

## STATIONARITY

TODO: not forget this useful link :) [jasss.soc.surrey.ac.uk/15/2/7.html](http://jasss.soc.surrey.ac.uk/15/2/7.html) !!

TODO: how do metrics vary with measurement length/through time...

## 1.1 Motivation

We look here at stationarity because it was seen in previous chapter that in certain circumstances the population dynamics become highly variable, as measured by the temporal variability metric.

Ecological relevance... There has been a strong tradition of understanding ecosystems as existing at or near some stochastic equilibrium or steady state [2] (issue of short time scale fluctuations and thermodynamics systems). This is motivated partly by observation of constancy in ecological communities through time [REFS] and also by theoretical considerations - for examples various stability metrics require the assumption of a stable equilibrium.. stability people ?? and section ??.

There is also the concern that high variability affects the results we present... previously results were collected.. but averaged over replicates... here we look for stationarity.. if stationary.. we also look at how results differ depending on how they are measured for a highly variable simulation... OUTLINE WHAT IS TO FOLLOW:

- Thermodynamic equilibrium is... therefore can talk about extrinsic state variables..
- Different concepts of equilibrium in ecology - IBT suggest that island communities exist in a dynamic equilibrium between immigration and local extinctions.. [5].

- Steady-state can mean different things in different contexts. In general we define it as a condition under which certain properties of the system can be defined which are unchanging in time. For example..a chaotic attractor can be considered a steady-state
- Whether or not real we consider...
- Chaotic attractor could be steady-state..

In the previous work ([4] and section ??) the results presented were mainly taken as a final value at the end of the simulation. This is true for the Diversity and Abundance metrics, taken from outputecosystem. This is also true of mutualistic subnetwork - which is only taken at the final iteration (over period of 200 iterations). And Network properties are all calculated from the final 200 iterations of the simulation i.e 4800 to 5000. Here we investigate this, the assumptions behind it and the potential affect on results. This investigation is in part inspired by the result that the CoV in the total biomass increases as the IR is decreased, bringing into question the assumptions behind taking 'snapshot results'.

## 1.2 Second order stationarity

- Justify choice of weak-stationarity
- discuss tests for weak stationarity e.g. Nason [? ]. PSR. ADF. KPSS. define them.
- mention use of wavelets in ecological time series - refer to PSR test
- Choice of metric to look at : total number of individuals, or by species?? (justify both..)

The first question we can ask is whether the time-series is stationary, which is defined as the time invariance of certain statistical properties of the time-series. We focus here on weak stationarity, or second-order stationarity, so-called because it requires the time invariance of the first and second moments of the series. Specifically it is defined as [? ]:

$$(1.1) \quad \mathbf{E}(x(t)) = m(t) = m(t + \tau) \tau \in \mathbf{R},$$

for a process  $x(t)$ . Intuitively this is the definition of stationarity we are looking for in the model output - if the generated time series is weakly stationary then the simulation has reached a steady state distribution, and the results that we take involve sampling from this distribution. The repeatability of the results then only depends on the properties of this distribution (we also need to look at individual species..see later).

### 1.2.1 Tests for stationarity

We compare three different tests of stationarity: KPSS, ADF, PSR.

The ADF test is...

The PSR test is...

The KPSS test is...

### 1.2.2 Characterising the tests

We run 100 repeats of long 50,000 iteration simulations, for the base case of zero HL and no mutualism. This is done for a high level of IR (0.001) and a low level of IR (0.0001). We will refer to these two cases here as high immigration (HI) and low immigration (LI) respectively. Since the temporal variation is greater, on average, for lower IR, we will first focus on the HI scenario - this is closer to the previous work that has been done using the model and is more likely to reach a steady-state as desired. For the initial analysis we select a single HI simulation run. The total abundance dynamics of this simulation is shown in panel (A) of figure 1.1. The total abundance has a mean and standard deviation of 15915.8 and 1545.6 respectively.

