

TABLE OF CONTENTS

	Page
1 Introduction	1
1.1 Introduction	1
1.1.1 Communities of single and multiple interaction types	3
1.1.2 Spatially explicit model and metrics	6
1.1.3 Modelling Habitat Loss	6
2 Habitat loss: Modelling approach	7
2.1 Agent based simulation	7
2.1.1 The interaction network	8
2.1.2 Model specification	11
2.1.3 Modelling habitat loss	14
2.1.4 Implementation	16
2.2 Dynamics of the model	16
2.2.1 Transience	17
2.2.2 Long term distribution	17
2.2.3 Diffusion behaviour	17
2.3 Ecological metrics and analysis methods	17
2.3.1 Biodiversity metrics	17
2.3.2 Stability metrics	17
2.3.3 Network metrics	17
2.3.4 Spatial metrics	17
2.3.5 Interaction strength metrics	17
3 Habitat loss with high immigration	19
3.1 Response of ecological metrics	19
3.2 More preliminary results	24
3.3 Further development and work to do	24
3.3.1 Work to do	24
3.3.2 Further development	25

TABLE OF CONTENTS

4 Varying Immigration rate	29
4.1 Assumptions of what has gone before	29
4.2 Literature review for immigration	30
4.3 Motivation	32
4.4 Exploration of parameter space: habitat loss and immigration	34
4.4.1 Summary heat-maps	38
4.4.2 Example dynamics	42
4.4.3 Relative abundances and abundance distributions	45
4.4.4 Rank abundance distributions	45
4.5 Points for discussion (Rough Notes)	49
4.6 Habitat loss with low immigration	50
4.7 Questions for Alan or Daniel	50
4.8 RADS with averaging	53
4.9 New results: without vegetarian predators	55
5 Stationarity	59
5.1 Motivation	59
5.2 Second-order stationarity	60
5.2.1 Tests for stationarity	61
5.2.2 Characterising the tests	61
5.2.3 Stationarity results	69
5.2.4 Discussion ('skeleton')	72
5.3 Reliability of results	73
5.3.1 Abundances	73
5.3.2 Variability: CoV	73
5.3.3 Repeatability	73
5.4 Determinism and Chaos	73
5.5 Discussion	76
6 Persistence in closed communities	81
6.1 Motivation	81
6.2 Stability in ecology	82
6.2.1 Dynamical systems and stability	83
6.2.2 Network structure and stability	83
6.2.3 Spatial stability	84
6.3 Model parameters versus persistence	85
6.3.1 Mutualistic to antagonistic interaction (MAI) ratio	85
6.3.2 Reproduction rate (RR)	88
6.3.3 Landscape size	90

TABLE OF CONTENTS

6.3.4	Number of initial species	91
6.4	Why are some communities more stable than others?	93
7	Inferring species interactions	103
7.1	Motivation	103
7.2	Methodology	104
7.2.1	Interaction strength	105
7.2.2	Population models	106
7.2.3	Simulation procedure	109
7.2.4	Numerical estimation method	110
7.2.5	Examples	113
7.3	Results	116
7.3.1	Linear model	116
7.3.2	Holling model	119
7.3.3	Range sampling	119
7.4	TEMP : other results	119
7.5	Application to IBM (optional)	120
7.5.1	Testing functionl response	120
7.5.2	Extend methodology (3 and 4 species)	120
7.5.3	Results	120
7.6	Discussion	120
Bibliography		125

INTRODUCTION

1.1 Introduction

This project focuses on the impact of habitat destruction on communities of species. A habitat may be defined as the environment containing an organism, or collection of organisms. It has both biotic and abiotic components. Therefore habitats are constantly changing due to ongoing environmental processes. These changes may make the habitat more or less hospitable to different organisms, generating emergent effects at the species and community levels. Human activity in particular creates pronounced and significant changes in habitat. There is good evidence [48] that anthropogenic climate change has affected living systems by changing regional habitat suitability. An example of this is the northward shift in butterfly species ranges attributed to rising temperatures [47]. Other activities such as agriculture, deforestation and urbanisation interfere directly with physical habitat components and with local flora. This alters the type of species and the community that can be supported [7, 36]. Globally the scale of these man-made effects is huge. Various studies have suggested that habitat modification is the leading cause of global species extinctions [17, 63]. Therefore an understanding of how ecological communities respond to changes in habitat is essential in order to mediate the destructive effects of human activity, and to create beneficial conservation, land management and restoration strategies. The subject has received much attention in the ecological literature, and this project is a continuation of that dialogue.

The destruction of habitats due to human activity has also received much attention in the media. This has done a lot to raise public awareness, and to fuel a growing number of campaign groups, charities and conservation organisations. In most cases the focus is on *single species effects*, especially on those threatened with extinction. The most notorious example of this may be the polar bear as the media face of global warming (see figure 1.1). Similarly the habitat loss



Figure 1.1: Stranded polar bears on Cross Island outside Prudhoe Bay, Alaska. The plight of the polar bear has received much attention in the media. The habitat loss it suffers from is very visible. However the focus of conservation strategies must be on the ecological communities, of which it is one member species. (Source: www.greenpeace.org.uk)

literature has largely focused on the loss of species [17, 62], and has reinforced the notion of *species richness*¹ as a measure of biodiversity and ecosystem health. This is perhaps because species level effects are the most visible results of ecosystem damage, and the easiest to study empirically. However they are symptomatic of underlying system processes. At least since Darwin's marvel at the complexity of the "Tangled Bank" [11] ecologists have understood that species exist in highly interdependent communities. Therefore the ecological impacts of habitat destruction, and other human activities, must be approached from a systems perspective.

In community ecology the system of study is the ecological community - a local collection of co-existing species. The focus is on the structure, patterns and processes within the community. A key aspect of this is the pattern of *interactions between species*, which underlies many of the processes that shape the community (for more detail refer Chapter 2). Recently the habitat loss literature has begun to move away from species level effects, towards community wide effects and especially inter-specific interactions [64]. This has been facilitated by the wider availability of ecological network data, improved methods for data collection, and the ability to simulate large

¹Simply defined as the number of different species present in a community.

ecological networks and communities. Advances in ecological network theory have also provided many new metrics for community stability, biodiversity and for analysis of network structure (section 2.3). Our approach to the study of habitat loss is situated in this context.

There is now a growing consensus that ecological interactions are the key to understanding the effects of habitat loss on ecological communities [23, 26, 42]. In addition to the loss species, it has long been known that habitat loss also leads to the important loss of inter-specific interactions. As Janzen remarked [30] in 1974: “what escapes the eye, however, is a much more insidious kind of extinction: the extinction of ecological interactions”. It has since been demonstrated that ecosystems experiencing habitat alteration often suffer loss of interactions *before* loss of species [2, 20, 64]. This can result in detectable changes in community structure, without any detectable change in species richness [63]. These structural changes have consequences for community stability, robustness and population dynamics. A significant part of the ongoing challenge is to identify meaningful measures for the structural (network) changes, and to generalise the ways in which they impact on the community. The bulk of the recent literature supports the belief of Valiente et al. [64] in “the importance of focusing on species interactions as the major biodiversity component on which the ‘health’ of ecosystems depends.”

1.1.1 Communities of single and multiple interaction types

In the habitat loss literature most studies have looked at communities with a single type of interaction. The same has been true for network ecology in general, with the bulk of the literature focused on either antagonistic or mutualistic networks. In these networks a node represents a species, and a directed link represents a certain type of interaction (for example predation). Such networks represent the interaction structure of an idealised and closed community. For example it is common to study mutualistic communities, such as plants and their pollinators, in isolation. This is represented as a bipartite network of plant and pollinator species, with mutualistic interactions between them. Both empirical and *in silico* studies have derived some apparently general results on the response of such single-interaction communities to habitat loss. We discuss some of these findings here. However in nature a single-interaction community is a subset of a larger group of species with multiple types of interaction (predation, mutualism, competition, parasitism). There has been a recent move towards studies of communities with multiple types of interaction [32], which are less simplistic models of natural systems. These hybrid communities are represented as networks with more than one type of link. We also discuss this body of work, some of which challenges previous finding based on single-interaction communities.

Perhaps most the general result, already discussed, is that habitat destruction leads to a loss of inter-specific interactions. This may be accompanied by lower interaction frequencies, changes in interaction strength, reduced connectivity, or other structural changes in the network due to rewiring. Tylianakis et al. [63] showed that empirical antagonistic communities (host-parasitoid) responded to habitat degradation with reduced evenness in interaction frequencies. This means

that certain interactions became relatively more frequent, so that energy flow through the community became concentrated along certain pathways. Also, importantly, the quantitative changes in network structure that they observed were not detectable by equivalent qualitative metrics. Neither were conventional diversity metrics, based on species abundance or richness, able to distinguish between habitats at different levels of degradation. Similarly Albrecht et al. [2] showed that insect food webs in a grassland system lost interaction diversity faster than species diversity, when subjected to habitat alteration. This suggests a biodiversity reduction in the interaction structure that is not measurable by metrics based on species abundance. Both of these examples highlight the sensitivity of results to the metrics used, when studying community response to habitat loss. Hence the large suite of metrics introduced and discussed in section 2.3.

An issue of particular interest is community stability, its response to habitat loss and its relationship to network structure. Mutualistic networks tend to have a highly nested structure and low modularity [4]. These properties are believed to improve the stability of the community [61]. It has been shown that habitat destruction can push mutualistic networks towards higher modularity, higher connectivity, and lower nestedness, thereby reducing stability [26, 58]. Conversely antagonistic networks tend to be modular in structure, which is believed to promote stability and robustness in these communities [61]. Habitat loss has been shown to destabilise antagonistic communities by lowering modularity and increasing interaction strengths [26]. Generally the literature suggests, as expected, that habitat loss reduces community stability, irrespective of the interaction type. However the underlying changes driving this loss in stability appears to differ between mutualistic and antagonistic communities. It should also be noted here that the definition and measurement of stability is non-trivial. Lurgi et al. [39] have shown that certain stability metrics may respond differently to a changing control variable, meaning that a combined, or multi-stability approach is required.

The above examples represent attempts to understand the structural changes that occur due to habitat loss, prior to the occurrence of species extinctions. From a conservation perspective this highlights the importance of targeting inter-specific interactions and the maintenance of network structure and function, rather than focusing on species level effects [42]. Fortuna and Bascompte [19] have demonstrated that real-world networks have better persistence against habitat loss than random networks assembled using null-models. This suggests that artificially managed ecosystems may be more vulnerable to perturbations than their ‘wild-type’ equivalents, unless careful attention is paid to those properties that promote stability and robustness. In food webs there appear to be certain simple properties that mediate the impacts of habitat destruction [41]. For example omnivory is shown to increase extinction thresholds, as is a reduction in top-down control by predators. However these numerical results are for small model networks and remain to be demonstrated empirically.

Recently ecologists have realised the importance of studying ecological networks that contain multiple types of inter-specific interaction [18, 32, 43]. It is known that mutualistic communities

have knock on effects on food webs, and vice versa. Indeed certain species are simultaneously involved in more than one type of network or community. A powerful example of this phenomenon was demonstrated empirically by Knight et al. [34]. They showed the presence of a trophic cascade, crossing ecosystem and habitat boundaries, by which freshwater fish were able to facilitate terrestrial plant reproduction. The inclusion of such indirect and cascading effects is one of the many strengths of the network paradigm in ecology. However this study highlights the limitations of focusing on localised community subsets and single-interaction types.

A large scale study by Pocock et al. [49] was one of the first to combine networks of different types into a network of ecological networks. They used empirical networks constructed over different habitats on a farm, to construct a whole farm network. This included host-parasitoid, seed-dispersal, plant-pollinator and predator-prey networks. Using quantitative robustness analysis (section 2.3), they were able to identify keystone plant species which generated significant cascading effects across networks, and also determined the most fragile components of the meta-network. This type of integrated analysis has different implications for conservation and restoration than an approach which looks at the individual networks in isolation.

The integration of multiple interaction types has begun to shed new light on the stability of ecological communities. This is because the conventional understanding is based on studies of communities with single-interaction types. In general complex antagonistic networks with strong interactions are thought to be unstable [46]. This presents a problem for ecological theory since natural food webs, which are inherently complex, appear to be stable. The problem may lie in the fact that antagonistic networks have been studied in isolation. It has been shown theoretically that introducing mutualistic interactions into the network can be stabilising [39, 44]. Specifically Lurgi et al. [39] propose that increasing the proportion of mutualistic interactions at the base of a food web reduces the overall strength of species interactions. They found that this improved the stability of their model communities, according to a spatial aggregation metric (section 2.3).

Recently Sauve et al. [55] have brought into question the established wisdom on the relationship between network structure and stability. As discussed previously, the structural properties believed to promote stability differ between antagonistic and mutualistic communities. High modularity and high nestedness are thought to promote stability in antagonistic and mutualistic networks respectively. However Sauve's work suggests that, for a combined network of mutualisms and antagonisms, modularity and nestedness do not strongly affect stability. The results of Lurgi et al. also support this finding [39]. Therefore new metrics, accounting for diversity in interaction type, may be required in order to understand community structure and stability in hybrid networks².

Since hybrid networks of multiple interaction type are relatively new, there are few studies relating them to habitat loss. One study, by Evans et al. [14], uses the same empirical network of networks as [49]. They employed a robustness algorithm to determine how vulnerable the

²See suggestions in the text of [55] and talk to Alix about possibly including these in our analysis?

hybrid network is to the loss of different habitats from the farm³. Aside from this study there is a lack of empirical and theoretical results on the response of hybrid networks to habitat loss. This project aims to make a contribution towards this area. We will extend on the work of Lurgi et al. [39] to simulate multi-trophic communities with mutualistic and antagonistic interactions. By investigating the response of these communities to simulated habitat destruction we will be generating novel results and predictions which can be tested empirically in the future. To do this we will employ a range of metrics to quantify structural changes and community stability. We will focus on the regime before species are lost from the community, with an interest in the underlying changes that occur as a result of habitat destruction.

1.1.2 Spatially explicit model and metrics

Another novel aspect of this work is the spatially explicit modelling approach... And some of the spatial analysis employed...

[57] - spatially explicit analysis.

[20] mutualistic interactions decrease non-linearly. Connectance increases? Abrupt change in number of interactions, spatial skewness in number of interactions.

[31] - quantitative food web metrics did not vary between fragmented habitat patches in different landscape contexts.

[46] - interaction strengths is focus, but also spatial stability. c.f. a,b,g stability and Lurgi et al.

1.1.3 Modelling Habitat Loss

Habitat loss has been modelled in various ways..Spatial auto-correlation..how does our approach fit in with the literature..

[16] - controlled habitat destruction, large empirical project

³Interestingly they reported that two of the most important habitats, relative to their sizes, were hedgerow and wasteland.

HABITAT LOSS: MODELLING APPROACH

2.1 Agent based simulation

We study the effects of habitat loss on ecological communities using a spatially explicit agent-based model. This simulation model was first published by Lurgi et al in [39] (section 2.1.2). The landscape consists of a homogeneous 2-dimensional grid (200×200 cells) on which individuals of 60 species move around and interact subject to bio-energetic constraints. Local rules define dispersal, demographic processes and interaction behaviour of the individuals. The potential for interaction between two individuals is governed by an underlying species interaction network, which is generated using the niche model [67] (section 2.1.1.1. (GIVE A BIT MORE ABOUT NICHE HERE))

Unlike most previous *in silico* studies, the model includes both trophic and mutualistic interactions. Species belong to four trophic levels. The niche model generates a trophic interaction network. Then a fraction of the links between species in the first two trophic levels are changed to define mutualisms ($-+ \rightarrow ++$). The fraction of links switched is called the mutualistic vs. antagonistic interaction (MAI) ratio.

To simulate habitat loss, a fraction of the grid cells are made inhospitable to all species. We compare two algorithms for choosing which cells to destroy: 1) random destruction and 2) contiguous destruction. For random destruction grid cells are destroyed uniformly at random, up to the desired fraction of the total landscape. For contiguous destruction a seed cell is chosen uniformly at random, then destruction spreads radially in all directions from this point.

We simulated communities with MAI values of $0, 0.1, 0.2, \dots, 1.0$, with habitat loss (HL) percentages of $0, 10, 20, \dots, 90$. For each combination of MAI and HL values, 25 replicates were simulated with different interaction networks (same species richness, same connectance).

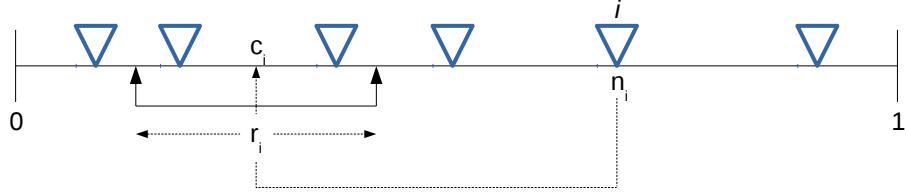


Figure 2.1: A representation of 1-dimensional *niche space* as visualised in the original publication [67], for number of species $S = 6$. The blue triangles represent the placement of species in niche space. The niche value of species i is given by n_i . The width and centre of the feeding range for species i are denoted by r_i and c_i respectively. Species i consumes all species whose niche values fall within the feeding range.

2.1.1 The interaction network

An underlying interaction network defines which direct interactions between species are allowed. This network contains two types of link: *antagonistic* (predator-prey) and *mutualistic* (plant-pollinator). To construct this network a food web, containing only antagonistic links, is generated using the niche model (section 2.1.1.1). This network consists of 60 species belonging to 4 trophic levels, with links that define the feeding relationships between them. Species in the basal trophic level represent plants and those in the trophic level above represent herbivores. To introduce mutualism a fraction of the herbivorous links are replaced by mutualistic links (section 2.1.1.2).

2.1.1.1 The niche model

We use the niche model (NM) of Williams & Martinez [67] is used to generate food webs. This simple model has been shown to produce network structures that closely resemble empirically derived food webs¹, and has become a standard tool for creation of model food webs [13, 59, 67]. The model has two parameters: the number of species S , and the desired connectance C . The model output is an adjacency matrix \mathbf{a} for which the element $a_{ij} = 1$ implies that species i consumes species j , and $a_{ij} = 0$ implies the absence of an interaction. Connectance is defined as the proportion of the maximum possible number of links that are realised i.e. $C = L/S^2$, where L is the number of links in the network.

Figure 2.1 illustrates the ideas of niche space, niche value n_i for a particular species i , its feeding range r_i . Niche space is the 1-dimensional range of real numbers $[0, 1]$. Each of the S

¹FIND MORE REFS HERE.

species is assigned a niche value n_i , drawn uniformly at random from the niche space. It is then assigned a feeding range with a central value c_i and a width r_i . Species i consumes all species, including itself, whose niche values fall within its feeding range.

To determine the width of the feeding range, a beta function with expectation $2C$ is used to draw a number from the range $[0, 1]$. This number is then multiplied by n_i to give the chosen value of r_i . Since $n_i \sim U(0, 1)$, we know that the expectation value $E(n_i) = 0.5$, and so $E(r_i) = C$. Therefore on average a species consumes a fraction C of the total number of species, resulting in a network with close to the desired connectance.

A beta function has two parameters: $\alpha, \beta \in \mathbb{R}^+$ [REF]. The choice of $\alpha = 1$ simplifies the probability density function to

$$f(x; 1, \beta) = \begin{cases} \beta(1-x)^{\beta-1} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

The cumulative distribution function is derived by:

$$\begin{aligned} P(x) &= \int_0^x \beta(1-x')^{\beta-1} dx' \\ &= 1 - (1-x)^\beta. \end{aligned}$$

Therefore, by choosing a probability value y uniformly at random from the interval $[0, 1]$, we can draw an x value from our beta distribution:

$$\begin{aligned} y &= 1 - (1-x)^\beta, \quad \text{such that} \\ x &= 1 - (1-y)^{1/\beta}. \end{aligned}$$

The expectation value of this beta distribution is given by $E(x) = \frac{1}{1+\beta}$, therefore we choose

$$\beta = \frac{1}{2C} - 1$$

to give the desired expectation of $E(x) = 2C$.

Once the width r_i has been chosen, the feeding range is placed in niche space by randomly drawing the range centre c_i from the interval $[r_i/2, n_i]$. Therefore cannibalism and looping are possible because up to half of the feeding range may contain niche values $\geq n_i$. In some cases the generated network may not be connected (i.e. contains one or more disconnected components), or two species may be trophically identical. In these cases the guilty species are deleted and replaced until the network is connected and without identical species. Also the species with the smallest niche value is given $r_i = 0$, such that there is at least one basal species (i.e. species with no prey).

INSERT NICE PICTURE OF A NETWORK!

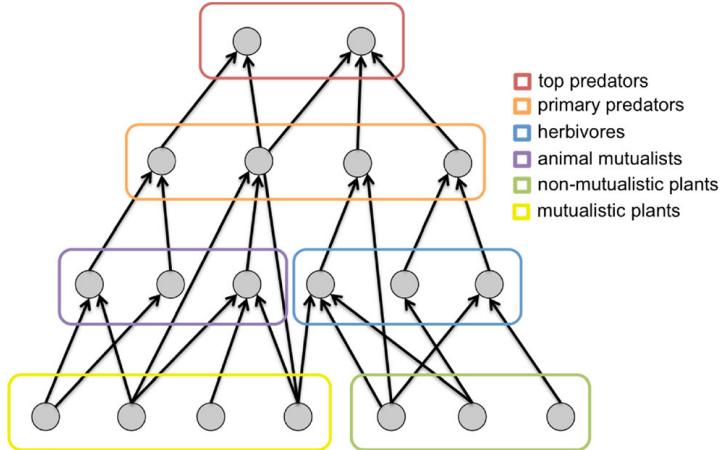


Figure 2.2: Schematic of an underlying interaction network (reproduced from [39]). Nodes correspond to species, and arrows to trophic links (antagonistic or mutualistic), from resource to consumer. The six functional groups of species are colour coded, and named in the legend. In this case there are twelve links between the first two trophic levels, six of these have been replaced by mutualistic links giving a MAI ratio of 0.5. Mutualistic plants and animal mutualists are defined by any species that has at least one mutualistic links. However both these groups of species may also have antagonistic links.

2.1.1.2 Link Replacement

Having generated a food web with only antagonistic links, we now introduce mutualism. The mutualistic interactions are trophic, as with antagonisms, since there is an energy flow from resource to consumer. For example pollinators receive nectar from flowering plants. However in a mutualistic interaction there is a benefit for both parties. In this example flowering plants are pollinated and can reproduce. In the simulation model the plants receive better dispersion abilities as a result of mutualisms (section 2.1.2.1). We impose the constraint that mutualisms can only exist between species of the first two trophic levels: plants and herbivores. Some of the antagonistic links between the first two trophic levels are replaced by mutualistic links. This changes the rules of interaction between individuals of these species in the cellular automata model (section 2.1.2.1). The fraction of these links switched is defined as the mutualistic vs. antagonistic interaction (MAI) ratio. Figure 2.2 is a schematic of a possible interaction network generated by this procedure, for a nineteen species community. In this case there are twelve links between the first two trophic levels, and six of these have been replaced by mutualistic links. The other six links remain antagonistic. Since half of the basal links have been replaced, the MAI ratio for this community is 0.5.

The result of link replacement is a hybrid network that defines two types of interaction between species. We can define two functional groups in each of the first two trophic levels. In the first trophic level *non-mutualistic plants* are basal species which do not have any mutualistic links. This group represents wind-dispersed plants which only have antagonistic interactions with

the trophic level above. *Mutualistic plants* are any basal species with at least one mutualistic link. This group are dispersed by species from the second trophic level via their mutualistic interactions and can no longer be wind dispersed. They may also be predated upon by herbivores, if they have such links. Similarly *herbivores* are members of the second trophic level which only predate on basal species, whereas *animal mutualists* may either predate or engage in mutualisms. See figure 2.2 for a visualisation of these groups.

For the simulations we generated networks with eleven different MAI ratios ($[0, 0.1, 0.2, \dots, 1.0]$). This is in accordance with the previous study [39], and allows us to look at how communities with different MAI ratios respond to habitat loss.

2.1.2 Model specification

We use the model of Lurgi, Montoya & Montoya [39] as the basis of our simulation model. It is a cellular automaton (CA) in which individuals belonging to different species move around, reproduce, die and interact. These actions are subject to bioenergetic constraint and the rules governing them are detailed in section 2.1.2.1 below. The CA landscape is a homogeneous 2D square lattice with toroidal boundaries. Each cell can contain up to two individuals: at most one animal and one plant individual. It may not contain more than one individual of either type. Types of individual are defined by its trophic position in the underlying interaction network (section 2.1.1). All basal species are plants, all other species are animals.

Distance on the lattice is defined as follows. The immediate neighbours of any given cell are the eight adjacent cells including diagonals (i.e. a Moore neighbourhood). These eight neighbours have a distance-1 from the central cell. This distance metric is used in the rules for movement and reproduction (SEE BELOW NUM?), and also in the calculation of various spatial metrics (section SECTIONNUM).

Initial conditions are defined randomly by the following setup procedure. A species is selected uniformly at random from the sixty species in the underlying network. A cell from the landscape is selected uniformly at random. If there is space in the cell, an individual belonging to the selected species is placed in the selected cell. This is repeated until the value of parameter *occupied cells* is reached.

Table 2.1 shows all the model parameters, their values and definitions. Where possible the parameters values are chosen to be biologically realistic. A discussion of values chosen for these parameters can be found in section 2.1.2.2.

2.1.2.1 Cellular-automata rules

In the following description italicised words refer to model parameters, which are defined in table 2.1. Each individual stores energy (or resource), which it expends to perform actions. If the energy of an individual drops below *min_resource* it dies and is removed from the landscape.

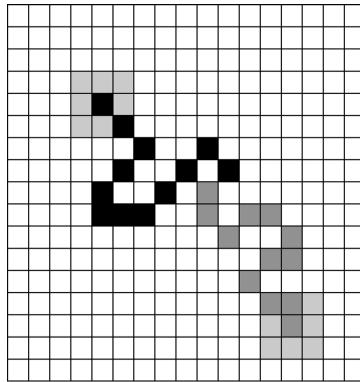


Figure 2.3: Example trajectories and neighbourhoods of two individuals.

On each iteration the basic demographic processes occur in the following order (FOR EACH SPECIES?):

1. Death
2. Movement
3. Reproduction
4. Feeding
5. Immigration

1) Death

As stated, if an individuals energy drops below *min_resource*, it is removed form the simulation.

2) Movement

For each individual, a neighbouring cell (distance 1) is selected uniformly at random. If the cell is available the individual moves there. Otherwise it remains stationary.

3) Reproduction

All species may only reproduce if their stored energy is greater than *mating_resource*. Animals reproduce sexually, plants reproduce asexually.

- **Sexual reproduction:** This occurs between two memebers of the same animal species if:
 - 1) There is a member of the same species in the immediate neighbourhood of the subject species; and 2) there is an available cell for the offspring in the distance 4 neiighborhood of the subject individual. When these two conditions are met both parents give a fraction of

their stored energy (*mating_energy*) to the offspring. The offspring is placed in a cell chosen uniformly at random from the available cells within distance 4 of the subject individual.

- **Asexual reproduction:** This occurs for plants via two possible mechanisms.
 1. Wind dispersal occurs for non-mutualistic plants, on each iteration with a probability equal to *reproduction_rate*. If reproduction occurs the offspring is placed in a randomly selected available cell in the distance 4 neighbourhood. For plants, available means empty or only occupied by an animal individual. If no cells are available the plant cannot reproduce. Again a fraction of the parent plant's stored energy (*mating_energy*) is given to the offspring to the offspring.
 2. Mutualistic dispersal occurs for mutualistic plants. This action is carried out by the animal partner, and is done in the 'feeding' phase (see below), since it is also a trophic interaction. The 'seed' of the parent plant is carried by the animal partner, so it may be placed beyond the distance 4 neighbourhood.

3) Feeding

For a trophic (feeding) interaction to occur, two individuals must belong to species that are connected in the interaction network. Also the individuals must explicitly find each other in space - in the 'movement' phase one must choose the cell occupied by the other. If this happens there are three possibilities:

1. **Predation:** If neither individual belongs to a basal species a predation event occurs with probability *capture_probability*. The prey species dies and a fraction of its energy *efficiency_transfer* is given to the predator.
2. **Herbivory:** If one individual is a non-mutualistic animal and the other is a plant, they interact. A fraction of the plant's energy *herb_fraction* is lost, and a fraction (*herb_efficiency*) of this energy is given to the herbivore. Both individuals continue living. If the animal is an omnivore and additional trade-off (*omni_tradeoff*) is applied to its energy gained, since omnivores are less efficient at digesting plant matter than straight herbivores.
3. **Mutualism:** If the individuals share a mutualistic interaction they interact. A fraction of the plant's energy (*mut_fraction*) is transferred to the animal. The animal also keeps track of which plant it interacted with. If it reaches an available cell in the landscape it creates an offspring of this plant with probability *mut_efficiency*. On each iteration that an offspring is not produced, the mutualistic efficiency is reduced by a fraction *mut_cooling*.

4) Immigration

At each iteration there is a probability (*immigration*) with which each empty cell may be colonised by an individual selected at random from the original species pool.

5) Energy update

On each iteration all animal individuals' energy stores are reduced by a fraction *living_expend*, to account for metabolic losses. Also all plant individuals autotrophically increase their energy stores by a fraction *synthesis_ability*. This is the only energy input to the system.

2.1.2.2 Model Parameters

The model parameters were chosen blablabla... reference Ings et al. Which parameters are most interesting, why? Also discuss sensitivity analysis from previous publications, findings. We I conduct my own version? With regard to which parameters? (Maybe discuss this lat bit somewhere else).

2.1.3 Modelling habitat loss

The current project extends the above defined model of Lurgi, Montoya and Montoya [39] by implementing habitat loss algorithms. The algorithms are simple. A fraction of the cells in the landscape are made uninhabitable to all species. We denote the fraction of destroyed cells by HL. The simulations are set up and run as detailed above (section 2.1.2), after 1000 iterations one of two habitat loss algorithms is applied to the landscape. The species inhabiting the destroyed cells are deleted. Species attempting to move into destroyed cell are unable to and remain stationary. Destroyed cells are counted as unavailable for the placement of offspring. We choose the cells to destroy using one of two habitat loss algorithms: 1) Random and 2) Contiguous.

2.1.3.1 Random Habitat Loss

Cells in the landscape may contain habitats in two states: pristine or destroyed. Pristine corresponds to the cells in the original model. To destroy habitat randomly, cells are selected and destroyed uniformly at random from the set of cells containing pristine landscape. This is repeated until the desired fraction of HL is achieved.

2.1.3.2 Contiguous Habitat Loss

This algorithm results in a contiguous region of cells with destroyed habitat. A 'seed cell' is selected uniformly at random from the fully pristine landscape. The seed cell is destroyed, and destruction proceeds radially from the seed cell until the desired fraction of HL is achieved. This process follows the same toroidal boundary conditions as the CA.

2.1. AGENT BASED SIMULATION

Parameter name	Value	Description
OCCUPIED_CELLS	0.4	Fraction of the grid initially occupied by individuals randomly placed on it.
MAX_RESOURCE	20	Maximum amount of resource an individual may possess at any given time.
MIN_RESOURCE	3	Death threshold: minimum amount of resource an individual may possess. Any individual possessing less than this amount at any given iteration will die (see text).
LIVING_EXPEND	0.01	Fraction of resource an individual spends in living every iteration of the model. Metabolic rate.
MATING_RESOURCE	0.5	Fraction of MAX_RESOURCE that is required for an individual to be able to reproduce.
MATING_ENERGY	0.2	Fraction of resource given to the offspring by the parent during reproduction. Each parent gives the same fraction. The total amount depends on how much resource the parent possesses at the time of reproduction.
IMMIGRATION	0.005	Probability that a new individual will appear in a cell of the grid each iteration. The species this individual belongs to is randomly chosen from the original species pool.
SYNTHESIS_ABILITY	0.1	Fraction of resource that is autotrophically created by each individual from the basal species every iteration. This is the only energy input to the system.
HERB_FRACTION	0.7	Fraction of resource lost to herbivores by individuals belonging to a basal species during a trophic event, i.e. a species in the first trophic level feeding on a species in the basal level.
OMNI_TRADEOFF	0.4	Fraction of resource that omnivores are effectively able to gather when feeding on a species from the basal level (a plant).
MUT_FRACTION	0.25	Fraction of resource of a primary producer (basal species individual) that a mutualistic partner obtains when an interaction of this type occurs.
CAPTURE_PROB	0.4	Probability that a predator individual embark upon a trophic relationship with one of its prey individuals when it encounters it.
EFFICIENCY_TRANS_F	0.2	Fraction of the resource the prey that is assimilated by the predator in a carnivorous interaction, i.e. trophic interaction not involving individuals from the basal species.
HERB EFFICIENCY	0.8	Fraction of the resource of the prey assimilated by the herbivore in an herbivorous interaction.
MUT_EFFICIENCY	0.8	Efficiency of an individual mutualist when dispersing a plant partner. In other words, the probability with which a mutualistic individual will facilitate the creation of a new individual of the last species of plant it visited when it is positioned on an empty cell immediately after it interacted with a mutualistic plant partner.
MUT_COOLING	0.9	Cooling factor for the mutualistic efficiency of plant dispersers (mutualists). This is the fraction of mutualistic efficiency that remains after each iteration.
REPROD_RATE	0.01	Reproduction rate of non-mutualistic plant species. Probability with which an individual belonging to a plant species that does not possess mutualistic partners for dispersal will create an offspring in any given iteration of the simulation run.

Table 2.1: The parameters of the model and what they mean.

2.1.4 Implementation

The code for the simulation model was originally written by Miguel Lurgi for research leading to the publication [39]. He and Daniel Montoya were responsible for the bulk of the model development, testing and parameter selection - a considerable task. My task was to take this legacy code and work with it to generate results that would allow us to study the effects of habitat loss.

The model is implemented in *Python*, with several switches that ensure portability between different versions of python. The programme makes extensive use of *numpy* and *networkx*, amongst other *Python* libraries. The original code was well written to allow for the easy implementation of new mechanisms. The intention was always to extend the model for further study. For example there is a parameter for *habitat type*, whereby the landscape would be heterogeneous with certain habitats best suited for different species. (In the current implementation this is not used and the landscape is homogeneous.) There was also a prototype algorithm for contiguous habitat loss.

Working with the legacy code I implemented the random habitat loss algorithm, and tested both the contiguous and random algorithms to ensure they were performing as desired. I also added methods to an already extensive library for saving simulation outputs, with a view to conducting more analysis on the spatial state of the system². I then ran numerous simulations without habitat loss (HL=0) to ensure that the model reproduced the results presented in [39].

Simulation ensembles were run on Blue Crystal Phase 3 (BC3), Bristol's computer cluster. Compatibility issues arose because only *networkx* version 1.9 is available on this system, which has significant differences in the return

My contribution to the model was the implementation of the random habitat loss algorithm, and running all the simulations on Bristol's computer cluster Blue Crystal. The former was trivial but key to this project. The latter involved a considerable amount of debugging and scripting and testing. I also made

I conducted post-simulation analyses in the statistical package *R*, and in *Python*. Analyses in *R* used adapted legacy code from the previous project. Analysis and plotting methods implemented in *Python* were written by myself.

Write here about code, my contributions, Blue Crystal. Repeats runs, result format, run times etc.

2.2 Dynamics of the model

Display and discuss several example full runs. Also with varied parameters?

Well mixed approximation?

²IMPORTANT: don't write about this if you are not going to do the analysis properly.

2.2.1 Transience

Discuss and analyse transience. How long is it? Are the general relationships? Perhaps use simplified modelling (ODEs) to try and predict the length of transience.

2.2.2 Long term distribution

Is the final dynamics (after transience) steady state? What can we say about this?

2.2.3 Diffusion behaviour

How is species movement/dispersion affected by habitat loss? Can we derive a diffusion coefficient?

2.3 Ecological metrics and analysis methods

Introduce, define and discuss each metric.

Stability - Jacobian, dynamic stability, multi-stability, CoV, reproductive stability.

Robustness - secondary extinctions, cascading effects [14]. Re-wiring algorithms?

2.3.1 Biodiversity metrics

Richness, Simplicons, Shannon Entropy.

2.3.2 Stability metrics

Coefficient of Variation, May Stability

2.3.3 Network metrics

GenSd, VulSd, Gq, Vq, MTP, H2', Connectance, Nestedness, Compartmentalisation

2.3.4 Spatial metrics

Moran's I, Geary's C. Spatial autocorrelation. Centroids.

2.3.5 Interaction strength metrics

IS1, IS2, IS3 [6].

CHAPTER



HABITAT LOSS WITH HIGH IMMIGRATION

3.1 Response of ecological metrics

For both habitat loss scenarios there is no loss of species richness up to 90% habitat loss. However significant changes are observed in metrics relating to community composition, network properties and stability. In general the qualitative response of these metrics is not governed by MAI ratio i.e. the direction of the trends are the same across MAI ratios, although the extent of the response may vary.

Although habitat destruction does not lead to extinctions, it does reduce the total biomass of the communities. This is measured by the total number of individuals of all species remaining at the end of a simulation, which we average across replicate communities. Figure 3.1 shows that, although mutualistic communities contain more biomass, the loss of biomass due to habitat destruction is ubiquitous. This is not surprising.

How community stability is affected by loss of habitat is of particular interest. Temporal stability is measured by the coefficient of variation (CV) of the total biomass of the system. (This metric is calculated over a number of iterations after the transient dynamics.) A higher CV indicates lower stability, because there are greater fluctuations in the dynamics. We observe that temporal stability is affected differently by the two habitat loss scenarios. As shown in figure 3.2, random destruction increases temporal stability, whereas contiguous destruction decreases it. What is driving this different response in the dynamics?

From figure 3.3 it is clear that the response in temporal stability is closely correlated with interaction strengths, as measured by the metric IS3. Random habitat destruction is characterised by a decrease in total interaction strength, whereas contiguous destruction results in a dramatic increase. It is reasonable to assume that these changes in IS3 are causing the different responses

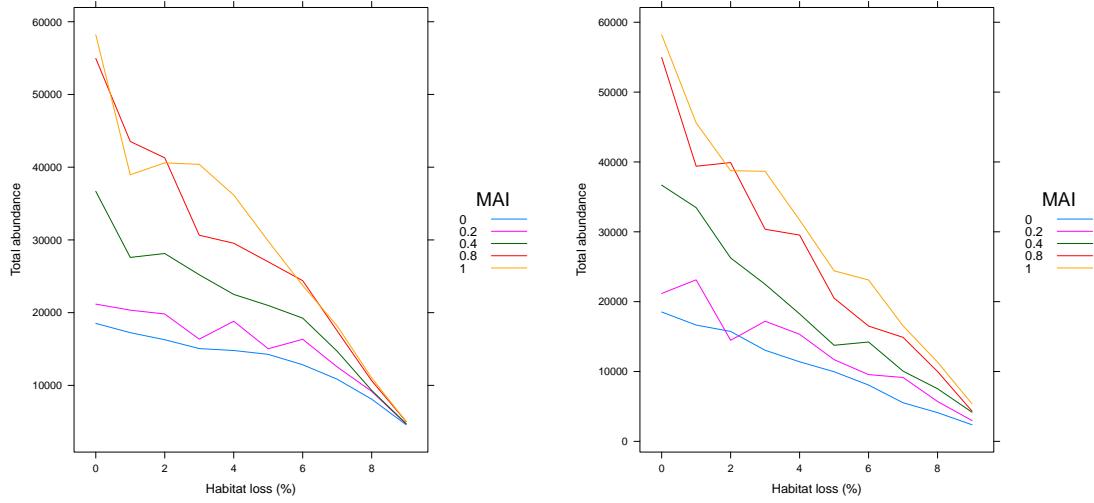


Figure 3.1: Mean total number of individuals across replicate communities decreases with habitat loss, for all MAI ratios. Left: random destruction. Right: contiguous destruction.

