

CHAPTER



PERSISTENCE AND STATIONARITY

4.1 Introduction

In the chapter 3 we analysed in detail how simulated communities responded to two habitat loss scenarios. These simulations used the *default parameter values* for the IBM model, as published in [68]. An important feature of these simulations was that they exhibited almost no species extinctions, even up to 90% habitat loss (HL). This due to a high immigration rate ($IR = 0.005$), providing a rescue effect for all species. In nature the destruction or degradation of habitat often leads to a loss of species [34?]. Therefore we also intend to use the IBM to investigate the regime where habitat loss causes extinctions. The obvious way to do this is to reduce the IR. Since immigration proved to be a key driver of the dynamics in chapter 3, in this chapter we conduct a detailed analysis into the consequences of varying the IR parameter. This is largely in preparation for chapter 5 where we will study, from a more ecological perspective, how changing IR affects the response of communities to HL. We restrict the scope of this chapter to pristine landscapes ($HL = 0$). In the first part of this chapter (section 4.2) we study *closed communities* by removing immigration from the model altogether ($IR = 0.0$). In the second part (sections 4.3 to 4.6) we address an issue that is found to arise from reduced IR, namely *increased temporal variability*. In section 4.4 we revisit the assumption made in the previous analysis, that simulations reach steady-state after 5000 time steps (see section 3.2.2). In section 4.5 we test for determinism in the simulated population dynamics, and in section 4.6 we apply different sampling regimes to compare the convergence and repeatability of our experimental results.

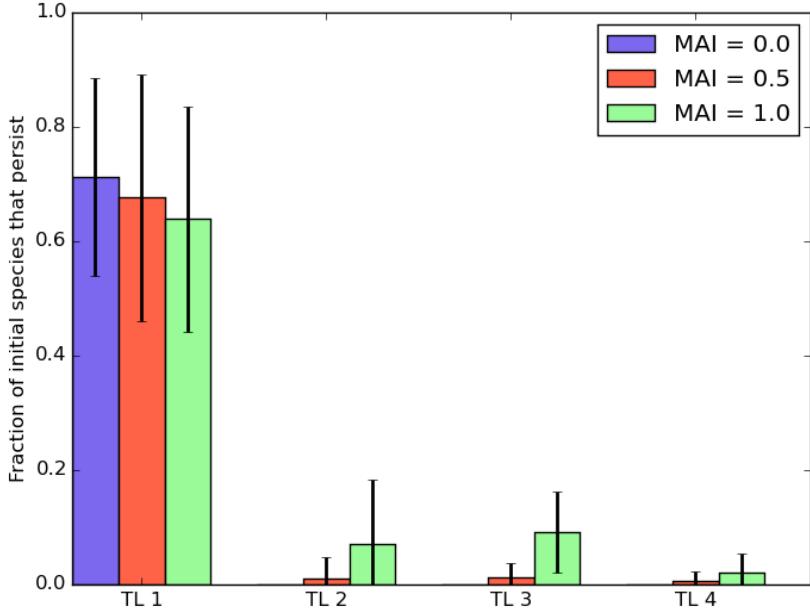


Figure 4.1: **Fractional persistence** by trophic level for three different MAI ratios, with **zero immigration** ($IR=0$). All other parameters set to default values. Fractional persistence measured by the fraction of species initially belonging to a trophic level which have not gone extinct by the end of a simulation (5000 time steps). Solid bars give the mean value over 22 repeat simulations. Error bars show ± 1 standard deviation.

4.2 Persistence in closed communities

We define a *closed-community* as one in which there is no inflow of individuals from an external source. In the model this is achieved by setting the immigration rate to zero ($IR=0$). Here we study such communities in a pristine landscape i.e. without habitat loss ($HL=0$). The simulation procedure is the same as in chapter 3. All parameters, except IR, take default values unless otherwise stated, and each simulation uses a distinct network topology generated with the method described in section 2.2. A species is said to *persist* if there is at least one individual belonging to that species present in the landscape on the final time step of the simulation. Therefore *fractional persistence* is defined as the fraction of the initial pool of species (or subset of that pool) that persist. We also use the *number of persistent species*, where absolute values are desirable. All abundance measurements given are calculated from the number of individuals belonging to each species on the final (5000th) time step of the simulation. Simulations were run with three MAI ratios ($MAI = 0.0, 0.5, 1.0$) to explore the full range between antagonism and mutualism. To reduce the number of figures we sometimes present results for only one or two MAI ratios only, but make it clear where results differ across MAI ratios.

An initial set of simulations with zero IR shows that species persistence is low (Figure 4.1). We see that for an antagonistic community ($MAI = 0$) all non-basal species go extinct, as do around 30% of basal species. Introducing mutualism into the communities slightly improves fractional persistence in the higher trophic levels. We see some persistence of non-basal species, but no more than 10% of the initial pool on average survive. Therefore, even in the *best case* at $MAI = 1.0$, we expect around 90% of non-basal species to go extinct. It is clear that, given the default parameters, immigration is a requirement for the maintenance of species richness. This is a troubling feature of the model. In nature a true closed-community may not exist, but there are systems which come close (for example remote island ecosystems). It is desirable for our model to be able to describe such systems, and for us to investigate the effects of habitat loss in this extreme case. It may also be informative to discover which factors in the IBM model contribute to persistence in closed-communities. Here we conduct a systematic review of certain model parameters and their impact on persistence. Due to the large parameter space of the model we study these parameters individually, with other parameter values held constant. We restrict the analysis to four parameters, namely: MAI ratio (section 4.2.1); reproduction rate (section 4.2.2); landscape size (section 4.2.3); and number of initial species (section 4.2.4). In addition we look at the effect of the interaction **network structure** (section 4.2.5). The choice of each parameter is justified in the corresponding subsections. As in chapter 3 we use linear models to detect trends in persistence in response to the parameter values varied. Again the *p-value* of the fit is used to test for significance in any trend identified (see section 3.2.2).

4.2.1 Mutualistic-to-antagonistic interaction ratio

In [68] it was shown that increasing the levels of mutualism (MAI ratio) in the IBM can have a stabilising effect. However in the previous chapter we saw that this stability did not translate into a greater robustness of mutualistic communities to habitat loss. From figure 4.1 we see that mutualism has a small but positive effect on fractional persistence at zero IR. An interesting feature of this effect is that **it cascades to higher trophic levels**, benefiting species other than mutualists. Here we explore the effect of mutualism on persistence in more detail.

Figure 4.2 shows the average abundance dynamics, by functional group (FG), for four different MAI ratios. These plots are produced by taking the mean abundance of each trophic level, at every time step, over 22 repeat simulations. In panel A we see that the abundance of producers rises to fill the whole landscape ($200 \times 200 = 40,000$), while 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-plants and mutualistic-animals. However this also appears to also confer some small benefit to primary-predators and top-predators, as their abundances are small but non-zero. Ecologically this makes sense. If mutualism strongly benefits mutualistic species, it will also benefit those species that feed on them. However the benefit may only be a transient effect as the abundance of producers in panel D is still rising, on average, at the end of the simulations.

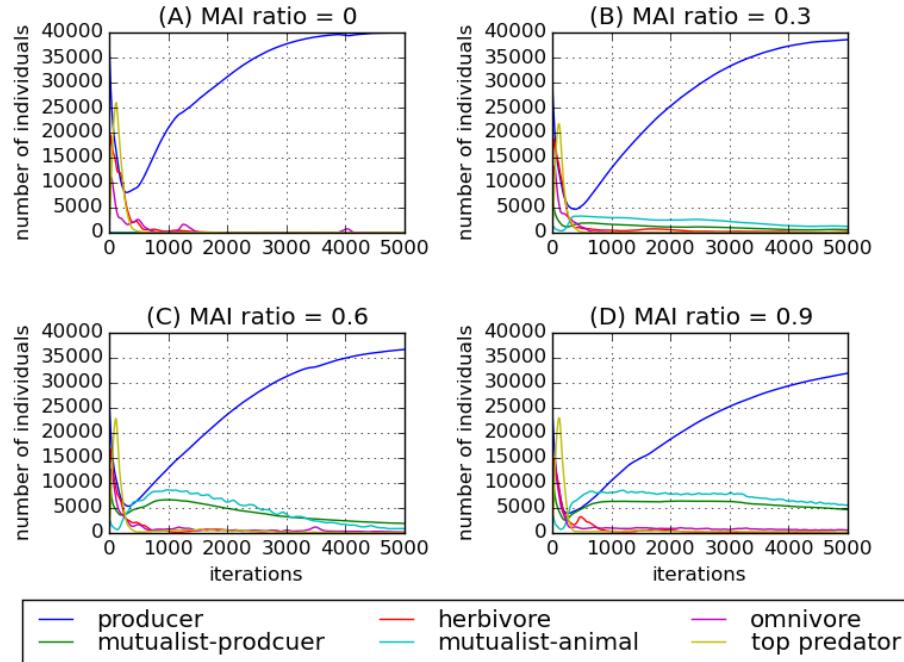


Figure 4.2: **Mean dynamics by functional group** for four different MAI ratios (A-D), without immigration. Lines given by number of individuals belonging to a functional group, averaged over 22 simulations. Colours indicate functional group (see legend).

Figure 4.3 shows the number of persistent species by trophic level for a range of MAI ratios. We see that increasing MAI ratio in fact has very little effect on the absolute number of persistent species in each trophic level. For example at MAI= 1.0 there are only two persistent species on average in the second trophic level. Therefore the increase in abundance seen in figure 4.2 must be due to a small number of dominant species. We conclude that mutualism has a negligible effect on overall persistence in terms of species richness.

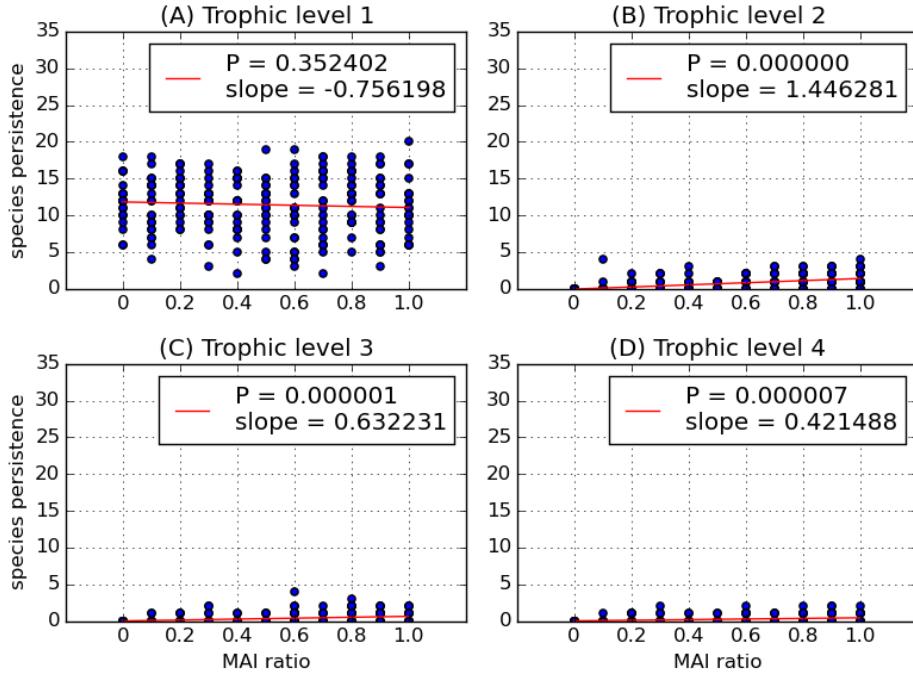


Figure 4.3: **Number of persistent species** plotted against MAI ratio, for each trophic level (A-D). Blue circles give number of persistent species, for a single simulation, on the 5000th time step. 22 repeat simulations at each MAI ratio. Red lines show linear regression fits, with slope and p-value of fit given.

4.2.2 Reproduction rate

The initial transience in the abundance dynamics at zero IR (figure 4.2) is characterised by a sharp decline in plant abundance (mutualist and non-mutualist), which reaches a minimum and then rises again. We hypothesise that this overconsumption of producers, and therefore limited availability of food for animal species, causes many of the extinctions. Indeed in these simulations ~85% of the extinctions occur during the first 500 time steps. Therefore we look at the possibility of improving persistence by increasing the reproduction rate (RR). The RR parameter defines that rate at which non-mutualist plants reproduce (via the wind-dispersal mechanism, see section 2.3). Therefore increasing this mechanism should improve the availability of plant biomass in the system, with potentially cascading effects. The RR parameter does not affect mutualistic-plants, which only reproduce via their interactions with mutualistic-animals and not via wind-dispersal. Here we vary the RR between 0.01 (default value) and 0.2 and look at the effect on persistence.