To check the functioning of these tests we also generate two time-series of known stationarity. The first (panel B figure 1.1) is a simple one-dimensional random walk defined by the series:

$$(1.2) \quad x(t) = \sum_{i=1}^t Z_i,$$

where  $Z_i$  are independent random variables that may take values of  $-10$  or  $+10$ , both with probability a half. An ensemble of random walks was generated and a single instance was chosen with mean and variance closest to the HI simulation run. These values are 15525.2 and 1549.8. We know that such a series is non-stationary in general (although it may appear stationary by chance?). The second series (panel C figure 1.1) is stationary. The value at each time point is independently drawn from a normal distribution with mean and variance equal to the original HI series.

The ADF test states for all three time-series we can reject the null-hypothesis of stationarity at 99% confidence. Looking at the ADF statistics in table 1.2.2 we see that the ranking of these series is as expected. The more negative the tests statistic, the more evidence there is to reject the null hypothesis of stationarity. According to the ADF test the NS series is most likely to be stationary, followed by the HI series and then the RW. This seems like the correct order although the evidence for stationarity of the random walk may be too strong. The KPSS test also ranks the series in this order. According to this test the NS is clearly stationary (accept  $h_0$ ) and the

	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 1.1: Results of applying the three stationarity tests to the example time series shown in figure 1.1. P-values that indicate evidence for stationarity at 95% are highlighted in green. The test statistics are also given for the ADF and KPSS tests.

RW is clearly non-stationary. However the HI series is borderline. There is sufficient evidence to reject the null-hypothesis at 95%, but not at 99%. The PSR test, which is based on the time variability of the frequency spectrum of the signal, appears to fail completely. Therefore we do not trust it. (But we may come back to this in the wavelet discussion below)..but it does suggest that... Ref Nason paper (above) - Grouse populations mode of oscillation varies..Also good way to see transience!

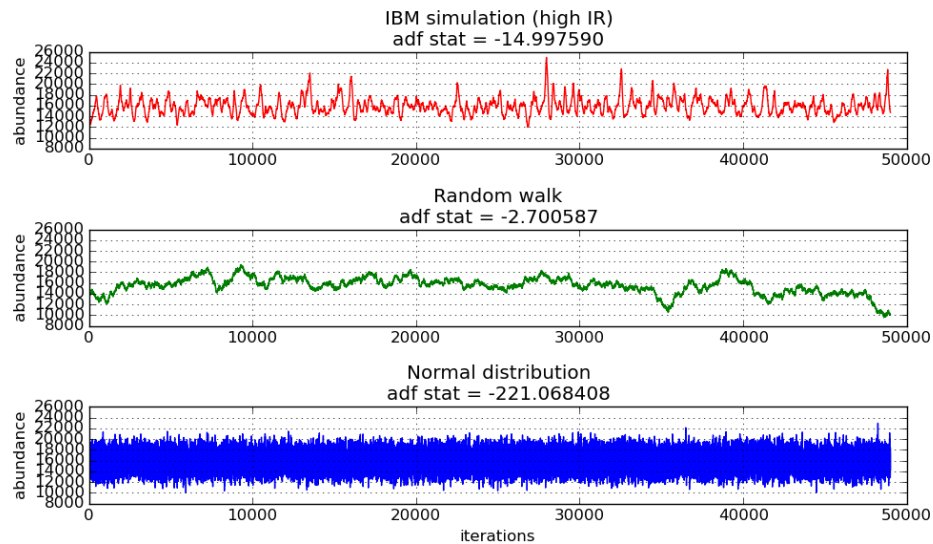


Figure 1.1: The three time series used to characterise the performance of the stationarity tests. The initial 1000 points removed such that all are 49,000 points long. (A) **HI**: total abundance dynamics of an IBM simulation with high immigration rate; (B) **RW**: a random walk without drift, as described in the text; and (C) **NS**: a series generated by independent sampling from a normal distribution.

