et al.* (1997) about the nature of environmental variation are far too simple. In order to detect any unimodal response, a sufficient range of environmental variation must be represented in the real world. Some combinations of particular environmental variables may simply not exist within the geographic range of a species. Similarly for this response to be translated into a unimodal pattern in space, the spatial structure of the environment would have take on a precise form.

The results of the holly leaf-miner surveys are consistent with the idea that in order to make any generalisations about the structure of species' ranges then ecologists need to be able to make more realistic generalisations about the spatial structure of the environment (see Bell *et al.*, 1993 for a preliminary attempt at this). Whether this is possible remains to be seen but much useful work may already have been done within the geological sciences. A closer integration between these two disciplines may help to increase our understanding of both regional and large scale ecological patterns.

### *P. ilicis* as a model system

The European leaf-miner survey was largely conducted by a team of two people in one field season. As such, it probably represents the limit of what can be learned about the distribution of a species over such a large area by a small team using ground surveying techniques. The task was certainly made a great deal easier by the characteristics of the holly leaf-miner system discussed in Chapter 2 and it is difficult to imagine collecting a similarly comprehensive data set for other groups of species without considerable human and financial resources. It is likely that exemplary systems such as that provided by the holly leaf-miner will be some of the only opportunities ecologists have for similar investigations into range structure. Therefore one should consider carefully whether such systems are likely to be typical of other taxonomic groups before attempting to make generalisations about range structure and demographics in general.

For example, since *P. ilicis* is strictly monophagous, holly trees represent the only potential habitat patches for the insect. For many other species, suitable habitat patches are far less clearly defined. Population densities for *P. ilicis* have been analysed and mapped as functions of leaf-miner density per tree whereas for North American birds they are necessarily mapped as density per unit area. Brown (1984) points out that differences in densities across the range may occur as a result of differences in the proportion of local habitat patches occupied, or of differences in average population density within a patch. An additional possibility is that the density of potential habitat patches may also vary across the range. In the case of *P. ilicis*, this is the density of holly trees. That there is variation in the density of holly trees across the range was strongly evident during the course of the survey. However, it is not immediately obvious how incorporating data on holly tree density across the range would affect the map of leaf-miner distribution expressed as density of leaf-miners per unit area. Producing a reasonably accurate map of holly tree distribution for the whole range may be possible by combining botanical records already available in many countries for limited areas with detailed environmental information in order to extrapolate to unsurveyed areas. The success that Sykes *et al.* (1996) have had in modelling the distribution of many tree species across Europe based on meteorological information is encouraging in this respect.

Another important feature of the holly leaf-miner system that makes it atypical is that the majority of its life-history is spent within another living host with all the implications this has for interactions between the herbivore and holly. In addition, leaf-miners are sedentary and have no opportunity for dispersal from unsuitable hosts until the short adult stage. Many other species have a much greater opportunity to move between habitat patches during the course of their life-histories resulting in a rather more dynamic definition of range structure both at regional and geographic scales. The range of the holly leaf-miner is strictly defined and can be measured at any time of the year whereas this is obviously not true for many migratory species. As a result, the North American bird surveys provide maps for specific periods of the year, for example the North American Breeding Bird Survey and the North American Wintering Bird Survey (Root, 1995; Price *et al.*, 1995). Information throughout the life-history of a species must be gathered before a comprehensive picture of its range structure can be obtained.

Finally, the natural enemies of *P. ilicis* are largely communities of parasitoids a feature restricted to invertebrate groups and most common for phytophagous insects. The implications of these aspects of the natural history of the holly leaf-miner and indeed the idiosyncrasies of any model system for the definition and causes of range structure must always be taken into account before attempting generalisations for other species.

### *Assessment of the geostatistical methods used*

The use of several geostatistical techniques has certainly facilitated the research reported here. Spatial autocorrelation analysis provides a relatively simple method of establishing the existence of any spatial dependence in the leaf-miner data and also helps to identify additional spatial trends present.

### 1. Spatial dependence and hypothesis testing

Spatial dependence, that is the dependence of the value of a variable on its location, comes in two forms. One is as a result of contagious processes. The response variable at a location takes a particular value because it is influenced by the values at other points nearby. The obvious contagious biological process in the context of local population dynamics is migration. A holly tree may support high densities of leaf-miners, not because it is a particularly suitable tree for infestation but because it is adjacent to trees which have a positive net emigration rate. This form of spatial dependence is of course related to the idea of source and sink populations (Pulliam, 1988; Pulliam & Danielson, 1991; Dias, 1996) and also to metapopulation models (Hanski, 1999). As spatial scale increases the importance of these contagious processes is likely to diminish.

Another form of spatial dependence is the dependence of the response variable on its location as a result of underlying spatially structured environmental variables. For example, local population densities may be spatially structured because they reflect the spatial structure of important limiting variables across the study area. However, shared spatial structure between two variables does not necessarily imply causation. Both variables may be responding to other unmeasured underlying processes which are themselves spatially-structured. The problem of spurious correlations arising between variables that share spatial structure has prompted the formulation of a number of statistical techniques which attempt to separate the effect of spatial location from the analysis.

For the leaf-miner data obtained from the European survey, partial regression analysis was used (Legendre, 1993; Legendre & Legendre, 1998). The details of the technique are given in both Chapters 4 and 5. This method has a particular appeal because it attempts to quantify the relative importance of each of the four components of variation in the response variable: (a) non-spatial environmental - the fraction that can be explained by the environmental variables independent of any spatial structure; (b) spatially structured environmental - spatial structuring in leaf-miner data that is shared with the environmental data; (c) non-environmental spatial - spatial structure in leaf-miner data that is not shared with the environmental variables; and (d) unexplained (residual) variation. If a significant relationship between the holly leaf-miner data and the environmental variables has been established, the incorporation of the spatial coordinates and subsequent calculation of these fractions allows distinctions to be made between four plausible causal models (figure 7.1).

The first model (figure 7.1a) is that the spatial structure in holly leaf-miner densities is independently caused by both the measured environmental variables and by other spatially-structured causal factors not measured in the survey. If this model is supported by the data, the regression of the environmental variables on the spatial coordinates matrix would not be significant - that is the component (b + c) of the partial fractions. The second model (figure 7.1b) is that the spatial structure in the holly leaf-miner densities is determined by both the spatial structure of the environmental variables and also other factors not measured in the survey. For this model to be supported by the data, fractions (a), (b) and (c) would all individually account for a significant proportion of the variation in leaf-miner densities. The third model (figure 7.1c) is that the spatial structure in holly leaf-miner densities is determined only by the spatial structure of the environmental variables measured. For this to be supported by the data, the fraction (c) (space only) from the partial regression analysis would not be significant. Finally, the fourth model (figure 7.1d) is that both leaf-miner densities and the environmental variables are independently controlled by unmeasured factors. That is, the correlation is spurious since the variables only appear to be related due to a shared spatial structure. This is supported by the data when fraction (a) (environment

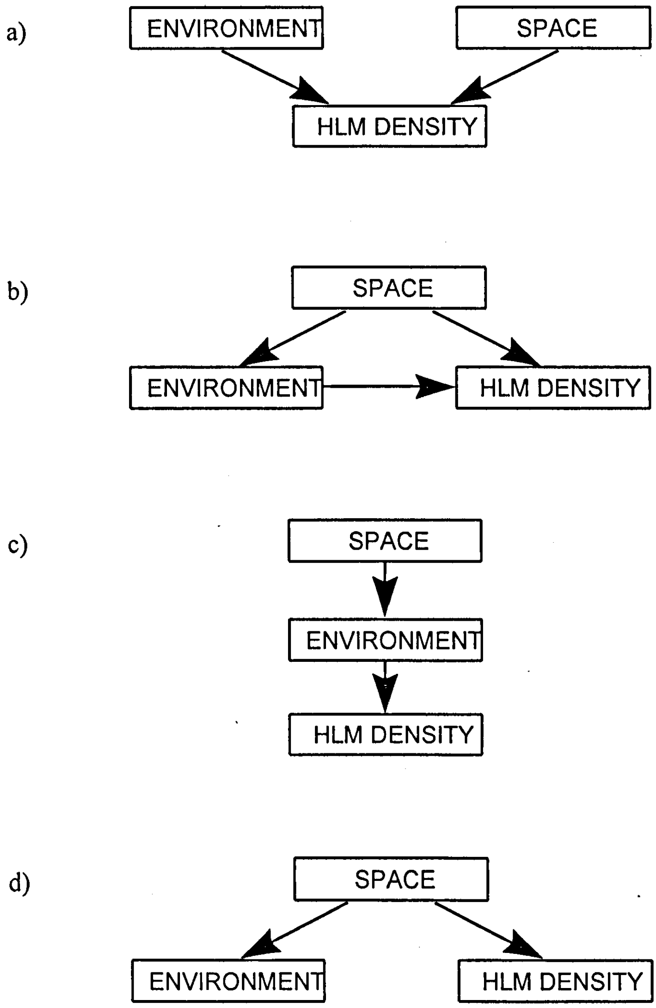


Figure 7.1 The potential causal relationships between environmental data, spatial structure and holly leaf-miner data that are tested when using partial regression analysis (see text for a full explanation)

only) is not significant but the fractions (c) and (d) are. Of course this does not mean that a causal relationship between these environmental factors and holly leaf-miner densities cannot exist. What it does indicate is that there is insufficient statistical evidence to support this hypothesis. In the case of the environmental variables that were measured in the European survey (altitude and the meteorological variables), it is difficult to imagine other underlying environmental variables that could be responsible for the shared spatial structure between them and the holly leaf-miners. However, at smaller spatial scales and with different environmental variables this may often be the case. More important, is that large scale range structure arising as a result of fixed patterns of environmental variation is only an hypothesis, plausible as it sounds. Alternative hypotheses do exist. For example, range structure might arise as a result of historical factors such as the radiation of individuals from geographically restricted populations. Alternatively, stochastic modelling has suggested that large scale patterns in abundance can be generated through temporal environmental variation with no fixed spatial component (Ives & Klopfer, 1997).

Partial regression analysis is a useful technique but it does have some limitations. Despite the appeal of quantifying the different components of the variation in the leaf-miner data, these components are not strictly additive, as was recently pointed out by Legendre & Legendre (1998). While this does not render the approach uninterpretable, its utility with respect to other techniques does diminish somewhat. Occasionally the analysis may yield a negative fraction for the spatially structured environmental variation which is difficult to interpret in the light of the four models described above. A negative fraction was obtained in the analysis of leaf-miner densities in occupied sites only (Chapter 4). Another aspect of the technique which should be considered is that the spatial structure in the variables is modelled explicitly using a trend surface derived from a third order polynomial expression in the *x* and *y* coordinates of the data. While such a surface may be sufficient to account for large scale trends (for example, gradients and troughs), it may not always provide a reasonable fit to data from intensive surveys where the spatial structure of local densities and the environment themselves may be complex (for example multiple peaks). Fitting trends to data is necessarily a simplification and care should be taken to examine this fit to ensure it is adequate before proceeding with the partial regression analysis itself.