Increasing the RR improves the fractional persistence of all trophic levels, as shown in figure 4.4. This effect is greatest for the two lowest trophic levels, but has a small cascading effect on the upper two trophic levels. However, as we saw with MAI ratio, the improvement in persistence due to RR is small. Even for high RRs we still find that many species ~~to~~ go extinct. Looking at the

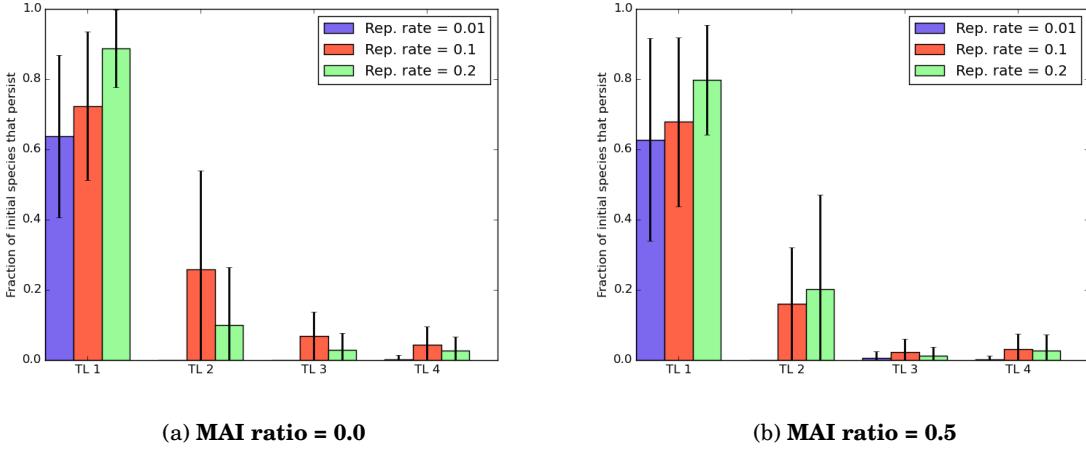


Figure 4.4: **Fractional persistence** by trophic level for three different reproduction rates (RR), at two MAI ratios (A-B). Fractional persistence measured by the fraction of species initially belonging to a trophic level which have not gone extinct by the end of a simulation (5000 time steps). Solid bars give mean value over 22 repeat simulations. Error bars show ± 1 standard deviation.

abundance dynamics in figures 4.5 and 4.6 we see that increasing the RR reduces the severity of the decline in plant abundance during transience. The resulting increase in the availability of plants does benefit all trophic levels by increasing the number of individuals present. However, because of the low fractional persistence, it must again be the case that these communities are dominated by just a couple of species in each trophic level, with all other species going extinct. The results for MAI= 1.0 (not shown) are qualitatively similar, although these communities contain fewer plants that reproduce via wind-dispersal and can therefore benefit from increased RR. In the remaining simulations for this section, as we continue to search for improvements in persistence, we choose to use RR= 0.1 . This choice reflects the resulting improvement in abundances across trophic levels compared with the default value. It was decided that this effect, combined with changes in other parameter values may help to improve persistence further.

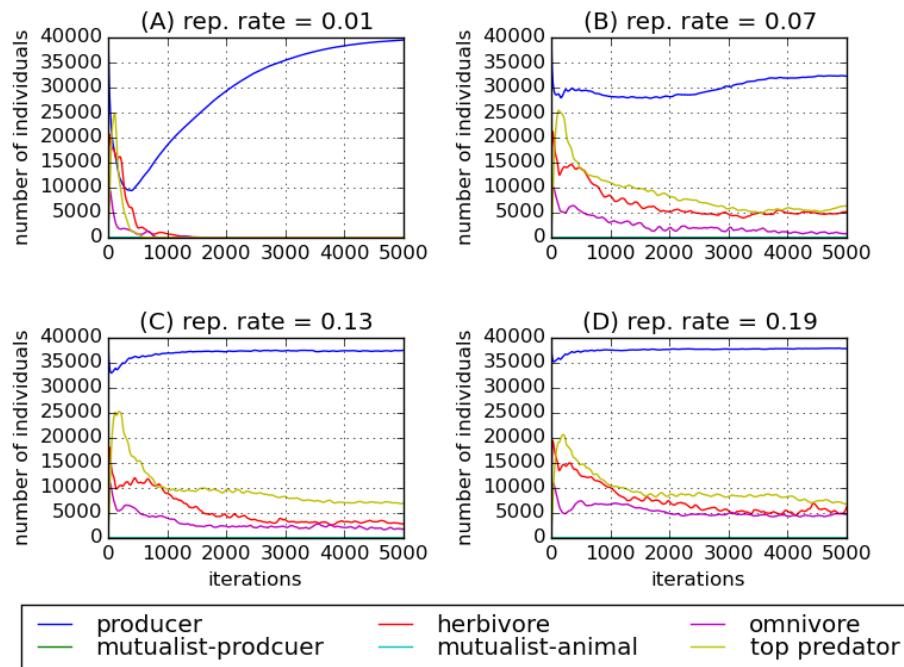


Figure 4.5: **Mean dynamics by functional group** for four different reproduction rates (A-D), at **MAI=0**. Lines given by number of individuals belonging to a functional group, averaged over 22 simulations. Colours indicate functional group (see legend).

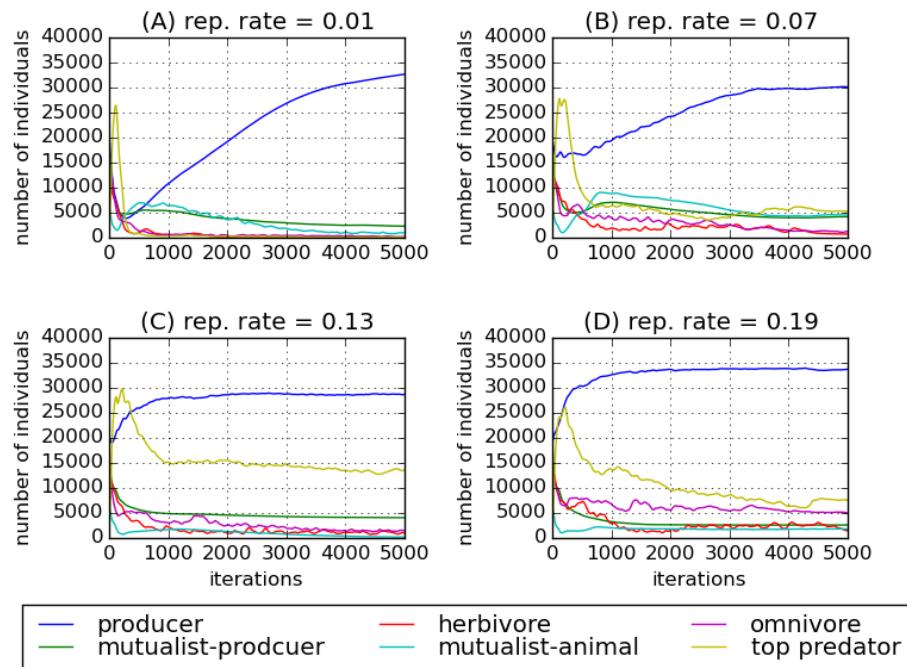


Figure 4.6: **Mean dynamics by functional group** for four different reproduction rates (A-D), at **MAI=0.5**. Lines given by number of individuals belonging to a functional group, averaged over 22 simulations. Colours indicate functional group (see legend).

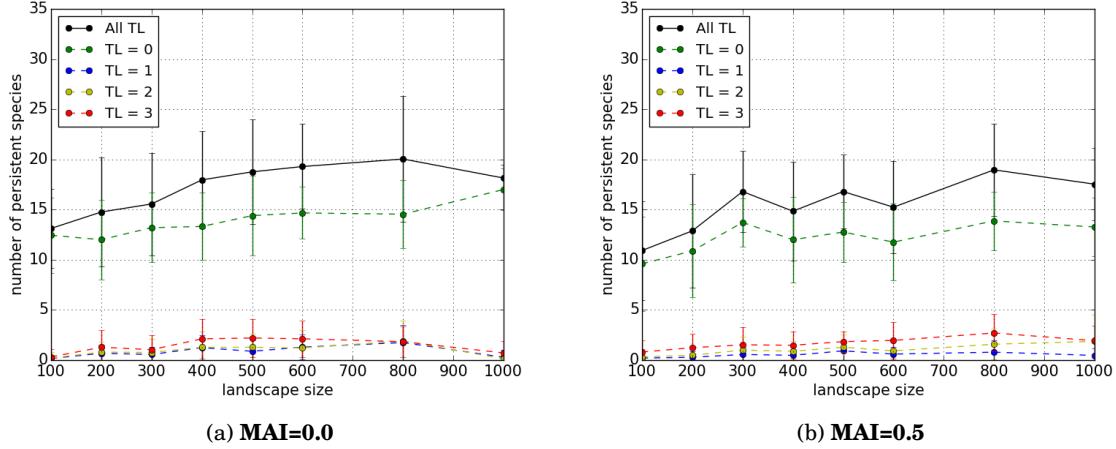


Figure 4.7: **Landscape size** against species persistence (on 5000th time step), for two MAI ratios (A,B). Total persistence, and persistence by trophic level. Points show mean value over 25 replicates. Error bars show ± 1 standard deviation.

4.2.3 Landscape size

One might hypothesise that competition for space is causing species extinctions, in which case increasing the size of the landscape may improve persistence. We varied the width of the square landscape between 100 and 1000 cells, running 25 repeat simulations at each width. These simulations were run for MAI= 0.0, 0.5, 1.0, and used the increased reproduction rate RR= 0.1. Figure 4.7 shows how the number of persistent species changes in response to landscape size for MAI= 0.0 and 0.5 (results qualitatively similar for MAI= 1.0). For both MAI ratios depicted there is an overall increase in species persistence up to a landscape width of 800. This increase in persistence is driven by small increases in the species richness of each trophic level. Beyond a width of 800 cells persistence appears to decline, although this may be a statistical anomaly and we do not explore this result further. Even with a landscape of 800 \times 800 the results show that we should not expect more than 20 persistent species, which represents the extinction of two thirds of the initial species pool. Therefore we conclude that landscape size does not resolve the problem of extinctions. Further simulations are therefore run using the default landscape width of 200, which requires significantly less CPU time than for a width of 800.

4.2.4 Number of initial species

All previous simulations have been run with an interaction network 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 this way we may hope to evolve to a stable network structure as extinctions kill off species that are not viable. Here we vary the number of species in the network between 30 and 240, and look at the effect this has on persistence. As in the previous section, 25 repeat simulations are run for each number of species at all three MAI ratios (0.0, 0.5, 1.0).

Using a network of more than around 70 species causes a problem with our network generation procedure. As described in section 2.2 we use the niche model to create interaction networks, but reject those that do not satisfy the trophic constraints. With a large number of species the probability of the niche model generating a network that meets the trophic constraints becomes low enough that the procedure cannot be run within a reasonable time (*time* > 2 days on Blue Crystal for 100 species). In particular, as discussed in section 2.2.2, the niche model overestimates the number of primary-predator species, and underestimates the number of herbivores. To overcome this problem we here construct networks in two different ways: 1) we use the standard niche model with no trophic constraints; and 2) we *rewire* networks generated using the niche model iteratively updating the network until it meets the trophic constraints. The procedure is as follows. A species is selected uniformly at random from the trophic levels that contain more than the desired number of species. This species is then moved to a trophic level that contains too few species, and is linked to other species selected at random from a pool of possible candidates. Possible candidates are defined by the new trophic level of the species that has been moved e.g. a herbivore can only eat plants, and only be eaten by species in higher levels. The number of new links created is chosen to preserve the degree of the moved species, and therefore the connectance of the network (in/out degree cannot necessarily be preserved because of the change in trophic level). The procedure is repeated, moving randomly selected species until all trophic constraints are satisfied. Mutualistic links are then introduced using the standard *link replacement* method (section 2.2.3). Here we use the same constraints as were applied in the original network generation procedure. That is the proportion of species belonging to trophic levels one, two and four must be at least 25%, 25% and 5% respectively. We refer to the networks generated using procedures 1 and 2 as *niche* and *rewired*, respectively. Simulations are run for both types of network.