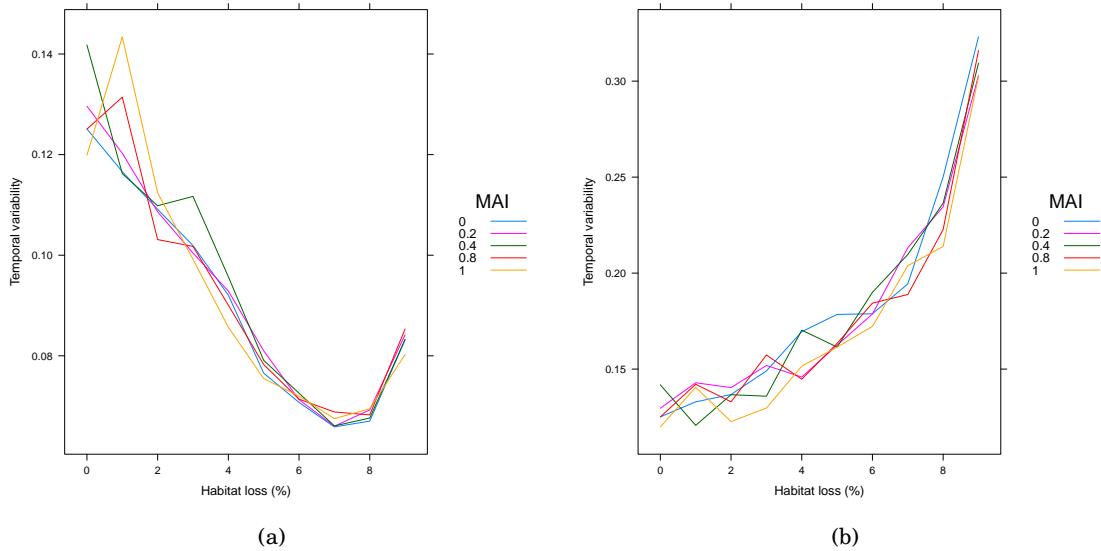


Figure 3.2: Mean CV in total biomass across replicate communities against habitat loss for selected MAI ratios. Left: random destruction leads to an increase in temporal stability, as indicated by a decrease in CV. Right: contiguous destruction dramatically reduces temporal stability.

3.1. RESPONSE OF ECOLOGICAL METRICS

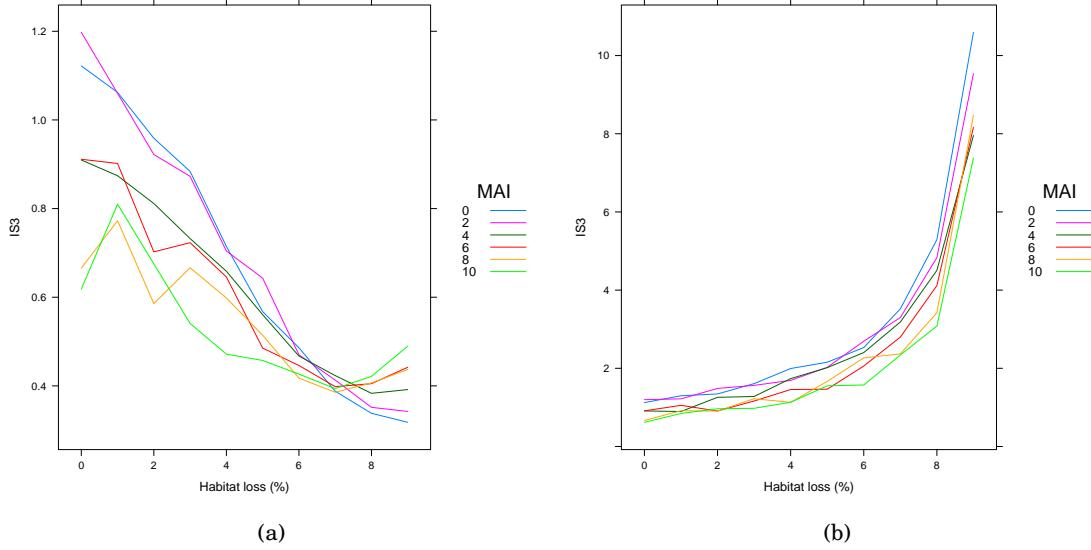


Figure 3.3: Interaction strengths: The sum of the elements of the interaction matrix, averaged over replicate communities, for selected MAI ratios. Left: random destruction reduces total interaction strength. Right: contiguous destruction dramatically increases total interaction strength.

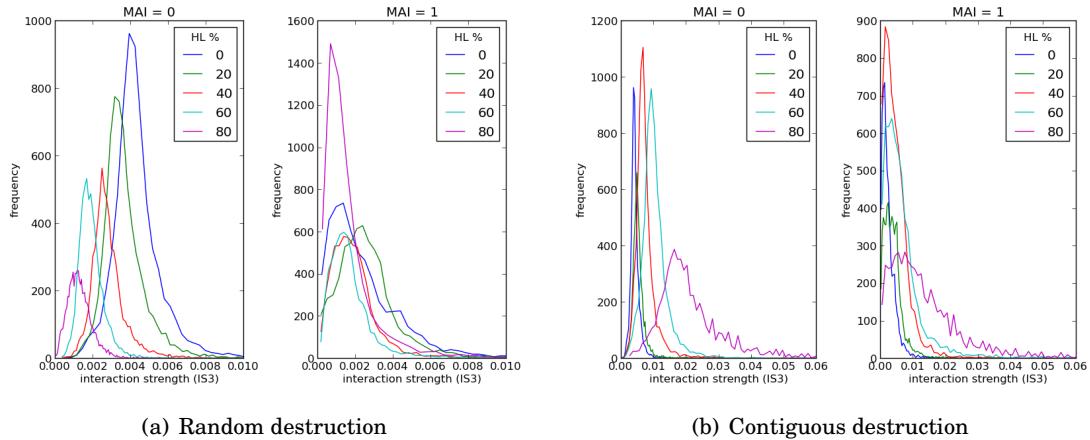


Figure 3.4: IS3 distributions: The interaction strength distribution shifts leftwards for random habitat loss, and rightwards for contiguous loss. The extent of this shift is mediated by MAI ratio - it is most visible for MAI= 0, and least visible for MAI= 1.

in temporal stability. This is supported by the literature, where it is well documented that strong trophic interactions destabilise population dynamics [REFERENCES]. It is not so obvious what affect on dynamics should be expected from an increase in the strength of mutualistic interactions, which are included in this metric. However, figures 3.2 and 3.3 suggest that this is also destabilising.

The trends in total interaction strength are due to underlying shifts in the distribution (figure 3.4). This is made visible by a shift in the modal peak to lower or higher interaction strengths, for random or contiguous loss respectively. The extent of this shift is mediated by MAI ratio, with greater shifts for lower levels of mutualism.

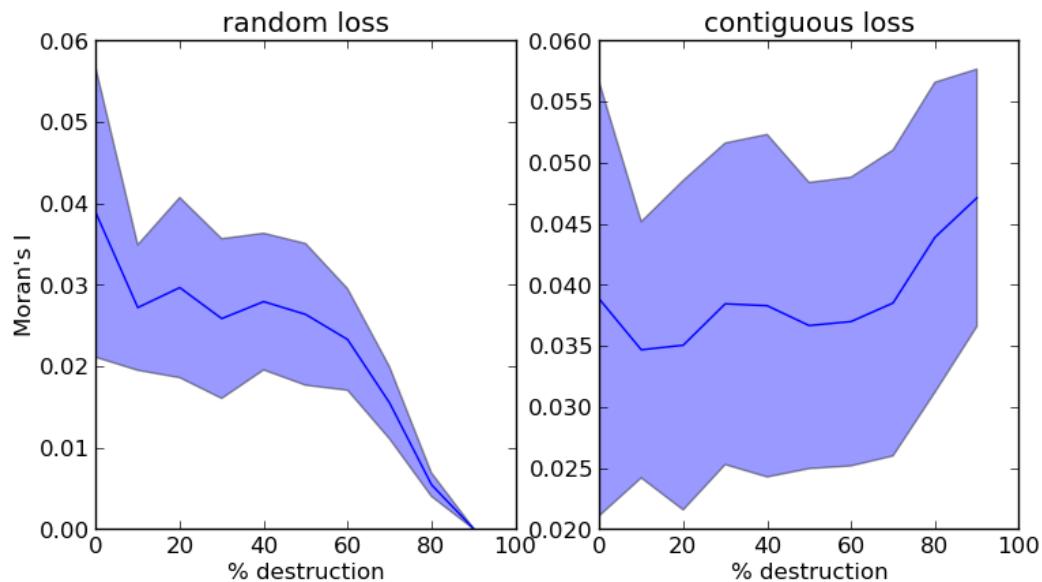


Figure 3.5: Spatial autocorrelation: Moran's I is calculated for all species distributions and averaged over the community. These plots show the results averaged over all replicates and over all MAI ratios. Solid lines indicate the mean value, and shaded areas indicate ± 1 standard deviation. This metric suggests that species distributions, on average, become less aggregated in space due to random destruction. Whereas they appear to become more spatially aggregated as a result of contiguous destruction.

What is causing the different responses of IS3 to the different habitat loss scenarios? Spatial aggregation of species distributions is quantified using Moran's I [REFERENCE]. According to this metric, species distributions become less aggregated in space as a result of random destruction (figure 3.5). Whereas they appear to become more aggregated in space under contiguous destruction. If species are more aggregated in space, it may be easier to find an interaction partner. This would benefit predators (and mutualists), potentially leading to stronger de-stabilising trophic interactions. The mean frequency of interactions (figure 3.6) shows some evidence for this effect. On average there are fewer interactions in communities suffering random habitat destruction, than those with contiguous destruction. However the difference is small, suggesting another mechanism is required to explain the strong responses of IS3 and temporal stability.

It is likely that other changes in network properties and community composition can explain the changes in IS3. The elements of the interaction matrix, used to calculate IS3 are given by [6, 68]:

3.1. RESPONSE OF ECOLOGICAL METRICS

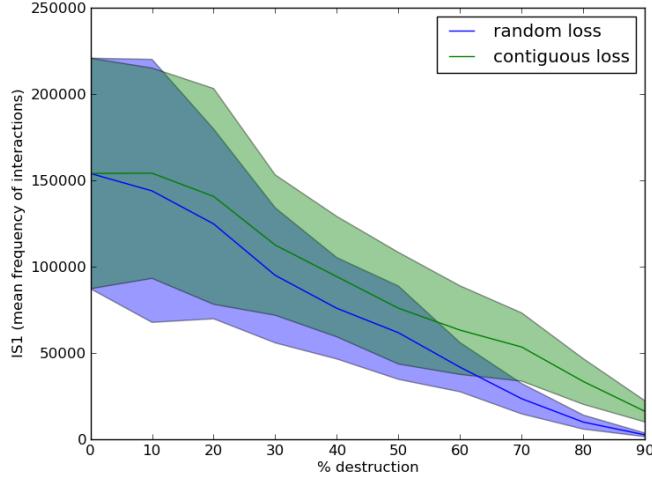


Figure 3.6: Interaction frequencies: The metric for interaction strength IS1 averaged over all replicates and all MAI values. This metric is the sum of the elements of the interaction frequency matrix i.e. the total number of interaction occurring within a given period of time. Solid lines indicate the mean value, and shaded areas indicate ± 1 standard deviation. For contiguous destruction interactions are slightly more frequent, on average, than for random destruction.

$$(3.1) \quad \alpha_{ij} = \frac{b_{ij}}{N_i N_j},$$

where α_{ij} is the effect of species j on species i ; b_{ij} is the biomass flow from i to j (here measured by the frequency of the interaction, equivalent to IS1) and N_i, N_j are the number of individuals belonging to species i and j respectively. Therefore the elements of the interaction matrix are dependent on the relative abundances of the interacting species, as well as the frequency of the interaction. A more detailed analysis of community composition will let us explain to observed trends.

3.2 More preliminary results

By averaging results over MAI ratios we are able to effectively obtain more replicate communities. When doing this certain metrics appear to display trends in response to habitat loss that are not clearly visible without this averaging. We perhaps need to justify this approach, and to fit statistical models to quantify the significance of the trends in certain metrics.

3.3 Further development and work to do

3.3.1 Work to do

- Write introduction
- Write up ecological metrics
- Amend figure captions
- Plot example networks
- Plot dynamics
- Analyse dynamics (transience, steady-state)
- Fix conflict in trophic level results

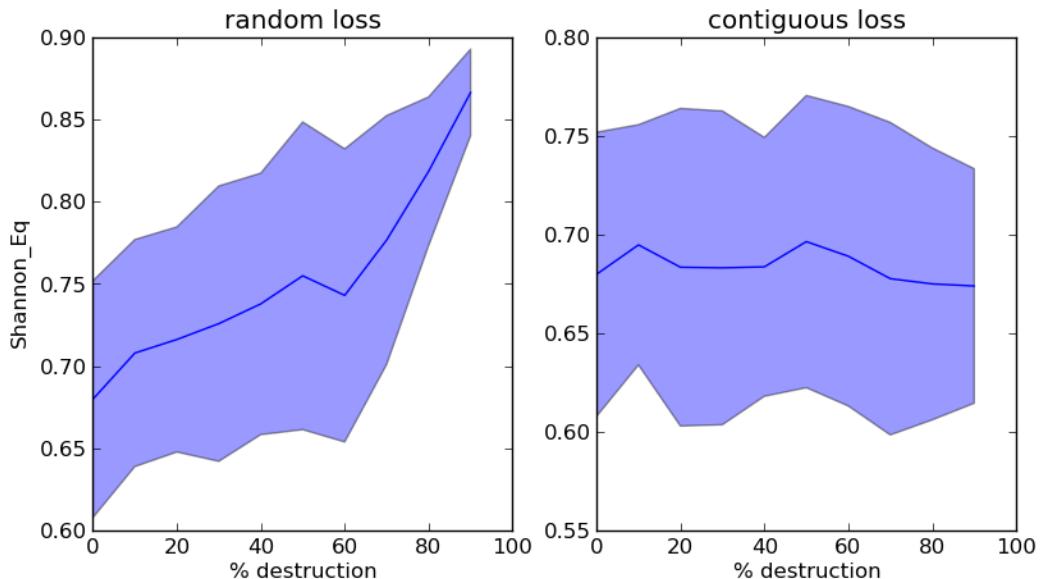


Figure 3.7: **Shannon evenness:** average over MAI ratios and replicate communities. Under random habitat destruction communities on average become more even. Under contiguous destruction there is no change.

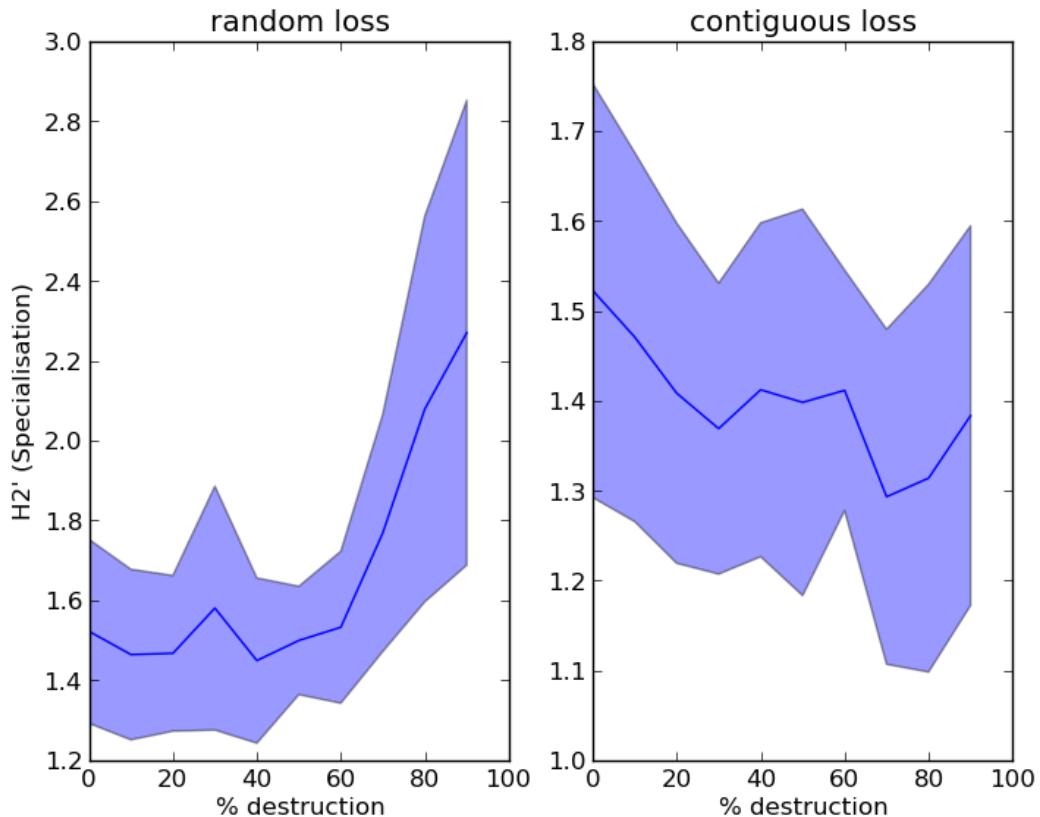


Figure 3.8: **Specialisation:** $H2'$ is a metric that quantifies the degree of specialisation of interactions in the mutualistic sub-network. Mutualists appear to become less specialised under random destruction, and more specialised under contiguous. (This appears to disagree with previous plots of $H2$, but I can't see why - the numbers used are taken from `output_network.csv`, column titled $H2$.)

- Analyse ecological metrics by trophic level (basal, intermediate, top)
- Understanding IS3 response: plot $IS1/N^2$, look at individual runs
- Compare movement of species to diffusion coefficient in porous medium
- Plot RADS with species coloured according to trophic level.

3.3.2 Further development

- Run simulations with lower immigration
- Well mixed approximation. Simplified model for analysis

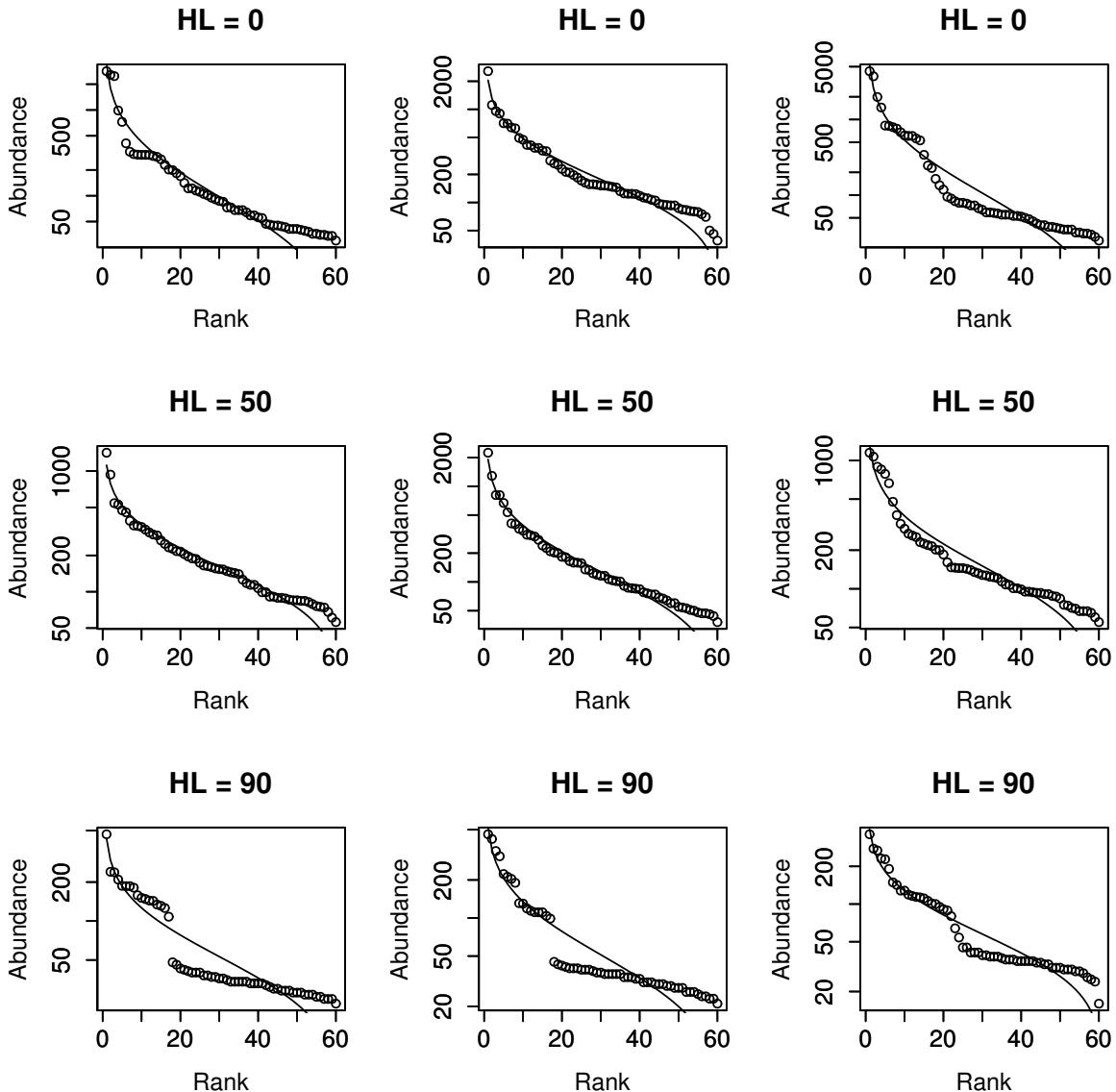


Figure 3.9: RADs for Random habitat loss: Rank abundance distributions for MAI= 0, three replicate communities shown for selected levels of habitat loss (0,50,90). In the case of random destruction we expect the communities to become more even on average, based on Shannon_Eq metric, (figure 3.7)

3.3. FURTHER DEVELOPMENT AND WORK TO DO

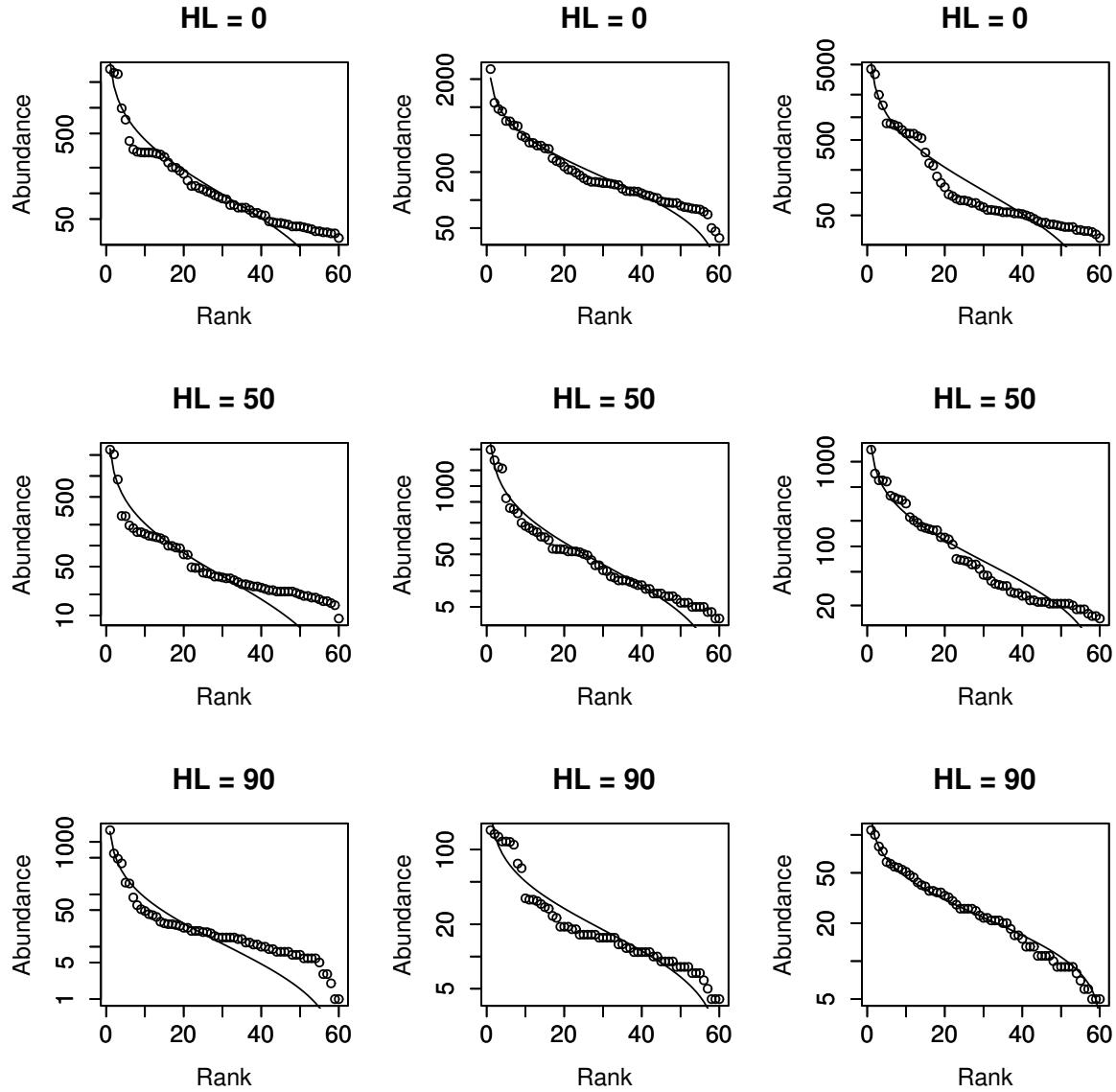


Figure 3.10: RADs for Contiguous habitat loss: Rank abundance distributions for MAI=0, three replicate communities shown for selected levels of habitat loss (0,50,90). In the case of contiguous destruction we expect no trend in evenness on average, based on Shannon_Eq metric, (figure 3.7)

CHAPTER 3. HABITAT LOSS WITH HIGH IMMIGRATION

- Can we calculate robustness from network, and test this against what extinctions we get?
(With lower immigration)

VARYING IMMIGRATION RATE

This chapter explores the interactions between immigration rate, habitat loss and MAI ratios - choose a better title.

TODO: link up these section references. (and add recent papers into .bib file)
TODO: finish discussion of results. But first redo analysis with averaging? TODO: look at "effective immigration" rates..? (inversely proportional to total biomass) TODO: show some results here for 0 HL TODO: look at longer runs (steady-state) TODO: look at over-abundance of top trophic level (re-run results) TODO: averaging of RADS (and others..) WAIT FOR STEADY STATE. TODO: either end each section with ecological context and implications, or have in separate section after results. TODO: find out about Kevin McCann on omnivory (biomass pyramids?)

4.1 Assumptions of what has gone before

- Discussion of what we will now refer to as "default parameter values" (see table ??)
- Conclusion that the default immigration rate is high and this is an open system. This represents a restricted scenario where the regional species pool is constant and high (60 species), all species are equally likely to immigrate and dispersal from outside our 'world' is not heavily constrained. In this chapter we look at varying immigration rates, habitat loss and MAI.
- A discussion of the "rescue effect" due to immigration. Probability of species spawning. Effective immigration, although currently that concept is first mentioned in this chapter.
- Context of 'viable species' as opposed to those that only remain due to immigration 'bubbling' along close to zero. (note that this is quite realistic)

- Snapshots of system versus averaging of metrics over a number of iterations (seems a bit late in the thesis to realise that this is a problem in chapter 4!) - averaging over replicates can justify this to some extent (25 in previous chapters, 100 here).
- Discussion of the functional groups and their names. These assumed below.
- It may be that this model is well approximated by GLV. If so it would make sense to discuss the results with reference to that? Should try to fit?

[53]

4.2 Literature review for immigration

Possibly include here a summary of recent work on immigration, including IBT and meta-community theory. (Alternative is that this goes in introductory chapter.)

IBT -> species area relationship (SAR), does high immigration reduce this effect. [66] - Search for immigration.

[33] Many studies of dispersal/immigration take a metacommunity approach. In general higher dispersal between communities is found to promote better species richness, and lower variation between communities. However we are interested in...

[38] shows, using a metapopulation model, that competitive plant communities benefit from high immigration. In these communities competition for space would lead to the extinction of all but one species. Therefore some immigration is required for diversity to exist. (Similar to our case of zero IR. Is there evidence for spatial competition in our simulations?) Some metrics change with IR some do not (check which ones, do they agree with their experiment, and ours?) They have species specific immigration rates, and an intensity parameter (otherwise same as ours with effective rate proportional to number of vacant sites...nice). Our simple case is simpler than their simple case 1 (IBT -size of regional pool and degree of isolation). They call the immigration a 'propagule rain'.

Their model displays classical competitive exclusion: "This result is different from that usually found in metapopulation competition models (Levins and Culver 1971; Horn and MacArthur 1972; Slatkin 1974; Hastings 1980; Nee and May 1992; Tilman 1994). These models implicitly or explicitly allow interference competition between species within sites, and coexistence is obtained when there is a trade-off between competitive superiority and colonization ability. Because we do not consider interference, there is no possibility for such a trade-off in our model and, hence, coexistence in a closed community is impossible. As mentioned earlier, we deliberately ignore trade-offs associated with interference competition because we wish to explore the effects of immigration from an external source on their own."

4.2. LITERATURE REVIEW FOR IMMIGRATION

They never get stochastic extinctions because their model is continuous, therefore use an extinction threshold.

"Thus, there is always stable coexistence in a community with a propagule rain." (Interesting dicussion at end : where does the propagule rain come from?)

As immigration goes to zero they find that most species have a density close to zero (therefore likely extinct) and that dominance is determined by basic reproductive rate ($r = c/m$). Also space is more fully occupied as IR increases. (Figure 1 nice plot of expected number of species versus immigration intensity). Interesting change in dominance relationships (relative abundances) with IR.

Experimental results by Mouquet and Loreau [45] suggest that immigration has a positive impact on plant community diversity. (Manipulation of seed rain.) However it also shows that certain community-level properties do not depend on IR e.g. total biomass (check others). Also local competition effects.¹

[21] importance of the propagule pool in determining local diversity, and shifting limitations hypothesis (SLH) [9] dispersal as a community structuring mechanism, immigration increasing local diversity, effect differs between plants and animals, intermediate dispersal rates best?

[69] Importance of dispersal in maintaining diversity in fragmented habitat patches (small tropical forest fragments). Dispersal limitation versus edges effects. Must be other REFS for this as well?

[10] Obviously proponents of neutral theory, but suggests the importance of immigration, and measures immigration rate close to ours..

[12] mathematical treatment of 3D prediction model, effect of immigration rates.

[51] "Although previous studies as far as we are aware did not investigate the effects of landscape context on immigration rates, immigration has consistently been shown to increase population size in patches both in field studies and simulation models [44], & [47]." Importance of landscape context.

[35] Changes in demographic rates due to Hfrag, in addition to population sizes and extinctions.

[25] 70 per cent of remaining forest within 1km of forest edges. Fragmentation effects. Smallest and most isolated patches most vulnerable. Reduce biodiversity by 13 to 75 per cent (metrics?) and reduce biomass, affecting ecosystem functions.

[60] Testing different HL scenarios on neutral and non-neutral communities. Habitat alteration reduces average level of specialisation - 'functional homogenisation'.

¹Nice structure here - theoretical results, followed up a few years later by experimental study. Talk about this somewhere.

- [1] Communities are more repeatable (similar between repeats) for high immigration rates. Use IBM model. Immigration from regional species pool. Zero sum netural theory dynamics.
- [27] SAR updated to account for habitat fragmentation i.e. habitat loss that is not contguious. Veyr relevant for discussion somewhere.
- [22] spatial competition.
- [15] food web plasticity, stabilising and predictable.

4.3 Motivation

In chapter ?? we saw that no species go extinct, even at extreme levels of habitat destruction ($HL = 90\%$), when using the default parameter set (see table ??). Motivated by [63], this allowed us to explore community responses to habitat loss that are not associated with, or may preceed, the loss of species. The lack of extinctions produced by the model were shown to be due to a rescue effect from immigration. Even species which go locally extinct from the landscape may be replaced by this immigraiton.

The simulations presented so far represent *open* communities with a strong influx of individuals belonging to all species.

the default immigration rate (IR), as given in the default parameter values is relatively high, corresponding to an open community. In particular, at the IR, we expect no extinction of species

This behaviour (control for species richness). However for such a heavily impacted community to not exhibit local extinctions would be unusual in nature [REF]. This may be considered an edge case - an open ecosystem with a strong influx of individuals from all species. Although the local habitat may be very close to total destruction the community is sustained by strong immigration from surrounding habitats. In reality such a strong and uniform rescue effect from immigration is unlikely due to spatial auto-correlation, differential dispersal rates and other effects (see discussion in section ??[references - York pollinator study]).

- propagule rain
- Require a uniform species pool to be maintained (trad IBT -> continent), or hetegeneity at landscape level [REF]

In this chapter we further investigate the impacts of habitat loss (HL) on multi-species communities with different proportions of mutualistic and antagonistic interactions (MAI ratios). We now consider other realistic scenarios by varying the immigration rate parameter. At one extreme we have the above scenario of high immigration, where extinction is

prevented. At the other extreme we have closed communities with zero immigration. In this case there is no rescue effect from surrounding habitats, and we may expect to see extinctions in response to habitat loss. Although a totally closed system does not exist in nature certain systems may come close to this ideal. For example an island community that is a sufficiently distant from other land (see discussion on Island bio-geography theory in section ??) will have very low immigration rates, and systems that are effectively closed may be artificially achieved in controlled situations (e.g. laboratory mesocosm). Although extremely open and extremely closed systems are possible, most real-world communities lie within these two extremes. By changing immigration rates, the approach followed here allow us to explore the entire range of possible responses² of biological communities to habitat loss.

For the default parameter values (see table ??) zero IR results in the inevitable extinction of all non-plant species, in our simulated communities. We will refer to this scenario as *community collapse*. Even for pristine habitats (0% HL) we do not see stable and persistent communities without some non-zero IR (some increase in persistence with MAI ratio - see chapter ??). This result is demonstrated in chapter ??, where we explore factors contributing to stability. For now we accept the general result that, with the default parameters, zero IR results in community collapse³. In this chapter we are interested in the regime between these two extremes of zero IR, where we see many extinctions even at 0%HL, and high IR where we see no extinctions even at 90%HL. We are particularly interested in finding IRs for which communities are stable at low levels of HL, but where collapse is initiated as HL is increased. This is a scenario that we see in real-world communities. We are also interested in how community composition and stability vary with HL and IR, and how this is mediated by MAI ratio. (And interaction strength distributions.)

²Not sure about this..

³weaken this statement, parameter dependent, or at least refer forwards again to next chapter

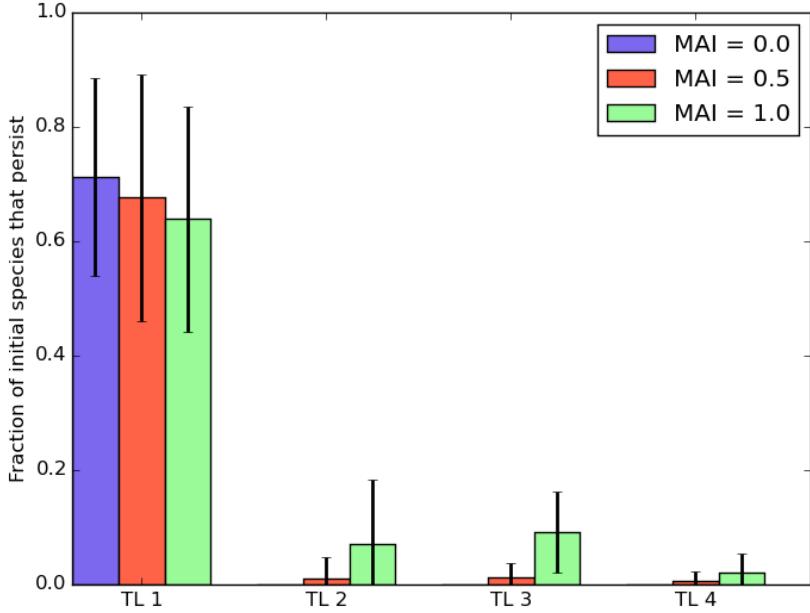


Figure 4.1: Fractional persistence by trophic level for three different MAI ratios. Fractional persistence is measured by the fraction of species initially belonging to a trophic level which have not gone extinct by the end of a simulation (5000 iterations). The solid bars give the mean value, taken from 22 repeat simulations. Error bars show \pm one standard deviation. (Simulations from chapter 6

4.4 Exploration of parameter space: habitat loss and immigration

Motivated by the above we explore a two dimensional slice of parameter space. The immigration rate (IR) and the level of habitat destruction (HL) are varied and one hundred repeat simulations are conducted at each pair value. Therefore we are able to estimate how the simulated communities are expected to behave across this region of parameter space, by averaging over the repeats. These simulations are run for three different MAI ratios: 0.0, 0.5 and 1.0. As in previous simulations each repeat uses a different interaction network topology, generated with the procedure described in section ???. All other parameters are held constant at their default values (table ??), including the number of iterations which remains at 5000.

In order to speed up the simulations certain metrics used in the previous analysis (chapter ??) are not calculated. Particularly the spatial stability metrics are computationally expensive. Only two pieces of information are saved as output from these simulations: the underlying network structure and the abundance time-series for each species. The abun-

4.4. EXPLORATION OF PARAMETER SPACE: HABITAT LOSS AND IMMIGRATION

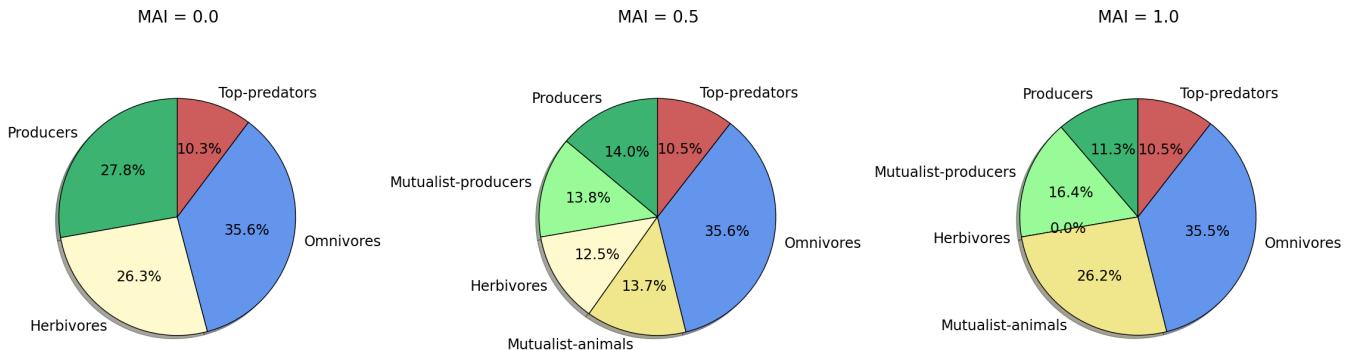


Figure 4.2: Mean number of species belonging to each functional group for the three MAI ratios in consideration (0.0, 0.5, 1.0). The results are averaged over one thousand simulations with the given MAI ratio, selected from the total ensemble of simulations that were run for this chapter (and used to generate e.g. fig 4.11). The species numbers depicted are independent of all simulation parameters, other than those that define the interaction network. That is the average number of species in each function group depends on the niche model parameters (connection and number of species), the MAI ratio, and the trophic constraints that we impose.

dance time-series is simply a record of the abundance of each species at every simulation iteration. (What do we calculate from these and why..) By limiting the simulation output the scope for analysis is restricted but the parameter space can be explored in more detail (higher resolution, greater number of repeats)⁴. This ‘first pass’ scan of the parameter space allows us to construct a general picture of how the model behaves in this region. It may also be used to identify subsets of the region of parameter space on which to focus further computational effort for e.g. spatial analysis.