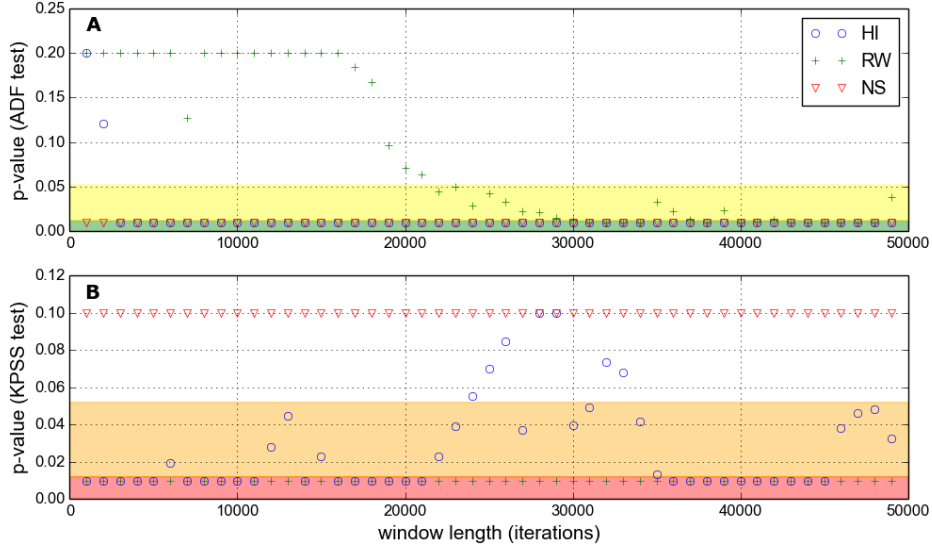


Figure 1.2: Two tests for stationarity applied to samples of varying size (window length). Samples are taken from the three time series shown in figure 1.1 - the total number of species of the high immigration (HI) simulation run, and the two control time series (RW and NS). All three time series contain 50,000 points. Sample windows begin at the 1000th time point (since this removes the initial transience from the HI simulation), and increase in length up to 49,000 points. (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 0.01. 95th and 99th percentile in orange and red respectively, indicating significant evidence for non-stationarity.

The ADF and KPSS are tested for varying sampling sizes, using the three example time series shown in figure 1.1. The two tests again to perform differently (refer to table 1.2.2). The KPSS test correctly identifies the RW and NS series as non-stationary and stationary respectively, for all sample sizes (1000 – 49,000). This is shown in panel B of figure 1.2. The ADF test (panel A figure 1.2) correctly identifies the NS series as stationary for all sample sizes. For short sample sizes it also identifies the RW series as non-stationary. However, for sample sizes much above 20,000 the ADF test finds sufficient evidence to reject the null hypothesis that the RW series is non-stationary, at 95%. This is an interesting result because the RW is generated by a non-stationary process. However it is a specific instance of this process, which to the eye does appear quite constant, especially over the central part of the time series. Therefore it is not clear if the time series should be classified as stationary or non-stationary. It is interesting that the result of the tests depends on the length of sample used, suggesting that..

There is mixed evidence for the stationarity of the HI series, as shown in figure 1.2. The ADF test, for all sample sizes above 2000, finds significant evidence that the HI series is stationary. Whereas the KPSS test, on the whole, gives significant evidence that the HI series 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.) It appears that the KPSS test is a more stringent test for stationarity, whereas the ADF test is more easy to fool by series that appear stationary [REPHRASE - USE ENSEMBLE TO JUSTIFY THIS?? ALAN?]. This knowledge is useful, and we will continue to use both tests.