Figures 4.8 and 4.9 show how the number of persistent species depends on the number of species in the initial network. For the *niche* networks we see that there is little change in species persistence. Whereas for the *rewired* networks we see a large increase in total persistence. However this increase is mainly due to plant species, with only a small change in persistence for trophic levels two, three and four. The difference between the *niche* and *rewired* network results is due to the fact that large niche model networks are dominated by omnivores. It appears

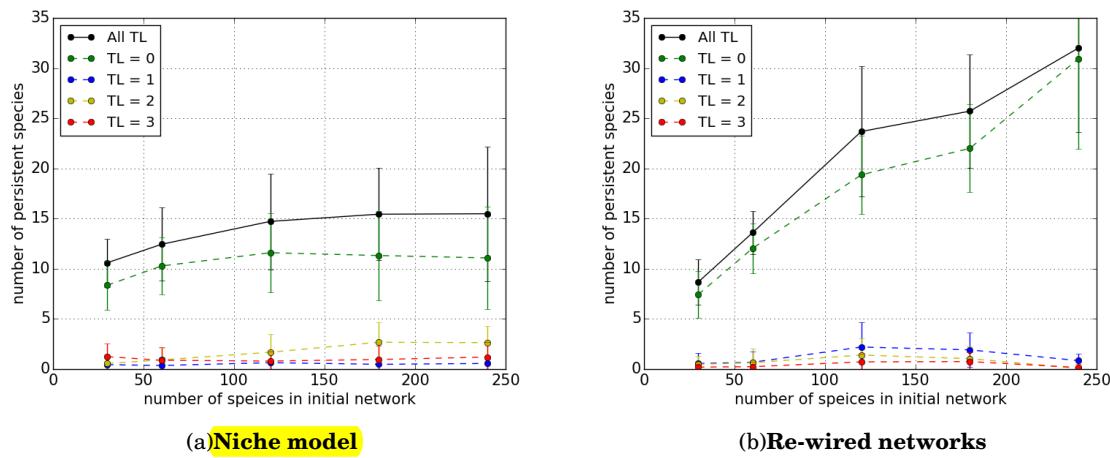


Figure 4.8: Number of initial species against species persistence at **MAI=0**, for two types of initial network: (A) Interaction network generated using standard niche model, and (B) generated using re-rewired niche model topology (see text). Total persistence, and persistence by trophic level. Points show mean value over 25 replicates. Error bars show ± 1 standard deviation.

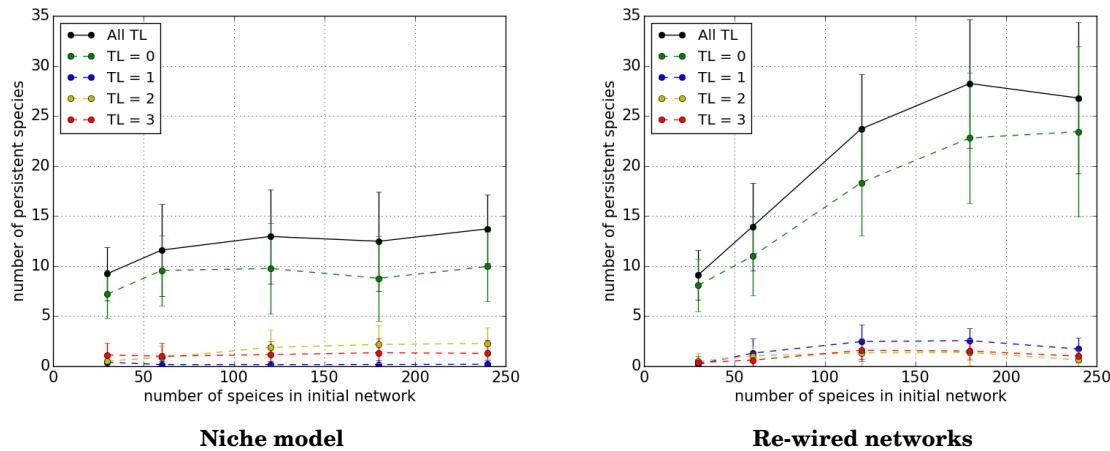


Figure 4.9: Similar to figure 4.8, but with **MAI=0.5**.

that the omnivore species out-compete each other and therefore we do not see a large change in persistence. Conversely, for the *rewired* networks the number of species in each trophic level grows proportionally with the total number of species. The only significant benefit is to plant species, which seem to be able to coexist in larger numbers. We conclude that increasing the number of initial species is not able to generate the diverse communities that we had hoped for. However the *rewiring algorithm* developed for this analysis represents a novel modification to the niche model, which may prove useful for the generation of realistic interaction networks.

4.2.5 Network structure

A notable feature of the results presented in this section is that there is significant variation in persistence between replicate simulations. Presumably some of this variation is due to noise. However some of this variation may be due to the structure of the interaction network used. As we know (section 1.3.2) the structure of an interaction is believed to impact on various aspects of community stability. Here we look for systematic differences in persistence that are due to network structure. To do this we choose one ‘good’ and one ‘bad’ network from the simulations used in the section 4.2.4. From the 25 repeat simulations using a *rewired* network of 120 species we select the two communities that display the best and worst persistence profiles. The best case, giving the ‘good’ network, had persistent species in all trophic levels and two persistent species in each of trophic levels three and four. In contrast the worst case, giving the ‘bad’ network, had no species persistence in any trophic level above one. With each of these two networks we ran 100 repeat simulations ($MAI = 0.0$) to test if the difference in persistence was repeatable and therefore due to the network structure.

Panel A in figures 4.10 and 4.11 shows the structure of the 120 species network, for the good and bad cases respectively. Visually there is little to distinguish the two networks. However we find that they have very different persistence properties. Panel C of these figures shows the average number of persistent species in each trophic level. It is clear that the good network performs better on average, in terms of persistence, than the bad network. We do not analyse the network structures here, however this could be an interesting avenue of further work. We simply conclude that the network structure is an important determinant of community dynamics in our simulation, with an effect on persistence that is repeatable.

4.2.6 Summary

1.

From the results presented in this section it appears that immigration is required in order for the IBM to generate communities with more than a handful of persistent species in the higher trophic levels. Having not exhaustively searched the parameter space we cannot be certain of this fact. However this finding is not in disagreement with our observations of the natural world. Indeed the first laboratory experiments to produce persistent predator and prey populations in a microcosm, conducted by Gause [41], required either spatial heterogeneity (to provide refuge for the prey) or immigration in order to prevent extinctions [24]. Also, in larger systems the principle of competitive exclusion is expected to cause extinctions. The principle states that the number of coexistent species cannot be greater than the number of niches (or alternatively, the number of limiting resources). The reasoning is that two species competing for the same resource is unstable, such that if either species happens to gain an advantage it drives the other

¹Note that this does not rule out competition as a cause of extinctions? Just that increasing the available space does not remove it.

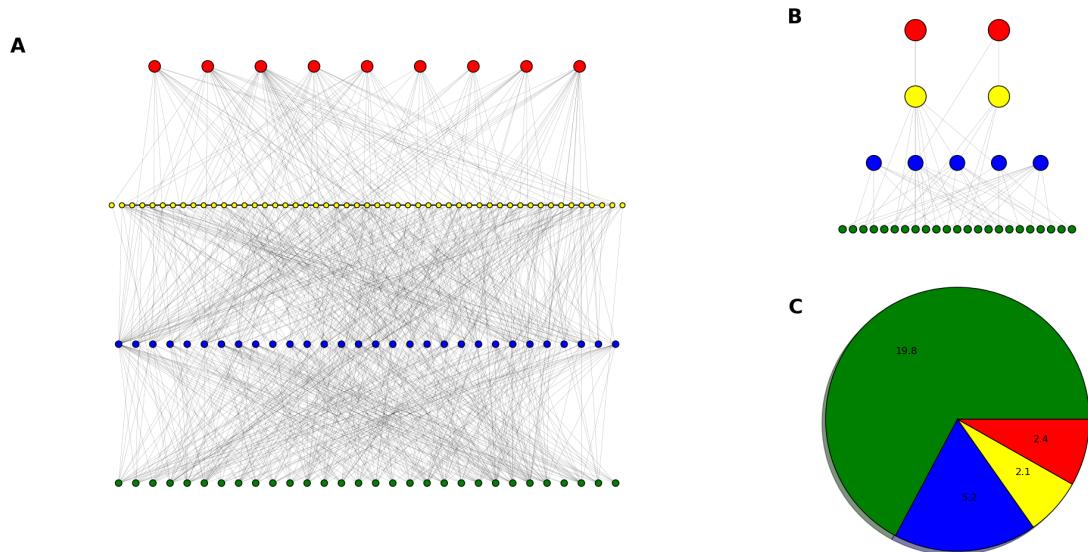


Figure 4.10: Example of a ‘**good network structure**’. Colours represent trophic level. (A) Antagonistic network of 120 species, rewired from the niche model (see text), used as input for simulations. (B) Example of ‘pruned’ network consisting of the species that persist after 5000 time steps. (C) Mean number of species belonging to each trophic level in the pruned network, averaged over 100 replicate simulations (all using A as the initial network).

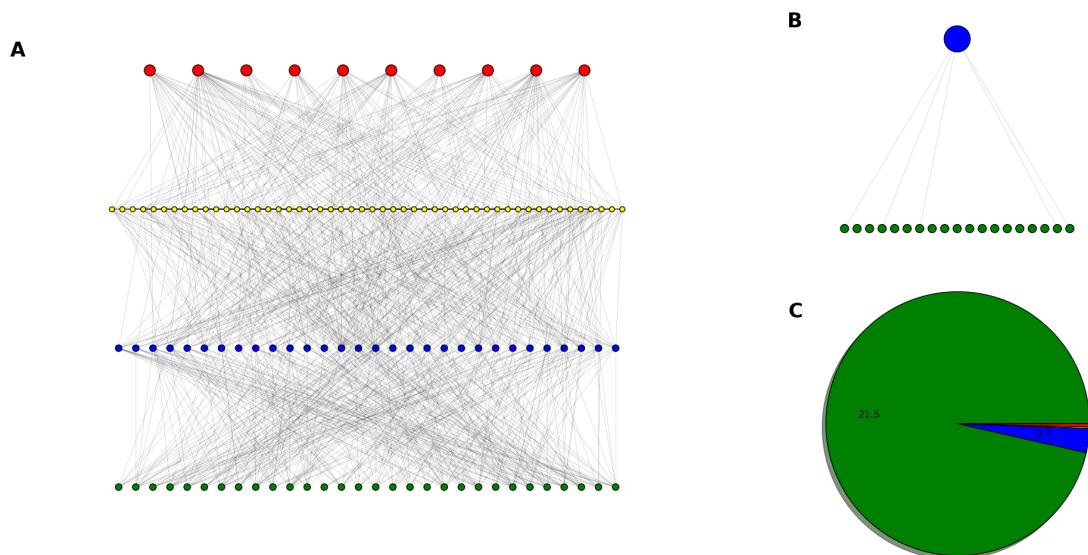


Figure 4.11: Example of a ‘**bad network structure**’. Similar to figure 4.10, but with a different initial network of 120 species used as input for the simulations.

to extinction. The principle of competitive exclusion presents an apparent paradox in Nature because many communities exhibit species richness much greater than the apparent number of limiting resources. A prime example of such communities are oceanic plankton, hence the *paradox of the plankton* [51]. One proposed solution to these paradoxes comes from meta-community theory (see section 1.3). That is, although local populations may be unstable in isolation, they exist as part of a meta-community. Various models predict that meta-community dynamics, which involves dispersal between local patches or communities, can stabilise local populations [49]. The key feature here is *dispersal*, which is not dissimilar to our immigration mechanism because it provides an external source of individuals. Therefore we propose that the extinctions we are seeing in the IBM at zero IR are due to competitive exclusion, an effect which is reduced by the introduction of immigration².

We demonstrated in section 4.2.5 that network structure *can* have an effect on persistence, and therefore on community dynamics also. This results implies that the IBM could be used to study the effect of various network properties on different aspects of stability (see sectionref:WHERE). We also presented a novel algorithm to modify niche model networks according to desired trophic constraints. In order to justify the use of this method further investigation would be required to compare the realism of the derived topologies to empirical webs.

²Also discuss differential traits?