The entire range of habitat destruction is explored from pristine landscape (0%*HL*) to near total destruction (90%*HL*) in steps of 10%. In the current chapter all habitat is destroyed using the contiguous algorithm since it was decided that this is more realistic (see discussion in section ??). Ranges for the IR were chosen based on previous simulations. Since *IR* = 0.005 is sufficiently high to prevent any extinctions, this was taken as the maximum of the range. Simulations using *IR* = 0 have already determined that this leads to community collapse, therefore these were not repeated. A value of *IR* = 0.0001 was heuristically selected as the lower bound, at which some non-zero extinction is expected in pristine habitat for all MAI ratios⁵.

⁴The changes to the model output reduced run times by up to a factor of 10. This required changes in how the interaction network is represented and therefore previously used network metrics could not be calculated. However it would be easy to modify the code to save some information on interaction frequencies and spatial states, which could be used in later analysis. This has not been done yet -ON THE WAY.

⁵Dani thinks we should maybe look at lower IR values.

The choice of MAI ratios allows us to compare purely antagonistic ($MAI = 0.0$), mixed ($MAI = 0.5$) and purely mutualistic ($MAI = 1.0$) communities. Figure 4.2 shows the expected fraction of species belonging to each of the six functional groups in the interaction networks for these communities. The constraints we place upon the niche model are that at least 25%, 25% and 5% belong to the first, second and fourth trophic levels respectively. In particular it is known that the unconstrained niche model struggles to generate realistic number of species in the second trophic level [REF]. The figure shows that the interaction networks meet these constraints and that, as expected, the largest number of species is found in the third trophic level⁶ i.e. the functional group labelled *omnivores*. Antagonistic communities are missing the two mutualist functional groups from the first two trophic levels, whereas the mixed communities have a roughly 50 : 50 split between mutualists and non-mutualists as expected (this split is not exact because it is links that are switched no species). Importantly although the purely mutualistic communities contain no herbivores (as all their links to plants have been switched), they do contain non-mutualist plants. These plants are those that share no interactions with the first trophic level⁷, therefore the link replacement procedure does not give them any mutualist partners. These plants remain wind dispersed and are predated upon by animals from trophic levels three and four⁸.

⁶Perhaps these constraints should be changed in future simulations - discuss with Daniel. - he thinks OK. Suggests look at RADS with colouring by TL. REQUIRES LOOKING AT SINGLE NETWORK.

⁷Is this realistic - Daniel? - RESTSAE QUESTION: REALISTIC THAT PLANTS ARE NOT EATEN BY FIRST TROPHIC LEVEL.

⁸Should there be a constraint that top predators do not consume plants? Not in original niche model. (Dani says no. Re-run. Does it make a difference?)

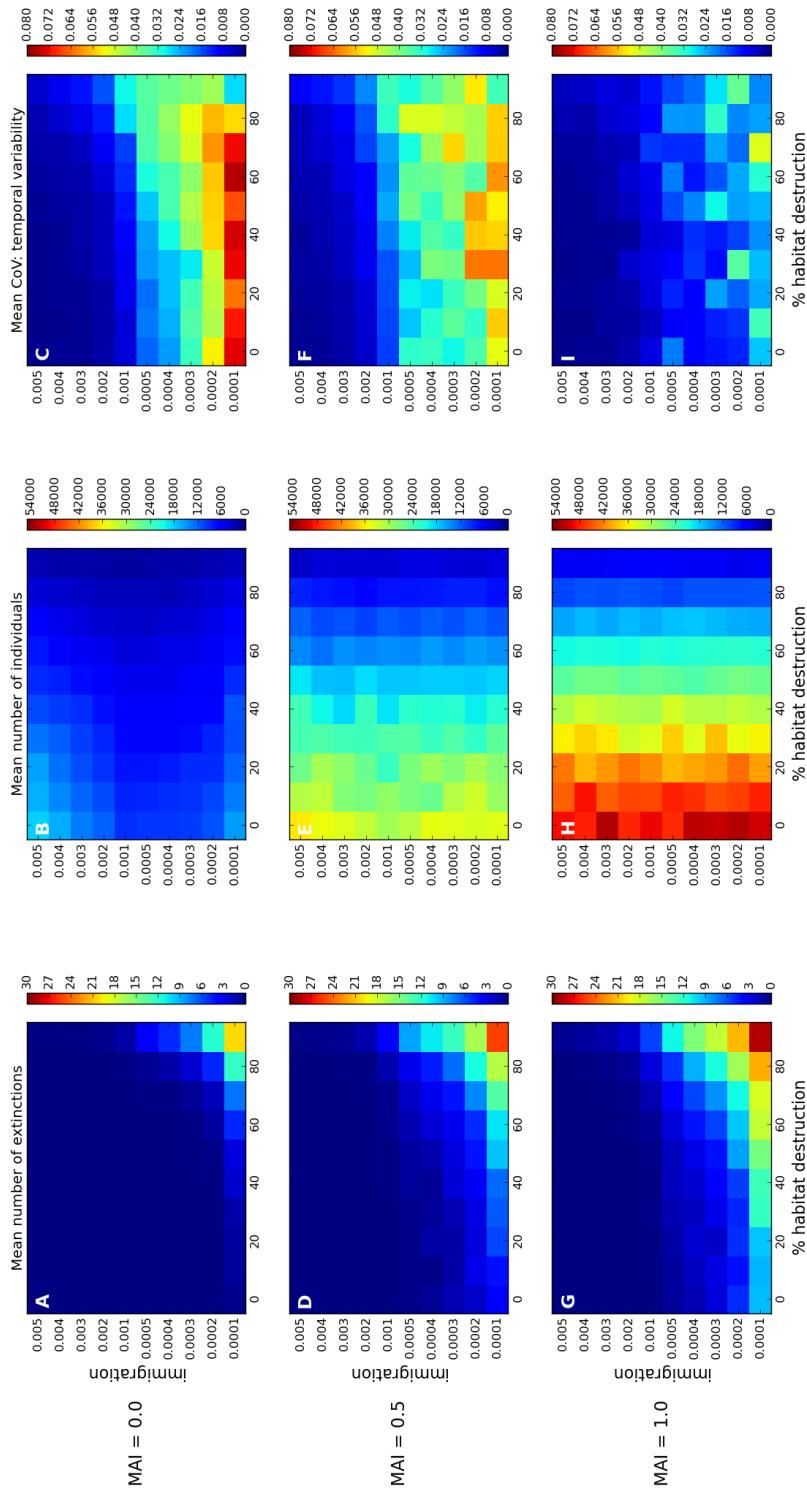


Figure 4.3: Summary heatmaps: Each heat map shows the value of a certain response metric across a 2-dimensional slice of parameter space. The parameters varied are immigration rate IR (y-axis) and percentage habitat destruction HL (x-axis). Each row of plots corresponds to a different MAI ratio as labelled. To construct the heatmaps one hundred repeat simulations were run for each combination of parameter values, with each simulation using a different underlying network. The mean value of the response metrics is taken over the hundred repeats. Therefore each pixel shows an estimate of the expectation value of the metric at those parameter values. The left column shows the expected number of species that are extinct at the end of a simulation; the central column shows the expected biomass (total number of individuals) at the end of a simulation; and the right column shows the expected temporal variability (coefficient of variation of total biomass) of the dynamics during the final thousand iterations of a simulation. The latter is used as a proxy for stability (see text).

4.4.1 Summary heat-maps

The results of these simulations can be concisely represented as heat maps over the region of parameter space explored. Figure 4.11 shows how the expected value of three summary metrics varies across this space: the number of extinct species, community biomass (total number of individuals) and temporal variability in community biomass. The response of each of these metrics is discussed individually below. The latter is used as a proxy for stability and is measured by the coefficient of variation (CoV) of the community biomass about its mean during the final thousand iterations of a simulation. (later..) This metric is often used to assess dynamic stability, but should not be confused with rigorous mathematical metrics relating to the stability of the equilibrium state of the system [REF]. It should be noted that the other two metrics, and all abundance measures in the following analysis, are calculated from a snapshot of the system state on the final iteration of a simulation. (Dani points out that average over replicates..., compare with averaged analysis, and discuss in context of steady state.)

Extinctions

No species extinctions are expected for sufficiently high levels of IR, across the whole range of HL values and for all MAI ratios. This results is visible in the left column of figure 4.11 and was already discussed in section ???. It is found that reducing the IR leads to an increasing number of extinctions. At low IR extinctions are possible, even in pristine landscape. This fits the previous observation that zero IR always leads to community collapse.

Increasing HL generally increases the number of expected extinctions. However nowhere in the parameter space do we see community collapse. In the most extreme case of low IR and high HL ($MAI = 1.0, HL = 90, IR = 0.0001$) an average of close to thirty extinctions may be expected. Although this expected loss of half of the species is fairly catastrophic, it does not guarantee total collapse of the community. The trophic constraints imposed in the food-web generation procedure ensure that at least 25% of species belong in the first (basal) trophic level (figure 4.2). In practice this very rarely (quantify) reaches above 30%. Therefore a loss of thirty species suggests that at least 40% of the remaining species are non-basal⁹. In other words, despite significant loss of species, there is some persistence in higher trophic levels.

For all three MAI ratios there exists an IR where the expected number of extinctions is zero in pristine landscape, but increases with HL. So although the immigration rescue effect prevents total community collapse, we do have a situation where HL can initiate species

⁹If all thirty species lost are non-basal we are left with 3/5 basal species to 2/5 non-basal. IN NATURE HIGHER TROPHIC LEVELS USUALLY MORE VULNERABLE [REF]. IS THIS THE CASE. OOO, COMMUNITY COLLAPSE.

extinctions. The IR at which extinctions are initiated is increased by increasing the MAI ratio. This effect of MAI ratio on extinctions is general. On average we expect a greater number of extinctions for high MAI (1.0) than for low MAI ratio (0.0), all else being equal. At the lowest IR and with pristine habitat we may expect about one extinction with a MAI ratio of 0.0, compared to about ten extinctions with a MAI ratio of 1.0. This can possibly be explained by looking at the second column in figure 4.11. On average a higher MAI ratio lead to a greater total number of individuals at the end of a simulation¹⁰. This means that there are fewer empty landscape cells into which an individual may immigrate at random. This reduces the *effective immigration rate* and so weakens the rescue effect. Any very rare species, only made viable by immigration, will be the ones hit by this and are likely to go extinct¹¹.

Community biomass

There are strong trends in expected community biomass. Increasing HL has a negative effect on community biomass. This is intuitive and has been seen before. Also previously discussed (chapter ??) is the result that, on average, communities with higher MAI ratio can support a greater biomass. However this effect is striking in these results, especially at low levels of HL. In a pristine habitat with an IR of 0.005, the expected number of individuals for a community with $MAI = 0.0$ is around 20,000, compared to around 50,000 for a community with $MAI = 1.0$. In fact, across the parameter space, purely mutualistic communities have around twice the biomass of purely antagonistic ones. Therefore in some sense mutualism appears to be ‘better’ for the community. (Dani: although having more individuals means rescue effect less likely, and perhaps increase competition for space) In section ?? we discuss whether it is better for the community as a whole, or only for those species that engage in mutualistic interactions.

For antagonistic communities ($MAI = 0.0$) the biomass is dependent on IR. Both very high and very low IRs support high biomass, whereas intermediate IRs support less (central panel, top row, figure 4.11). The effect of high IR is intuitive - births due to high immigration supplement births due to reproduction in the local community. This supplementary effect is greater at higer IR. However the increase in biomass at very low IR is harder to explain. We know that at zero IR all non-plant species go extinct [REF CHPAT/SEC]. So we may expect

¹⁰Mechanism behind this? - From the Theoretical Ecology paper: "communities with larger MAI ratios hosted a larger number of individuals ($F(1273) = 98.69, p < 0.001$) (Fig. 4). In spite of a decline in the abundance of non-mutualistic primary producers and herbivores with increasing MAI ratios (as expected due to a larger fraction of mutualistic species), the increase in mutualistic plants and animals overcompensated for this loss, causing an overall increase in abundance. This overcompensation was due to mutualistic plants becoming more abundant than non-mutualistic ones since mutualistic consumers do not consume as much resources from them and are, additionally, beneficial for their reproduction.

¹¹To determine if this is what is happening need to look at total abundances?

that in the region of low IR non-plant species become increasingly rare¹². In an antagonistic community this means a reduction in the number of herbivores and omnivores, which will benefit plant species. Therefore we may propose that the increase in the biomass at low IR is accounted for by an increased abundance of plants¹³. This reasoning suggests that we should expect a difference in composition between the abundant antagonistic communities seen at low and high IR (see section ?? - ABUNDANCE DISTS.).

Mutualism removes the dependence of community biomass on IR. Although the total biomass does not vary with IR for these communities ($MAI = 0.5, 1.0$) there may be changes in community composition. For example it is still reasonable to suspect that non-plant species become increasingly rare at low IR. However in a mutualistic community this has a different effect. It will benefit those plants that still have antagonistic interactions, but it will be detrimental to mutualist plants since they will be less likely to interact with a partner and therefore less likely to reproduce. So we may expect a shift in the relative abundances of the two functional groups of plants in favour of the antagonists at low IR (see section 4.4.3).

Temporal variation

In general increasing HL increases the temporal variability of the dynamics. That is, communities are less stable in damaged landscapes. This result is only seen in the case of contiguous habitat destruction, as opposed to random destruction, and is discussed in more detail in section ?? where it was shown to be associated with changes in the distribution of interaction strengths. Also communities are less stable at lower IR. This fits with previous results. It has been shown that communities are very stable and resistant to HL at high IR (section ??). It has also been shown that they are unstable at zero IR, exhibiting community collapse (section ??). This suggests that the model has a stable and an unstable regime, and that there must be a transition between the two when moving from high to low IR. The right-hand column of figure 4.11 shows a signature of this. Interestingly the loss of dynamic stability is greatest for antagonistic communities and weakest for purely mutualistic communities. This suggests that mutualism has a stabilising effect on community dynamics. It appears to confer better dynamic stability in the face of HL and changing IR (but there are also more extinctions as discussed..).

Another interesting feature of the CoV plots is that the trends described above appear to be broken at very low IR and high HL, where an increase in stability is visible. One potential mechanism is that this is an averaging effect. If some communities are totally collapsing in this region they would exhibit stable dominance of plant species, which would

¹²This can be checked later.

¹³This proposed mechanism may be working in reverse in the $MAI=1.0$ communities.

4.4. EXPLORATION OF PARAMETER SPACE: HABITAT LOSS AND IMMIGRATION

contribute positively to average community stability. However it may be that this effect is due to another mechanism.

As mentioned previously the loss of dynamic stability is troubling since it calls into question the way that we calculate abundance metrics. Therefore the conclusions drawn in the following discussion may not be general and may not hold if the metrics were averaged over a number of iterations.(Dani: don't stress this here, put in limitations section.)

4.4.2 Example dynamics

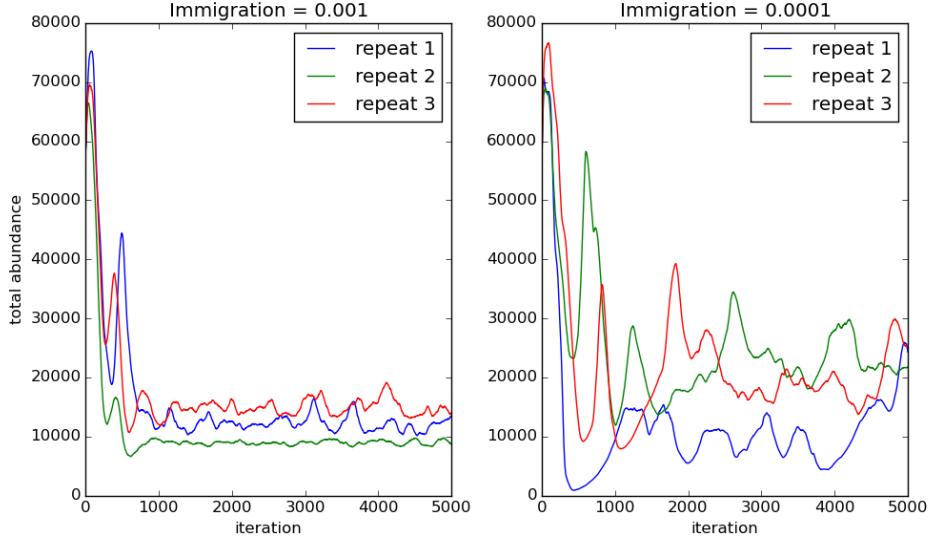


Figure 4.4: Temporal dynamics of the total biomass of communities over the course of six simulations. Each panels shows the dynamics for three distinct simulations, each in a different colour. The left panels shows communities with a high immigration rate, and the right panel for a low immigration rate. In all cases there is no habitat destruction $HL = 0$.

Figure 4.4 illustrates the loss of stability in passing from a high to a low IR regime. This transition was proposed in section ???. The dynamics of three example antagonistic communities are depicted for each regime. These communities were selected at random from the one hundred repeat simulations at these parameter values. Antagonistic communities are shown because the increase in temporal variability is greater for these than for those with mutualism (see figure 4.11).

In the high IR regime, shown in the left-hand panel of figure 4.4, we see that the total biomass of each community undergoes an initial transience followed by a period of relative stability. It appears that, during this second period, the system is undergoing stochastic fluctuations about its stable equilibrium¹⁴. In the low IR regime, shown in the right-hand panel, we see that the community biomass exhibits large scale fluctuations throughout the course of the simulations. It is not clear from inspection that the system is being perturbed about a stable equilibrium.¹⁵ It may be that the reduction in IR increases the length of the initial transience, and that the communities illustrated are yet to reach steady-state after 5000 iterations. Or it may be that these communities reach their steady-state, but that the

¹⁴Test for this?

¹⁵I would rephrase this. For example: there are different explanations for this pattern: (i) Explanation 1, (ii) Explanation 2.

4.4. EXPLORATION OF PARAMETER SPACE: HABITAT LOSS AND IMMIGRATION

stochastic fluctuations are amplified because the equilibria are less stable¹⁶.

Figure 4.5 shows example dynamics by trophic level of four antagonistic communities in the high and low IR regimes. The left-hand panels depict two communities in the high IR regime. Again the initial transience is followed by a period of relative stability, which is consistent across trophic levels. It is clear from these two plots that the positions of the system's equilibria and the size of the fluctuations about it vary between simulations.

The right-hand panel of figure 4.5 depicts two communities in the high IR regime. It is clear from inspection that the mean and the variance of the biomasses varies between trophic level, and between simulation. The lower plot shows dynamics dominated by species from the first trophic level, with large scale but decreasing amplitude fluctuations in the second trophic level. The upper plot shows perhaps even less stable dynamics with increasing amplitude fluctuations in the first and fourth trophic levels, and very low abundances in the intermediate trophic levels. In both simulations there are several instances where the biomass of an entire trophic level comes close to zero. However, as figure 4.11 shows, we should only expect around one extinct species at the end of a simulation at this IR. It must be that that immigration is preventing stochastic extinctions here¹⁷, by providing some buffering to populations at the low end of their biomass fluctuations and by rescuing those species that do go extinct¹⁸.

The breakdown of dynamics by trophic level demonstrates that the timing of measurement will affect the calculation of relative abundance metrics, and not just that of the aggregate community biomass. If the fluctuation in trophic biomass were more synchronous between levels, the timing of the measurement would be less significant. However the figure shows that even the ordering of trophic levels by abundance is dependent on time¹⁹. Therefore further analysis should attempt to remove this time dependence by averaging biomasses over a number of iterations. The plots suggest that the increase in community biomass at low IR (discussed in section ??) may be a genuine effect. However it is hard to determine the contribution of the increased fluctuations without averaging the abundances over time. (There may be other points in parameter space where it would be informative to plot the dynamics...e.g. high mutualism region, temporally stable) (Could plot biomass dynamics, averaged over replicates?)

¹⁶Further mathematical analysis to try and determine this? - Final chapter on model fitting?

¹⁷At this IR we would expect on average four immigrations per iteration, if the landscape were empty.

¹⁸It would be interesting to look at the breakdown of these trophic dynamics by species - e.g. how synchronous are the different species in the same trophic level with each other.

¹⁹This is beginning to look make the results seem invalid.

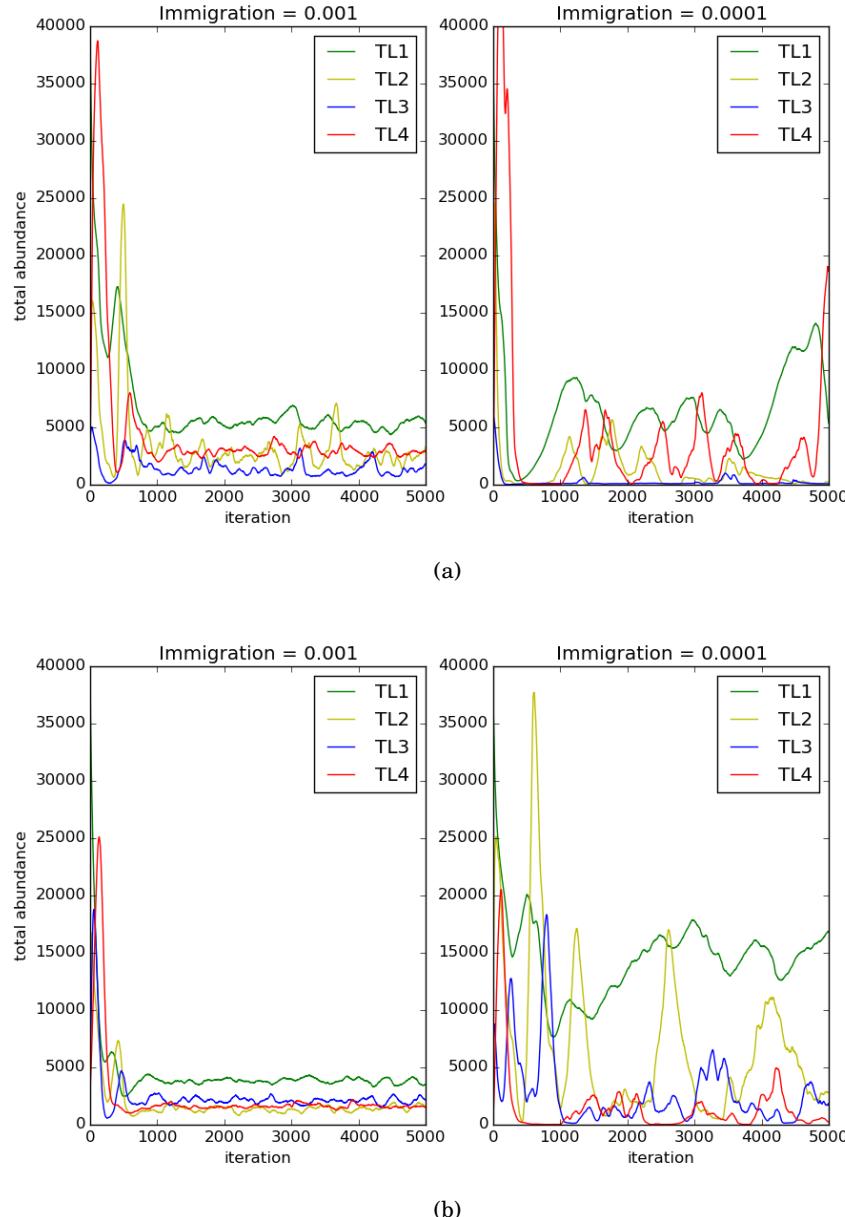


Figure 4.5: Dynamics from four individual simulation runs, with biomasses aggregated by trophic level. Each panel represents the dynamics of a single simulation run. In all cases the MAI ratio is 0.0, and there is **no habitat destruction** ($HL = 0$). The coloured lines represent the temporal dynamics of the biomass of each trophic level, as indicated in the legends. Two immigration scenarios are presented. **Left column: high immigration. Right column: low immigration.**

4.4.3 Relative abundances and abundance distributions

Contextualise - begin this section with the points made above that suggest useful to look at RADS.

Figure 4.6 shows the mean relative abundance of each trophic level for antagonistic and mutualistic communities, across the parameter space. For purely antagonistic communities the proportion of individuals in each trophic level varies strongly with IR and weakly with HL. At low IR antagonistic communities become dominated by plant species. This is in agreement with the mechanism proposed in section ??, whereby plants benefit from a scarcity of animal consumers at low IR. At high IR the distribution of biomass is much more even across trophic levels. In this region of parameter space the biomass of trophic levels one and four are roughly equal at around 30%, with the remaining 40% of the biomass split fairly evenly between trophic levels two and three. This biomass distribution is not necessarily unrealistic for a community in nature, however it does not conform to the classic *biomass pyramid* (see discussion in section ??). In fact the distribution at low IR is much closer to the standard pyramid.

Mutualistic communities ($MAI=1.0$) show much less variation in their trophic composition across the parameter space. The first two trophic levels are most abundant, with slightly more biomass in the first trophic level than the second. The third and the fourth trophic levels are much less abundant with around 20 – 30% of the biomass split fairly evenly between them. This distribution is remarkably constant over the parameter space. Only at extreme levels of disturbance ($IR = 0.0001, HL \geq 70\%$) do the communities begin to be dominated by plants.

Figure 4.7 shows the mean relative abundance of each functional group for mutualistic communities with $MAI = 0.5$ and $MAI = 1.0$. As expected purely mutualistic communities are dominated by functional groups two and four (mutualistic producer and animals) across the whole region of parameter space. Functional groups five and six do relatively better at high IR and low HL. Whereas at low IR and high HL the relative abundance functional group 1 increases significantly. This is an indication of the shift in favour of antagonists, suggested in section ??, due to the low biomass making it hard for mutualists to reproduce and less likely that plants will be eaten. The same patterns are seen in the case of $MAI = 0.5$, however the trends appear stronger since the relative abundances are less robust to changes in IR and HL.

4.4.4 Rank abundance distributions

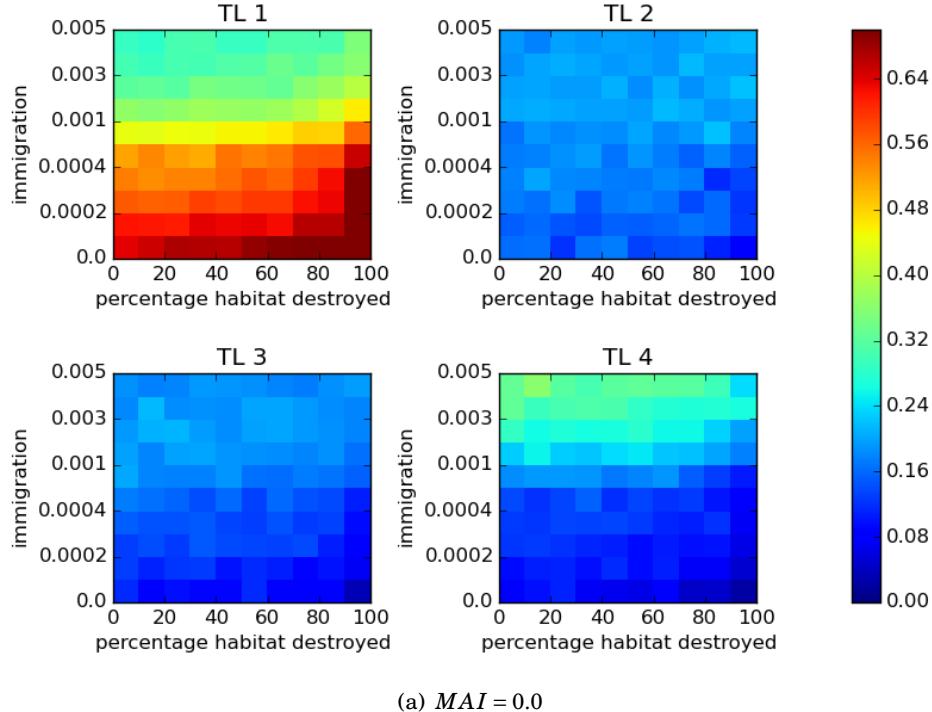
These results, as with the other should be recalculated using averaged metrics.

Figure 4.9 shows the mean rank abundance distributions for a range of IR and HL values. Communities with all three MAI ratios are shown in different colours. Across the parame-

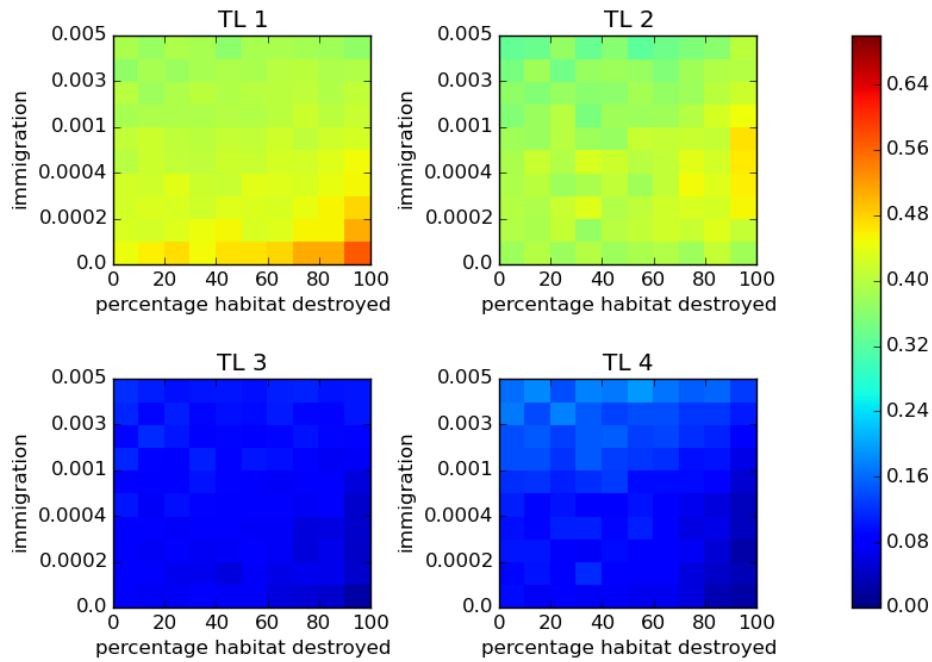
ter space mutualistic communities (blue) have less even distributions than antagonistic communities. This difference is more pronounced at low IR and high HL.

An interesting feature of the RADs is that some of them display an apparent discontinuity in the distribution. This is perhaps most pronounced in the bottom left panel of figure 4.9 ($IM = 0.0001, HL = 0, MAI = 1.0$). A sigmoidal shape is a feature of log-normal abundance distributions and is often observed in natural communities [REF]. However this extreme case does not appear to fit. What is driving this distribution? The ‘flat’ section of low abundance species could be those species whose presence is sustained only by continuous immigration and which are therefore present in roughly equal abundances?

4.4. EXPLORATION OF PARAMETER SPACE: HABITAT LOSS AND IMMIGRATION

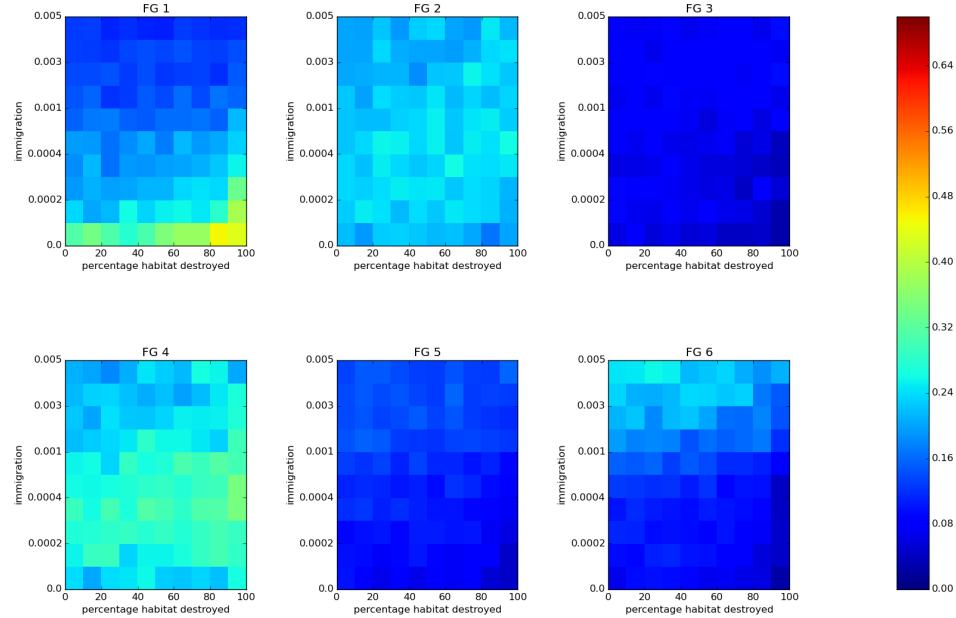


(a) $MAI = 0.0$

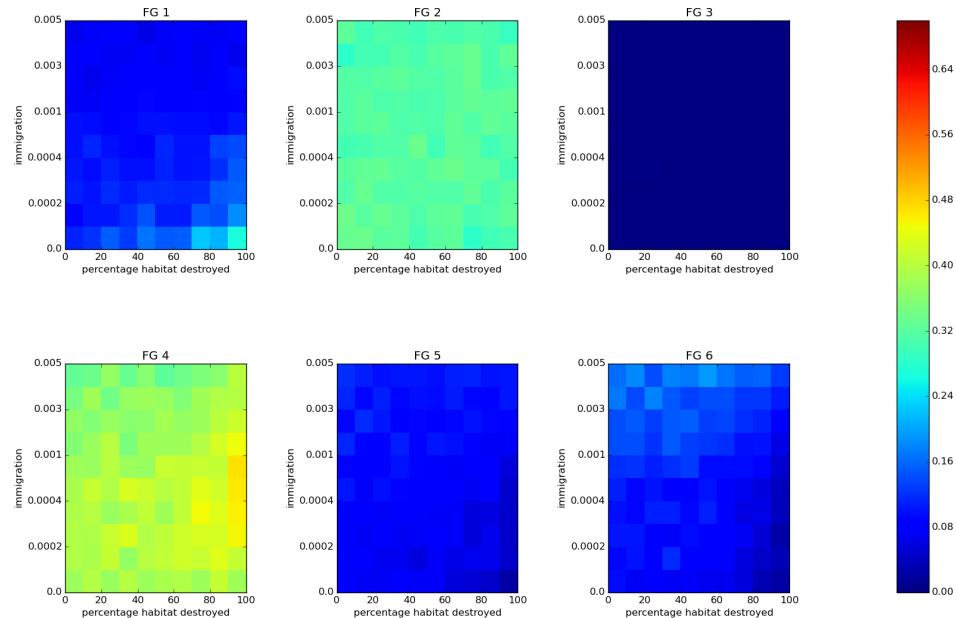


(b) $MAI = 1.0$

Figure 4.6: The relative abundance of species belonging to each of the four trophic levels. Above: $MAI = 0.0$. Below: $MAI = 1.0$. Each pixel on the heat maps corresponds to an average over one hundred repeat simulations at those parameter values. The abundances are measured at the end of each simulation.



(a) $MAI = 0.5$



(b) $MAI = 1.0$

Figure 4.7: The relative abundance of species belonging to each of the six functional groups. Above: $MAI = 0.5$. Below: $MAI = 1.0$. Each pixel on the heat maps corresponds to an average over one hundred repeat simulations at those parameter values. The abundances are measured at the end of each simulation.

4.5 Points for discussion (Rough Notes)

A comparison of the relative merits of being a mutualist versus a non-mutualist is worthwhile. Importantly it must be remembered that mutualistic interactions are also trophic interactions. In our case, energy is transferred from producer to animal. In nature for example the bee receives energy from the nectar, but also carries pollen to fertilise other flowers. So there is some loss/detriment to the producer as well as the benefit of reproduction (These mechanisms are in place in the model through the bioenergetic parameters. Traditionally, simulating mutualistic communities has failed because the simulations ended in an 'orgy of mutual benefaction'). It is an interesting strategy from an evolutionary perspective...discuss this (with relation to link switching)?

In the model species become mutualistic by having at least one of their links, in the antagonistic interaction network, switched for a mutualistic link. Table ?? shows the default parameter values used for most simulations. Lets consider the potential benefit of switching a single herbivorous link for a mutualistic link, for either party. If the plant is a non-mutualist it must impart 20% of its energy to the offspring when reproducing (this happens with a probability of 0.01 on each iteration). It is also subject to lose 70% when it is encountered by this herbivore. If it were to switch this herbivorous link for a mutualistic link it would only lose 25% of its energy in the interaction, and it would pass on a seed that is almost guaranteed²⁰ to generate an offspring. Therefore the cost of reproducing is slightly increased for a mutualist, but the cost of interacting with an individual from the trophic level above is dramatically reduced. There is an additional benefit that the mutualistic reproduction can occur over a greater distance. The net gain/loss of this change depends on the probability/rate of interactions. We should investigate this, however the results suggest that being a mutualist is of significant benefit to plants. (These mechanisms are in place in the model through the bioenergetic parameters. Traditionally, simulating mutualistic communities has failed because the simulations ended in an 'orgy of mutual benefaction')

Question: in the above analysis are mutualistic plants are relatively more abundant than non-mutualistic ones, except in the case of high habitat loss or low immigration (when there are few enough mutualistic partners that interactions become infrequent?)

For animals there is no cost to carrying and spawning the seed of their mutualistic partners. The only change in the switching of mutualistic links is the amount of energy that they receive from the interaction. During a herbivorous interaction, the herbivore takes 70% of the plant's energy, and assimilates it with an efficiency of 80%. Therefore it obtains 60% of the plant's energy. During a mutualistic interaction the animal-mutualist takes and

²⁰Really? We could look at how many mutualistic interactions lead to a new individual. It would only not occur in very crowded situations.

assimilates 25% of the plants energy. Therefore on an interaction by interaction basis there is a negative trade off for an animal in switching its link to mutualistic. However there may be an emergent benefit in that this type of interaction is much better for plants, therefore increasing the plant biomass and therefore indirectly benefiting animal (mutualists and non-mutualists?) due to the increased frequency of interactions (density of plants).

Mutualism in general stabilises dynamics, and leads to communities with more realistic biomass pyramids - i.e. dominated by the first two trophic levels, with fewer individuals in TL2/3.

It could be argued that the RADS are realistic, and the some immigration is a requirement to prevent stochastic extinction of the very rare species, which are found in nature. This begs the question as to what mechanism prevents their extinction in nature? And are they the most vulnerable to extinction?