An alternative approach that does not incorporate an explicit model of spatial structure is the use of partial Mantel tests (Legendre, 1993). The Mantel test (Mantel,

1967) is based on distance matrices and its statistic is the sum of the cross-products of the corresponding values of two matrices, for example holly leaf-miner data (A) and an environmental variable (B). The resulting Mantel coefficient, which can be normalised to take values in the range -1 to +1, yields a correlation between the two matrices. It is not identical to the correlation between the variables since it measures the extent to which the variation within the first matrix corresponds to that of the second. A partial correlation between the two matrices A and B can be computed while controlling for the effect of another matrix C (denoted (A.B).C). The significance of both simple and partial Mantel coefficients is checked by computing their associated probabilities by permutation (Legendre & Fortin, 1989). If Matrix C represents the geographic distances between the data points, then this partial Mantel coefficient would yield the correlation between the holly leaf-miner data and the environmental variable, while controlling for spatial structure. Rossi & Quénéhervé (1998) provide a good example of the use of partial Mantel tests in the consideration of nematode distribution with respect to soil properties within two different plots. The technique may also prove useful in analysing the leaf-miner data, particularly at regional scales where the sample units are individual trees and there is complex spatial structure which may not be adequately modelled using a trend surface. In order to do this, additional software would have to be written and thus the approach is beyond the scope of this thesis.

### 2. Mapping and interpolation

For all of the maps produced from the survey data sets, a simple method of interpolation was used. The point estimates derived from the holly leaf-miner data across the whole of the range were calculated by taking a weighted mean value from a number of nearest neighbours. The weights are proportional to the inverse squared distances between the actual data points and the point to be estimated. For a visual summary of the spatial structure of a data set this method is often adequate. However, this method of interpolation can be sensitive to the distribution of points across space. With an irregular distribution, such as the sampling sites during the European survey, the resulting surface would not likely provide accurate estimates of the data at unknown sites. Of course, this was not the purpose of the maps produced here. However, kriging is an alternative geostatistical method which is less sensitive to the distribution of sample points should a more accurate predictive surface be required. It incorporates statistical summaries of the spatial dependence of the data set through the use of variograms (closely allied to the correlograms presented in this thesis) and the identification of larger scale trends in the data. More sophisticated forms of kriging will also incorporate additional environmental information when producing each point estimate (see Isaaks & Srivistava, 1989, for a summary of these techniques). Kriging is available in some comprehensive GIS software packages.

The geostatistical techniques used in this thesis illustrate how the incorporation of spatial information in the analysis of ecological data can help to identify structure and improve both hypothesis testing and visualisation of spatial structure. This thesis has necessarily utilised only a small number of the techniques available and it is likely that further exploration of geostatistical methods will prove useful.

### *Conclusion*

The holly leaf-miner system has provided an unusual opportunity to examine range structure in some detail in terms of underlying population processes using relatively simple surveying techniques. This thesis has helped to highlight a number of important issues when using such an approach, for example the importance of spatial scale, the consideration of density-dependent aspects of local population dynamics and accounting for spatial dependence when testing hypotheses about the causes of range structure. Surveying methods alone cannot be sufficient to fully test current macroecological models of range structure but they are an essential first step. Given the current paucity of reliable data, with its limited taxonomic and geographic domain, it has yet to be established whether generalities can be made about the pattern of range structure let alone the processes that may be responsible for creating them. It is hoped that the work presented here for *P. ilicis* has demonstrated that field investigation of large scale pattern is tractable even with rather limited resources and this will encourage further research into other suitable systems across a variety of geographic regions.

## References

Alonso C. (1999) Variation in herbivory by *Yponomeuta mahalebella* on its only host plant *Prunus mahaleb* along an elevational gradient. Ecological Entomology **24**:371-379.

Anstead J. (1996) Phenolics in holly: Their distribution with respect to aspect and the presence of holly leaf miners. Unpublished BSc. Thesis. University of East Anglia.

Auerbach M.J., Connor E.F. & Mopper S. (1995) Minor miners and major miners: population dynamics of leaf-mining insects. In Cappuccino N. & Price P.W. (editors) Population Dynamics: New Approaches and Synthesis. Academic Press. San Diego

Athey L.A. & Connor E.F. (1989) The relationship between foliar nitrogen-content and feeding by *Odonata dorsalis* Thun on *Robinia pseudocacia* L. Oecologia **79**:390-394.

Austin M.P. (1985) Continuum concept, ordination methods, and niche theory. Annual Review of Ecological Systemtatics. **16**:39-61

Bart, J. & Klosiewski, S.P. (1989) Use of presence-absence to measure changes in avian density. Journal of Wildlife Management **53**:847-852.

Bell, C.P. (1996) The relationship between geographic variation in clutch size and migration pattern in the Yellow Wagtail. Bird Study **43**:333-341.

Bell G., Lechowicz M.J., Appenzeller A., Chandler M., DeBlois E., Jackson L., MacKenzie B., Preziosi R., Schanllenberg M. & Tinker N. (1993) The spatial structure of the physical environment. Oecologia **96**:114-121.

Bellows Jr. T.S., Van Driesche R.G. & Elkinton J.S. (1992) Life-table construction and analysis in the evaluation of natural enemies. Annual Review of Entomology **37**:587-614.

Blackburn, T.M., Gaston, K.J., Quinn, R.M. & Gregory, R.D. (1999) Do local abundances of British birds change with proximity to range edge? Journal of Biogeography 26, 493-505.

Blackburn T.M., Lawton J.H. & Perry J.N. (1992) A method of estimating the slope and upper bounds of plots of body size and abundance in natural animal assemblages. Oikos **65**: 107-112.

Blackman R.L. & Eastop V.F. (1994) Aphids on the World's Trees: An Identification and Information Guide. CABI Publishing, CAB International

Bock C.E., Bock J.H. & Lepthien L.W. (1977) Abundance patterns of some bird species wintering on the Great Plains of the USA Journal of Biogeography **4**:101-110.

Borcard D., Legendre P. & Drapeau P. (1992) Partialling out the spatial component of ecological variation. Ecology **73**:1045-1055.

Brown J.H. (1984) On the relationship between abundance and distribution of species. The American Naturalist. **124**: 255-279.

Brown J.H. (1995) Macroecology. The University of Chicago Press. Chicago.

Brown J.H. & Lomolino M.V. (1998) Biogeography. Second Edition. Sinauer Associates, Sunderland, Massachusetts.

Brown J.H. & Maurer B.A. (1989) Macroecology - the division of food and space among species on continents. Science **243**:1145-1150.

Brown, J.H. Mehlman D.W. & Stevens G.C. (1995) Spatial variation in abundance. Ecology **76**: 2028-2043.

Brown J.H., Stevens G.C. & Kaufman D.M. (1996) The geographic range: size, shape, boundaries and internal structure. Annual Review of Ecology and Systematics **27**:597-623.

Brussard P.F. (1984) Geographic patterns and environmental gradients: the central-marginal model in *Drosophila* revisited. Annual Review of Ecology and Systematics **15**:25-64.

Bultman T.L. and Faeth S.H. (1985) Patterns of intra- and interspecific association in leaf-mining insect on three oak host species. Ecological Entomology **10**:121-129.

Bylund H. & Tenow O. (1994) Long-term dynamics of leaf-miners, *Eriocrania* spp., on mountain birch - alternate year fluctuations and interaction with *Epirrita autumnata*. Ecological Entomology **19**:310-318.

Cameron E. (1939) The holly leaf miner (*Phytomyza ilicis*, Curtis) and its parasites. *Bulletin of Entomological Research*, **30**:173-208.

Cammell M.E., Tatchell G.M. & Woiwod I.P. (1989) Spatial pattern of abundance of the blackbean aphid, *Aphis fabae*, in Britain. Journal of Applied Ecology **26**:463-472.

Carey P.D., Watkinson A.R. & Gerard F.F.O. (1995) The determinants of the distribution and abundance of the winter annual grass *Vulpia ciliata* ssp. *ambigua*. Journal of Ecology **83**:177-187.

Carroll S.S. & Pearson D.L. (1998) The effects of scale and sample size on the prediction accuracy of spatial estimates of tiger beetle species diversity Ecography **21**:404-414.

Carter R.N. & Prince S.D. (1985) The geographical distribution of prickly lettuce (*Lactuca serriola*). I. A general survey of its habitats and performance in Britain. Journal of Ecology **73**: 7-38.

Caughley G., Grice D., Barker R. & Brown B. (1988) The edge of the range. Journal of Animal Ecology **57**:771-785.

Cliff A.D. & Ord J.K. (1973) Spatial autocorrelation. Pion. London.

Cliff A.D. & Ord J.K. (1981) Spatial processes: models and applications. Pion, London.

Clifford P. & Richardson S. (1985) Testing the association between two spatial processes. Statistics and Decisions Supplement **2**:155-160.

Clifford P., Richardson, S. & Hémon, D. (1989) Assessing the significance of the correlation between two spatial processes. Biometrics **45**:123-134.

Collett D. (1991) Modelling Binary Data. Chapman & Hall, London.

Connor E.F. (1984) The causes of overwintering mortality in *Phyllonorycter* on *Quercus robur*. Ecological Entomology **9**:23-28.

Connor E.F. & Taverner M.P. (1997) The evolution and adaptive significance of the leaf-mining habit. Oikos **79**:6-25.

Connor E.F., Adams Manson R.J., Carr T.G. & Beck M.W. (1994) The effects of host plant phenology on the demography and population dynamics of the leaf-mining moth, *Cameraria hamadryadella* (Lepidoptera: Gracillariidae). Ecological Entomology **19**:111-120.

Corbacho C., Sánchez J.M. & Sánchez A. (1997) Breeding biology of Montagu's Harrier *Circus pygargus* L. in agricultural environments of southwest Spain; comparison with other populations in the western Palearctic. Bird Study **44**:166-175.

Cornell H.V. (1989) Endophage-ectophage ratios and plant defense. Evolutionary Ecology **3**:64-76.

Crawley M.J. (1993) GLIM for Ecologists. Blackwell Scientific Publications. London.

Curnutt J.L., Pimm S.L. & Maurer B.A. (1996) Population variability of sparrows in space and time. Oikos **76**:131-144.

Dennett D.C. (1995) Darwin's Dangerous Idea. Penguin Books, London.

Denno R.F. and McClure M.S. eds. (1983) Variable plants and herbivores in natural and managed systems. Academic Press, New York.

Dias P.C. (1996) Sources and sinks in population biology. Trends in Ecology and Evolution **11**:326-330.

Dormer K.J. & Hucker J. (1957) Observations on the occurrence of prickles on the leaves of *Ilex aquifolium*. Annals of Botany **21**:385-398

Dunn, P.O., Thusius, K.J., Kimber, K. & Winkler, D.W. (2000) Geographic and ecological variation in clutch size of tree swallows. Auk **117**:15-221.

Eber S. & Brandl R. (1994) Ecological and genetic spatial patterns of *Urophora cardui* (Diptera: Tephritidae) as evidence for population structure and biogeographical processes. Journal of Animal Ecology **63**:187-199.

Emlen J.T., De Jong M.J., Jaeger M.J., Moermond T.C., Rusterholz K.A., White R.P. (1986) Density trends and range boundary constraints of forest birds along a latitudinal gradient. Auk **105**:791-803.

Enquist B.J., Jordan M.A. & Brown J.H. (1995) Connections between ecology, biogeography, and paleobiology: relationship between local abundance and geographic distribution in fossil and recent molluscs. Evolutionary Ecology **9**:586-604.

Environmental Systems Research Institute, Inc. (1997) ArcInfo GIS Version 7.1.2, Redlands, California USA http://www.esri.com

Environmental Systems Research Institute, Inc. (1997) ArcView GIS Version 3.0, Redlands, California, USA http://www.esri.com

Fitton M.G., Graham M.W.R.de V., Boucek Z.R.J., Ferguson N.D.M., Huddleston T., Quinlan J. & Richards O.W. (1978) Handbooks for the identification of British Insects. Vol. XI, Part 4: Hymenoptera. Royal Entomological Society. London.