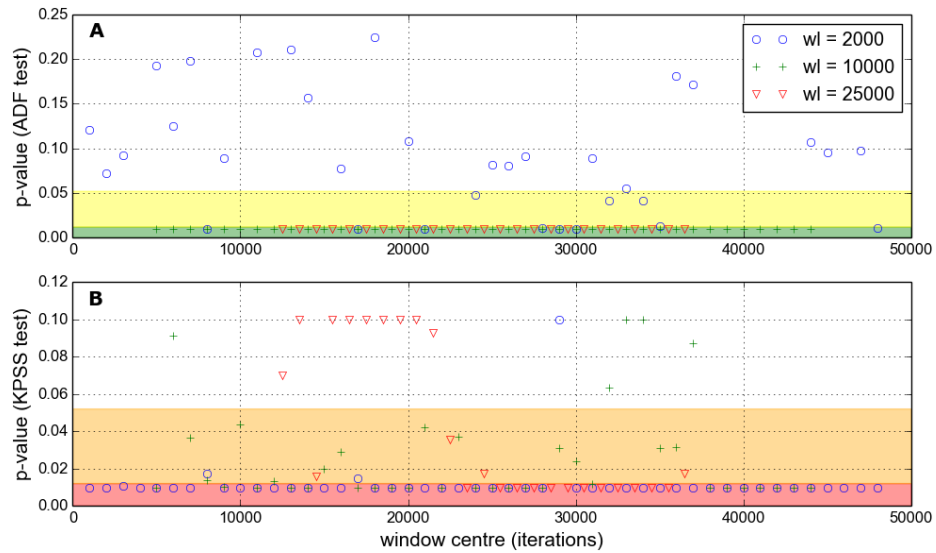


Figure 1.3: Similar to figure 1.2 but with samples of fixed sample size taken from different parts of the time series. To sample windows of given length (wl) are moved along the series and the tests are applied to the sub-series that falls within the window. Results are plotted against the mid-point of the window.

From inspection of the dynamics, there is a quick initial transience ( $\sim 500$  iterations) followed by a long period of relatively constant number of individuals. This pattern appears to be ubiquitous at high immigration rates, however it does not necessarily hold for low immigration rates. In these cases we noted a high temporal variability in abundances. Therefore we look next at the case of low IR..

In the case of high IR we saw that the evidence for stationarity increases with window length. We also tested for the possibility that the simulation only reaches stationarity after a certain number of iterations - the results were qualitatively similar when sampling in reverse. Figure 1.3 shows the tests applied to the HI series for three different window lengths. The hypothesis would be shown by initially no evidence for stationarity, followed by increased evidence towards the end. This is not seen.

The time series is further broken down by trophic level and by species. It may be the case that the total number of individuals is stationary, but that the trophic or species level dynamics cause changes in community composition over time. It is also likely that different species may exhibit different types of dynamics, and therefore it makes more sense to test individual species. Furthermore...biomass argument.

There is an apparent difference in the stationarity of the four trophic levels. Figure 1.4 shows  
 There is also apparent differen between species...as in figure 1.7..

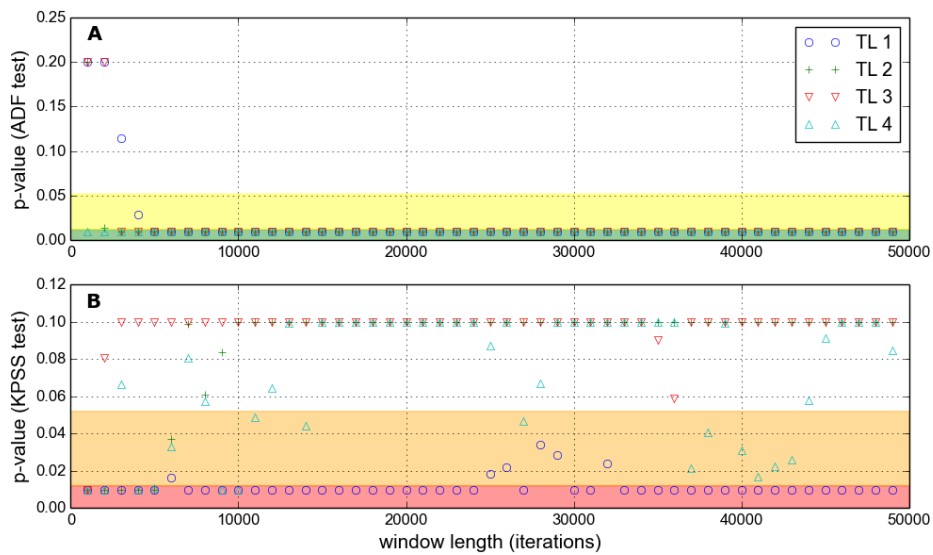


Figure 1.4: Similar to figure 1.2, but here the tests are applied separately to each trophic level of the HI simulation. The four time series (TL1 – 4) represent the total number of individual belonging to each trophic level at every iteration.