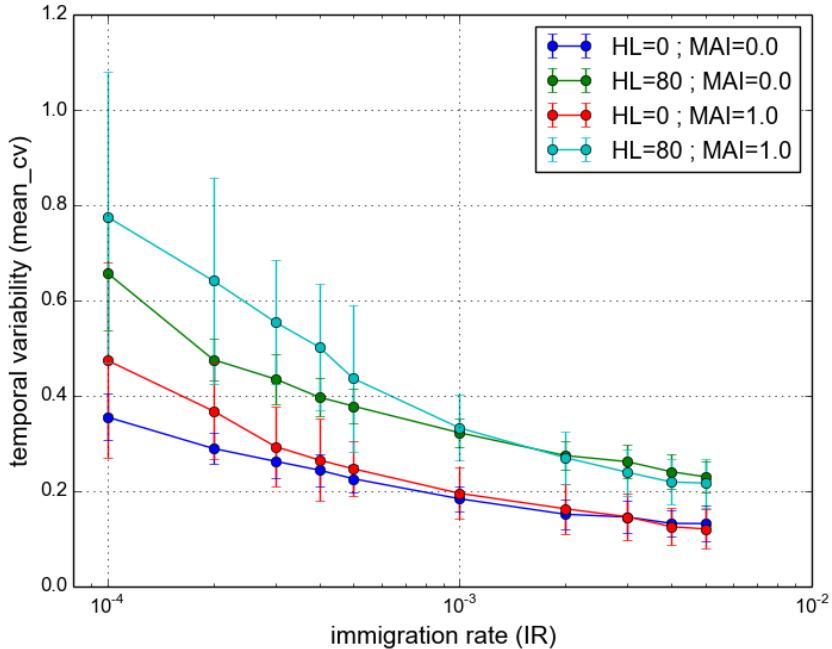


Figure 4.12: **Temporal variability** against immigration rate (IR). Variability calculated as *mean_CV* (defined in section 2.7.2). Points show mean value over 50 repeat simulations. Error bars show ± 1 standard deviation. $HL = 0$, all other parameters take *default values*.

4.3 Temporal variability and immigration

In section 4.2 we concluded that immigration is required for our model to produce persistent communities. Therefore to study communities that are persistent in pristine landscape, but exhibit HL *induce* extinctions, we must investigate the region between zero immigration and the *default rate* ($IR \sim 0.005$). As we began to investigate this region we found a robust feature of the model: reducing immigration rate *increases the temporal variability* of the dynamics. We again measure temporal variability using *mean_CV* (section 2.7.2) i.e. the coefficient of variation in species population dynamics, averaged over all species in the community. In the previous chapter *contiguous* habitat loss was shown to increase temporal variability, up to a average value of ~ 0.3 at $HL = 90\%$ (figure 3.16). Figure 4.12 illustrates that reducing the immigration rate can push the temporal variability well above this level, even in pristine landscape.

This increased variability may bring into question the validity of results derived from the simulation output. Implicit in our analysis so far is the assumption that the system reaches *steady-state* by the end of the simulation (see section 3.2.2). If the system is not at steady-state then our results, especially those calculated from snapshots or short samples, may not be reliable. In the high immigration regime (HIR) the steady-state assumption may be reasonable. From inspection, the HIR simulations contain transient dynamics within the first 1000 time steps,

followed by a period of relatively constant abundance (see for example figure 2.5). However the high level of temporal variability illustrated in figure 4.12 now motivates us to question whether the system does reach steady-state, especially at low IR. Therefore, in section 4.4, we conduct a detailed analysis of the *stationarity* of population dynamics generated by the IBM model. We then consider whether the increased variability is due to stochastic or deterministic effects. In section 4.5, we use *recurrence quantification analysis* to look for signatures of determinism in the simulation output. Finally we close the chapter (section 4.6) by looking at how the accuracy and reliability of our numerical results are affected by increased variability, and how sensitive the results are to the sampling regime used.

4.4 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 [50] states that a random process $X(t), t \in \mathbf{Z}^+$ is second-order stationary if:

$$(4.1) \quad \mathbf{E}[X(t)] = \mu \text{ (constant)},$$

$$(4.2) \quad R_X(t,s) = \mathbf{E}[X(t)X(s)] = R_X(|s - t|),$$

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 an autocorrelation dependent only on time separation. From now on we refer to a time series satisfying 4.1 and 4.2 as *stationary*. If the conditions are not met we call the time series *non-stationary*. Non-stationary time series cannot be fit to a constant distribution. 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 the dynamics is characterised by large amplitude oscillations in abundance, which produce non-stationarity. Again such oscillations have not been tested for. Alternatively a steady-state equilibrium may exist, but a long period of transience means it is not reached during the time frame of our simulations (5000 time steps previously).

4.4.1 Tests for stationarity

We compare three different tests of stationarity: the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) [65]; the Augmented Dickey-Fuller (ADF) [97]; and the the Priestley-Subba Rao (PSR) [92]

tests. These tests were chosen for their popularity in the time series literature. All three are implemented in the programming language *R* [94]. The KPSS and ADF are provided in the *tseries* package, and the PSR is in the *fractal* package. All three tests are applicable to univariate time series.

The ADF test has a null hypothesis that the time series is non-stationary. The test models the data as an auto-regressive process (see section ??), 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 [92]. 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.

4.4.2 Characterising the tests

To understand the performance of the stationarity tests (section 4.4.1) 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 IR= 0.001, zero mutualism (MAI= 0.0) and otherwise default parameters (table 2.1). The series represents the total number of individuals of all species at each time step. The simulation was run for 50,000 time steps, 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 time steps were discarded, since these contain clearly transient dynamics (see figure 4.15B), leaving a time series of 49,000 points. The IR used is lower than the *default value*, but an order or magnitude higher than the lowest IR used in this chapter (see figure 4.12). The rate $IR = 0.001$ was chosen because it gives slightly increased temporal variability over the default value and therefore increased chance of non-stationarity. In the context of what follows we refer to IR as high immigration (HI).

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:

$$(4.3) \quad x(t) = \sum_{i=1}^{t-1} Z_i,$$

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 those of the HI series. RW has 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 4.13.

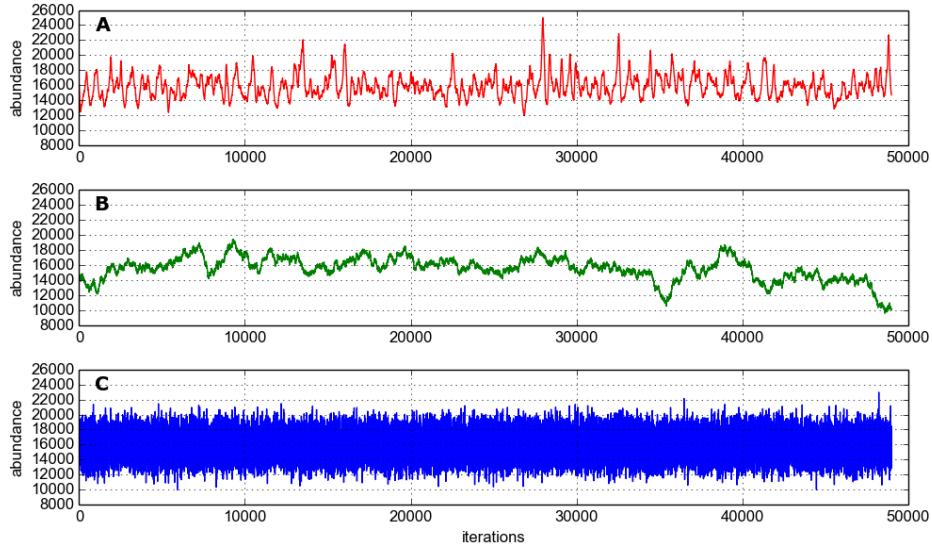


Figure 4.13: The three time series used to characterise the performance of the stationarity tests (see text for details of how they are generated). The initial 1000 points are removed such that all are 49,000 points long. (A) HI: total abundance dynamics of a simulation with high immigration rate; (B) RW: a random walk without drift; and (C) NS: a series generated by independent sampling from a normal distribution.

	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 4.1: Results of applying the three stationarity tests to the three time series shown in figure 4.13. 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.

Initially we apply the three stationarity tests to the entire length of the time series. The results are shown in table 4.1. 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 RW is a particular instance of a random walk, chosen from several thousand to closely match the mean and variance of HI. Therefore this random walk is likely to appear more stationary than others in the ensemble generated. 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, based on the magnitude of the test statistic. According to this test NS is clearly stationary (accept h0), and RW is clearly non-stationary (reject h0 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 data generators that produced the time series. Therefore we do not use this test in the analysis that follows. Nevertheless the apparently erroneous result may contain interesting information about the HI series and the process that generated it (see the discussion in section ??).

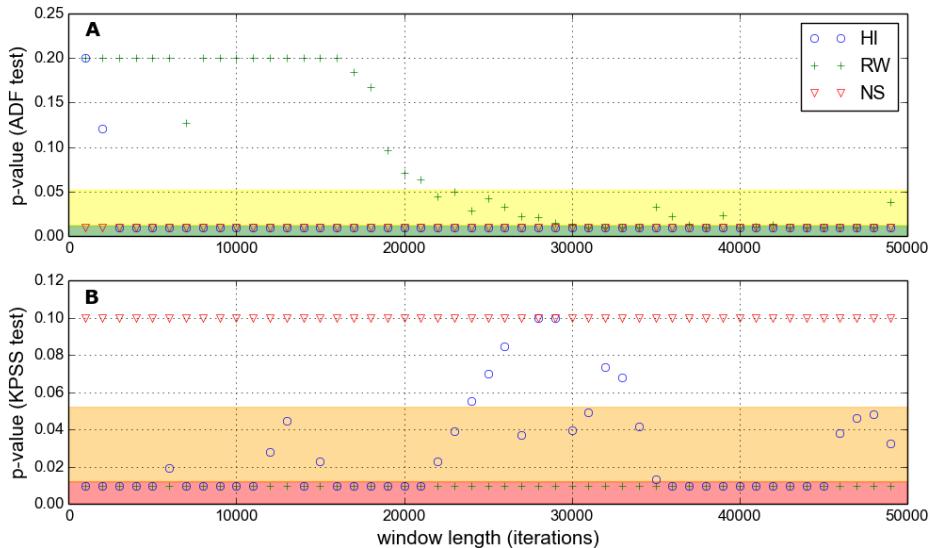


Figure 4.14: Two tests for stationarity applied to samples of varying size (window length). Samples are taken from the three time series (HI,RW,NS) shown in figure 4.13. All three time series contain 49,000 points. Sample windows begin at the first point and increase in length from 1000 to 49,000 points. Minimum p-value plotted is 0.01, actual values may be lower. (A) ADF test, with p-values capped at 0.20. The 95th and 99th percentile shaded yellow and green respectively, indicate significant evidence for stationarity. (B) KPSS test, with p-values capped at 0.1. The 95th and 99th percentile shaded orange and red respectively, indicate 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 sample length. Sample lengths range from 1000 to

49,000 data points. Again, as we saw in table 4.1, 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 4.14B. The ADF test (figure 4.14A) 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 4.14. 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, the test statistic does order the series correctly in the above examples (table 4.1) 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 sensitivity to sampling is investigated by *reversing* the time series and repeating the above analysis (**results not shown**). 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 run.³

HI simulation. ³ 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 due to the phase differences between species oscillations that we might expect due to trophic interactions. Such oscillations may mean that temporal variability is cancelled out when aggregating abundances in this way (see discussion on ecosystem synchrony, section 3.3.4). It

³Trophic dynamics figure: refer to comparison with figures from methods chapter..

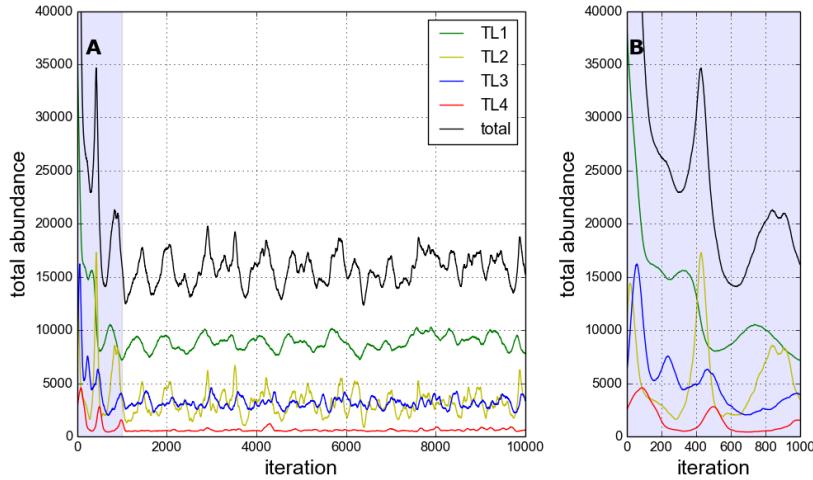


Figure 4.15: Population dynamics from HI ($Ir = 0.001$) simulation: total number of individuals, and by trophic level. (A) First 10,000 time steps of simulation run (total = 50,000 time steps). (B) Enlargement of first 1000 tim steps, showing transience.

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 TL1 – 4. These *trophic dynamics* are plotted in figure 4.15, and display visibly more variability than the equivalent plot produced using the default IR (figure 2.5). The initial period of transience is expanded in panel B. As in the previous analysis this part of the time series (first 1000 time steps) is removed. The ADF and KPSS tests are applied to the four trophic series separately and the results are shown in figure 4.16. According to ADF all trophic levels are stationary for samples sizes greater than 4000. TL1 appears to be least stationary according to ADF, requiring a sample size of at least 4000 to reject the null hypothesis at 95% confidence. According to KPSS TL1 is non-stationary for all sample sizes, whilst TL2 and 3 are stationary for samples sizes above 6000 and 1000 respectively. KPSS gives mixed results for TL4, with no clear dependence on sample size. It is hard to reconcile these results with an observation of the dynamics in figure 4.15, indicating the usefulness of the statistical tests.

