Estimating Natural Scales of Ecological Systems

by

Rebecca Lynn Habeeb, MSc
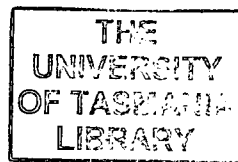
Submitted in fulfillment of the requirements for the Degree of

Doctor of Philosophy

University of Tasmania
June 2005

Statement of Originality

This thesis contains no material that has been accepted for a degree or diploma by the University or any other institution, and to the best of my knowledge and belief contains no material previously published or written by another person except where due acknowledgement is made in the text of the thesis.

Statement of Authority of Access

This thesis may be made available for loan and limited copying in accordance with the *Copyright Act 1968*.

Rebecca Lynn Habeeb

Acknowledgements

Dedication

I would like to dedicate this work to the memory of my father, the honorable Shouphie Habeeb, who passed away during the second year of completion of this thesis. His brilliance inspired me, his encouragement led me, and his love nourished me. I hope that you are proud, Daddy. May your light continue to shine through me. Shmula.

And to my mom, thank you for your continual support and love throughout this thesis. Now it is my turn to support you. I'll be home soon. I love you.

Abstract

A fundamental goal in ecology is to identify the appropriate scales at which to observe trends in ecosystem behavior. Researchers often rely on intuition to choose a sampling scale, which if too large or too small potentially obscures the real system trends. A characteristic length scale (CLS) is a natural scale of a system at which the underlying deterministic dynamics are most clearly observed over stochasticity. The overarching aim of this research was to develop, examine and apply a robust technique to detect CLSs of real ecological systems.

In this thesis, I first compared the robustness of two CLS methods, both of which account for complex oscillatory dynamics of ecological systems using attractor reconstruction from long time series of data. I applied these techniques to estimate CLSs of spatial multispecies systems of varying complexity, showing that for more complex models, the prediction r^2 metric of Pascual and Levin (1999) is a robust method. I then used an alternative method of CLS estimation, based on prediction r^2 but where repetition in space is largely substituted for repetition in time in attractor reconstruction, to determine the CLS of a natural marine fouling system. The new technique, requiring only a short temporal sequence of as few as three highly resolved spatial maps, enabled unambiguous length scales to be estimated for this system. Importantly, the estimated CLS was similarly based on analysis of several species representing a spectrum of phyla and life history patterns, indicating the adequacy of this method for objectively determining optimal scales of observation for real systems. Moreover, the CLS estimates of this system remained surprisingly consistent despite changes in time intervals between the spatial maps, changes in the number of maps used in the temporal sequence, and varying start dates of map sequences.

When the field results were compared to those derived from a spatially explicit individual-based model of a similar marine community nearby, the average CLSs of the two communities were strikingly similar (~ 0.35 m), suggesting that dynamical trends of like systems may be best observed on similar scales.

I also considered whether the approach could be extended to identify trends based on the dynamics of interactions among habitat types as opposed to interactions among species groups. Analysis of maps produced from remote sensing data at the habitat level successfully revealed unambiguous characteristic length scales of a coral reef. Habitats distinct in their species assemblages, abundances and morphologies provided similar length scales (~ 300 m), suggesting that the system-level CLS detected is independent of the habitat type used for its estimation.

Different species within the same system indicated dissimilar CLSs in some spatial model communities where species were only weakly connected, either due to the topology of network interactions or through spatial isolation as a result of spatial self-organization and patchiness. I finally evaluated the sensitivity of CLS estimates to varying levels of species connectivity and found that when species were weakly connected in terms of their network topology, the estimates of scale were likely to be dissimilar.

The results contained within this thesis illustrate that the new method developed to detect natural scales of complex systems (1) can be applied to natural dynamical systems with a reasonable quantity of data, (2) is robust to changes in parameters that should not affect the scale of observation (e.g. initial configuration of the landscape), but (3) is sensitive to variations that would likely affect the spatial scale at which trends can be observed (e.g. connectivity). Species or habitat level data can be used to estimate CLSs, which will be critical for choosing the scale at which to sample real ecosystems to distinguish trends from random variation. Characteristic length scales can now be implemented to objectively guide the choice of observation scale, and I expect that in the future, they will become a part of every assiduous ecologist's toolbox.

Table of Contents

Chapter 1

General Introduction

1.1 Determining appropriate scales to observe ecological systems

One of the fundamental challenges in ecology is to determine the dynamical processes underlying observed patterns (Levin 1992). The greatest difficulty confronting this objective is that any pattern detected, and ultimately the understanding of underlying dynamical processes, depends on the spatial scale at which observations are made (Wiens 1989, Levin 1992, Schneider 1994, Levin 2000). The challenge, which is becoming increasingly significant as new work brings the scale-dependence of fundamental functional relationships into clearer focus (Crawley and Harral 2001, Chase and Leibold 2002, Chalcraft et al. 2004), is to identify appropriate scales of observation for ecological investigation (Levin 1992). In determining the scales for examining an ecosystem's dynamics, a crucial question to ask is whether ecological systems have natural or "characteristic" scales at which observations provide unambiguous information about the dynamic, while minimizing noise in the signal measured. If so, then the capacity to objectively identify these optimum scales of observation will be important in the monitoring of an ecosystem to detect meaningful changes in its state (Rand 1994, Bishop et al. 2002).

Characteristic length scales (CLSs) are, in theory, intrinsic to the system, and are the scales at which the stochastic signal is minimized and the deterministic, or ecologically meaningful, signal is maximized. The underlying tenet is that if the scale used to sample an ecosystem is too small, observations are swamped by noise due to strong correlations among individual samples (Durrett and Levin 2000, Wilson and Keeling 2000). If the sampling scale is too large, the

nontrivial dynamics will be averaged out because distant parts of the landscape begin to act independently (Keeling et al. 1997, Pascual and Levin 1999, Wilson and Keeling 2000). The CLS is an intermediate scale which most clearly reflects the underlying deterministic signal (Pascual and Levin 1999, Wilson and Keeling 2000; for illustration, see Figure 1.1). Use of this natural scale as the size of sampling units (e.g. for monitoring the abundance of a particular harvested organism) will enable detection of the true system trends.

The idea that these intrinsic scales exist for ecological systems has a long history, and several methods have been proposed to estimate these scales (e.g. Grieg-Smith 1952, Kershaw 1957, Carlile et al. 1989, De Roos et al. 1991, Schneider 1994, Rand and Wilson 1995). Each of these approaches employs some form of variance spectrum, expressing a measure of variance among observations as a function of the size of the 'window' of observation. In early vegetation studies, scientists sought the inherent spatial scale of a landscape, or the intermediate scale that best described the processes driving the distribution of species (Greig-Smith 1952, Carlile et al. 1989). However, most approaches assumed that ecological systems are stationary in space and time, that fluctuations are random around a stationary global average (e.g. Rand and Wilson 1995), or that any dynamical behaviors are linear (see Turner et al. 1991).

These kinds of approaches fail to take into account the complex dynamical nature of ecological systems in space and time, which is characterized by nonlinear oscillatory behaviors (e.g. Hastings et al. 1993, Ellner and Turchin 1995, Sole and Bascompte 1995, Little et al. 1996, Pascual and Ellner 2000). More recent approaches have suggested that characteristic length scales (CLSs) are detectable in spatial models of oscillating systems, such as predator-prey or
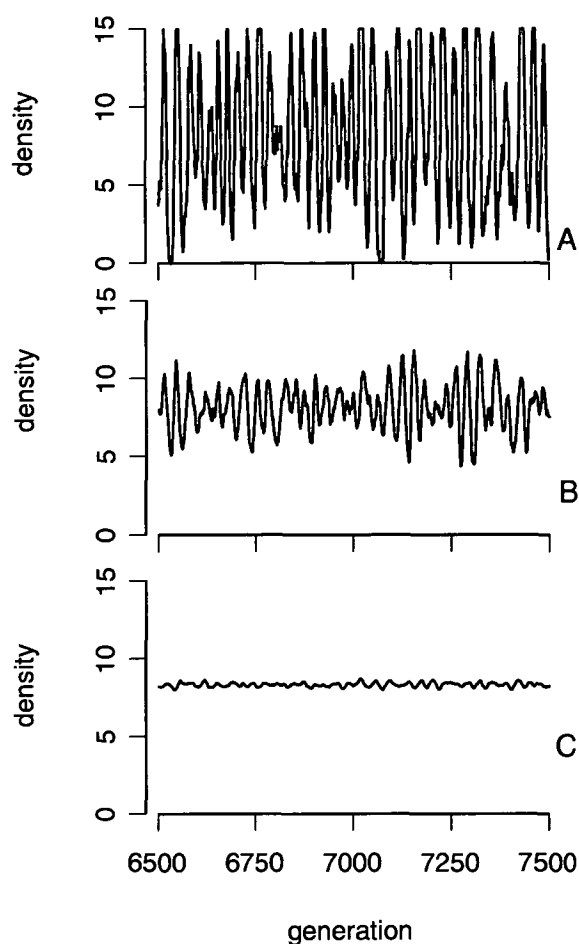
Figure 1.1. Representative temporal trajectories of the density of a single species in a 3-species model system (see Chapter 2) determined by observing the system at three different window sizes on a 500 x 500 landscape. Observations are made (A) through a small window (5 x 5 cells), (B) through an intermediate sized window similar to the CLS of this system (30 x 30 cells), and (C) at the size of the landscape. Observing the system through a small window emphasizes noise in the signal, while averaging over larger windows decreases the amplitude of the trajectory. The trajectory at the intermediate window size best illustrates the smooth yet oscillatory dynamics of the system. It is this feature of change in the characteristics of the signal with window size that CLS methods attempt to quantify (Pascual and Levin 1999).

host-parasite systems (Rand and Wilson 1995, Keeling et al. 1997, Pascual and

Levin 1999). In particular, the techniques of Keeling et al. (1997) and Pascual and

Levin (1999) to identify CLSs are exceptional in that they accommodate both the

complex nonlinear oscillations and non-uniform patterns in spatial variance that

are pervasive in real ecosystems. The crucial development of these new approaches is the application of attractor reconstruction (Takens 1981) and prediction algorithms (Kaplan and Glass 1995) from nonlinear time series analysis to characterize dynamics at particular scales of observation. In general, the method allows the phase space of the ecosystem to be built from time delay coordinates of a single variable, which acts as a surrogate for the unobserved variables of the system. The trajectories in phase space are then used to construct a model of algorithms, which predict the trajectories in n-dimensional space some time ahead (Kaplan and Glass 1995, Abarbanel 1996). This procedure is described in detail in Appendix A.

Using long time series of the density of a single species, the attractor is reconstructed for each scale, generating a variance spectrum from the prediction errors, plotted across the scales of observation. The region of plateau in prediction error dictates the CLS of the system, an estimate or range of scale within which the deterministic dynamics can most clearly be observed. From this range, the most appropriate scales at which to sample a system can be inferred. Implicit in the original techniques is the assumption that measurement of a single variable in the system over time can be used to illustrate the dynamics of the whole. For example, in a predator-prey system, the prey (or resource) density can be measured to illustrate the deterministic signal of the entire system (Rand and Wilson 1995). This assumption implies that the system-level CLS can be estimated by observations taken on a single species, and arises from the mathematics of attractor reconstruction (Takens, 1981; Kaplan and Glass, 1995; Little et al., 1996). This method is summarized in Appendix A, as it is fundamental to the approaches proposed by Keeling et al. (1997) and Pascual and

Levin (1999) to estimate CLSs.

1.2 Methods of estimating CLSs in oscillating systems

Both Keeling et al. (1997) and Pascual and Levin (1999) use attractor

reconstruction to identify the spatial scale that best distinguishes the deterministic

dynamic, or trend, from noise. Keeling et al. (1997) extended the approach of

Rand and Wilson (1995), by using attractor reconstruction in a method they called

fluctuation analysis, which models and then calculates deviations around the

underlying deterministic behavior within the system. The variance measure used

by Keeling et al. (1997), termed error X, is calculated as:

$$\text{error X} = L\sqrt{E_t[(X_L^t - \hat{X}_L^t)^2]}$$

where the density of a particular species X in a window of side length L is

$$X_L^t = \frac{1}{L^2}\sum_{i=1}^{L}\sum_{j=1}^{L}L_{ij}^t \quad \text{and} \quad E_t(X_L^t) = \pi \text{ is the probability of observing a particular}$$

species X. The predicted value of X_L^t is $\hat{X}_L^t$, as determined by the nonlinear time

series analysis. Where the variance measure (error X) plateaus with increasing

window size is the length scale where correlations have decayed and values at

different sites behave like independent random variables. Keeling et al. (1997)

used the plot of error X against L not only to estimate the CLS, but also to

identify scales of aggregation.

Pascual and Levin (1999) developed a 'determinism test', which uses

methods of nonlinear time series analysis, similar to those of Keeling et al.

(1997), to reconstruct the system attractor, but they focus on how the prediction

error of the trajectories changes with scale. They obtain a statistic called

prediction r^2, similar to the r^2 in linear regression, for each window size where:

$$\text{prediction } r^2 = 1 - \frac{E_t[(X_L^t - \hat{X}_L^t)^2]}{Var(X_L^t)}$$

The scale where the r^2 (or degree of determinism) is maximized is the CLS of the system. This measure is likely to indicate a CLS smaller than that identified by error X because maximum predictability may occur at a smaller scale than the onset of independence (Pascual and Levin 1999).

While the techniques of both Keeling et al. (1997) and Pascual and Levin (1999) show promise because of their capacity to accommodate complex oscillatory dynamics, they have thus far been applied only to simple model systems. Moreover, they require unrealistically long time series so that their application to most real ecosystems is impractical. Thus, while the theory of natural length scales has mathematical merit, more effort is necessary if it is to emerge as a tool that is practical for applied ecologists. Objective estimation of the optimal scales at which to sample a real ecosystem to detect trends is highly attractive to both ecologists and managers of natural resources, since sampling at inappropriate scales can lead to potential bias and/or misinterpretation of variation occurring at unstudied scales (Denny et al. 2004). Accordingly, the fundamental goal of this thesis is to develop, test and apply a modified method of CLS estimation that can provide an objective means to identify the natural scales of real ecological systems.

1.3 This thesis

With the goal of developing a technique to identify CLSs in real ecosystems, I first assess the robustness of the approaches of Keeling et al. (1997) and Pascual and Levin (1999) so that we can be confident in applying the

techniques to dynamical systems more complex than the simple models

investigated to date (Chapter 2). I examine the capacity of the two techniques to

identify unambiguous length scales for model systems spanning a spectrum of

complexity, and determine the sensitivity of CLS estimates to (1) the initial

spatial arrangement of individuals in a system, (2) parameter choices in attractor

reconstruction (i.e., τ, d_e, and k) and (3) the choice of species used in attractor

reconstruction. If these techniques are to be usefully applied to real systems, then

they should be robust to the spatial arrangement of individuals on the first

landscape of a time series, and to subjective choices of parameters used in

attractor reconstruction. Sensitivity to the choice of the single species from the

system that is studied is assessed to test the underlying assumption that

information from a single species can be used to identify the CLSs of the entire

system. If each species provides the same length scale, then the theory that the

CLS is a system-level measure is supported.

I then use a modified technique of attractor reconstruction, which obviates the

need for long time series, to determine the CLSs of a natural marine fouling

community (Chapter 3). By largely substituting repetition in space for repetition

in time, CLSs can be estimated with a series of as few as three highly resolved

landscape maps through time, rather than requiring tens of thousands of time steps

(Habeeb et al. 2005, Appendix B). I compare the CLSs indicated by 10 different

species groups representing several phyla and a range of life histories to test the

implicit assumption that dynamical information of a single species, irrespective of

its life history characteristics, within a multi-species system can be used to

estimate the unmeasured whole-system dynamic. To assess whether CLS

estimates are sensitive to particular site-specific characteristics, I compare the

results based on field data to those derived from a spatially explicit individual-

based model of a similar, nearby fouling community (Dunstan and Johnson 2005).

Given the novelty of the CLS metric, in Chapter 3, I also attempt to place the

measure in a broader ecological context by comparing it to other, more

established and better-understood ecological scaling measures. The intent is to

illuminate what, if anything, the CLS might reveal about the broader ecological

properties of a system. I chose conceptually simple spatial scales for comparison,

namely the region of asymptote in species-area curves (Conner and McCoy 1979)

and, with reference to patterns of species abundance, scales at which troughs and

peaks are evident in the relationship between variance-to-mean ratios and size of

the sampling unit. Interestingly, the spatial scale of the asymptote in species-area

curves was initially proposed as the appropriate scale of observation of

communities, with a motivation to capture most species in each sampling unit

(Greig-Smith 1964). Variance-to-mean ratio spectra have been used to describe

patterns of scale dependent aggregation in species distributions (Pielou 1977).

In the fourth chapter, I assess the sensitivity of the real fouling community

CLS estimates to arbitrary choices of the nature of the sampling regime by

varying the number of time steps in the time series (3, 4, or 5), the interval

between observations (3 month, 6 month, or 9 month), and the start date of the

time series (month 0, month 3, or month 6). If the CLS is, in fact, an inherent

characteristic of an ecosystem (Carlile et al. 1989, Wilson and Keeling 2000), and

providing that the attractor describing the system dynamic is stationary, then if

this concept is to be applied usefully to real ecological systems, the CLS should

not be sensitive to the particular protocol by which a system is sampled. Using

spatially detailed maps of the fouling community collected over 18 months, I

assess the robustness of the estimates to changes in varying combinations of these factors.

Because a great deal of monitoring real ecosystems is based on the dynamics of habitats rather than single species (e.g. McNeill 1994, Mumby et al. 2001, Rouget 2003), an important question is whether natural length scales are evident at the level of habitat types. The question is germane to several agendas in applied ecology, including reserve design and management zoning based on high-resolution habitat mapping that requires a measure of the appropriate scale at which to sample the habitat maps (Loehle and Wein 1994, Ward et al. 1999, Mumby et al. 2001). In Chapter 5, I apply the new length scale technique to remote sensing data collected at the habitat rather than species level, to extend the application of characteristic length scales to data based on interactions among habitat types. If CLSs can be determined at this level, sampling at the CLSs will optimize detection of critical ecological trends when monitoring changes in abundances of habitats or species assemblages, which is now becoming more prevalent than individual species monitoring (Peterson and Estes 2001). Using spatial maps of benthic habitats within a Caribbean coral reef over 21 years, I attempt to determine the system-level CLS based on the dynamics of three distinctly different coral reef habitats. The two broad aims of this component of the work are to ascertain (1) whether unambiguous length scales are evident from dynamics at this scale of biological organization, and (2) whether the different habitat types indicate similar CLSs.

In Chapter 6, I attempt to determine if the CLSs indicated by distinct species are influenced by the connectivity between those species. According to the mathematical theory, the system-level CLS can be determined by any species

within the system, provided that all species follow the same system dynamic, or attractor (Takens 1981). However, occasionally a species will provide an outlying CLS estimate (Chapter 3). Thus in this chapter, I attempt to determine if weak species connectivity, due to network topology, interaction strength, or spatial proximity and patchiness, leads to dissimilar CLS estimates between the constituent species of a system. Six species spatial model systems are run with varying levels of connectivity, in terms of both network topology and interaction strength. In each system, two groups of spatially connected species form and CLSs are compared between the two groups as connectivity changes, to evaluate the hypothesis that weak connectedness between species produces dissimilar CLS estimates. If the CLS is strongly influenced by species connectivity, then it may be potentially used as an indicator of relative connectivity between species within a system.

In concluding (Chapter 7), the combined results of this thesis are considered, and the utility of estimating the natural length scales of real ecological systems is assessed.

This thesis[*] provides the first investigation of identifying natural length scales of real ecological systems. Careful analysis of issues of robustness and sensitivity is an essential step in the application of any useful theory. The results affirm that characteristic length scales do exist for natural systems, that they can be detected using reasonable data sets, and that their potential for application in ecology, whether for management or basic research, is significant.

---

[*] Readers should note that the chapters contained in this thesis are written for submission as stand-alone manuscripts and therefore some repetition in their introductions is unavoidable.

Chapter 2

Characteristic Length Scales of Ecological Systems: Robustness of Estimates

2.1 Abstract

The scale used to view an ecosystem largely influences the patterns, and therefore the underlying processes that are detected. A key issue in ecology is identifying the appropriate scale(s) at which to observe ecosystem trends, namely the system's deterministic dynamic. The characteristic length scale (CLS) is the inherent or natural scale(s) of the system, at which the underlying deterministic dynamics can be best observed with minimal noise. Here, we compare the robustness of two recently developed methods to estimate CLSs. Both approaches use attractor reconstruction to account for the complex oscillatory dynamics of ecological systems, but they have been applied only to simple model systems. We used the Compete© software to model spatial multispecies systems of varying complexity, and applied these techniques to estimate CLSs in each system. We examined sensitivity of the CLS estimates to the choice of species and initial spatial conditions, across the range of model complexity. Robustness to changes in parameters used in attractor reconstruction, such as the time delay and embedding dimension, was also assessed. Both methods of CLS estimation were robust for the simplest model system, but as model complexity increased, the Pascual & Levin (1999) measure was more robust than that of Keeling et al. (1997) to changes in reconstruction parameters and initial conditions. Both

methods were sensitive to the choice of species used in the analysis of complex model systems. The connectivity of species appears to influence the CLS found, with closely connected species producing more similar CLSs than loosely connected species. In this sense, connectivity is determined both by the topology of the interaction network and spatial organization in the system. Notably, more complex systems that show spatial self-organization can yield multiple CLSs, with larger length scales indicating emergent dynamics resulting from interactions at the patch level. Overall, the approach of Pascual & Levin (1999) appears to be potentially useful for estimating system-level CLSs, but its greatest limitation is a requirement for long time series. We suggest that application of this technique, or some modification of it, can provide improved objectivity to important decisions about scaling in ecology.

2.2 Introduction

One of the fundamental challenges in ecology is to determine the dynamical processes underlying observed patterns (Levin 1992, Chave et al. 2002). However, the patterns detected depend upon the scale at which the system is viewed (Wiens 1989). The principle of spatial scale permeates ecology, with patterns of species interactions, coexistence, and exclusion varying from local to regional to global scales (Wiens 1989, Levin et al. 1997, Levin 2000, Bishop et al. 2002). Many scales of interest likely exist for a given application, and the challenge for scientists is to properly identify the most relevant scales for investigation (Levin 1992). In determining appropriate scales for examining an ecosystem's dynamics, a crucial question is whether natural or "characteristic" scales exist. If they do, then they may provide appropriate scales to observe

system behavior, retaining essential information about the dynamic without excessive noise. In addition to indicating optimal scales for monitoring dynamical change in ecosystem structure, characteristic scales may be helpful in defining minimum sizes for ecosystem conservation reserves and management units in harvested systems (Rand 1994, Rand and Wilson 1995, Wilson and Keeling 2000). Here, we address techniques designed to detect the characteristic scales of ecological systems.

Early studies of vegetation dynamics sought to identify the inherent 'intermediate' scale of a landscape that best reflected the underlying processes driving the heterogeneous distribution of species into patches (Greig-Smith 1952, Kershaw 1957, Carlile et al. 1989). Methods used in these studies (summarized in Turner et al. 1991) usually involved the development of variance spectra by sampling a landscape through "windows" of different sizes, either by combining small windows into larger windows (zooming) or by incrementally increasing the distance between windows of a given size (lagging, Schneider 1994). These early studies explicitly assumed that the system was both linear and uniform in space (Turner et al. 1991). Uniformity in time was assumed implicitly, as variance spectra were derived from the spatial pattern assessed on a single occasion. These approaches therefore ignored the spatially and temporally dynamical nature of ecological systems. However, recent spatio-temporal models of ecological systems suggest the prevalence of nonlinear dynamics in nature (e.g. Ellner and Turchin 1995, Little et al. 1996, Pascual and Ellner 2000). Nonlinear population dynamics caused by density-dependent and other processes can lead to chaos and/or other oscillatory patterns, such as frequency-locking or quasiperiodicity (Turchin and Taylor 1992, Pascual and Ellner 2000), in both deterministic and

stochastic systems (May 1976, Ellner and Turchin 1995). For these and other reasons, previous approaches to identifying 'intermediate' spatial scales are inadequate for ecological application. Contemporary ecology is in need of a more comprehensive approach to detect the "characteristic length scale" (CLS) of a system while accounting for the presence of chaotic, or other oscillatory dynamics (Hastings et al. 1993, Sole and Bascompte 1995, Little et al. 1996, Pascual et al. 2001).

Keeling et al. (1997) and Pascual and Levin (1999) proposed techniques that detect CLSs in spatial models of simple 2-species systems that oscillate, such as predator-prey or host-parasite systems (Rand and Wilson 1995, Keeling et al. 1997, Pascual and Levin 1999, Wilson and Keeling 2000). These techniques identify CLSs by distinguishing structural change driven by deterministic properties of the system from demographic noise, based on the prediction method of attractor reconstruction. The underlying tenet is that if the scale used to sample an ecosystem is too small, observations are clouded by noise due to strong correlations among individual samples (Durrett and Levin 2000, Wilson and Keeling 2000). If the sampling scale is too large, the non-trivial dynamics will be averaged out because distant parts of the landscape begin to act independently (Keeling et al. 1997, Pascual and Levin 1999, Wilson and Keeling 2000). An intermediate scale, the CLS, is the ideal scale at which to sample a complex natural ecosystem to maximize the underlying deterministic signal (Wilson and Keeling 2000; see Fig 1.1 for illustration).

While promising, the techniques of Keeling et al. (1997) and Pascual and Levin (1999) have been applied only to a small number of very simple oscillatory model systems. To estimate the CLSs of natural ecological systems, the methods

must produce robust estimates of CLSs for highly complex systems. Ideally, a robust measure would produce similar estimates of CLSs for different states of the same system as, for example, might arise from observing the system over different time periods, or from dissimilar spatial arrangement of recruits on an initial landscape. Thus, our aims are: 1) to determine whether the techniques of Keeling et al. (1997) and Pascual and Levin (1999) indicate unambiguous length scales for systems of a range of complexity; 2) to assess the sensitivity of CLS estimates to changes in the initial spatial arrangements of individuals; 3) to distinguish whether an ecological system has a single CLS, or several CLSs which depend on the identity of the species investigated; and 4) to analyze the robustness of both methods of CLS estimation to parameter choices in attractor reconstruction (Ellner 1989, Turchin and Taylor 1992). Using spatial models of increasing complexity, we compare the two approaches with respect to each aim.

2.3 Methods to determine CLSs of dynamic oscillating systems

Both the methods of Keeling et al. (1997) and Pascual and Levin (1999) employ the technique of sampling windows of increasing size across a landscape and measuring the temporal variance in abundance or density as a function of window size (L). Assuming there is a deterministic system underlying the dynamics, observations of a single variable in the system over time is used as a proxy to illustrate the dynamics of the whole system. For example, in a predator-prey system, the prey (or resource) density can be measured to illustrate the deterministic signal of the system (Rand and Wilson 1995). Both Keeling et al. (1997) and Pascual and Levin (1999) use the attractor reconstruction method of prediction from nonlinear time series analysis (Takens 1981, Kaplan and Glass

1995, Little et al. 1996; Appendix A). In summary, the phase space of the ecosystem dynamic is built from time delay coordinates of a single species, which act as surrogates for the unobserved variables of the system (Casdagli 1989, Abarbanel 1996, Kantz and Schreiber 1997). The trajectories in phase space are then used to construct the finite-dimensional attractor in a manner that preserves the topological properties of the original state space (Farmer 1982, Sugihara et al. 1990, Kaplan and Glass 1995, Abarbanel 1996).

While both methods use attractor reconstruction to estimate CLSs, the approaches differ in their measures of prediction error. Keeling et al. (1997) use fluctuation analysis to extend an earlier approach (Rand and Wilson 1995) to more complex ecosystems. Keeling et al. (1997) apply the attractor prediction methods to calculate deviations around the underlying deterministic behavior and then to plot error variance (termed 'error X') as a function of window size. For sufficiently large windows, the relative variance initially increases at a rate proportional to window size, and then plateaus. The window size at which variance reaches the plateau, or its first major point of inflection where very little increase in error is observed with an increase in scale, is the length scale where correlations have decayed and values at different sites behave like independent random variables. This scale, where the windows become statistically independent and the full spatial dynamics of the system can be observed, is identified as the CLS.

Pascual and Levin's (1999) method is a variant of the approach of Keeling et al. (1997) and aims to extract the scale where the determinism to noise ratio is maximized. This scale can be smaller than that required for the onset of independence defined by Keeling et al. (1997), as this CLS is the minimum

window size where the dynamics of the system can be accurately predicted

(Pascual and Levin 1999). At each window size, the degree of determinism is

evaluated from the prediction accuracy of the algorithm derived from attractor

reconstruction. Pascual and Levin (1999) then examine how the prediction error

of the trajectories changes with spatial scale (Kaplan and Glass 1995). They

obtain a statistic (termed the prediction r^2, or degree of determinism) for each

window size and the scale that emerges with maximum determinism is the CLS of

the system. Thus, this scale occurs where the prediction r^2 first attains a plateau

with respect to window size.

2.3.1 Models

Models of varying complexity were developed using the Compete©

software (Johnson and Seinen 2002), which is a probabilistic individual-based

system to model spatial competition between sessile colonial organisms. The

models follow the fates of competing individuals in a 2-dimensional landscape,

and can demonstrate complex behaviors indicating nonlinear dynamics and spatial

self-organization (Bascompte and Sole 1995, Johnson 1997, Johnson and Seinen

2002). Any network topology among S-species is possible, including intransitive

loops (e.g., where $S_1 > S_2$, $S_2 > S_3$, $S_3 > S_1$; with $S_x > S_y$ indicating that species x

outcompetes and displaces species y), which arise in benthic marine systems

(Johnson and Seinen 2002). We used models with intransitivities in their network

topology because they enable persistence stability (*sensu* Johnson and Mann

1988) of the system without the need for elaborate model closures and forcing

functions. We update the system synchronously, and use periodic (torodial)

boundary conditions.

Five model systems were implemented (in order of complexity of spatial pattern): symmetric networks of 3 species, 6 species (two different network topologies), and 12 species, and a model of the dynamics of the benthos of a coral reef (Fig 2.1). The 3-species system is the simplest intransitive loop as described above, i.e., a circular network with binary interaction outcomes, so that each interaction has one winner and one loser (Fig 2.1a). The first 6-species system (denoted 6(1)) involves a network in which each species overgrows one other species, in a similar structure to the 3-species loop ($S_1 > S_2$, $S_2 > S_3$,...,$S_6 > S_1$), with all other interactions as standoffs. This system forms a self-organized pattern of incipient spirals (Fig 2.1b). The second, more complex, 6-species system (denoted 6(2)) involves a symmetrical network in which each species overgrows and is overgrown by two species ($S_1 > (S_2, S_3)$, $S_2 > (S_3, S_4)$,...,$S_6 > (S_1, S_2)$), again with all other interactions as standoffs. In this system, the species spatially organize into two distinct groups of 3 species, and, if the model is run for sufficient time, either group may eventually dominate, depending on the landscape dimensions (Fig 2.1c). The 12-species system has a network structure of ($S_1 > (S_2, S_3, S_4)$, $S_2 > (S_3, S_4, S_5)$,...,$S_{12} > (S_1, S_2, S_3)$) and organizes into three groups of four species, any of which may begin to dominate, depending on landscape size and the number of generations (Fig 2.1d). In all four models the growth rates of all species are identical. For simplicity, there was no disturbance or mortality, and no recruitment of propagules to unoccupied sites. For the 6(2) and 12 species models, two scales of self-organization are emergent: that of the colony and that of the patch, which is a distinctive group of colonies of several species.