### 1.3 Chaotic dynamics?

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 [3]. We also draw inspiration from the demonstration that plankton communities may undergo chaotic dynamics - [1], and their focus on the Lyapunov exponent.

### 1.4 Discussion

This behaviour may or may not be seen in real communities - chaotic dynamics have been demonstrated in plankton, how about terrestrial ecosystems? However we come back again to limitations - 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 approaches of taking snapshots and averaging over many iterations...DISCUSS WITH ALAN.

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

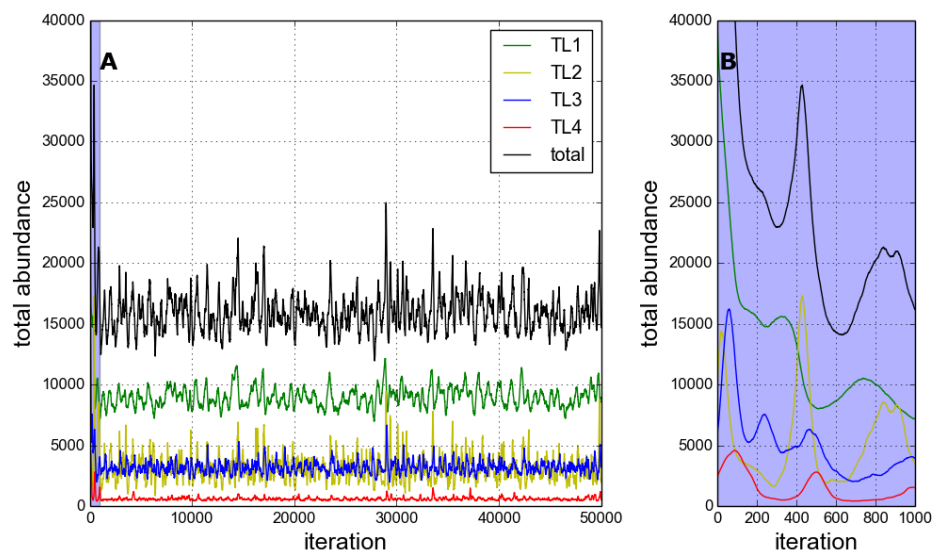
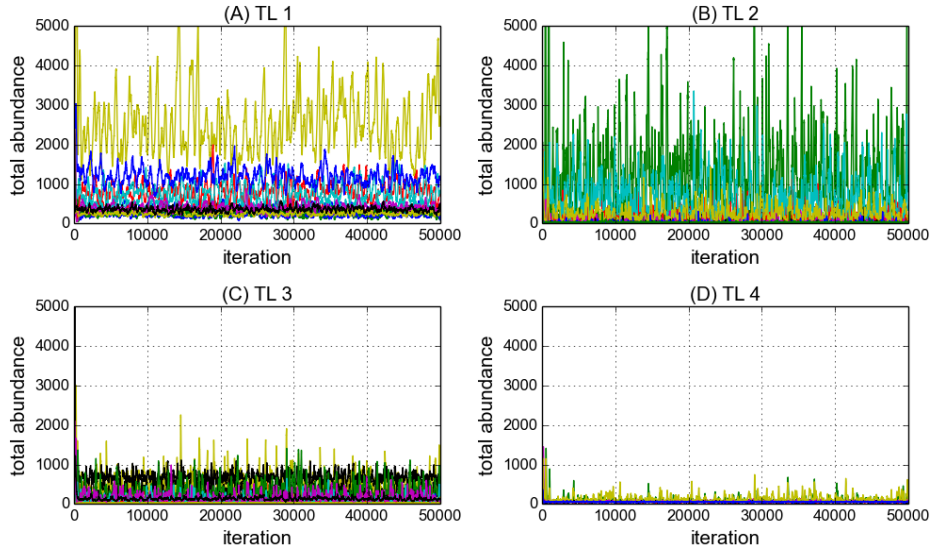
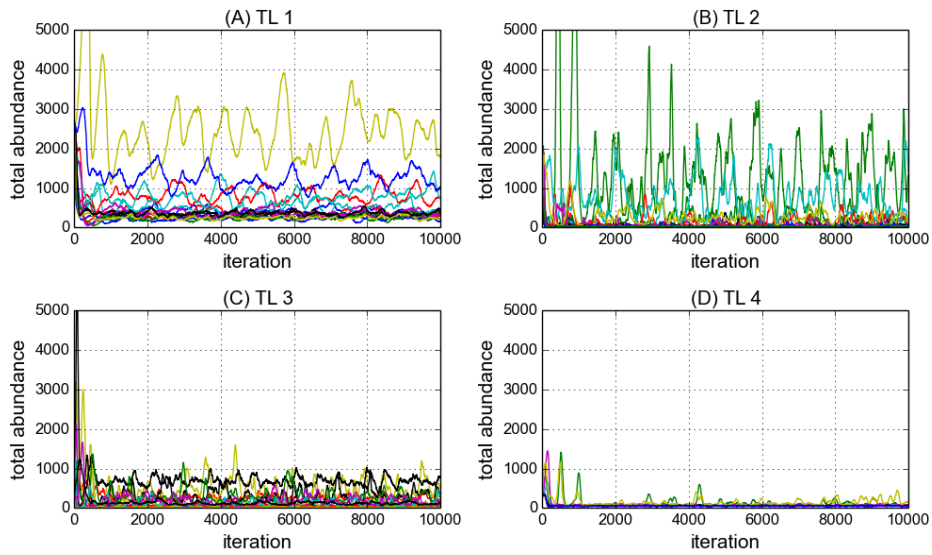