4.6 Habitat loss with low immigration

4.7 Questions for Alan or Daniel

- Worried about general flow and structure of discussion. Feels like trying to present too much information all at once. How to not turn into a list of facts, where the relevance gets lost?
- Since the dynamics do not necessarily reach steady state should I re-do analysis with average over a time window? (We need a discussion section where the results are discussed in the context of current literature in the field, real-world communities, etc. Also, contextualizing (see a previous comment) is important so the reader does not feel like we include all these metrics because we can. For doing this it is always helpful to write down the main findings as bullet points and develop them; also, the limitations of the model should be presented here as well as the ways forward)
- Can I use "we"??
- Tense?
- Figure 1.1 summary heatmaps: too much information in one figure? (feels that way from discussion).
- OK to use plant, basal and producer interchangeably?
- The ability of the top predator to survive almost entirely on plant matter is troubling.
- Is it in fact OK to use biomass and number of individuals interchangeably? (We need a discussion section where the results are discussed in the context of current literature in the field, real-world communities, etc. Also, contextualizing (see a previous comment)

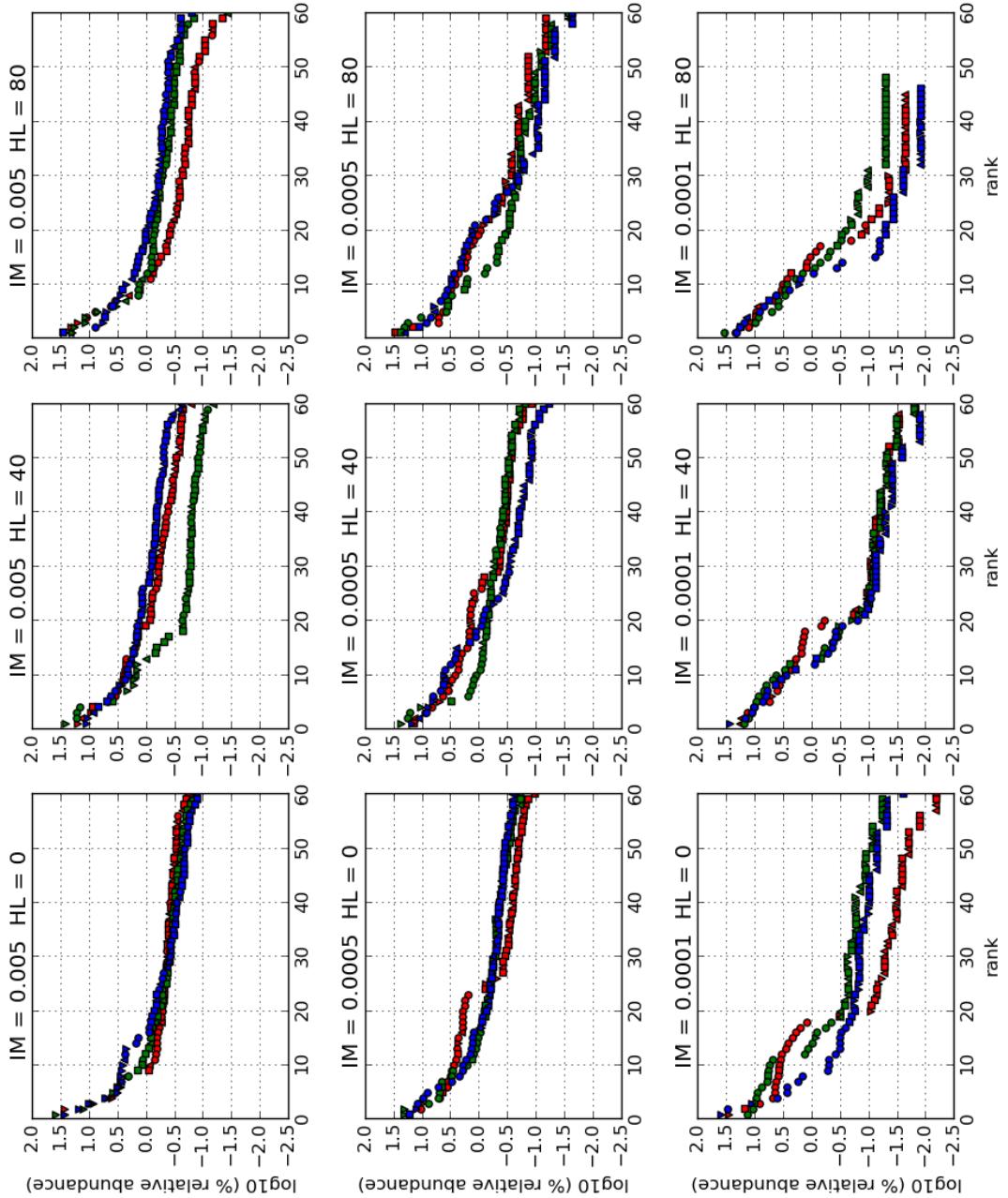


Figure 4.8: **Rank abundance distributions** for individual simulation runs, for nine different pair values of immigration rate and habitat destruction. Each distribution is for a single community at the end of an individual simulation run. The different colours correspond to different MAI ratios: red = 0.0; green = 0.5; blue = 1.0. And the different symbols correspond to different trophic levels: circle = 0; upwards triangle = 1; square = 2; downwards triangle = 3.

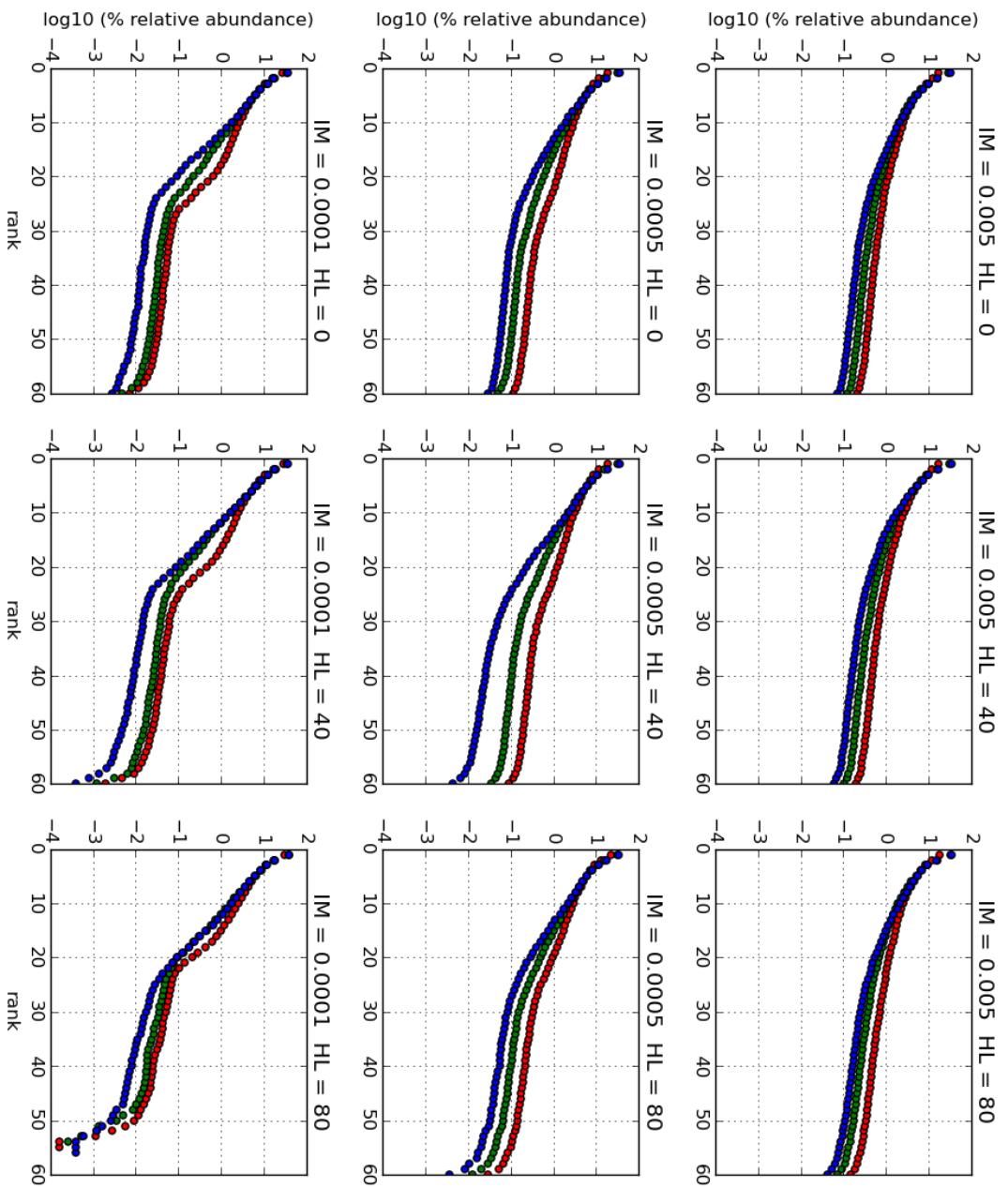


Figure 4.9: Average rank abundance distributions over one hundred simulation runs, for nine different pair values of immigration rate and habitat destruction. Each distribution is calculated using the mean relative abundance of the ranked species, averaged over the final abundances of one hundred repeat simulations. The different colours correspond to different MAI ratios: red = 0.0; green = 0.5; blue = 1.0.

is important so the reader does not feel like we include all these metrics because we can. For doing this it is always helpful to write down the main findings as bullet points and develop them; also, the limitations of the model should be presented here as well as the ways forward)

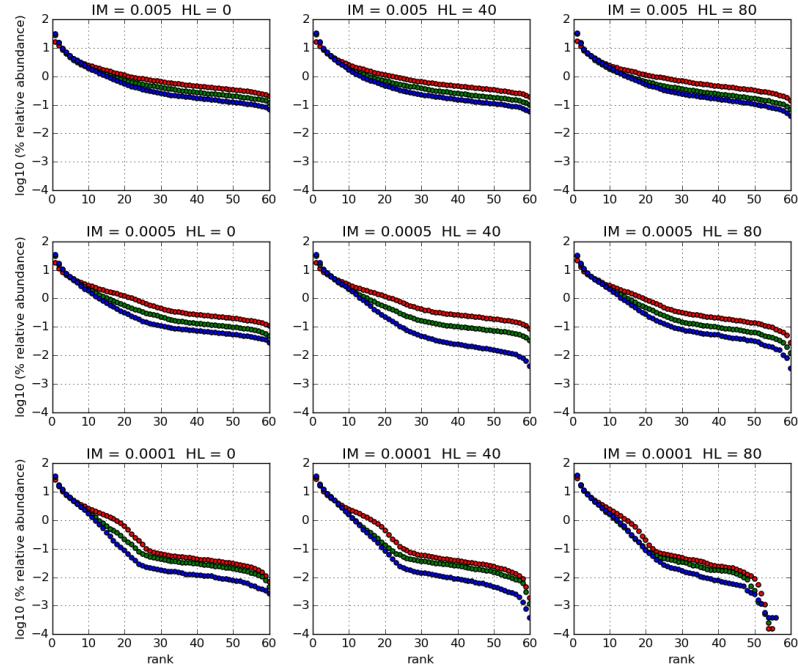
- Theme for discussion seems to be developing: point out a feature of the results, explain what could be causing it (in the model), relate this to other results. (should add to this - comment on how this may relate to the real world??) (Agree. The discussion section to summarize and explain, contextualize the results is needed.)

RADS: Good. As the main changes are observed for different IR it is good to include plots of abundance vs rank with the lines representing different IRs. Also, it could be nice to see if they deviate from a classical lognormal or lognormal family type distribution (which is usually found in nature)

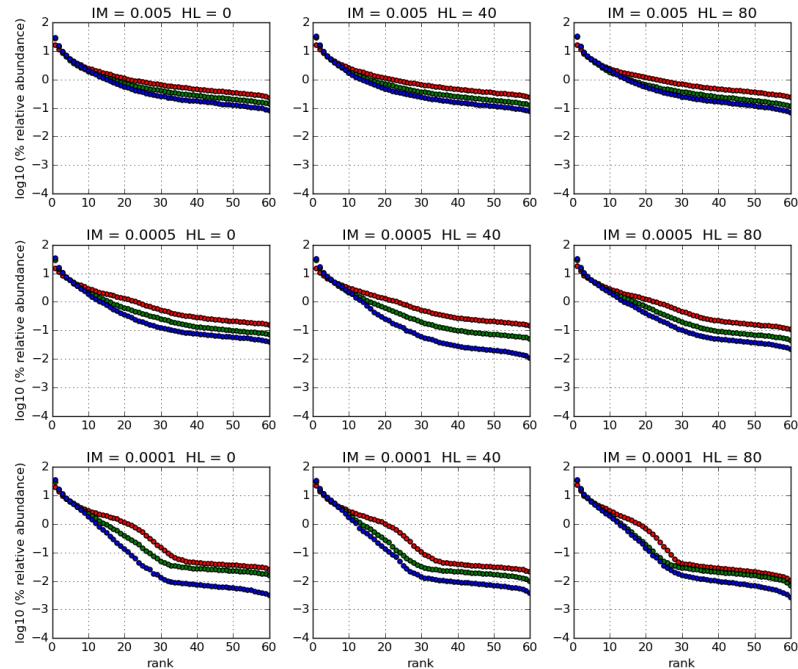
Are the legends correct? The plots next to this note are similar. I think they have different IR, right?

4.8 RADS with averaging

Here we show the difference between rank abundance distributions calculated using two different samplings from the simulations..



(a) From snapshot of final simulation state



(b) Abundance averaged over final 1000 iterations

Figure 4.10: Average rank abundance distributions over one hundred simulation runs, for nine different pair values of immigration rate and habitat destruction. Each distribution is calculated using the mean relative abundance of the ranked species. The different colours correspond to different MAI ratios: red = 0.0; green = 0.5; blue = 1.0.

4.9 New results: without vegetarian predators

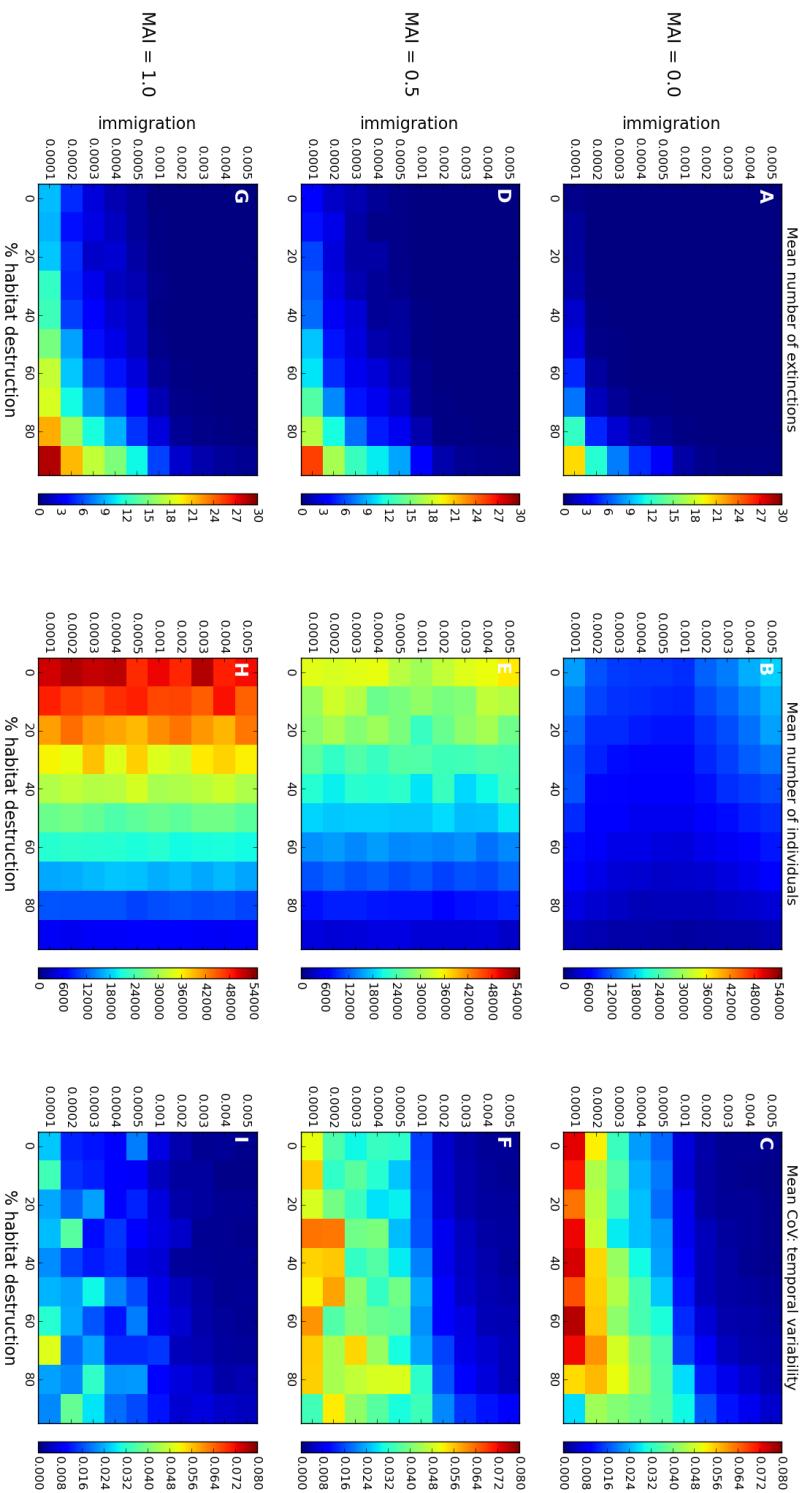


Figure 4.11: VEGETARIAN PREDATORS. Summary heat maps: Each heat map shows the value of a certain response metric across a 2-dimensional slice of parameter space. The parameters varied are immigration rate IR (y-axis) and percentage habitat destruction HL (x-axis). Each row of plots corresponds to a different MAI ratio as labelled. To construct the heatmaps one hundred repeat simulations were run for each combination of parameter values, with each simulation using a different underlying network. The mean value of the response metrics is taken over the hundred repeats. Therefore each pixel shows an estimate of the expectation value of the metric at those parameter values. The left column shows the expected number of species that are extinct at the end of a simulation; the central column shows the expected biomass (total number of individuals) at the end of a simulation; and the right column shows the expected temporal variability (coefficient of variation of total biomass) of the dynamics during the final thousand iterations of a simulation. The latter is used as a proxy for stability (see text).

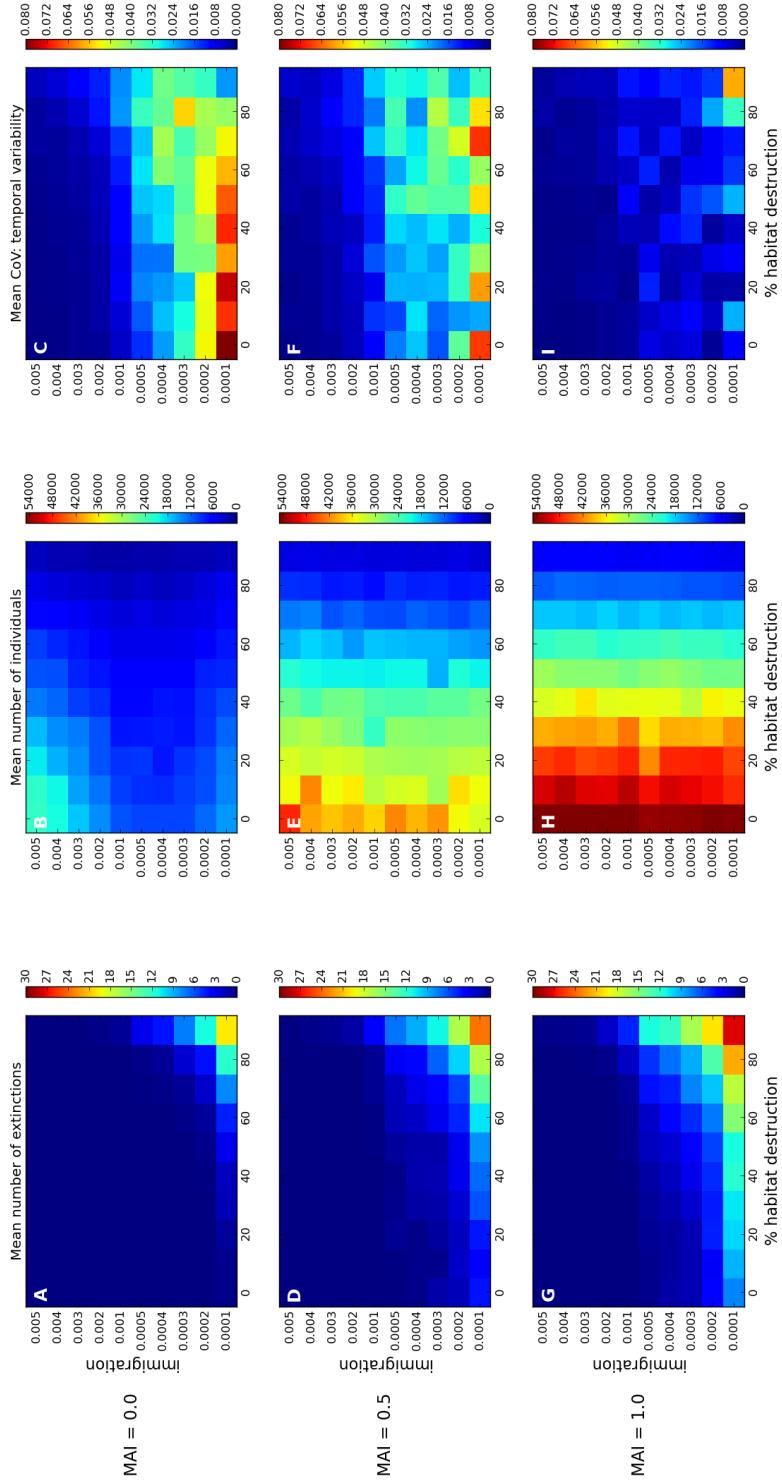


Figure 4.12: NO VEGETARIAN PREDATORS. Summary heat maps: Each heat map shows the value of a certain response metric across a 2-dimensional slice of parameter space. The parameters varied are immigration rate IR (y-axis) and percentage habitat destruction HL (x-axis). Each row of plots corresponds to a different MAI ratio as labelled. To construct the heatmaps one hundred repeat simulations were run for each combination of parameter values, with each simulation using a different underlying network. The mean value of the response metrics is taken over the hundred repeats. Therefore each pixel shows an estimate of the expectation value of the metric at those parameter values. The left column shows the expected number of species that are extinct at the end of a simulation; the central column shows the expected biomass (total number of individuals) at the end of a simulation; and the right column shows the expected temporal variability (coefficient of variation of total biomass) of the dynamics during the final thousand iterations of a simulation. The latter is used as a proxy for stability (see text).

CHAPTER



STATIONARITY

5.1 Motivation

In the previous chapter we saw that under certain circumstances the population dynamics generated by the IBM become highly variable (see section ??). This variability is measured by CoV, the temporal variability metric described in section ?? . In particular reducing the immigration rate (IR) is found to dramatically increase temporal variability (section ??). These immigration induced changes in CoV are much larger than those seen in chapter ??, which were due to habitat destruction alone (and were restricted to a high immigration regime - IR=0.005)¹. These findings bring into question certain assumptions implicit in the previous analyses, which are discussed below, and motivate the investigation of temporal variability which we undertake in this chapter.

The primary concern of increased variability in the population dynamics is that it brings into question the validity of our results. In the previous analysis and in [39] the results are either calculated from a ‘snapshot’ of the simulation state on the final (5000th) iteration or, in the case of the network metrics, are calculated from aggregate interaction frequencies during the final 200 iterations. For results obtained in this way to be reliable the simulation must reach a sufficiently steady state after 5000 iterations. In the high immigration regime (HIR) it may not be unreasonable to assume that this is the case. In the HIR simulations that we inspected the main transient dynamics are contained within the first 1000 iterations, which is followed by a period of relatively constant abundance (see for example figure 5.3). However we are now motivated to investigate ‘how constant’ the simulations are after the initial transience, and how

¹There is also the as yet not fully explained result from two chapters ago that habitat destruction either increased or decreased temporal variability depending on the type of destruction (random or contiguous). This chapter does not yet address this but probably should.

this might affect our results.

The topic of this chapter is also relevant to our understanding of ‘real world’ communities. The assumption that an ecosystem is in a *steady-state* has often been often made [8]. It is clear that ecosystems are dynamic, but they are remarkably robust and persistent in the face of environmental variability. For those who approach ecological theory from a dynamical systems perspective these properties are related to the dynamic stability of the system. Indeed to model an ecosystem as a dynamical system presupposes the existence of a stable equilibrium or attractor - otherwise the model could not explain observed species persistence. However it is not clear how to relate the concept of dynamic stability to the temporal variability of population dynamics, which is frequently used as a proxy for stability (there is an interesting and in depth treatment of this issue in [3]). In an extreme example there may exist a chaotic attractor, which is highly stable but which creates highly variable population dynamics. In such a case the steady-state assumption does not seem appropriate.

In this chapter we will see how our simulated communities relate to the theoretical debate on steady-states and temporal variability. Specifically we conduct a detailed analysis of the *stationarity* of population dynamics generated by the IBM model (section 5.2). We then move on to look at how our numerical results are affected by a lack of stationarity, and how this might be remedied (section 5.3). Finally we close the chapter with some methods that attempt to uncover determinism and signatures of chaos from noisy dynamics (section 5.4).

5.2 Second-order stationarity

We introduce here three tests for second-order (or ‘weak’) stationarity in time series. Second-order stationarity may be defined as the time invariance of the first and second moments of the data. Specifically Hsu [29] states that a random process $X(t), t \in \mathbf{Z}^+$ is second-order stationary if:

$$\begin{aligned}\mathbf{E}[X(t)] &= \mu \text{ (constant)}, \\ R_X(t,s) &= \mathbf{E}[X(t)X(s)] = R_X(|s-t|),\end{aligned}$$

where $R_X(t,s)$ is the *autocorrelation* function of the process. Conceptually these conditions state that a second-order stationary time series has constant mean and autocorrelation dependent only on time separation. From now on we will refer to this just as *stationary*. If the conditions are not met we call the time series *non-stationary*, and we cannot parameterise a constant distribution for the data. Non-stationarity may be due, for example, to a trend in the data or a change in the parameters of the data generator.

In our case the data generator is the IBM model and there are several possible causes of non-stationarity. It may be that there is no steady-state equilibrium in the model. For example the number of individuals may undergo a random-walk. From previous analysis this situation

seems unlikely, since we have observed what appear to be deterministic population cycles. However randomness has not been explicitly tested for. Another possibility is that a steady-state equilibrium exists, but that a long transience means it is not reached during the time frame of our simulations.

5.2.1 Tests for stationarity

We compare three different tests of stationarity: the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) [37]; the Augmented Dickey-Fuller (ADF) [54]; and the the Priestley-Subba Rao (PSR) [50] tests. These tests were chosen for their popularity in the time series literature. All three are implemented in the programming language *R* [52] - the former two in the package *tseries*, and the latter in the package *fractal*.

The ADF test has null hypothesis that the time series is non-stationary. The test models the data as an auto-regressive process (see section 5.2.4), and the null hypothesis is that this process has a *unit root*. The test produces a statistic that is negative. The greater the magnitude of the test statistic the more evidence there is to reject the null hypothesis in favour of stationarity.

The KPSS test complements the ADF test in that the null hypothesis is stationarity. The data is modelled as the sum of a random-walk and an error component, and tests the hypothesis that the variance of the random walk is zero. The test statistic is always positive, and the greater its magnitude the more evidence there is to reject the null hypothesis in favour of non-stationarity.

The null hypothesis of the PSR test is also that the series is stationary. The test is based on the idea that non-stationary processes have power spectra that change over time [50]. These are called *evolutionary spectra*. The test, as implemented in *R*, returns several statistics. We quote the ‘p-value for T’ which can be thought of as the confidence that the estimated spectral density functions are constant in time.

5.2.2 Characterising the tests

To understand the performance of the stationarity tests (section 5.2.1) we apply them to three example time series, which we refer to as HI, RW and NS. The first series, HI, is taken from a single IBM simulation run with high immigration rate (IR= 0.001), zero mutualism (MAI= 0.0) and otherwise default parameters (table ??). The series represents the total number of individuals of all species at each iteration. The simulation was run for 50,000 iterations, compared with the 5000 used in previous chapters. The increased length allows more time for the simulation to reach steady-state, and allows comparison of tests applied to different sections of the series. The first 1000 iterations were discarded, since these contain clearly transient dynamics (see figure 5.3B), leaving a time series of 49,000 points. A high immigration rate was chosen because it reduces the temporal variability of the dynamics, as was discussed in chapter 4. Therefore the HI series is more likely to be stationary than the output of a simulation with a lower IR (section 5.2.3).

The series RW and NS are chosen as a negative and a positive control respectively. Both have the same length as HI. RW is a non-stationary series generated by a one-dimensional *random-walk*, defined as:

$$(5.1) \quad x(t) = \sum_{t=1}^{i=1} Z_i,$$

where Z_i are independent random variables that may take a value of either -10 or $+10$, both with probability half. An ensemble of such random walks was generated and a single instance was chosen with mean and variance closest to the HI series. RW has a mean and standard deviation of 15525.2 and 1549.8 respectively, compared to 15915.8 and 1545.6 for HI. For comparison, NS is a stationary series generated by drawing each value independently from a normal distribution with mean and variance equal to that of HI. The three series are plotted in figure 5.1.

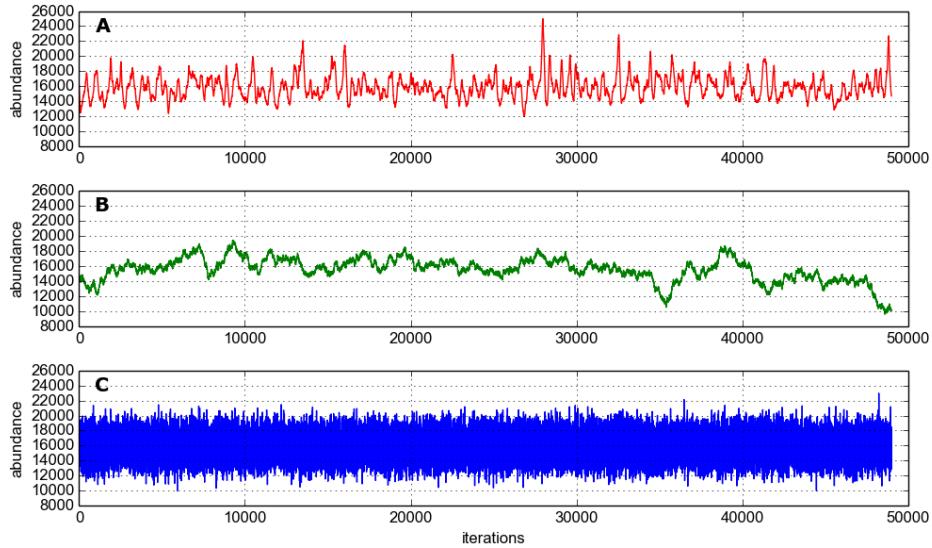


Figure 5.1: The three time series used to characterise the performance of the stationarity tests. The initial 1000 points removed such that all are 49,000 points long. (A) HI: total abundance dynamics of an IBM simulation with high immigration rate; (B) RW: a random walk without drift, as described in the text; and (C) NS: a series generated by independent sampling from a normal distribution.

Initially we apply the three stationarity tests to the entire length of the time series. The results are shown in table 5.2.2. ADF finds significant evidence that all three series are stationary, at 99% confidence. We may be suspicious of this result since we know that RW is generated by a non-stationary process. However this is a special case of a random walk, chosen from several thousand to closely match the mean and variance of HI. Therefore it may not be unreasonable that it can pass as stationary. The test statistic for ADF indicates that there is most evidence for NS to be stationary, followed by HI, then RW. The KPSS test ranks the series in the same order,

	A.D.F.		P.S.R.		K.P.S.S.	
	stat	p-value	stat	p-value	stat	p-value
HI	-15.401	<0.01	-	0.0004782808	0.5395	0.03277
RW	-4.0386	<0.01	-	0.9929773	18.7453	<0.01
NS	-37.5348	<0.01	-	0.811097	0.0466	>0.1

Table 5.1: Results of applying the three stationarity tests to the example time series shown in figure 5.1. P-values that indicate evidence for stationarity at 95% confidence are highlighted in green. The test statistics are also given for the ADF and KPSS tests.

based on the magnitude of the test statistic. According to this test NS is clearly stationary (accept h_0), and RW is clearly non-stationary (reject h_0 at 99% confidence), whilst HI is borderline. For HI we would accept the null-hypothesis of stationarity at 95% confidence, but reject it at 99%.

The PSR test gives unexpected results. It concludes that RW and NS are both stationary, whilst HI is non-stationary with a high degree of confidence ($p\text{-value} < 0.001$). In fact, according to PSR, RW is more likely to be stationary than NS. This result contradicts what we know about the series. Therefore we do not use this test in the analysis that follows. However the apparently erroneous result may contain interesting information about the HI series and the process that generated it (see discussion in section 5.2.4).

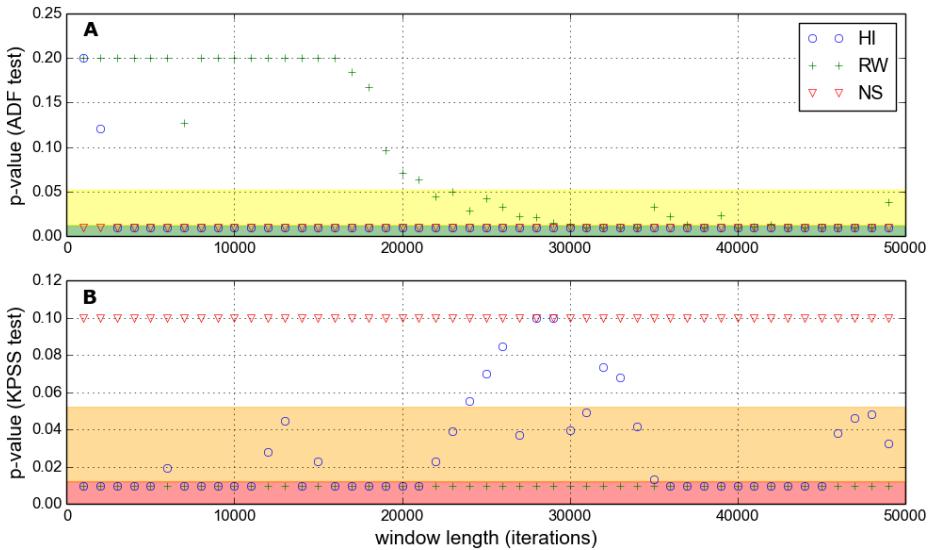


Figure 5.2: Two tests for stationarity applied to samples of varying size (window length). Samples are taken from the three time series (HI,RW,NS) shown in figure 5.1. All three time series contain 49,000 points. Sample windows begin at the first point and increase in length from 1000 to 49,000 points. Points plotted at 0.01 indicate p-values less than or equal to this. (A) ADF test, with p-values capped at 0.20. 95th and 99th percentile in yellow and green respectively, indicating significant evidence for stationarity. (B) KPSS test, with p-values capped at 0.1. 95th and 99th percentile in orange and red respectively, indicating significant evidence for non-stationarity.

Having discarded the PSR test, we now apply ADF and KPSS to samples of varying sizes, taken from the three series (HI,RW,NS). Sampling begins at the first point of the series and takes consecutive points up to the desired length of samples. Sample lengths range from 1000 to 49,000 data points. Again, as we saw in table 5.2.2, the two tests perform differently. The KPSS test correctly identifies RW and NS as non-stationary and stationary respectively, for all sample sizes. This is shown in figure 5.2B. The ADF test (figure 5.2A) correctly identifies NS as stationary for all sample sizes. For short sample sizes it also correctly identifies RW as non-stationary. However, for sample sizes much above 20,000, ADF finds significant evidence that RW is stationary at 95% confidence. This is an interesting result. Although RW is generated by a non-stationary process, it appears to fool the ADF test by staying ‘stationary enough’ over many time points.

There is mixed evidence for the stationarity of HI, as shown in figure 5.2. ADF, for all sample sizes above 2000, finds significant evidence that HI is stationary. Whereas KPSS, on the whole, gives significant evidence that HI is non-stationary - There are only seven cases where there is insufficient evidence to reject the null hypothesis that the HI series is stationary, and these occur at sample sizes between 24,000 and 34,000. From these results it appears that the KPSS test is a stricter test of stationarity, and is less sensitive to the size of the sample. Although it appears that the ADF test is biased in favour of stationarity, it does order the series correctly in the above examples and is a useful complement to KPSS. Also it may be that the sensitivity of ADF to sample length is useful, since processes may appear stationary/non-stationary at different scales.

We consider the possibility that the method of sampling from the time series affects the results of the stationarity tests. For example samples taken near the beginning of an IBM simulation run may be more likely to give the non-stationary series because of transient dynamics. Alternatively a non-stationary data generator may produce sections of time series that appear stationary purely by chance. This sensitivtiy to sampling is investigated by *reversing* the time series and repeating the above analysis. For HI, RW and NS we see no qualitative change in the results presented above. We also scan sampling windows of fixed length along the series to look for time dependence in the test results. The time at which samples are taken appears to make no qualitative difference, and there is no systematic change in the results that would suggest the simulation becomes more stationary the longer it is run².

HI simulation. We now focus on the simulation data used to generate HI and look in more depth at whether this dataset can be considered stationary. We use the same two tests, ADF and KPSS, for the stationarity of univariate time series. Since our abundance vector is 60-dimensional ($N = 60$ species), it is necessary to perform some manipulation to get a time series we can test. Previously we used the total number of individuals as our time series. However simply summing over species (l1-norm) is not necessarily the most informative metric to use. One possible issue is that the phase differences between species oscillations that we would expect

²Note that this is not necessarily the case when we move to lower IRs.

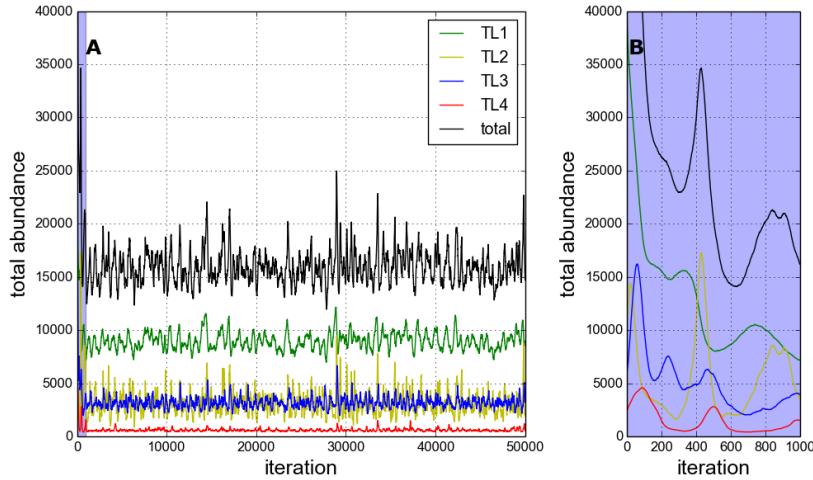


Figure 5.3: Dynamics for the HI simulation, broken down by trophic level ($TL_1 - 4$). Abundance is measured by the number of individuals. (A) The whole simulation run of 50,000 iterations. (B) Enlargement of first 1000 iterations, showing transience.

due to trophic interactions (see chapter ??) may mean that temporal variability is cancelled out when aggregating abundances in this way. It is possible that simulations which appear stationary according to some aggregate metric (e.g. total number of species) may have non-stationary underlying dynamics. This suggests that it is most informative to consider stationarity at the species level. We also consider the stationarity of abundances by trophic level, as an alternative aggregate metric.

The dynamics of the HI simulation are aggregated by trophic level to create four new time series $TL_1 - 4$. These *trophic dynamics* are plotted in figure 5.3. The initial period of transience is expanded in panel B. As in the previous analysis this part of the time series (first 1000 iterations) is removed. The ADF and KPSS tests are applied to the four trophic series separately and the results are shown in figure 5.4. According to ADF all trophic levels are stationary for sample sizes greater than 4000. TL_1 appears to be least stationary according to ADF, requiring a sample size of at least 4000 before to reject the null hypothesis at 95% confidence. According to KPSS TL_1 is non-stationary for all sample sizes, whilst TL_2 and 3 are stationary for sample sizes above 6000 and 1000 respectively. KPSS gives mixed results for TL_4 , with no clear dependence on sample size. It is hard to reconcile these results with an observation of the dynamics in figure 5.3, indicating the usefulness of the statistical tests. It may be informative to consider if there are general trends in the stationarity of trophic levels.