Fritz & Sims (1992)

Galle F.C. (1997) Hollies. The Genus *Ilex*. Timber Press. Portland. Oregon.

Gaston K.J. (1991) How large is a species' geographic range? Oikos **61**: 434-438.

Gaston K.J. (1994) Rarity. Chapman & Hall. London.

Gaston K.J. (1996) The multiple forms of the interspecific abundance-distribution relationship. Oikos **76**: 211-220.

Gaston K.J. (1999) Implications of interspecific and intraspecific abundance-occupancy relationships. Oikos **86**: 195-207.

Gaston K.J., Reavey D. & Valladares G.R. (1991) Changes in feeding habit as caterpillars grow. Ecological Entomology **16**:339-344.

Gaston K.J., Blackburn T.M. & Lawton J.H. (1997) Interspecific abundance-range size relationships: an appraisal of mechanisms. Journal of Animal Ecology **66**: 579-601.

Gauch H.G. (1982) Multivariate analysis in community ecology. Cambridge University

Press, Cambridge.

Gauch H.G. & Whittaker R.H. (1972) Coenocline simulation. Ecology **53**:1382-1390.

Getis A. (1990) Screening for spatial dependence in regression analysis. Papers of the Regional Science Association **69**: 69-81.

GLIM 3.77 (1985) Royal Statistical Society. London.

Grice D., Caughley G. & Short J. (1986) Density and distribution of the Australian Bustard *Ardeotis australis*. Biological Conservation. **35**:259-267.

Griffiths G.C.D. & Piercey-Normore M.D. (1995) A New Agromyzid (Diptera) Leaf-miner of Mountain Holly (*Nemopanthus*, Aquifoliaceae) from the Avalon Peninsula, Newfoundland. The Canadian Field-Naturalist **109**:23-26

Griggs R.F. (1914) Observations on the behaviour of some species at the edges of their ranges. Bulletin of the Torrey Botanical Club **41**:25-49.

Hairston N.G., Smith F.E., and Slobodkin L.B. (1960). Community structure, population control, and competition. American Naturalist **94**:421-425.

Hall C.A.S., Stanford J.A. & Hauer F.R. (1992) The distribution and abundance of organisms as a consequence of energy balances along multiple environmental gradients. Oikos **65**:377-390.

Hanks L.M. & Denno R.F. (1993) Natural enemies and plant water relations influence the distribution of an armoured scale insect. Ecology **74**;1081-1091.

Hanski I. (1999) Metapopulation Ecology (Oxford Series in Ecology and Evolution) Oxford University Press.

Hansson C. (1985) Taxonomy and biology of the Palearctic species of *Chrysocharis* Förster, 1856 (Hymenoptera: Eulophidae). Entomologica Scandinavica Supplement No. 26.

Harrison S. (1997) Persistent, localized outbreaks in the western tussock moth *Orgyia vetusta*: the roles of resource quantity, predation and poor dispersal. Ecological Entomology **22**:158-166.

Harrison S. and Maron J.L. (1995) Impacts of defoliation by tussock moths (*Orgyia vetusta*) on the growth and reproduction of bush lupine (*Lupinus arboreus*). Ecological Entomology **20**:223-229.

Heads P.A. and Lawton J.H. (1983) Studies on the natural enemy complex of the holly leaf miner: the effects of scale on the detection of aggregative responses and the implications for biological control. *Oikos* **40**:267-276.

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

Hengeveld, R. (1990) Dynamic biogeography. Cambridge University Press, Cambridge.

Hengeveld R., Kooijman S.A.L.M. & Taillie C. (1979) A spatial model explaining species-abundance curves. In: Statistical distributions in ecological work (eds J.K. Ord, G.P. Patil & C. Taillie), pp. 333-347. International Co-operative Publishing House, Fairland, Maryland.

Hengeveld, R. & Haeck, J. (1981) The distribution of abundance. II. Models and implications. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen C84, 257-284.

Hengeveld R. & Haeck J. (1982) The distribution of abundance. I. Measurements. Journal of Biogeography **9**:303-316.

Hoffmann A.A. & Blows M.W. (1994) Species borders: ecological and evolutionary perspectives. Trends in Ecology and Evolution **9**:223-227.

Holt R.D., Lawton J.H., Gaston K.J. & Blackburn T.M. (1997) On the relationship between range size and local abundance: back to basics. Oikos **78**:183-190

Hultén, E. & Fries, M. (1986) Atlas of North European vascular plants north of the Tropic of Cancer. Königstain: Koeltz Scientific Books.

Isaaks, E.H. & Srivastava, R.M. (1989) An introduction to applied geostatistics. Oxford University Press, Oxford.

Ives & Klopfer (1997) Spatial variation in abundance created by stochastic temporal variation. Ecology **78**:1907-1913.

Iversen J. (1944) *Viscum*, *Hedera* and *Ilex* as climatic indicators. Geol. For. Stockh. Forh. **66**:463-483

James R. & Pritchard I.M. (1988) Influence of the holly leaf miner, *Phytomyza ilicis* (Diptera Agromyzidae), on leaf abscission.

Kahn D.M. and H.V. Cornell. (1983) Early leaf abscission and folivores: comments and considerations. American Naturalist **122**:428-432.

Karban, R. (1989)

Kendeigh, S.C. (1974) Ecology with special reference to animals and man. Prentice Hall, Englewood Cliffs, New Jersey.

Kitron U., Otieno L.H., Hungerford L.L., Odulaja A., Brigham W.U., Okello O.O., Joselyn M., Mohamed-Ahmed M.M. and Cook E. (1996) Spatial analysis of the distribution of tsetse flies in the Lambwe Valley, Kenya, using Landsat TM satellite imagery and GIS. Journal of Animal Ecology **65**:371-380.

Koenig, W.D. (1984) Geographic variation in clutch size in the northern flicker (*Colaptes auratus*): support for Ashmole's hypothesis. Auk **101**:698-706

Koenig, W.D. (1986) Geographical ecology of clutch size variation in North American woodpeckers. Condor **88**:499-504.

Koenig, W.D. (1998) Spatial autocorrelation California land birds. Conservation Biology **12**:612-620.

Koenig, W.D. (1999) Spatial autocorrelation of ecological phenomena. Trends in Ecology and Evolution **14**:22-26.

Kulp L.A. (1968) The taxonomic status of dipterous holly leaf miners (Diptera: Agromyzidae). University of Maryland Agricultural Experimental Station, Bulletin A-155.

Krebs C.J. (1989) Ecological methodology. HarperCollinsPublishers, NewYork.

Lawton J.H. (1996) Population abundances, geographic ranges and conservation: 1994 Witherby Memorial Lecture. Bird Study **43**:3-19.

Lawton J.H. (1993) Range, population abundance and conservation. Trends in Ecology and Evolution **8**:409-413.

Legendre P. (1990) Quantitative methods and biogeographic analysis. Evolutionary biogeography of the marine algae of the North Atlantic. NATO ASI Series. Volume G 22 - Garbary D.J. & South R.R. (eds.) Springer-Verlag. Berlin.

Legendre P. (1993) Spatial autocorrelation: trouble or new paradigm? Ecology **74**:1659-1673.

Legendre P. & Fortin M-J. (1989) Spatial pattern and ecological analysis. Vegetatio **80**:107-138.

Legendre P. & Legendre L. (1998) Numerical Ecology. Developments in Environmental Modelling, 20. Second English Edition. Elsevier. Amsterdam.

Lewis, T and Taylor, L.R. (1967) Introduction to experimental ecology. N.B. pp. 202-208. London: Academic Press.

MacArthur R.H. (1972) Geographical ecology: patterns in the distribution of species. Harper & Row, New York.

Mantel N. (1967) The detection of disease clustering and a generalised regression approach. Cancer Research **27**:209-220.

Mattson W.J. & Addy N.D. (1975) Phytophagous insects as regulators of forest primary production. Science **190**:515-522.

Maurer B.A. (1994) Geographical Population Analysis: Tools for the Analysis of Biodiversity. Blackwell Scientific Publications. Oxford

Maurer B.A. (1999) Untangling Ecological Complexity: The Macroscopic Perspective. The University of Chicago Press. Chicago.

Maurer B.A. & Brown J.H. (1989) Distributional consequences of spatial variation in local demographic processes. Annales Zoologici Fennici **26**:121-131.

Maurer B.A. & Villard M. (1994) Geographic variation in abundance of North American Birds. Research and Exploration.

Mayr E. (1963) Animal species and evolution. The Belknap press of the Harvard university press, Cambridge Massachusetts & Oxford university press, London.

McClure M.S. & Price P.W. (1976) Ecotope characteristics of coexisting *Erythroneura* leafhoppers (Homoptera: Cicadellidae) on sycamore. Ecology **57**:928-940.

McGeoch and Gaston (2000) Edge effects on the prevalence and mortality factors of *Phytomyza ilicis* (Diptera, Agromyzidae) in a suburban woodland.

Minkenberg O.P.J. & Ottenheim J.J.G.W. (1990) Effects of leaf nitrogen content of tomato plants on preference and performance of a leafmining fly. Oecologia **83**:291-298.

Mopper S., Whitham T.G. & Price P.W. (1990) Plant phenotype and interspecific competition between insects determine sawfly performance and density. Ecology **71**:2135-2144.

Nestel D., Dickschen F. & Altien M.A. (1994) Seasonal and spatial population loads of a tropical insect: the case of the coffee leaf-miner in Mexico. Ecological Entomology **19**:159-167.

New M. G., Hulme M. & Jones P. D. (2000) Representing twentieth-century space-time climate variability. Part II: Development of 1901-1996 monthly grids of terrestrial surface climate. Journal of Climate **13**:2217-2238.

Obeso J.R. (1997) The induction of spinescence in European holly leaves by browsing ungulates. Plant Ecology **129**:149-156

Peakall D.B. (1970) The eastern bluebird: its breeding season, clutch size, and nesting success. The Living Bird **: 239-256.

Perring F.H. & Walters S.M. (1962) Atlas of the British Flora. London.

Perry J.N. & Taylor L.R. (1985) Adès: New ecological families of species-specific frequency distributions that describe repeated spatial samples with an intrinsic power-law variance-mean property. Journal of Animal Ecology **54**:931-953.

Perry J.N. & Taylor L.R. (1986) Stability of real interacting populations in space and time: implications, alternatives and the negative binomial $k_c$. Journal of Animal Ecology **55**:1053-1068.

Peterken G.F. & Lloyd P.S. (1967) Biological Flora of the British Isles: *Ilex aquifolium* L. Journal of Ecology **55**:841-858.

Potter D.A. (1985) Population regulation of the holly leafminer, *Phytomyza ilicicola* Loew (Diptera: Agromyzidae), on American holly. Oecologia **66**:499-505.

Potter D.A. & Kimmerer T.W. (1986) Seasonal allocation of defense investment in *Ilex opaca* Aiton and constraints on a specialist leafminer. Oecologia **69**:217-224.

Price J., Droege S. & Price A. (1995) The Summer Atlas of North American Birds. Academic Press, London.

Pulliam H.R. (1988) Sources, sinks, and population regulation. American Naturalist **132**:652-661.

Pulliam H.R. & Danielson B.J. (1991) Sources, sinks and habitat selection: a landscape perspective on population dynamics. The American Naturalist **137 Supplement**:550-566.

Quinn R.M., Gaston K.J. & Roy D.B. (1997) Coincidence between consumer and host occurrence: macrolepidoptera in Britain. Ecological Entomology **22**:197-208.