Figure 1.5: Dynamics by trophic level for HI simulation.



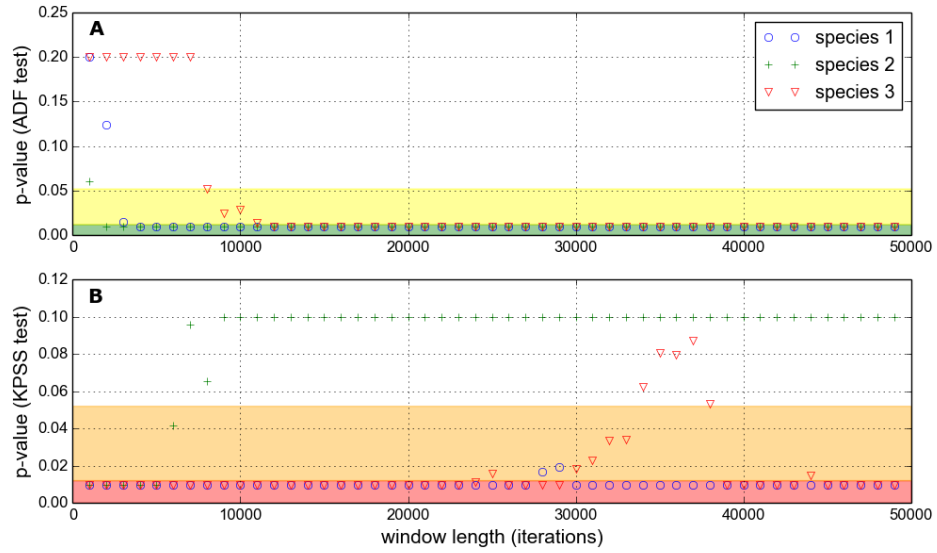


(a) Sample size = 1000 iterations

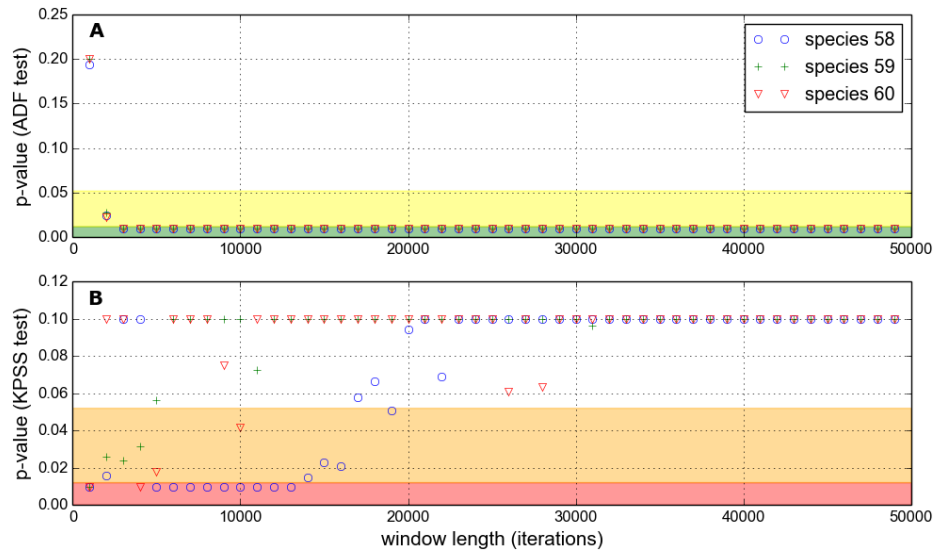


(b) Sample size = 5000 iterations

Figure 1.6: The dynamics of individual species.

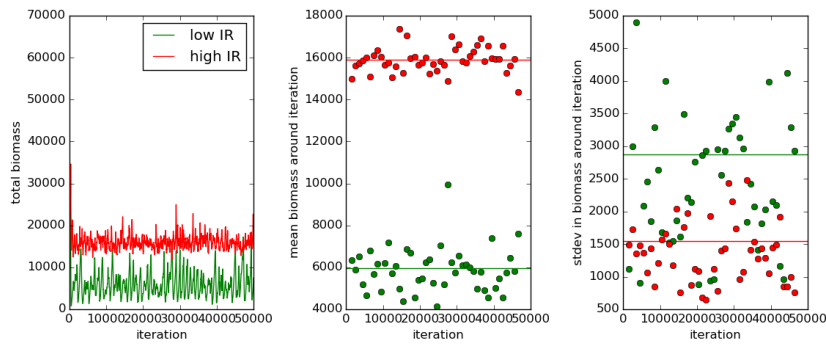


(a) Three most abundant species

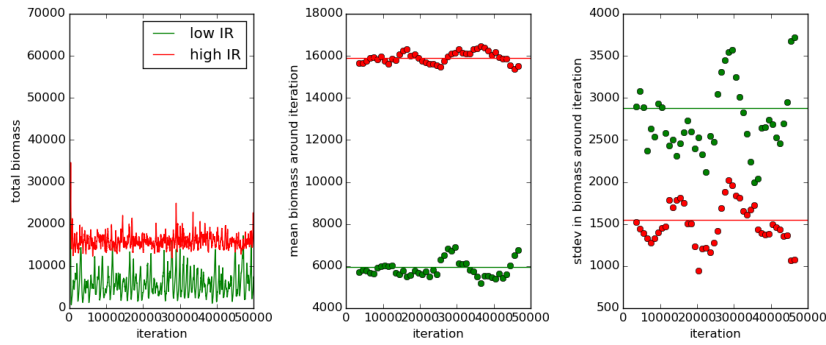


(b) Three least abundant species

Figure 1.7: Similar to figure 1.2, 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.



(a) Sample size = 1000 iterations



(b) Sample size = 5000 iterations

Figure 1.8: The effect of using different sample sizes on the sample mean and standard deviation. Dynamics generated using IBM simulation model with low and high IR (green and red respectively).



## BIBLIOGRAPHY

- [1] 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.
- [2] T. D. BROCK, *The ecosystem and the steady state*, Bioscience, 17 (1967), pp. 166–169.
- [3] S. HAZLEDINE, *The Analysis and modelling of chaotic calcium oscillations*, PhD thesis, Department for Computational Systems Biology, John Innes Centre, 2009.
- [4] 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.
- [5] D. S. SIMBERLOFF, *Equilibrium theory of island biogeography and ecology*, Annual review of Ecology and Systematics, (1974), pp. 161–182.