The coral reef model is based on parameters describing the recruitment,

mortality, neighbor-specific growth rates, and outcomes of interactions between

13 benthic phsyiognomic groups (Fig 2.1e). Parameters were estimated from

observation of a natural reef assemblage on the Great Barrier Reef over 3 years.

a)

b)

c)

d)

e)
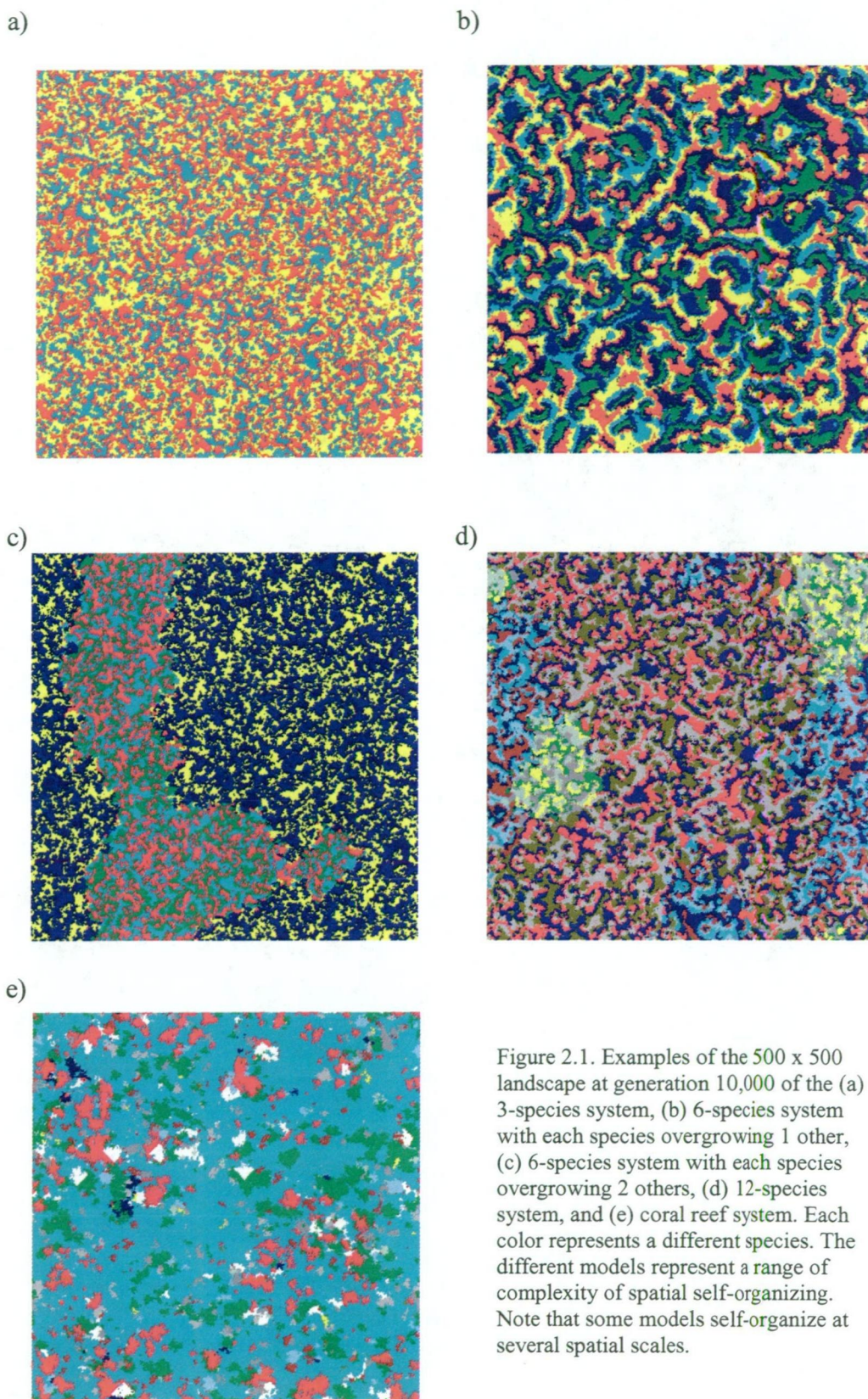
Figure 2.1. Examples of the 500 x 500 landscape at generation 10,000 of the (a) 3-species system, (b) 6-species system with each species overgrowing 1 other, (c) 6-species system with each species overgrowing 2 others, (d) 12-species system, and (e) coral reef system. Each color represents a different species. The different models represent a range of complexity of spatial self-organizing. Note that some models self-organize at several spatial scales.

The model accurately predicts the global community structure of the reef from

local processes (C. Johnson, *in prep*) and, in this context, is the model that most

closely simulates a natural system.

2.3.2 Attractor reconstruction using nonlinear time series analysis

We ran all five model systems for 10,000 generations on landscapes of 500

x 500 cells, and sampled landscapes from generation 201 to 10,000 (the first 200

time steps are ignored while the system self-structures). For each generation, we

sampled different window sizes L (5 to 495 in steps of 5) within the 500 x 500

landscape. We determined the density of each species for each window size L x L,

generating a separate time series for each L. For a selected species, the attractor

of the system in d dimensional space was estimated for each time series. Using

the Takens (1981) theorem for attractor reconstruction, we construct from the

original time series $x(t_i)$ time series vectors of dimension d_e,

$$X(t_i) = \{x(t_i), x(t_i + \tau), x(t_i + 2\tau), ..., x(t_i + (d_e - 1)\tau)\}$$

where $X(t_i)$ is the observable state variable at discrete time (t_i), τ is the time delay,

and d_e is the embedding dimension. These points are then assumed to

approximate the reconstructed attractor.

The time delay (τ) is some multiple of the sampling time, describing how

lagged in time the coordinates of the attractor will be (Abarbanel 1996). Here, τ

was chosen using the first minimum point in the time delayed mutual information

(MI; see Abarbanel 1996, Kantz and Schreiber 1997, Nichols and Nichols 2001).

This technique measures the amount of information shared between two measurements a and b. When the amount of information learned from a about b is at a minimum, the two time points are taken to be sufficiently independent (Abarbanel 1996). Of several possible techniques, we considered MI as the preferred approach to determine the time delay because it takes into account nonlinear dynamical correlations (Liebert and Schuster 1989). The alternative method for choosing the delay with an autocorrelation function was used by Pascual and Levin (1999), but has been criticized for being based on linear statistics (Nichols and Nichols 2001).

The embedding dimension (d_e) is the minimum dimension in phase space needed to capture the system dynamics (Farmer 1982). To estimate d_e, we used the false nearest neighbors approach (Kantz and Schreiber 1997). Each point on the attractor has some nearest neighbor in Euclidean space. False projections occur when the attractor is embedded in too few dimensions, and therefore is folded. The embedding dimension required to unfold the attractor is estimated to be where the number of false nearest neighbors drops below the level of noise and each point's nearest neighbors remain the same when the embedding dimension is increased (Liebert et al. 1991, Nichols and Nichols 2001).

Once τ and d_e are chosen, the prediction of the attractor must be made. Nearby points on the attractor are followed to determine their location after some t time steps (Sugihara et al. 1990). To predict $x(t_i + \tau)$ from $x(t_i)$, a list of all the states of x previously visited is searched for those closest to $x(t_i)$. If the time series is long enough, then some past states will be close to the present and the prediction will be close to the true state of the system (Kantz and Schreiber 1997). Some number of points (k) around the point we are trying to predict is used for

the prediction. These points, or nearest neighbors, are chosen based on their proximity, and then are averaged to determine the prediction in phase space (Kantz and Schreiber 1997). Our approach was to choose a fixed number of k nearest neighbors (neighborhoods of *fixed mass*) and then weight the average of the neighbors by inverse distance (Schreiber 1995). The predicted value of each point is then the average of the observed values of its k nearest neighbors, and the average is weighted by inverse distance so that neighbors further away contribute less (Casdagli 1989).

2.3.3 Determining the robustness of CLS estimates

To be robust measures, both prediction r^2 (Pascual and Levin 1999) and error X (Keeling et al. 1997) should indicate no dependence on window size when there is no deterministic signal in the dynamics of the system, and no oscillatory or self-organizing behavior. In such systems, the spatial and temporal patterns are completely random. This result has been shown in Appendix B.

To determine whether interpretable length scales can be obtained for systems of a range of complexity, we derived both the Keeling et al. (1997) and Pascual and Levin (1999) CLS estimates for each of the five models. Several aspects of robustness were considered: 1) capacity to identify similar length scales for different runs of the same system, 2) capacity to identify similar length scales from different species in the same system, and 3) for any one model run, sensitivity of the CLS estimate to different choices of parameters required for attractor reconstruction, namely τ, d_e, and k.

2.3.3.1 Robustness to initial conditions

Each model was begun with a random spatial arrangement of 'recruits' (10% total cover) on the initial landscape, with identical amounts of each species. For each model system, we assessed the variability of the CLS estimated from the dynamics of a given species over 100 runs of the model, each with a different initial random configuration.

2.3.3.2 Robustness to choice of species

A system's attractor is built with observations on a single species. In theory, the choice of species to reconstruct the entire system's attractor is arbitrary, as every species reflects the same underlying attractor (Abarbanel 1996). Previously, CLSs have been generated from only a single species within the system, implicitly assuming that all species will indicate the same CLS (Rand and Wilson 1995, Keeling et al. 1997, Pascual and Levin 1999). We tested this assumption by determining the CLS of each model system using different species in the system.

2.3.3.3 Robustness to parameters of attractor reconstruction

The estimated CLS will depend in part on the accuracy of the attractor reconstruction, which itself depends on appropriate choices of the reconstruction parameters (τ, d_e, and k) (Buzug and Pfister 1992, Kantz and Schreiber 1997). No single unambiguous value exists for any of these parameters for a particular reconstruction, and indeed different techniques commonly yield different estimates (Buzug and Pfister 1992, Schreiber 1995, Kantz and Schreiber 1997, Schreiber 1999). Thus, we examined the sensitivity of the estimated CLS to a

range of sensible potential choices of these parameters.

(i) Time delay (τ)

From a mathematical perspective, the choice of delay is arbitrary because the data set is assumed to be infinitely precise (Kantz and Schreiber 1997). However, for a finite set of data, the choice of τ dictates the quality of the reconstructed trajectory (Liebert and Schuster 1989). If τ is too small, the coordinates $x(t_i)$ and $x(t_i + \tau)$ will be almost identical, offering redundant information about the state space. Alternatively, if τ is too large, the coordinates will be almost uncorrelated and therefore their connection to one another is no different from random (Abarbanel 1996). The goal is to determine the delay where coordinates are independent while preserving their dynamical relationship (Nichols and Nichols 2001). However, because identifying the first minimum point from the plot of MI vs τ is somewhat subjective, we assessed the robustness of CLS estimates for a variety of choices of τ that might be considered reasonable using the MI approach.

(ii) Embedding dimension (d_e)

If an attractor is projected in too few dimensions, the observed orbits will overlap themselves and distinct segments on the attractor become confused (Abarbanel 1996). The appropriate d_e allows the attractor to be sufficiently unfolded in space such that this overlap no longer occurs. Over-embedding (embedding in too many dimensions) requires larger numbers of coordinates, and may enhance the possibility for noise in the dimensions of the embedded space where no dynamics are operating (Kennel et al. 1992). The d_e is chosen as the smallest dimension required to sufficiently unfold the attractor, and is indicated as

the first minimum of the false nearest neighbors *versus* dimension curve. We assessed the robustness of a CLS for a range of embedding dimensions around the value suggested by the false nearest neighbor method.

(iii) Number of nearest neighbors (k)

The number of nearest neighbors (k) is a tunable parameter that influences the quality of the prediction. If too few neighbors are picked, then important non-random information may be missed. If too many are picked, the points may be widely spread in space, decreasing the accuracy of the prediction (Kantz and Schreiber 1997). We assessed the sensitivity of CLS estimates over a range of reasonable choices of k.

2.3.4 Estimating the CLSs

In calculating the CLSs, abundances of the selected species were scaled by window area to convert each time series of absolute counts to a time series of species density. Attractor reconstruction was undertaken as described above for particular choices of τ, d_e, and k ($k = 10$ unless otherwise specified). A small amount of noise (measurement error with a given standard deviation) was added to the density values, because, with the discrete nature of the cellular models, a point can be identical to its nearest neighbor, producing a zero in the denominator of the inverse distance weighting. The measures of Keeling et al. (1997) and Pascual and Levin (1999) were computed, and error X and prediction r^2 spectra were plotted respectively. The CLS was estimated as the approximate window size, L, at which the spectrum reached a plateau, or where the slope exhibited an abrupt shallowing. Ninety-five percent confidence intervals were calculated from the mean curve of 100 independent runs of each model.

2.4 Results

For each model system, the CLS determined using the two methods is compared. We indicate the length scale determined using the method of Keeling et al. (1997) as CLS_k, while that of Pascual and Levin (1999) as CLS_p.

2.4.1 Robustness of CLS estimates to initial conditions

Using the simplest model system (3 species), the two methods produced curves of the expected shape with both error X and prediction r^2 increasing to a plateau as a function of window length, therefore demonstrating a single length scale (Fig 2.2a). As expected, CLS_k (ca. 60 cells) was slightly larger than CLS_p (ca. 40 cells; Pascual and Levin 1999), demonstrating the difference between the two methods. There was little variation between runs with different initial conditions. However, with increasing complexity of the model system, CLS_k became more difficult to determine and confidence intervals around the error X spectrum broadened (Fig 2.2). The error X curves for the 6(1) and 12-species systems were not the expected positive asymptotic shape, and were not readily interpretable (Fig 2.2b, d). Similarly, based on error X, no length scale could be determined for the coral reef model because the curve deviated dramatically from the expected shape with no inflection or plateau (Fig 2.2e). Conversely, CLS_p was not highly sensitive to changes in initial conditions for any of the models, and produced an interpretable curve for the full range of models we examined (Fig 2.2).
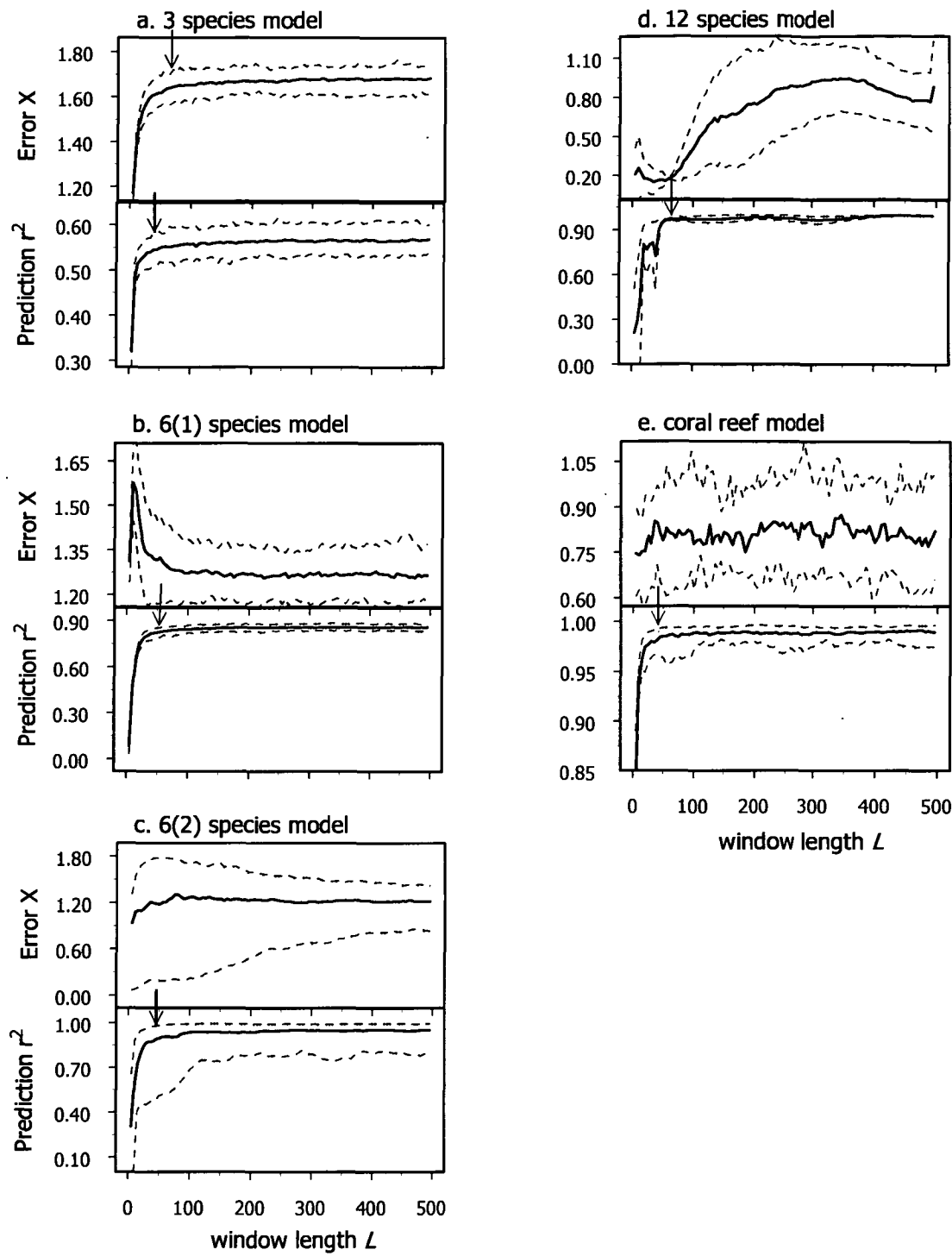
Figure 2.2. Average values (solid lines) of error X (Keeling et al. 1997) and prediction r^2 (Pascual and Levin 1999) and the 95% confidence intervals (dotted lines) for a single species, calculated from a Monte Carlo of 100 independent runs for (a) 3-species model, (b) 6(1)-species model, (c) 6(2)-species model, (d) 12-species model, and (e) coral reef model. CLS estimates are indicated by arrows where curves were interpretable.

2.4.2 Robustness of CLS estimates to choice of species

Estimates of CLS_k and CLS_p in the 3-species (Fig 2.3) and 6(1) model

systems were not dependent on species identity. In contrast, curves for the more

complex model systems were often highly sensitive to the choice of species (Figs

2.4 – 2.6). CLS_k depended heavily on the species used in the attractor

reconstruction for all three complex model systems. For example, in the 12-

species system, CLS_k ranged from 50 cells for one species to 250 cells for another

(Fig 2.5a). For error X, species in the same self-organized patch demonstrated

similar curves (Figs 2.4a, 2.5a,c). However, not only were error X curves

dissimilar for species from different patches in the same run (compare Figs 2.4a,

2.5a,c), but the overall shape of the curve, and therefore the CLS_k indicated,

changed markedly among runs for the same species. For example, one species in

the 12-species system had a CLS_k of approximately 80 cells in one run (Fig 2.5a),

but 300 cells in another (Fig 2.5c). The error X curves from the coral reef model

displayed little variation from run to run (Fig 2.6a versus 2.6c), but were sensitive

to species, with only some interpretable spectra. It should be noted, however, that

when error X curves were interpretable (e.g., see arrow in Fig 2.6a), CLS_k was

larger than the corresponding CLS_p, as expected.

As model complexity increased, prediction r^2 curves generally separated

into groups that corresponded to species within the same spatially self-organized

patch, but CLS_p of the different groups of species were more similar than

indicated by the Keeling et al. (1997) method (e.g., Fig 2.4). In the two most

complex systems, CLS_p was more sensitive to species, with the estimate differing

by up to 40 cells between groups (Figs 2.5b,d). For the 6(2) and 12-species

systems, the curves of prediction r^2 for a given species demonstrated multiple

peaks (Figs 2.4b, 2.5b), indicating the presence of more than one length scale.
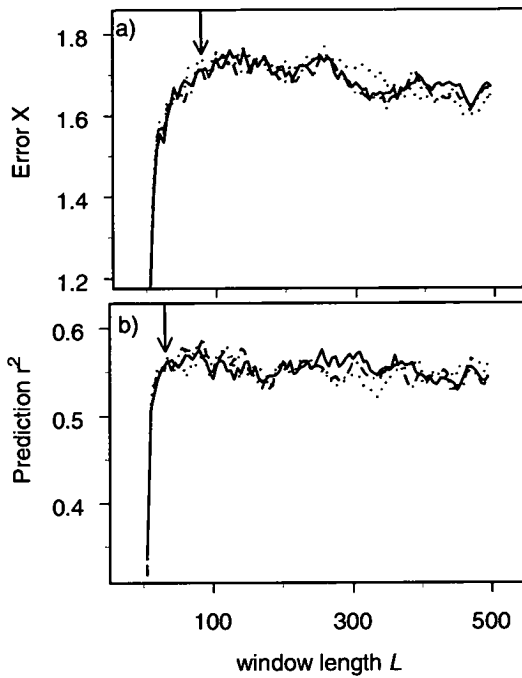
Figure 2.3. CLS curves for each species of the 3-species model system for the measures of (a) Keeling et al. (1997) and (b) Pascual and Levin (1999) methods. Data are for a single run of the model. Arrows indicate the CLS estimates.
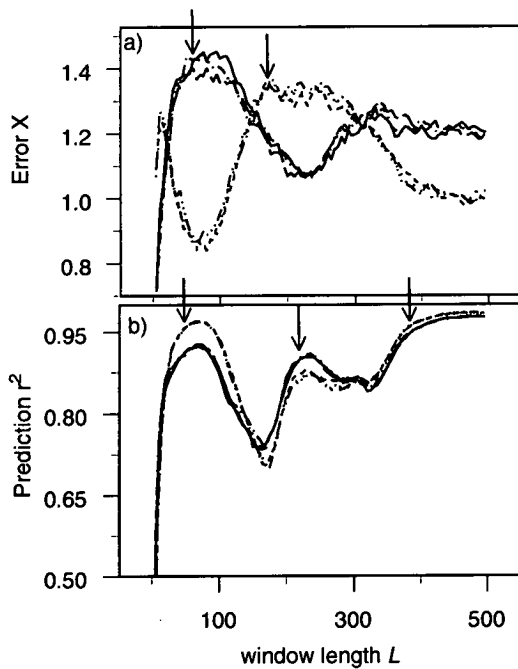
Figure 2.4. CLS curves for each species of the 6(2)-species model system for (a) Keeling et al. (1997) and (b) Pascual and Levin (1999) methods. Species spatially self-organize into two groups of three, which is reflected as two groups of curves on the graphs. Data are for a single run of the model. Potential CLS estimates are indicated by arrows where curves were interpretable.
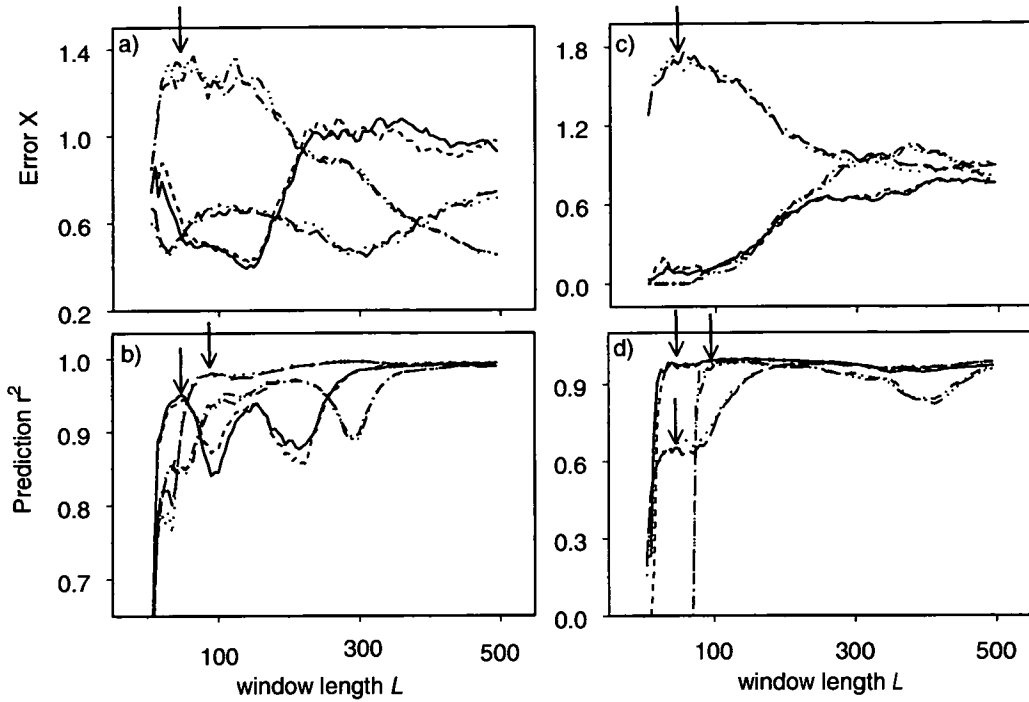
Figure 2.5. CLS curves for 6 species of the 12-species model system shown from two separate runs. For run 1: (a) Keeling et al. (1997) and (b) Pascual and Levin (1999) methods and run 2: (c) Keeling et al. (1997) and (d) Pascual and Levin (1999). Species organize into 3 groups of 4 on the landscape, and species in the same patch on the landscape show similar CLS curves. Primary CLS estimates are indicated by arrows where curves were interpretable. Multiple peaks in the curves (b) may be evidence of multiple length scales.

2.4.3 Robustness to choices of time delay in attractor reconstruction

Changes in time delay of the 3-species system shifted the CLS curves on the y-axis, but did not change the magnitude of the CLS estimate for either method (Fig 2.7). Similarly, time delay had little effect on estimates of the CLS interpreted from curves of the 6(1) and 6(2)-species systems. In the two most complex model systems, CLS_k was notably less robust, with the shape of the curve changing with delay. For the coral reef model in particular, the error X curves were ambiguous at best, and varied with the delays (Fig 2.8a). Conversely, CLS_p was robust to changes in delay in all of the model systems (Fig 2.7b – 2.8b).
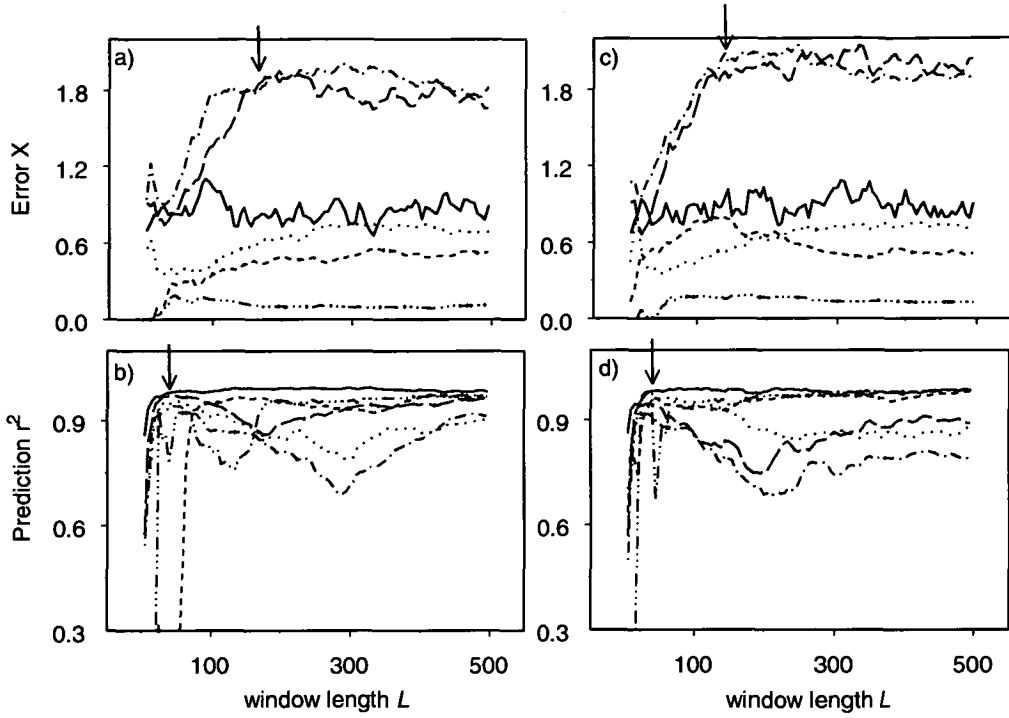
Figure 2.6. CLS curves for 6 species groups of the coral reef model system shown from two separate runs. For run 1: (a) Keeling et al. (1997) and (b) Pascual and Levin (1999) methods and run 2: (c) Keeling et al. (1997) and (d) Pascual and Levin (1999). CLS estimates are indicated by arrows where curves were interpretable.