Quinn R.M., Gaston K.J. & Roy D.B. (1998) Coincidence in the distributions of butterflies and their foodplants. Ecography **21**:279-288.

Randall M.G.M. (1982) The dynamics of an insect population throughout its altitudinal distribution: *Coleophora alticolella* (Lepidoptera) in northern England. Journal of Animal Ecology **51**:993-1016.

Rapoport E.H. (1982) Areography: geographical strategies of species. Pergamon, Oxford.

Rausher M.D. (1983) Ecology of host-selection behaviour in phytophagous insects. In Denno and McClure (1983) pp. 223-257.

Redfearn M. and Cameron A.D. (1993) Population dynamics of the yew gall midge *Taxomyia taxi* and its chalcid parasitoids: a 24 year study. Ecological Entomology **18**:365-378.

Ricklefs R.E. (1972) Latitudinal variation in breeding productivity of the rough-winged swallow. Auk **89**:826-836.

Rogers D.J. & Randolph S.E. (1986) Distribution and abundance of tsetse flies (*Glossina* spp.). Journal of Animal Ecology **55**:1007-1025.

Root T. (1988) Atlas of Wintering North American birds. University of Chicago Press, Chicago.

Root (1988c) quoted in Wiens (1989) (reference unavailable).

Root T. (1995) Atlas of Wintering North American Birds. An Analysis of the Christmas Bird Count Data. The University of Chicago Press. Chicago.

Rosenthal G.A. and Janzen D.H. eds. (1979) Herbivores: their interaction with secondary plant metabolites. Academic Press, New York.

Rossi R.E., Mulla D.J. Journel A.G. & Franz E.H. (1992) Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs **62**:277-314.

Safriel U.N., Volis S. & Kark S. (1994) Core and peripheral populations and global climate change. Israel Journal of Plant Sciences **42**:331-345.

Sanz J.J. (1997) Geographic variation in breeding parameters of the Pied Flycatcher *Ficedula hypoleuca.* Ibis **139**:107-114.

Shelford V.E. (1911) Physiological animal geography. J. Morphol. **22**:551-618.

Sokal R.R. and Oden N.L. (1978) Spatial autocorrelation in biology. 1.Methodology. Biological Journal of the Linnean Society **10**:199-228.

Stiling P.D., Simberloff D. S. & Broadbeck B.V. (1991) Variation in rates of leaf abscission between plants may affect the distribution patterns of sessile insects. *Oecologia* **88**:367-370.

Speight M.R., Hails R.S., Gilbert M. & Foggo A. (1998) Horse chestnut scale (*Pulivinaria regalis*) (Homoptera: Coccidae) and urban host tree environment. Ecology **79**:1503-1513.

Strong D.R., Lawton J.H. & Southwood T.R.E. (1984) Insects on Plants. Harvard University Press, Cambridge, MA.

Strayer D.L. (1991) Projected distribution of the zebra mussel, *Dreissena polymorpha*, in North America. Canadian Journal of Fisheries and Aquatic Sciences **48**:1389-1395.

Svensson B.W. (1992) Changes in occupancy, niche breadth and abundance of three *Gyrinus* species as their respective range limits are approached. Oikos **63**:147-156.

Sykes M.T., Prentice I.C. & Cramer W. (1996) A bioclimatic model for the potential distributions of north European tree species under present and future climates. Journal of Biogeography **23**:203-233

Taylor L.R. (1986) Synoptic dynamics, migration and the Rothamsted insect survey. Journal of Animal Ecology **55**:1-38.

Taylor R.A.J. & Taylor L.R. (1979) A behavioural model for the evolution of spatial dynamics, in Population Dynamics (eds Anderson R.M., Turner B.D. & Taylor L.R.). Blackwell Scientific, Oxford.

Taylor L.R. & Woiwod I.P. (1982) Comparative synoptic dynamics. I. Relationships between inter- and intra-specific spatial and temporal variance/mean population parameters. Journal of Animal Ecology **51**:879-906.

Taylor L.R., Woiwod I.P. & Perry J.N. (1978) The density dependence of spatial behaviour and the rarity of randomness. Journal of Animal Ecology **47**:383-406.

Taylor L.R., Woiwod I.P. & Perry J.N. (1980) Variance and the large scale spatial stability of aphids, moths and birds. Journal of Animal Ecology **49**:831-854.

Tellería J.L. & Santos T. (1993) Distributional patterns of insectivorous passerines in the Iberian forests: does abundance decrease near the border? Journal of Biogeography **20**:235-240.

Ter Braak C.J.F. & Prentice I.C. (1988) A theory of gradient analysis. Advances in Ecological Research. **18**:271-317.

Thompson J.N. (1988) Coevolution and alternative hypotheses on insect/plant interactions. Ecology **69**:893-895.

Thomson J.D., Weiblen G., Thomson B.A., Alfaro S. & Legendre P. (1996) Untangling multiple factors in spatial distributions: lilies, gophers, and rocks. Ecology 77:1698-1715

Tukey, J.W. (1977) Exploratory Data Analysis (First Edition). Reading, MA: Addison-Wesley Publishing Co.

Valladares, G. and Lawton, J.H. (1991) Host-plant selection in the holly leaf miner: does mother know best? Journal of Animal Ecology **60**:227-240.

Virtanen T. & Neuvonen S. (1999) Performance of moth larvae on birch in relation to altitude, climate, host quality and parasitoids. Oecologia **120**:92-101.

Whitcomb R.F., Hicks A.L., Blocker H.D. & Lynn D.E. (1994) Biogeography of leafhopper specialists of the shortgrass prairie: evidence for the roles of phenology and phylogeny in determination of biological diversity. American Entomologist (Spring), 19-35.

Whittaker J.B. (1971) Population changes in *Neophilaenus lineatus* (L.) (Homoptera: Cercopidae) in different parts of its range. Journal of Animal Ecology **40**:425-443.

Wiens J.A. (1989) The Ecology of Bird Communities. Volume 1 Foundations and Patterns. Cambridge University Press. Cambridge.

Williams P.H. (1988) Habitat use by bumblebees (*Bombus* spp.) Ecological Entomology **13**:223-237.

Woods K.D. & Davis M.B. (1989) Paleoecology of range limits: beech in the upper peninsula of Michigan. Ecology **70**:681-696.

# Appendices

## Program source code and documentation

These are edited versions of the source code and internal documentation for the programs referred to in this thesis.

Visual Basic for Applications (VBA) in Microsoft Excel 95 requires comments to preceded by a ' symbol and all code and comments to be pure text (no formatting). In addition, there is no limitation on line length in VBA, although lines may be broken up with the _ symbol. However, the program has been reformatted here for clarity and a single line of code may wrap onto the next line. The following formatting conventions are used:

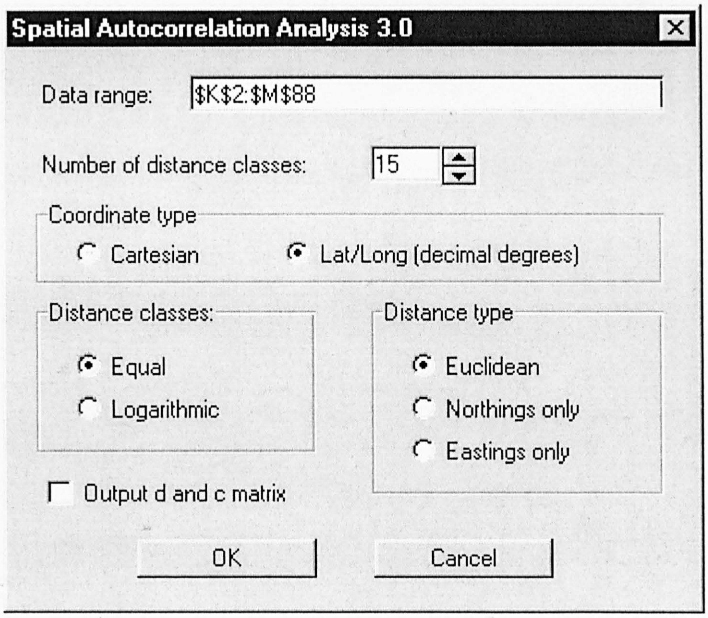
`original source code`

*comments*

## Appendix 1

## Source code and documentation for the spatial autocorrelation analysis program

This program was compiled from the workbook format to an add-in, which automatically added the spatial autocorrelation analysis menu and functionality to Microsoft Excel95. The figure below shows the dialog box presented when the program is run. The data range is three columns of a worksheet that contain the *x* (x-coordinates), *y* (y-coordinates) and *z* (response variable) values for the analysis respectively. Options are given to treat distances between points as though they are on a plane (Cartesian coordinates) or to take into account the curvature of the Earth's surface (Lat/Long (decimal degrees)). Distance classes may be equal or logarithmic. An option is also given for the output of both the distance and similarity matrices from the analysis.


Spatial Autocorrelation Analysis 3.0

Data range: $K$2:$M$88

Number of distance classes: 15

Coordinate type
○ Cartesian
● Lat/Long (decimal degrees)

Distance classes:
● Equal
○ Logarithmic

Distance type
● Euclidean
○ Northings only
○ Eastings only

☐ Output d and c matrix

---

*Spatial Autocorrelation Analysis Program*
*Version 3*
*Author: Andrew M. Brewer*
*Last Revision: 17/3/99*

*This program conducts a spatial autocorrelation analysis on a data set in Microsoft Excel 95.*
*The data should be in 3 columns in the following order: x y z*
*where*
*x is the x co-ordinate (or decimal degrees longitude) for the point,*
*y is the y co-ordinate (or decimal degrees latitude) for the point*
*z is the value at this point.*
*format for latitude and longitude:*
*latitude: north positive, south negative*
*longitude: east positive, west negative*
*note: only Euclidean distances can currently be calculated when using the decimal degrees coordinates system in this version of the software.*

*Changes since version 2.01:*
*User is given the option of choosing between Cartesian and decimal degree coordinates system.*

---

## *Module 1*
## *Auto Open and Close*

*The procedures in this module manage the menus that appear when the add-in is loaded.*

```
Option Explicit
Option Private Module
```

*Auto_Open() adds the Sheffield menu to the menubars*

```
Sub Auto_Open()

     Application.StatusBar = "Adding Sheffield menu ..."

     With MenuBars(xlModule)
          .Menus.Add Caption:="&Sheffield"
          .Menus("Sheffield").MenuItems.Add _
               Caption:="&Spatial Autocorrelation Analysis ...", _
               OnAction:="Spatial Autocorrelation Analysis (version
               3.0).xls!main.main"
     End With

     With MenuBars(xlWorksheet)
          .Menus.Add Caption:="&Sheffield"
          .Menus("Sheffield").MenuItems.Add _
               Caption:="&Spatial Autocorrelation Analysis ...", _
               OnAction:="Spatial Autocorrelation Analysis (version
               3.0).xls!main.main"
     End With

     With MenuBars(xlChart)
          .Menus.Add Caption:="&Sheffield"
          .Menus("Sheffield").MenuItems.Add _
               Caption:="&Spatial Autocorrelation Analysis ...", _
               OnAction:="Spatial Autocorrelation Analysis (version
               3.0).xls!main.main"
     End With

     With MenuBars(xlNoDocuments)
          .Menus.Add Caption:="&Sheffield"
          .Menus("Sheffield").MenuItems.Add _
               Caption:="&Spatial Autocorrelation Analysis ...", _
               OnAction:="Spatial Autocorrelation Analysis (version
               3.0).xls!main.main"
     End With

     Application.StatusBar = False

End Sub
```