The dynamics of all the species belonging to each trophic level are plotted in figure 4.17. 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 chapter 3. 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.

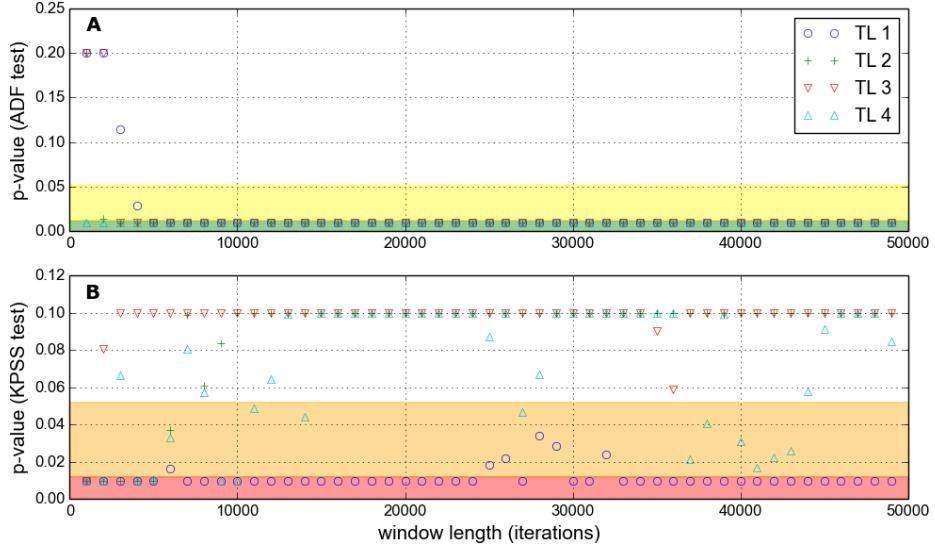


Figure 4.16: Similar to figure 4.14, but here the tests are applied separately to each trophic level of the HI simulation shown in figure 4.15.

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 initial transience).

We see from figure 4.18 that all six species are stationary according to ADF, given a 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 4.17 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 lack of stationarity is largely due to the apparent strictness of the KPSS test. We also assert that 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 due to aggregation. In the following analysis we apply stationarity tests to species dynamics individually, and calculate the number of stationary species (NSSP) as an aggregate statistic. If NSSP equals the total number of species, then the community dynamics is fully stationary.

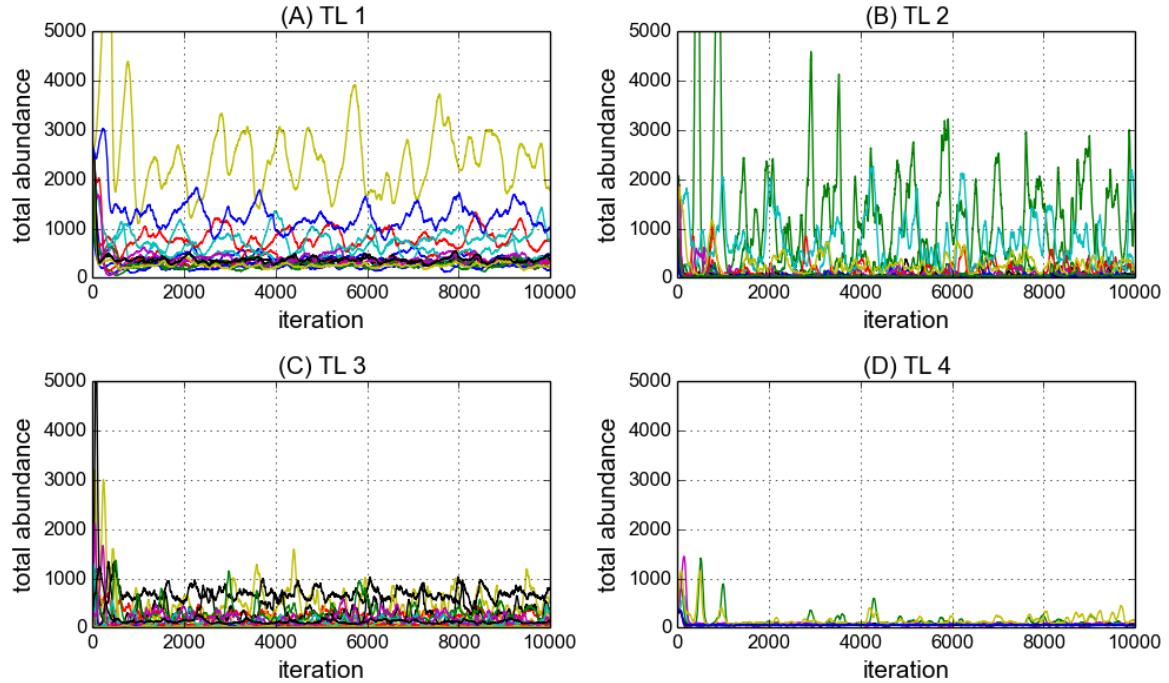


Figure 4.17: Dynamics of all species in the first 10,000 time steps of the HI simulation shown in figure 4.15. Each colour represents different species, each panel (A-D) shows a different trophic level (TL1 – 4).

(according to the test used).

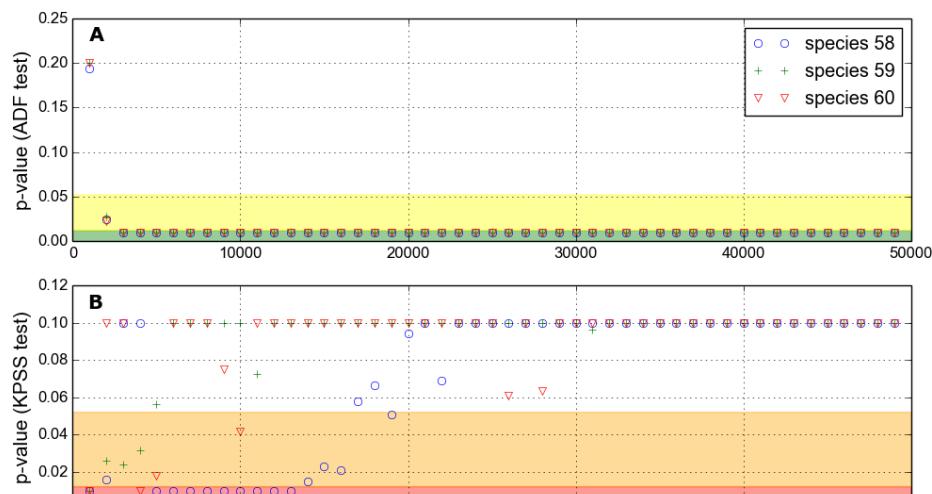
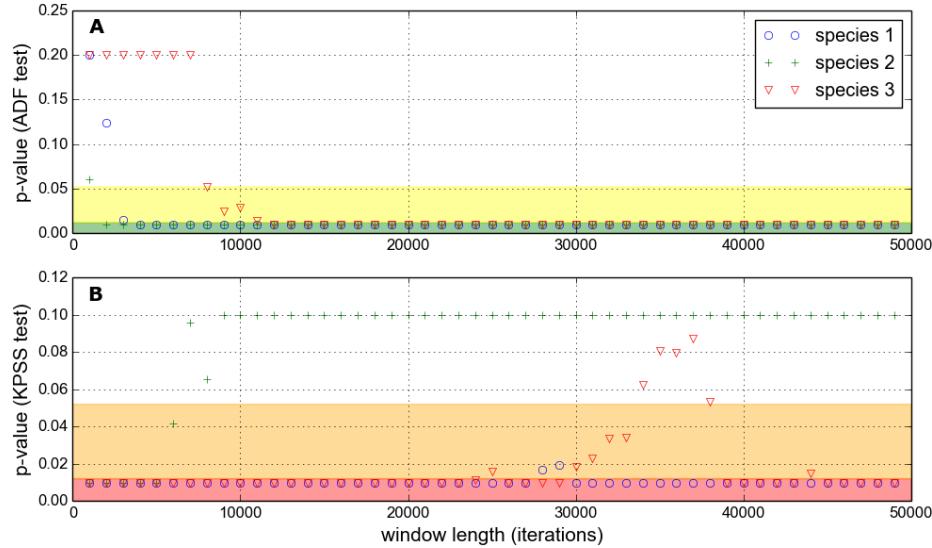


Figure 4.18: Similar to figure 4.14, but here the tests are applied separately to individual species from the HI simulation shown in figure 4.15. (A) The three species with highest long-term average abundance. (B) The three species with lowest long-term average abundance.

4.4.3 General stationarity results

We now conduct 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 4.19). The new simulations use default parameters, unless otherwise specified, and are run for 50,000 time steps. 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. All networks are generated using the niche model as described in section 2.2, with zero mutualism ($MAI=0.0$), and each simulation uses a uniquely generated network. A third ensemble of 50 shorter simulations (10,000 time steps) is run, at high immigration rate, using a fixed interaction network selected at random from the HI ensemble. This third ensemble we call NM1. Stationarity testing is done using the ADF and KPSS tests characterised in section 4.4.2 above. As standard the initial 1000 time steps 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 4.19A. 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 time steps. 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 towards the end of a simulation) samples of length 3000 were taken from different points along the time series. From figure 4.19B we can see that there is no clear trend in stationarity over 49,000 time steps. The average NSSP is almost the same whether the sample is taken from time steps 1000-3000 or 46,000-49,000. This lack of time dependence in NSSP also holds for windows of different length (results not shown).

On average we see that the LI ensemble is less stationary than the HI ensemble. This we expected based on the previous observation that reducing IR increases temporal variability. However we cannot be confident that either ensemble contains communities in which all species are stationary. This observation 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. It may be that stationarity of the simulation output is not dependent on the interaction network structure. Alternatively the observation may be anomalous, owing to the fact that we have accidentally chosen NM1 to closely resemble the average of this ensemble (see both panels of figure 4.19).

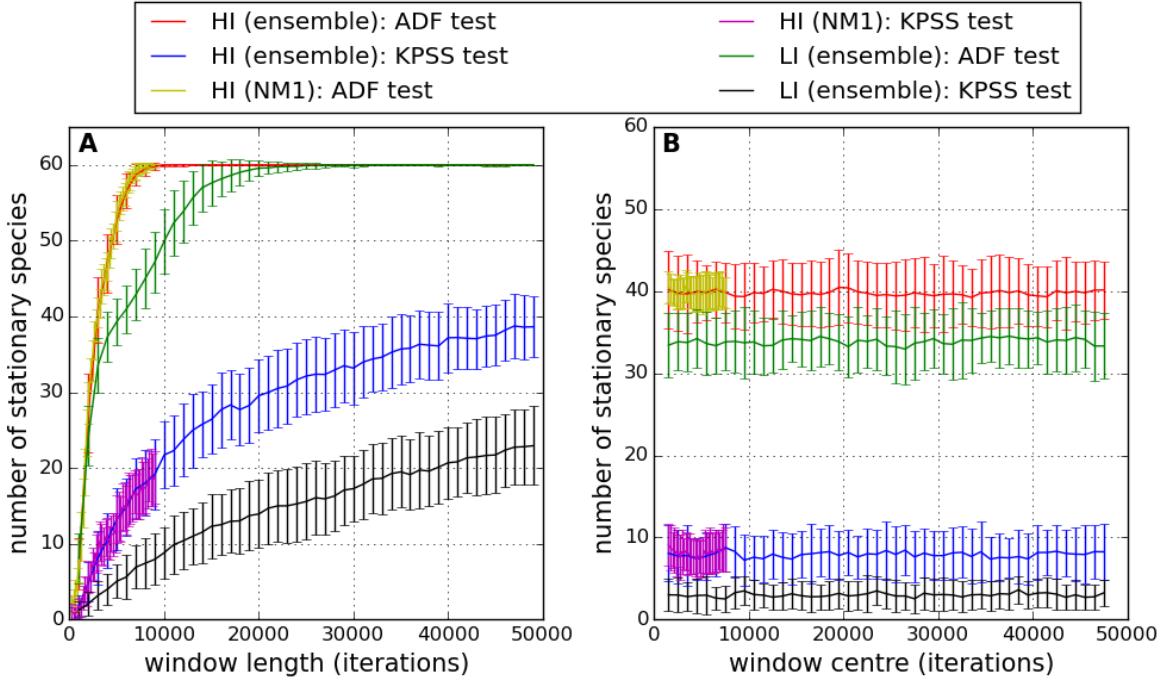


Figure 4.19: The number of stationary species (NSSP) according to the two stationarity tests (ADF and KPSS) at the 95% confidence level. Results averaged over three different ensembles of simulations: HI; NM1 (high immigration) and LI as described in the text. The first two are high immigration ensembles, whilst the latter is low immigration. Solid lines indicate mean results for the ensemble. Error bars indicate ± 1 standard deviation. (A) Each species abundance time series is sampled with a window of increasing length, as in figure 4.18. (B) Each species time series is sampled with a window of length 3000 time steps, which is scanned along the series to check for changes in stationarity over time.