2.4.4 Robustness to choices of embedding dimension in attractor reconstruction

Changes in embedding dimension of the 3-species and 6(2)-species systems shifted the curves on the y-axis, but had no effect on estimates of CLS_k or CLS_p. In the 6(1)-species system, the error X curve shifted from an inverted shape to the expected shape with the increase of dimension from 3 to 7 (Fig 2.9a), while CLS_p was robust to changes in dimension (Fig 2.9b). In the 12-species system, the shape of the error X curve remained robust to increasing dimension, but interpreting the curves was difficult (Fig 2.10a). For the coral reef system, CLS_k
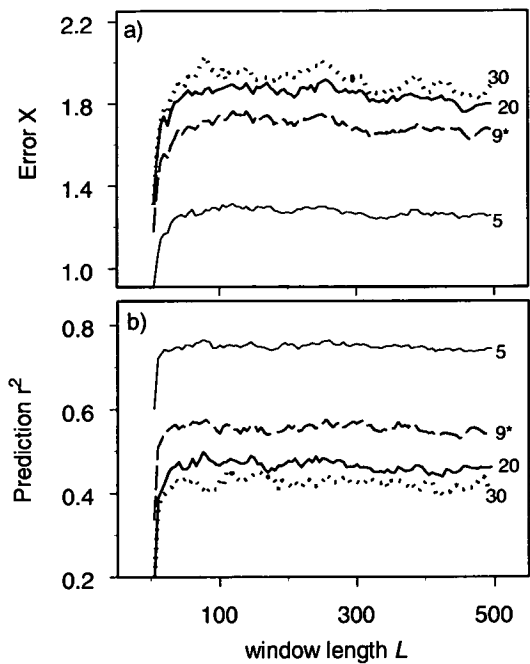
Figure 2.7. Robustness of CLS curves constructed from one species of the 3-species system as a function of time delay. Different lines denote the delays (shown to the right of each curve) used in attractor reconstruction with methods of (a) Keeling et al. (1997) and (b) Pascual and Levin (1999). * denotes the delay indicated using the mutual information method.
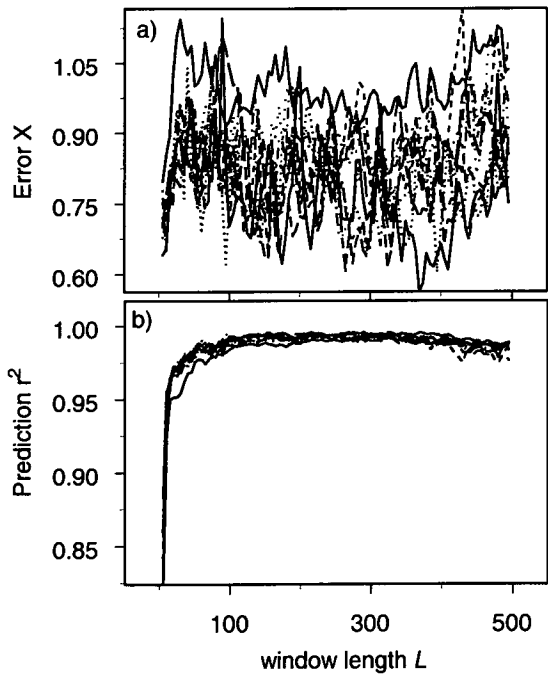
Figure 2.8. Robustness of CLS curves constructed from the dominant species group of the coral reef model system as a function of time delay. Different lines denote the delays used in attractor reconstruction from 20 to 200 in steps of 20 with the methods of (a) Keeling et al. (1997) and (b) Pascual and Levin (1999).

became interpretable only with overly large embedding dimension (Fig 2.11a). In both of these more complex systems, the estimates of CLS_p were reasonably robust to changing dimension (Figs 2.10b, 2.11b).
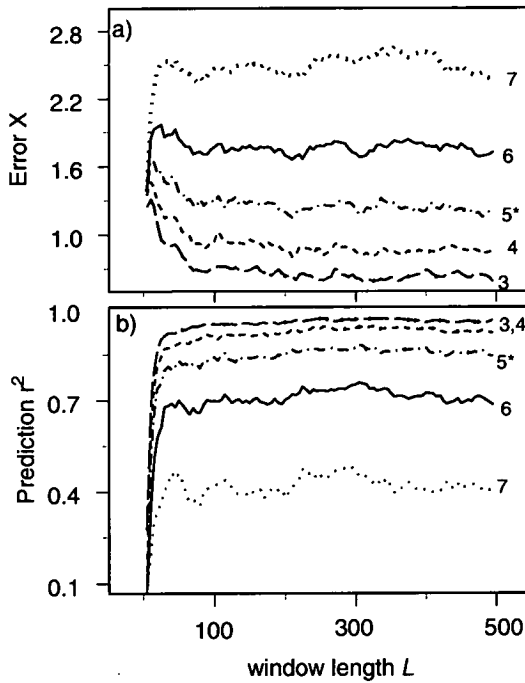
Figure 2.9. Robustness of CLS embedding dimension for the 6(1)-species system with the methods of (a) Keeling et al. (1997) and (b) Pascual and Levin (1999). * denotes the dimension indicated using the false nearest neighbors method.
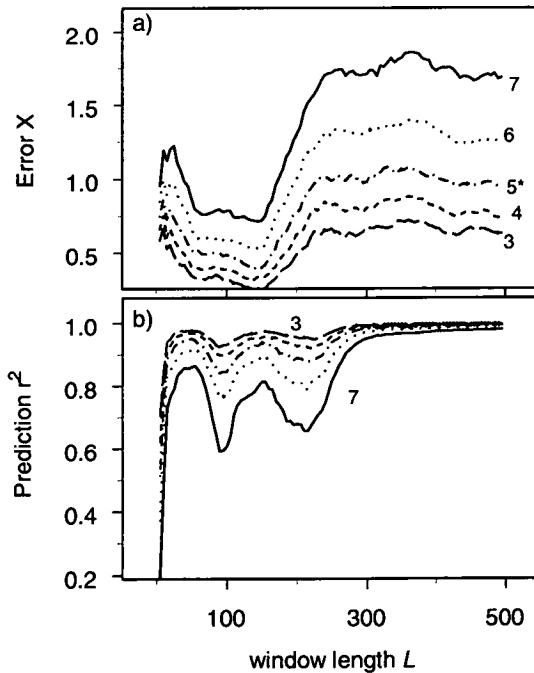
Figure 2.10. Robustness of CLS embedding dimension (shown to the right of the curves) for the 12-species system with the methods of (a) Keeling et al. (1997) and (b) Pascual and Levin (1999). * denotes the dimension indicated using the false nearest neighbors method.

2.4.5 Robustness to choices of k nearest neighbors in attractor reconstruction

The number of k nearest neighbors used in reconstruction of the attractor of

the 3-species system had no effect on estimates of CLS_k or CLS_p. However, with

the more complex systems, the error X curves shifted from an inverted shape with

low numbers of neighbors to a curve of the expected shape with excessively high

numbers of neighbors (for example, Fig 2.12a). CLS_p for the same systems was

robust to varying numbers of neighbors in attractor reconstruction (Fig 2.12b).
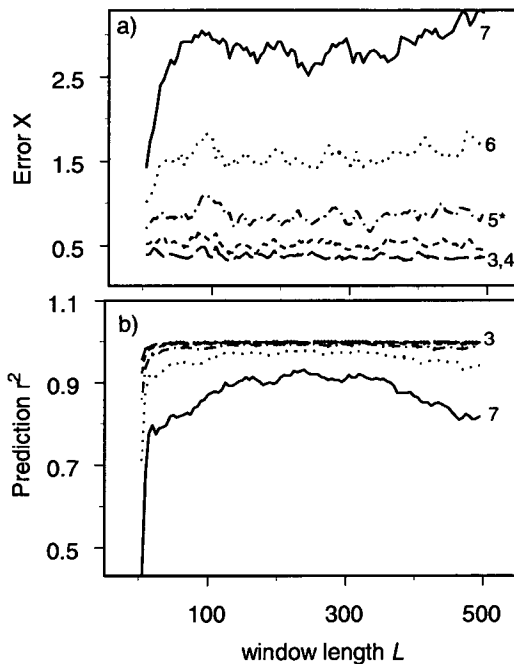
Figure 2.11. Robustness of CLS embedding dimension (shown to the right of the curves) for the coral reef model system with the methods of (a) Keeling et al. (1997) and (b) Pascual and Levin (1999). * denotes the dimension indicated using the false nearest neighbors method.

2.5 Discussion

2.5.1 General

The issue of spatial scale is a central theme in ecology (Levin 1992, Levin

et al. 1997, Tyre et al. 1997, Levin 2000, Wilson and Keeling 2000, Molofsky et

al. 2002). While investigators have acknowledged the need to address questions at
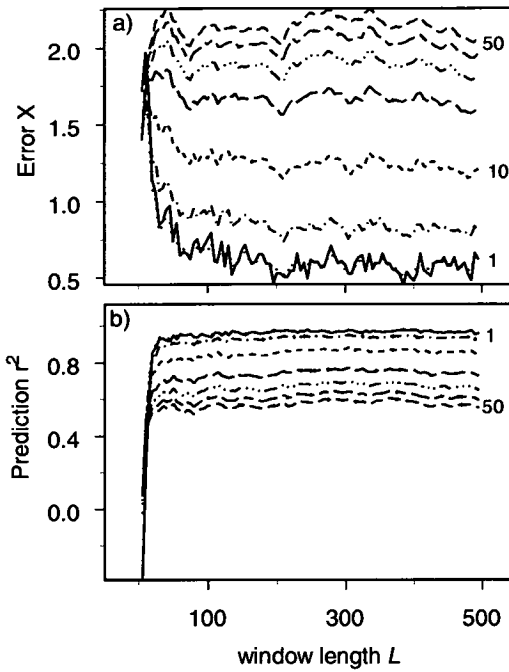
Figure 2.12. Robustness of CLS curves for one species of the 6(1)-species model system shown for the (a) Keeling et al. (1997) and (b) Pascual and Levin (1999) methods to increasing numbers of k nearest neighbors from 1 to 3 to 5, and then from 10 to 50 in increments of 10.

specific and appropriate scales (Carlile et al. 1989, Wiens 1989, De Roos et al. 1991), how this "appropriate scale" is identified has often been ambiguous. Recent application of methods from nonlinear time series analysis has refined a crucial aspect of the study of scale in ecology, allowing a shift in focus from observing mean behaviors to extracting the deterministic signal from dynamical systems (Rand and Wilson 1995, Keeling et al. 1997, Pascual and Levin 1999). Towards an ultimate goal of estimating CLSs of natural systems, here we evaluate the robustness of recent methods used to estimate CLSs.

Both the methods of Keeling et al. (1997) and Pascual and Levin (1999) maintained high robustness and low sensitivity to parameter choices in the simplest model systems. CLS_k was consistently larger than CLS_p, indicating the need for larger windows to detect the overall spatial dynamic. However, as the complexity of the model systems increased, the error X curves of Keeling et al.

(1997) became more difficult to interpret. With increasing model complexity, error X and therefore CLS_k, were increasingly sensitive to changes in initial conditions, to the species of focus, and to values of parameters used in attractor reconstruction. Error X curves for the 6(1)-species, 12-species and coral reef model systems deviated from the expected shape so that interpretations were ambiguous, clouding any attempt to assess robustness (Fig 2.8a).

The behavior of the error X curves for the 6(1)-species system was of particular interest. With low embedding dimension, the curve was the inversion of the expected shape (Fig 2.9). However, when the dimension was increased to 6 and above, the curve changed to the expected shape, with a decreasing positive slope leading to a plateau (Fig 2.9). Similarly, the shape of the error X curve inverted from negative to positive slope between 10 and 20 k nearest neighbors (Fig 2.12). The meaning of these inversions is unclear. However, because the inflection points of the curves remain consistent whether their slopes increase or decrease, they may be tentatively regarded as the system's CLS.

In contrast, the Pascual and Levin (1999) method provided an interpretable measure of the CLS with increasing model complexity. Prediction r^2 of all model systems maintained an unambiguous shape with changes in attractor reconstruction parameters, and with varying initial conditions. Thus, evaluation of prediction r^2 using the approach of Pascual and Levin (1999) appears to be the more robust means of estimating CLSs of complex ecological systems.

The reason for the ambiguity and, in some cases, failure (Fig 2.8a) of CLS_k is unclear. Error X may be more susceptible to random noise than prediction r^2. Alternatively, error X may be simply more sensitive to a system's complexity than prediction r^2. This raises the question of how much data would be required

for reliable CLS estimates to be obtained using the error X measure. Further research may demonstrate that the sensitivity of error X reveals useful spatial information for other applications, such as identifying features of the spatial pattern, but we conclude that it is inadequate for estimating CLSs of complex systems.

2.5.2 Do different species indicate different length scales?

Estimates of CLS_p showed dependence on the species used in attractor reconstruction in the more complex models. Differences among species were most notable in systems that were strongly spatially self-organizing, but these differences were not related to abundances. In the 12-species model, the curves for the species form three distinct groups, each with a different CLS_p (Fig 2.5). While all species in this model are topologically equivalent with respect to network structure, the system spatially self-organizes into 3 distinct patch types, each with 4 species. Because species within patches are more likely to interact with each other than with species from other patches, actual connectivity is higher among species within patches than it is among species between patches, despite topological symmetry in the interaction network.

Not surprisingly, highly connected species with tightly coupled dynamics (i.e., within patches) indicate similar length scales, while species whose dynamics are more weakly linked (because they are spatially separated in different patches) can manifest dissimilar length scales, even though they have identical 'life history' types. In our model examples (with the exception of the coral reef system), differential connectivity among species arises through spatial self-organizing. We anticipate that other factors which influence species' connectivity,

such as the topology of food webs in which some groups of species are tightly coupled trophically while others manifest low connectivity (O'Neill et al. 1986, Johnson et al. 1995), will have a similar effect. Because the approach is based on reconstruction of deterministic dynamics, species that are poorly linked dynamically can indicate different length scales for their different behaviors. Thus, the spatially separated and therefore loosely connected species in the 12-species model produce different estimates of CLS_p (Fig 2.5), while the highly (and equivalently) connected species in the 3 and 6(1)-species models produce almost identical estimates of CLS_p (Fig 2.3). In the coral reef model, different species provide similar but slightly different CLS_p estimates (Fig 2.6). Despite notable differences among species in their recruitment, growth, and mortality, marked differences in CLS_p among species in this model do not arise because species do not self-organize into distinct patches. The system is maximally connected with each species competing with all others for space (Johnson and Seinen 2002). However, the species are not equally connected because they have different interaction strengths and neighbor specific growth rates, which may be the reason for the divergence of one species in particular (Fig 2.6b). Because CLS estimates reflect the strength of dynamical connectivity among species, we predict that for complex real systems, different species or functional groups which are loosely connected may indicate dissimilar length scales (see Chapter 6).

2.5.3 Multiple length scales and spatial pattern

For complex model systems, the technique of Pascual and Levin (1999) may detect several different length scales within the system. For example, the prediction r^2 curve of a single group in the 6(2)-system (Fig 2.4b) displays several

critical points, which we interpret as multiple length scales. The smallest length scale, at approximately 60 cells (the first peak in the curve), remains consistent among runs. This 'primary scale' is the scale at which the local dynamic is best predicted, and it reflects the scale of interaction between colonies of different species within patches. Scales larger than the primary scale likely reflect the system's emergent dynamics resulting from the interactions between spatially distinct patches of species, which vary between runs. Note that while any single run often demonstrates several length scales, the average prediction r^2 curve for a given species in the 6(2)-species model indicated only a single CLS, likely due to variations in its shape between runs (Fig 2.2c). The widening of the 95% confidence intervals around the average curve suggests that the second peaks are absorbed as noise (Fig 2.2c). We suggest that the secondary peaks that we interpret to indicate larger length scales are more variable (e.g., Fig 2.5b versus 2.5d) than the primary CLS because the higher level emergent dynamics of the system are unpredictable, with the shapes and sizes of patches changing between runs. Thus, for any one species, we might expect two CLSs reflecting:

1.) interactions within patches, influenced by individual species abundances, colony size, and the nature of local interactions between species within patches. We expect this CLS to be larger than the mean colony size but smaller than the patch size. For example, for the 6(2)-system on a 500 x 500 landscape, the size of individual colonies is approximately 20 x 20 cells, while the size of the patches of 3 species is about 200 x 200 cells. The primary CLS_p of this system is approximately 60 cells (Fig 2.4b).

2.) the emergent dynamics among patches. This CLS is larger than the

mean patch size. In the 6(2)-system, the scales larger than the primary CLS_p are greater than 200 cells (Fig 2.4b).

2.5.4 Do CLSs have a future in ecology?

While research of the robustness of methods to define CLSs is an important step towards their application to natural systems, several conceptual and other theoretical questions arise. Does a system have a single length scale or are there many? Our results indicate that different species can display dissimilar CLSs if their dynamics are weakly linked, and that the same species may also display several length scales if its abundance is determined by both local interactions and emergent behavior at larger scales. This reflects that in natural systems (and, by inference, complex model systems) some species interactions will function on different scales (Levin 2000, Bishop et al. 2002). Teasing apart the factors that underpin different length scales in dynamic ecological systems will improve the potential for application of this technique to real ecosystems. Regardless, the use of nonlinear analysis to determine the length scale or scales of a system will allow more objective scaling decisions to be made.

That a single species may demonstrate several length scales has important implications for the application of these techniques to applied scaling issues in ecology. For our model systems, where more than one length scale was indicated by the prediction r^2 spectra, the smallest or primary length scale was unambiguous and consistent between replicate runs. This is the optimum scale at which behaviors, as a result of local deterministic dynamics can be observed, and would be a useful scale for monitoring particular components such as key indicator species. The larger length scales were elusive, indicative of the variable

manifestation of higher-level emergent behaviors among runs (Bascompte and Sole 1995). When CLSs larger than the primary CLS were apparent for a given system, the magnitudes fell within a consistent range between runs. These larger CLSs may prove to be useful towards objectively defining minimum sizes for conservation areas and management units in harvested systems.

While the robustness of the Pascual and Levin (1999) approach indicates it is promising in assessing characteristic length scales in natural systems, the analysis requires spatially resolved data over long time series (thousands of generations), which are difficult, if not impossible, to obtain for natural ecological systems. The imperative is to modify this method to reduce data requirements while maintaining the integrity of the estimate. With new developments of the Pascual and Levin (1999) approach that require much smaller quantities of data (Trebilco honours thesis 2002, Habeeb et al. 2005 (see Appendix C)), estimating CLSs of natural ecological systems now appears feasible. Our next focus will be to assess what CLSs inform us about natural systems, and to evaluate their utility in providing objective estimates for scaling issues in applied ecology.

Chapter 3

Natural scales of real ecosystems: Choosing optimal spatial scales for observation

.

3.1 Abstract

Choosing an appropriate spatial scale to observe an ecosystem is a universal challenge for ecologists, largely because the scale used to view the system strongly influences the patterns and processes that are observed. Characteristic length scales (CLSs) are the scales at which the non-trivial deterministic signal of a system's temporal dynamic can be best observed, while taking into account stochastic demographic fluctuations and inherent nonlinear behaviors in the system dynamic. However, thus far, CLSs have been identified only for model systems. Here, for the first time, we apply a modified technique to detect CLSs of a natural system. This new technique requires much less information of the temporal dynamics than does the original approach used on model systems. We created a short time series of three digital maps, each three months apart, for a marine fouling community in Tasmania, Australia, and analyzed these maps for the CLSs. The new technique provided length scales for ten different species groups, indicating the adequacy of this method for objectively determining optimal scales of observation. For nine out of the ten species groups, the primary CLS was $\approx 0.30 - 0.40$ m. Species from several phyla with dissimilar life histories provided remarkably similar length scales, confirming the system-level nature of the estimate. To assess whether CLS estimates are sensitive to particular site-

specific characteristics, we compared these field-based results to those derived from a spatially explicit individual-based model of a similar, nearby fouling community. The average CLS of the model community was strikingly similar (0.35 – 0.45 m) to that of the natural community, suggesting that dynamical trends of like systems may be best observed on similar scales. Secondary length scales, larger than the initial CLS and indicative of higher-level emergent properties of the system, were indicated by several species in the natural system but not in the model system. These scales reflect the more complex emergent dynamics of the system and so are less likely to be identified than the primary length scales. This new technique may allow definition of optimal scales of observation not only in ecological experiments, but also in monitoring studies for conservation, natural resource management, and environmental impact assessment.

3.2 Introduction

Choosing an appropriate spatial scale of observation is a universal challenge for ecologists. Indeed, the spatial scale at which a system should be measured has long been identified as a "central problem in ecology" (Levin 1992). Measures of landscape diversity and pattern, for example, were highly sensitive to the size of sampling units, even when the variation in scale was much less than an order of magnitude (Turner et al 1989, Qi and Wu 1996). How, then, should the sampling scale be chosen? Most researchers assign the sampling scale intuitively, often depending on how the species or community in question is perceived, or the sampling scale is set largely by logistic constraints (Levin 1992, Denny et al. 2004). However, if the sampling scale is inappropriate for the ecological system

being studied, the observed variability may be misleading (Wiens 1989). The effect of the chosen spatial scale on patterns and processes observed is now well documented (e.g., Levin 1992 and references therein, Bissonette 1997, Chave and Levin 2003), so that a widespread dilemma is how to choose objectively a sampling scale or 'window size' that best detects biologically meaningful trends while minimizing noise (Wiens 1989, Carlile et al. 1989, Tyre et al. 1997, Keeling et al. 1997, Wilson and Keeling 2000). Thus, a key goal in ecology is to objectively determine optimal scales of observation while accounting for the complexities of nonlinear dynamics and spatial variability (Keeling et al. 1997, Pascual and Levin 1999, Wilson and Keeling 2000, Lundquist and Sommerfeld 2002, Petrovskii et al. 2003). The issue of observational scale is relevant to levels of biological organization from populations of particular species to the dynamics of entire ecosystems (Pascual and Levin 1999).

Recent studies have shown that simple models of ecological systems that exhibit demographic stochasticity and inherent nonlinear behaviors demonstrate a natural spatial scale at which the non-trivial deterministic signal of the system's temporal dynamic can be best observed (Pascual and Levin 1999). This spatial scale is termed the natural or characteristic length scale (CLS) of the system. An estimate of the characteristic length scale is derived from the nonlinear time series technique of attractor reconstruction (Takens 1981) applied to highly spatially resolved density data for a single species in the system. Thus, this system-level measure of scale is obtained by reconstructing the dynamic of the entire system in phase space based on information of the dynamics at the species level. In theory, at the CLS, deterministic information is maximized with respect to stochastic fluctuations (Rand and Wilson 1995), thereby providing an estimate of the

optimal scale at which to observe the ecosystem.

Characteristic length scales based on the method used in Pascual and Levin (1999) have been shown to be robust to parameters chosen in attractor reconstruction, to the complexity of the model system used to simulate the data and, in most cases, to the species in the model chosen for the time series analysis (Chapter 2, Habeeb et al. 2005). This level of robustness indicates the potential for this kind of metric to be a widely useful tool in ecology. However, studies of characteristic length scales have until now been theoretical, based only on data generated from models (Rand and Wilson 1995, Keeling et al. 1997, Pascual and Levin 1999, Petrovskii et al. 2003, Habeeb et al. 2005). Hence, the technique generally relies on extremely long time series data (from models) unlikely to be available for any natural system (Marcos-Nikolaus et al. 2002). Habeeb et al. (2005) modified Pascual and Levin's (1999) method to derive CLSs from very short time series of highly spatially resolved 'maps' of species abundances (see Appendix C). Their approach uses both temporal and spatial information (as in Bascompte and Sole 1995), and largely substitutes replication in space for replication in time. This drastic reduction in data requirements enables examination of real ecosystems for CLSs.

Here we apply the modified CLS technique to a natural system for the first time, examining whether interpretable CLSs can be derived from observations of a marine subtidal epibenthic community. If natural systems display interpretable CLSs, then application of the technique should help to objectively determine optimal scales of observation not only for ecological experiments, but also for monitoring studies for conservation, natural resource management, and environmental impact assessment (Lewis et al. 1996, Pressey and Logan 1998,

Schwartz 1999). Takens' (1981) theory of attractor reconstruction suggests that the dynamical information of a single species within a multi-species system should act as a substitute for the unmeasured whole-system dynamic, therefore in this context, implying that each species will reflect the same CLS. However, ecologists might intuitively expect that different species with distinct behaviors and life history strategies may operate on different spatial scales. Therefore, we also compare the CLS indicated by 10 different species representing several phyla and a range of life histories.

To assess whether CLS estimates are sensitive to particular site-specific characteristics, we compare the results based on field data to those derived from a spatially explicit individual-based model of a similar fouling community (Dunstan and Johnson 2005). The model generates landscapes based on empirically determined parameters measured from a similar fouling community developed on a jetty wall, also on the east coast of Tasmania. Notably, the model accurately predicts the structure and dynamics of the real community (Dunstan and Johnson 2005). We determine the CLSs of 13 species within this model community and compare these estimates to those from the real community we studied.

Given the novelty of the CLS metric, we also attempt to place the measure in a broader ecological context by comparing it to other more established and better understood ecological scaling measures. This comparison should also help to illuminate what, if anything, the CLS might reveal about the broader ecological properties of a system. We chose conceptually simple spatial scales for comparison, namely the region of asymptote in species-area curves (Conner and McCoy 1979) and, with reference to patterns of species abundance, scales at which troughs and peaks are evident in the relationship between variance-to-mean

ratios and size of the sampling unit. Interestingly, the spatial scale of the

asymptote in species-area curves was initially proposed as the appropriate scale of

observation of communities, with a motivation to capture most species in each

sampling unit (Greig-Smith 1964). Variance-to-mean ratio spectra have been used

to describe patterns of scale dependent aggregation in species distributions

(Pielou 1977).

3.3 Methods

3.3.1 Estimating the CLS

Because sufficiently detailed spatial data are rarely available in a long time

series, Habeeb et al. (2005) modified the original nonlinear time series technique

as outlined by Pascual and Levin (1999) to enable its application to real

ecosystems (Appendix C). By largely substituting space for time, CLSs can be

estimated with a series of as few as three highly resolved landscapes through

time, rather than requiring tens of thousands of time steps (Habeeb et al. 2005).

This modified technique is based on the original 'prediction r^2' variance spectrum

proposed by Pascual and Levin (1999), but the method of attractor reconstruction

uses observations of short time series at many points in space rather than a single

long time series. We used this modified method to evaluate whether a natural

characteristic length scale emerges from a real system.

3.3.2 The empirical system

We chose a marine fouling community as the study system because it

supports a wide diversity of species/phyla, it has relatively rapid dynamics, and it

is spatially tractable in two dimensions (Dunstan and Johnson 2005).

Furthermore, it is a suitable 'model' where the whole system can be captured

accurately on a relatively small areal extent (Dunstan and Johnson 2004, 2005). A

3.0 m x 3.3 m permanent quadrat was established at 7 – 10 m depth on a vertical

rock wall at the entrance to Cathedral Cave in Waterfall Bay, Tasmania, Australia

(43°3'43S, 147°57'1E). This dolerite wall has low topographic complexity, and

virtually complete cover, predominantly of sessile invertebrates but with some

algae. The wall supported 140 species, but we selected for analysis ten distinct

species/functional groups, representing a diversity of abundances and life

histories (Fig 3.1). The species groups included five species of sponges

(specimens lodged with the Tasmanian Museum; K430, K431, K432, K433), one

group of arborescent bryozoans, an ascidian, a cnidarian, and two groups of algae

(Table 3.1), with an overall average colony size of 390 mm^2 ± 10 mm^2. The

permanent quadrat was partitioned into a grid of ninety 0.33 m x 0.33 m squares.

Digital photographs were taken of each square at three-month intervals, and

digital maps with a resolution of 10 mm pixels were created of the entire wall at

each point in time. Because a three-month interval was adequate to detect change

in this community (21% decrease in average colony size after the first three

months), we produced maps for three consecutive three-month time steps. Each

map consisted of a grid of 83,904 pixels, and the presence and absence of the ten

species groups was examined by hand for every pixel. The resulting matrices

(November 2002, February 2003, and May 2003) were used to reconstruct the

system's attractor. Separate reconstructions were undertaken for each species,

from which variance spectra were generated (Pascual and Levin 1999, Habeeb et

al. 2005) to determine length scales. To describe the ten species in context of the

entire community, their rank was identified on a rank abundance curve generated

with all 140 species.

3.3.3 A model community

The individual-based spatial model developed by Dunstan and Johnson (2005) was used to generate data similar to those collected on the Cathedral Cave wall. One hundred Monte Carlo iterations were run (using 600 x 600 landscapes)

Species Identifications
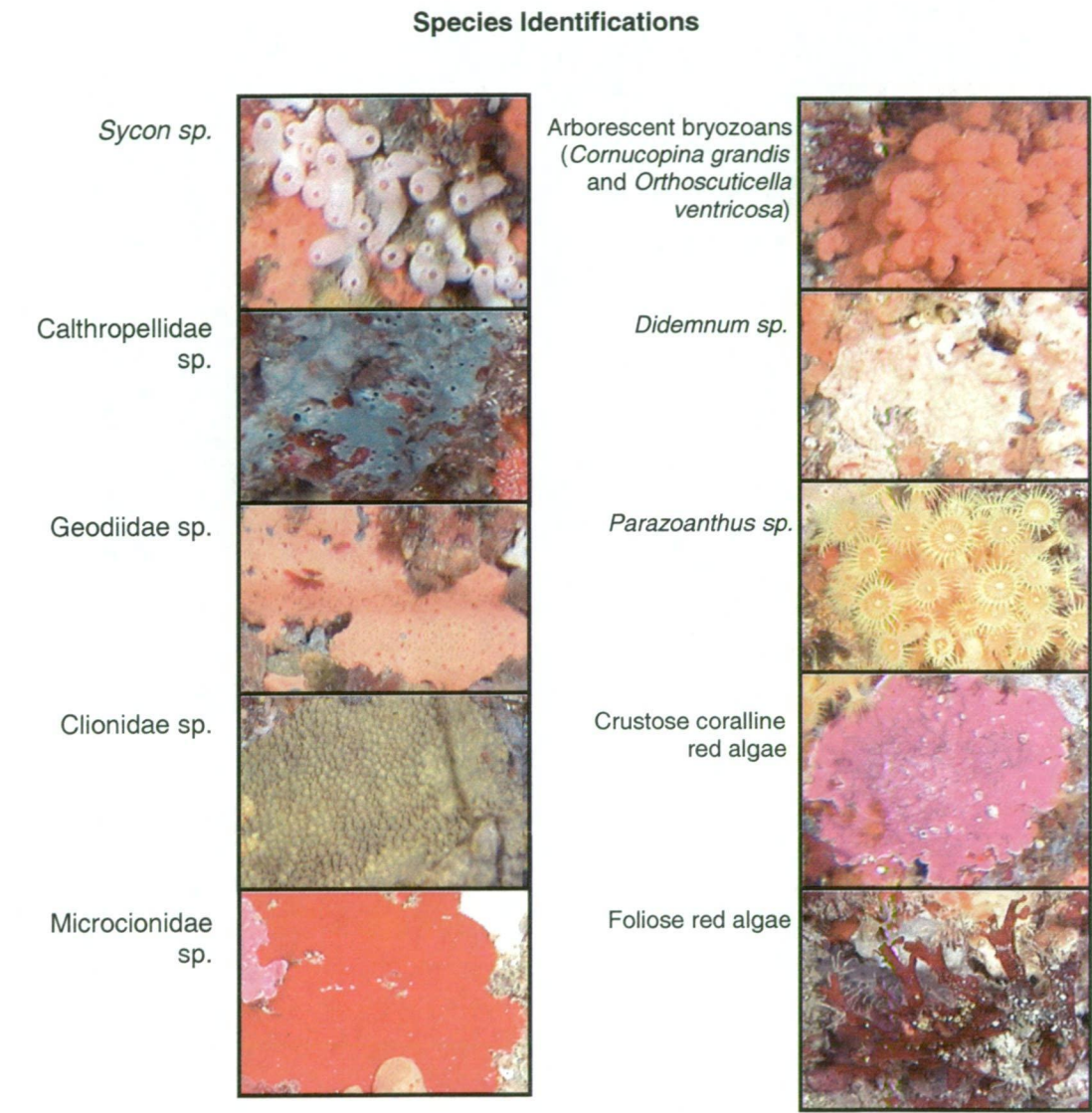
Figure 3.1. Example photos of the ten species found within the Cathedral Cave fouling community and selected for use to estimate length scales of the system. Most sponge species were only taxonomically identified to the family level, as these particular species are unnamed, but species could nonetheless be unambiguously identified from the high-resolution photographs taken of the wall.

for 5000 time steps, where each time step represented one day. This model was used to simulate the dynamics of a shallow water fouling community developed on a jetty wall ca. 30 km to the north of Cathedral Cave. The model closely captures the community structure of the jetty fouling community at 3900 time steps, representing the age of the jetty (Dunstan and Johnson 2005). Thus, we analyzed community assemblages at three time steps, approximately one year apart, that closely replicate the natural system (times 3510, 3870, and 4230). While the Cathedral Cave system changed more rapidly, a one-year interval was necessary to capture adequate change in this particular system. The three frames were sampled for each of 13 species representing a range of abundances and life histories comprising seven sponges, three bryozoans, two ascidians and one cnidarian (Table 3.2).