*Auto_Close() removes the Sheffield menu from the menubars*

```
Sub Auto_Close()

     Application.StatusBar = "Restoring menus ..."

     MenuBars(xlModule).Reset
     MenuBars(xlWorksheet).Reset
     MenuBars(xlChart).Reset
     MenuBars(xlNoDocuments).Reset

     Application.StatusBar = False
```

```
End Sub
```

---

### *Module 2*
### *Main*

```
Option Explicit
Option Base 1

Type data_record
     x_coord As Double
     y_coord As Double
     z_value As Double
End Type
Data type for each point of the data set

Type I_record
     pairs As Double
     I As Double
     exp_I As Double
     sd_I As Double
     sig_I As Double
End Type
Data type for the Moran's I statistics calculated for each distance class

Public data() As data_record
Matrix holding all the data
Public dc_number As Integer
Number of distance classes
Public dc_type As Integer
Type of distance class to be used
Public n As Integer
Number of data points
Public distances() As Double
Distance matrix for data
Public dc_matrix() As Double
Array of distance classes
Public mean_z As Double
Mean value of z
Public variance_z As Double
Variance of z
Public c_matrix() As Double
Similarity matrix for the data
Public Morans_I() As I_record
Array holding values and significances of I
Public bon_sig_level As Double
Bonferroni significance level

Public Const equal_dc = 1
Public Const log_dc = 2
Distance class types

Sub Main()

     Dim data_range As Range
     Worksheet range containing the data
     Dim distance_type As Integer
     Dim wkbook As Workbook
     Workbook containing the data

     Display dialog box
     With ThisWorkbook.DialogSheets("Main Dialog")
     '.EditBoxes("dr_editbox").Text = ""
```

*If the above line is commented out, then the data range edit box saves the range from the last time the program was run. This is useful when conducting several analyses on worksheets with the same size data range.*

```
          If .Show Then
               Initialise variables from dialog
               dc_number = .EditBoxes("dc_editbox").Text
               Set data_range = Range(.EditBoxes("dr_editbox").Text)
               If .OptionButtons("option_dc_equal").Value = xlOn Then
                    dc_type = equal_dc
               Else
                    Option_dc_log is on
                    dc_type = log_dc
               End If
          Else
               Program cancelled
               End
          End If
     End With

     Application.ScreenUpdating = False

     Application.StatusBar = "Reading data from spreadsheet ..."
     n = read_data(data_range)
     Read the data into the array
     Application.StatusBar = "Calculating distance matrix ..."
     calc_distances
     Calculate the distance matrix
     calc_dc_matrix
     Calculate the distance class matrix
     mean_z = calc_mean
     Calculate the mean value of z
     variance_z = calc_variance
     Calculate the variance of z
     Application.StatusBar = "Calculating similarity matrix ..."
     calc_c_matrix
     Calculate the similarity matrix
     calc_I
     Calculate Moran's I

     Application.StatusBar = "Writing results to spreadsheet ..."
     output_results data_range

     Application.StatusBar = False
     Application.ScreenUpdating = True

End Sub Main
```

*read_data*

*This function reads in data from the chosen range on a worksheet into the data array. It dynamically allocates memory to this array to accommodate the data and returns the number of points read in.*

```
Function read_data(ByVal datarange As Range) As Integer

     Dim RowNumber As Integer
     Dim counter As Integer

     RowNumber = datarange.Rows.Count
     ReDim data(RowNumber)

     For counter = 1 To RowNumber
          data(counter).x_coord = datarange.Cells(counter, 1).Value
          data(counter).y_coord = datarange.Cells(counter, 2).Value
          data(counter).z_value = datarange.Cells(counter, 3).Value
     Next counter

     read_data = RowNumber
```

```
End Function read_data
```

*calc_distances*
*This procedure creates the distance matrix for the data, using the formula:*
*d12 = ((x1-x2)^2+(y1-y2)^2)^0.5 (Pythagorus) for Cartesian coordinates and an algorithm for decimal degrees coordinates to take into account the earth's curvature.*
*Note only Euclidean distances for decimal degrees coordinates system in the current version.*
*It dynamically allocates memory to this matrix.*

```
Sub calc_distances()

    Dim counter1 As Integer
    Dim counter2 As Integer

    ReDim distances(n, n)
    Allocate enough memory to accommodate matrix

    With ThisWorkbook.DialogSheets("Main Dialog")
        If .OptionButtons("option_coord_cartesian").Value = xlOn Then
        Co-ordinate type is Cartesian

        If .OptionButtons("option_d_euclidean") = xlOn Then
            For counter1 = 1 To n
                For counter2 = 1 To n
                    distances(counter1, counter2) =
                    ((data(counter1).x_coord -
                    data(counter2).x_coord) ^ 2 +
                    (data(counter1).y_coord -
                    data(counter2).y_coord) ^ 2) ^ 0.5
                Next counter2
            Next counter1
        ElseIf .OptionButtons("option_d_northing") = xlOn Then
            For counter1 = 1 To n
                For counter2 = 1 To n
                    distances(counter1, counter2) =
                    Abs(data(counter1).y_coord -
                    data(counter2).y_coord)
                Next counter2
            Next counter1
        Else
        Eastings is on
            For counter1 = 1 To n
                For counter2 = 1 To n
                    distances(counter1, counter2) =
                    Abs(data(counter1).x_coord -
                    data(counter2).x_coord)
                Next counter2
            Next counter1
        End If Cartesian coordinates

        Else
            co-ordinate type is decimal degrees
            MsgBox "Only Euclidean distances can be used for decimal
            degrees coordinates in this version of the software"
            For counter1 = 1 To n
                For counter2 = 1 To n
                    If counter1 = counter2 Then
                    If it's the same point
                        distances(counter1, counter2) = 0
                    Else
                    Use the algorithm for calculating the distances
                        distances(counter1, counter2) =
                        calc_dd_distance(data(counter1).x_coord,
                        data(counter1).y_coord,
                        data(counter2).x_coord,
                        data(counter2).y_coord)
                    End If counter1 = counter2
```

```
                    Next counter2
               Next counter1

          End If co-ordinate types
     End With

End Sub calc_distances
```

*calc_dc_matrix*
*This procedure creates the distance class matrix for the analysis, finding the maximu distance between points and using the number of classes specified by the user. It dynamically allocates memory to distance class matrix.*

```
Sub calc_dc_matrix()

     Dim max_distance As Double
     Dim counter1 As Integer
     Dim counter2 As Integer

     Find the maximum distance in the distance matrix
     max_distance = 0
     For counter1 = 1 To n
          For counter2 = 1 To n
               If distances(counter1, counter2) > max_distance Then
                    max_distance = distances(counter1, counter2)
               End If
          Next counter2
     Next counter1

     Allocate memory to the distance class matrix
     ReDim dc_matrix(dc_number)

     Insert distance classses into matrix
     If dc_type = equal_dc Then
     distance classes are equal
          For counter1 = 1 To dc_number
               dc_matrix(counter1) = counter1 * max_distance /
               dc_number
          Next counter1
     Else
     distance classes are logarithmic (dc_type = log_dc)
          For counter1 = 1 To dc_number
               dc_matrix(counter1) = counter1 * Log(max_distance) /
               dc_number
               dc_matrix(counter1) = Exp(dc_matrix(counter1))
          Next counter1
     End If
     dc_matrix(dc_number) = max_distance
     (Just in case of rounding error)

End Sub calc_dc_matrix
```

*calc_mean*
*This function calculates the arithmetic mean of z.*

```
Function calc_mean() As Double

     Dim mean As Double
     Dim counter1 As Integer

     mean = 0
     For counter1 = 1 To n
          mean = mean + data(counter1).z_value
     Next counter1

     calc_mean = mean / n
```

```
End Function calc_mean
```

*calc_variance*
*This function calculates the variance of z.*

```
Function calc_variance() As Double

     Dim variance_z As Double
     Dim counter1 As Integer

     variance_z = 0
     For counter1 = 1 To n
          variance_z = variance_z + (data(counter1).z_value - mean_z) ^
          2
     Next counter1

     calc_variance = variance_z / n

End Function calc_variance
```

*calc_c_matrix*
*This procedure creates the similarity matrix for the analysis. Each cij is the covariance between the value of z at one place and its value at another.*

```
Sub calc_c_matrix()

     Dim counter1 As Integer
     Dim counter2 As Integer

     Allocate memory for the similarity matrix
     ReDim c_matrix(n, n)

     calculate cij's
     For counter1 = 1 To n
          For counter2 = 1 To n
               c_matrix(counter1, counter2) = (data(counter1).z_value -
               mean_z) * (data(counter2).z_value - mean_z)
          Next counter2
     Next counter1

End Sub calc_c_matrix
```

*calc_dd_distance*
*Calculates the surface distance between the points(long1, lat1) and (long2, lat2), expressed as decimal degrees. The code here is based on a standard algorithm.*

```
Function calc_dd_distance(long1 As Double, lat1 As Double, long2 As
Double, lat2 As Double) As Double

Const cEarthRadius = 6367          radius of the earth in kilometres
Const cPi = 3.14159265358979       pi

Dim theta1 As Double
Dim theta2 As Double
Dim phi1 As Double
Dim phi2 As Double
Dim x1 As Double
Dim x2 As Double
Dim y1 As Double
Dim y2 As Double
Dim z1 As Double
Dim z2 As Double
Dim dist As Double

     Calculte theta, the angle from the prime meridian
     theta = longitude
```

```
    theta1 = long1
    theta2 = long2

    Calculate phi, the angle from the north pole
    phi = 90 - latitude
    phi1 = 90 - lat1
    phi2 = 90 - lat2

    Convert to radians
    radians = (degrees*2*pi)/360
    theta1 = (theta1 * 2 * cPi) / 360
    theta2 = (theta2 * 2 * cPi) / 360
    phi1 = (phi1 * 2 * cPi) / 360
    phi2 = (phi2 * 2 * cPi) / 360

    Convert to spherical cartesian coords
    north pole = (0,0,6367) in cartesian coords
    x = 6367 (cos theta) (sin phi)
    y = 6367 (sin theta) (sin phi)
    z = 6367 (cos phi)
    x1 = cEarthRadius * Cos(theta1) * Sin(phi1)
    x2 = cEarthRadius * Cos(theta2) * Sin(phi2)
    y1 = cEarthRadius * Sin(theta1) * Sin(phi1)
    y2 = cEarthRadius * Sin(theta2) * Sin(phi2)
    z1 = cEarthRadius * Cos(phi1)
    z2 = cEarthRadius * Cos(phi2)

    Calculate straight line distance between points
    dist = ((x1 - x2) ^ 2 + (y1 - y2) ^ 2 + (z1 - z2) ^ 2) ^ 0.5
    Convert to straight line distance to surface distance
```

surface distance = $2R\sin^{-1}\left(\dfrac{\text{straight line distance}}{2R}\right)$

```
    where R is the radius of the earth.

    dist = 2 * cEarthRadius * Application.Asin(dist / (2 *
    cEarthRadius))

    calc_dd_distance = dist

End Function calc_dd_distance
```