In chapter ?? we will explore a slice of the IBM parameter space by varying both immigration rate and habitat loss. Here we study the stationarity of the simulations that we will use in that chapter. All simulations are 5000 time steps long. The initial 1000 time steps are discarded and the ADF and KPSS tests applied, species by species, to the remaining 4000. Figure 4.20 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, but still visible, effect 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 4.21 summarises these trends using cross sections taken from the ADF heat maps (panels A,C,E) in figure 4.20, 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).

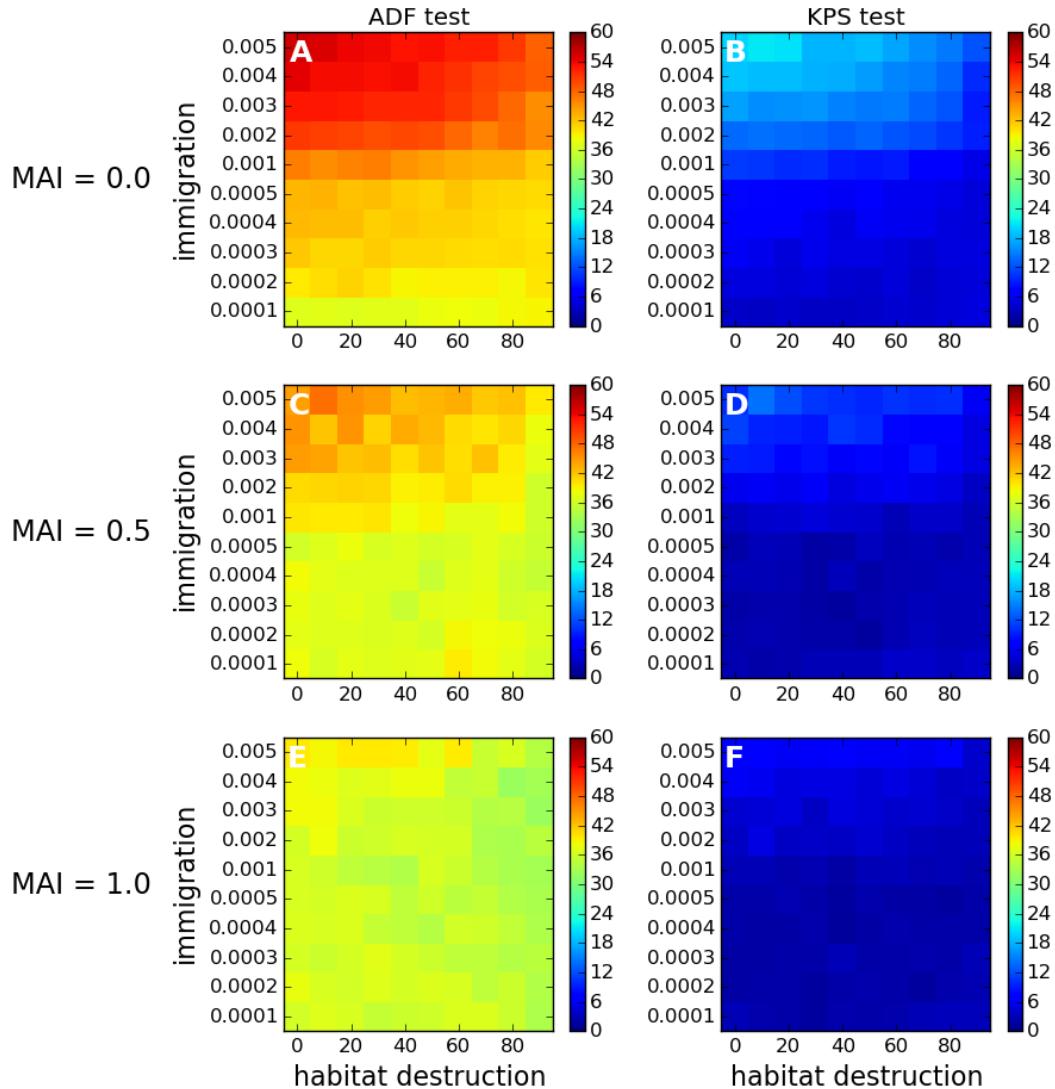


Figure 4.20: The average number of stationary species (NSSP) according to the two stationarity tests (ADF and KPSS) at the 95% confidence level. Each cell on the grid represents the mean value over 50 repeat simulations. All simulations are 5000 time steps with default parameter values. Tests are applied to the final 4000 time steps of each species abundance time series. These correspond to the simulations presented in chapter ??.

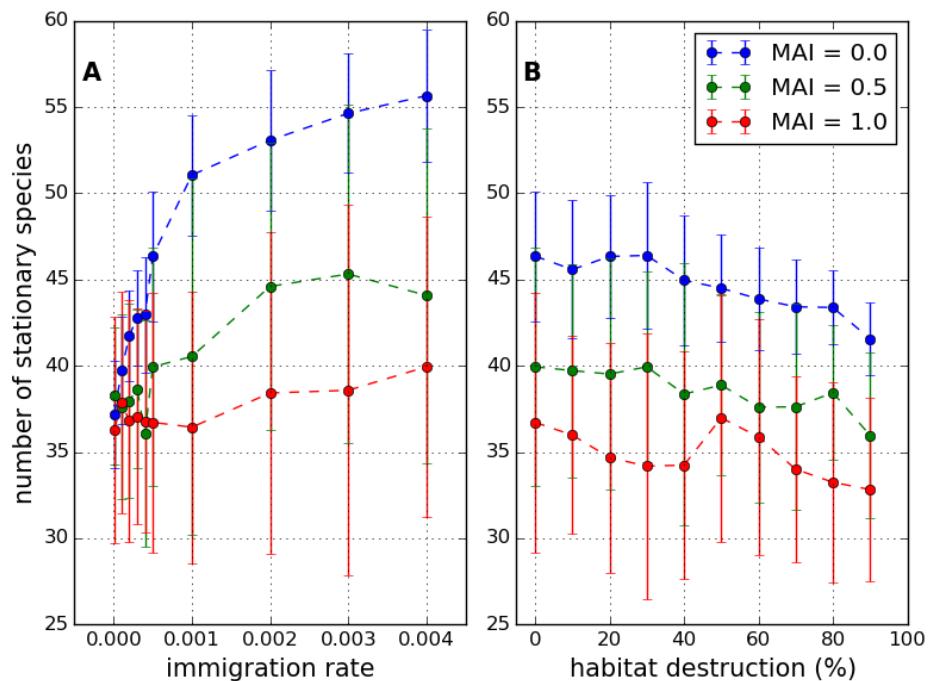


Figure 4.21: The number of stationary species (NSSP) according to the ADF test at the 95% confidence level. Points show mean value over 50 replicates. Error bars show ± 1 standard deviation. Tests are performed on the same simulations depicted in figure 4.20. (A) Plotted against immigration rate (IR), with zero habitat destruction. (B) Plotted against habitat destruction, with IR= 0.001 (high immigration regime).

4.5 Determinism tests

We have seen that population dynamics becomes highly variable, especially at low IR, and indeed may not be stationary. This leads us to speculate that in some instances the simulation output may become dominated by stochastic effects. This randomness is an interesting feature of the model, and indeed noise is ubiquitous in natural systems. However much of our analysis (chapters 3 and 5) involves interpreting the simulation results in terms of the ecological mechanisms built into the model, in particular the patterns and structures generated by species interactions. If the model output lacks determinism it may not be meaningful to conduct such analyses. Therefore we present here a test for determinism based on *recurrence quantification analysis* (section 4.5.1), which we then apply to the IBM output (section 4.5.1.1).

4.5.1 Recurrence quantification analysis

Recurrence quantification analysis (RQA) *maybe* be used to detect signatures of determinism [6, 47, 70]. The analysis is based on the idea of *recurrence* - deterministic dynamical systems tend to return to similar regions of phase space. Moreover, when they do, their trajectories tend to remain close in phase space for some amount of time. In the case of chaotic systems the trajectories will diverge, at a rate which is broadly determined by the maximal Lyapunov exponent [47]. However, even for chaotic systems, there is some tendency for neighbouring trajectories to remain close. RQA aims to detect the presence of this feature in the dynamics. The analysis is often used with univariate time series (such as in [47]), in which case time-delay embedding must be used to increase the dimensionality of the phase space. However, since we have a time series for each of the 60 species, our phase space is already high dimensional. Therefore our first step is to construct a recurrence matrix (RM). The RM is a binary matrix whose elements, for a system $x(t) \in \mathbb{R}^N$, are given by the function:

$$(4.4) \quad d_{ij} = \begin{cases} 1 & \text{if } \|x(i) - x(j)\| < r \\ 0 & \text{if } \|x(i) - x(j)\| \geq r \end{cases},$$

where r is a threshold distance that defines the measure of ‘closeness’ in phase space. Various methods have been proposed to choose the value of r . These methods are discussed in [6], but we use the method which they adopt for the results in that paper: r is chosen such that $\sim 10\%$ of the elements of the RM are equal to one.

Having constructed the RM it can be visualised as a *recurrence plot* (Figure 4.22). Such plots can be visually striking (search recurrence plots for examples), but are difficult to interpret without experience. An important feature, in the search for determinism, are *diagonal lines* - lines parallel to the leading diagonal. Such lines indicate that trajectories which find themselves ‘close’ in phase space remain close, for a period of time given by the length of the line. Visual

inspection can detect the presence of diagonal lines but it is better to use a quantitative pattern detection method. A common metric that quantifies the relative abundance of diagonal lines is the *percentage of determinism* (%DET) [6, 70], and is given by:

$$(4.5) \quad \%DET = \frac{NPD}{NREC} \times 100,$$

where NREC is the number of entries in the RM equal to one; and NPD is the total number of points found on diagonal lines of length greater than or equal to two. The %DET allows quantitative comparison of the level of determinism between different RMs. In [6] they develop three statistical tests, based on %DET and two similar metrics, for the null hypothesis of pure randomness in the time series. However in the current analysis we make do with a comparison of the value of %DET between different test cases.

4.5.1.1 Results

We test for determinism in the simulation ensembles HI and LI used in section 4.4 (see for example Figure 4.19). Each ensemble consists of 100 repeat simulations with different network structures. These simulations are used because one ensemble is from the high immigration regime (HI: IR= 0.001) and one is from the low immigration regime (LI: IR= 0.0001). Therefore the temporal variability is higher on average in the LI ensemble than the HI ensemble. The simulations in these ensembles are also longer than other simulations - being run for 50,000 time steps, they allow for more chance of recurrent dynamics and therefore provide more information from which to detect a signature of determinism. In both ensembles the communities are antagonistic (MAI= 0) and the landscape is pristine (HL= 0). Therefore we do not cover the full range of parameters explored in other analyses, but we at least consider one highly variable and one less variable scenario.

We construct randomised control data in two different ways. The control data is used to compare %DET results from the IBM with %DET values generated by randomness. In [6] they state a random process would generate a RM with NREC points distributed uniformly at random. We construct such matrices by randomly permuting the elements of another RM, to obtain a randomised RM with the same number (NREC) of points. The second method is to construct RMs by randomising the order of the time series of each species. This preserves the mean and variance of the population dynamics, but removes any determinism. The randomised time series are then used to construct an RM.

Figure 4.22 shows four example recurrence plots. Panels A and B show the RMs of single simulations from ensembles HI and LI respectively. Panels C and D show randomised versions of panel B generated by randomising the matrix, and by randomising the time series respectively. Panels A and B clearly have some structure, whilst the structure is lost in panel C as expected. Panel D retains some structure. In particular it contains the leading diagonal, since even the

randomised time series is identically equal to itself. It also contains horizontal and vertical bands, created by points in the time series that are unusually distant (or close) to a large number of other points in the phase space. A subtle but detectable feature of panels A and B is the presence of diagonal lines (parallel to the leading diagonal) suggesting determinism. This feature is lost in panel D, as we would expect due to the randomisation of the time series. Interestingly panel A looks very similar to figure 1(c) in [6]: an RM generated from the dynamics of a noisy chaotic Hénon attractor. This leads us to speculate that the deterministic component of the dynamics may be chaotic. We do not search for signatures of chaos, partly because the presence of noise severely complicates the available methods for doing so (such as estimation of the Lyapunov exponent)[47].