3.3.4 Species-area curves and variance-to-mean spectra

A species-area curve was constructed from a detailed map of all species present on the cave wall in May 2003. The number of species present in randomly placed 'quadrats' of different sizes was counted, and the region of asymptote identified from the resultant species-area curve. One hundred such curves were generated, and the mean of these curves was presented as the species-area curve.

The abundance of each of the ten selected species was determined in 100 randomly selected 'quadrats' of varying size from 0.05 to 0.50 m in length (i.e., 0.0025 – 0.25 m^2 in area), and the variance-to-mean ratio was plotted against the linear scale ($\sqrt{area}$) of the quadrat.

3.4 Results

3.4.1 Estimating CLSs of a real fouling community

CLSs were estimated by visual inspection of prediction r^2 variance spectra determined for each target species on the cave wall (Fig 3.2). Every species analyzed provided prediction r^2 spectra from which length scales could be determined, though the spectrum for species 10 was somewhat ambiguous. Nine of the ten species demonstrated primary length scales within the range of 0.20 – 0.45 m (Table 3.1), while the CLS estimated for *Didemnum sp.* was larger at 0.60 – 0.70 m.
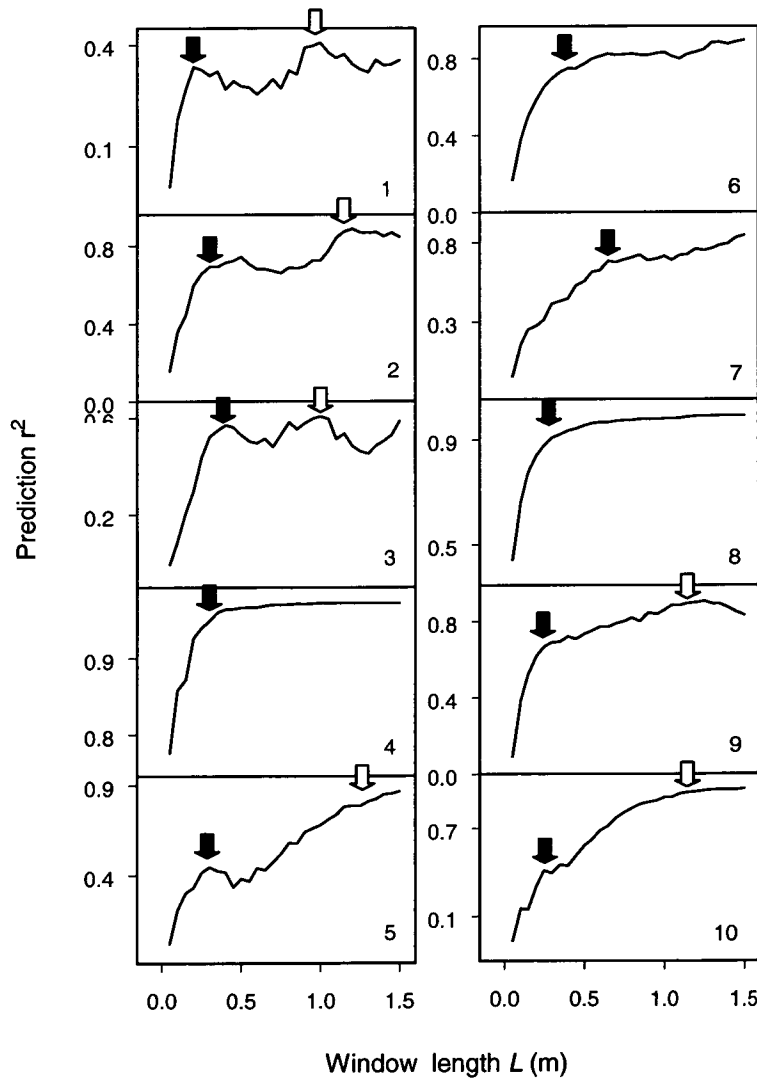
Figure 3.2. Prediction r^2 as a function of window length L for the ten species examined on the wall of Cathedral Cave. Solid arrows indicate approximate CLSs for each species, while open arrows indicate likely secondary CLSs. Numbers 1 – 10 refer to species: (1) *Sycon sp.*, (2) Calthropellidae sp., (3) Geodiidae sp., (4) Clionidae sp., (5) Microcionidae sp., (6) Arborescent bryozoans *Cornucopina grandis* and *Orthoscuticella ventricosa*, (7) *Didemnum sp.*, (8) *Parazoanthus sp.*, (9) non-geniculate Crustose coralline red algae, (10) Foliose red algae.

The average CLS range was approximately 0.30 – 0.40 m. For some

spectra secondary length scales were indicated by a second peak and subsequent

plateau of prediction r^2 (Fig 3.2, open arrows). These secondary CLSs averaged

1.1 – 1.2 m, or almost 3 times the average primary CLS (Table 3.1).

Approximately 140 species were identified within the Cathedral Cave

fouling community. The rank abundance plot from map of the wall in May 2003

(Fig 3.3) indicates that nine of the ten species selected were among the top 15

most abundant species on the wall, with five of these achieving 1 – 3% cover, and

the other five attaining 3 – 10% cover. One selected species of sponge (Clionidae

sp.) was relatively rare (with a rank of 77 out of 140), covering < 1% of the

available area.

Taxonomic group	Identification	Primary CLS range (m)	Secondary CLS range (m)
Sponges	*Sycon sp.*	0.20-0.30	0.95-1.05
	Calthropellidae sp.	0.30-0.40	1.15-1.25
	Geodiidae sp.	0.30-0.40	0.95-1.05
	Clionidae sp.	0.25-0.35	
	Microcionidae sp.	0.25-0.35	1.35-1.45
Bryozoans	Arborescent bryozoans (*Cornucopina grandis* and *Orthoscuticella ventricosa*)	0.35-0.45	
Ascidians	*Didemnum sp.*	0.60-0.70	
Cnidarians	*Parazoanthus sp.*	0.25-0.35	
Algae	Crustose coralline red algae	0.25-0.35	1.20-1.30
	Foliose red algae	0.20-0.30	1.05-1.15
	Mean	0.30-0.40	1.10-1.20

Table 3.1. Estimates of the system CLS based on analysis of ten different species groups of the fouling community at Cathedral cave. Secondary CLSs are included where applicable.

Species	Identification	Mean % cover (Nov 2002-May2003)
1	*Sycon sp.*	2.0
2	Calthropellidae sp.	3.7
3	Geodiidae sp.	1.6
4	Clionidae sp.	0.90
5	Microcionidae sp.	1.7
6	Arborescent bryozoans (*Cornucopina grandis* and *Orthoscuticella ventricosa*)	9.2
7	*Didemnum sp.*	6.1
8	*Parazoanthus sp.*	6.7
9	Crustose coralline red algae	2.1
10	Foliose red algae	4.4
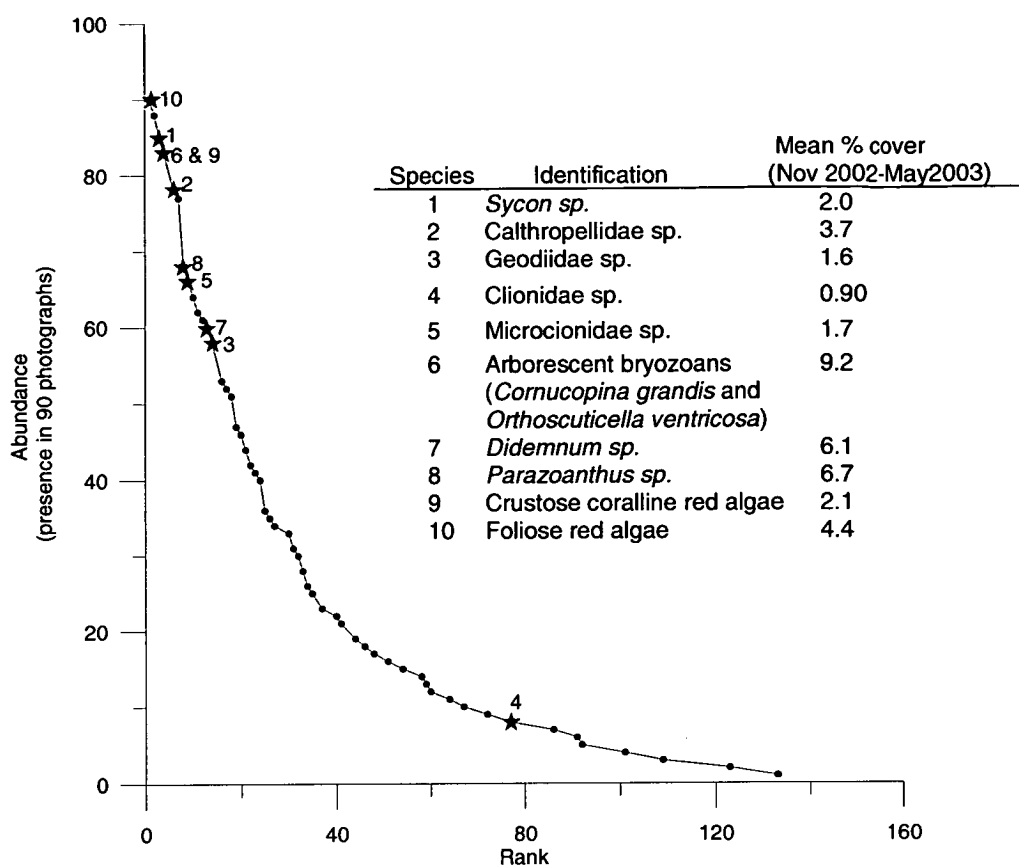
Figure 3.3. Rank abundance of all species found in the Cathedral Cave fouling community. Numbers to the right of the curve denote the ten species groups (identified by star symbols) chosen for CLS estimation, and correspond to the species identifications in the table. The table indicates mean percentage cover of the selected 10 species in the study area over three seasons (n = 3).

3.4.2 Estimating CLSs of a model fouling community

Mean prediction r^2 curves from 100 runs of the spatial model were plotted for each species, bounded by 95% confidence intervals (Fig 3.4). CLSs were estimated from the curves where possible (Table 3.2). For 12 of the 13 species in the model, prediction r^2 rapidly increased with scale before reaching a plateau. Though some curves demonstrated gradual plateaus that were challenging to interpret (e.g. Fig 3.4 K), the CLSs estimated for each species consistently ranged
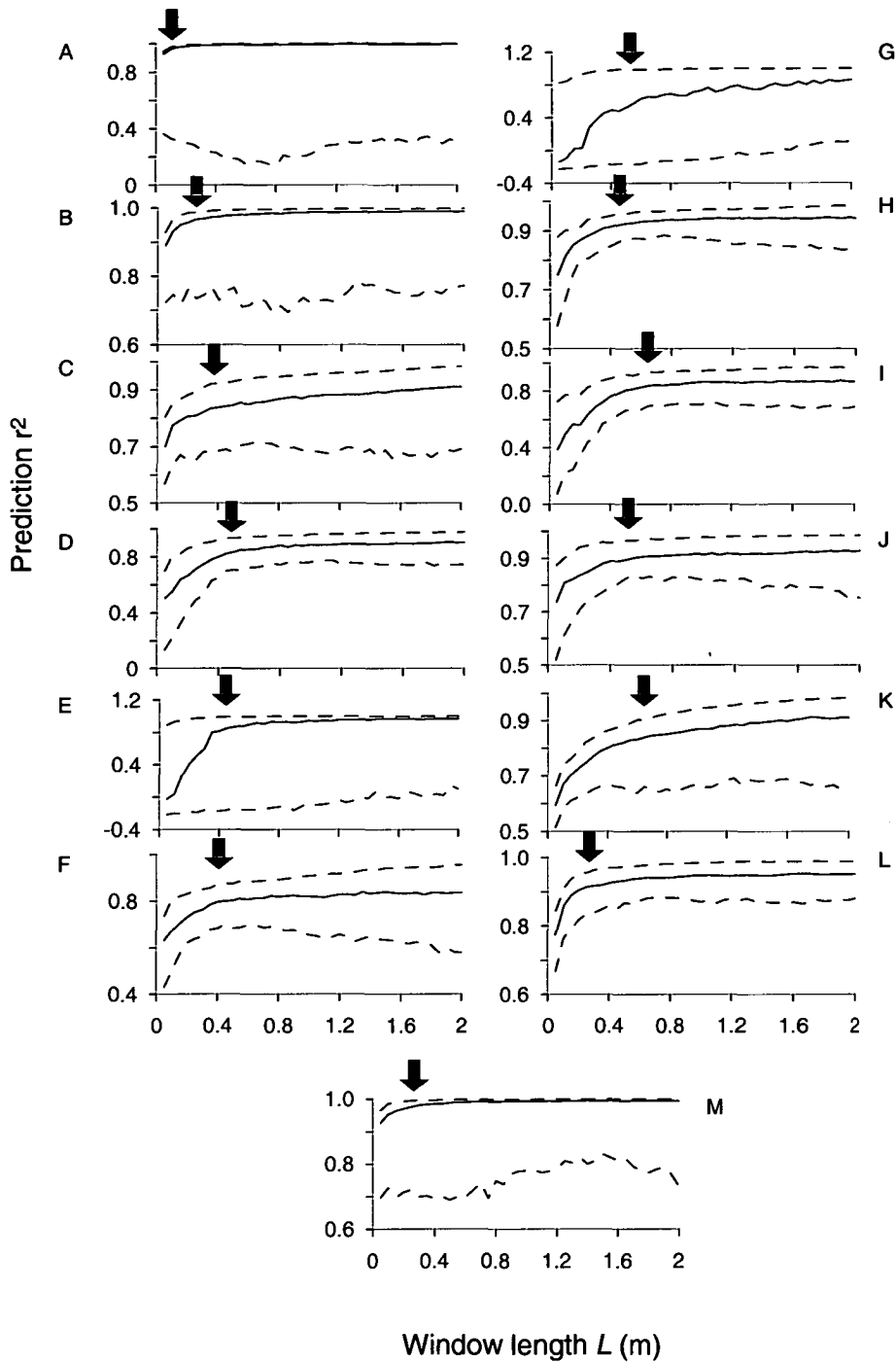
Figure 3.4. Prediction r² as a function of window length L for the 13 species of the model fouling community. Solid lines indicate mean curves of 100 Monte Carlo runs, and dashed lines indicate the 95% confidence intervals. Letters A – M refer to species: (A) Halichondriidae sp1, (B) Halichondriidae sp2, (C) Halichondriidae sp3, (D) Microcionidae sp., (E) Leucetiidae sp1, (F) Leucetiidae sp2, (G) Phloeodictyidae sp., (H) *Watersipora torquata*, (I) *Celleporaria sp.*, (J) *Parasmittina sp.*, (K) *Didemnum sp.*, (L) *Botrylloides leachi*, (M) *Corynactis australis*.

from 0.25 m to 0.70 m. The mean prediction r^2 curve for Halichondriidae sp1.

reached a plateau early and demonstrated exceptionally wide confidence intervals.

These attributes precluded the determination of an accurate estimate of scale for

this species. The average CLS range of all 13 species from the model was

approximately 0.35 – 0.45 m. The CLS range for *Didemnum sp.*, indicated as $\approx$

0.60 – 0.70 m for the cave wall fouling community, was estimated at 0.55 – 0.65

m in the model community.

Taxonomic group	Identification	Primary CLS range (m)
Sponges	Halichondriidae sp1	0.10-0.20?
	Halichondriidae sp2	0.25-0.35
	Halichondriidae sp3	0.35-0.45
	Microcionidae sp.	0.45-0.55
	Leucetiidae sp1	0.40-0.50
	Leucetiidae sp2	0.35-0.45
	Phloeodictyidae sp.	0.45-0.55
Bryozoans	*Watersipora torquata*	0.40-0.50
	Celleporaria sp.	0.60-0.70
	Parasmittina sp.	0.45-0.55
Ascidians	*Didemnum sp.*	0.55-0.65
	Botrylloides leachi	0.25-0.35
Cnidarians	*Corynactis australis*	0.25-0.35
	Mean	0. 35-0.45

Table 3.2. CLS estimates obtained from analysis of 13 species groups of the model fouling community. (?) indicates species where interpretation of the prediction r^2 spectrum is ambiguous. Secondary CLSs are included where applicable.

3.4.3 Comparison of CLS with other scales

The mean species-area curve for the fouling community was calculated

from 100 curve estimates, with means and standard errors plotted (Fig 3.5). The

average curve began to flatten at ≈ 8 m^2, indicating a length scale of ≈ 2.8 m. In

contrast, the average CLS for the fouling community occurred at an area of

approximately 0.12 m^2 (i.e., CLS ≈ 0.35 m), which captures only 35 (or 25%) of

the total species present within the community (Fig 3.5 inset).
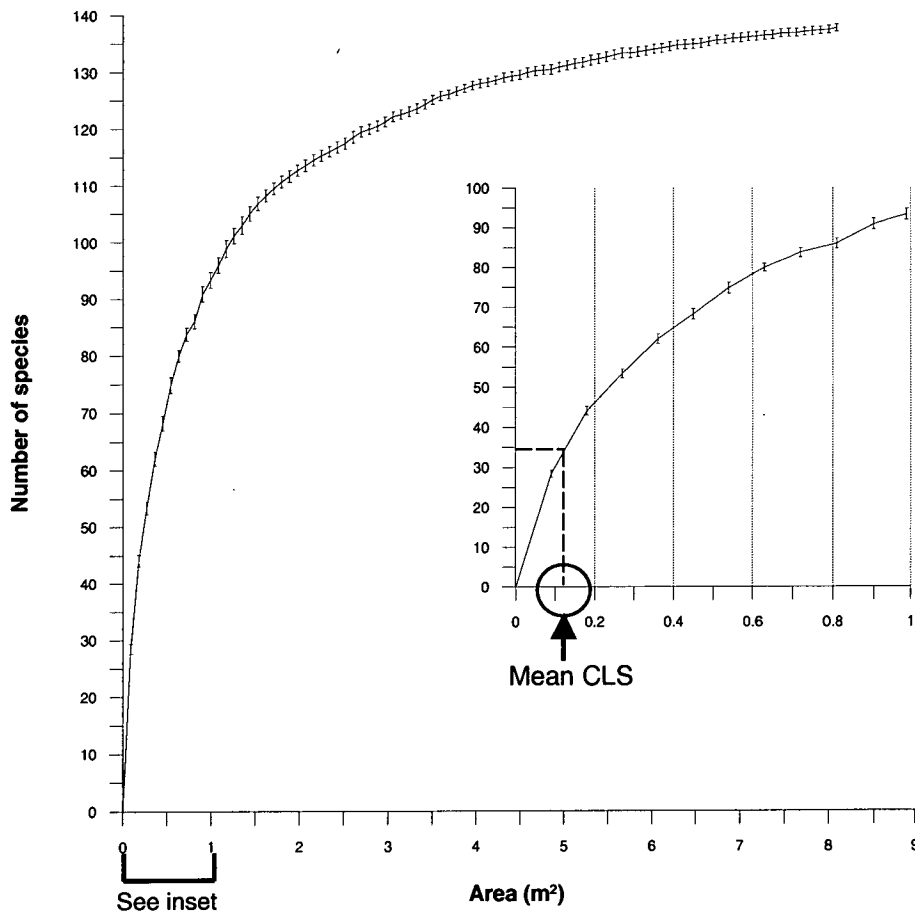
Figure 3.5. Species-area curve for the Cathedral Cave fouling community. The curve shows the mean of 100 curves generated from placing square sampling areas (ranging 0.9 – 8.1 m^2) randomly on the landscape, (bars indicate standard error). (Inset) Enlarged section of the species-area curve, with the mean CLS estimate indicated (0.35 m x 0.35 m in area). Sampling at the scale of the CLS, on average only 35 out of 140 species detected on the wall will be captured.

The abundance of each species within windows of increasing size was

identified and the variance-to-mean ratio for 100 samples positioned randomly on

the rock wall was calculated and plotted against window length (Fig 3.6). Curves

attain their maxima approximately at the CLS (ca. 0.35 m).
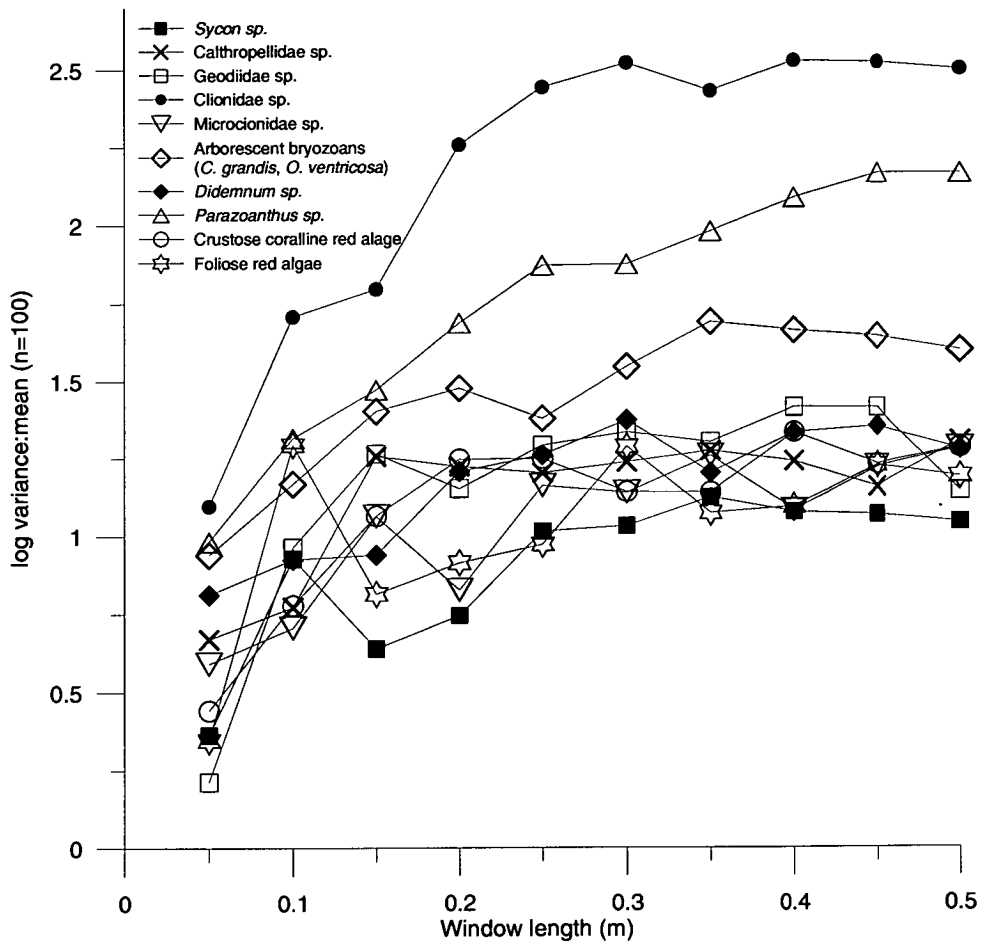
Figure 3.6. Log of the variance-to-mean ratio of abundance of the ten
target species groups at Cathedral Cave plotted as a function of window
length. Mean abundances of the species groups were recorded for 100
randomly positioned windows of each size.

3.5 Discussion

3.5.1 Length scales of a natural system

This study marks the first attempt to apply nonlinear time series analysis to detect the characteristic length scales of a natural ecological system. Determining the characteristic length scale of a system using this approach objectively indicates the scale of observation where the deterministic dynamics of the system are maximized with respect to stochastic fluctuations. Thus, using sampling units at this scale yields the most ecologically meaningful information about the trends of the system.

Our results indicate that the technique of substituting spatial replication for temporal replication in attractor reconstruction (Habeeb et al. 2005) yields interpretable prediction r^2 spectra for most species, and thus length scales, for a natural system. Using this short time series method, we successfully determined the length scale of the system from each of the 10 species (or species groups) on a rock wall fouling community. The method provided interpretable spectra for 9 out of 10 target species with only a fraction of the data that are required for the original technique (Pascual and Levin 1999). One species provided a slightly ambiguous curve (species 10, Fig 3.2), but when analyzed in association with other species' curves, the small change in prediction r^2 at the scale of approximately 0.25 m could be interpreted as its CLS.

Highly resolved photographs from only three time steps were sufficient to provide estimates of the CLS, which for this system was of the order of 0.30 – 0.40 m. Notably, a relatively rare species (a clionid sponge) indicated similar length scales to the most abundant species in this system (e.g. foliose red algae and *Sycon sp.*, Fig 3.3). Thus, the deterministic dynamics of the system would be

most evident using a quadrat size of ca. 0.35 x 0.35 m ≈ 0.12 m^2. Sampling the

system with smaller quadrats would potentially result in data inundated with

noise, while sampling with larger units would provide data with dampened

oscillations, unlikely to capture the intrinsic dynamic of the system (Pascual et al.

2001). When monitoring system trends is the goal, the optimal sampling scale

defined by the CLS should be used for the aggregation of data.

Curves of some species indicated the presence of a secondary length scale

as a second peak in the prediction r^2 spectrum (Fig 3.2, open arrows, Table 3.1).

Secondary length scales, larger than the primary or smallest length scale, have

previously been detected in model systems, and are interpreted as the scales of

emergent dynamics reflecting interactions at the patch level within the system

(Habeeb et al. 2005). Although the likelihood of identifying the scale(s) of higher-

level emergent dynamics is greater when using the long time series method of

attractor reconstruction (Habeeb et al. 2005), 60% of the species (most species of

sponge and both algal groups) displayed these larger scales. Secondary CLSs

ranged from ca. 1.0 – 1.4 m, with an average length of ca. 1.1 – 1.2 m (≈ 1.3 m^2 in

area). Most second peaks were 3 – 4 times the scales of the first peaks. When

detectable, these secondary CLSs may be useful indicators of the dynamics at the

patch level. The only sponge which did not indicate a secondary CLS was the

clionid. This particular sponge was relatively rare on the cave wall (Fig 3.3),

thereby providing little opportunity for patch level emergent dynamics.

It is not sensible to discuss whether primary or secondary length scales are

more 'valid,' useful, or reliable. They indicate different things, namely the scales

at which local and emergent dynamics are best observed. Because patch level

emergent phenomena are more variable and less predictable than local dynamics,

secondary length scales may not always be evident. The capacity to detect the scales of the patch level emergent dynamics is significant, since it provides a more complete picture of spatial complexity, such as that manifest as spatial self-organizing (Johnson 1997).

3.5.2 Comparison of the CLS with other scaling measures

Because the length scale is such a new measure, it is useful to compare it to other important scales that capture essential information about structure and dynamics of ecosystems, such as the scale of asymptote of a species-area curve, and the scales of change in variance-to-mean ratios of species' abundances. If the community at Cathedral Cave is sampled at the optimal quadrat size as estimated by CLSs, on average approximately 25% of the entire species complement on the wall will be captured in a single sampling unit (Fig 3.5). The area where all species are captured (ca. < 3 m in length) is considerably larger than that indicated by the CLS (ca. 0.35 m in length) in this case. The optimal scale at which to observe the community dynamics of this system is approximately 100 fold smaller in area than the scale at which most species in the system are captured. If this relationship holds true for all systems, it would impact the sampling scales of empirical experiments conducted to assess functional relationships within ecosystems, as results obtained are often scale-dependent (Weitz and Rothman 2003).

For example, sampling units of experiments addressing the 'diversity-stability debate' (McCann 2000) should be determined carefully, because our results indicate that the experiments observing a system's stability would need to be carried out on a much smaller spatial scale (almost 100 fold smaller in area) than the scale necessary to detect the entire suite of species present in the

community. Indeed, extrapolating results from microcosm experiments testing multitrophic stability to ecosystem scale diversity, for example, may be ambitious, and a particular CLS may not be appropriate for such cross-scale phenomena. Instead, investigations that actually link mechanistic, small-scale phenomena to higher-scale stabilizing processes will likely be necessary.

Even with high replication of sampling units taken at the CLS, it is unlikely that all species in this particular system would be detected, suggesting that "complete" knowledge of the system is not required for detecting trends. Interestingly, sampling at an area dictated by the average secondary length scale (ca. 1.3 m^2) would capture approximately 100 (out of 140) species within this system. Even at length scales indicating higher-level emergent dynamics, not all species would be captured.

The colonies living on this particular cave wall are relatively small (average colony area 390 mm^2 ± 10 mm^2), with a maximum colony area (0.04 m^2) still well below the area estimated by the CLS (0.12 m^2). The optimal scale to observe the trends of this densely packed fouling community is thus between the maximum colony area and the scale needed to detect the majority of the species present (8.0 m^2) as determined by the species-area curve. Further research on different systems is necessary to elucidate the relationship between colony size, species packing density (or richness), and the CLS, but we would expect that non-sessile communities will prove to have larger CLSs due to increased colony sizes, larger inhabited areas, and decreased species packing densities. Indeed, models of predator-prey systems have shown that as mobility of organisms within a community increases, the magnitude of the characteristic length scale also increases (De Roos et al. 1991).

The variance-to-mean ratio is the basis of several indices to describe the intensity of spatial patterns of aggregation (Pielou 1977). For the Cathedral Cave community, the variance-to-mean ratio was > 1 for all species across all window sizes, suggesting clumped distributions. As window length increased from 0.05 m to 0.50 m, the ratio approached its maximum for all species, with most species reaching a plateau at, or slightly smaller than, the CLS (approximately 0.35 m, Fig 3.6). Thus, though the variance-to-mean ratio is an entirely linear prediction and prediction r^2 is highly nonlinear, the CLS may be related to that scale where the change in variance-to-mean of species abundances reaches a maximum. For the large sample size of 100, the mean density of each species is relatively stable, thus the variance-to-mean ratio reflects predominantly the variance in the species' densities. The variances plateau at scales similar to the average CLS of the system, suggesting that the CLS approximates the average scale of aggregation in this system. A similar finding, that the CLS estimation technique could be used for identifying the spatial scale of aggregation, has been explored on simple model systems previously (Keeling et al. 1997). Whether this relationship between the two metrics holds true for other more complex systems is unknown, but is an interesting focus for further evaluation.