---

**Module 3**
**Moran's I**

```
Option Explicit
Option Base 1
Option Private Module

Sub calc_I()

    Dim weights() As Integer
    Binary weights for the analysis
    Dim W As Double
    Dim W_i_dot() As Double
    Dim W_dot_i() As Double
    Dim S1 As Double
    Dim S2 As Double
    Dim counter As Integer
    Dim counter1 As Integer
    Dim counter2 As Integer
    Dim Expected_I As Double
    Dim Var_I As Double
    Dim b2 As Double
```

```
Dim m2 As Double
Dim m4 As Double

ReDim Morans_I(dc_number)

For counter = 1 To dc_number
    Application.StatusBar = "Calculating Moran's I for distance
    class " & counter
    m2 is the variance of z
    m2 = variance_z

    Allocate memory to weights matrices
    ReDim weights(n, n)
    ReDim W_i_dot(n)
    ReDim W_dot_i(n)

    calc_weights weights, counter

    Initialize variables
    W = 0
    S1 = 0
    S2 = 0
    For counter1 = 1 To n
        W_i_dot(counter1) = 0
        W_dot_i(counter1) = 0
    Next counter1

    Calculate marginal totals and W
    For counter1 = 1 To n
        For counter2 = 1 To n
            W_i_dot(counter1) = W_i_dot(counter1) +
            weights(counter1, counter2)
            W_dot_i(counter2) = W_dot_i(counter2) +
            weights(counter1, counter2)
            W = W + weights(counter1, counter2)
        Next counter2
    Next counter1
    Morans_I(counter).pairs = W / 2

    Calculate Moran's I
    Morans_I(counter).I = 0
    For counter1 = 1 To n
        For counter2 = 1 To n
            Morans_I(counter).I = Morans_I(counter).I +
            weights(counter1, counter2) * c_matrix(counter1,
            counter2)
        Next counter2
    Next counter1

    If m2 = 0 Then
        MsgBox "Variance = 0 for whole data set." & Chr(13) &
        "Program terminated."
    End
    ElseIf W = 0 Then
        No data in distance class
        Morans_I(counter).pairs = 0
        Morans_I(counter).I = 0
        Morans_I(counter).exp_I = -1 / (n - 1)
        Morans_I(counter).sd_I = 0
        Morans_I(counter).sig_I = 0
    Else
        Morans_I(counter).I = Morans_I(counter).I / (W * m2)

        Calculate Expected value of I
        Morans_I(counter).exp_I = -1 / (n - 1)

        Calculate Variance of I
        For counter1 = 1 To n
```

```
                For counter2 = 1 To n
                    S1 = S1 + (weights(counter1, counter2) +
                    weights(counter2, counter1)) ^ 2
                Next counter2
            Next counter1
            S1 = S1 / 2

            For counter1 = 1 To n
                S2 = S2 + (W_i_dot(counter1) + W_dot_i(counter1)) ^
                2
            Next counter1

            m4 = 0
            For counter1 = 1 To n
                m4 = m4 + (data(counter1).z_value - mean_z) ^ 4
            Next counter1
            m4 = m4 / n

            b2 = m4 / (m2 ^ 2)

            Variance of I if assumptions of normality
            Var_I = (n ^ 2 * S1 - n * S2 + 3 * W ^ 2) / (W ^ 2 * (n ^ 2 - 1)) - Expected_I ^ 2
            Var_I = n * ((n ^ 2 - 3 * n + 3) * S1 - n * S2 + 3 * W ^
            2) - b2 * ((n ^ 2 - n) * S1 - 2 * n * S2 + 6 * W ^ 2)
            Var_I = Var_I / (CDbl(n - 1) * CDbl(n - 2))
            Var_I = Var_I / (CDbl(n - 3) * W ^ 2) - Expected_I ^ 2
            Morans_I(counter).sd_I = Var_I ^ 0.5

            Morans_I(counter).sig_I = Abs(Morans_I(counter).I -
            Morans_I(counter).exp_I) / Morans_I(counter).sd_I
            Morans_I(counter).sig_I = 1 -
            (Application.NormSDist(Morans_I(counter).sig_I))
        End If
    Next counter

    bon_sig_level = Bonferroni

End Sub calc_I

calc_weights

Sub calc_weights(ByRef weights() As Integer, ByVal dc As Integer)

Dim counter1 As Integer
Dim counter2 As Integer


    If dc = 1 Then
    If its the 1st distance class min distance = 0
        For counter1 = 1 To n
            For counter2 = 1 To n
                If counter1 <> counter2 Then
                    If distances(counter1, counter2) >= 0 And
                    distances(counter1, counter2) <= dc_matrix(dc)
                    Then
                        weights(counter1, counter2) = 1
                    Else
                        weights(counter1, counter2) = 0
                    End If
                Else
                    Pair is at the same locality, ignore
                    weights(counter1, counter2) = 0
                End If
            Next counter2
        Next counter1
    Else
        For counter1 = 1 To n
            For counter2 = 1 To n
```

```
                If distances(counter1, counter2) > dc_matrix(dc -
                1) And distances(counter1, counter2) <=
                dc_matrix(dc) Then
                    weights(counter1, counter2) = 1
                Else
                    weights(counter1, counter2) = 0
                End If
            Next counter2
        Next counter1
    End If

End Sub calc_weights
```

*Bonferroni*
*Carries out the Bonferroni test for multiple (x) comparisons. For the correlogram to be significant at the alpha level, at least one value of Moran's I should be significant at the alpha/x level.*
*This procedure calculates the Bonferroni significance for the whole correlogram*

```
Function Bonferroni() As Double

Dim counter As Integer
Dim min_alpha As Double

min_alpha = 1
Get minimum alpha from indivdual alphas of Moran's_I
    For counter = 1 To dc_number
        If Morans_I(counter).pairs <> 0 Then
            Check that sig_i is valid
            If Morans_I(counter).sig_I < min_alpha Then
                min_alpha = Morans_I(counter).sig_I
            End If
        End If
    Next counter

    If min_alpha <= (0.001 / dc_number) Then
        Bonferroni = 0.001
    ElseIf min_alpha <= (0.01 / dc_number) Then
        Bonferroni = 0.01
    ElseIf min_alpha <= (0.05 / dc_number) Then
        Bonferroni = 0.05
    Else
        Bonferroni = 1
    End If

End Function Bonferroni
```

---

***Module 4***
***Output Procedure***

```
Option Explicit
Option Base 1
Option Private Module
```

*output_results*
*This procedure outputs all the results from the analysis onto a new worksheet.*

```
Sub output_results(datarange As Range)

Dim wkbook As Workbook
Dim wksheet As Worksheet
Dim newwksheet As Worksheet
Dim sheet As Worksheet
Dim counter As Integer
Dim found As Boolean
Dim datastart As Range
Dim counter1 As Integer
```

```
Dim counter2 As Integer

     Insert a new worksheet in front of the data worksheet and name it "Resultsn"
     Access the workbook from datarange
     Set wksheet = datarange.Worksheet
     Set wkbook = wksheet.Parent
     Add the worksheet
     Set newwksheet = wkbook.Worksheets.Add(wksheet)
     Check the results sheet name isn't in use
     counter = 0
     Do
          counter = counter + 1
          found = False
          For Each sheet In wkbook.Worksheets
               If InStr(1, sheet.Name, "Results" & counter) Then
                    found = True
               End If
          Next sheet
     Loop Until found = False
     newwksheet.Name = "Results" & counter

     Format column widths on results worksheet
     With newwksheet
          .Columns(1).ColumnWidth = 25
          .Columns(2).ColumnWidth = 10
     End With

     Set datastart = newwksheet.Range("A1")
     With datastart
          .Value = "Spatial Autocorrelation Analysis Results"
          With .Font
               .Name = "Arial"
               .Size = 12
          End With

          .Offset(1, 0).Value = "Number of localities"
          .Offset(1, 0).Font.Size = 11
          .Offset(1, 1).Value = n
          .Offset(2, 0).Value = "Mean value of z"
          .Offset(2, 0).Font.Size = 11
          .Offset(2, 1).Value = Application.Round(mean_z, 3)
          .Offset(3, 0).Value = "Sample Variance of z"
          .Offset(3, 0).Font.Size = 11
          .Offset(3, 1).Value = Application.Round(variance_z * (n / (n
          - 1)), 3)
          .Offset(4, 0).Value = "Number of distance classes"
          .Offset(4, 0).Font.Size = 11
          .Offset(4, 1).Value = dc_number
          .Offset(5, 0).Value = "Distance class type"
          .Offset(5, 0).Font.Size = 11
          If dc_type = equal_dc Then
               .Offset(5, 1).Value = "Equal"
          Else
          dc_type = log_dc
               .Offset(5, 1).Value = "Logarithmic"
          End If
          .Offset(6, 0).Value = "Distance type"
          .Offset(6, 0).Font.Size = 11
          With ThisWorkbook.DialogSheets("Main Dialog")
               If .OptionButtons("option_d_euclidean") = xlOn Then
                    datastart.Offset(6, 1).Value = "Euclidean"
               ElseIf .OptionButtons("option_d_northing") = xlOn Then
                    datastart.Offset(6, 1).Value = "Northings only"
               Else
               eastings
                    datastart.Offset(6, 1).Value = "Eastings only"
               End If
          End With
```

```
End With datastart

Output distance matrix
Check that it's not too big to fit on to a worksheet
Set datastart = datastart.Offset(8, 0)
If ThisWorkbook.DialogSheets("Main
Dialog").CheckBoxes("cb_fulloutput") = xlOn Then
    If n <= 256 Then
        With datastart
            .Value = "Distance matrix"
            .Font.Size = 11
            For counter1 = 1 To n
                For counter2 = 1 To n
                    .Offset(counter1, counter2 - 1) =
                    distances(counter1, counter2)
                Next counter2
            Next counter1
        End With datastart
        Set datastart = datastart.Offset(n + 2, 0)
    Else
        With datastart
            .Value = "Distance matrix too large to output to
            worksheet"
            .Font.Size = 11
        End With datastart
        Set datastart = datastart.Offset(2, 0)
    End If

    Output similarity matrix
    If n <= 256 Then
        With datastart
            .Value = "Similarity matrix"
            .Font.Size = 11
            For counter1 = 1 To n
                For counter2 = 1 To n
                    .Offset(counter1, counter2 - 1) =
                    c_matrix(counter1, counter2)
                Next counter2
            Next counter1
        End With datastart
        Set datastart = datastart.Offset(n + 2, 0)
    Else
        With datastart
            .Value = "Similarity matrix too large to output to
            worksheet"
            .Font.Size = 11
        End With datastart
        Set datastart = datastart.Offset(2, 0)
    End If
End If full output

Output distance class matrix
With datastart
    .Value = "Distance classes"
    .Font.Size = 11
    .Offset(1, 0).Value = "Number"
    .Offset(1, 1).Value = "Minimum"
    .Offset(1, 2).Value = "Maximum"
    .Offset(1, 3).Value = "Pairs"
    .Offset(0, 4).Value = "Moran's I"
    .Offset(1, 4).Value = "E(I)"
    .Offset(1, 5).Value = "I"
    .Offset(1, 6).Value = "SD(I)"
    .Offset(1, 7).Value = "SIG(I)"
End With datastart

Output the results of the analysis
```

```
    Set datastart = datastart.Offset(2, 0)
    With datastart
        Put in the distance class numbers and maximum distances
        For counter1 = 1 To dc_number
            .Offset(counter1 - 1, 0).Value = counter1
            .Offset(counter1 - 1, 2).Value =
            Application.Round(dc_matrix(counter1), 2)
        Next counter1
        Put in the minimum distances
        .Offset(0, 1).Value = 0
        For counter1 = 1 To dc_number - 1
            .Offset(counter1, 1).Value =
            Application.Round(dc_matrix(counter1), 2)
        Next counter1
        Put in Moran's I and significance
        For counter1 = 1 To dc_number
            .Offset(counter1 - 1, 3).Value =
            Morans_I(counter1).pairs
            If Morans_I(counter1).pairs <> 0 Then
                .Offset(counter1 - 1, 4).Value =
                Application.Round(Morans_I(counter1).exp_I, 3)
                .Offset(counter1 - 1, 5).Value =
                Application.Round(Morans_I(counter1).I, 3)
                .Offset(counter1 - 1, 6).Value =
                Application.Round(Morans_I(counter1).sd_I, 3)
                .Offset(counter1 - 1, 7).Value =
                Application.Round(Morans_I(counter1).sig_I, 3)
            Else
                .Offset(counter1 - 1, 5).Value = "No data"
            End If
        Next counter1

    End With datastart

    Set datastart = datastart.Offset(dc_number + 1, 0)
    With datastart
        .Value = "Overall Correlogram Significance (Bonferroni
        approximation)"
        .Font.Size = 11
        If bon_sig_level <> 1 Then
            .Offset(0, 4).Value = Application.Round(bon_sig_level,
            3)
            .Offset(0, 4).NumberFormat = "0.000"
        Else
            .Offset(0, 4).Value = "not significant at 5% level"
        End If
    End With

    Set datastart = datastart.Offset(2, 0)
    With datastart
        .Value = "End"
        .Font.Size = 11
    End With 'datastart

End Sub output_results
```