We calculate the %DET for each simulation in the HI and LI ensemble. In each case, for comparison, we also calculate the %DET for the dynamics of the ten most and ten least abundant species, and for the two randomised versions of the data. These results are shown in figure 4.23. The randomised data shows little variation, with a $\%DET \approx 10$ in all cases. The %DET for the least abundant species is slightly greater than for random data, suggesting some evidence of determinism but also a strong stochastic component. In both the HI and LI ensembles there is good evidence for determinism at the whole community level, with %DET consistently > 30 and > 40 respectively. It is interesting that there is more evidence of determinism in the LI ensemble than in the HI ensemble, as we were previously concerned that the high temporal variability of the LI dynamics could mean that it lacked determinism. It is also clear for these results that the deterministic component of the dynamics is dominated by the most abundant species. This, together with the results from previous sections tells us a lot about the dynamics of the model (see discussion in section 4.7). We conclude here that the analysis so far suggests that non-stationarity may be largely due to deterministic population dynamics/oscillations of the most abundant species, while the dynamics of the less abundant species are more random and more stationary although they do have a deterministic component.

4.6 Convergence and repeatability

In chapter 3, and at the beginning of this chapter, we used ‘snapshots’ of the simulation state to calculate species abundances. The snapshot method was justified by the assumption that simulations reached stationarity and therefore we were sampling from a steady-state distribution. However, as we saw in section 4.3, stationarity cannot be guaranteed. The results of section 4.5 suggest that this lack of stationarity is due to deterministic population dynamics, especially those of the most abundant species. With oscillatory dynamics it is clear that snapshot sampling will yield different results depending on when the measurement is taken. We make the assumption that the best way to characterise the system is to take the long-term average of the metric in question. This approach is also justified by the observation that stationarity does not increase

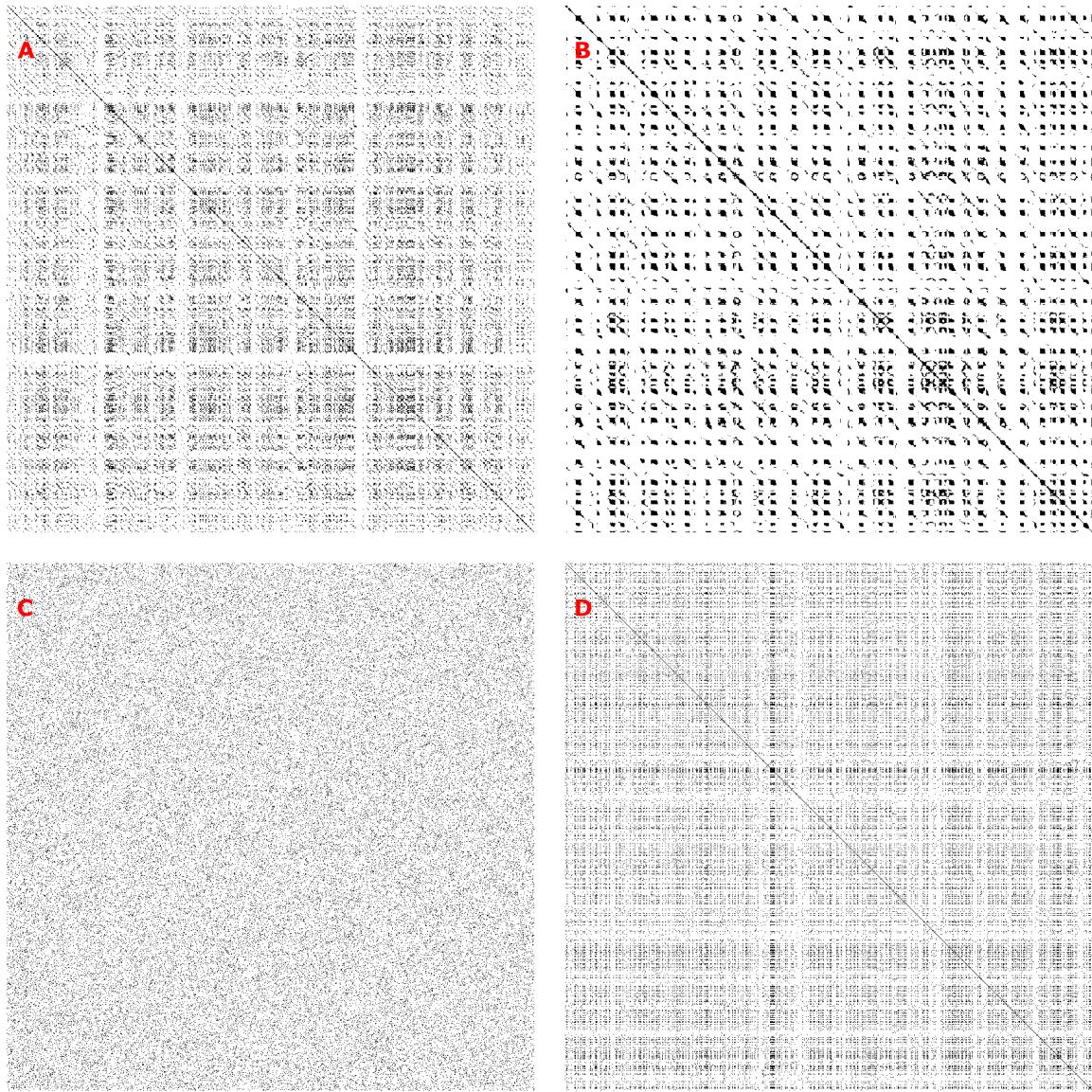


Figure 4.22: Recurrence plots defined by equation 4.4. (A) High immigration simulation ($IR=0.001$). (B) Low immigration simulation ($IR=0.0001$). (C) Randomised recurrence matrix (permutation of the elements of the matrix in B). (D) Randomised time series (permutation of species population time series from B, such that mean and variance for each species are preserved). In simulations for both A and B: $HL=0$ and $MAI=0.0$; number of time steps = 50,000; sampled every 50 time steps to construct plots.

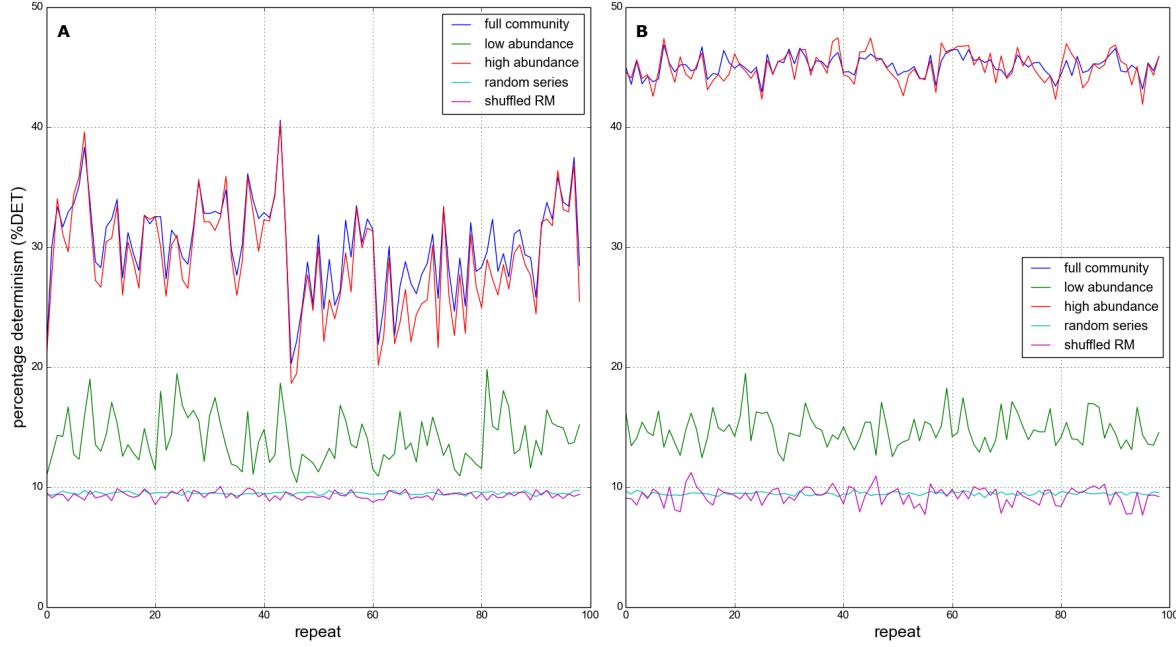


Figure 4.23: Percentage determinism (%DET) defined by equation 4.5. Value calculated for whole community, ten least and ten most abundant species, and for two randomised versions of the data (see text). For 100 repeat simulations at: (A) High immigration rate ($IR=0.001$); and (B) Low immigration rate ($IR=0.001$). In simulations for both A and B: $HL=0$ and $MAI=0.0$

of decrease over the course of a simulation, but remains relatively constant⁴ (figure 4.19B). In section 4.6.1 we look at how temporal variability affects the *convergence* of our results on the long term average. In section 4.6.2 we consider the *repeatability* of our results by running replicate simulations with the same network structure and parameter values.

4.6.1 Convergence

In this section we compare the performance of different estimators for long-term mean species abundance. In all that follows the initial 1000 time steps are removed from the simulated time series, as has been done throughout the chapter. Therefore when we refer to a time series below, we refer to this truncated version of the time series. We define the long-term average abundance of species i as:

$$(4.6) \quad \bar{X}_i = \frac{1}{T} \sum_{t=1}^T x_{i,t},$$

where T is the number of time steps in the time series, and $x_{i,t}$ is the abundance of species i at time point t . We denote by e_i an estimator of the long-term average abundance \bar{X}_i . The snapshot estimator e_{1i} for species i is defined as the abundance of species i on the 5000th time step. Other

⁴Really we should also check for trends.

estimators are calculated by taking the mean species abundance over different sample lengths. Sample are always taken from the beginning of the time series.

Figure 4.24 shows three different estimators plotted against long-term average abundances. Two simulations are shown, one with high IR (0.001) and one with low IR (0.0001). Also plotted are linear regression fits to the data, which for a perfect estimator would lie on the line $y = x$ and therefore have a regression coefficient (R^2) equal to one. In this way R^2 can be used to assess the quality of any given estimator as a measure of the long-term abundances of all species in a community. From the plots it is clear that the longer samples produce better estimates than the snapshot. In particular at low IR, where there is increased temporal variability, the performance of the snapshot estimator is poor. We define another metric for estimator quality, which we call *mean relative error* MRE:

$$(4.7) \quad MRE = \frac{1}{N} \sum_{i=1}^N |\bar{X}_i - e_i|,$$

where N is the number of species, and e_i is the value of the estimator in question for species i . This metric is equal to zero for a perfect estimator. Figure 4.25 shows the performance of estimators, calculated from increasing sample lengths, applied to the HI and the LI ensembles. Performance is measured by both the R^2 value and by MRE. In all cases the estimators converge on the long term average as expected. However for the LI ensemble the MRE converges very slowly. Even with sample lengths of 30,000 time steps the MRE is still around 0.1. Furthermore the performance of the snapshot estimators is very poor, with MREs of around 0.3 and 0.8 for the HI and LI ensembles respectively. From these results it appears that the snapshot estimator would introduce large errors into our results in cases where there is high temporal variability. However it is also clear that some trade-off must be made, because the length of sample required to guarantee convergence is very large.

FINISH: The estimator convergence depicted in figure 4.25 suggests that very long samples are required in order to reduce the expected deviation in abundance estimates from the long term average (measured by MRE (4.7)). However, based on the regression analysis, much shorter samples...suggesting that samples of length 4000 should be sufficient to characterise the system with a high level of accuracy.

4.6.2 Repeatability

Here we briefly consider the issue of repeatability in our experiments. I have heard IBMs (or *agent-based model*) referred to as ‘*very complicated random number generators*’. Although I can find no reference to such a critique in the literature, it is certainly possible that sufficient stochasticity could result in replicates of the same model producing wildly different outputs. Various results presented in this chapter suggest that this is not the case here. In section 4.2.5 we saw that replicate simulations with two different network structures generated consistently different persistence profiles. Also the RQA analysis (section 4.5) revealed a strong deterministic

signature in species population dynamics. The detection of determinism suggests that the local rules of the IBM successfully define mechanisms responsible for the observed patterns and dynamics, rather than simply producing randomness. As a final demonstration of repeatability we selected three replicate simulations at random from the NM1 ensemble (section 4.4.3). These three simulations use the same interaction network structure. We compare the abundance distributions of the three replicates using *rank-abundance spectra* (RAS) [69]. RAS are a useful tool for the comparison of abundances in communities that contain the same species. Species are ranked according to their abundance in one community (as in RADs - section 2.7.1.1). The rank of each species is then fixed, giving its spectral location, which allows the comparison of the initial community with several others. [RAS](#) are plotted for the three NM1 replicates in figure 4.26, using the long-term average abundance for each species (4.7). We provide no more than a qualitative description of the RAS, which clearly illustrate a high level of repeatability in the abundance distribution. There are only a few significant deviations from the rank ordering in the initial community (panel A) in the replicate communities (panels B,C). In particular the species at rank 8 has a much lower abundance in panel C than in A and B. It appears that this is due to competition from species in the same trophic level (ranks 38 and 54).