3.5.3 Dependence of CLS on species characteristics

The characteristic length scale is, in theory, independent of the species used for attractor reconstruction, and indeed we found that nine out of ten species, covering a range of distinct life histories, indicated similar CLSs. For example, while the clionid sponge is distinct from the other species in that it is rare (Fig 3.3) and grows slowly to a very large colony size with low mortality, it indicated

a similar primary CLS to other species. Furthermore, species within the different

taxonomic groups of sponges, bryozoans, cnidarians, and algae all indicated

similar length scales. Thus, the system-level CLS seems largely independent of

life history characteristics or taxonomy when the species dynamics are linked, as

predicted by theory. In this system, species are strongly connected because they

all compete for space, which is the limiting resource.

While most CLS estimates ranged from 0.20 – 0.45 m, *Didemnum sp.* was

the exception, with an estimate considerably larger (0.60 – 0.70 m). This may

reflect a species dynamic strongly independent of the system dynamic. The

degree of connectedness between species likely influences the solidarity of their

CLS estimates; thus, we hypothesize that the outlying CLS estimate of *Didemnum*

may reflect its lack of connectedness to the other species in the community

(Habeeb et al. 2005). Notably, *Didemnum sp.* is a particularly rapid colonizer with

a rapid turnover rate, and forms many small patches of colonies able to quickly

overgrow other species (Habeeb, personal observation; Oren and Benayahu

1998). It is arguably less influenced by the availability of unoccupied space than

other species.

While our evidence suggests that most species indicate a similar length

scale at the system level, it is clear that exceptions may arise, as in the example of

Didemnum. Our recommendation is to validate the system-level CLS based on

estimates obtained from several species.

3.5.4 Length scales of a model fouling community

If estimates of length scales are to be useful in facilitating objective

decisions about scales of observing real systems, then the estimates should not be

location specific for the same kind of system. Comparing results for the real

fouling community we examined with those from a model based on a similar

community in a different location (Dunstan and Johnson 2005) allowed us to

verify the CLS estimates and further explore the utility of CLS estimates to

ecologists. Prediction r^2 curves for the 13 species derived from the spatial model

increased and then reached a plateau, as expected, and yielded CLSs in the

average range of 0.35 – 0.45 m, which overlapped with the average range of the

real fouling community (0.30 – 0.40 m). The similarity of CLS estimates for the

model and real fouling communities suggests that the results can be generalized

across fouling communities despite that the most abundant species in these two

communities are distinctly different (Tables 3.1 and 3.2). Thus, when sampling

these kinds of temperate marine fouling communities, a length scale of ca. 0.35 –

0.45 m (i.e. using sampling units of area 0.12 – 0.16 m^2) is likely to be optimal in

distinguishing real trends from noise.

While ten of thirteen species indicated length scales in the range of 0.35 –

0.45 m, some species within the model provided larger CLSs, with the bryozoan

Celleporaria sp. and the ascidian *Didemnum sp.* at the top of the range. It is

interesting that *Didemnum sp.* also indicated the largest CLS estimate for the cave

wall system. Again, because some species like *Didemnum sp.* may deviate from

the majority, examining several species is important for accurate detection of the

system CLS. Notably, no secondary CLSs were detected for the model fouling

community, possibly indicating that the model system is not strongly self-

structuring, i.e. demonstrates weak patch level emergent dynamics.

3.5.5 Conclusion

An appropriate choice of sampling scale is critical to accurately detect trends in ecological systems. While the CLS is not a precise measure, it provides an objective estimate upon which to base sampling designs. Without this, we arbitrarily or intuitively design experiments and potentially miss crucial information due to the scale dependence of many ecological patterns and processes. This work marks the first attempt to estimate the CLS of a natural ecosystem, and the method proved successful. The CLS of the cave wall system was detectable, and similar to that indicated for a similar model community. We are now confident that meaningful natural length scales can be estimated for more complex real ecological systems, and we anticipate that as the technique becomes more widely used, the relationship between the CLS and other more established ecological scaling phenomena will become apparent.

Chapter 4

Characteristic length scale estimates: Robustness to the sampling regime

4.1 Abstract

Choosing the appropriate scales at which to observe an ecosystem is an important decision because the pattern that is detected will likely depend on the scale of observation. Researchers often rely on intuition to decide the sampling scale, but if their scale is too large or too small, they potentially miss the real system trends, which may lead to mistakes in management. For this reason, an objective approach to identifying the optimal scale at which to observe an ecological system is appealing. Sampling at this 'characteristic length scale' (CLS) maximizes the potential to detect the deterministic dynamics of an ecological system. A technique to estimate CLSs of complex dynamical systems has recently been developed and tested on a natural system, but has not yet been rigorously assessed for sensitivity to several choices about the nature of the data used to derive the estimate. The approach is based on a short temporal sequence of spatial maps of the landscape, but whether the CLS estimate is influenced by the time interval between maps and the number of successive maps has yet to be studied. Here, we assessed the robustness of the CLS estimate of a real marine fouling community to different time intervals between the maps, to changes in the number of maps used, and to varying start dates of map sequences. The CLS estimate for this system remained consistently around 0.30 – 0.40 m for most species despite changes in these choices, and despite that species represented a

diversity of taxa and life histories. The high level of robustness of this estimate suggests that the CLS is a suitable replacement for scientists' intuition as a tool to select appropriate scales of observation of an ecosystem. Use of this new approach to objectively determine optimal scales of observation should instill confidence that sampling programs will most effectively indicate deterministic trends against a background of system noise.

4.2 Introduction

The scale at which a system is viewed largely influences perceptions of that system's dynamic (Wiens 1989, Levin 1992). Thus, choosing the appropriate scale to best detect ecosystem trends is an important step in any task designed to monitor ecosystem dynamics. Currently, scientists and managers rely largely on their intuition and experience to determine scales of observation. However, subjective choices of sampling scale may lead to potential bias and/or misinterpretation of variability (Denny et al. 2004), while observations of the same system taken at different scales may provide dissimilar or even contradictory findings (Allen et al. 1984, Wiens 1989, Castilla 2000, Hill and Hamer 2004, Chalcraft et al. 2004). Not surprisingly then, determining an appropriate sampling scale is of utmost importance to managers of natural resources (Stalmans et al. 2001, Nicholson and Jennings 2004).

While more than one natural scale likely exists for complex real systems, resource managers often must choose a scale at which to sample their system. At this scale, the determinism to noise ratio should be maximized such that the dynamical trends can most clearly be observed. Sampling at smaller scales is likely to only capture noise within the system, while sampling at scales that are

too large is not only likely to be expensive but may render trends undetectable because independent system dynamics will overlap and may be averaged out (Fig 1.1; Pascual and Levin 1999, Pascual et al. 2001). It can be argued that appropriate management practice will be best informed by the clearest possible detection of system trends, which in turn requires an objective means to determine the optimal scale of observation to reflect those trends.

The characteristic length scale (CLS) is that optimal scale which maximizes the ratio of determinism to noise (Keeling et al. 1997, Pascual and Levin 1999). For fluctuating systems with complex nonlinear dynamics, this scale is estimated by reconstructing the system's attractor in phase space (Takens 1981). The approach is unique in that it accommodates complex nonlinear behaviors and, because it is based on the dynamical attractor which captures all possible states of the system, it yields a system-level scale that is, in theory, independent of the species being sampled. Based on a modification of this approach, Habeeb et al. (2005) showed recently that meaningful CLSs of natural ecological systems can be detected, and that variation in the estimate based on observations of dissimilar taxa is very low, supporting the notion that the estimate provides ecosystem-level information.

The minimum data requirement to estimate a system's CLS is three spatially resolved maps, consecutive through time (Habeeb et al. 2005). However, if the idea of CLSs is to be usefully applied in ecology, then the CLS should not be too sensitive to arbitrary choices of the time interval between maps or the number of maps. In this paper, we examine the sensitivity of a system's CLS to the frequency of sampling (time interval), the number of maps used to reconstruct the attractor (time steps), and the date that sampling is started. The aim is to

assess the robustness of the technique developed recently by Habeeb et al. (2005)

for estimating the CLSs of real ecosystems. Our results show that CLSs are

consistent across a range of choices about the sampling regime, further promoting

the utility of CLSs as useful metrics for ecologists and ecosystem managers.

4.3 Methods

4.3.1 CLS estimates

We studied a marine fouling community at the entrance to Cathedral Cave

(43°3'43S, 147°57'1E) on the southeast coast of Tasmania, Australia. The

vertical rock wall entrance to the cave is entirely covered by an assemblage of

sessile invertebrates and some algae, totaling approximately 140 species, which

compete for space on the dolerite substratum. A 3.0 m x 3.3 m permanent quadrat

established at 7 - 10 m depth on the rock wall was photographed initially in

November 2002, and every 3 months thereafter for 18 months (note: no sample

could be taken at 15 months due to inclement weather). High-resolution digital

maps (10 mm pixels) were created for each time step. Each of these six maps

consisted of grids of 83,904 pixels, and each map was analyzed manually for ten

target species groups, chosen to represent a range of phyla, life history traits and

abundances. The species groups included: five species of sponges (four unnamed

specimens lodged with the Tasmanian Museum; K430, K431, K432, K433), one

group of arborescent bryozoans, one ascidian, one cnidarian, and two groups of

algae (Table 1). The digital maps so created provided a short time series of spatial

landscapes from which the CLS(s) could be estimated (Habeeb et al. 2005,

Chapter 3). The primary CLS was determined as the plateau, or point of first

major inflection in the spectrum of prediction r^2 versus scale (Pascual and Levin

Species number	Identification	Taxonomic group
1	*Sycon* sp.	Sponge
2	Calthropellidae sp. (K431)	Sponge
3	Geodiidae sp. (K433)	Sponge
4	Clionidae sp. (K430)	Sponge
5	Microcionidae sp. (K432)	Sponge
6	Arborescent bryozoans (*Cornucopina grandis* and *Orthoscuticella ventricosa*)	Bryozoan
7	*Didemnum* sp.	Ascidian
8	*Parazoanthus* sp.	Cnidarian
9	Crustose coralline red algae	Algae
10	Foliose red algae	Algae

Table 4.1. Identification and taxonomic groupings of the ten selected species from the Cathedral Cave rock wall. Species numbers reflect identification codes used in the figures and text throughout. Corresponding Tasmanian Museum specimen numbers are indicated in parenthesis by species whose precise identities were unknown.

1999, Habeeb et al. 2005). Secondary CLSs were identified as subsequent peaks in the spectrum (Habeeb et al. 2005). In some cases, there may be uncertainty as to the location of the asymptotes or inflections in the prediction r^2 spectrum (see Fig. 3.2, sp 10). Hence, multiple species within the system should be evaluated so that any ambiguity can be resolved. Previous results (for the fouling community, Chapter 3) have shown that 80-90% of species will indicate similar CLSs.

We have already shown that the spatial time series we obtained from the rock wall at Cathedral Cave yielded prediction r^2 spectra from which CLSs could be interpreted, and that most of the dissimilar species yielded similar CLSs

(Chapter 3). Here, we examined the sensitivity of CLS estimates to arbitrary choices of the nature of the spatial series by varying the number of time steps in the time series (3, 4, or 5), the interval between time steps (3 month, 6 month, or 9 month), and the start date (i.e. initial configuration) of the landscape (month 0, month 3, or month 6). The CLS was evaluated for all possible combinations of these parameters using the short time series technique (Habeeb et al. 2005).

4.3.2 Species' life history differences

Simple metrics were calculated to summarize some features of the dynamics of the ten 'target' species from which we are able to infer (dis)similarities in life history strategies. The change in the total area that a species occupies is one indicator of its dynamics through time, but this measure will not accurately reflect colony turnover between time steps. Thus, species were assessed for changes in both their abundance (percentage cover), and absolute and relative turnover rates between time steps. Absolute turnover (t) of species s from time x to time y is defined as

$$t_{s,xy} = \frac{(m_{s,xy} + g_{s,xy})}{2}$$

where $m_{s,xy}$ is mortality, or the number of pixels occupied by species s at time x but not at time y; and $g_{s,xy}$ is growth, or the number of pixels occupied by species s at time y that were not occupied by s at time x. Relative turnover, or percentage turnover (%t) between times x and y is defined as

$$\%t_{s,xy} = \frac{(m_{s,xy} + g_{s,xy})}{2s_{ix}} \times 100$$

where s_{ix} is the initial abundance of species s at time x.

4.4 Results

4.4.1 CLS estimates

Prediction r^2 spectra were produced for each species for all possible combinations of numbers of time steps, intervals between time steps, and start dates. Particular combinations are indicated with reference to the month (0 – 18) of sampling. For example, spectra derived from landscape maps taken 3 months apart, with month 0 as the start date, and with 3, 4, and 5 time steps are indicated as times 0-3-6, 0-3-6-9, and 0-3-6-9-12 respectively (see Fig 4.1, first row). For brevity, we only present the prediction r^2 spectra of four species groups, which encompass the species showing the greatest variability among spectra (the Calthropellid sponge, Fig 4.1A), the species with the most consistent spectra (the Clionid sponge, Fig 4.1B), and two other average representative species (bryozoans, Fig 4.1C; *Parazoanthus* sp. Fig 4.1D). Prediction r^2 spectra of the other species are given in Appendix D. Characteristic length scales were estimated visually as the point (or region) of the first plateau of the prediction r^2 spectrum, where the slope exhibited an abrupt shallowing (Table 4.1; Pascual and Levin 1999, Habeeb et al. 2005).

For most species, the primary CLS estimated remained between 0.30 m and 0.40 m regardless of the manner in which the system was sampled (Figs 4.1, 4.2). Changing the time intervals between consecutive landscapes from 3 to 6 to 9 months had little effect on the CLSs indicated by most species. Similarly, changing the number of landscapes in the time series, and the starting configuration of those landscapes had little effect on the CLS. However, for the calthropellid and geodid sponge species (species 2 and 3), both the number of

time steps and the start date affected the CLS indicated, with estimates ranging

between 0.25 and 0.60 m, depending on parameters of the sampling (summarized

in Fig 4.2). Most of this fluctuation occurred when sampling intervals were 3

months apart (Fig 4.2). The CLS indicated by the calthropellid sponge, which was

the species showing most variation in the estimate, increased to approximately 0.6

m when the 0-3-6-9 and 3-6-9 month samples were used in the reconstruction.

The estimate indicated by the geodid sponge increased to approximately 0.45 m

using the 0-3-6 month sample, and to 0.6 m using the 3-6-9 month sample. For

these two sponge species, increasing the time intervals between spatial maps

increased the consistency in estimates indicated. Increasing the intervals between

maps from 3 months to 6 and 9 months resulted in more consistent estimates

(calthropellid, 0.2 – 0.3 m; geodid, 0.25 m; Fig 4.2), which were similar to

estimates indicated by the other species.

We also examined the effect of unequal time intervals between successive

landscapes (for example, 0-3-6-12 months) for reconstruction of the attractor, but

found very little effect on the CLS estimated (Fig 4.3).

4.4.2 Dynamics and life histories of target species

In theory, any species in a given system can be used in attractor

reconstruction to indicate the characteristic length scale of the system (Takens

1981). However, since some ecologists suggest intuitively that species with

strikingly different dynamics and life histories might yield dissimilar scales, we

compared the CLS estimates indicated by different species. We used the

Figure 4.1. Prediction r² versus scale spectra derived from (A) Calthropellidae — species 2, (B) Clionidae — species 4, (C) arborescent bryozoans — species 6, (D) *Parazoanthus* sp.— species 8 (spectra for the other six species can be viewed in Appendix A). The different spectra for each species represents different combinations of the time interval between successive landscapes, the start date of sampling, and the number of time steps in the series. Filled arrows indicate primary CLS estimates, while open arrows (D) indicate secondary length scales. Wide arrows are used to emphasize that the CLS estimates are in reality a range rather than a precise value. "Times" are the months of sampling the rock wall used to derive the spectrum, with the initial sampling at time = 0.
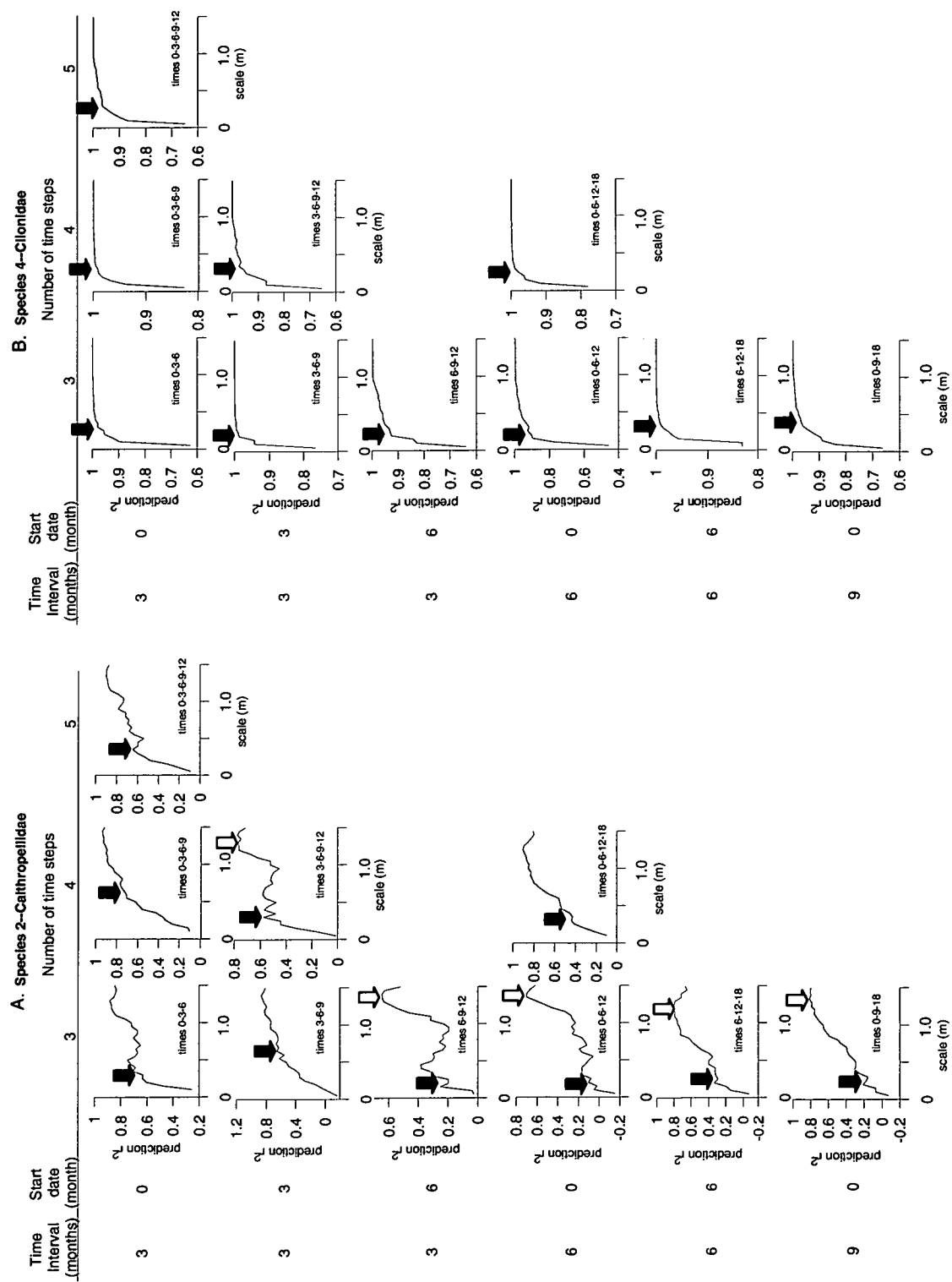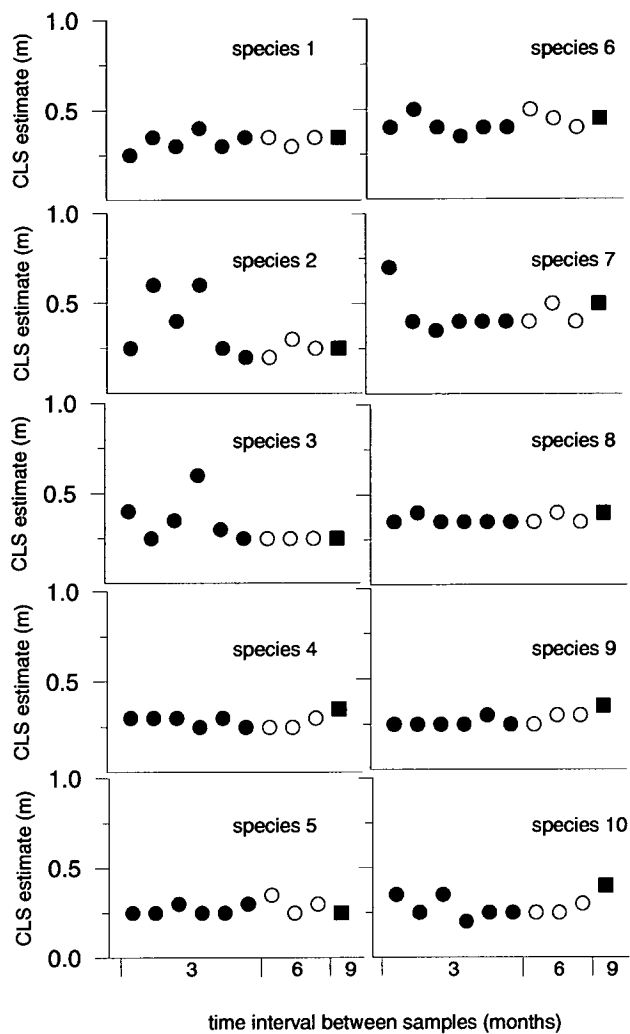
Figure 4.2. Summary of variability in CLS estimates determined from species 1 to 10, for the various combinations of sampling pattern we examined. Filled circles indicate 3-month intervals between time steps, with varying start dates ,and numbers of landscapes in the time series (corresponding to the top 6 plots of each species in Fig 4.1); open circles indicate 6-month intervals, and filled squares indicate 9-month intervals. Species identifications are given in Table 4.1.

dynamics of the ten species representing five phyla, to infer their life history strategies, which we then compared with the CLS estimates indicated. The species for which CLS estimates showed sensitivity to changes in the sampling

regime could then be assessed for particular differences in life history strategies, which might explain the variability in the CLS estimates they produced.
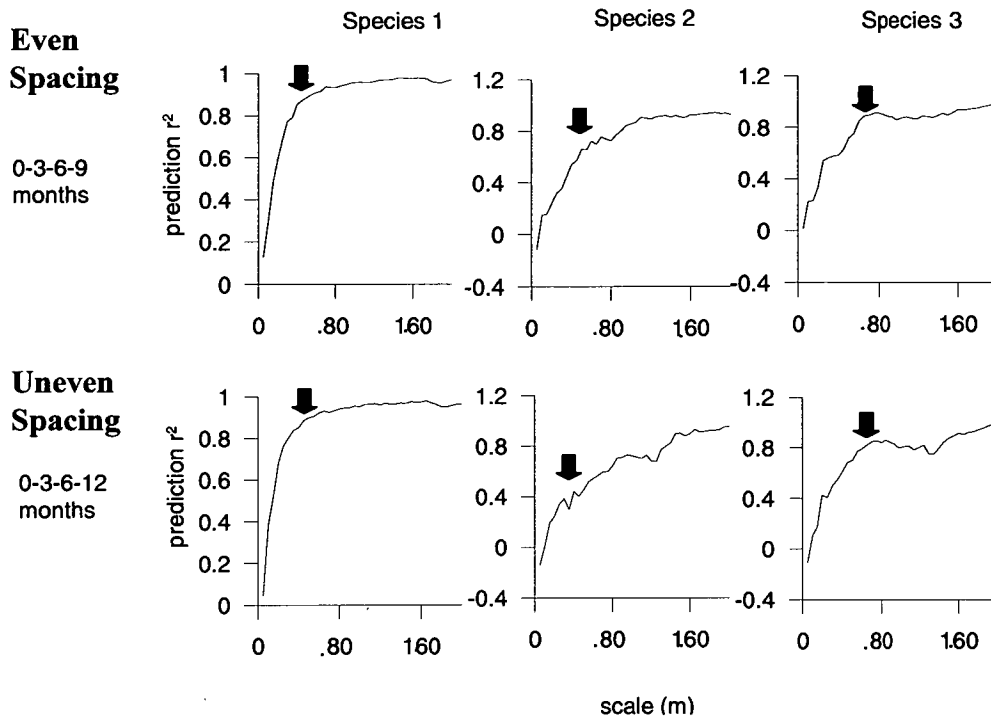
Figure 4.3. Spectra of prediction r^2 versus scale for the estimation of characteristic length scales, using the short time series method of Habeeb et al. 2005. Upper graphs are based on an evenly spaced sequence (0-3-6-9 months) and lower graphs on an unevenly spaced sequence (0-3-6-9 months), for three species within the Cathedral cave fouling community (Table 4.1). Arrows indicate approximate CLS estimates.

The range of space-time dynamics represented by the ten species groups is summarized by both the change in cover and the amount of turnover of the species experienced between two time steps. Variation in the characteristics of these dynamics reflects a range of life history strategies, as illustrated by the absolute (Fig 4.4A) and relative (Fig 4.4B) changes in area of each species on the cave wall versus turnover between two 12-month intervals (0 – 12 month and 6 – 18 month; Fig 4.4). Turnover indicates mortality in one position that is replaced
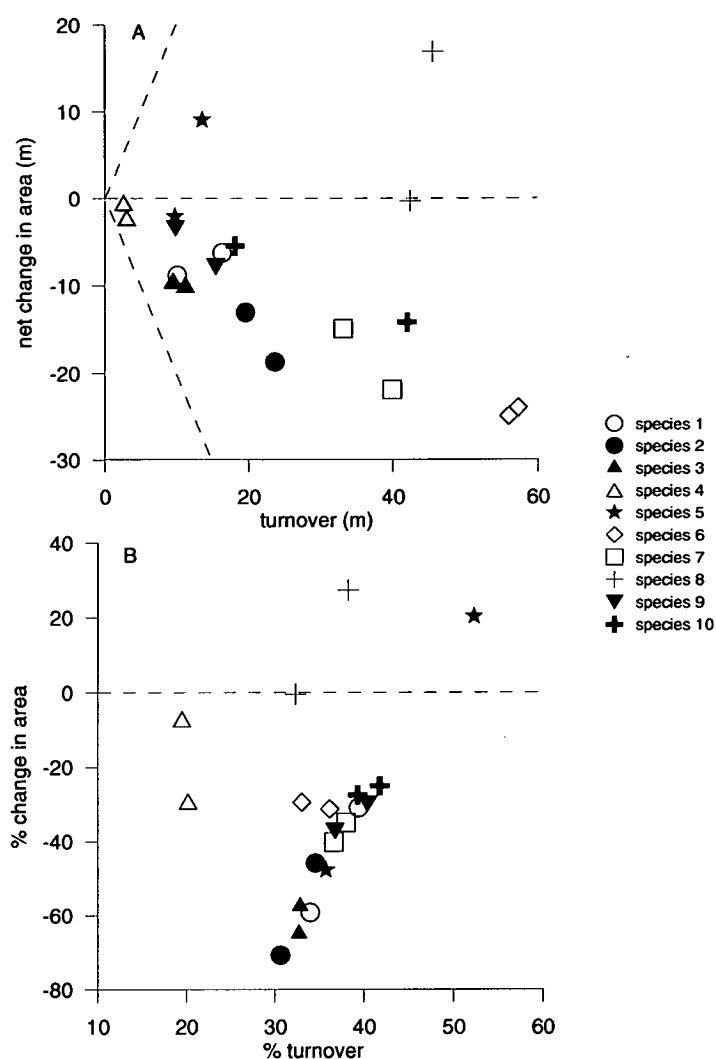
Figure 4.4. (A) Net change in area versus absolute turnover between two 12-month intervals (0-12 month, and 6-18 month) for each species. The horizontal dashed line indicates the space where growth and mortality are equal. Above the upper dashed line is the space where no mortality occurs, and below the lower dashed line is the space where no growth occurs. Species vary across a range of space-time dynamics, with more k-selected species close to the horizontal line and to the left of the graph with low turnover, and more r-selected species further from the horizontal line, but to the right of the graph, with high turnover. (B) Relative change in area versus relative turnover between two 12-month intervals (0-12 month, and 6-18 month) for each species. The dashed line indicates the space where growth and mortality are equal. Some target species demonstrated over 50% turnover in 12 months, while turnover in other species was less than 20% over the same period. Changes in the pattern from above (A) reflect the influence of colony size on the relationship. Species identifications are as in Table 4.1.

by new growth elsewhere on a spatial landscape. Species with low absolute turnover and a small change in the area occupied represent the more k-selected species, which are relatively long-lived and stable through time. Species with high

absolute turnover and a large (positive or negative) change in area reflect dynamics at the *r*-selected end of the life history spectrum, with relatively high rates of mortalities, colonization, and usually, growth. The dynamics of the target species on the cave wall vary across a continuum of these two strategies (Fig 4.4A). Some of our target species demonstrated in excess of 50% turnover in 12 months, while turnover in other species was less than 20% over the same period (Fig 4.4B). Comparing this relative measure of turnover with the absolute measure elucidates that the net change in area is related to body size for some species (Fig 4.4B). For example, microcionid sponges (species 5) indicated high relative turnover (~ 44%) but low absolute turnover, which reveals that they form small colonies which quickly compensate for mortality by further growth. Conversely, arborescent bryozoans (species 6), which form large rapidly changing colonies, exhibited the highest absolute turnover, but average relative turnover. Investigating both relative and absolute measures is necessary for a comprehensive picture of the different species' dynamics. Notably, the ten species observed are generally highly dynamic, with an average relative turnover rate of approximately 35% of a colony in 12 months.