---

***Module 5***
***Dialog functions***

```
Option Explicit
Option Base 1
```

*Dialog box functions*
*dc_spinner_Change and dc_editbox_Change link the spinner and editbox values*

```
Sub dc_spinner_Change()
    With ThisWorkbook.DialogSheets("Main Dialog")
        .EditBoxes("dc_editbox").Text = .Spinners("dc_spinner").Value
    End With
End Sub

Sub dc_editbox_Change()
    With ThisWorkbook.DialogSheets("Main Dialog")
        .Spinners("dc_spinner").Value = .EditBoxes("dc_editbox").Text
    End With
End Sub
```

# Appendix 2

## Source code and internal documentation for the vital rates model simulation (Chapter 6)

Note that a 'wrap-around' program was used to run this code the large number of times needed to obtain the results presented in Chapter 6.

---

*RA_MODEL*
*Version 6*
*Last revision 6/7/97*
*Author: Andrew M. Brewer*

*This program calculates the interspecific range abundance relationship between species with different demographic rates along an environmental gradient, based upon the vital rates model (Holt et al. 1997).*

```
Option Explicit
Option Base 1
```

*Global variables. Since these variables are used by many of the subroutines within the program, they have been declared global to simplify the parameter passing.*

```
Const MaxSites  As Integer = 2001
```
*The number of sites in the simulation. This must be an odd number for the allocation of vital rates across sites to work correctly*
```
Dim births() As Double
```
*Site specific birth rates for each species*
```
Dim deaths() As Double
```
*Site specific death rates for each species*
```
Dim r() As Double
```
*Intrinsic rate of increase for each site and species*
```
Dim Meanr() As Double
```
*Mean r for each species*
```
Dim AreaOfOccupancy() As Double
```
*Area of occupancy for each species*
```
Dim ExtentOfOccurence() As Double
```
*Extent of occurrence for each species*
```
Dim MaxSpecies As Integer
```
*Number of species in simulation*
```
Dim Species As Integer
```
*Counter variable for each species*
```
Dim Site As Integer
```
*Counter variable for each site*
```
Dim Datastart As Object
```
*Range variable for data input/output*
```
Dim catchzeros As Boolean
```
*Debug variable for setting negative values for a vital rate to zero*

---

*Main presents the dialog box to the user, allocates memory to the arrays and coordinates the rest of the simulation subroutines.*

```
Sub Main()
With ThisWorkbook
     If .DialogSheets("Main Dialog").Show Then
```
*If the user presses the OK button on the dialog then the simulation starts*
```
          Application.ScreenUpdating = False
          If .DialogSheets("Main Dialog").CheckBoxes("Catch Zeros") =
          xlOn Then
```
*This establishes whether to prevent negative values being used in the simulation (a debug option)*
```
               catchzeros = True
```

```
        Else
            catchzeros = False
        End If
```

*Get the number of species in the simulation from the dialog box and allocate memory for the arrays*

```
        MaxSpecies = .DialogSheets("Main Dialog").EditBoxes("Species Number").Text
        ReDim births(MaxSpecies, MaxSites)
        ReDim deaths(MaxSpecies, MaxSites)
        ReDim r(MaxSpecies, MaxSites)
        ReDim Meanr(MaxSpecies)
        ReDim AreaOfOccupancy(MaxSpecies)
        ReDim ExtentOfOccurence(MaxSpecies)
```

*Initialize the vital rates arrays*

```
        Randomize
        InitializeRates
```

*Calculate r across environmental gradient for each species*

```
        For Species = 1 To MaxSpecies
            Application.StatusBar = "Calculating r for species " & Species
            For Site = 1 To MaxSites
                r(Species, Site) = births(Species, Site) - deaths(Species, Site)
                If r(Species, Site) < 0 And catchzeros = True Then
                    r(Species, Site) = 0
                End If
            Next Site
        Next Species
```

*Calculate mean r and return number of sites occupied*

```
        CalculateMeanrAndRange
        CalculateExtentOfOccurence
```

*Test to see if there is a significant relationship between local abundance and range size*

```
        CalculateRegression
```

*Output the results of the simulation*

```
        OutputData

        Application.StatusBar = ""
    End If

End With

Beep
```

*Wake up the user!*

```
End Sub
```

*Main*

---

*InitializeRates takes values from the main dialog box and initialises the arrays for birth rates, death rates and adds noise to them.*

```
Sub InitializeRates()

Dim BirthSlopes() As Double
```

*Slope of birth rate with respect to sites for each species*

```
Dim DeathSlopes() As Double
```

*Slope of death rate with respect to sites for each species*

```
Dim counter As Integer
Dim BirthSlopeVar As Double
```

*Standard deviation of the variation in birth rates slopes between species ($\sigma_b$ in Chapter 6)*

```
Dim DeathSlopeVar As Double
```

*Standard deviation of the variation in death rates slopes between species ($\sigma_d$ in Chapter 6)*

```
Dim BirthSlope As Double
```

*The mean birth rate slope specified by the user*

```
Dim DeathSlope As Double
```

*The mean death rate slope specified by the user*

```
Dim SigmaBirths As Double
```

*Standard deviation of the noise to be added to site-specific birth rates. Note this is not equivalent to $\sigma_b$ in Chapter 6.*

```
Dim SigmaDeaths As Double
```

*Standard deviation of the noise to be added to site-specific death rates. Note this is not equivalent to $\sigma_d$ in Chapter 6.*

```
Dim minslope As Double
```

*The minimum value of the birth slope to prevent it going past the limits of the environmental gradient.*

```
Dim upperbound As Integer
Dim lowerbound As Integer
```

*upperbound and lowerbound specify the limits for random number generation*

```
Dim middle As Integer
```

*Stores the index for the middle of the arrays containing the vital rates*

*Allocate memory to arrays and find the middle site in the simulation*

```
ReDim BirthSlopes(MaxSpecies)
ReDim DeathSlopes(MaxSpecies)
middle = (MaxSites \ 2) + 1
```

*Integer division to find the middle of the range*

```
Randomize
With ThisWorkbook.DialogSheets("Main Dialog")
```

*Initialise the mean and standard deviation for the birth and death rates slopes specified by the user*

```
    BirthSlope = .EditBoxes("Births Slope").Text
    DeathSlope = .EditBoxes("Deaths Slope").Text
    BirthSlopeVar = .EditBoxes("Birth Slope Variation").Text
    DeathSlopeVar = .EditBoxes("Death Slope Variation").Text

    For Species = 1 To MaxSpecies
```

*Assign a birth rate slope to each species, based on the mean value and variation specified by the user*

```
        Application.StatusBar = "Initializing birth slopes for
        species " & Species
        BirthSlopes(Species) = BirthSlope
        AddNoise BirthSlopes(Species), 0, BirthSlopeVar
        minslope = (100 / ((MaxSites - 1) / 2)) + 0.01
```

*0.01 is added so that the slope cannot go right to the very end of the environmental gradient*

```
        If BirthSlopes(Species) < minslope Then BirthSlopes(Species)
        = minslope
```

*this prevents an inverted births slope and one which goes off the ends of the graph*

```
        If BirthSlopes(Species) > 1 Then BirthSlopes(Species) = 1
```

*prevents numerical errors from cut-off points on x-axis*

*Assign a death rate slope to each species, based on the mean value and variation specified by the user.*

```
        DeathSlopes(Species) = DeathSlope
        AddNoise DeathSlopes(Species), 0, DeathSlopeVar
    Next Species
```

*Set the levels for birth and death rates for each species and at each site based on the option chosen by the user (births and deaths variable, just births, just deaths, neither)*

```
If .OptionButtons("Both Variable").Value = xlOn Then
    For Species = 1 To MaxSpecies
        upperbound = 100
        lowerbound = 25
```