The dynamics of all the species belonging to each trophic level are plotted in figure 5.5. It is clear here that the community is dominated by a few abundant species, mainly in the lower trophic levels, with a large number of relatively scarce species. This agrees with the rank abundance plots from chapters ??, and with the long tailed distributions seen in real world

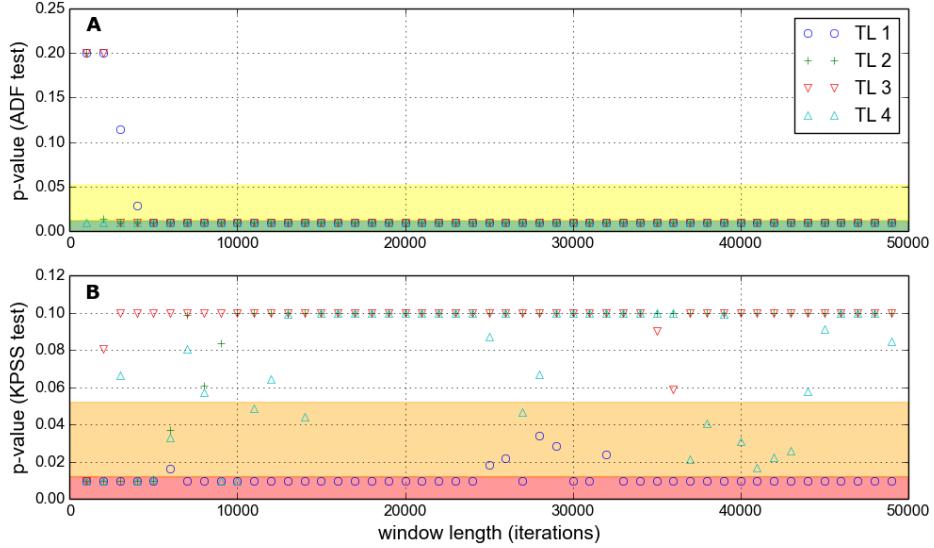


Figure 5.4: Similar to figure 5.2, but here the tests are applied separately to each trophic level of the HI simulation. The four time series (TL1 – 4) represent the total number of individuals belonging to each trophic level at every iteration.

communities. It also appears from this figure that the more abundant species exhibit large amplitude oscillations in their dynamics. This leads us to hypothesise that the most abundant species may be non-stationary, whereas the least abundant species may be stationary. We test this hypothesis by applying the ADF and KPSS tests to the three most abundant and three least abundant species in the HI simulation. Species are selected based on their average abundances over the whole simulation (minus the intial transience).

We see from figure 5.6 that all six species are stationary according to ADF, given sufficiently large sample size. However the sample size for all three of the abundant species to be stationary is greater (panel A: $\geq 9,000$ points) than for the three least abundant species (panel C: $\geq 2,000$ points). This suggests that the most abundant species are indeed ‘less stationary’ than the least abundant species. The KPSS test supports this conclusion. KPSS finds that the three least abundant species are stationary above samples sizes of $\sim 18,000$, whereas two of the most abundant species are non-stationary for almost all sample sizes. Inspecting the dynamics in figure 5.5 we see that these non-stationary species are those with largest amplitude fluctuations in their abundances.

In general we conclude that the choice of metric used to generate the time series does affect the conclusions about stationarity. Overall we cannot be confident that the HI simulation is stationary, based on the results presented above for species, trophic and total abundances. This is largely due to the apparent strictness of the KPSS test. Considering species dynamics individually is the most informative. It allows for the possibility that some species abundances may be more variable than others, and information is not lost by aggregating. In the following analysis we

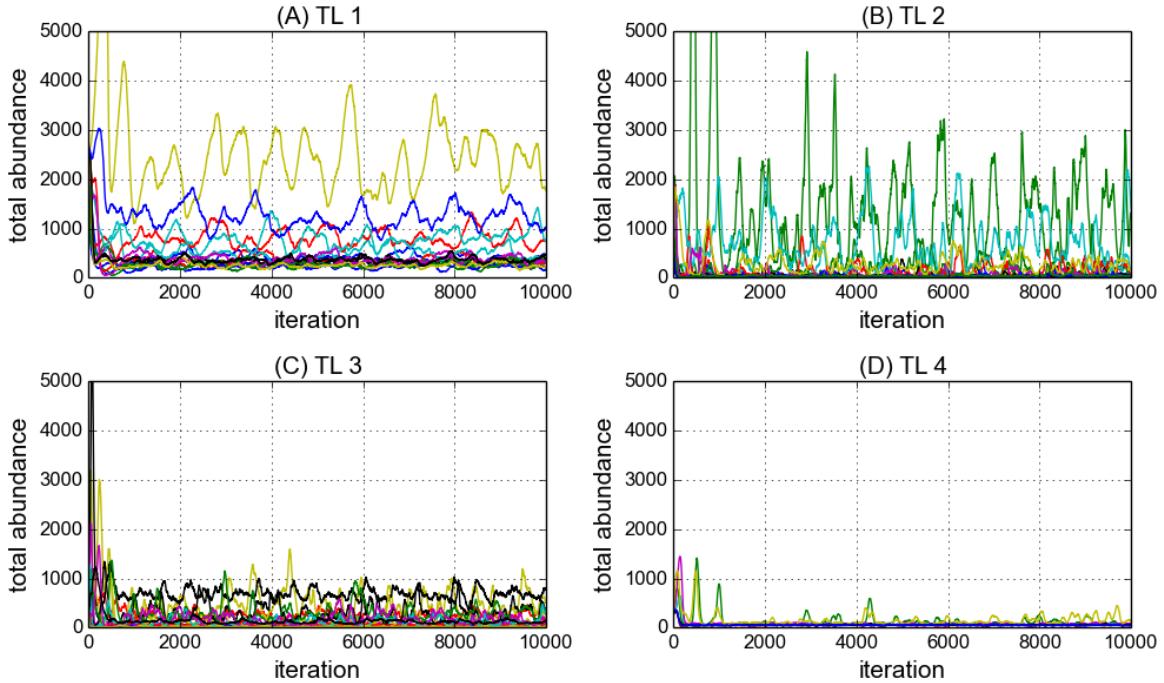


Figure 5.5: Dynamics of every species in the first 10,000 iterations of the HI simulation, broken down by trophic level. Panels (A)-(D) show all the species belonging to each trophic level (TL1 – 4).

propose that stationarity tests should be applied to species dynamics, and then the number of stationary species (NSSP) used as an aggregate statistic. If NSSP equals the total number of species, then the community dynamics is fully stationary (according to the test used).

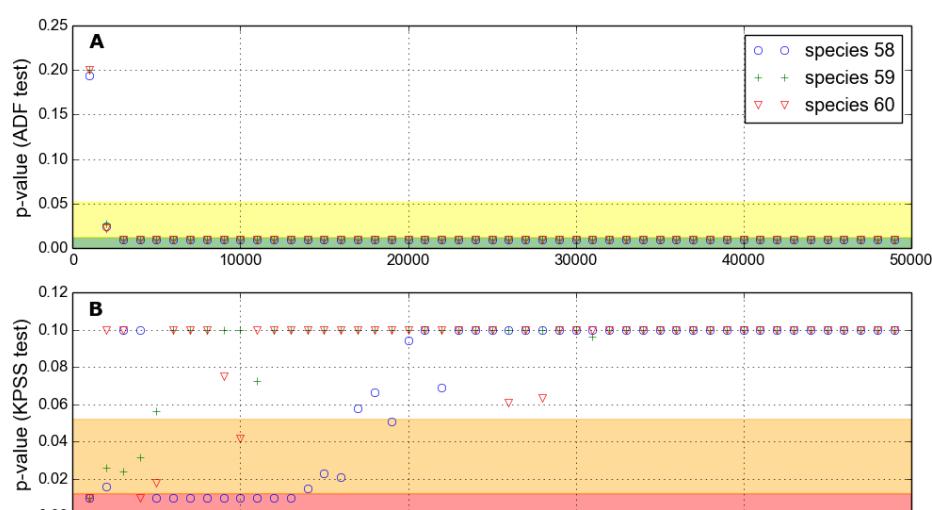
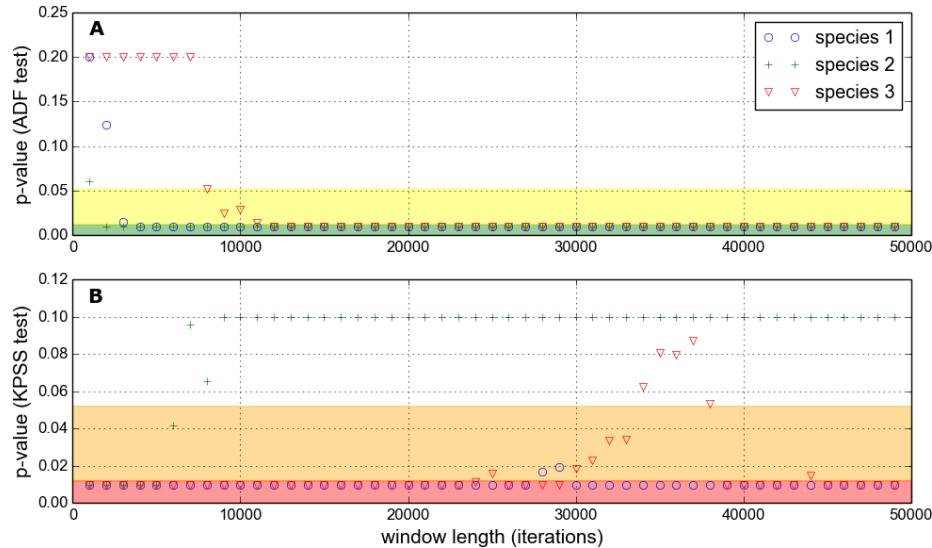


Figure 5.6: Similar to figure 5.2, but here the tests are applied separately to individual species from the HI simulation. (A) The abundance time series of three species with highest average abundances. (B) The three species with lowest average abundance.

5.2.3 Stationarity results

We now pursue a general investigation of the stationarity of communities simulated with the IBM model. First we test the stationarity of three ensembles of new simulation runs (figure 5.7). Secondly we test the simulations used to generate the results presented in the previous chapter (figures 5.8,5.9). The new simulations use default parameters, unless otherwise specified, and are run for 50,000 iterations. Two ensembles of 100 simulations each are run for high ($IR= 0.001$) and low ($IR= 0.0001$) immigration rates, which we refer to here as the HI and LI ensembles respectively. A third ensemble of 50 shorter simulations (10,000 iterations) is run, at high immigration rate, using a fixed interaction network, which we call the NM1 ensemble. All networks are generated using the niche model as described in section ??, with zero mutualism ($MAI=0.0$). Each simulation uses a uniquely generated network, except for those of the NM1 ensemble. Stationarity testing is done using the ADF and KPSS tests characterised in section 5.2.2 above. As standard the initial 1000 iterations are discarded in an attempt to remove transience. The tests are then applied to a sample taken from the abundance time series of each species. The results presented in this section give the number of stationary species (NSSP) in the community, at the 95% confidence level.

As the length of the samples taken from the abundance time series increases, the average NSSP also increases. This is true of both tests and for all three ensembles, as we can see from figure 5.7A. According to ADF all species are stationary, on average, for sufficiently large sample length. The required length of sample is larger for the LI ensemble than for HI. For KPSS, although the NSSP does increase with sample length, it is not clear that it will asymptotically approach 60 species in the limit of many iterations. The average NSSP at 49,000 sample points is just under 40 and just over 20 for HI and LI ensembles respectively.

To check the time dependence of stationarity (i.e. are species more likely to be stationary after many iteration of a simulation?) samples of length 3000 were taken from different points along the time series. From figure 5.7B we can see that there is no clear trend in stationarity over 49,000 iterations. The average NSSP is almost the same whether the sample is taken from iterations 1000-3000 or 46,000-49,000. This result also holds for windows of different length, which are not plotted here.

On average we see that the LI ensemble is less stationary than the HI ensemble. This we expected from the results of chapter ?. However we cannot be confident that either ensemble contains communities with stationary species distributions. This may be problematic for the interpretation of our results, and we discuss this further below. Interestingly the NM1 ensemble gives very similar results to the HI ensemble. This may be because we have accidentally chosen NM1 to closely resemble the average of this ensemble (see both panels of figure 5.7). Alternatively it may be that stationarity of the simulation output is not dependent on the interaction network structure. Again, we will return to this issue in what follows.

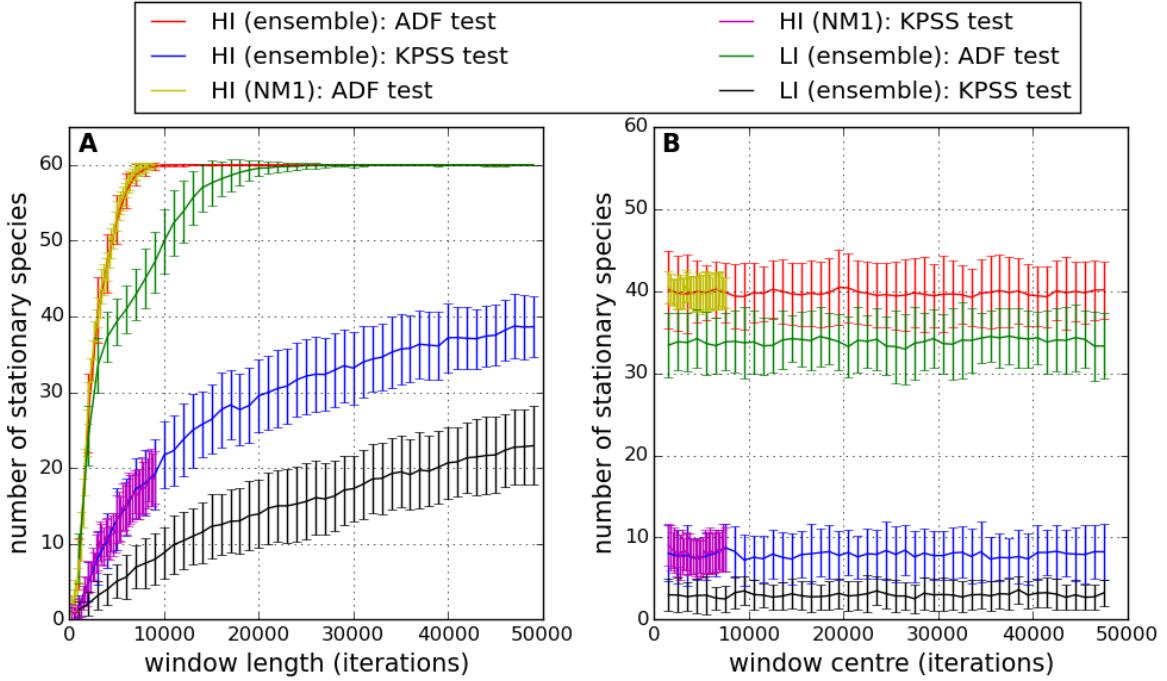


Figure 5.7: The number of stationary species (NSSP) according to the two stationarity tests (ADF and KPSS), averaged over three different ensembles of simulations: HI(ensemble); HI(NM1) and LI(ensemble) as described in the text. The first two are high immigration runs, whilst the latter is low immigration. Solid lines indicate the mean results for the ensemble, and error bars indicate ± 1 standard deviation from the mean. (A) Each species abundance time series is sampled with a window of increasing length, as in figure 5.6. (B) Each species series is sampled with a window of length $wl=3000$, which is scanned along the series as in figure ???. For both tests results are interpreted at the 95% confidence level.

Previous simulations. The simulations from chapter ?? are tested for stationarity. All simulations are 5000 iterations long. The initial 1000 iterations are discarded and the ADF and KPSS tests applied, species by species, to the remaining 4000. Figure 5.8 shows the average NSSP across the region of parameter space investigated, for three MAI ratios ($MAI = 0.0, 0.5, 1.0$). The results are qualitatively the same for both tests, although NSSP is higher for ADF than for KPSS as expected. On average reducing IR reduces the NSSP. A weaker effect, but still visible is that increasing HL reduces the NSSP. Most striking is the effect of MAI ratio on stationarity - the average NSSP is greater across the whole parameter region at $MAI = 0.0$ than at $MAI = 1.0$, with $MAI = 0.5$ in between the two. Increasing mutualism also appears to reduce the dependence of NSSP on IR. Figure 5.9 summarises these trends using cross sections taken from the heat maps in figure 5.8, with error bars added. It is clear that there is high variability across replicate simulations, and this variability appears to be greatest for high mutualism ($MAI = 1.0$).

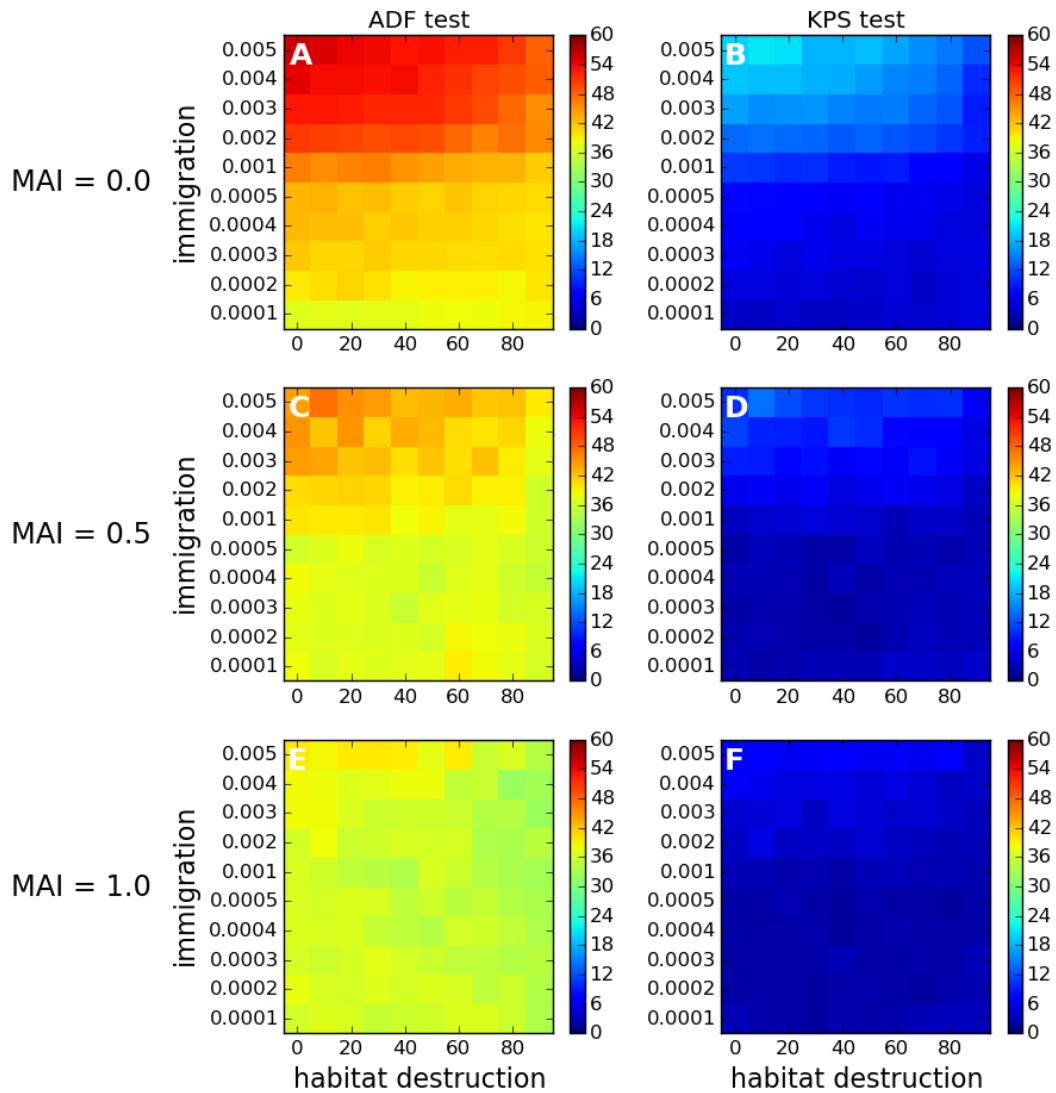


Figure 5.8: The average number of stationary species (NSSP) according to the two stationarity tests (ADF and KPSS), across the slice of parameter space explored in chapter ???. All simulations are 5000 iterations. Tests are applied to final 4000 iterations.

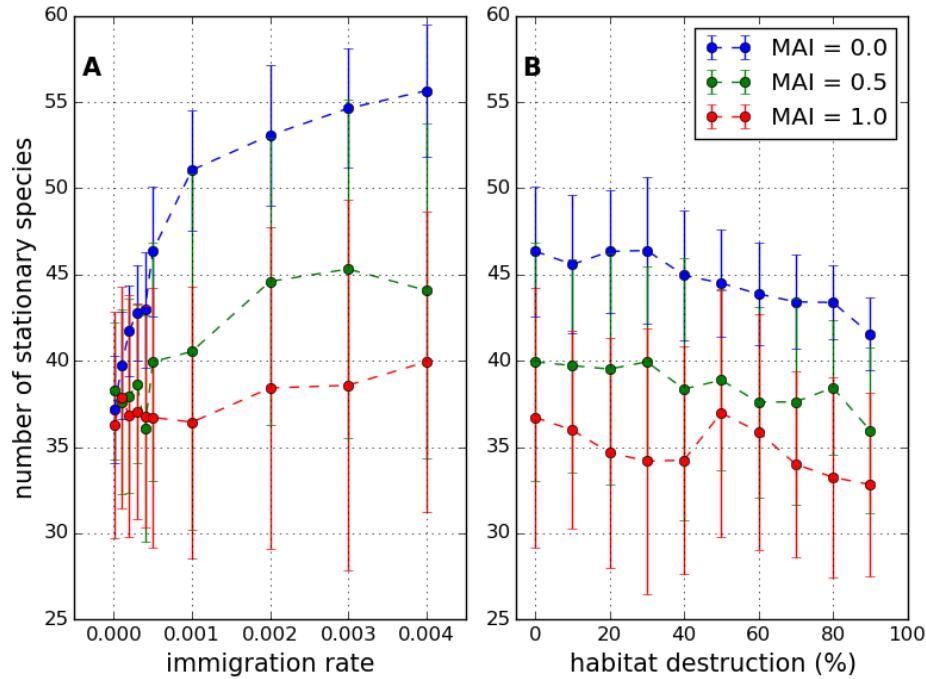


Figure 5.9: The number of stationary species (NSSP) according to the ADF test. The points show mean numbers and error bars show \pm one standard deviation. Tests are performed on the same simulations depicted in figure 5.8. (A) Plotted against immigration rate (IR), with zero habitat destruction. (B) Plotted against habitat destruction, with IR = 0.001.

5.2.4 Discussion (“skeleton”)

(From here on the document is not complete. This discussion skeleton synthesises the ideas about stationarity, before the narrative turns to look at how results are affected by temporal variability (section 5.3), and then to test for determinism and chaos (section 5.4).)

Main conclusion: communities not guaranteed stationary, even in high immigration regime. Most important question - how does this affect our results? Second most important question: why are they not stationary? Hypothesis: deterministic chaos? Alternative hypothesis: stochastic fluctuations about a stable equilibrium (but is this deterministic? and why would this not appear stationary?)

Importantly there is no evidence that the simulations are getting more stationary as they go on (i.e. 5000 iterations is probably enough) This means we do not need to throw away all previous results. But may need to reconsider how to calculate them.

Also to discuss:

- Stationary in real world communities (plankton, and look for more examples).
- Other possible tests for stationarity (parametric vs non-parametric, are the tests we used

OK??)

- PSR test: time dependent frequency spectra, possible use of wavelets (signature of aperiodic dynamics??)

5.3 Reliability of results

Since we cannot be certain over stationarity, especially within 5000 iterations with variable IR!..we now look at what this means for reliability our results..

5.3.1 Abundances

This analysis uses two metrics, both of which require the hypothesis that each species has an equilibrium abundance or long-term average abundance, which is well approximated by the mean abundance between 1000 and 50,000 iteration in the long simulations. (This is related to the idea of repeatability, which we look at below - is this long term average the same...RAS etc. for NM1)

Ideally we need to look at HI ensemble for IR=0.005 to justify original ‘snapshot’ sampling - ask Alan.

**Need a plot to show how MRE is calculated..

In this section we look at how reliable results are, how sensitive they are to different ways of measuring, and different lengths of averaging.

At the end of the section we also look at repeatability using the ensemble NM1. And introduce RAS spectra.

5.3.2 Variability: CoV

Here we look at the CoV - how does it depends on the length of window used? 200 was used initially..

We may also look at CoV (or stationarity again) versus mean species abundance - do the more abundant species have higher temporal variability?

5.3.3 Repeatability

<<< HEAD

Repeats communities - do they always do the same things...NM1 ===== Repeat communities (simulate with same interaction networks) - do they always do the same things...NM1 >>>
205e25b4ac2e2d84c24bee8ea4a528f60ad29aaa

5.4 Determinism and Chaos

We use two tests for determinism....cite cite...

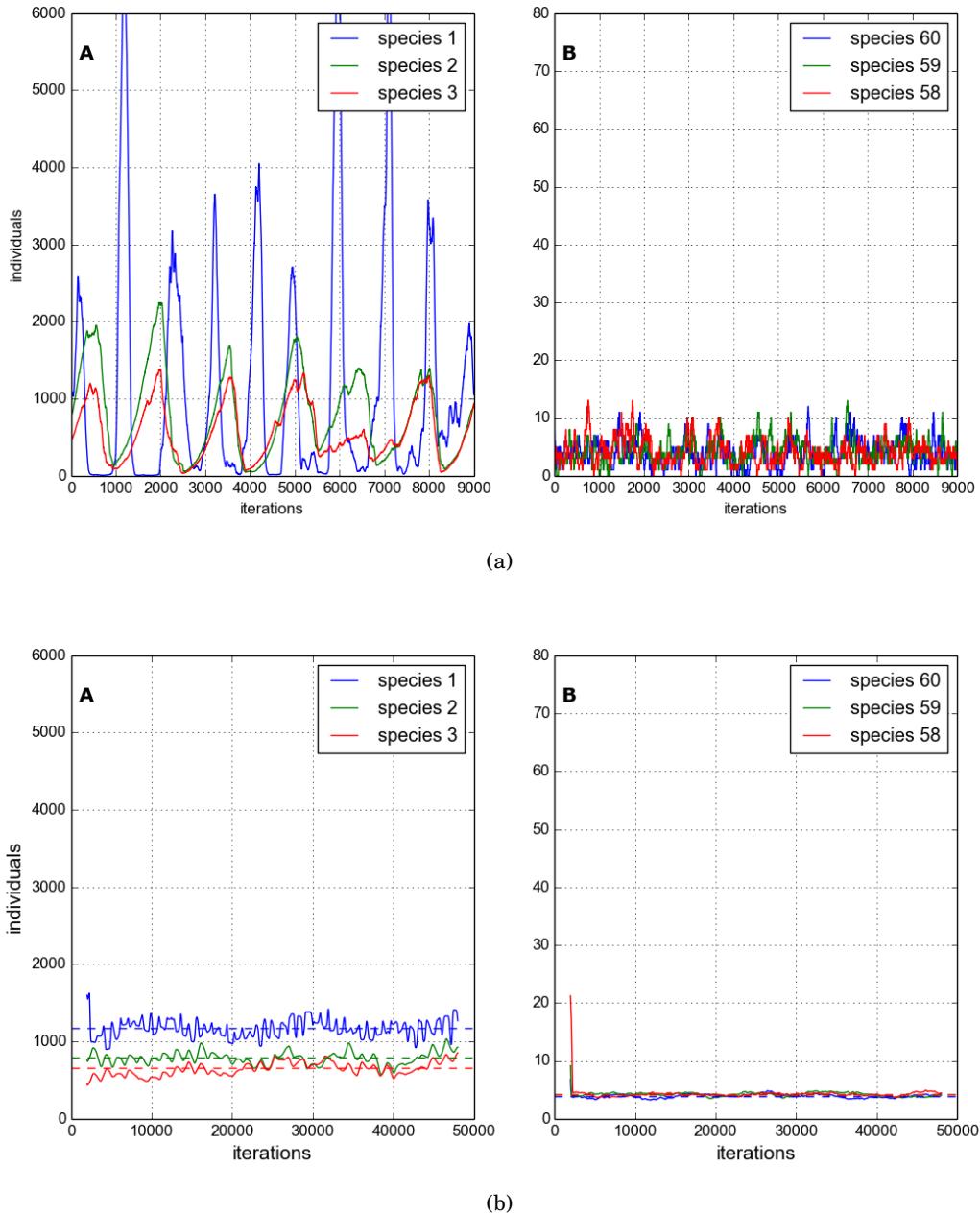


Figure 5.10: Example of speices dynamics with and without averaging.

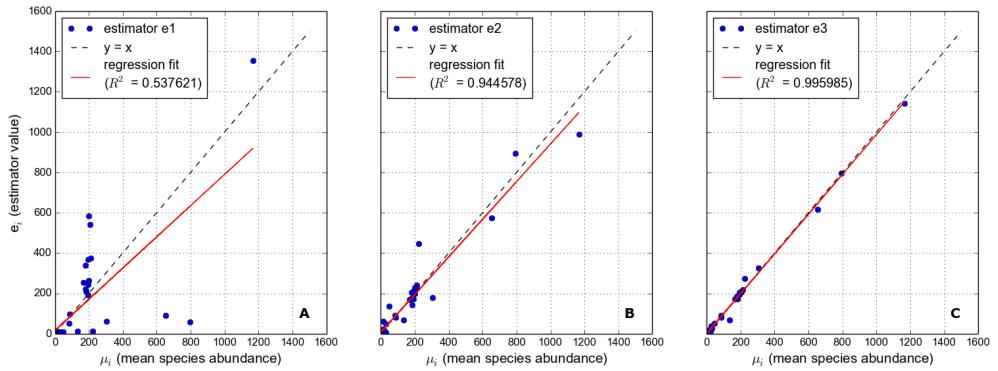
Chaos - Although in many cases a statistical steady-state appears to be reached, there are complex dynamics and fluctuations within that state (see section ??). Here we look at if these are due to noise or deterministic dynamics. We follow work presented by Saul in his PhD thesis [28]. We also draw inspiration from the demonstration that plankton communities may undergo chaotic dynamics - [5], and their focus on the Lyapunov exponent.

5.5 Discussion

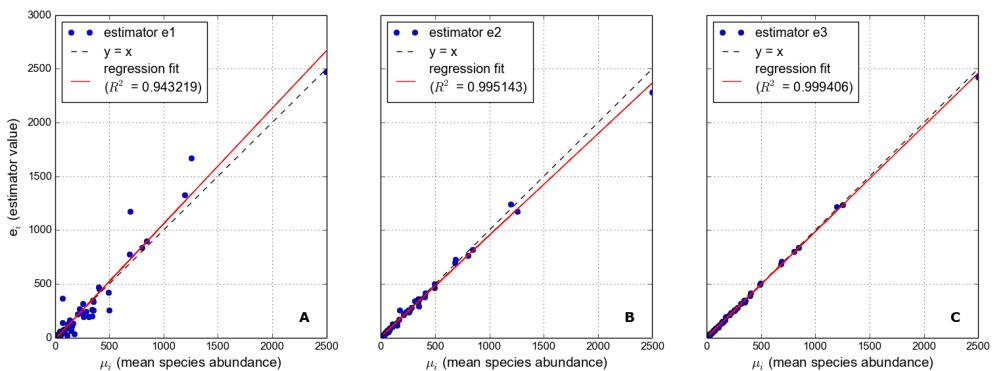
This behaviour may or may not be seen in real communitites - chaotic dynamics have been demonstrated in plankton, how about terrestrial ecosystems? HOwever we come back again to limitiations - snapshot measurements are taken - with replicate in time. Average over these? Check for differences between them - what is the actual procedure? Can we comment here?

Computationally we should perhaps compare the appraoches of taking snapshots and averaging over many iterations...DISCUSS WTIH ALAN.

Other question - does it reach the same steady-ish state every time? Is it always the same species that dominate/just bubble along.



(a) HI ensemble



(b) LI ensemble

Figure 5.11: Example of linear regression with the three estimators.

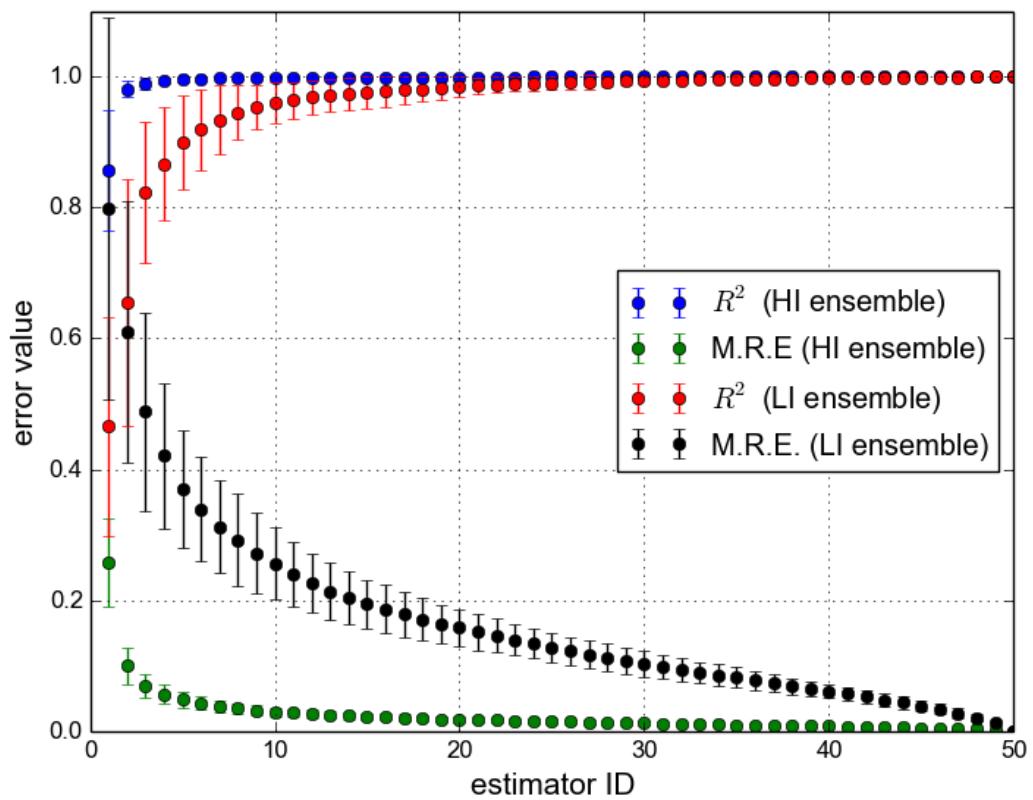


Figure 5.12: Abundance estimator performance - use this to justify the way we take our measurements from simulations! The first estimator is a ‘snapshot’ at 5000th iteration, the rest are averages of windows ranging from length 1000 to 50,000 in steps of 1000.

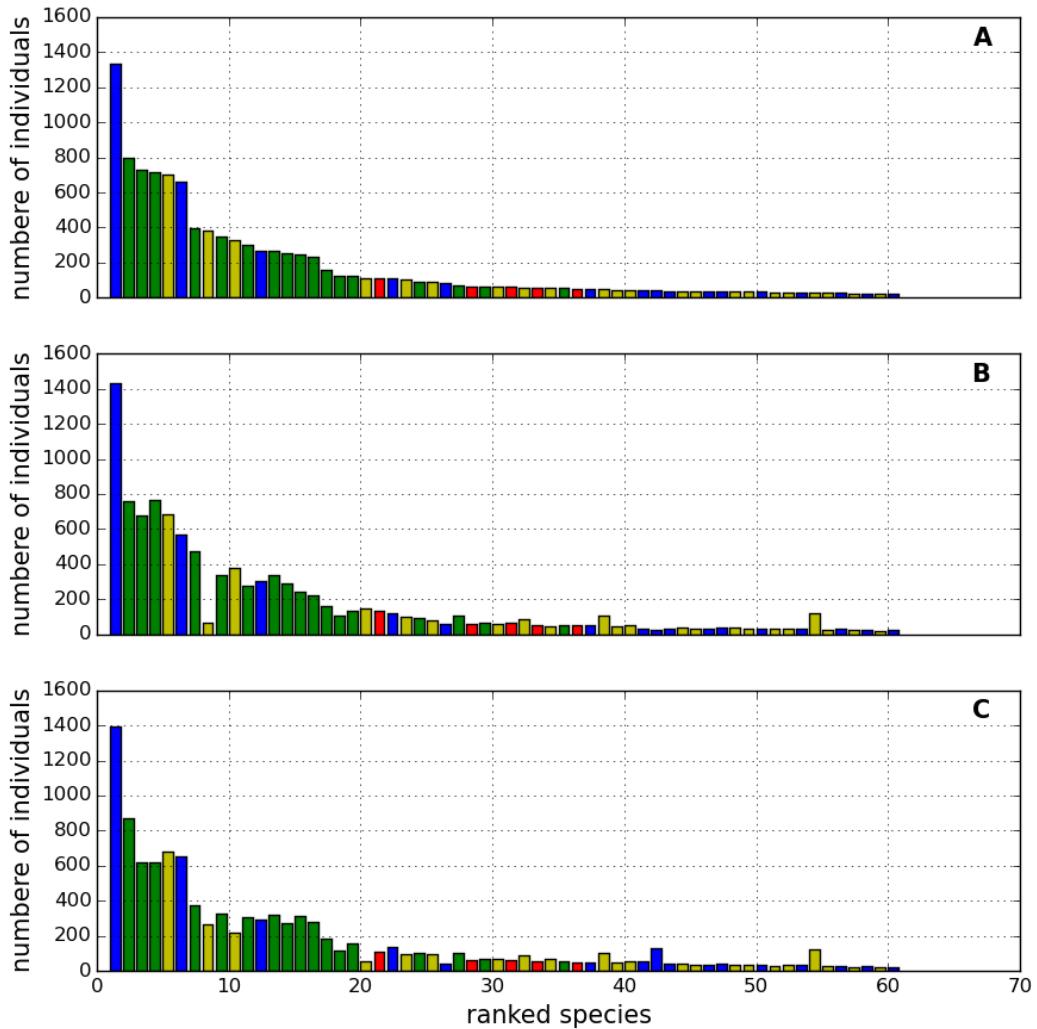


Figure 5.13: Rank abundance spectra (RAS) for three simulations run using the interaction network NM1 (see text). Species abundances are measured by taking the mean abundance over the final 1000 iterations of the simulation. The species are ranked according to their abundances in the first simulation (panel (A)). This ordering is retained in panels (B) and (C), which represent different simulations. Colouring of species by trophic level is consistent with previous figures.

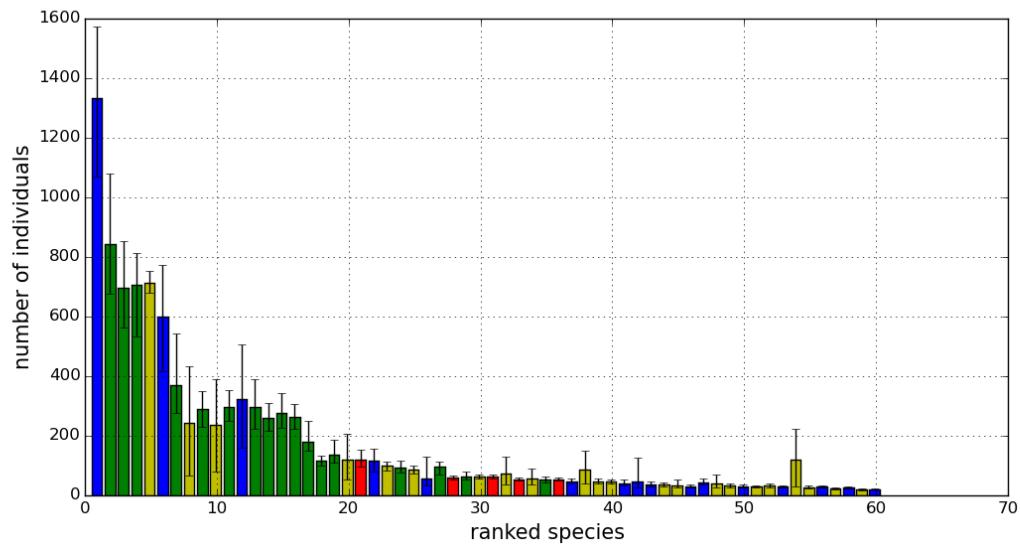


Figure 5.14: The average rank abundance spectrum (RAS) for the ensemble of 50 simulations run using interaction network NM1 (see text). Species abundances measured as in figure 5.13, and ranked as in panel (A) of that figure. The main bars indicate mean abundance values, whilst the error bars indicate the minimum and maximum abundances over the ensemble.