Species can be separated arbitrarily into four groups of life history types when variance in total cover as a function of absolute turnover is plotted for the three-month intervals (0 – 3 month, 3 – 6 month, 6 – 9 month, and 9 – 12 month; linear relationship of variance to the mean, $y = 0.69x - 45.85$, $R^2 = 0.865$; Fig 4.5). Species such as *Sycon* sp. and nongeniculate coralline red algae (species 1 and 9), with low turnover and high variance in total cover over the three-month intervals, appear to grow and recede seasonally. Arborescent bryozoans (species 6) are outliers with high absolute turnover but low variance in total cover between

three-month periods, indicating rapid formation of new colonies balanced by

rapid mortality (Fig 4.5). Species in the group nearest the origin grow slowly and

maintain a relatively constant total cover (Fig 4.5).
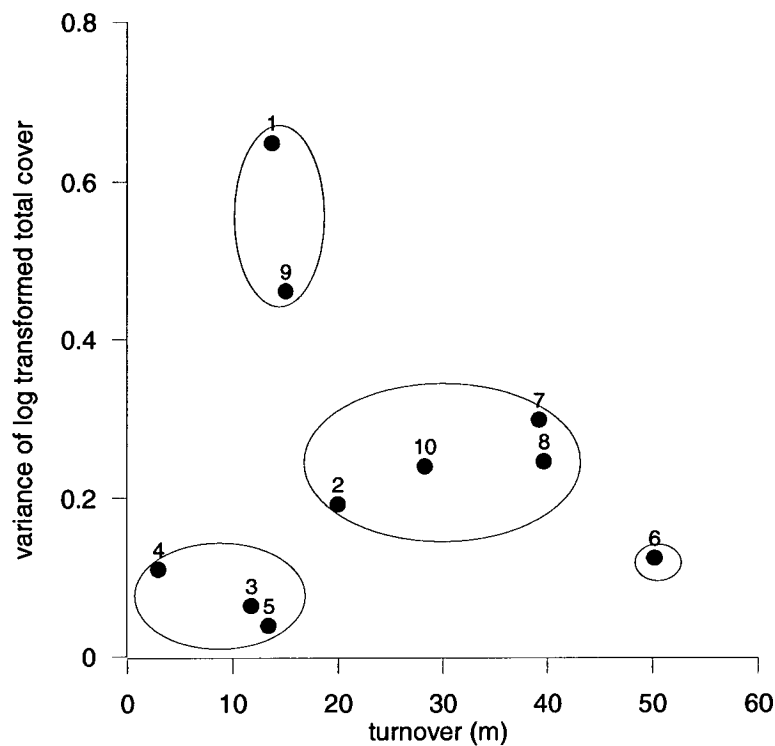
Figure 4.5. The variance of log transformed total cover between all 3-month intervals (n=4; 0-3 month, 3-6 month, 6-9 month, and 9-12 month) versus absolute turnover for each of the ten species. Species numbers are indicated above the points, and correspond to the identifications in Table 4.1. Species can be separated arbitrarily into four groups of distinct space-time dynamics, identified by circles, which reflect dissimilar life history strategies.

4.5 Discussion

4.5.1 Robustness of CLSs to arbitrary choices in sampling regimes

A CLS is, in theory, an inherent characteristic of an ecosystem (Carlile et al.

1989, Wilson and Keeling 2000). If this is true, and providing that the attractor

describing the system dynamic is stationary, then in principle, the CLS should not be sensitive to the particular protocol by which a system is sampled. If techniques to estimate CLSs are robust to arbitrary choices in the sampling regime, then determining CLSs should provide objective estimates of the optimum scales to best observe the deterministic trends in a system, which is of critical importance to inform scientists and managers of changes in natural systems and their resources.

We are encouraged that the estimate of CLS is largely consistent, at approximately 0.30 – 0.40 m for this system, across a range of dissimilar species (Fig 4.2; see also Chapter 3), and across dissimilar patterns of sampling the system. For most species, changing the time interval between successive spatial maps from 3 months to 9 months had very little effect on the CLS estimate (Fig 4.1B, 4.1C, 4.1D). Similarly, commencing the sampling in different seasons (November [month 0] versus May [month 6]) had little effect on the CLS estimates (Fig 1).

Importantly, the CLS indicated by most species was not overly sensitive to large differences in the quantity of data used to derive the estimate, i.e. by changing the number of maps in the time sequence from 3 to 4 or 5 (Fig 4.1). Systems that show little change with successively higher embeddings are low-dimensional and are considered 'nearly deterministic,' while a stochastic system has infinite dimensions (Wilson and Keeling 2000). This robustness to the change in number of time steps therefore not only reflects robustness of the CLS estimate, but also indicates low-dimensionality of the attractor. This level of robustness of the CLS estimate suggests the approach to be an adequate and objective means of choosing the appropriate scale of observation of an ecosystem.

Two of the ten species sampled showed some sensitivity to the time interval between successive spatial landscapes. Spectra derived from calthropellid (Fig 4.1A) and geodiid (Appendix D Fig B) sponges varied among start dates and different lengths of the time series when samples were taken at 3-month intervals (Fig 4.1A, top 6 plots; Fig 4.2, closed circles). However, when the interval was increased to 6 or 9 months, the variation in CLS estimates indicated by these sponges declined and the length scale became more similar to that indicated by the other species (Fig 4.1B, bottom 4 plots; Fig 4.2, open circles and squares). These two sponge families are thought to be k-strategists, with broadcast spawning (R.Van Soest, personal communication) and are likely to demonstrate strong seasonal patterns in growth, as occurs in other temperate sponge species (Duckworth and Battershill 2001, Tanaka 2002). If seasonality affects the dynamics of these species, the 3-month time intervals may be too short for robust estimation of the CLS. What, then, is the most appropriate time interval to use between the maps? In general, we would suggest that practitioners consider the turnover of species in the system to choose a time interval that allows at least some (~ 20%) but not complete (100%) turnover. In most cases, the CLS estimate will be robust to variations within this time range.

Interestingly, secondary length scales were detectable for the calthropellid sponges (Fig 4.1A, open arrows), reflecting the patch level emergent dynamics of this species (Chapter 3). For this species, these higher-level emergent dynamics are most apparent at a scale near 1.3 m in length, and this scale also seems robust to the changes in the sampling regime. Secondary length scales were apparent for few other species (Fig 4.1 and Appendix D), indicating that the patch level emergent dynamics were often obscured in the time steps sampled. This result

concurs with previous results from model systems that emergent dynamics

resulting from interactions at the patch level are more difficult to detect from

spatial pattern than from time series (Habeeb et al. 2005); therefore, these larger

length scales may prove unpredictable when using the short time series method.

4.5.2 Life histories

The different species we studied demonstrate a broad range of spatio-

temporal dynamics, reflecting a diverse spectrum of life history strategies. No free

space was visible on the wall at any time in the study, signifying that competition

for space is a key process underpinning the dynamics of the > 140 species. The

ten chosen species groups represent life history characteristics ranging from

small, rapidly growing, high-turnover species to slowly growing massive species

(Fig 4.5). Overall, this system is highly dynamic with, on average, an absolute

turnover of 23.4 m in any three-month period (Fig 4.5). The large variation in life

history inferred from the space-time dynamics (Fig 4.4), however, has no

recognizable effect on the CLS estimates obtained from analysis of the different

species. The calthropellid and geodiid sponges (species 2 and 3), which show

more sensitivity to the time interval of sampling than the other species, do not

manifest dynamics that separate them distinctly from the spectrum of life histories

indicated by the other species we examined (Fig 4.4).

Similarly, the life history groupings based on the variance of total cover as a

function of turnover (Fig 4.5) have no apparent relationship to the variations in

estimated CLSs obtained from analyses of the calthropellid and geodiid sponges.

Species with dissimilar life history strategies offer the same approximate CLS, as

the theory predicts for a system of connected species (Takens 1981, Habeeb et al.

2005). Moreover, the two sponge species (species 2 and 3) that showed some sensitivity to the time interval sampled, do not group together. This lack of relationship provides further evidence for the use of the CLS as an ecosystem-level measure. While our ten target species may perform functionally different roles within the ecosystem and have distinctly different patch sizes and/or life history traits, it is encouraging that variations in the CLS estimate among species were small and unlikely to affect decisions about the optimal scale at which to observe the system.

4.5.3 Conclusion

The emergence of robust techniques to estimate characteristic length scale(s) provides a valuable tool to objectively ascertain optimal scales of sampling for identifying dynamical trends in real ecosystems. Results from this paper and our previous work (Habeeb et al. 2005 & Chapter 3) collectively suggest that the short time series technique for estimating CLSs is valid, robust to arbitrary choices in sampling, and suitable for implementation. Only three sequential highly resolved spatial maps of a single species are necessary to estimate the CLS, and our results suggest that the CLS is not sensitive to the choice of the start date in sampling. The time interval should be long enough to capture some but not complete turnover within the system, but will not significantly affect the CLS estimated by most species. As this technique requires a relatively low level of initial assessment, determining optimal sampling scales should be achievable for a majority of ecosystems.

Chapter 5

Estimating natural scales of a coral reef using habitat level data

5.1 Abstract

A new technique to estimate the characteristic length scales (CLSs) of real ecological systems provides a means to identify the optimal scale(s) of observation to best detect the underlying dynamical trends of systems. Application of the technique to natural systems has focused on identifying appropriate scales to measure the dynamics of species as descriptors of ecosystem variability. However, monitoring of ecosystems is often based not on assessing single species, but on species assemblages, functional groups or habitats types. We examined whether the concept of CLSs based on dynamic interactions among species could be extended to determine scales based on interactions among habitat types. A time series of three spatial maps of benthic habitats on a Caribbean coral reef over 21 years was constructed from aerial photographs, Compact Airborne Spectrographic Imager (CASI) images and IKONOS satellite images, providing the short time sequence of maps required for this technique. We estimated the CLS of the system based on the dynamics of three distinctly different habitat types, namely dense stands of seagrass, sparse stands of seagrass and *Montastrea* patch reefs. Despite notable differences in the areal extent of and relative change in these habitats over the 21-year observation period, analyses based on each habitat type indicated a similar CLS of approximately 300 m. We interpret the consistency of CLSs among habitats to indicate that the dynamics of the three habitat types are strongly linked, and suggest that this may be the result of hurricanes as a strong external forcing impacting all habitat types. The results are

encouraging, and indicate that the CLS techniques can be used to identify the appropriate scale at which to monitor ecosystem trends on the basis of the dynamics of only one of a disparate suite of habitat types.

5.2 Introduction

The focus of ecosystem monitoring has shifted in recent years from spatial information based on individual level detail to that based on details of dynamics at a habitat level. This shift has been driven by the emergence of automated technologies such as remote sensing, which provide a means of discerning patterns in ecological systems with continuous data over larger areas than has previously been practical with ground-based field work (Roughgarden et al. 1991, Green et al. 1996). Detailed classification of habitat types using remotely sensed imagery has enabled high-resolution landscape mapping, which is playing an increasingly important role in assessment of biodiversity, reserve design, and management zoning (Loehle and Wein 1994, Green et al. 1996, Ward et al. 1999, Mumby et al. 2001, Stevens 2002). However, the natural scale at which to analyze the maps is ambiguous; what is the optimum scale to distinguish critical ecological trends from noise (Habeeb et al. 2005)? A technique for objectively defining an appropriate scale of sampling habitat maps is necessary for detecting signals in the varying abundances of the habitat types through time (Rouget 2003).

Identifying the level of detail or scale needed to detect the mechanisms that generate patterns has long been an important goal in ecology (Carlile et al. 1989, Wiens 1989, Levin 1992), one that is becoming increasingly pertinent as the scale-dependence of fundamental ecological relationships is established

(Crawley and Harral 2001, Chase and Leibold 2002, Chalcraft et al. 2004). The characteristic length scale (CLS) has emerged as a new tool to estimate objectively the optimal scales of observation to detect underlying dynamical trends in the behavior of real ecological systems, even when dynamics show complex nonlinear oscillatory behaviors (Habeeb et al. 2005, Chapters 3 & 4). Sampling a system at its characteristic length scales maximizes the potential to observe the true deterministic nature of the system least influenced by stochastic variability. Based on nonlinear time series analysis, the estimation of these ecosystem-level length scales requires only a short time series of highly resolved spatial maps. The estimates are largely robust to arbitrary choices about sampling patterns such as the time interval between maps (Chapter 4), and to the choice of species to be monitored, with species across several phyla indicating similar scales (Habeeb et al. 2005 & Chapter 3).

Thus far, characteristic length scales have been detected for a variety of models (De Roos et al. 1991, Keeling et al. 1997, Pascual and Levin 1999, Wilson and Keeling 2000, Habeeb et al. 2005) and most recently, for a real marine system (Chapter 3), but their estimation has always been based on the dynamics of component species. Here we examine, for the first time, whether larger natural length scales are manifest that reflect spatial dynamics among habitat types, as opposed to among species. If such scales exist, then their detection could be extremely valuable when monitoring of community or ecosystem change is based on changes in the extent and/or location of habitats. Monitoring habitat dynamics with landscape maps derived from increasingly available remote sensing techniques can be useful because habitats are often good proxies for more finely resolved information that is difficult to collect (Mumby et al. 2001, Rouget 2003).

Here, we use data from remote sensing to search for characteristic length scales at the habitat level.

Using spatial maps of benthic habitats within a Caribbean coral reef over 21 years, we attempt to determine the CLSs of coral reef habitats. While only one habitat type is necessary to estimate length scales, we classify the spatial maps of three different lagoonal habitats to determine whether they provide similar estimates, if any, of the CLSs for the reef ecosystem. If they exist, these scales will be useful only if disparate habitat types indicate similar scales (Habeeb et al. 2005).

In substituting habitat data into the nonlinear time series method of Habeeb et al. 2005, the underlying assumption is that the attractor of the system can be reconstructed from a short time series of habitat level information. This assumption is reasonable if changes in the habitats reflect the deterministic dynamics of the system (Ward et al. 1999).

5.3 Methods

5.3.1 Data requirements

To estimate CLSs, at least three consecutive maps of the same area are required. The time intervals between these maps can vary, but should be long enough to allow between 20% and 80% turnover in the species or habitats sampled (Chapter 4). On a coral reef where massive coral colony sizes may exceed 3 m^2 (Meesters et al. 2001), intervals need to be of the order of years to decades rather than months, to allow large-scale habitat change (Lewis 2002, Pandolfi 2002). For each map, the data need to be spatially contiguous, with no breaks in the landscape as, for example, might arise with cloud cover. Finally, to

rebuild the attractor with the short time series method, the data must have high

spatial resolution relative to the entire landscape mapped.

5.3.2 High-resolution maps from the Turks and Caicos, West Indies

The study site chosen was a shallow reef ecosystem near Cockburn Harbour

within the Turks and Caicos Islands, West Indies (Fig 5.1), whose reefs have been

mapped intensively (Mumby et al. 1998a and b, Mumby and Edwards 2002). An

aerial photograph from 1980 (2 m resolution), a Compact Airborne

Spectrographic Imager (CASI) image from 1995 (1 m resolution) and an

IKONOS satellite image from 2001 (4 m resolution), all of the same area,

provided the short time sequence of spatial maps necessary for CLS estimation.
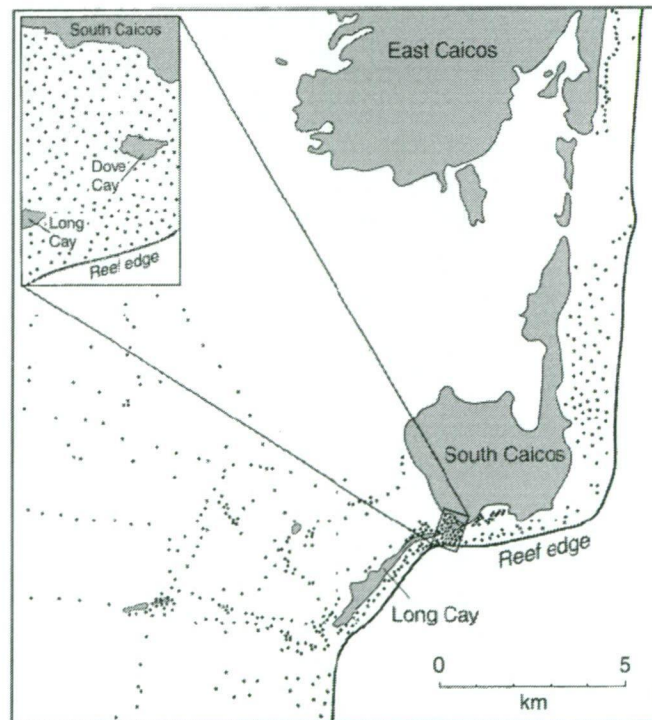
Figure 5.1. Map of the Turks and Caicos Islands, West Indies. The inset
shows the study area on the southern side of South Caicos Island.

The benthic habitats captured in these images comprised reef-building corals,

gorgonian communities, seagrass beds, and algal dominated substrate within a

depth range of 3 – 25 m. Three maps, each with an area of approximately 1.4 km^2

were constructed (Fig 5.2), based on the classification system of Mumby and

Harborne (1999). Standard image processing was achieved through the ERDAS

image software. Images were geometrically corrected and co-registered, and

supervised classification, which was directed by field surveys of 60 sites within

the harbor, was applied to spectral bands.
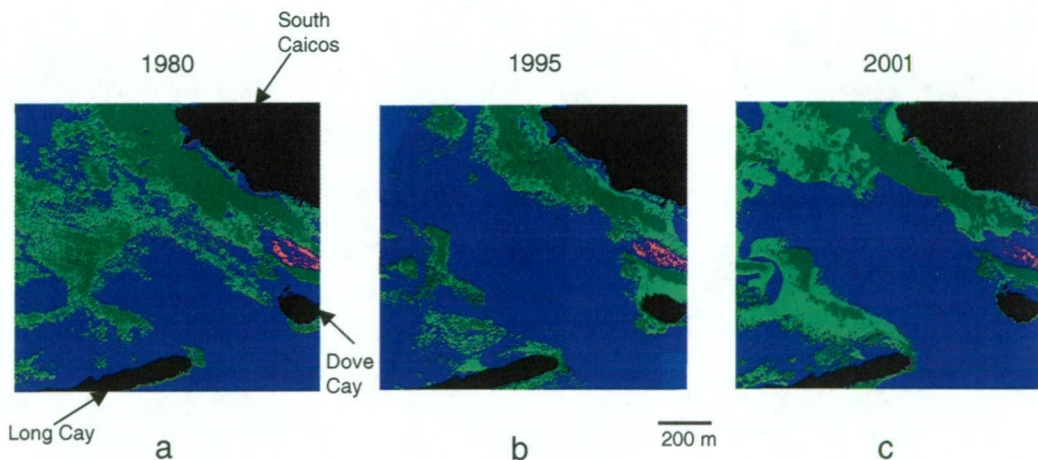
Figure 5.2. Digitized maps of the benthic habitats at Cockburn Harbour (South Caicos) taken from (a) 1980 aerial photography (b) 1995 CASI imagery, and (c) 2001 IKONOS imagery. Dark green regions indicate areas of dense seagrass, light green regions indicate areas of sparse seagrass, red regions indicate *Montastrea* patch reef, blue indicates deep ocean and unclassifiable reef habitats, and black regions indicate landmasses, which are named for reference in (a). Each map has an end resolution of 4 m pixels.

Because of their different pixel sizes, the images were re-scaled to the

lowest resolution of 4 m pixels by taking means. This provided a conservative

estimate of resolution, which accounted for most of the error in the rectification of high-resolution airborne imagery (errors of ~ 3 m).

Three distinct habitats were examined in Cockburn Harbour. The first, aggregated colonies of the massive coral *Montastraea annularis* (*senso stricto*), bridge the zone between the main, hard-bottom forereef and the softer-sediment lagoon habitats in the harbor (which included the other 2 habitat types). The *Montastraea* zone straddles the main inlet to the harbor and attracts many fish because of its high structural complexity (the diameter of individual colonies often exceeds 1.5 m and is the result of decades to centuries of growth). The second and third habitat types, dense and sparse seagrass beds, are found within hundreds of meters of the *Montastraea* zone and are dominated by two seagrass species; *Thalassia testudinum* and *Syringodium fileforme* (see Mumby et al. 1997). The standing crop of the dense beds ranges from 51-230 g m^{-2} with a mean species composition of 72% *Thalassia* and 28% *Syringodium*. The low-density beds typically have 80% *Thalassia*, 20% *Syringodium* and a lower standing crop of 1-50 g m^{-2}. These two seagrass habitat types have been shown to be distinct, harboring different assemblages of species (Mumby et al. 1997).

5.3.3 Estimating the CLSs of reef habitats

The short time series technique (Habeeb et al. 2005) was employed, using the sequence of three maps created for the lagoon habitats to reconstruct the attractor and generate variance spectra for each habitat type (n = 3). Primary CLSs were estimated as the first plateau of the prediction r^2 versus scale spectrum, and secondary CLSs were detected where subsequent peaks in prediction r^2 were evident.

5.4 Results

5.4.1 Habitat level CLSs of a coral reef ecosystem

Prediction r^2 versus scale spectra were produced from analysis of each habitat type, and primary and secondary CLSs were determined by visual inspection (Fig 5.3). Each habitat type displayed a spectrum of the expected shape, with prediction r^2 sharply increasing to a plateau as the scale of observation increased. The CLS was interpreted as the scale beyond which the slope shallows, or virtually no increase in prediction occurs. Primary CLSs estimated by all three habitats were similar at approximately 300 m. Secondary CLSs, indicative of patch level emergent dynamics (Habeeb et al. 2005 & Chapter 3), were evident when either type of seagrass habitat was used to reconstruct the attractor, with both habitats indicating a secondary length scale at approximately 600 m.

5.5 Discussion

5.5.1 CLS estimation using habitat level data

The choice of sampling scales in ecology is critical for clear detection of ecosystem trends. Results can be clouded by noise if the sampling scale is too small, and signals averaged out if the scale is too large (Pascual and Levin 1999, Wilson and Keeling 2000). The characteristic length scale is a quantitative tool that offers an objective estimate of the sampling scales, which best identify the nontrivial trends in an ecological system. Here, we establish that CLSs can be estimated using data collected at the habitat level. This result demonstrates the versatility and potential applicability of this new estimate for use in ecosystem
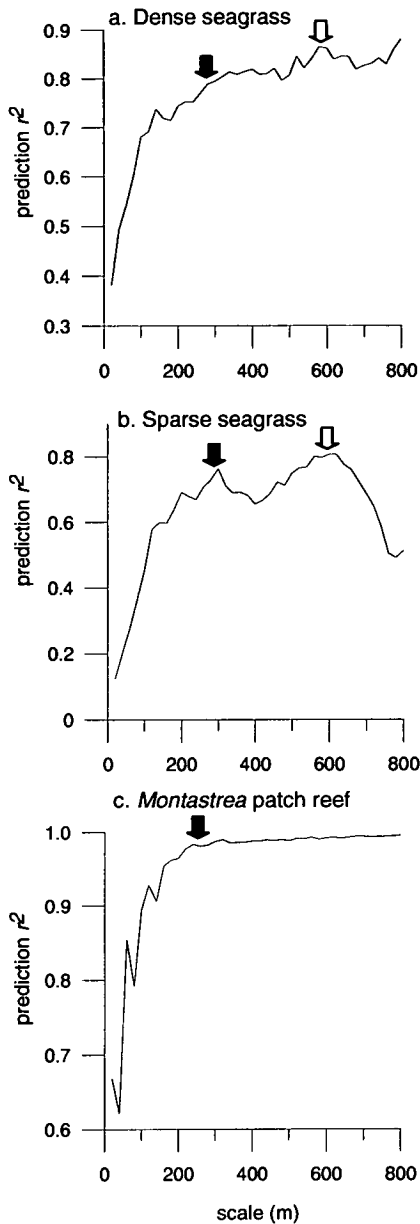
Figure 5.3. Spectra of prediction r^2 as a function of scale determined using the short time series method of attractor reconstruction, based on abundances of (a) dense seagrass, (b) sparse seagrass, and (c) *Montastrea* patch reef. Filled arrows indicate the approximate primary CLS, and open arrows indicate secondary CLSs.

monitoring that employs remote sensing to map systems at the habitat level. Once landscape maps are acquired for a system, the scale at which to analyze those maps must be chosen. Thus, establishing a guide to defining the appropriate scales of observation at which the deterministic dynamics of habitat change are most clearly in focus is a significant advance in ecology. The choice of scale at which to analyze habitat maps can now be guided by the CLS of the system.

Furthermore, if shortages in funding prevent the comprehensive sampling of a region, knowing the optimal size of sampling units for detecting trends in habitat abundances as determined with the CLS technique will be valuable (Pressy and Logan 1998, Rouget 2003).

All three habitats provided interpretable spectra of prediction r^2 versus sampling scale (Fig 5.3), and these spectra yielded similar length scale estimates (ca. 300 m). This similarity occurred despite vastly dissimilar aeral coverages and habitat dynamics over the 21-year observation period (Fig 5.2). Thus, it appears that habitat change is indicative of the underlying dynamics of the system over the time scale assessed. The analysis indicates that sampling landscape maps of this reef at a scale of ca. 300 x 300 m (0.09 km^2 in area) will yield data reflecting the biologically significant dynamics of the system at the habitat level. These results clearly demonstrate that natural scales can be estimated using spatial information not only based on species abundances, but also based on habitat abundances, which will be useful for monitoring the dynamics of habitat level change.

5.5.2 Similarity of the habitat estimates

Given that the 3 habitat types studied here are comprised of taxonomically and functionally distinct species, manifest large differences in absolute cover, and change distinctly in relative spatial extent and arrangement (Fig 5.2), it is encouraging that all habitats indicated a similar primary CLS at ≈ 300 m. This result suggests that the 3 habitat types are components of a connected dynamic on this landscape, thereby allowing estimation of the system-level length scale. We suggest that at least part of this connected dynamic reflects that the habitats are subject to similar disturbance regimes. Five tropical storms or hurricanes were

recorded in the study region during the 21 years of investigation

(http://hurricane.csc.noaa.gov/hurricanes/viewer.htm). Hurricanes are major

sources of disturbance to corals (Done 1992, Bythell et al. 2000) and may

physically remove moderate-sized massive corals such as *Montastraea* (Massel

and Done 1993). Likewise, recent work on comparable seagrass beds from

elsewhere in the Caribbean illuminates the role of hurricanes in seagrass

dynamics. For example, hurricanes may facilitate the spread of *Syringodium* by

increasing seed dispersal (Kendall et al. 2004) and freeing space for growth of all

habitat types (Fourqurean and Rutten 2004). The consistency of CLSs among

habitats suggests that hurricanes may bring about similar deterministic dynamics

for strikingly different habitats when these habitats occur in close proximity.

5.5.3 Secondary characteristic length scales

Spectra of both dense and sparse seagrass habitats indicated secondary

CLSs at approximately 600 m, twice the primary CLS (Fig 5.3). Conversely, the

spectrum for the *Montastrea* reef patches reached a plateau after the primary CLS

and remained constant with increasing scale, indicating that detection of the

secondary CLS is dependent on the habitat studied. This secondary scale likely

reflects the large-scale loss and subsequent return of seagrass beds over the 21-

year time period (Fig 5.2), dynamics that were not evident for the more consistent

Montastrea patch reef. Secondary length scales may not be useful as sampling

scales for monitoring due to their unpredictable nature and their apparent

dependence on the habitat sampled, but when evident, these secondary CLSs are

consistent in magnitude and likely indicate the more patch level emergent

phenomena occurring within the system (Habeeb et al. 2005).

5.5.4 Importance of estimating CLSs at a habitat level

Natural scales may be determined for reefs and other ecosystems worldwide using this method in combination with increasingly available archives of satellite images (e.g. for coral reef satellite images: http://seawifs.gsfc.nasa.gov/cgi/landsat.pl), providing for the first time an objective means of scale determination by which to sample habitat maps for real trends. Such large-scale implementation of the method may enable the use of CLSs not only as observational scales, but also as baselines in themselves to detect dynamic shifts in ecosystems. If CLSs are calculated regularly for a given system, then a change in the length scale from one time to the next is likely to indicate a shift in the attractor (Habeeb et al. 2005), warranting further investigation into the cause of such a fundamental shift.

In addition to providing the optimal scales of sampling landscape maps for detecting trends, CLSs estimated according to habitat changes within a system might be compared in themselves to potentially identify functionally linked areas and watersheds for management. Within a single system, those habitats that provide similar CLSs are likely governed by similar processes, whereas habitats with distinctly different CLSs are unlikely to be linked dynamically, either by 'internal' patterns of interaction or by external forcing such as disturbance. Similarly, comparing habitat level CLSs between regions may also be beneficial. Similar habitat types in different regions that indicate vastly different CLSs may reveal fundamental differences in the dynamic processes influencing the two systems, such as anthropogenic effects in one region that do not occur in the other. Using the CLS as an index in and of itself for these and other comparisons

may provide valuable clues about the underlying dynamical behaviors of the complex systems being monitored.

5.5.5 Conclusions

We have demonstrated that natural length scales can be estimated using data taken at the habitat level rather than at the species level. Distinct, but spatially connected habitats provided similar length scales, reflecting the ecosystem nature of the estimate and the functional linkage of the habitats. This development provides new possibilities for large-scale use of this quantitative estimate in ecology, as ecosystem monitoring is often based on assessment of suites of species or habitat types rather than on individual species, and as data requirements are more easily met with long-term sequences of remote sensing imagery than with local, field-based data. Estimating the characteristic length scale of a landscape, whether terrestrial or marine, will enable sampling at appropriate scales for detecting dynamical ecosystem trends.

Chapter 6

Influence of species connectivity on characteristic length scale estimates

6.1 Abstract

The choice of scale at which to observe an ecological system is critical. Recently, quantitative methods for choosing optimal sampling scales have been developed, with the underlying theory that observations of only a single species in the system are necessary for estimation of the system-level characteristic length scales (CLSs). However, in spatial model communities where the species are poorly connected either due to the nature of the network topology or to the spatial self-structuring which results in species rarely interacting, different species can indicate dissimilar length scales. Here we examine the role of poor connectivity in producing dissimilar CLS estimates between the constituent species of a system. We define connectivity in terms of the network topology and the interaction strengths. In a simple spatial model of six interacting species, a highly connected system is one in which all species interact with the same strength and at the same rate, and weak connection between two groups occurs when their interactions, such as overgrowth in a competition model, are infrequent, weak or non-existent. We use spatial models of varying connectivity to generate long time series for each species within the system, and then estimate CLSs. We demonstrate that for a simple model system, when species groups are weakly connected in their network interactions, the CLS estimated can depend on the species used to reconstruct the attractor. Models of loosely connected species groups were more likely to provide CLSs that were distinct between the groups, while CLSs

estimated from strongly connected groups were always clearly similar. Thus, when utilizing the new CLS method on a system where species connectivity is particularly low, different length scales may be indicated by the species, thereby necessitating determination of CLSs with several representative species. This result may prove beneficial, however, in that CLSs may potentially be useful as qualitative indicators of poorly connected species within an ecological system.

6.2 Introduction

Characteristic length scales (CLSs) are, according to theory, inherent properties of dynamical systems and are not features of the individual component 'agents' that comprise a dynamical system (Takens 1981). In the context of ecological systems, this means that the length scale of a particular system should be independent of the species that is observed to calculate the length scale (Habeeb et al. 2005). The implicit assumption underpinning this notion is that all of the component agents of a dynamical system are well connected, either directly or indirectly. While this may be the case for many physical systems, it is unlikely to be true for complex ecological systems in which the dynamics of some species are likely to be strongly linked to some subset of species in the system but not to others (e.g. Dunne et al. 2002). This raises the possibility that the CLS of an ecological system estimated from one particular species x would be similar to that estimated from the dynamics of other species in the system with which it was well connected, but dissimilar to that calculated from species whose dynamics are largely independent of species x.

The results of a recent study to determine the length scale a natural marine fouling community growing on a subtidal rock wall lend support to this notion

(Chapter 3). It was found that most species, representing a diverse spectrum of life histories, indicated very similar system-level length scales (0.20-0.45 m, average ca. 0.35 m), while a single species indicated a CLS considerably larger (0.60 – 0.70 m). Results of Chapter 3 suggested that the divergent CLS estimate of this single species may reflect its lack of connectedness to the other species that were analyzed in the community, arguing that this species may be less influenced by the availability of unoccupied space than others (Habeeb et al. in review).