*set middle value for birth rate*

```
      births(Species, middle) = Int((upperbound - lowerbound + 1) *
      Rnd + lowerbound)
      fill in the rest of the array
      For counter = 1 To (middle - 1)
         births(Species, middle + counter) = births(Species,
         middle) - (counter * BirthSlopes(Species))
         births(Species, middle - counter) = births(Species,
         middle + counter)
      Next counter
      upperbound = births(Species, middle) - 1
      maximum death rate must not exceed maximum birth rate
      lowerbound = 1
      set middle value for death rate
      deaths(Species, middle) = Int((upperbound - lowerbound + 1) *
      Rnd + lowerbound)
      fill in the rest of the array
      For counter = 1 To (middle - 1)
         deaths(Species, middle + counter) = deaths(Species,
         middle) - (counter * DeathSlopes(Species))
         deaths(Species, middle - counter) = deaths(Species,
         middle + counter)
      Next counter
   Next Species

ElseIf .OptionButtons("Variable Births").Value = xlOn Then
   For Species = 1 To MaxSpecies
      Application.StatusBar = "Calculating vital rates for species
      " & Species
      upperbound = 100
      lowerbound = 25
      set middle value for birth rate
      births(Species, middle) = Int((upperbound - lowerbound + 1) *
      Rnd + lowerbound)
      fill in the rest of the array
      For counter = 1 To (middle - 1)
         births(Species, middle + counter) = births(Species,
         middle) - (counter * BirthSlopes(Species))
         births(Species, middle - counter) = births(Species,
         middle + counter)
      Next counter
      set middle value for death rate (constant)
      deaths(Species, middle) = lowerbound - 1
      fill in the rest of the array
      For counter = 1 To (middle - 1)
         deaths(Species, middle + counter) = deaths(Species,
         middle) - (counter * DeathSlopes(Species))
         deaths(Species, middle - counter) = deaths(Species,
         middle + counter)
      Next counter
   Next Species

ElseIf .OptionButtons("Variable deaths").Value = xlOn Then
   For Species = 1 To MaxSpecies
      Application.StatusBar = "Calculating vital rates for species
      " & Species
      upperbound = 100
      lowerbound = 25
      set middle value for birth rate (constant)
      births(Species, middle) = upperbound 'Int((upperbound -
      lowerbound + 1) * Rnd + lowerbound)
      fill in the rest of the array
      For counter = 1 To (middle - 1)
         births(Species, middle + counter) = births(Species,
         middle) - (counter * BirthSlopes(Species))
         births(Species, middle - counter) = births(Species,
         middle + counter)
      Next counter
```

```
          upperbound = upperbound - 1
          maximum death rate must not exceed maximum birth rate
          lowerbound = 1
          set middle value for death rate
          deaths(Species, middle) = Int((upperbound - lowerbound + 1) *
          Rnd + lowerbound)
          fill in the rest of the array
          For counter = 1 To (middle - 1)
               deaths(Species, middle + counter) = deaths(Species,
               middle) - (counter * DeathSlopes(Species))
               deaths(Species, middle - counter) = deaths(Species,
               middle + counter)
     Next counter
Next Species

ElseIf .OptionButtons("Neither Variable").Value = xlOn Then

     For Species = 1 To MaxSpecies
          Application.StatusBar = "Calculating vital rates for species
          " & Species
          upperbound = 100
          lowerbound = 25
          set middle value for birth rate (constant)
          births(Species, middle) = upperbound
          fill in the rest of the array
          For counter = 1 To (middle - 1)
               births(Species, middle + counter) = births(Species,
               middle) - (counter * BirthSlopes(Species))
               births(Species, middle - counter) = births(Species,
               middle + counter)
          Next counter
          set middle value for death rate (constant)
          deaths(Species, middle) = lowerbound
          fill in the rest of the array
          For counter = 1 To (middle - 1)
               deaths(Species, middle + counter) = deaths(Species,
               middle) - (counter * DeathSlopes(Species))
               deaths(Species, middle - counter) = deaths(Species,
               middle + counter)
          Next counter
     Next Species

End If OptionButtons

Add noise to birth rates
SigmaBirths = .EditBoxes("Birth Noise").Text
If SigmaBirths > 0 Then
     For Species = 1 To MaxSpecies
          Application.StatusBar = "Adding noise to birth rates for
          species " & Species
          For Site = 1 To MaxSites
               AddNoise births(Species, Site), 0, SigmaBirths
               If catchzeros = True And births(Species, Site) < 0 Then
               births(Species, Site) = 0
          End If
          Next Site
     Next Species
End If

Add noise to death rates
SigmaDeaths = .EditBoxes("Death Noise").Text
If SigmaDeaths > 0 Then
     For Species = 1 To MaxSpecies
          Application.StatusBar = "Adding noise to death rates for
          species " & Species
          For Site = 1 To MaxSites
               AddNoise deaths(Species, Site), 0, SigmaDeaths
```

```
               If catchzeros = True And deaths(Species, Site) < 0 Then
               deaths(Species, Site) = 0
               End If
          Next Site
     Next Species
End If

End With The dialog sheet

End Sub InitializeRates
```

---

*AddNoise adds a random number sampled from a normal distribution of mean, mu, and standard deviation, sigma, to the argument data. This is based on a standard algorithm.*

```
Sub AddNoise(data As Double, mu As Double, Sigma As Double)

Dim U1 As Double
Dim U2 As Double
Dim V1 As Double
Dim V2 As Double
Dim S As Double
Dim X1 As Double
Dim X2 As Double
Dim Y1 As Double
Dim Y2 As Double

Do
     U1 = Rnd
     U2 = Rnd
     V1 = 2 * U1 - 1
     V2 = 2 * U2 - 1
     S = V1 ^ 2 + V2 ^ 2
Loop While S >= 1

X1 = V1 * ((-2 * Application.Ln(S)) / S) ^ 0.5
X2 = V2 * ((-2 * Application.Ln(S)) / S) ^ 0.5
Y1 = mu + Sigma * X1
Y2 = mu + Sigma * X2

data = data + Y1

End Sub AddNoise
```

---

*CalculateMeanrAndRange calculates the mean value of r, and the area of occupancy for each species. The area of occupancy is the number of occupied sites. The mean value of r is the sum of all the r values (births - deaths) divided by the area of occupancy.*

```
Sub CalculateMeanrAndRange()

For Species = 1 To MaxSpecies
     Application.StatusBar = "Calculating mean r and area of occupancy
     for species " & Species
     Meanr(Species) = 0
     AreaOfOccupancy(Species) = 0
     For Site = 1 To MaxSites
          If r(Species, Site) > 0 Then
               AreaOfOccupancy(Species) = AreaOfOccupancy(Species) + 1
               Meanr(Species) = Meanr(Species) + r(Species, Site)
          End If
     Next Site
     If AreaOfOccupancy(Species) = 0 Then
          Meanr(Species) = 0
     Else
          Meanr(Species) = Meanr(Species) / AreaOfOccupancy(Species)
     End If
```

```
Next Species

End Sub CalculateMeanrAndRange
```

---

*CalculateExtentOfOccurence calculates the extent of occurence for each species. This is done by finding the lowest and highest points on the gradient where r is greater than zero, i.e. where the species is present, and measuring the distance in-between.*

```
Sub CalculateExtentOfOccurence()

Dim StartOfRange As Integer
Dim EndOfRange As Integer

For Species = 1 To MaxSpecies
     Application.StatusBar = "Calculating extent of occurrence for
species " & Species
     If AreaOfOccupancy(Species) = 0 Then
          ExtentOfOccurence(Species) = 0
     Else
          StartOfRange = 1
          EndOfRange = MaxSites
          Do While r(Species, StartOfRange) <= 0
               StartOfRange = StartOfRange + 1
          Loop
          Do While r(Species, EndOfRange) <= 0
               EndOfRange = EndOfRange - 1
          Loop
          ExtentOfOccurence(Species) = EndOfRange - StartOfRange + 1
     End If
Next Species

End Sub CalculateExtentOfOccurence
```

---

*CalculateRegression tests for a significant relationship between abundance and range size and outputs the results to the worksheet "Results". It uses four worksheet functions, INDEX, ROUND, LINEST and FDIST, to perform the calculations. Note that if there is no noise added to the simulation, the F-value for the regression will be undefined (but the regression will obviously be significant).*

```
Sub CalculateRegression()

Dim counter As Integer
Dim f1 As Double
Dim r_square1 As Double
Dim df1 As Integer
Dim prob1 As Double
Dim coeff1 As Double
Dim f2 As Double
Dim r_square2 As Double
Dim df2 As Integer
Dim prob2 As Double
Dim coeff2 As Double
Dim msg As String

Dim CalculateF As Boolean

Application.StatusBar = "Calculating regression ..."
With ThisWorkbook.DialogSheets("Main Dialog")
     If .EditBoxes("Death Noise").Text > 0 Or .EditBoxes("Birth
     Noise").Text > 0 Or .EditBoxes("Birth Slope Variation").Text > 0
     Or .EditBoxes("Death Slope Variation").Text > 0 Then
          CalculateF = True
     End If
End With

If CalculateF Then
```

```
    calculate statistics for area of occupancy
    r_square1 = Application.Index(Application.LinEst(AreaOfOccupancy,
    Meanr, True, True), 3, 1)
    f1 = Application.Index(Application.LinEst(AreaOfOccupancy, Meanr,
    True, True), 4, 1)
    df1 = Application.Index(Application.LinEst(AreaOfOccupancy, Meanr,
    True, True), 4, 2)
    prob1 = Application.FDist(f1, 1, df1)
Else
    r_square1 = 1
    prob1 = 0
End If

coeff1 = Application.Index(Application.LinEst(AreaOfOccupancy, Meanr,
True, True), 1, 1)

If CalculateF Then
    calculate statistics for extent of occurence
    r_square2 =
    Application.Index(Application.LinEst(ExtentOfOccurence, Meanr,
    True, True), 3, 1)
    f2 = Application.Index(Application.LinEst(ExtentOfOccurence,
    Meanr, True, True), 4, 1)
    df2 = Application.Index(Application.LinEst(ExtentOfOccurence,
    Meanr, True, True), 4, 2)
    prob2 = Application.FDist(f2, 1, df2)
Else
    r_square2 = 1
    prob2 = 0
End If

coeff2 = Application.Index(Application.LinEst(ExtentOfOccurence,
Meanr, True, True), 1, 1)

With ThisWorkbook.Worksheets("Results")
    .Cells.ClearContents
    Set Datastart = .Range("F1")
    Datastart.Offset(0, 0).Value = "Regression results"
    Datastart.Offset(1, 0).Value = "Regresson coefficient"
    Datastart.Offset(2, 0).Value = "R-square"
    Datastart.Offset(3, 0).Value = "Probability"
    Datastart.Offset(0, 1).Value = "Area of occupancy"
    Datastart.Offset(1, 1).Value = Application.Round(coeff1, 3)
    Datastart.Offset(2, 1).Value = Application.Round(r_square1, 3)
    Datastart.Offset(3, 1).Value = Application.Round(prob1, 3)
    Datastart.Offset(0, 2).Value = "Extent of occurrence"
    Datastart.Offset(1, 2).Value = Application.Round(coeff2, 3)
    Datastart.Offset(2, 2).Value = Application.Round(r_square2, 3)
    Datastart.Offset(3, 2).Value = Application.Round(prob2, 3)
End With

End Sub CalculateRegression
```

---

*OutputData sends the vital rates data for each species and site to the worksheet "Scratch" which displays this data as charts to give a visual impression of the whole data set of the simulation. In addition the simulated values for each species' mean r, extent of occurrence and area of occupancy are output to the worksheet "Results".*

```
Sub OutputData()

Application.StatusBar = "Outputting results ... "
With ThisWorkbook
    With .Worksheets("Scratch")
        .Cells.ClearContents
        Set Datastart = .Range("A1")
        If ThisWorkbook.DialogSheets("Main
        Dialog").CheckBoxes("Output all data") = xlOn Then
```

```
            Datastart.Offset(0, 0).Value = "Births"
            For Species = 1 To MaxSpecies
                For Site = 1 To MaxSites
                    Datastart.Offset(Site, Species).Value =
                    births(Species, Site)
                Next Site
            Next Species

            Set Datastart = Datastart.Offset(0, MaxSpecies + 1)
            Datastart.Offset(0, 0).Value = "Deaths"
            For Species = 1 To MaxSpecies
                For Site = 1 To MaxSites
                    Datastart.Offset(Site, Species).Value =
                    deaths(Species, Site)
                Next Site
            Next Species

            Set Datastart = Datastart.Offset(0, MaxSpecies + 1)
            Datastart.Offset(0, 0).Value = "r"
            For Species = 1 To MaxSpecies
                For Site = 1 To MaxSites
                    Datastart.Offset(Site, Species).Value =
                    r(Species, Site)
                Next Site
            Next Species

            Set Datastart = Datastart.Offset(0, MaxSpecies + 1)
        End If
    End With Scratch worksheet

    With .Worksheets("Results")

        Set Datastart = .Range("A1")
        Datastart.Offset(0, 1).Value = "Mean r"
        For Species = 1 To MaxSpecies
            Datastart.Offset(Species, 1).Value = Meanr(Species)
        Next Species

        Set Datastart = Datastart.Offset(0, 1)
        Datastart.Offset(0, 1).Value = "Area Of Occupancy"
        For Species = 1 To MaxSpecies
            Datastart.Offset(Species, 1).Value =
            AreaOfOccupancy(Species)
        Next Species

        Set Datastart = Datastart.Offset(0, 1)
        Datastart.Offset(0, 1).Value = "Extent Of Occurence"
        For Species = 1 To MaxSpecies
            Datastart.Offset(Species, 1).Value =
            ExtentOfOccurence(Species)
        Next Species

    End With Results worksheet

End With ThisWorkbook

End Sub OutputData
```