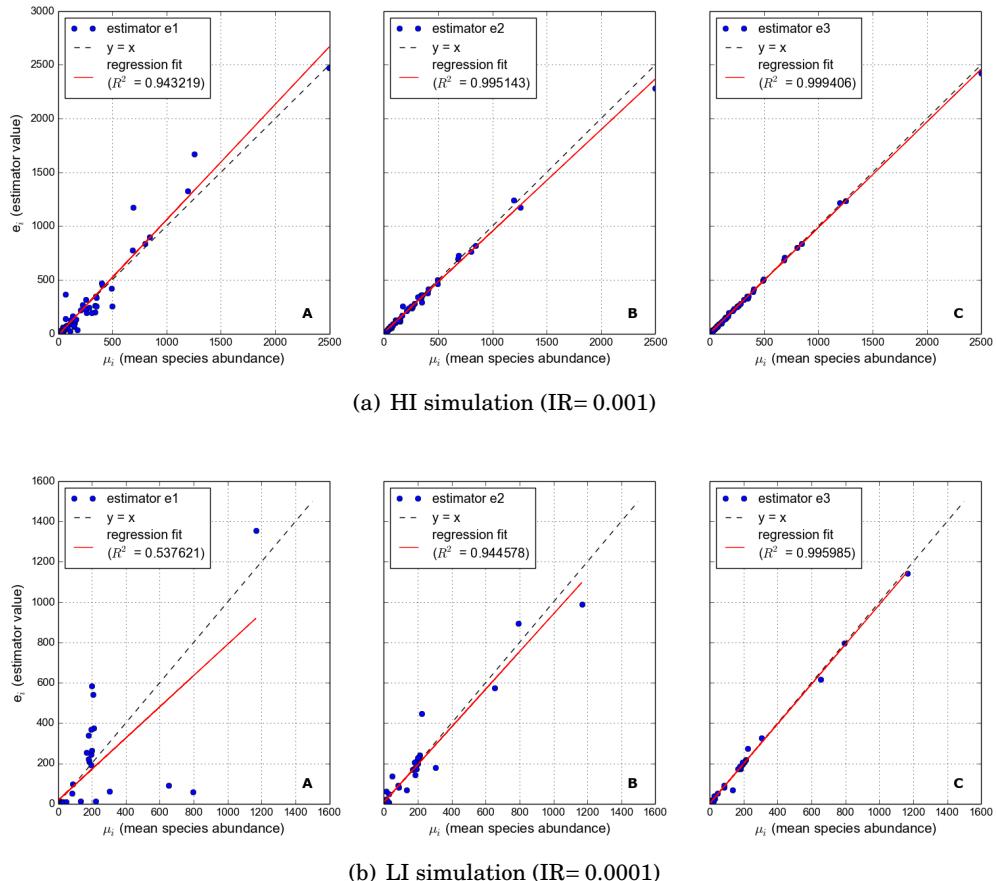


Figure 4.24: Performance of three different estimators of species abundance, applied to a (A) single HI simulation and (B) a single LI simulation. Estimator ‘e1’ is a snapshot of abundance on 5000th time step; ‘e2’ is an average over 4000 time steps; and ‘e3’ is an average over 29,000 time steps. Points are the long-term mean abundance of a species, plotted against the estimator value for that species. Red lines show linear regression fits for the estimator, and how close the estimator is to modelling the ‘true value’ of species abundances (dashed blue line).

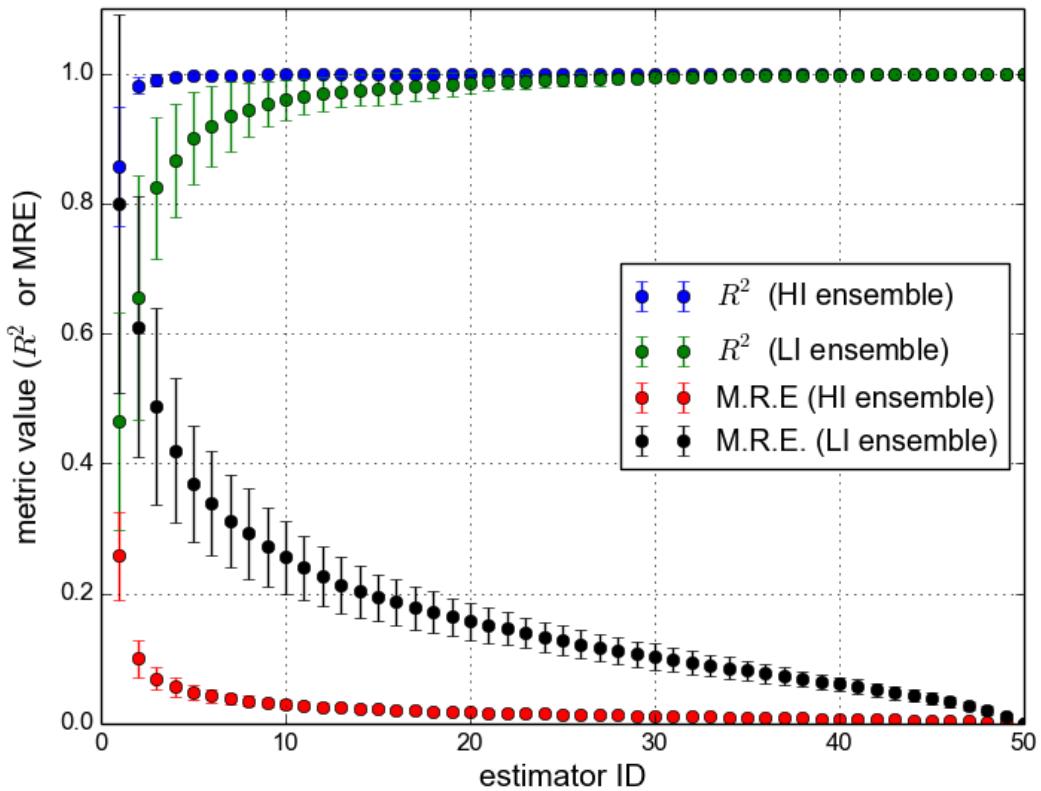


Figure 4.25: Performance of 50 different estimators for species abundance, measured by R^2 value and mean relative error MRE (see text for definitions). The first estimator (estimator ID=1) is uses ‘snapshot’ of simulation state at 5000th time step (as used in previous chapter), the remaining estimators (IDs= 2,..,50) use average abundances over sample windows ranging from length 1000 to 50,000 in steps of 1000. Points indicate mean value of metric over ensemble of simulations. Error bars indicate ± 1 standard deviation.

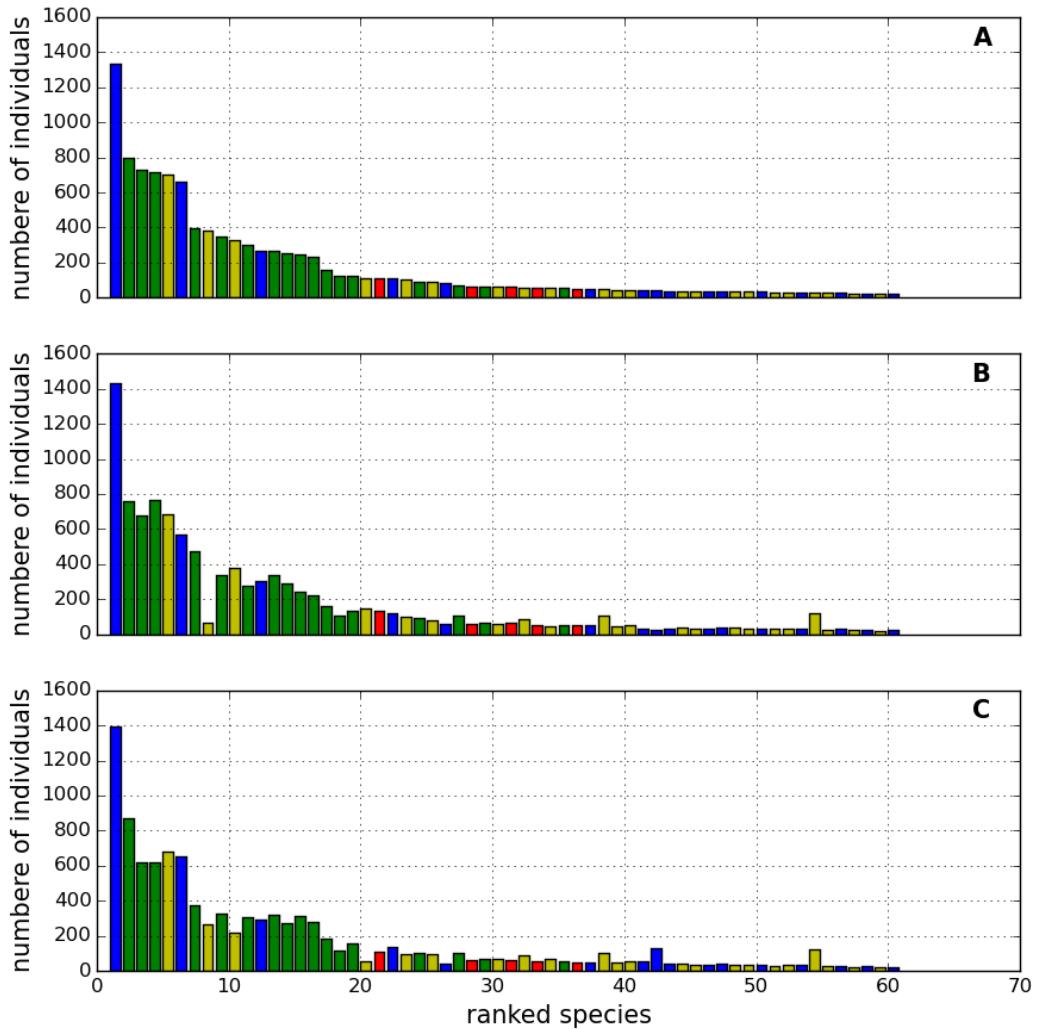


Figure 4.26: Rank abundance spectra (RAS) for three simulations from the NM1 ensemble (see text). Species abundances measured as long term average (over time steps 1000-9,000). Species are ranked according to their abundances in the first simulation (panel (A)). Ranking is retained in panels (B) and (C), which show abundances from two different simulations. Colouring of species by trophic level is consistent with previous figures.

4.7 Conclusion

In this chapter we have discovered that immigration is required for the IBM to generate persistent communities. In the absence of immigration the majority of species in higher trophic levels go extinct, and we have argued that this is due to competition effects. We have also demonstrated the population dynamics becomes more variable as the immigration rate is reduced. This may be a feature of the system passing between a regime of stable persistence at high IR into an unstable regime at low IR. Furthermore we have demonstrated that species population dynamics may be non-stationary, especially at low IR. However results from the RQA suggest that this non-stationarity is due to deterministic population dynamics rather than stochastic effects. We have also shown that the dynamics of species with low abundance tend to be more stationary and less deterministic, whereas species with high abundance tend to be less stationary and more deterministic. Together these results suggest that the abundant species are undergoing larger amplitude trophic dynamics, generated by their interactions with other species. The least abundant species, we propose, are those maintained mostly by immigration (even at lower IR values) because their relative scarcity makes them less likely to interact with other species. As we argued in chapter 3, immigration is a random process. So the dependence of low abundance species on immigration would explain their lower %DET values and, together with their reduced ability to interact with other species explains their higher stationarity.

We have addressed the impact of increased variability on our experimental results. We showed that, especially at low IR, snapshot sampling can produce abundances measures with large deviation from the long term average. However it is also clear that, for highly variable communities, the length of sample required for the measured abundance to converge on the long term average is *infeasible*. Therefore we accept that increase variability may introduce random error into our results. In chapter ?? we address this issue by comparing two different sampling regimes in the analysis. We have also briefly considered the issue of repeatability in our experimental results. In particular we saw that replicate simulations using the same interaction network produced very similar rank-abundance spectra (RAS). The reproducible RAS, together with the observed signature of determinism in the IBM output, give us confidence in the repeatability of our results. Although clearly the strong stochastic component requires that replicates are part of the experimental procedure.