There are several ways to define connectivity among species in ecological systems, ranging from interaction coefficients that define per capita effects of one species on the growth and reproduction of another, to the landscape ecology concept which combines the effects of landscape structure and species' behaviors to define movement rates among habitats (Tischendorf and Fahrig 2000). Here we are concerned with the likelihood and outcome of direct interactions between each pair of species competing for space on a landscape (Anderson and Jensen 2005). We, therefore, define connectivity by (1) the network topology describing where interspecific interactions exist, (2) the strength of interactions, and (3) the likelihood of species' co-occurrences in space as defined by their relative abundance and patterns of spatial organization on the landscape.

We use simple spatial models of six interacting species, and examine whether there is any relationship between connectivity among species and the magnitude of the length scales they indicate for the same system. If there is, then comparison of system-level length scales estimated from the dynamics of different species may provide a proxy for connectivity among species. In this context, a highly connected system is one in which all species show strong

interactions with all others and at the maximum possible rate. A weak connection

between two species occurs when their interaction is infrequent and weak. We

estimate CLSs based on the original method of Pascual and Levin (1999), which

is robust to varying model complexity (Chapter 2, Habeeb et al. 2005).

6.3 Methods

6.3.1 Spatial models

Models of varying connectivity, defining the number and strength of direct

and indirect interspecific interactions in the system (e.g. Johnson and Seinen

2002) were developed using the Compete© software (available at

http://www.zoo.utas.edu.au/CJPblist/PubListCJohnson2.htm#compete). This

probabilistic individual-based system models spatial competition between sessile

colonial organisms, follows the fate of competing individuals in a 2-dimensional

landscape, and demonstrates complex behaviors indicating nonlinear dynamics

and spatial self-organization (Johnson 1997, Johnson and Seinen 2002). We

studied two model systems with different network topologies describing

interactions among 6 species (Fig 6.1). Both systems spatially self-organize to

form two distinct groups of 3 species on the landscape. In the first, less connected

system (denoted 1:1), each species overgrows one species within its group, and

one species from the other self-organized group (thus $S_1 > (S_2, S_5)$, $S_2 > (S_3, S_6)$,

..., $S_6 > (S_4, S_1)$; with $S_x > S_y$ indicating that species x outcompetes and displaces

species y, and where S_1, S_2, S_3 and S_4, S_5, S_6 form the two spatially distinct groups).

In the second, more connected system (denoted 1:3), each species overgrows one

species within its group, and all three species from the opposite group (thus $S_1 >$

(S_2, S_4, S_5, S_6), $S_2 > (S_3, S_4, S_5, S_6)$, ..., $S_6 > (S_4, S_1, S_2, S_3)$). Thus, in this model

The 1:1 model system
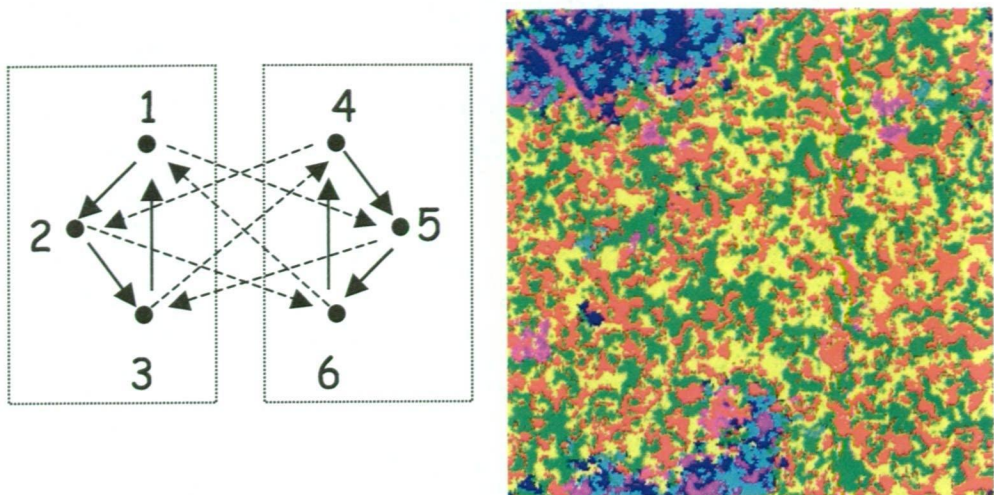
The 1:3 model system
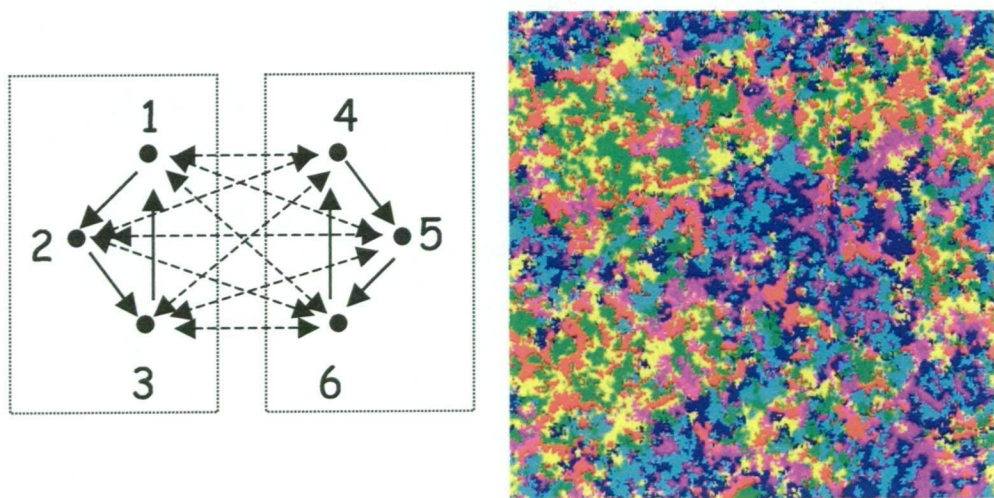
Figure 6.1. Interactive networks for the 1:1 and 1:3 model systems, and representative landscapes from each model system (generation 10,000). Species are denoted by numbers, and boxes indicate the distinct groups that are formed. Species within each group are equivalent. Dark-lined arrows indicate interaction between two species within the same spatially self-organizing group, and dotted arrows indicate interaction between two species within different groups. Arrowheads indicate the direction of overgrowth. Note that in the 1:1 model system, all dotted arrows indicate overgrowth in only a single direction, while in the 1:3 model system, all dotted arrows have double-headed arrows, indicating overgrowth in both directions.

system, each species in one group can overgrow, and be overgrown by, all species in the other. Within these qualitative network structures, connectivity was further varied by changing the rates and probabilities of overgrowth between species both within and between groups. In both models, ties (or mutual overgrowth) were randomly split.

All model systems were run for 10,000 time steps on landscapes of 300 x 300 cells, and landscapes were sampled from times 201 to 10,000 (the first 200 time steps were ignored while the system self-structured). Each model was begun with a random spatial arrangement of 'recruits' (10% total cover) on the initial landscape. The system updates synchronously, and uses periodic (toroidal) boundary conditions.

6.3.2 Levels of connectivity

The two network topologies define two 6-species systems with different degrees of connectivity among species (Fig 6.1). Within each of these systems connectivity among species was varied by changing the strengths of non-zero interactions. Probabilities (or rates) of overgrowth of species in one spatially self-organized group by those in the other were increased from 0.2 to 0.4 to 0.8, reflecting an increase in the groups' connectedness to each other. In addition, probabilities of overgrowth among species within each spatially self-organized group were either set to be identical in the two groups (both groups at 0.8) or set to be different between the two groups, with species in one group overgrowing each other with a probability of 0.8, and in the other at either 0.4 or 0.2. The net effect of species in one group overgrowing each other at the same rate that species in the other group overgrow each other is that both groups of species form 'spirals' that rotate at the same rate. We predict that similar within group

Factors (# of levels)	Levels
1. Network topology (2)	1:1, 1:3
2. Between groups overgrowth rates (3)	0.2, 0.4, 0.8
3. Within groups overgrowth rates (3)	Same (0.8-0.8), Different (0.8-0.4, 0.8-0.2)

Table 6.1. The factors, and levels within each factor, used to vary connectivity between model runs. Overgrowth rates are given as probabilities. See text for an explanation of the 1:1 and 1:3 network topologies. System-level CLSs were determined from the dynamics of every species for all 18 possible combinations of the 3 factors.

overgrowth rates provide more highly connected groups than dissimilar rates, as the rates dictate the rotation of the patches through the landscape. When groups spiral at the same rate, we predict that they are more likely to encounter those species (from the other group) with which they interact.

Eighteen scenarios of varying connectivity among species in the 6-species systems were evaluated for CLS estimates (Table 6.1). In the context of this paper, the most connected system is the 1:3 system with each species overgrowing within and between groups at a probability of 0.8, and the least connected system occurs in the 1:1 model with between groups overgrowing 20% of the time, and within the groups overgrowing at different probabilities of 0.8-0.2. Note that there are several aspects that define the realized connectivity among species, namely the network topology and interaction strengths, as just described, but also the nature of spatial self-organizing, which cannot be controlled directly. Spatial structure is important because it can determine the likelihood that two

species are in spatial proximity and are able to interact at all.

6.3.3 The 1:1 model stability

The space-time trajectories of the 1:1 model system indicated a shifting attractor, often with one group overgrowing the other to extinction after 20,000 time steps. As in previous chapters we have shown that species following nonstationary attractors often indicate different CLSs, and in fact that CLS estimates might be used to identify shifting attractors, this became a confounding factor. Thus, we stabilized the model by adding low levels of recruitment (probability of influx = 0.05, < 2 pixels of recruits of each species for each time step) and disturbance (probability of 0.0001 with an area of 25 pixels), so that we could unambiguously attribute any differences in CLS estimates to differences in connectivity. The stabilized model was then run with the previously described interaction strengths to assess the change in CLS due to changes in connectivity. Additionally, within group rates of 0.2-0.2 were assessed to determine if the similar rotation rates of the spirals formed would produce similar CLSs even when connectivities were weak.

6.3.4 CLS estimates

Time series for each species within the model were generated and analyzed for length scale with the original long time series method of attractor reconstruction, which is the method most appropriate for use with model output (Habeeb et al. 2005, Pascual and Levin 1999). Attractor reconstruction parameters were chosen as per Habeeb et al. (2005). If CLS estimates reflect connectivities among species in the system, then we expect that strongly connected species in a

system will yield more similar CLSs than those estimated from the dynamics of weakly connected species.

6.4 Results

Prediction r^2 spectra were produced from the space-time dynamics of each species within each model scenario (Fig 6.1), and the resulting CLSs compared. All species within spatially self-organized groups indicated similar CLSs in all of the models examined. Species from different self-organized groups within the 1:3 system, where each species overgrew all species within the opposite group, consistently indicated more similar length scales than those within the 1:1 system, where species overgrew only one species within the opposite group. Notably, groups of species in the 1:1 system formed more distinct self-organized patches than groups in the 1:3 system. Little variation in the size of self-organized patches was evident with changes in interaction strength for the 1:3 model. Patch sizes in the 1:1 model system varied, but changes were not systematic with the changes in within group probabilities, and may simply reflect the between run variation.

6.4.1 The 1:1 system

When species in one self-organized group overgrew only a single species from the other group, prediction r^2 spectra produced by models of all levels of connectivity separated into two distinguishable sets of curves that corresponded to the self-organized patches, reflecting that with this network structure, the technique is sensitive to the formation of two groups of species (Fig 6.2). In some instances these curves, though separated, did not indicate different CLSs (e.g., Fig 6.2, within = 0.8-0.8 and between = 0.8), but in most cases the CLS estimates

The 1:1 model system

Between group overgrowth rate

Figure 6.2. Prediction r^2 spectra for the 1:1 model system, with varying within group probabilities (down) and between group probabilities of overgrowth (across). Arrows indicate the CLSs estimated for each run. In this system, each species in a group overgrows 1 species within its own group and 1 species within the other group. Note that models self-organize at several spatial scales. All model runs used random initial configurations of recruits covering 10% of the 300 x 300 landscape at time step zero.

were different between the two groups.

No significant effect of varying the between-group overgrowth rates was

evident. When the between-group overgrowth was 0.4, CLS estimates between the two groups were most dissimilar, with one set of curves indicating a CLS of approximately 30 cells and the other set indicating a CLS of 150 cells (with overgrowth rates of species within both groups at 0.8, Fig 6.2). When the between group rates were increased to 0.8, CLSs between the groups remained dissimilar rather than converging (Fig 6.2).

As expected, some effect of changing from similar to dissimilar within group overgrowth rates was observed. When the two groups were connected by a 0.8 probability of overgrowth, for example, and the within group probabilities were identical, CLSs for both groups were similar at approximately 30 cells. CLSs then separated to 30 cells versus 70 cells when they were dissimilar at 0.8-0.4, and 0.8-0.2 (Fig 6.2). Such was also the case when between groups overgrew at 0.2 (Fig 6.2). However, the results were different when between groups overgrew at 0.4 (Fig 6.2). Dissimilarity between the two groups' CLS estimates could not be conclusively attributed to varying the within group overgrowth rates, as when the between groups overgrew at 0.4, CLS estimates of the two groups were very different both when within groups were most similar and when they were least similar (Fig 6.2).

6.4.2 The 1:3 system

When the groups were linked more closely in the network topology, i.e., each species overgrowing all species in the other group, CLS estimates were identical for both groups, regardless of the between and within group interaction strengths (Fig 6.3). In most cases, prediction r^2 spectra overlapped and curves of the two groups were visually indistinguishable. In two cases, when the between

The 1:3 model system
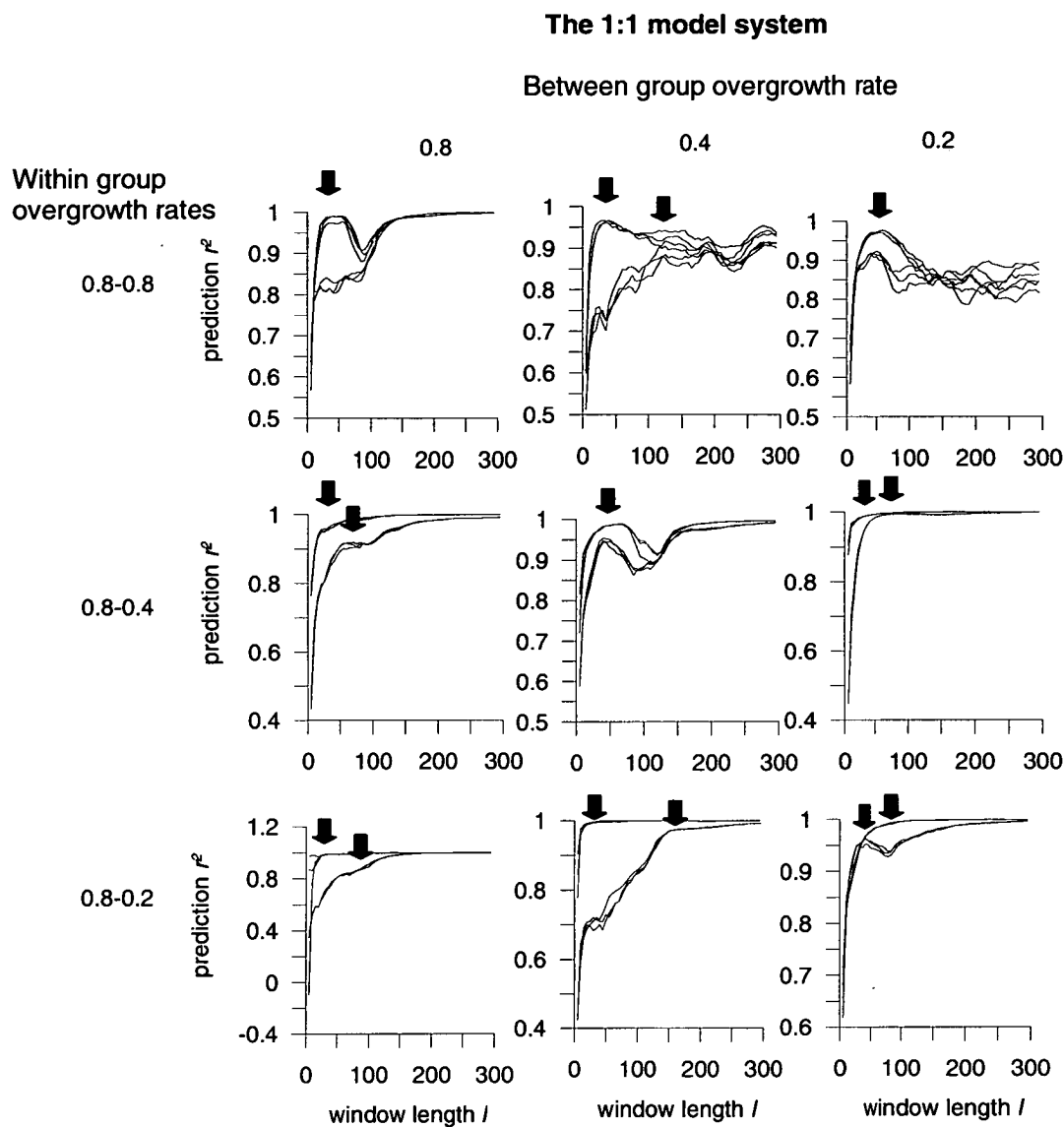
Between group overgrowth rate

Figure 3. Prediction r^2 spectra for the 1:1 model system, with varying within group probabilities (down) and between group probabilities of overgrowth (across). Arrows indicate the CLSs estimated for each run. In this system, each species in a group overgrows 1 species within its own group and all 3 species within the other group. Note that models self-organize at several spatial scales. All model runs used random initial configurations of recruits covering 10% of the 300 x 300 landscape at time step zero.

groups overgrew each other at 0.2 and 0.4 and within groups interacted at the same rates, the curves separated for the two groups, but the CLS estimates

remained identical (Fig 6.3).

6.4.3 The 1:1 system stabilized

When the attractor of the 1:1 model system was stabilized with low levels of recruitment and disturbance, the prediction r^2 spectra remained similar to those from the unstable system (Figs 6.2 & 6.4). Estimates between the two spatially self-organizing groups were always dissimilar (Fig 6.4) when the within group overgrowth rates were different between the two spatially self-organized groups (0.8-0.2, for example), but were more likely to converge when rates were similar (0.8-0.8, for example). When within group connectivity was lowered to 0.2 for both groups, estimates of CLS remained identical for the two groups (Fig 6.4, bottom row). Again, we found that between group overgrowth rates had little effect on the CLSs.

6.5 Discussion

Implicit in the theory of attractor reconstruction (Takens 1981) is that any species can be used to reconstruct the attractor, and therefore that any species in a system will indicate the same CLS, irrespective of its life history parameters. This theory assumes that all species in a system are connected. However, we have shown that different species can indicate dissimilar length scales. For a simple model system, when species groups are weakly connected in their network interactions, the CLS estimate can depend on the species used to reconstruct the attractor. In models where species self-organize into discrete groups where there are weak connections among species between groups, and where the different groups have dissimilar dynamics, then species in the different groups indicated

The 1:1 model system stabilized
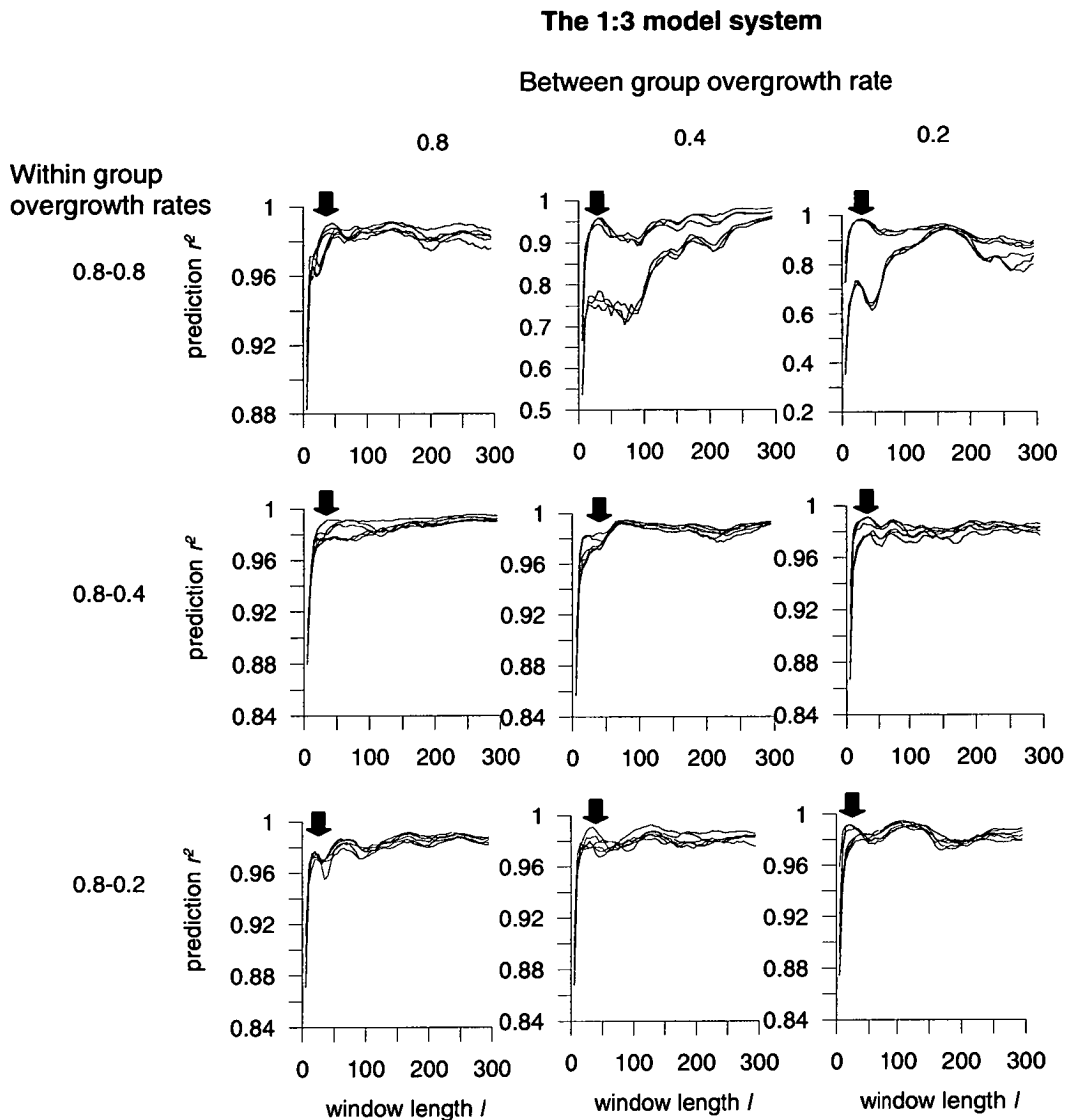
Between group overgrowth rate
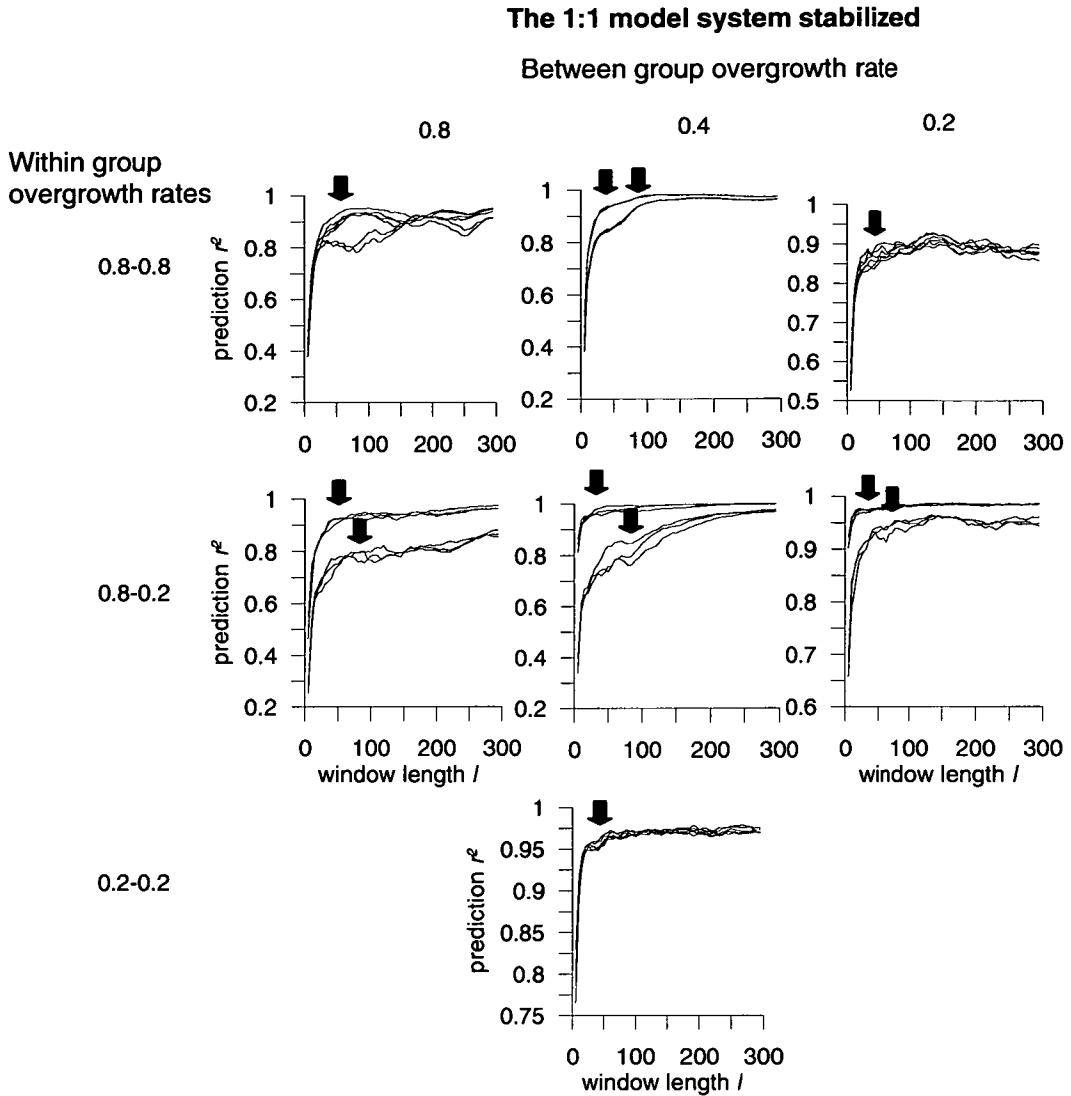
Figure 4. Prediction r^2 spectra for the stabilized 1:1 model system, with varying within group probabilities (down) and between group probabilities of overgrowth (across). To stabilize the nonstationary attractor, low levels of recruitment and disturbance were added to the model. Arrows indicate the CLSs estimated for each run. Note that only the 0.8-0.8 and 0.8-0.2 within group probabilities are shown, as before, with the additional 0.2-0.2 interaction inserted. All model runs used random initial configurations of recruits covering 10% of the 300 x 300 landscape at time step zero.

dissimilar length scales. In contrast, CLSs estimated from species that are strongly connected were always similar. As predicted, in the most connected system (our

1:3 system with each species overgrowing within and between groups at a

probability of 0.8) all species provided similar CLS estimates, while in the least

connected system (1:1 model with between groups overgrowing at 0.2, and within

groups overgrowing at different probabilities of 0.8-0.2) the CLSs depended on

the identity of the group from which the species for analysis was selected. This

result provides preliminary evidence that different 'sub communities' in a system

that have dissimilar dynamics may require different scales of observation to

optimally define those dynamics.

There are several definitions of connectivity, and one of the simplest is

characterized by the network topology and its interaction strengths (Anderson and

Jensen 2005). Here, network topology had a greater influence on the length scales

estimated by species groups than did interaction strengths. Different spatially self-

organized species groups that are highly connected in terms of the number of

links between the groups indicate similar CLSs, while species from groups linked

by only a single connection between each species are more likely to indicate

dissimilar length scales. When the interaction strengths within the 1:3 model

network were varied, all species indicated similar CLSs, implying that species do

not need strong associations to indicate similar scales. The presence of such

frequent interactions between species groups holds the groups on the same

dynamical attractor, therefore only a single CLS is indicated.

Previous results have shown that when systems are nonstationary with one

group overgrowing the other to eventual exclusion, the CLSs indicated by the

groups are disparate (Chapter 2, Habeeb et al. 2005). Such instability, or

nonstationarity, can influence the connectivity between species groups (Haydon

2000), thus, it was necessary to stabilize the 1:1 model system with very low

levels of disturbance and recruitment. In the 1:1 network, prediction r^2 spectra

separated and species from the two different groups indicated dissimilar CLSs when the within group overgrowth rates were different. This is likely due to both the interaction strengths, and the spatial self-organization of this loosely connected system. Species within patches were more likely to interact with each other than with species from the other patches simply because of their spatial proximities. When the two distinct groups rotated at similar rates, as dictated by similar within group overgrowth rates, either at high levels of 0.8-0.8 or at low levels of 0.2-0.2, CLS estimates were similar across all species. This result suggests that the similarity of the rotation rates of the spirals more strongly affects the CLS estimates than does the strength or magnitude of those rates. The influence that spatial self-organization has on the CLS estimate might in the future be differentiated from the other aspects of connectivity (topology and interaction strengths) by comparing results based on species from spatially explicit models to those from mean field models, using identical network topologies and interaction strengths.

This research provides preliminary evidence that connectivity can influence the CLSs estimated by different species within a system. Is the CLS, then, a system-level measure? While only a single run of each model was evaluated, broad patterns were evident. In systems where species are connected and interacting on the landscape, the CLS will remain similar for all species and can be treated as a system-level measure. Indeed, the CLS has been established as a measure robust to distinctly different species used to reconstruct the attractor (Chapter 3, Habeeb et al. in review). However, CLSs are not always independent of the species used in their estimation, and these results indicate that low connectivity within the network topology and in spatial proximity due to self-

organization may contribute to disparity in estimates between species. When connectivity is extremely poor, different CLSs will likely arise due to divergence of the 'system' components, i.e., when the sub-communities begin to act as different systems. Thus, the CLS is indeed a system-level measure, but caution must be applied when estimating length scales of systems with potentially divergent dynamics. Such systems will likely need to be observed on different spatial scales.