CHAPTER



PERSISTENCE IN CLOSED COMMUNITIES

TODO: These results also have trophic links between bottom and top levels. Need to re-run these simulations? TODO: Come back to this chapter having done work on stability... e.g. Are the persistent communities stationary? TODO: Look at Thilo's reference on constructing stable GLM topologies (distinct pairs approach..)

TODO: Are the 'good net' persistent net works here stationary?? TODO: Find smaller stationary networks (4-10 species, and fit GLV!!) [w or w/o immigration] TODO: Finish All chapters on the IBM!!

6.1 Motivation

At the begining of chapter 4 we saw that species persistence is low in communities without immigration (figure 4.1). We defined the term *community collapse* as the extinction of all non-basal species from the landscape. With a zero immigration rate (IR) antagonistic communities collapsed in all simulations, whilst mutualistic communities ($MAI = 0.5, 1.0$) sometimes show persistence of a few species in higher trophic levels. It appears that our model struggles to produce stable and persistent communities without immigration. This may also be the case in the real-world. It certainly appears that immigration plays an important role in shaping natural communities [REFS]. In the previous chapter we saw that the IR can alter the structure of simulated communities and affect how they respond to habitat loss. In this chapter we fix the IR at zero, and explore other factors contributing to the stability of simulated communities. In particular we ask whether it is possible to obtain stable and persistent communities without immigration. We begin with a discussion (section 6.2) on the concept of stability and why it is

important in ecology. We then explore how certain model parameters¹ impact on persistence (Section 6.3), and ask what role is played by the structure of the interaction network (section ??).

6.2 Stability in ecology

Stability as a fundamental property of ecological community? - but we do not know how they manage to be stable - our mathematical representations do not always suggest that they are stable - examples. Also do not even really know in which ways they are stable, and to what extent they are - there are clearly points of instability - mass extinctions, community collapse...

Stability is an important concept in ecology. However there is no single coherent definition that ecologists use. In the work up to this point stability has been a key consideration - we have defined, and used in our analysis, a number of different metrics relating to it. In what follows we explicitly compare the different stability concepts used in ecology. Before going into the details we note the key point that real-world ecosystems generally appear to be stable, in some sense. A lack of stability may suggest wildly varying population dynamics, extreme responses to extrinsic disturbances, or the extinction of species, depending on your definition of stability. Although there are examples of these phenomena in nature (e.g. spring bloom in plankton abundances, desertification, mass extinction events), luckily they are the exception not the norm. Therefore stability has become a key property of ecological communities (and meta-communities) that we must try to understand.

To include here:

- Influence of dynamical systems (population modelling) - Lyapunov stability. Steady-state/Equilibrium assumption vs Plasticity, adaptivity.
- Limitations with what can be measured in the field. Other metrics used instead:
 - Resilience, Robustness, Persistence, Temporal variability.
 - Spatial stability metrics (alpha,beta,gamma). Issue of scale. [65]
 - Immigration as an important mechanism behind stability/persistence - see previous chapter
 - Competitive exclusion
 - Network structure as a contributing factor to stability (large body of literature, also hybrid-networks)
 - Biodiversity thought to be good for stability - insurance hypothesis [ref], others?
 - Functional stability - see iee paper. Is this a widespread concept?

¹We could include more parameters in this analysis e.g. synthesis ability

Argue that the metric/concept of stability used depends on the context/questions you are asking. Here we choose to focus mainly on persistence and on species extinctions. Specifically we try to overcome the problem that most non-plant species go extinct with there is no immigration.

Resilience

Robustness

6.2.1 Dynamical systems and stability

The mathematics of dynamical systems has made a major contribution to the field of ecology, and the paradigm of an ecological community as a complex dynamical system remains prevalent amongst theoretical ecologists. As discussed in section ??, there are some theoretical ecologists who question the use of the concept. In this context stability has a rigorous meaning that relates to the dynamics of the mathematical system. This stability concept is often referred to as dynamical stability or asymptotic stability, and is derived from the eigenvalues of the linearised system close to equilibrium. Conceptually it tells us if the system will return to the equilibrium if perturbed a small way. Related concepts tell us...[3]

[40] - complexity versus stability debate. Random matrix theory. Important misinterpretation of this equation..

[24] - the ensemble of possible networks scales combinatorially with number of species. Variation in interaction strengths is shown to destabilise large networks. For a 50 species network probability of stability is $\sim 10^{-3}$ with connectance 0.1. This suggests that our sample size for repeat simulations is very small indeed. (Does it explain why dynamics without immigration is unstable?)

6.2.2 Network structure and stability

The popularity of the network perspective in community ecology makes it natural to ask what effect the interaction structure has on stability. This question represents a large body of literature. As we saw, the dynamical systems approach makes explicit the relationship between network structure and stability (section ??). This is not so easy with natural communities and empirical data..for various reasons..lack of information, do not have knowledge of underlying dynamics, unclear if there is an underlying dynamic model with fixed stability [REFs]. Therefore some community properties must be understood in terms of plasticity..

As discussed in section ??, the bulk of network ecology has focused on networks of single interaction types. The same is true when it comes to stability. The general consensus has been that the properties conferring stability on antagonistic and mutualistic networks differs. This is discussed in the introduction.

[61] simulated 7200 mutualistic and trophic networks, with varying diversity, connectance, nestedness, and modularity. During simulations species become extinct, which changes the initial network structure, and therefore changes these metrics also. Results confirmed by looking at empirical networks. Attributed to perhaps indirect effects: apparent competition and apparent facilitation.

[55] later shows that these findings do not hold when combining interaction types into a single network.

[44] and [39] ...

[46] - experimental manipulation of strong and weak interactors.

6.2.3 Spatial stability

alpha beta gamma stability metrics?? ref??

6.3 Model parameters versus persistence

We first attempt to improve the persistence of our simulated communities by varying certain model parameters. The parameter space of the model is large (see table ??), therefore we do not attempt to explore all of it. Previous work by Lurgi et al. [39] in developing the model has ensured the realism of the bio-energetic parameters (where possible they are derived from literature - more on this). Therefore we restrict our exploration to the region of the default parameters. It may be that there exists somewhere a region of stable coexistence of all species for zero IR. If there does, we will not find it. However we may attempt to improve persistence and assess the impact of varying sensibly chosen model parameters.

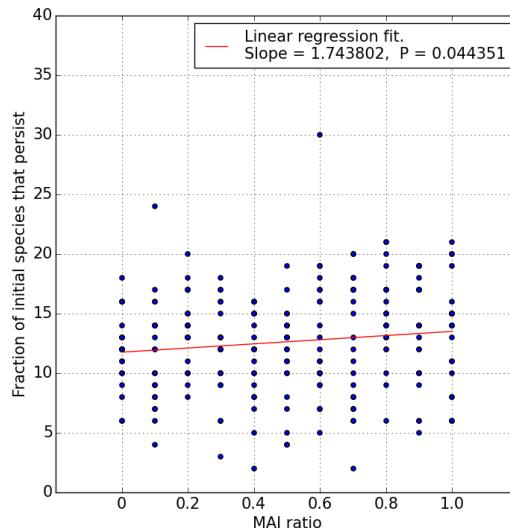


Figure 6.1: Species persistence is plotted for 22 repeat simulations at each MAI ratio. Persistence is measured as the fraction of the initial 60 species that have not gone extinct by the end of a simulation (5000 iterations). Each blue dot gives the number of persistent species for a single simulation. The red line shows a linear regression fit with slope and p-value given in the inset.

6.3.1 Mutualistic to antagonistic interaction (MAI) ratio

A key theme throughout this thesis, and one of the main novel aspects of this research, is the inclusion of mutualistic interactions into simulations of trophic dynamics. In some cases we have seen that these mutualistic interactions play a stabilising role in the community (in contrast to May's classic 'orgy of mutual benefaction'). Therefore it seems natural to ask what role the MAI ratio plays in the persistence of communities at zero IR.

Figure 6.1 shows that there is an increase in overall species persistence with MAI ratio. Although the trend is statistically significant it is small, with an expected increase of about twelve

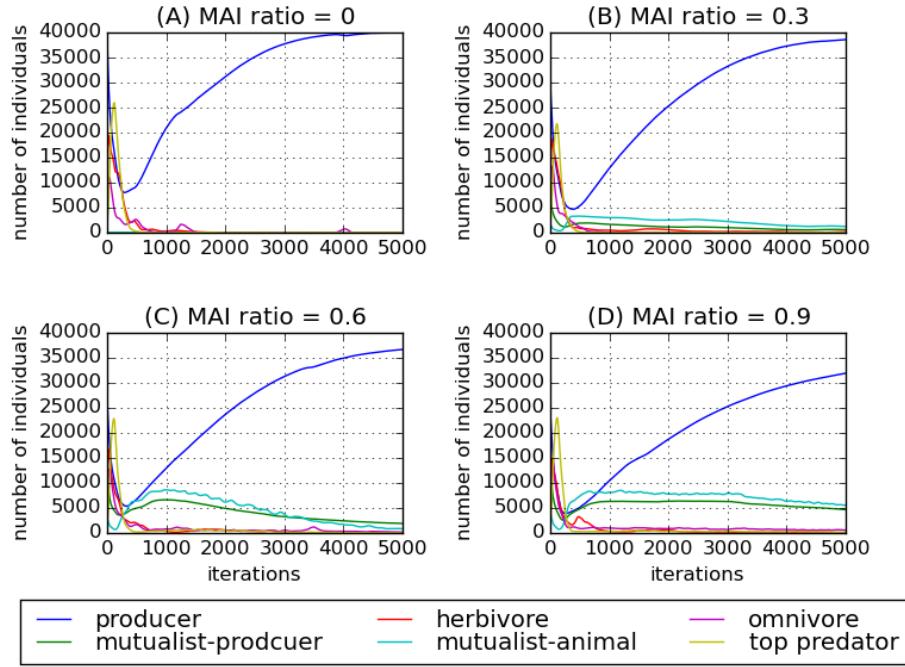


Figure 6.2: Mean dynamics by functional group for four different MAI ratios. The coloured lines represent the abundance dynamics of the different functional groups, as indicated in the legend. Abundance at each simulation iteration is measured by the total number of individuals belonging to all species of a given functional group.

to fourteen species over the whole range of MAI ratios. For antagonistic communities ($MAI = 0.0$) we know to expect community collapse. This is observed in figure 6.2, which shows the expected abundance dynamics of each functional group (FG). In panel A we see that the abundance of producers rises to fill the whole landscape ($200 \times 200 = 40,000$), whilst the abundance of all other FGs is at or near zero. From the other panels (B-D) we see that increasing the MAI ratio particularly benefits the mutualistic species (both producers and animals) as expected, and appears to also confer some benefit to members of the other FGs. Ecologically this makes sense - if mutualism strongly benefits mutualistic species, it will also benefit those species that feed on them. (It also appears that increasing the MAI ratio increases the time taken to reach steady state - the abundance of producers in panel D is clearly still rising.)

As we have seen previously (sections ??, 4.4.3) the MAI ratio affects community composition as measured by the relative abundances of the FGs. Figure 6.4 shows us that the relative abundance of non-mutualist producers falls sharply as the relative abundance of mutualist species, both plants and animals, increases. It appears that the mutualist-producers outcompete the non-mutualists, thanks to the benefit gained by a plant in switching to mutualism (section ??). Interestingly this also benefits the mutualist-animals, but not the herbivores, which show no

6.3. MODEL PARAMETERS VERSUS PERSISTENCE

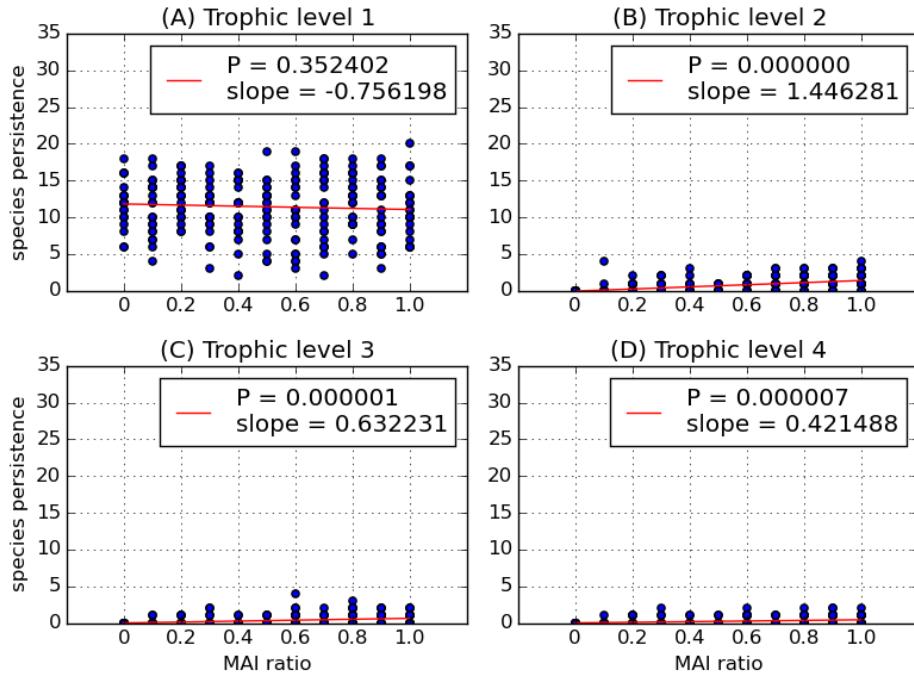


Figure 6.3: Species persistence by trophic level over a range of MAI ratios. Each blue dot corresponds to the number of remaining species from that trophic level at the end of a single simulation. There are 22 repeat simulations at each MAI ratio, each using a different interaction network. The red lines show linear regression fits, with slopes and p-values given in the insets.

significant increase in relative abundance ².

Despite the changes in the distribution of abundances, there is little change in the species richness of each trophic level. Figure 6.3 shows that the overall increase in species persistence is due to an increase in the species richness from zero to about one or two, in trophic levels two, three and four (panels B,C,D). We may have expected a greater increase in persistence, especially in the second trophic level, where the expected absolute and relative abundance increases considerably. The fact that less than two species are expected in this trophic level at $MAI = 1.0$ suggests that either competition or stochastic effects are important³ here.

²perhaps there is competition between these two FGs? Would have thought that those herbivores which feed on mutualistic plants would benefit from their increased availability? - only in absolute numbers, some suggestion of this in panel D of figure 6.4. Why do they then die out?

³Since cannot recover from extinctions.

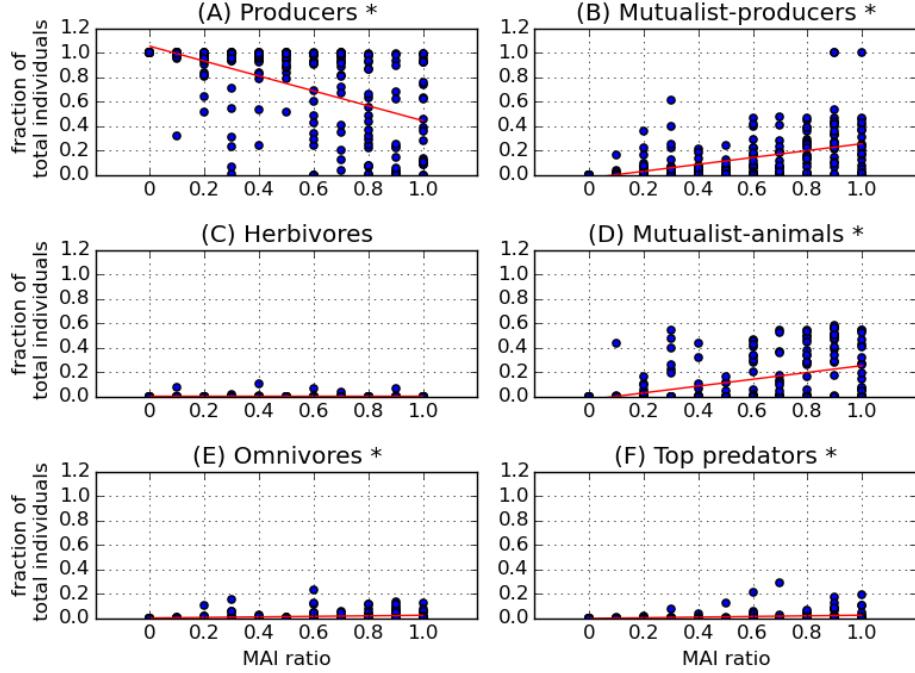


Figure 6.4: Relative abundance (RA) by functional group for a range of MAI ratios. RA is measured as the fraction of the total individuals that belong to a given functional group. Each blue dot corresponds to the RA at the end of a single simulation. There are 22 repeat simulations at each MAI ratio, each using a different interaction network. Red lines show linear regression fits. Panels marked with an asterix * (A,B,D,E,F) have fits with $p\text{-value} < 0.0001$.

6.3.2 Reproduction rate (RR)

The problem of low persistence in high trophic levels remains. Mutualism has some small effect, but even at $MAI = 1.0$ we expect only one or two species on average in the non-basal trophic level. The initial transience in the abundance dynamics (figure 6.2) is characterised by a sharp decline in plant abundance (mutualist and non-mutualist), which reaches a minimum and then rises again. It was hypothesised that this overconsumption and therefore limited availability of plant individuals, causes many of the extinctions. Indeed in these simulations $\sim 85\%$ of the extinctions occur during the first 500 iterations. Therefore we look at the possibility of improving persistence by increasing the reproduction rate (RR). This parameter defines the rate at which non-mutualist producers reproduce (via the wind-dispersal mechanism, see chapter ??). Therefore this increasing this mechanism should improve the availability of plant biomass in the system, with potentially cascading effects. The RR parameter does not affect mutualist-animals, which only reproduce via their interactions with mutualist-animals and not via wind-dispersal.

Simulation results are presented for $MAI = 0.0$ and $MAI = 0.5$. The main results are as follows:

6.3. MODEL PARAMETERS VERSUS PERSISTENCE

- Increasing RR increases overall species persistence (figure 6.5). The effect is greater for antagonistic communities.
- The sharp decline in plant abundance during the transience is reduced (figures 6.10, 6.11). As we reasoned, this does result in increased absolute abundances of all FGs at both MAI ratios. This is visible in these figures.⁴
- The relative abundances by FG indicate that top-predators do very well out of the increase in RR (figures 6.8, 6.8). This is due to the flaw in the niche model already discussed.
- As before the increased abundance by FG does not necessarily translate into increased species richness (figures 6.6, 6.7). Again there is a weak trend - increasing the RR by a factor of twenty, results in one or two more species on average in the higher trophic levels. In the mutualistic communities ($MAI = 0.5$) increasing the reproduction rate is bad for persistence in the second trophic level.
- We choose a higher reproduction rate for the further simulations in this chapter because overall it improves persistence in all FGs. It is not unrealistic to improve the reproductive ability of plants. Importantly it does affect the trade-off between mutualism/non-mutualism for plants.

⁴Look at when extinctions occur? Plot cumulative extinctions against time?

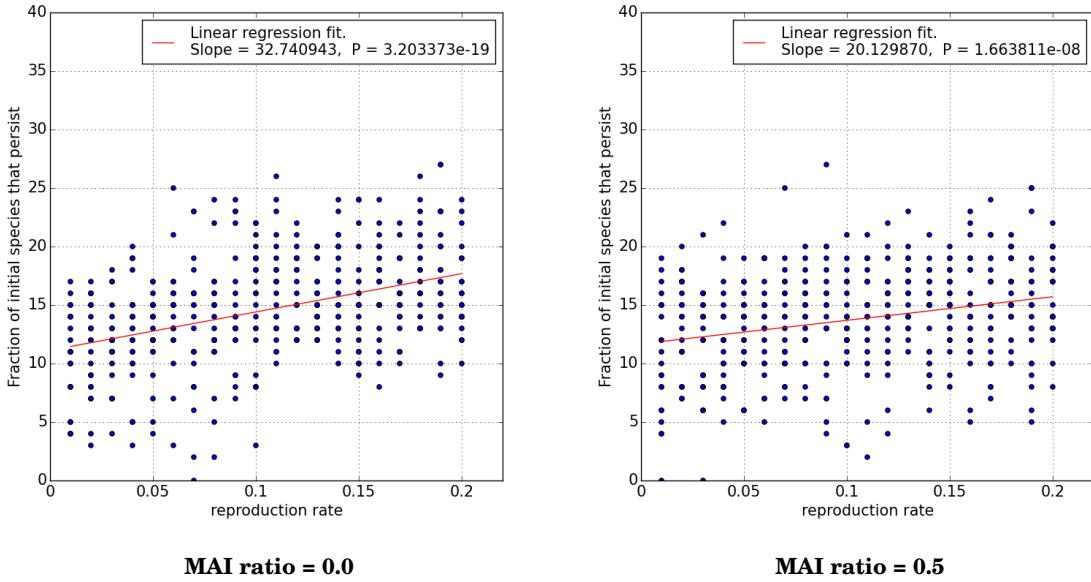


Figure 6.5: Species persistence against reproduction rate (RR), with 22 repeat simulations at each RR. Persistence is measured as the fraction of the initial 60 species that have not gone extinct by the end of a simulation (5000 iterations). Each blue dot gives the number of persistent species for a single simulation. The red line shows a linear regression fit with slope and p-value given in the inset.

6.3.3 Landscape size

Another hypothesis was that spatial competition was causing the community collapse. Simulations run to test persistence response to both the size of the landscape. This was done for MAI = 0.0 and 0.5, with 25 repeat networks at each landscape size. Increasing the size of the landscape should reduce the effect of spatial competition and therefore increase persistence.

The total abundance increases by around 100 fold, as the width of the landscape is increased from 100 to 1000. For MAI=0.0 it goes from 10586 to 1015475 individuals on average. Figures 6.12 and 6.13 summarise the results for antagonistic and mutualistic communities respectively. In both cases there is an overall increase in species persistence with landscape size, driven by small increases in the species richness of all trophic levels. However the effect is small and it does not appear that it would resolve the species persistence problem for a landscape of a size we could realistically simulate. Therefore further simulations use the same landscape size of 200.

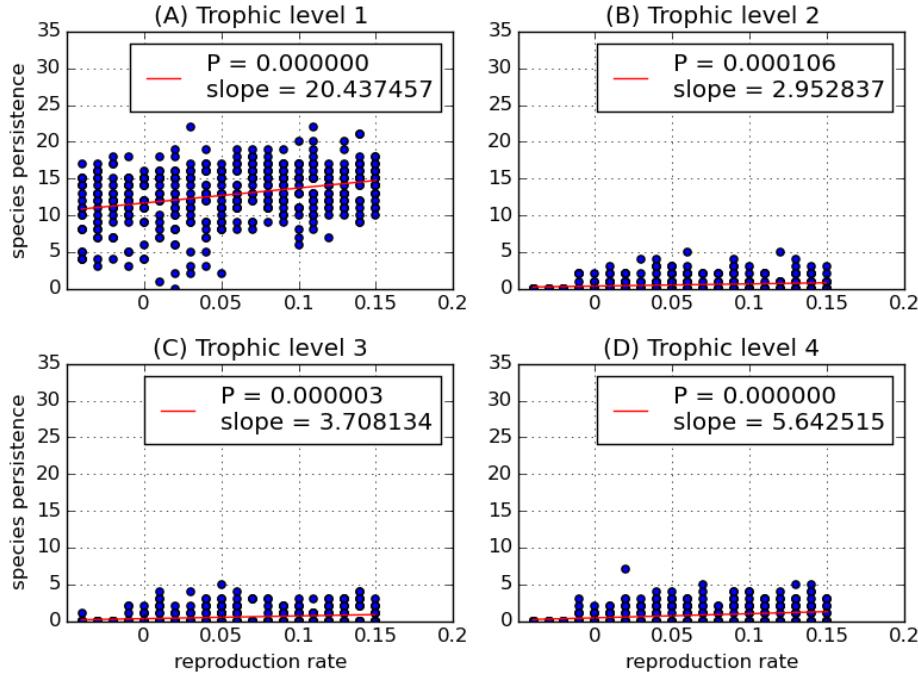


Figure 6.6: **MAI = 0.0.** Species persistence by trophic level over a range of reproduction rates. Each blue dot corresponds to the number of remaining species from that trophic level at the end of a single simulation. There are 22 repeat simulations at each MAI ratio, each using a different interaction network. The red lines show linear regression fits, with slopes and p-values given in the insets.

6.3.4 Number of initial species

All previous simulations have been run with an interaction network and initial species pool consisting of 60 species. We consider the possibility that beginning the simulation with a larger network may result in a greater number of persistent species. (In fact increasing the number of possible interaction networks and therefore making it less likely to find a stable one?).

With more than around 70 species it became impossible to use the network generation procedure (discuss why). Therefore we used the pure niche model, and implemented an algorithm to re-wire niche model networks such that the original trophic constraints could be met. We refer to the pure niche model networks and the re-wired networks as *niche* and *rewired* respectively.

The number of initial species does not appear to effect the total abundance ⁵. There is an increase in overall persistence, in the case of the rewired networks. As shown in figures 6.14 and 6.15 this increase is almost entirely due to plants. Therefore this does not overcome the problem that very few species persist at higher trophic level. We are still include to propose that this is

⁵Need to plot this? Could include some plots in an appendix.

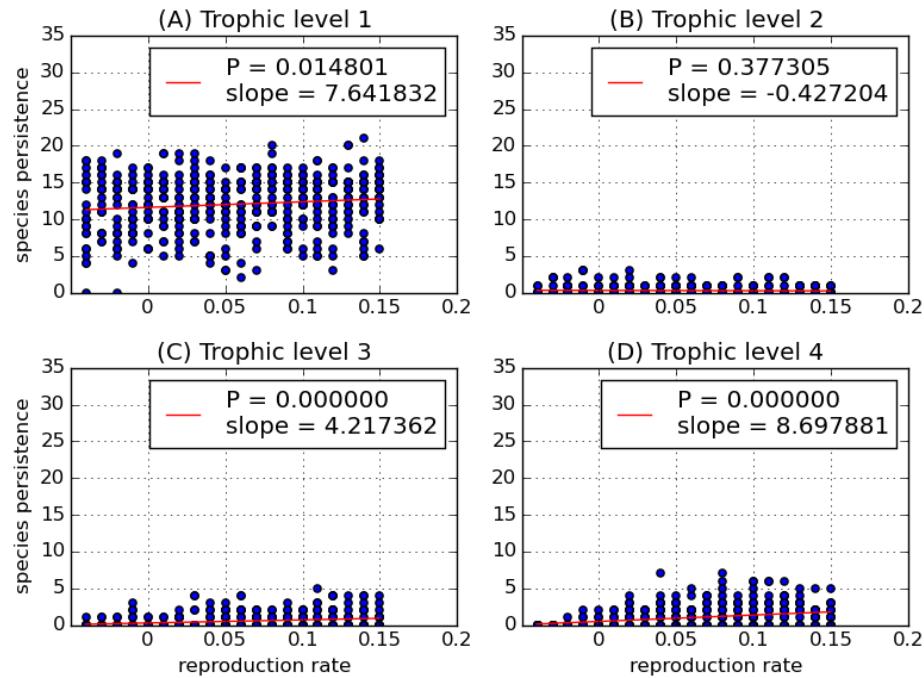


Figure 6.7: MAI = 0.5. Species persistence by trophic level over a range of reproduction rates. Each blue dot corresponds to the number of remaining species from that trophic level at the end of a single simulation. There are 22 repeat simulations at each MAI ratio, each using a different interaction network. The red lines show linear regression fits, with slopes and p-values given in the insets.

due to competition, combined with stochastic extinctions during transience.

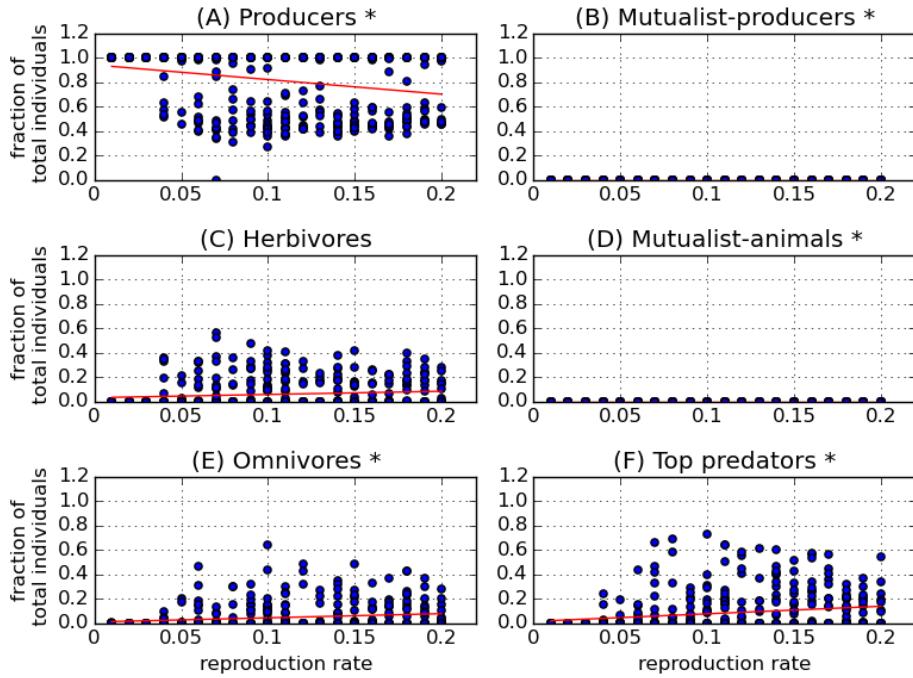


Figure 6.8: **MAI = 0.0**. Relative abundance (RA) by functional group for a range of reproduction rates. RA is measured as the fraction of the total individuals that belong to a given functional group. Each blue dot corresponds to the RA at the end of a single simulation. There are 22 repeat simulations at each MAI ratio, each using a different interaction network. Red lines show linear regression fits. Panels marked with an asterisk * (A,B,D,E,F) have fits with $p\text{-value} < 0.0001$.

6.4 Why are some communities more stable than others?

Despite extensive investigation of model parameters persistence at zero IR is remains poor. However in the results presented above one thing stands out - there considerable variation in the results between simulations.

- Is this variation due to noise, or inherent differences between each simulation? (Below we show that it is inherent by repeating simulations with the same network.)
- Therefore the difference must be due to interaction network. Accepted in the field that network structure affects stability (see section 6.2). Is this true in the real world?
- We choose one 'good' and one 'bad' network each with 120 species. Chosen by looking at the 25 repeats and picking the one with the highest/lowest number of species in all trophic levels. We show that this difference between the two networks is repeatable - that there is evidence for a systematic difference between how many species persist and therefore one is better than the other (see figure 6.17).

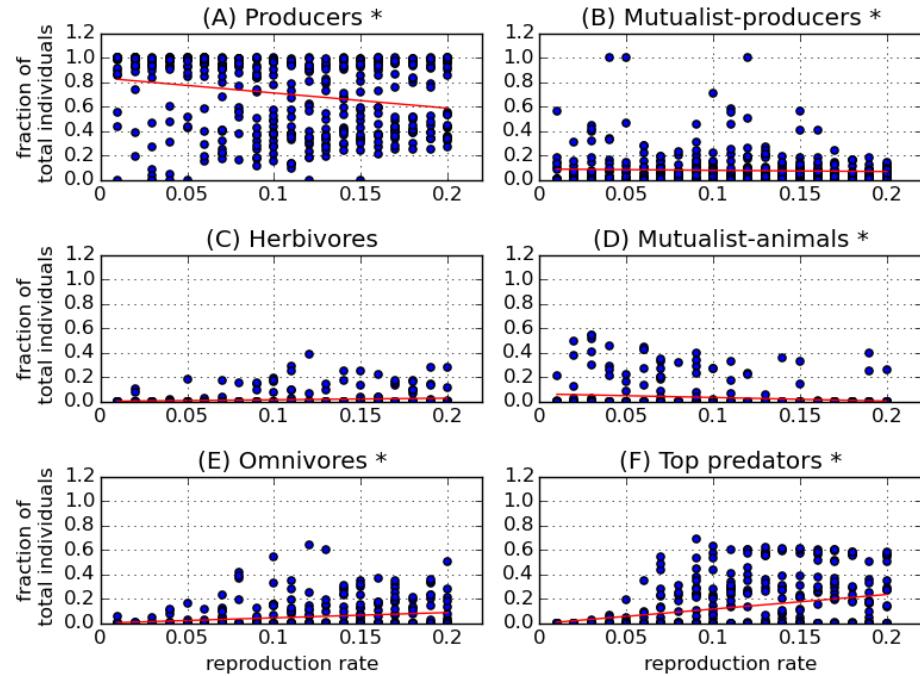


Figure 6.9: MAI = 0.5. Relative abundance (RA) by functional group for a range of reproduction rates. RA is measured as the fraction of the total individuals that belong to a given functional group. Each blue dot corresponds to the RA at the end of a single simulation. There are 22 repeat simulations at each MAI ratio, each using a different interaction network. Red lines show linear regression fits. Panels marked with an asterisk * (A,B,D,E,F) have fits with $p\text{-value} < 0.0001$.

We need to extend this analysis and crucially work out **where this chapter is going!**. ToDo:

- Re-do this analysis for 30 species (easier to visualise than 120)
- Look at individual species. Do they go extinct in the same order?
- Look at network metrics that have been associated with stability. E.g. are modular networks more stable?

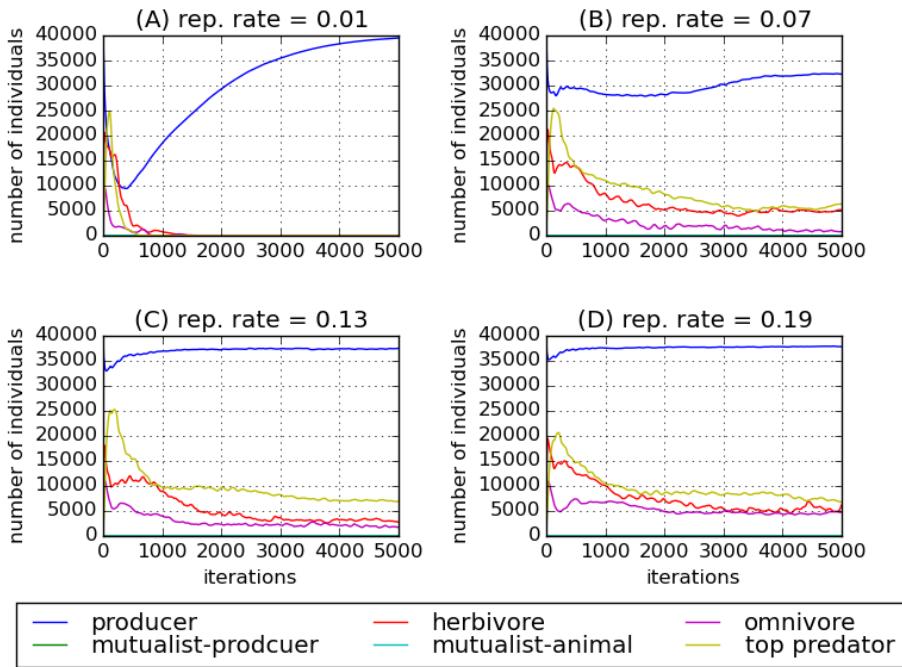


Figure 6.10: **MAI = 0.0.** Mean dynamics by functional group for four different reproduction rates. The coloured lines represent the abundance dynamics of the different functional groups, as indicated in the legend. Abundance at each simulation iteration is measured by the total number of individuals belonging to all species of a given functional group.

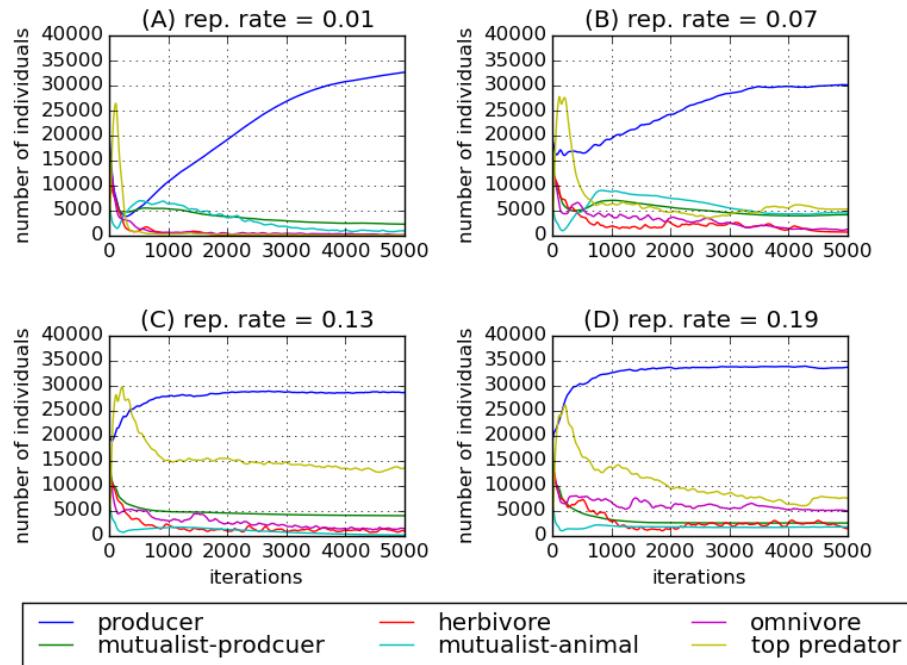


Figure 6.11: **MAI = 0.5**. Mean dynamics by functional group for four different reproduction rates. The coloured lines represent the abundance dynamics of the different functional groups, as indicated in the legend. Abundance at each simulation iteration is measured by the total number of individuals belonging to all species of a given functional group.

6.4. WHY ARE SOME COMMUNITIES MORE STABLE THAN OTHERS?

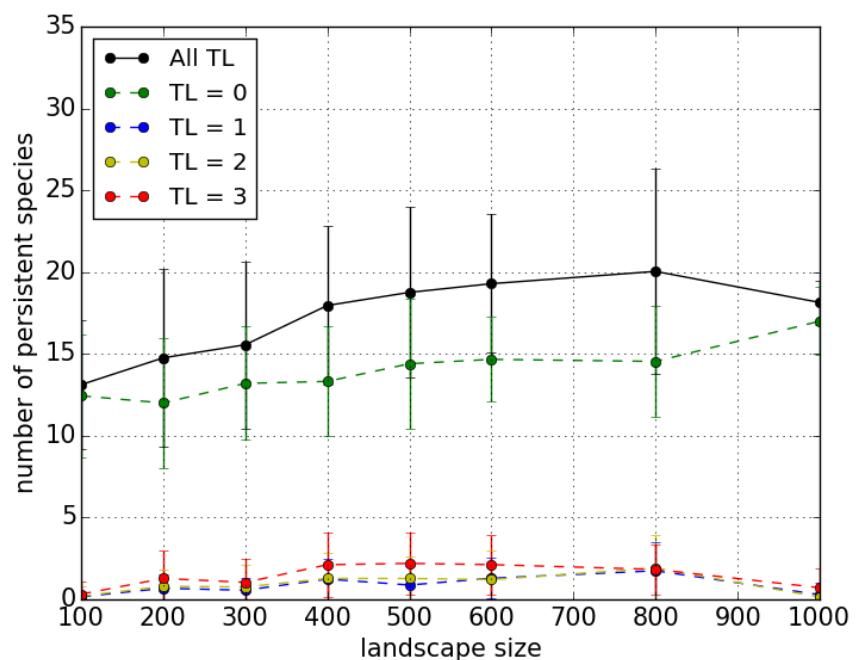


Figure 6.12: **MAI = 0.0**. Landscape size against species persistence, total and broken down by trophic level. The points show the mean value over 25 replicates and the error bars show \pm one standard deviation.

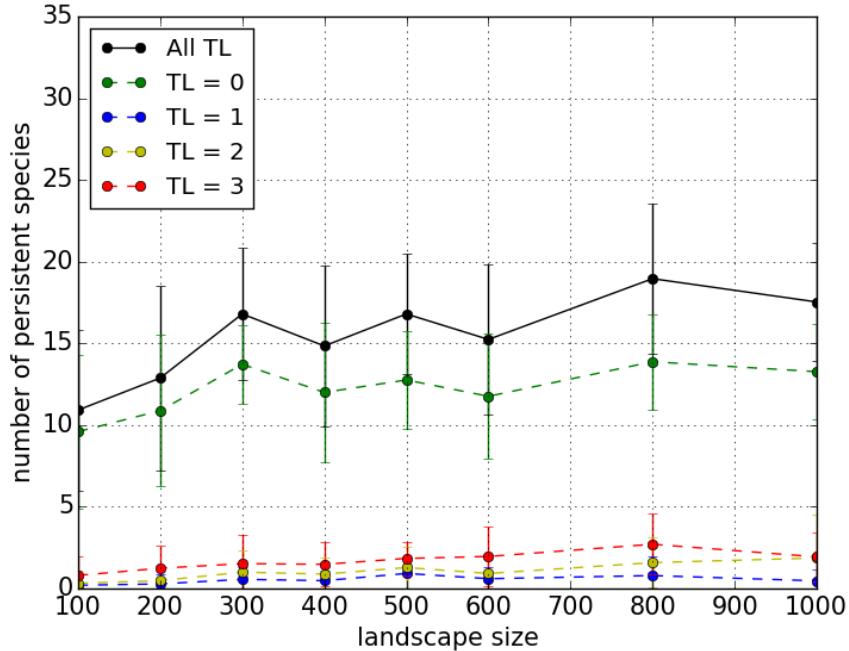


Figure 6.13: **MAI = 0.5.** Landscape size against species persistence, total and broken down by trophic level. The points show the mean value over 25 replicates and the error bars show \pm one standard deviation.

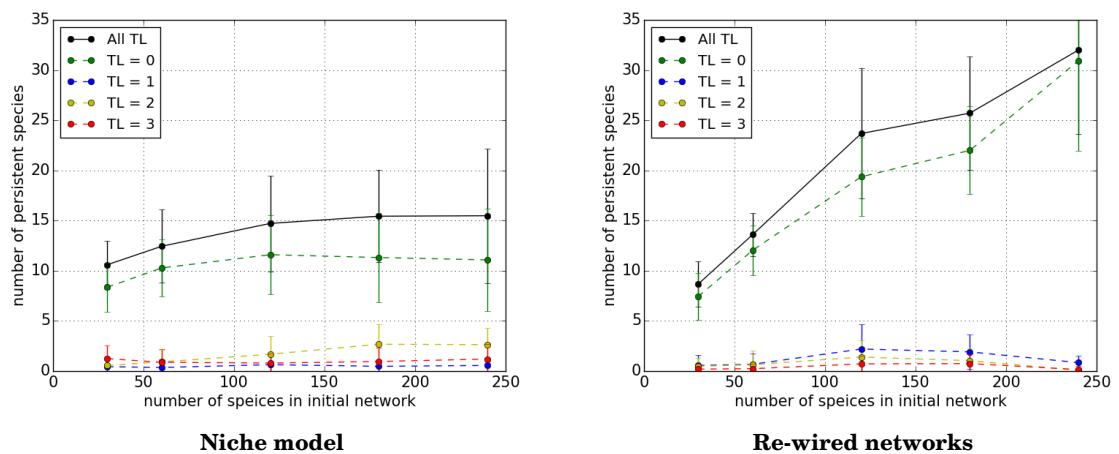


Figure 6.14: **MAI = 0.0.** Number of initial species against species persistence, total and broken down by trophic level. The points show the mean value over 25 replicates and the error bars show \pm one standard deviation.

6.4. WHY ARE SOME COMMUNITIES MORE STABLE THAN OTHERS?

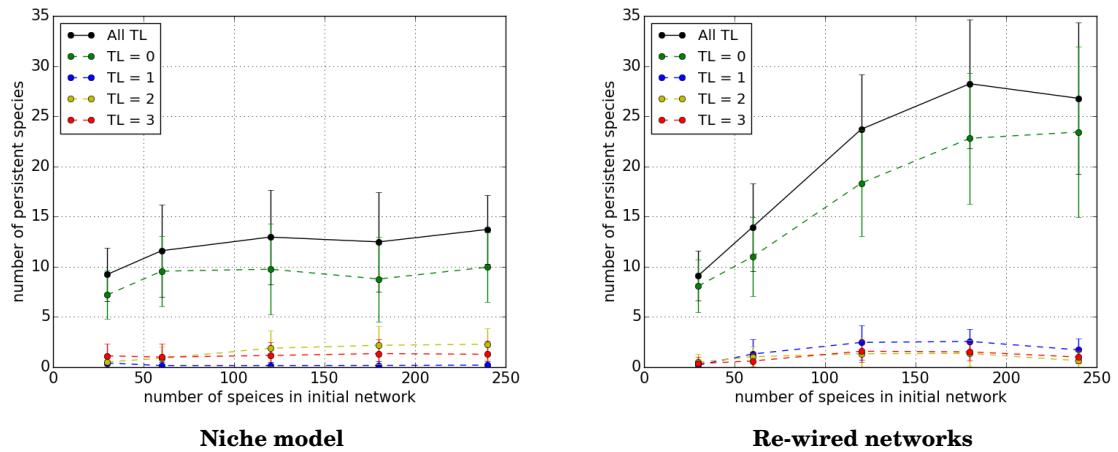
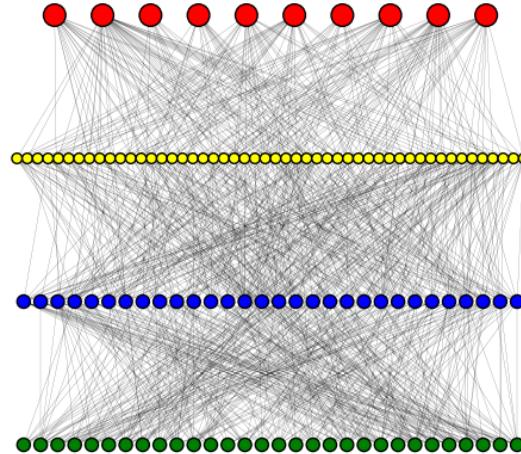


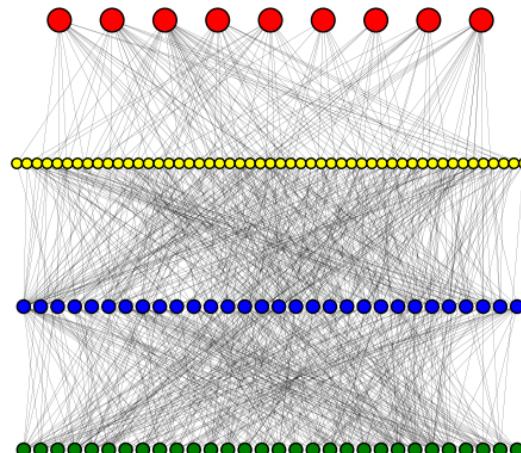
Figure 6.15: **MAI = 0.5.** Number of initial species against species persistence, total and broken down by trophic level. The points show the mean value over 25 replicates and the error bars show \pm one standard deviation.

Example 'bad' network (ID=59). Re-wired from niche model.



Bad network

Example 'good' network (ID=52). Re-wired from niche model.



Good network

Figure 6.16: Two 120 species interaction network. One produces better persistence than the other, although in both cases most species go extinct.

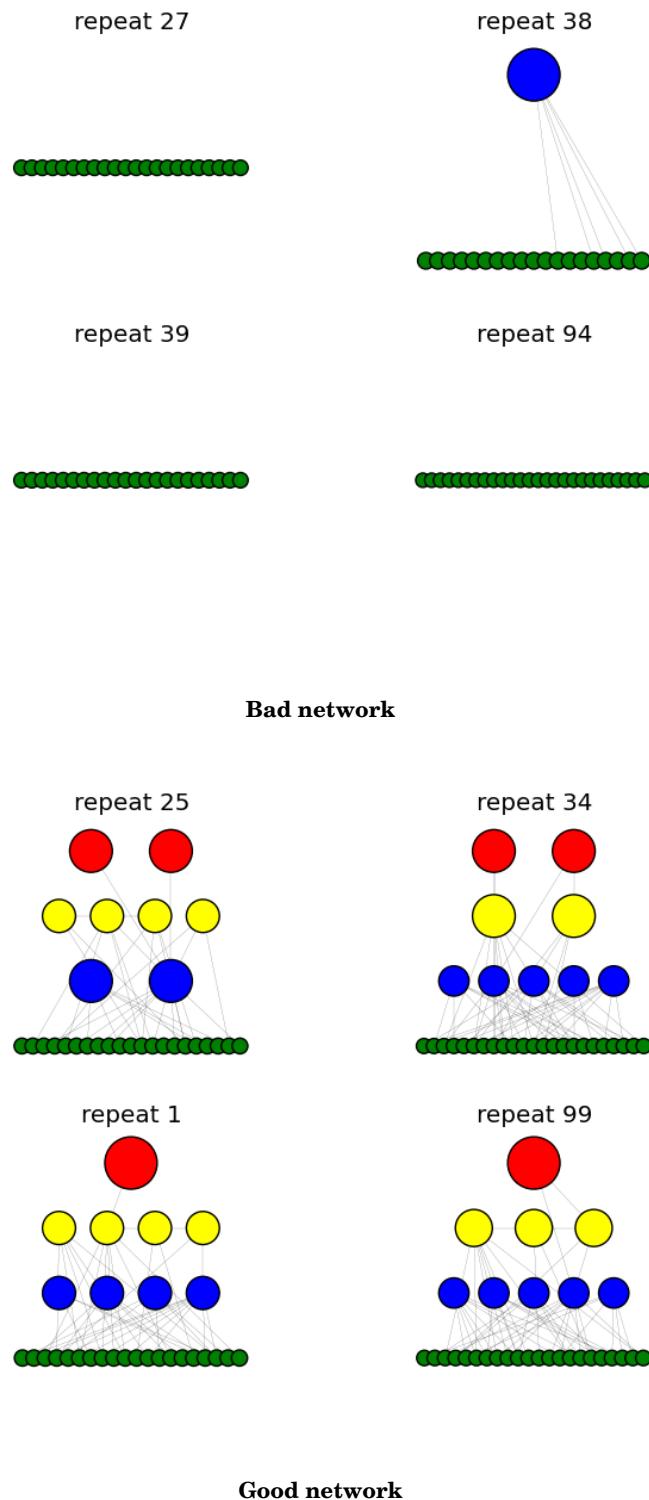


Figure 6.17: Examples of the networks of species that remain at the end of a simulation.

INFERRING SPECIES INTERACTIONS

7.1 Motivation

In this chapter we investigate the possibility of quantifying species interaction strengths from observed population dynamics. This was discussed at length in the introduction (section ??), however we now reiterate some of the important points here and to motivate our approach to the problem.

- metrics for interaction strength (choice of IM metric)
- hare-lynx dataset (ongoing debate!)
- Availability of time-series data. Plankton system (complications! e.g. seasonality)
- Motivate simulation approach (known interactions)
- Timme approach to fit GLV (use of GLV in general - constant interaction strengths)
- Simplify the problem (two species predator prey, but see extension)

We implement and test a novel method for quantifying species interactions from population dynamics. Initially we test this method using two distinct ODE simulation models to generate the population dynamics. These two models are referred to as the *linear* and *holling* models. The GLV can fit the *linear* model exactly. Therefore the results presented in section 7.3.1 serve as a test of our numerical method. We characterise the robustness of the method to the addition of noise, and under sparsity of sampling. The *holling* model cannot be fitted exactly by the GLV because it uses a different form of *functional response*. Therefore applying our method to population dynamics simulated using the *holling* model provides a test of its ability to give approximate

estimates of interaction strength, when underlying structure of the system is not of GLV type. This is also tested in the presence noise and under sparse sampling, and the results are presented in section 7.3.2. In section 7.3.3 we provide a preliminary look at how the method could be used to determine the type of functional response present, by observing the dynamics. Having characterised the performance of our method using ODE models, we then apply it, in section 7.5, to dynamics generated using the IBM model of previous chapters. This represents a first step towards applying the method to empirical data, since it involves a spatial system and a larger number of species. We conclude the chapter with a discussion of how the method could be developed further towards empirical application.

7.2 Methodology

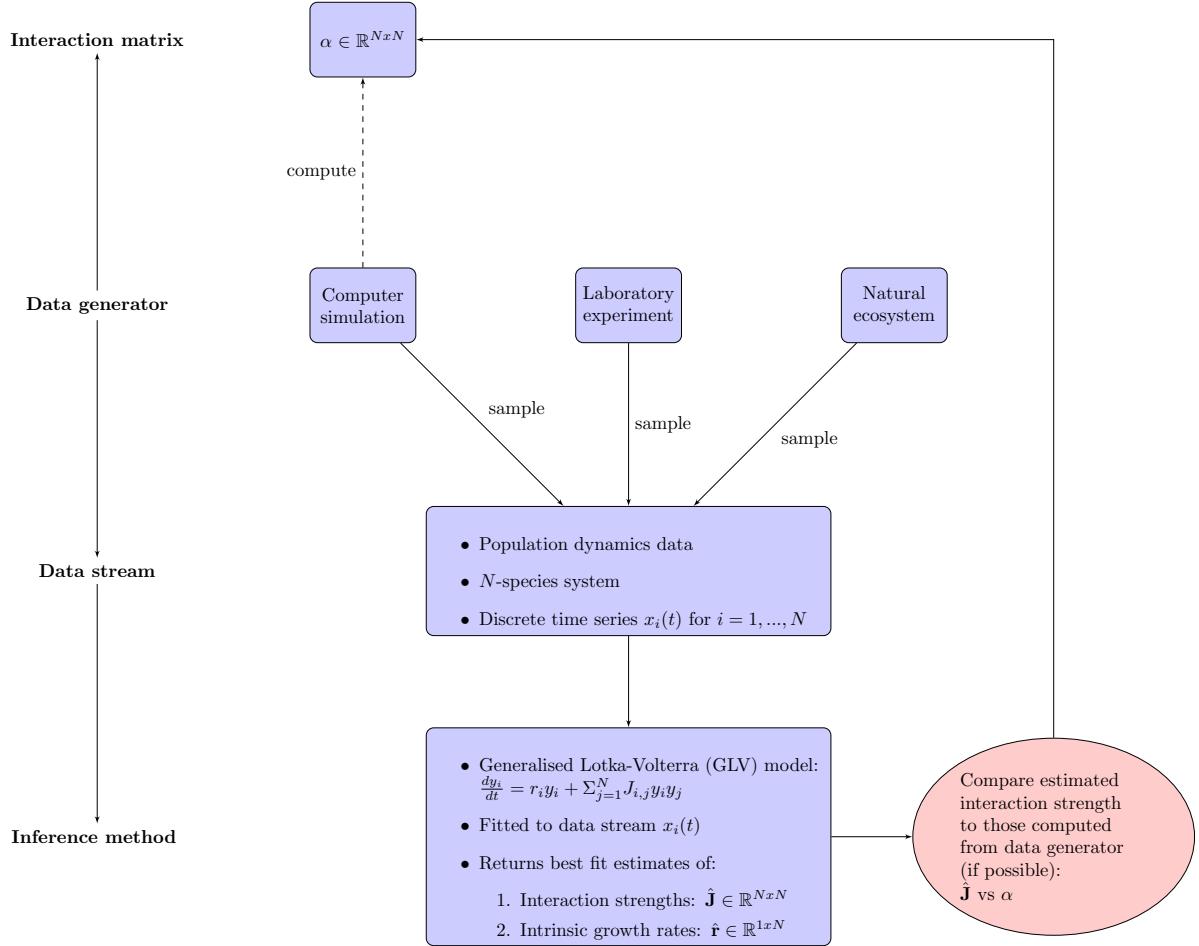


Figure 7.1: This is what we do.

To summarise our methodology, we simulate population dynamics then sample these dynamics and fit a *generalised Lotka-Volterra* (GLV) model. The fitted GLV parameters give us estimates

of the species interaction strengths (and other parameters), which we then compare to those used in the original simulation. The details of all the stages are given below. In section 7.2.1 the *interaction matrix* (IM) is introduced. The IM is the metric used to quantify the strength of species interactions and is key to this chapter. We also introduce the *generalised Lotka-Volterra* (GLV) model, and show that this model has constant interaction strengths, given by the coupling matrix J . In section 7.2.2 we give a general framework for ODE predator-prey modelling, and derive the two models that we use to simulate population dynamics. We then discuss, in section 7.2.3 the details of how these models are simulated *in silico*, including the selection of model parameters. Section 7.2.4 gives the details of the numerical method we use for fitting the GLV model to sampled population dynamics. In section 7.2.5 we give an example of the full methodology in action.

7.2.1 Interaction strength

The metric used for species interaction strength is key to this chapter. As discussed in section 7.1 there are several metrics available [REF]. However there is one that is a natural choice, given our methodology. This metric allows us to calculate the interaction strengths from our simulation models, and to directly compare them to the estimates obtained by fitting the GLV model. The metric is called the *interaction matrix* (IM). The elements of the IM, α_{ij} , quantify the effect of a small change in the population density of species j on the per capita growth rate of species i [REF]. Therefore the IM elements are given by:

$$(7.1) \quad \alpha_{ij} = \frac{\partial}{\partial x_j} \left(\frac{1}{x_i} \frac{dx_i}{dt} \right),$$

where x_i and x_j are the population densities of species i and j respectively. In the case of our two species systems the IM is a 2×2 matrix, but trivially extends to quantify all N^2 pair-wise interactions between species in a N -species system (including self-interactions).

Using the IM we are able to calculate the interaction strengths exactly from the models that we use for simulation. This is because they are ODE models with explicit expressions for dx_i/dt , so we can evaluate the partial derivative in equation 7.1 to obtain analytic forms for all the IM elements ($\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11}$). Depending on the model used the IM elements are either constants, or are functions of prey density. The interaction strengths for our simulation models are given at the end of section 7.2.2, and are illustrated in figure 7.2.

Generalised Lotka-Volterra. The GLV model is the extension of the Lotka-Volterra equations to N species. We use this model to fit to simulated population dynamics and obtain estimates of the underlying interaction strengths. The GLV model is given by:

$$(7.2) \quad \frac{dx_i}{dt} = r_i x_i + \sum_{j=1}^N J_{ij} x_i x_j,$$

where x_i is the population density of species i ; r_i is the intrinsic growth rate; N is the number of species and J_{ij} is the coupling between species i and j . Applying equation 7.1 to equation 7.2 we see that $\alpha_{ij} = J_{ij}$, such that the IM for the GLV model is equal to the coupling matrix. Therefore, by fitting the GLV model to population dynamics, the fitted parameters J_{ij} give us numeric estimates of interaction strength. To perform the fit we use the method detailed in section 7.2.4.

7.2.2 Population models

We use ordinary differential equation (ODE) models to simulate two species predator-prey dynamics. Therefore the mathematics presented in this section focuses on this case. However the framework may be easily extended to include other interaction types (competition and mutualism), and to model larger systems in a similar way. For a two species system the dynamics are governed by two coupled equations, one for each species. Importantly it is possible, using the IM introduced in section 7.2.1, to calculate the species interaction strengths from the model. The model equations take the general form:

$$(7.3) \quad \begin{aligned} \frac{dx_0}{dt} &= G_0(x_0) + \alpha_{01}x_1 H(x_0, x_1), \\ \frac{dx_1}{dt} &= G_1(x_1) + \alpha_{10}x_0 H(x_0, x_1) \end{aligned}$$

where species x_0 and x_1 are the population densities of the prey and the predator species respectively; the $G_i(x_i)$ are the intrinsic growth functions of each species; the α_{ij} are constant coefficients and $H(x_0, x_1)$ is the functional response (FR) of the predator. This form is standard in the literature [REFS] and many models may be expressed in this way by choosing different functional forms for G and H . The coefficients α_{01} and α_{10} are negative and positive respectively, such that the prey losses biomass, and the predator gains biomass as a result of the interaction. These coefficients may be used to introduce asymmetry into the interaction terms. For example it is common to choose $|\alpha_{01}| > |\alpha_{10}|$, to model the inefficiency of the predator in the conversion of biomass from the prey. For the intrinsic growth functions we use the functional forms:

$$(7.4) \quad G_0(x_0) = r_0 x_0 \left(1 - \frac{x_0}{K_c}\right)$$

$$(7.5) \quad G_1(x_1) = -r_1 x_1,$$

where $r_i \in \mathbb{R}^+$ is the intrinsic growth rate of species i , and K_c is the carrying capacity of the prey species. Therefore the predator has an exponential intrinsic mortality, whereas the prey species has logistic intrinsic growth. These use of these functional forms in predator-prey modelling was made popular by Rosenzweig and MacArthur [REF]. They are now widely used [REFS].

The FR defines the per-predator rate of consumption of prey. We focus on the forms proposed by Holling in the 1950s [REFs], which are also widely used [REFS]. However it is worth noting that various other forms have been proposed and there is an ongoing debate about which to use [REFS] (see discussion in section 7.6). There are three types of Holling FR, referred to as types I, II, and III. These can be expressed as:

$$(7.6) \quad H_I(x_0, x_1) = x_0,$$

$$(7.7) \quad H_{II}(x_0, x_1) = \frac{x_0}{x_0 + K_s},$$

$$(7.8) \quad H_{III}(x_0, x_1) = \frac{x_0^2}{x_0^2 + K_s^2},$$

where x_0 is the prey density, and K_s is the saturation constant for the predator, giving the prey density at which the per-predator consumption rate reaches half-maximum. We choose to narrow our investigation by focusing here on the first two forms: Holling type I and type II. Therefore we have two distinct simulation models, which we refer to as the *linear* and *holling* models.

Linear model. This model uses the type I FR. As is clear from equation 7.6, type I is the simplest of the Holling functions: the per-predator predation rate increases linearly as the abundance of available prey increases. The slope of this linear relationship given by the a_{ij} parameters in equations 7.5. The linear FR is the same as is used in the famous Lotka-Volterra equations [REF], meaning that our linear model may be expressed in GLV form (equation 7.2). We may rescale the equations of the linear model in order to reduce the number of parameters. This makes the local stability analysis simpler, and reduces the dimension of the search space when probing the equations numerically via simulation. The re-scaled equations are given by:

$$(7.9) \quad \frac{d\chi_0}{dt} = A\chi_0(1 - \chi_0) - B\chi_0\chi_1$$

$$(7.10) \quad \frac{d\chi_1}{dt} = -\chi_1 + C\chi_0\chi_1,$$

where χ_0 and χ_1 are the re-scaled prey and predator population densities respectively; the parameters $A, B, C \in \mathbb{R}^+$. The equilibrium population densities are given by:

$$(7.11) \quad \chi_0^* = \frac{1}{C}$$

$$(7.12) \quad \chi_1^* = \frac{A}{B} \left(1 - \frac{1}{C}\right),$$

Therefore χ_0^* is always positive, and χ_1^* is positive if $c > 0$. This is a requirement for physical realism, since it is not possible to have negative populations of species. In most applications it

is also required that this equilibrium is stable, to allow for the coexistence of species. We use these conditions on the equilibrium for parameter selection, which is discussed in section 7.2.3. By applying equation 7.1 to equations 7.9 and 7.10 we can evaluate the elements of the IM for the linear model. This gives:

$$(7.13) \quad \alpha_{linear} = \begin{bmatrix} -A & -B \\ C & 0 \end{bmatrix},$$

such that all the interaction strengths are constants, which is illustrated in figure 7.3.

Holling model. This model uses the type II FR, which is a non-linear function of prey density. As we can see from equation 7.7 and figure 7.2, this FR models predator saturation - individuals take a certain amount of time to process and digest prey - such that the response curve flattens out at high prey densities. The difference between the type I and II functions is illustrated in figure 7.2. We may perform a similar rescaling as we did with the linear model to reduce the number of parameters. The resulting equations for the holling model are given by:

$$(7.14) \quad \frac{d\chi_0}{dt} = A\chi_0(1-\chi_0) - \frac{B\chi_0\chi_1}{\chi_0+D}$$

$$(7.15) \quad \frac{d\chi_1}{dt} = -\chi_1 + \frac{C\chi_0\chi_1}{\chi_0+D},$$

where the saturation constant $D \in \mathbb{R}^+$, and the other symbols are the same as in equations 7.9 and 7.10. The equilibrium populations for this model are given by:

$$(7.16) \quad \chi_0^* = \frac{D}{C-1}$$

$$(7.17) \quad \chi_1^* = \frac{ACD(C-1-D)}{B(C-1)^2},$$

Therefore χ_0^* is positive if $C > 1$, χ_1^* is positive if $C - D > 1$. Again we use these conditions, and the requirement of stability, to constrain our choice of parameters (7.2.3). As we did for the linear model, we can evaluate the IM for the holling model, giving:

$$(7.18) \quad \alpha_{holling} = \begin{bmatrix} -A + \frac{B\chi_1}{(\chi_0+D)^2} & \frac{-B}{\chi_0+D} \\ \frac{CD}{(\chi_0+D)^2} & 0 \end{bmatrix},$$

such that the three non-zero elements of the IM are functions of prey density χ_0 , instead of constants. The shape of these interaction functions is shown in figure 7.2, and we will return to them in section 7.2.5.

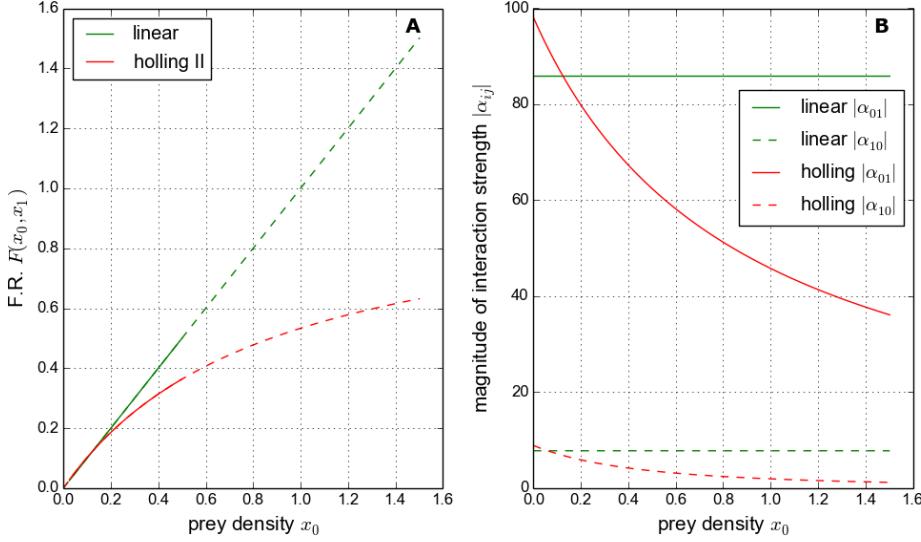


Figure 7.2: Example of (A) the functional response curve, and (B) the corresponding inter-specific interaction strengths for one parameter set of the *linear* model, and one of the *holling* model.

	A	B	C	D
linear	0.1 - 100	0.1 - 100	1 - 100	N/A
holling	0.1 - 100	0.1 - 100	1.1 - 100	0.1-99

Table 7.1: Ranges from which parameters were selected uniformly at random for the two ODE simulation models. The parameters are all allowed to vary over at least three orders of magnitude, to ensure that our investigation covers a large region of parameter space. The restrictions on parameters C and D ensure that it is always possible to achieve an equilibrium population of both species that is strictly positive (see equations 7.11, 7.12, 7.16, 7.17)

7.2.3 Simulation procedure

We apply a strict recipe when running simulations in order to ensure consistency and to allow comparison of our numerical results. Key to this is the control of certain variables across simulations, and also our method for parameter selection, both of which are discussed below. All simulations are run using the first-order forward Euler approximation to the ODE model. We use additive Gaussian noise to simulate *process error*. Therefore we implement the stochastic difference equation :

$$(7.19) \quad \chi_{i,t+1} = \chi_{i,t} + \Delta t \Delta \chi_{i,t} + \xi_{i,t},$$

where Δt is the integration time step, $\Delta \chi_i$ is given by the right hand side of ODE model being simulated (e.g. equations 7.9, 7.10 for the linear model); and the additive noise term

$\xi_{i,t} \sim \mathcal{N}(0, \sigma_{noise} \chi_{i,t} \Delta t)$. The value of σ_{noise} is quoted as *noise intensity* in what follows. In the event of stochastic extinction of either species, both population densities are reset to their initial conditions. The case where $\sigma_{noise} = 0$ is referred to as the *deterministic* model. In all the results presented the simulations were run with a time step $\Delta t = 10^{-4}$. All code was implemented in the language *Python*, and large computations were performed on the UoB HPC cluster *Blue Crystal* [REF].

The goal of fitting the GLV model to simulated population dynamics requires that the dynamics contain enough information to perform the fit - it is not possible to fit the a model if species populations are sitting at equilibrium. Therefore we follow the precedent set in [REF], such that all simulated dynamics of the *deterministic models* exhibit two ‘large amplitude’ oscillations about a stable equilibrium (see condition 2 below). Every simulation is run with the initial population densities set to half of their equilibrium value. This ensures that all systems start consistently away from equilibrium.

Parameter selection. We select an ensemble of 100 parameter sets for both simulation models (*linear* and *holling*). Parameters are selected uniformly at random from predefined ranges, which are given in table 7.1. This range ensures that a positive equilibrium population is possible (see equations 7.11, 7.12, 7.16, 7.17), but also allows for parameters to vary over at least three orders of magnitude so that our investigation covers a large region of parameter space. The selected parameters are then accepted if they meet the following conditions:

1. The equilibrium population is positive, and is locally a stable spiral (eigenvalues of the Jacobian have negative real part and complex conjugate imaginary part).
2. The deterministic dynamics exhibit at least two full rotations in the phase plane before relaxing to within 5% of the equilibrium (Euclidean distance in the phase plane).
3. The population densities do not differ by more than an order of magnitude, in the deterministic case.

The two parameter sets generated by the above procedure are used for the simulation results presented in section 7.3. All simulations , including those with $\sigma_{noise} \neq 0$, are run for the length of time T_{2P} required to achieve two full oscillations in the deterministic case, for that parameter set.

7.2.4 Numerical estimation method

To estimate the inter-specific interaction strengths we use a numerical method, adapted from [56], to fit the GLV model to the population dynamics. The method gives ‘best fit’ estimates of the GLV parameters, which include constant coefficients for the interaction strengths as we saw in section 7.2.1. We include here a derivation of the method, slightly adapted and simplified for

our purposes. Say the dynamics of the population density of a species x_i is governed by coupled differential equations of the form:

$$(7.20) \quad \dot{x}_i = r_i f_i(x_i) + \sum_{j=1}^N J_{i,j} g_{ij}(x_i, x_j),$$

where $\dot{x}_i = \frac{d}{dt}x_i$; N is the number of species in the system and i, j index the species. The $r_i, J_{i,j} \in \mathbb{R}^+$ are constants, and the functions f_i and g_{ij} are known. This form looks familiar, indeed all of the ODE models discussed so far may be expressed in this form - there is an intrinsic growth term, and a linear sum of interaction terms. To express the GLV model (equation 7.2) in this form, we have:

$$(7.21) \quad f_i(x_i) = x_i$$

$$(7.22) \quad g_{ij}(x_i, x_j) = x_i x_j,$$

It would be possible to use this method to fit models other than the GLV, so long as the functions f_i and g_{ij} are *known and parametrised*. Since the functions f_i and g_{ij} are known there are $N+1$ unknowns in equation 7.20: r_i and $J_{i,j}$ for $j = 1, \dots, N$. Therefore, if we knew the exact values of \dot{x}_i, x_i and the x_j 's, at $N+1$ time points, then we could solve the equation for r_i and the $J_{i,j}$'s. However in any practical application our knowledge of the system is not *exact*; the system is subject to noise; and the model may be an imperfect description of the dynamics. So the equation cannot be solved exactly. We must look for an approximate solution. To do this the full state of the system is sampled at $M+1$ time points t_m for $m \in 1, \dots, M, M+1$. These samples are used to construct estimates for the states x_i and their time-derivatives \dot{x}_i at M intermediate time points, for every species i .

The simplest way to estimate the time-derivatives is to take the finite difference between observations at two consecutive time points, giving estimates:

$$(7.23) \quad \hat{x}_i(\tau_m) := \frac{x_i(t_m) - x_i(t_{m-1})}{t_m - t_{m-1}},$$

where $\tau_m \in \mathbb{R}, m \in \{1, \dots, M\}$ is the midpoint of the two time-points:

$$(7.24) \quad \tau_m := \frac{t_{m-1} + t_m}{2}.$$

To evaluate the functions f_i, g_{ij} at these new time-points we must estimate the states $x_i(\tau_m)$ from our observations. We use the linear interpolation:

$$(7.25) \quad \hat{x}_i(\tau_m) := \frac{x_i(t_{m-1}) + x_i(t_m)}{2}.$$

So from equation 7.20 we can now construct M equations using our $M + 1$ samples:

$$(7.26) \quad \hat{x}_i(\tau_m) = r_i f_i(\hat{x}_i(\tau_m)) + \sum_{j=1}^N J_{i,j} g_{ij}(\hat{x}_i(\tau_m), \hat{x}_j(\tau_m)).$$

We now simplify the notation such that equation 7.26 may be written

$$(7.27) \quad \hat{x}_{i,m} = r_i f_{i,m} + \sum_{j=1}^N J_{i,j} g_{ij,m},$$

where the subscripts i, j indicate the species, and m indicates the time-point τ_m for which the equation holds. This system of M equations can be expressed in matrix form:

$$(7.28) \quad X_i = J_i G_i,$$

where we have

$$(7.29) \quad X_i = \begin{pmatrix} \hat{x}_{i,1} & \hat{x}_{i,2} & \cdots & \hat{x}_{i,M} \end{pmatrix} \in \mathbb{R}^{1 \times M},$$

$$(7.30) \quad J_i = \begin{pmatrix} r_i & J_{i,1} & J_{i,2} & \cdots & J_{i,N} \end{pmatrix} \in \mathbb{R}^{1 \times (N+1)},$$

$$(7.31) \quad G_i = \begin{pmatrix} f_{i,1} & f_{i,2} & \cdots & f_{i,M} \\ g_{i,1,1} & g_{i,1,2} & \cdots & g_{i,1,M} \\ g_{i,2,1} & g_{i,2,2} & \cdots & g_{i,2,M} \\ \vdots & \vdots & \ddots & \vdots \\ g_{i,N,1} & g_{i,N,2} & \cdots & g_{i,N,M} \end{pmatrix} \in \mathbb{R}^{(N+1) \times M}.$$

The system 7.28 has $N + 1$ unknowns ($J_{i,k}$ for $k = 1, \dots, N + 1$) and M equations. In the case when $M > N + 1$ the system is over constrained and there is no exact solution in general. We look for an approximate solution \hat{J}_i that minimises the error between the LHS and RHS of equation 7.28. We take the error function:

$$(7.32) \quad E_i(\hat{J}_i) = \sum_{m=1}^M (X_{i,m} - \sum_{k=1}^{N+1} \hat{J}_{i,k} G_{i,k,m})^2,$$

which we want to minimise with respect to the matrix elements $\hat{J}_{i,k}$. That is

$$(7.33) \quad \frac{\partial}{\partial \hat{J}_{i,k}} E_i(\hat{J}_i) \stackrel{!}{=} 0.$$

By taking the derivative of the RHS of equation 7.32 we have that:

$$\begin{aligned}\frac{\partial}{\partial \hat{J}_{i,k'}} E_i(\hat{J}_i) &= \frac{\partial}{\partial \hat{J}_{i,k'}} [\sum_{m=1}^M (X_{i,m} - \sum_{k=1}^N \hat{J}_{i,k} G_{i,k,m})^2] \\ &= -2 \sum_{m=1}^M [(X_{i,m} - \sum_{k=1}^N \hat{J}_{i,k} G_{i,k,m}) G_{i,k',m}]\end{aligned}$$

To find the minimum of the error function we equate this to zero, giving:

$$\begin{aligned}0 &= \sum_{m=1}^M (-X_{i,m} G_{i,k',m} + G_{i,k',m} \sum_{k=1}^N \hat{J}_{i,k} G_{i,k,m}) \\ &= (-X_i G_i^T)_{k'} + \sum_{m=1}^M G_{i,k',m} (\hat{J}_i G_i)_m \\ &= (-X_i G_i^T)_{k'} + \sum_{m=1}^M (\hat{J}_i G_i)_m G_{i,m,k'}^T \\ (7.34) \quad &= -X_i G_i^T + \hat{J}_i G_i G_i^T\end{aligned}$$

Therefore we conclude that:

$$(7.35) \quad \hat{J}_i = X_i G_i^T (G_i G_i^T)^{-1},$$

which, in our case, is the analytic form for the best estimate of the row corresponding to species i in parameter matrix of the GLV model. For a two species system, by applying equation 7.35 to each species, we can obtain the full set of GLV parameter estimates:

$$(7.36) \quad \hat{J} = \begin{pmatrix} \hat{J}_{0,0} & \hat{J}_{0,1} \\ \hat{J}_{1,0} & \hat{J}_{1,1} \end{pmatrix},$$

and

$$(7.37) \quad \hat{r} = \begin{pmatrix} \hat{r}_0 & \hat{r}_1 \end{pmatrix}.$$

We perform the computation by constructing the matrices X_i and G_i for both species, and do the matrix multiplication using the *Python* package *numpy*[REF]. The fact that the error minimisation has an analytic solution makes it very computationally efficient, allowing us to perform many replicate calculations. However the performance of the method may be lower than other, more computationally expensive, model fitting algorithms (see discussion section 7.6). It is possible to assess the goodness of fit achieved by evaluating the error function (equation 7.32).

7.2.5 Examples

Here we show examples of the dynamics of both models, both with and without noise..For the holling model we also show the variability in interaction strengths during the simulations..We also present a table with the results of the GLV, comparing them to the simulation interaction strengths..

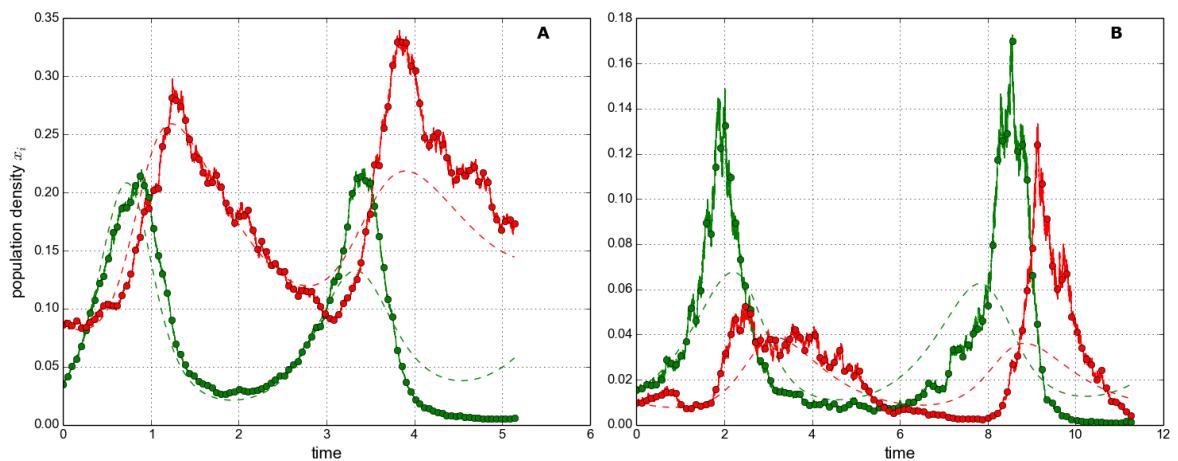


Figure 7.3: Example linear dynamics. 100 sampels. Two different parameter sets. A: noise=20. B:noise=50.

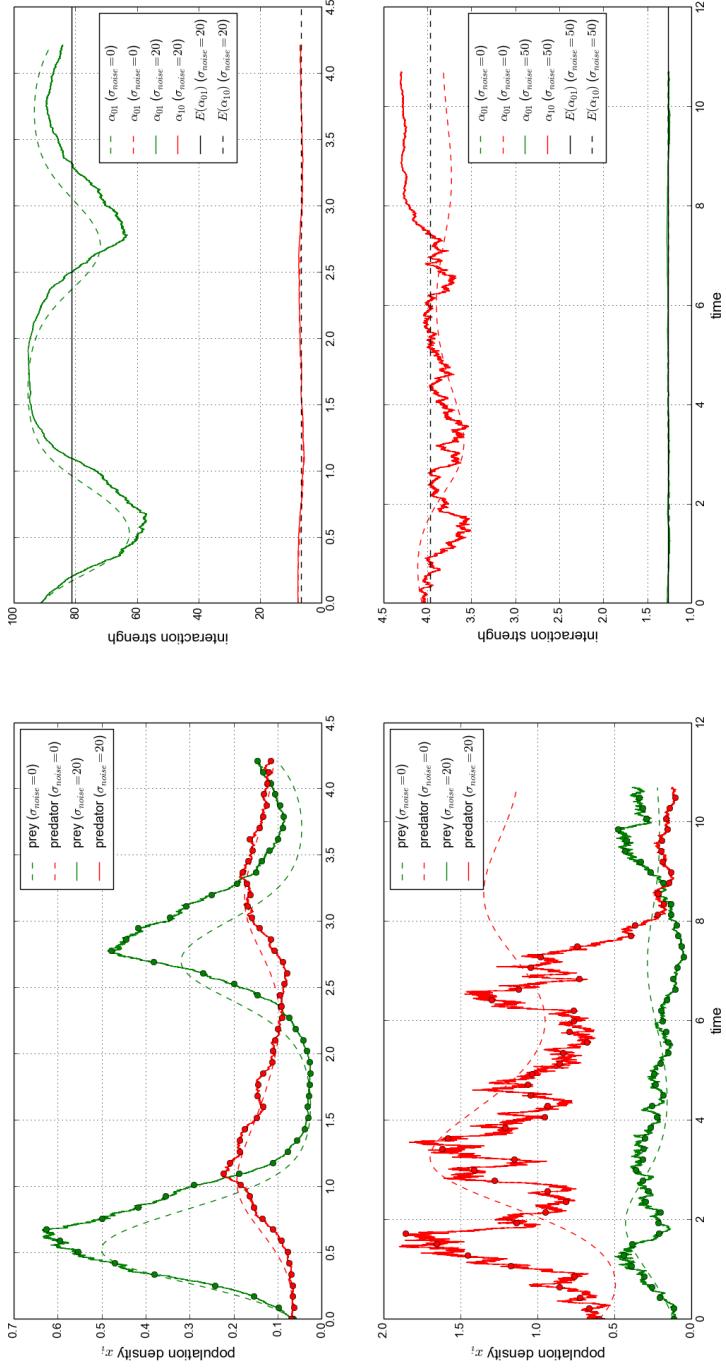


Figure 7.4: Example Holling II dynamics.

7.3 Results

In this sections we characterise the numerical performance of the method, described in section 7.2.4, for estimating the strength of interactions between species. The method is tested on the dynamics of two different ODE systems: a Lotka-Volterra (LV) and a Holling type II (HII) system. In the first case it is simply a test of a model fitting procedure. This is because the method works by fitting a generalised Lotka-Volterra (GLV) model to the dynamics, and he LV systems can be expressed as a GLV systems. Therefore we are simply simulating using one model, and then testing a method of estimating the model parameters form the simulated dynamics. We test the effects of noise and sampling frequency. In the second case, the HII system cannot be expressed as a GLV system. Therefore the GLV model that we fit can only approximate the dynamics and we cannot make a direct comparison between the parameters of the simulation model and the GLV model used for estimation. In this case we comapre to the mean interaction strengths (see section 7.3.2).

7.3.1 Linear model

Initially we run repeated simulations of the LV model using a single parameter set. We investigate how the numerical estimates of the model parameters respond to two variables: the level of noise in the simulations; and the number of samples used for estimation. Other variables are held constant using the simulation procedure described in section 7.2.3. We then generalise these results by looking at the relative error in the estimates, for repeated simulations using an ensemble of 100 selected parameter sets (as described in section 7.2.3).

Single parameter set. Here we can make direct comparison between model parameters. The GLV model for two species has six parameters: $r_0, r_1, J_{00}, J_{01}, J_{10}, J_{11}$. These correspond respectively to the following constant values of the LV system used for simulations: $A, -1, -A, -B, C, 0$ (see equation ??). In general we find that the numerical estimates perform well at low noise intensities and poorly at high noise intensities. This is illustrated in figures 7.5 and 7.6. We also find that the estimates improve with the number of samples used, up to a point. Beyond this point the use of more samples does little to improve to estimates, and in some cases makes them worse. This behaviour is illustrated in figures 7.7 and 7.8. These patterns were found to hold across all parameter sets investiagted, but are only shown using a single parameter set here for clarity.

In panel **A** of figures 7.5 and 7.6 we see that the mean value of the estimates approaches the true value for low noise and, in panel **B** that the variance in the estimates approaches zero. This tells us that the method consistently gives a good fit of the GLV to the dynamics of the LV system, even when only 100 sample points are used (figure 7.5). As the noise intesitiy is increased the mean values of the estimates deviate from the true values, and the standard deviation in the estimates increases. Comparing the two figures we see that the response to noise is very similar

whether 100 or 10,000 samples are used. A notable exception to this is a spike in the variance in panel **B** of figure. However this appears to be a single statistically anomalous result and not part of the trend. Panel **C** of both figures shows that the error function, which is minimised by the estimation method, increases with noise for both species. This cannot be directly compared between the two plots because of the different number of samples used. However it indicates that in both cases (100 and 10,000 samples) the quality of the fit is high in the deterministic case, and decreases with noise.

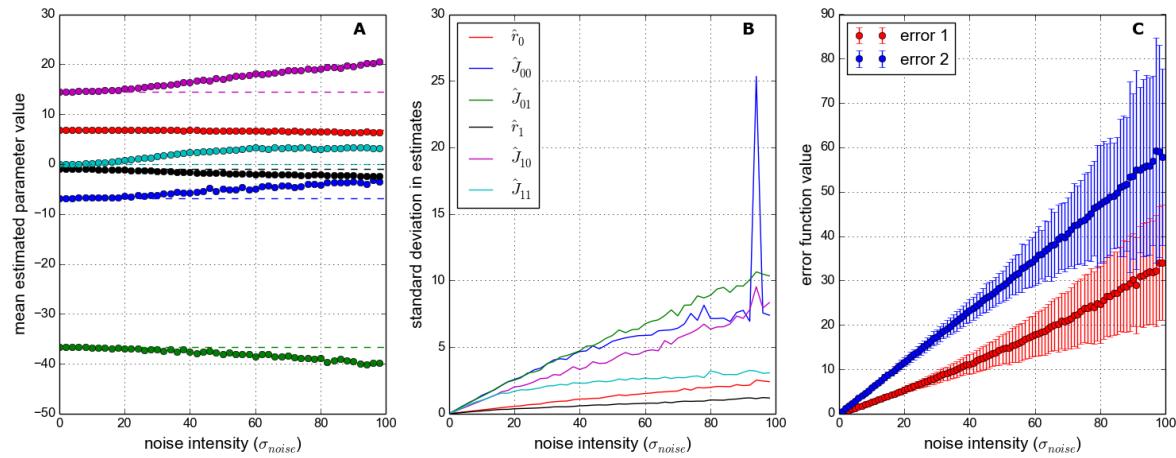


Figure 7.5: Effect of noise on numerical estimates. Here the method uses 100 samples from simulated dynamics. All simulations run using the LV model with a single parameter set. The noise intensity varies between 0 (deterministic) and 100. See section 7.2.5 for an intuition of how noisy this is. 1000 repeat simulations run at each noise level. **Panel A:** Mean estimated parameter values (each dot representing mean over 1000 repeats). The ‘true’ parameter values of the simulation model are shown by dashed lines. **Panel B:** Standard deviation in estimates. **Panel C:** Value of the error functions used in the estimation method, one for each species. The dots show the mean error, and the bars show \pm one standard deviation.

We now look at how the estimates respond to the number of samples used, in the cases of low and high noise intensity. Figure 7.7 shows the low noise case, with $\sigma_{noise} = 10$. In panel **A** we see that the mean value of the estimates quickly converges to close to the true parameter values, as the number of samples increases. Panel **B** shows that the standard deviation in the estimates quickly becomes small, but non-zero. Above about 32 samples there is no visible improvement in the estimates, as measured by the mean or standard deviation. In figure 7.8 we see the effect of a higher noise intensity. Here we have $\sigma_{noise} = 50$. Panel **A** shows that the estimates do not converge on the true parameter values, even for large numbers of samples. Also the standard deviation in the estimates, shown in panel **B**, is higher than in the low noise case. Again we find that there is little, if any, improvement in the estimates beyond about 32 samples.

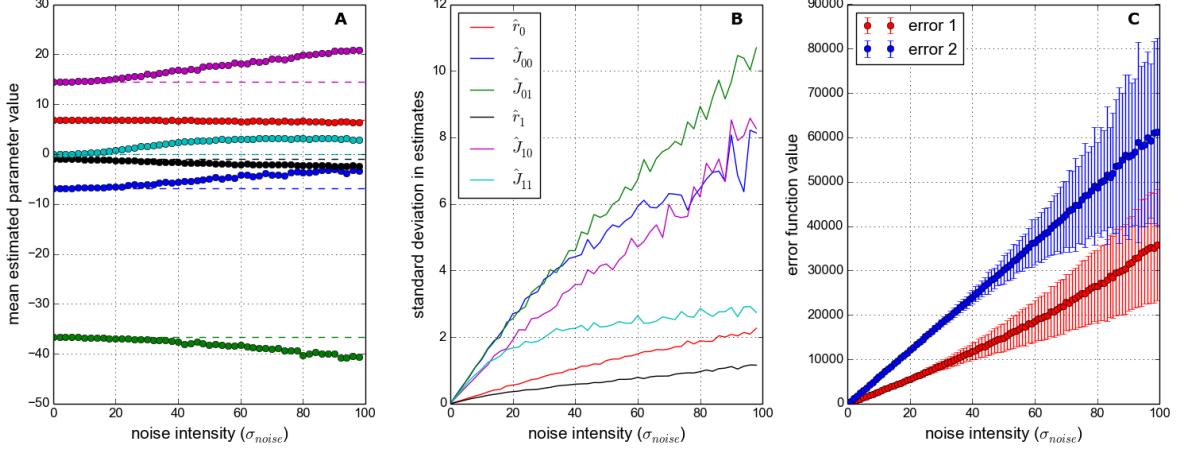


Figure 7.6: Exactly as in figure 7.5 but using 10,000 samples from the simulated dynamics.

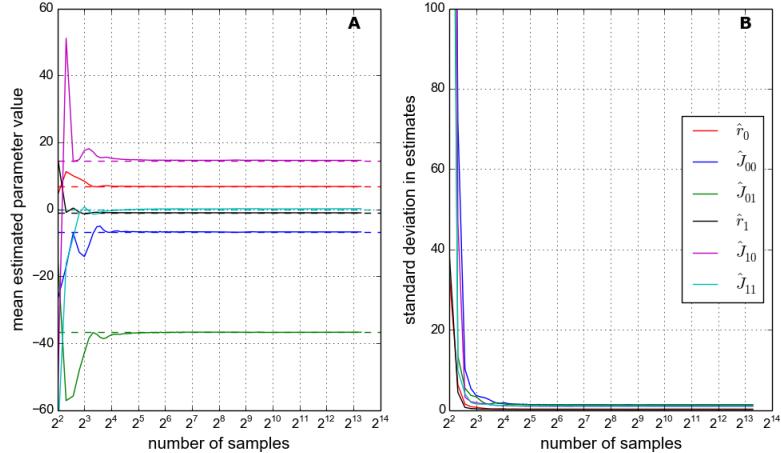


Figure 7.7: Effect of the number of samples on numerical estimates. All simulations run using the LV model with a single parameter set. The noise intensity $\sigma_{noise} = 10$. Number of samples ranges from 4 to 10,000. Samples drawn from simulated dynamics at equal intervals. 1000 repeat simulations for each number of samples. **Panel A:** Solid lines show mean estimated parameter values. Dashed lines show the ‘true’ parameter values of the simulation model. **Panel B:** Standard deviation in estimates.

Ensemble of parameter sets. Run 10 repeats for each of 100 parameter sets. In general the trends described above hold across the ensemble..

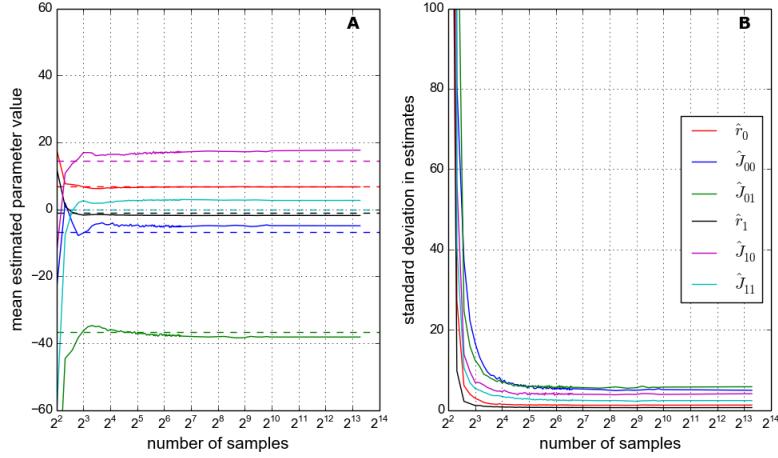


Figure 7.8: Exactly as in figure 7.7 but with noise intensity $\sigma_{noise} = 50$.

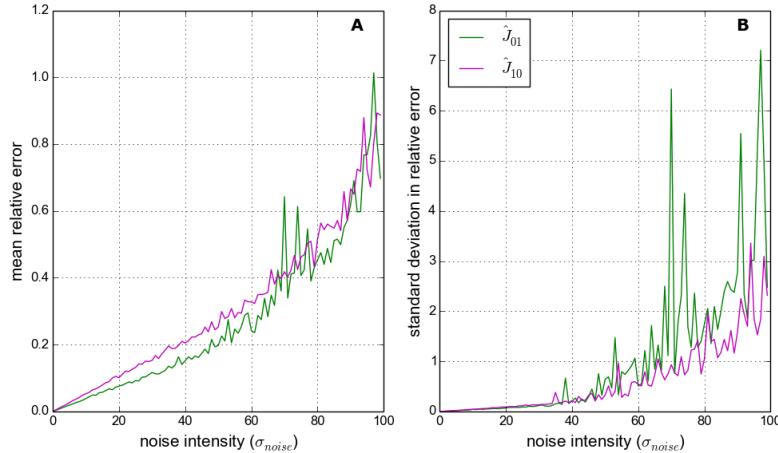


Figure 7.9: Nonsense. 1000 samples used.

7.3.2 Holling model

7.3.3 Range sampling

7.4 TEMP : other results

This section shows some plots which I was not planning to put into the thesis but are worth discussing..

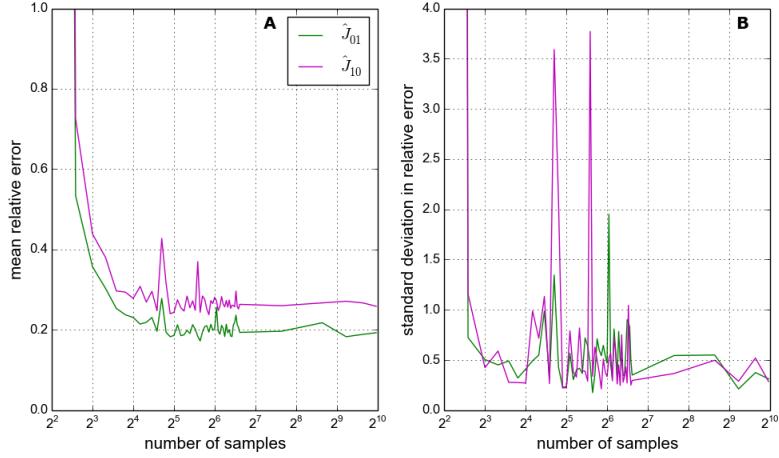


Figure 7.10: Nonsense. Noise is 50.

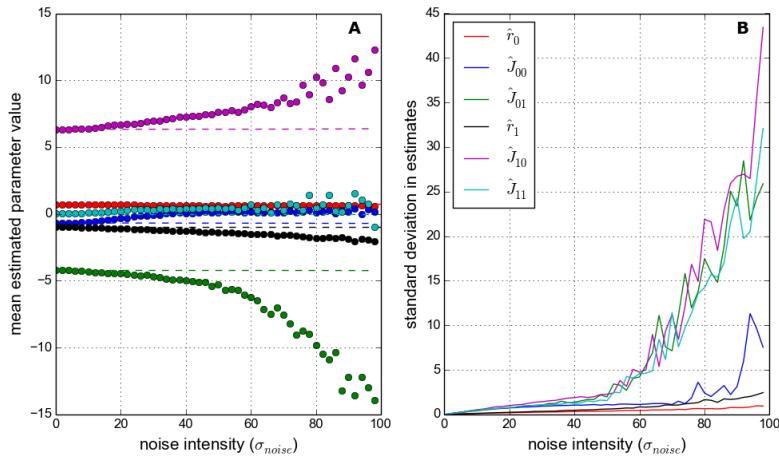


Figure 7.11: Nsamples is 10,000.

7.5 Application to IBM (optional)

7.5.1 Testing functionl response

7.5.2 Extend methodology (3 and 4 species)

7.5.3 Results

7.6 Discussion

Points referenced in text above, make sure to discuss them!..

- Discuss how this methodology could be used on empirical data...

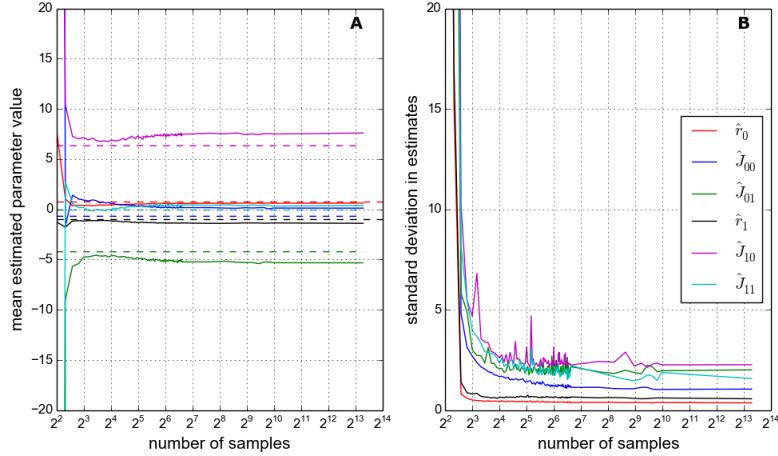


Figure 7.12: Noise is 50.

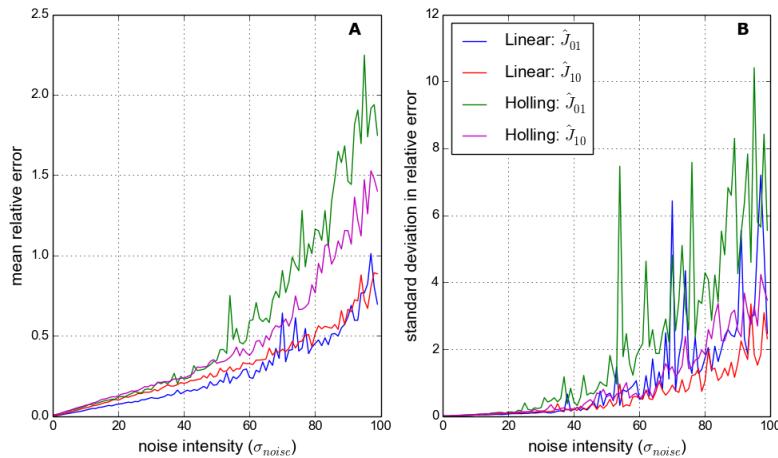


Figure 7.13: Noise is 50.

- Limitations of ODE models (non-spatial, response to debate on FR)
- Possibility of extending to more than two species systems (if this is actually done then change ref in text above)
- discussion of other forms of FR (not H) - or discussed already in intro?
- good GLV fit to LV even with 100 sample points - realistic?
- how good is the method of model fitting. Discuss more computationally expensive options (mentioned in section on Timme method)

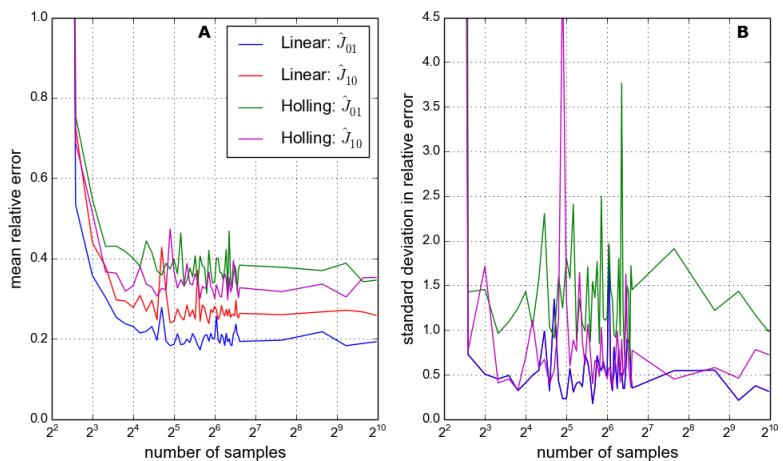


Figure 7.14: Noise is 50.

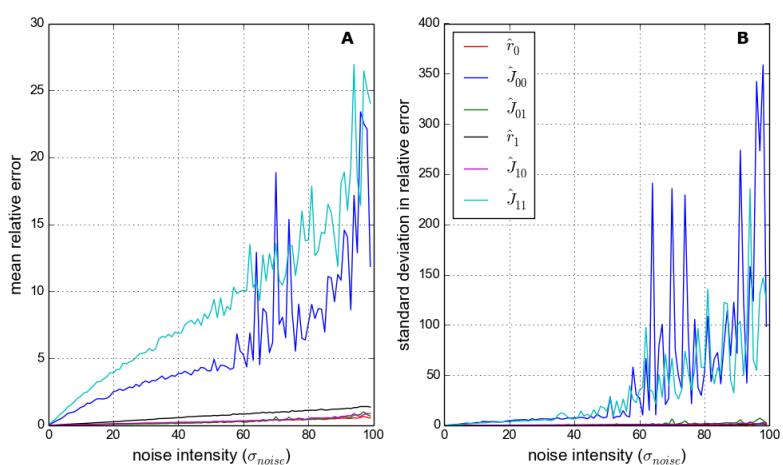


Figure 7.15: Nonsense. 1000 samples used.

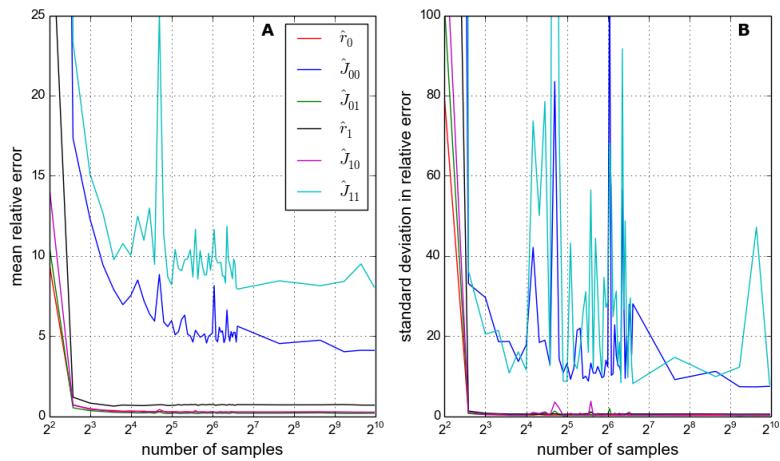


Figure 7.16: Nonsense. Noise is 50.

BIBLIOGRAPHY

- [1] D. AI, P. DESJARDINS-PROULX, C. CHU, AND G. WANG, *Immigration, local dispersal limitation, and the repeatability of community composition under neutral and niche dynamics*, (2012).
- [2] M. ALBRECHT, P. DUELLI, B. SCHMID, AND C. B. MÜLLER, *Interaction diversity within quantified insect food webs in restored and adjacent intensively managed meadows*, Journal of Animal Ecology, 76 (2007), pp. 1015–1025.
- [3] J.-F. ARNOLDI, M. LOREAU, AND B. HAEGEMAN, *Resilience, reactivity and variability: A mathematical comparison of ecological stability measures*, Journal of theoretical biology, 389 (2016), pp. 47–59.
- [4] J. BASCOMPTE AND P. JORDANO, *Plant-animal mutualistic networks: the architecture of biodiversity*, Annual Review of Ecology, Evolution, and Systematics, (2007), pp. 567–593.
- [5] E. BENINCÀ, J. HUISMAN, R. HEERKLOSS, K. D. JÖHNK, P. BRANCO, E. H. VAN NES, M. SCHEFFER, AND S. P. ELLNER, *Chaos in a long-term experiment with a plankton community*, Nature, 451 (2008), pp. 822–825.
- [6] E. L. BERLOW, A.-M. NEUTEL, J. E. COHEN, P. C. DE RUITER, B. EBENMAN, M. EMMERSON, J. W. FOX, V. A. JANSEN, J. IWAN JONES, G. D. KOKKORIS, ET AL., *Interaction strengths in food webs: issues and opportunities*, Journal of animal ecology, 73 (2004), pp. 585–598.
- [7] D. BOSSIO, M. S. GIRVAN, L. VERCHOT, J. BULLIMORE, T. BORELLI, A. ALBRECHT, K. SCOW, A. S. BALL, J. PRETTY, AND A. M. OSBORN, *Soil microbial community response to land use change in an agricultural landscape of western kenya*, Microbial ecology, 49 (2005), pp. 50–62.
- [8] T. D. BROCK, *The ecosystem and the steady state*, Bioscience, 17 (1967), pp. 166–169.
- [9] M. W. CADOTTE, *Dispersal and species diversity: A meta-analysis*, The American Naturalist, 167 (2006), pp. 913–924.
- [10] R. CONDIT, R. A. CHISHOLM, AND S. P. HUBBELL, *Thirty years of forest census at barro colorado and the importance of immigration in maintaining diversity*, (2012).

BIBLIOGRAPHY

- [11] C. DARWIN AND W. F. BYNUM, *The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life*, AL Burt, 2009.
- [12] K. L. DRURY, J. D. SUTER, J. B. RENDALL, A. M. KRAMER, AND J. M. DRAKE, *Immigration can destabilize tri-trophic interactions: implications for conservation of top predators*, Theoretical Ecology, (2015), pp. 1–12.
- [13] J. A. DUNNE, R. J. WILLIAMS, AND N. D. MARTINEZ, *Food-web structure and network theory: the role of connectance and size*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 12917–12922.
- [14] D. M. EVANS, M. J. POCOCK, AND J. MEMMOTT, *The robustness of a network of ecological networks to habitat loss*, Ecology letters, 16 (2013), pp. 844–852.
- [15] E. S. EVELEIGH, K. S. MCCANN, P. C. MCCARTHY, S. J. POLLOCK, C. J. LUCAROTTI, B. MORIN, G. A. McDougall, D. B. STRONGMAN, J. T. HUBER, J. UMBANHOWAR, ET AL., *Fluctuations in density of an outbreak species drive diversity cascades in food webs*, Proceedings of the National Academy of Sciences, 104 (2007), pp. 16976–16981.
- [16] R. M. EWERS, R. K. DIDHAM, L. FAHRIG, G. FERRAZ, A. HECTOR, R. D. HOLT, V. KAPOS, G. REYNOLDS, W. SINUN, J. L. SNADDON, ET AL., *A large-scale forest fragmentation experiment: the stability of altered forest ecosystems project*, Philosophical Transactions of the Royal Society of London B: Biological Sciences, 366 (2011), pp. 3292–3302.
- [17] J. A. FOLEY, R. DEFRIES, G. P. ASNER, C. BARFORD, G. BONAN, S. R. CARPENTER, F. S. CHAPIN, M. T. COE, G. C. DAILY, H. K. GIBBS, ET AL., *Global consequences of land use*, science, 309 (2005), pp. 570–574.
- [18] C. FONTAINE, P. R. GUIMARÃES, S. KÉFI, N. LOEUILLE, J. MEMMOTT, W. H. VAN DER PUTTEN, F. J. VAN VEEN, AND E. THÉBAULT, *The ecological and evolutionary implications of merging different types of networks*, Ecology Letters, 14 (2011), pp. 1170–1181.
- [19] M. A. FORTUNA AND J. BASCOMPTE, *Habitat loss and the structure of plant–animal mutualistic networks*, Ecology Letters, 9 (2006), pp. 281–286.
- [20] M. A. FORTUNA, A. KRISHNA, AND J. BASCOMPTE, *Habitat loss and the disassembly of mutualistic networks*, Oikos, 122 (2013), pp. 938–942.
- [21] B. L. FOSTER, T. L. DICKSON, C. A. MURPHY, I. S. KAREL, AND V. H. SMITH, *Propagule pools mediate community assembly and diversity-ecosystem regulation along a grassland productivity gradient*, Journal of Ecology, 92 (2004), pp. 435–449.

BIBLIOGRAPHY

- [22] R. P. FRECKLETON, A. R. WATKINSON, AND M. REES, *Measuring the importance of competition in plant communities*, Journal of ecology, 97 (2009), pp. 379–384.
- [23] A. GONZALEZ, B. RAYFIELD, AND Z. LINDO, *The disentangled bank: how loss of habitat fragments and disassembles ecological networks*, American Journal of Botany, 98 (2011), pp. 503–516.
- [24] T. GROSS, L. RUDOLF, S. A. LEVIN, AND U. DIECKMANN, *Generalized models reveal stabilizing factors in food webs*, Science, 325 (2009), pp. 747–750.
- [25] N. M. HADDAD, L. A. BRUDVIG, J. CLOBERT, K. F. DAVIES, A. GONZALEZ, R. D. HOLT, T. E. LOVEJOY, J. O. SEXTON, M. P. AUSTIN, C. D. COLLINS, ET AL., *Habitat fragmentation and its lasting impact on earth’s ecosystems*, Science Advances, 1 (2015), p. e1500052.
- [26] M. HAGEN, W. D. KISSLING, C. RASMUSSEN, D. CARSTENSEN, Y. DUPONT, C. KAISER-BUNBURY, E. O’GORMAN, J. OLESEN, M. DE AGUIAR, L. BROWN, ET AL., *Biodiversity, species interactions and ecological networks in a fragmented world*, Advances in Ecological Research, 46 (2012), pp. 89–120.
- [27] I. HANSKI, G. A. ZURITA, M. I. BELLOCQ, AND J. RYBICKI, *Species–fragmented area relationship*, Proceedings of the National Academy of Sciences, 110 (2013), pp. 12715–12720.
- [28] S. HAZLEDINE, *The Analysis and modelling of chaotic calcium oscillations*, PhD thesis, Department for Computational Systems Biology, John Innes Centre, 2009.
- [29] H. P. HSU, *Schaum’s outline of theory and problems of probability, random variables, and random processes*, McGraw-Hill, 1997.
- [30] D. H. JANZEN, *The deflowering of central America*, Nat. Hist, 83 (1974), pp. 48–53.
- [31] R. KAARTINEN AND T. ROSLIN, *Shrinking by numbers: landscape context affects the species composition but not the quantitative structure of local food webs*, Journal of Animal Ecology, 80 (2011), pp. 622–631.
- [32] S. KÉFI, E. L. BERLOW, E. A. WIETERS, S. A. NAVARRETE, O. L. PETCHEY, S. A. WOOD, A. BOIT, L. N. JOPPA, K. D. LAFFERTY, R. J. WILLIAMS, ET AL., *More than a meal—integrating non-feeding interactions into food webs*, Ecology letters, 15 (2012), pp. 291–300.
- [33] C. KLAUSMEIER, *Habitat destruction and extinction in competitive and mutualistic meta-communities*, Ecology Letters, 4 (2001), pp. 57–63.
- [34] T. M. KNIGHT, M. W. MCCOY, J. M. CHASE, K. A. MCCOY, AND R. D. HOLT, *Trophic cascades across ecosystems*, Nature, 437 (2005), pp. 880–883.

BIBLIOGRAPHY

- [35] N. M. KORFANTA, W. D. NEWMARK, AND M. J. KAUFFMAN, *Long-term demographic consequences of habitat fragmentation to a tropical understory bird community*, Ecology, 93 (2012), pp. 2548–2559.
- [36] C. KREMEN, N. M. WILLIAMS, M. A. AIZEN, B. GEMMILL-HERREN, G. LEBUHN, R. MINCKLEY, L. PACKER, S. G. POTTS, I. STEFFAN-DEWENTER, D. P. VAZQUEZ, ET AL., *Pollination and other ecosystem services produced by mobile organisms: a conceptual framework for the effects of land-use change*, Ecology Letters, 10 (2007), pp. 299–314.
- [37] D. KWIATKOWSKI, P. C. PHILLIPS, P. SCHMIDT, AND Y. SHIN, *Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?*, Journal of econometrics, 54 (1992), pp. 159–178.
- [38] M. LOREAU AND N. MOUQUET, *Immigration and the maintenance of local species diversity*, The American Naturalist, 154 (1999), pp. 427–440.
- [39] M. LURGI, D. MONTOYA, AND J. M. MONTOYA, *The effects of space and diversity of interaction types on the stability of complex ecological networks*, Theoretical Ecology, (2015), pp. 1–11.
- [40] R. M. MAY, *Will a large complex system be stable?*, Nature, 238 (1972), pp. 413–414.
- [41] C. J. MELIÁN AND J. BASCOMPTE, *Food web structure and habitat loss*, Ecology Letters, 5 (2002), pp. 37–46.
- [42] J. MEMMOTT, R. GIBSON, L. CARVALHEIRO, K. HENSON, R. HELENO, M. LOPEZARAIZA, S. PEARCE, AND S. PEARCE, *The conservation of ecological interactions*, Insect Conservation Biology. The Royal Entomological Society, London, (2007), pp. 226–44.
- [43] D. MONTOYA, M. YALLOP, AND J. MEMMOTT, *Functional group diversity increases with modularity in complex food webs*, Nature communications, 6 (2015).
- [44] A. MOUGI AND M. KONDOH, *Diversity of interaction types and ecological community stability*, Science, 337 (2012), pp. 349–351.
- [45] N. MOUQUET, P. LEADLEY, J. MÉRIGUET, AND M. LOREAU, *Immigration and local competition in herbaceous plant communities: a three-year seed-sowing experiment*, Oikos, 104 (2004), pp. 77–90.
- [46] E. J. O'GORMAN AND M. C. EMMERSON, *Perturbations to trophic interactions and the stability of complex food webs*, Proceedings of the National Academy of Sciences, 106 (2009), pp. 13393–13398.

- [47] C. PARMESAN, N. RYRHOLM, C. STEFANESCU, J. K. HILL, C. D. THOMAS, H. DESCIMON, B. HUNTLEY, L. KAILA, J. KULLBERG, T. TAMMARU, ET AL., *Poleward shifts in geographical ranges of butterfly species associated with regional warming*, Nature, 399 (1999), pp. 579–583.
- [48] C. PARMESAN AND G. YOHE, *A globally coherent fingerprint of climate change impacts across natural systems*, Nature, 421 (2003), pp. 37–42.
- [49] M. J. POCOCK, D. M. EVANS, AND J. MEMMOTT, *The robustness and restoration of a network of ecological networks*, Science, 335 (2012), pp. 973–977.
- [50] M. PRIESTLEY AND T. S. RAO, *A test for non-stationarity of time-series*, Journal of the Royal Statistical Society. Series B (Methodological), (1969), pp. 140–149.
- [51] T. PUETTKER, A. A. BUENO, C. DOS SANTOS DE BARROS, S. SOMMER, AND R. PARDINI, *Immigration rates in fragmented landscapes—empirical evidence for the importance of habitat amount for species persistence*, PLoS One, 6 (2011), p. e27963.
- [52] R CORE TEAM, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013.
- [53] W. J. RIPPLE AND R. L. BESCHTA, *Trophic cascades in yellowstone: The first 15 years after wolf reintroduction*, Biological Conservation, 145 (2012), pp. 205–213.
- [54] S. E. SAID AND D. A. DICKEY, *Testing for unit roots in autoregressive-moving average models of unknown order*, Biometrika, 71 (1984), pp. 599–607.
- [55] A. SAUVE, C. FONTAINE, AND E. THÉBAULT, *Structure–stability relationships in networks combining mutualistic and antagonistic interactions*, Oikos, 123 (2014), pp. 378–384.
- [56] S. G. SHANDILYA AND M. TIMME, *Inferring network topology from complex dynamics*, New Journal of Physics, 13 (2011), p. 013004.
- [57] R. V. SOLE AND J. M. MONTOYA, *Ecological network meltdown from habitat loss and fragmentation*, Ecological Networks: Linking Structure to Dynamics in Food Webs, (2006), pp. 305–323.
- [58] B. J. SPIESMAN AND B. D. INOUYE, *Habitat loss alters the architecture of plant-pollinator interaction networks*, Ecology, 94 (2013), pp. 2688–2696.
- [59] D. STOUFFER, J. CAMACHO, R. GUIMERA, C. NG, AND L. NUNES AMARAL, *Quantitative patterns in the structure of model and empirical food webs*, Ecology, 86 (2005), pp. 1301–1311.

BIBLIOGRAPHY

- [60] A. TEYSSÈDRE AND A. ROBERT, *Contrasting effects of habitat reduction, conversion and alteration on neutral and non neutral biological communities*, Oikos, 123 (2014), pp. 857–865.
- [61] E. THÉBAULT AND C. FONTAINE, *Stability of ecological communities and the architecture of mutualistic and trophic networks*, Science, 329 (2010), pp. 853–856.
- [62] D. TILMAN, R. M. MAY, C. L. LEHMAN, AND M. A. NOWAK, *Habitat destruction and the extinction debt*, (1994).
- [63] J. M. TYLIANAKIS, T. TSCHARNTKE, AND O. T. LEWIS, *Habitat modification alters the structure of tropical host-parasitoid food webs*, Nature, 445 (2007), pp. 202–205.
- [64] A. VALIENTE-BANUET, M. A. AIZEN, J. M. ALCÁNTARA, J. ARROYO, A. COCUCCI, M. GALETTI, M. B. GARCÍA, D. GARCÍA, J. M. GÓMEZ, P. JORDANO, ET AL., *Beyond species loss: the extinction of ecological interactions in a changing world*, Functional Ecology, 29 (2015), pp. 299–307.
- [65] S. WANG AND M. LOREAU, *Ecosystem stability in space: α , β and γ variability*, Ecology letters, 17 (2014), pp. 891–901.
- [66] B. H. WARREN, D. SIMBERLOFF, R. E. RICKLEFS, R. AGUILÉE, F. L. CONDAMINE, D. GRAVEL, H. MORLON, N. MOUQUET, J. ROSINDELL, J. CASQUET, ET AL., *Islands as model systems in ecology and evolution: prospects fifty years after macarthur-wilson*, Ecology letters, 18 (2015), pp. 200–217.
- [67] R. J. WILLIAMS AND N. D. MARTINEZ, *Simple rules yield complex food webs*, Nature, 404 (2000), pp. 180–183.
- [68] J. T. WOOTTON AND M. EMMERSON, *Measurement of interaction strength in nature*, Annual Review of Ecology, Evolution, and Systematics, (2005), pp. 419–444.
- [69] C. E. ZARTMAN AND H. E. NASCIMENTO, *Are habitat-tracking metacommunities dispersal limited? inferences from abundance-occupancy patterns of epiphylls in amazonian forest fragments*, Biological Conservation, 127 (2006), pp. 46–54.