In addition to using characteristic length scales to uncover the deterministic dynamics of a system (Pascual and Levin 1999), this research opens the possibility of using CLSs as a proxy for connectivity. Though the prediction r^2 curves may separate when connectivity changes (e.g., Fig 3, within = 0.8-0.8 and between = 0.4), CLSs estimated by different species remain similar until species connectivity is particularly weak (e.g., Fig 2, within = 0.8-0.8 and between = 0.4 but for the less connected system). This result demonstrates that the CLS is robust to slight variation in connectivity, and therefore may provide an indicator of the extremely disconnected species within a system, i.e., the species that show disparate CLSs. Such species might be following a trajectory different from that of the system of interest. In practice, CLSs might be estimated for a system using data from all species, and any species providing dissimilar estimates can be flagged as the weakly connected species. Using CLS estimates as such indicators might also enable assessment of habitat connectivity (Shumaker 1996) and identification of weakly connected habitats by comparing the CLSs estimated by different habitats (Chapter 5). This preliminary research, although limited in scope, provides an initial framework for exploring the relationship between the system-level measure of spatial scale and species connectivity, and extrapolations

of this work to identify the weakly connected agents, whether species or habitats,

within the system may be further explored in the future.

Chapter 7

General Discussion

7.1 The spatial scale of sampling

All scientific observations must be made at some spatial scale, and the selection of this scale usually requires little justification. For decades researchers have attempted to uncover the optimal scale at which to make observations of system dynamics (Greig-Smith 1964, Carlile et al. 1989, Turner et al. 1991, Rand and Wilson 1995), but few methods have been appropriate for application to the complexity of real ecological systems. Though systematic methods for defining spatial scales have long been recommended (Wiens 1989), and an increasing awareness of scale dependence has highlighted the importance of the choice of scale (Levin 1992, Crawley and Harral 2001, Rietkerk et al. 2002), sampling and monitoring scales continue to be defined intuitively, leading to potential bias and misunderstanding of variability within the system (Denny et al. 2004).

Not only does changing the scale of aggregation reflect a tradeoff between detail and statistical predictability as we move from observing individual variation to collections of behaviors that can be generalized (Levin 1992), but ultimately, the patterns and the mechanisms generating and maintaining those patterns that we detect depend on the scale at which the system is viewed (e.g., Chase and Leibold 2002). The same question addressed at different scales can provide different results (Qi and Wu 1996, Wu et al. 2002, Hill and Hamer 2004), which then have to be teased apart; thus, the choice of scale is critical. Even the most sophisticated statistics cannot mitigate noisy results based on data collected at an inappropriate spatial scale, and conflicting results based on data collected at

multiple scales are not only expensive to acquire but difficult to interpret, as the scale dependence of the variables renders indistinguishable the real underlying trends from the variability due to noise.

A more refined and objective approach to choosing the spatial scale is obviously necessary. Are there appropriate, natural scales of observation for a given system, such that trends can most likely be detected? Natural scales do exist; they are defined by the domain of attraction, or space through which a system rotates in time, and this research has focused on establishing and evaluating a technique to estimate those natural or characteristic spatial scales of ecosystems. Building on a method initiated by Rand and Wilson (1995), and then advanced by Keeling et al. (1997) and Pascual and Levin (1999), we have investigated a robust method that, requiring only a minimal amount of spatially resolved data through time, can be used to estimate the optimal sampling scales of natural ecological systems. These characteristic scales define the most appropriate scales at which densities of individuals can be averaged to sample maximum determinism (Pascual and Levin 1999).

7.2 Attributes of characteristic length scale estimates

7.2.1 Nonlinear and system-level

The various attributes of these natural scale estimates promote their widespread implementation in ecology. Characteristic length scales (CLSs) are derived with methods from nonlinear time series analysis, thereby becoming the first measures of the optimal sampling scale which take into account the nonlinear, dynamical, if not chaotic complexity of natural systems. CLSs are system-level measures and therefore can be estimated from the dynamics of a

single species within the system. Indeed, for the system studied here, species from a range of phyla including Rhodophyta, Cnidaria, and Porifera provided similar estimates of scale (Chapter 3), confirming that measurement of only a single species' dynamic is adequate for the determination of sampling scale for a whole system, given that the dynamics of these species competing for space are interdependent (Chapter 6). It is important to note, however, that if the species under investigation provides a spectrum that is even slightly ambiguous, other species from the system should be evaluated to confirm the CLS estimates.

7.2.2 Tractable data requirements

Using the new method based on short time series where replication is through space rather than through time, CLSs can be estimated from a set of spatially resolved maps collected over only three consecutive time steps. The modification of the method allows use of only a small fraction of the data required for the original method of Pascual and Levin (1999), which was developed for model systems from which time series of thousands of time steps were readily obtained. The data requirements are now suitable for application to natural systems.

7.2.3 Robust to arbitrary parameter changes

CLS estimates have been shown to be robust to arbitrary choices of parameters used in attractor reconstruction (Chapter 2) and the form of sampling regimes (Chapter 4). Importantly, CLS estimates are not sensitive to these arbitrary choices, but they are sensitive to weak connectivity among species, and to shifting attractors. Both properties are highly desirable. Robustness of the estimate is essential for its implementation; sensitivity to slight parameter or

sampling regime changes would significantly decrease its value as a tool for ecologists.

7.2.4 Estimated with both species and habitat level data

Another attribute that makes CLSs attractive is that in addition to their successful estimation at the local level of quadrats with species level data, characteristic length scales have been successfully estimated at the broader landscape scale, using habitat level data mapped from remotely sensed images (Chapter 5). A CLS was detectable for a coral reef lagoon using highly resolved habitat level data. Very distinct habitat types differing in morphology, abundance, and their component species indicated identical estimates of the CLS, reinforcing the conclusion that CLSs are properties intrinsic to the system observed. With a widening availability of remotely sensed imagery, applications for use of the estimate with habitat level data range from determination of the scale at which to monitor for ecosystem based management (Lewis et al. 1996), to using the CLS as an index in itself as a baseline for detecting long-term change.

7.2.5 Ability to identify weakly connected species

Finally, the CLS shows sensitivity to poorly connected species within the system, a factor that many ecologists might expect to alter the appropriate scale of observation of the system dynamic. Sub-communities that are only weakly connected to each other and which differ in the properties of their dynamics tend to indicate dissimilar CLS estimates. This suggests that the system-level nature of the CLS, i.e., that the same length scale will be defined by all species within the system, relies upon species with interlinked and therefore correlated dynamics. Using models with varying levels of connectivity between two spatially self-

organized groups of species we found that only when different species groups show dissimilar dynamics and are weakly connected do they indicate different CLSs (Chapter 6), a result which may eventually prove extremely valuable for using CLSs as indicators of poor species connectivity.

7.3 CLS implementation

Assessment of the reliability, usability, and flexibility of this estimate was necessary to establish it as a quantitative measure, and its performance now provides confidence that it is ready for application to other natural ecosystems. Characteristic length scales have been evaluated for various systems, both model and real, and we have established their potential as valuable tools for ecologists seeking to identify dynamical trends above stochastic variability. Estimating the natural scales of the study system at the commencement of monitoring will best facilitate observation of the dynamical trends that every investigator seeks to detect in a timely manner.

Almost two decades ago, Wiens (1989) recommended that rather than asking how results vary as a function of scale, we consider scaling as a primary focus of research. This goal has now been achieved and Wiens' request for "nonarbitrary, operational ways of defining and detecting scales" has, in part, been realized. There is now a tractable method that can objectively guide the scaling of ecological studies towards detecting the dynamical trends that underlie the variability seen in nature, and no longer will the scale of observation need to be defined intuitively.

7.4 The future of CLSs: possibilities and limitations

What is the future of characteristic length scales? As rates of collecting remotely sensed and other kinds of digital data increase, the technique for estimating the natural scales of a study system should become a part of every ecologist's toolbox. In the future, length scales will likely be determined for systems of all descriptions, from terrestrial to marine and pelagic to benthic, and for both one- and two-dimensional sampling data. However, this technique will not replace familiarization with the study system, which is essential regardless, nor will it offer a precise magic number that dictates the scale of the study. Instead the CLS will serve as a quantitative guide, to assist ecologists with the task of choosing the spatial scale of observation, above which the fluctuations average out, and below which noise dominates. It is encouraging that scientists have in the past sampled subtidal marine fouling communities at scales (area = approximately 0.1 m^2, e.g., Dunstan and Johnson 2004) similar to the optimal sampling scale estimated for a fouling community with this new CLS method (0.35 m in length).

Intuition can be a valuable tool, and if it is used in combination with this objective measure of sampling scale, studies will be more likely to detect the biologically driven variability in the system. Because the CLS method does not define the size of the landscape, nor the spatial resolution required for its estimation, an intuitive feel for the scale of the system will still be required to implement the technique. The size of the landscape mapped must be larger than the CLS of the system. If the landscape size used in the short time series method is too small or the resolution is too low, then the prediction r^2 spectra will be uninterpretable. Similarly, an appropriate time scale over which the three maps

are spaced must be defined based on an understanding of the change or turnover in the system. Small variations in the length of intervals will not affect the estimate (Chapter 4), but the choice of months versus years lies with the investigator, until a technique to determine characteristic time scales is developed at some future date. The decay of the autocorrelation function may be a candidate for normalizing the time interval between the maps, but the relationship between this function and the CLS needs further investigation before it can be implemented.

Originally, we mused whether CLSs might be useful in the design and planning of reserves for conservation. We now know that the local CLSs estimated are too small to be of use in reserve design, but will instead assist in scaling the monitoring of the reserves. Managers seek the optimal scale for monitoring trends in systems (Stalmans et al. 2001, Nicholson and Jennings 2004), and this method will be able to guide their efforts towards the most appropriate sampling scale, such that trends will be detected efficiently and effectively.

Relative to other scaling measures such as the area at which all species in the system can be observed, the primary CLS appears to be small in magnitude. Only a fraction of the diversity of the system is captured in the optimal sampling scale estimated for the fouling community system studied here (Chapter 3). An avenue of future research will be to look at other scaling phenomena in more detail to clarify the place of the CLS relative to other measures.

Multiple natural scales are likely to occur for real dynamical systems. The method used in the work presented here is able to detect at least two scales, the smallest of which defines the scale for observation of trends in the local system

dynamic, and the second, which we have termed the secondary length scale, indicates the scale at which the patch level emergent dynamic appears. These secondary scales are less predictable, more dependent on the species used to reconstruct the attractor, and arise from the collective behavior of the smaller scale processes. Such patch level emergent dynamics may manifest as spatial self-organization, for example (Bascompte and Sole 1995, Johnson 1997). Though we now have the ability to measure this scale, which is an accomplishment in itself, its use is a question that needs to be explored further.

Environmental stochasticity in the form of physical disturbances may affect CLSs detected, on occasions when disturbance changes the system's pattern formation. In the coral reef model evaluated in Chapter 2, mortality experienced by the colonies included that caused by environmental disturbance (Johnson 1997). However, the CLS was not influenced by the level of mortality because the spatial pattern driven by the dynamic of the system was maintained. Large-scale environmental disturbances such as hurricanes may even contribute to connectivity, driving the similarity in CLSs estimated for different species (Chapter 5). A direction of future studies may be to further investigate the role of environmental, as opposed to demographic, stochasticity on CLSs estimated for natural systems.

Finally, because of the method of graphical analysis used to visually estimate natural length scales, researchers should note that the scale is an estimate and not an absolute value. Where curves are difficult to interpret, or error in the prediction r^2 metric is difficult to distinguish from a region of inflection, multiple species should be evaluated for verification of the scale estimate. When comparing CLSs of systems evaluated by different researchers, care must be

taken to assure that similar, consistent criteria were used for choosing the region

of plateau in the curve.

7.5 Conclusion

The results contained within this thesis are exciting, and potentially relevant

to a wide spectrum of scientists with a management focus. This research likely

marks the introduction of what will become an accepted systematic means of

defining the sampling scale of observation.

Literature Cited

Abarbanel, H. 1996. Analysis of Observed Chaotic Data. Springer-Verlag, New York.

Allen, T. F. H., D. A. Sadowsky, and N. Woodhead. 1984. Data transformation as a scaling operation in ordination of plankton. Vegetatio 56:147-160.

Anderson, P. E., and H. J. Jensen. 2005. Network properties, species abundance and evolution in a model of evolutionary ecology. Journal of Theoretical Biology 232:551-558.

Bascompte, J., and R. V. Sole. 1995. Rethinking complexity: modelling spatiotemporal dynamics in ecology. Trends in Ecology and Evolution 10:361-366.

Bishop, M. J., A. J. Underwood, and P. Archambault. 2002. Sewage and environmental impacts on rocky shores: necessity of identifying relevant spatial scales. Marine Ecology Progress Series 236:121-128.

Bissonette, J. A. 1997. Scale-sensitive ecological properties: Historical context, current meaning. Pages 3-31 *in* J. A. Bissonette, editor. Wildlife and Landscape Ecology: Effects of Pattern and Scale. Springer-Verlad, New York.

Buzug, T., and G. Pfister. 1992. Comparison of aalgorithms calculating optimal embedding parameters for delay time coordinates. Physica D 58:127-137.

Bythell, J. C., Z. M. Hillis-Starr, and C. S. Rogers. 2000. Local variability but landscape stability in coral reef communities following repeated hurricane impacts. Marine Ecology Progress Series 204:93-100.

Carlile, D., J. Skalski, J. Batker, J. Thomas, and V. Cullinan. 1989. Determination of ecological scale. Landscape Ecology 2:203-213.

Casdagli, M. 1989. Nonlinear prediction of chaotic time series. Physica D **35**:335-356.

Castilla, J. C. 2000. Roles of experimental marine ecology in coastal management and conservation. Journal of Experimental Marine Biology and Ecology **250**:3-21.

Chalcraft, D. R., J. W. Williams, M. D. Smith, and M. R. Willig. 2004. Scale dependence in the species-richness-productivity relationship: The role of species turnover. Ecology **85**:2701-2708.

Chase, J. M., and M. A. Leibold. 2002. Spatial scale dictates the productivity-biodiversity relationship. Nature **416**:427-430.

Chave, J., H. Muller-Landau, and S. Levin. 2002. Comparing classical community models: Theoretical consequences for patterns of diversity. American Naturalist **159**:1-23.

Connor, E. F., and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. American Naturalist **113**:791-833.

Crawley, M. J., and J. E. Harral. 2001. Scale dependence in plant biodiversity. Science **291**:864-868.

De Roos, A., E. McCauley, and W. Wilson. 1991. Mobility versus density-limited predator-prey dynamics on different spatial scales. Proceedings of the Royal Society of London B **246**:117-122.

Denny, M. W., B. Helmuth, G. H. Leonard, C. D. G. Harley, L. J. H. Hunt, and E. K. Nelson. 2004. Quantifying scale in ecology: Lessons from a wave-swept shore. Ecological Monographs **74**:513-532.

Done, T. J. 1992. Effects of tropical cyclone waves on ecological and geomorphological structures on the Great Barrier Reef. Continental Shelf

Research **12**:859-872.

Duckworth, A. R., and C. N. Battershill. 2001. Population dynamics and chemical

ecology of New Zealand demospongiae Latrunculia sp nov and

Polymastia croceus (Poecilosclerida: Latrunculiidae: Polymastiidae). New

Zealand Journal of Marine and Freshwater Research **35**:935-949.

Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002. Food-web structure and

network theory: The role of connectance and size. Proceedings of the

National Academy of Science **99**:12917-12922.

Dunstan, P. K., and C. R. Johnson. 2004. Invasion rates increase with species

richness in a marine epibenthic community by two mechanisms.

Oecologia **138**:285-292.

Dunstan, P. K., and C. R. Johnson. 2005. Predicting global dynamics from local

interactions: individual-based models predict complex features of marine

epibenthic communities. Ecological Modelling, in press.

Durrett, R., and S. Levin. 2000. Lessons on pattern formation from planet

WATOR. Journal of Theoretical Biology **205**:201-214.

Ellner, S. 1989. Inferring the causes of population fluctuations. *in* C. Castillo-

Chavez, S. A. Levin, and C. A. Shoemaker, editors. Mathematical

approaches to problems in resource management and epidemiology. Lecture

Notes in Biomathematics. Springer-Verlag, Berlin, Germany.

Ellner, S., and P. Turchin. 1995. Chaos in a noisy world: new methods and

evidence from time-series analysis. American Naturalist **145**:343-375.

Farmer, J. 1982. Chaotic attractors of an infinite-dimensional dynamical system.

Physica D **4**:366-393.

Fourqurean, J. W., and L. M. Rutten. 2004. The impact of hurricane Georges on

soft-bottom, back reef communities: site- and species-specific effects in

South Florida seagrass beds. Bulletin of Marine Science **75**:239-257.

Fraser, A. M., and H. L. Swinney. 1986. Independent coordinates for strange

attractors from mutual information. Physical Review A **33**:1134-1140.

Green, E. P., P. J. Mumby, A. J. Edwards, and C. D. Clark. 1996. A review of

remote sensing for the assessment and management of tropical coastal

resources. Coastal Management **24**:1-40.

Greig-Smith, P. 1952. The use of random and contiguous quadrats in the study of

the structure of plant communities. Annals of Botany **16**:293-316.

Greig-Smith, P. 1964. Quantitative plant ecology, 2nd edition. Butterworths,

London.

Habeeb, R. L., J. Trebilco, C. R. Johnson, and S. Wotherspoon. 2005.

Determining natural scales of ecological systems. Ecological Monographs

75:467-487.

Hastings, A., C. L. Hom, S. Ellner, P. Turchin, and H. C. J. Godfray. 1993. Chaos

in ecology: Is mother nature a strange attractor? Annual Review of Ecology

and Systematics **24**:1-33.

Haydon, D. T. 2000. Maximally stable model ecosystems can be highly

connected. Ecology **81**:2631-2636.

Hill, J. K., and K. C. Hamer. 2004. Determining impacts of habitat modification

on diversity of tropical forest fauna: the importance of spatial scale. Journal

of Applied Ecology **41**:744-754.

Johnson, C. R. 1997. Self-organising in spatial competition systems. Pages 245-

263 *in* N. Klomp, and I. Lunt, editors. Frontiers in ecology: Building the

links. Elsevier, Oxford.

Johnson, C. R., D. Klumpp, F. J., and R. Bradbury. 1995. Carbon flux on coral

reefs: effects of large shifts in community structure. Marine Ecology

Progress Series **126**:123-143.

Johnson, C. R., and K.H. Mann. 1988. Diversity, patterns of adaptation, and

stability of Nova Scotian kelp beds. Ecological Monographs **58**:129-154.

Johnson, C. R., and I. Seinen. 2002. Selection for restraint in competitive ability

in spatial competition systems. Proceedings of the Royal Society of London

B **269**:655-663.

Kantz, H., and T. Schreiber. 1997. Nonlinear Time Series Analysis. Cambridge

University Press, Cambridge.

Kaplan, D., and L. Glass. 1995. Understanding Nonlinear Dynamics. Springer-

Verlag, New York.

Keeling, M., I. Mezic, R. Hendry, J. Mcglade, and D. Rand. 1997. Characteristic

length scales of spatial models in ecology via fluctuation analysis.

Philisophical Transactions of the Royal Society of London B **352**:1589-

1601.

Kendall, M. S., T. Battista, and Z. Hillis-Starr. 2004. Long term expansion of a

deep *Syringodium filiforme* meadow in St. Croix, US Virgin Islands: the

potential role of hurricanes in the dispersal of seeds. Aquatic Botany **78**:15-

25.

Kennel, M. B., R. Brown, and H. D. I. Abarbanel. 1994. Determining embedding

dimension for phase-space reconstruction using a geometrical construction.

in E. Ott, T. Sauer, and J. A. Yorke, editors. Coping with Chaos. John

Wiley & Sons, Inc., New York.

Kershaw, K. A. 1957. The use of cover and frequency in the detection of pattern

in plant communities. Ecology **38**:291-299.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology **73**:1943-1967.

Levin, S. A. 2000. Multiple scales and the maintenance of biodiversity. Ecosystems **3**:498-506.

Levin, S. A., B. Grenfell, A. Hastings, and A. S. Perelson. 1997. Mathematical and computational challenges in population biology and ecosystems science. Science **275**:334-343.

Lewis, C. A., N. P. Lester, A. D. Bradshaw, J. E. Fitzgibbon, K. Fuller, L. Hakanson, and C. Richards. 1996. Considerations of scale in habitat conservation and restoration. Canadian Journal of Fisheries and Aquatic Sciences **53**:440-445.

Lewis, J. B. 2002. Evidence from aerial photography of structural loss of coral reefs at Barbados, West Indies. Coral Reefs **21**:49-56.

Liebert, W., K. Pawelzik, and H. Schuster. 1991. Optimal embeddings of chaotic attractors from topological considerations. Europhysics Letters **14**:521-526.

Liebert, W., and H. Schuster. 1989. Proper choice of the time delay for the analysis of chaotic time series. Physics Letters A **142**:107-111.

Little, S., S. Ellner, M. Pascual, M. Neubert, D. Kaplan, T. Sauer, H. Caswell, and A. Solow. 1996. Detecting nonlinear dynamics in spatio-temporal systems, examples from ecological models. Physica D **96**:321-333.

Loehle, C., and G. Wein. 1994. Landscape habitat diversity: a multiscale information theory approach. Ecological Modelling **73**:311-329.

Lundquist, J. E., and R. A. Sommerfeld. 2002. Use of fourier transforms to define landscape scales of analysis for disturbances: a case study of thinned and

unthinned forest stands. Landscape Ecology **17**:445-454.

Marcos-Nikolaus, P., J. M. Martin-Gonzalez, and R. V. Sole. 2002. Spatial

forecasting: Detecting determinism from single snapshots. International

Journal of Bifurcation and Chaos **12**:369-376.

Massel, S. R., and T. J. Done. 1993. Effects of cyclone waves on massive coral

assemblages on the Great Barrier Reef: Meteorology, hydrodynamics and

demography. Coral Reefs **12**:153-166.

May, R. M. 1976. Simple mathematical models with very complicated dynamics.

Nature **261**:459-467.

McCann, K. S. 2000. The diversity-stability debate. Nature **405**:228-233.

McNeill, S. E. 1994. The selection and design of marine protected areas:

Australia as a case study. Biodiversity and Conservation **3**:586-605.

Meesters, E. H., M. Hilterman, E. Kardinaal, M. Keetman, M. de Vries, and R. P.

M. Bak. 2001. Colony size-frequency distributions of scleractinian coral

populations: spatial and interspecific variation. Marine Ecology Progress

Series **209**:43-54.

Mumby, P. J., A. J. Edwards, E. P. Green, C. W. Anderson, A. C. Ellis, and C. D.

Clark. 1997. A visual assessment technique for estimating seagrass standing

crop. Aquatic Conservation Marine and Freshwater Ecosystems **7**:239-251.

Mumby, P. J., E. P. Green, C. D. Clark, and A. J. Edwards. 1998a. Digital

analysis of multispectral airborne imagery of coral reefs. Coral Reefs **17**:59-

69.

Mumby, P. J., C. D. Clark, E. P. Green, and A. J. Edwards. 1998b. Benefits of

water column correction and contextual editing for mapping coral reefs.

International Journal of Remote Sensing **19**:203-210.

Mumby, P. J., and A. R. Harborne. 1999. Development of a systematic

classification scheme of marine habitats to facilitate regional management

and mapping of Caribbean coral reefs. Biological Conservation **88**:155-163.

Mumby, P. J., J. R. M. Chisholm, C. D. Clark, J. D. Hedley, and J. Jaubert. 2001.

A bird's-eye view of the health of coral reefs. Nature **413**:36-36.

Mumby, P. J., and A. J. Edwards. 2002. Mapping marine environments with

IKONOS imagery: enhanced spatial resolution can deliver greater thematic

accuracy. Remote Sensing of Environment **82**:248-257.

Molofsky, J., J. Bever, J. Antonovics, and T. Newmaan. 2002. Negative

frequency dependence and the importance of spatial scale. Ecology **83**:21-

27.

Nichols, J. M., and J. D. Nichols. 2001. Attractor reconstruction for non-linear

systems: a methodological note. Mathematical Biosciences **171**:21-32.

Nicholson, M. D., and S. Jennings. 2004. Testing candidate indicators to support

ecosystem-based management: the power of monitoring surveys to detect

temporal trends in fish community metrics. ICES Journal of Marine

Science **61**:35-42.

O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen. 1986. A

Hierarchical Concept of Ecosystems. Princeton University Press, Princeton,

NJ.

Oren, U., and Y. Benayahu. 1998. Didemnid ascidians: Rapid colonizers of

artificial reefs in Eilat (Red Sea). Bulletin of Marine Science **63**:199-206.

Palumbi, S. R. 2001. The ecology of marine protected areas. Pages 509-530 *in* M.

D. Bertness, S. D. Gaines, and M. E. Hay, editors. Marine Community

Ecology. Sinauer Associates, Inc., Sunderland, Massachusetts.

Pandolfi, J. M. 2002. Coral community dynamics at multiple scales. Coral Reefs **21**:13-23.

Pascual, M., and S. A. Levin. 1999. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. Ecology **80**:2225-2236.

Pascual, M., and S. Ellner. 2000. Linking ecological patterns to environmental forcing via nonlinear time series models. Ecology **81**:2767-2780.

Pascual, M., P. Mazzega, and S. A.Levin. 2001. Oscillatory dynamics and spatial scale: the role of noise and unresolved pattern. Ecology **82**:2357-2369.

Peterson, C. H., and J. A. Estes. 2001. Conservation and management of marine communities. Pages 469-507 *in* M. D. Bertness, S. D. Gaines, and M. E. Hay, editors. Marine Community Ecology. Sinauer Associates, Inc., Sunderland, Massachusetts.

Petrovskii, S., L. Bai-Lian, and H. Malchow. 2003. Quantification of the spatial aspect of chaotic dynamics in biological and chemical systems. Bulletin of Mathematical Biology **65**:425-446.

Pielou, E. C. 1977. Mathematical Ecology, 2nd edition. John Wiley & Sons, New York.

Pressey, R. L., and V. S. Logan. 1998. Size of selection units for future reserves and its influence on actual vs targeted representation of features: a case study in western New South Wales. Biological Conservation **85**:305-319.

Qi, Y., and J. Wu. 1996. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indicies. Landscape Ecology **11**:39-49.

Rand, D. 1994. Measuring and characterizing spatial patterns, dynamics and

chaos in spatially extended dynamical systems and ecologies. Philisophical

Transactions of the Royal Society of London A **348**:497-514.

Rand, D., and H. Wilson. 1995. Using spatio-temporal chaos and intermediate-

scale determinism to quantify spatially extended ecosystems. Proceedings

of the Royal Society of London B **259**:111-117.

Rouget, M. 2003. Measuring conservation value at fine and broad scales:

implications for a diverse and fragmented region, the Agulhas Plain.

Biological Conservation **112**:217-232.

Roughgarden, J., S. W. Running, and P. A. Matson. 1991. What does remote

sensing do for ecology? Ecology **72**:1918-1922.

Schneider, D. C. 1994. Quantitative Ecology: Spatial and Temporal Scaling.

Academic Press, San Diego.

Schreiber, T. 1995. Efficient neighbor searching in nonlinear time series analysis.

International Journal of Bifurcation and Chaos **5**:349-358.

Schreiber, T. 1999. Interdisciplinary application of nonlinear time series methods.

Physics Reports **308**:2-64.

Schumaker, N. H. 1996. Using landscape indices to predict habitat connectivity.

Ecology **77**:1210-1225.

Schwartz, M. W. 1999. Choosing the appropriate scale of reserves for

conservation. Annual Review of Ecology and Systematics **30**:83-108.

Sole, R. V., and J. Bascompte. 1995. Measuring chaos from spatial information.

Journal of Theoretical Biology **175**:139-147.

Stalmans, M., K. Balkwill, E. T. F. Witkowski, and K. H. Rogers. 2001. A

landscape ecological approach to address scaling problems in conservation

management and monitoring. Environmental Management **28**:389-401.

Stevens, T. 2002. Rigor and representativeness in marine protected area design. Coastal Management **30**:237-248.

Sugihara, G., B. Grenfell, and R. M. May. 1990. Distinguishing error from chaos in ecological time series. Philisophical Transactions of the Royal Society of London B **330**:235-251.

Takens, F. 1981. Detecting strange attractors in turbulence. Pages 366-381 *in* D. Rand and L. Young, editors. Dynamical systems and turbulence, Warwick 1980. Lecture Notes in Mathematics. Springer-Verlag, New York.

Tanaka, K. 2002. Growth dynamics and mortality of the intertidal encrusting sponge *Halichondria okadai* (Demospongiae, Halichondrida). Marine Biology **140**:383-389.

Tischendorf, L., and L. Fahrig. 2000. On the usage and measurement of landscape connectivity. OIKOS **90**:7-19.

Tong, H. 1990. Non-linear Time Series Analysis. Oxford University Press, Oxford, UK.

Turchin, P., and A. D. Taylor. 1992. Complex dynamics in ecological time series. Ecology **73**:289-305.

Turner, M. G., R. V. O'Neill, R. H. Gardner, and B. T. Milne. 1989. Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology **3**:153-162.

Turner, S., R. V. O'Neill, W. Conley, M. Conley, and H. Humphries. 1991. Pattern and scale: Statistics for landscape ecology. Pages 17-47 *in* S. J. Turner and R. H. Gardner, editors. Quantitative Methods in Landscape Ecology. Springer Verlag, New York.

Tyre, A., H. Possingham, and C. Bull. 1997. Characteristic scales in ecology: fact, fiction or futility. Pages 233-243 *in* N. a. L. Klomp, I., editor. Frontiers in Ecology. Elsevier Science Ltd, New York.

Ward, T. J., M. A. Vanderklift, A. O. Nicholls, and R. A. Kenchington. 1999. Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. Ecological Applications 9:691-698.

Weitz, J. S., and D. H. Rothman. 2003. Scale-dependence of resource-biodiversity relationships. Journal of Theoretical Biology 225:205-214.

Wiens, J. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.

Wilson, H. B., and M. J. Keeling. 2000. Spatial scales and low dimensional deterministic dynamics. Pages 209-226 *in* U. Dieckmann, R. Law, and J. A. J. Metz, editors. The geometry of ecological interactions: Simplifying spatial complexity. Cambridge University Press, Cambridge.

Appendix A

Attractor reconstruction

Takens (1981) showed that the general shape of a multi-species system dynamic (the attractor) can be reconstructed in n-dimensional phase space using time series data from a single observed variable X_t. A reconstructed attractor is qualitatively similar to the real state space attractor that could be obtained by measuring all variables in the system. While there are several approaches to attractor reconstruction (Abarbanel 1996, Kantz and Schreiber 1997) the method used here, and in previous CLS estimation techniques, is time-delay embedding. Using the Takens (1981) theorem for attractor reconstruction, vectors of dimension d_e are constructed from the original time series $x(t_i)$,

$$X(t_i) = \{x(t_i), x(t_i + \tau), x(t_i + 2\tau), ..., x(t_i + (d_e - 1)\tau)\}$$

where $X(t_i)$ is the observable state variable at discrete time (t_i), τ is the time delay, and d_e is the embedding dimension. These points are then assumed to approximate the reconstructed attractor. For example, for the time series X_1, X_2, X_3, ..., X_{10}, the reconstructed attractor with $\tau = 3$ and $d_e = 2$ has points (X_1, X_4), (X_2, X_5), (X_3, X_6), ..., (X_7, X_{10}) in a 2-dimensional phase space.

The time delay (τ) is some multiple of the sampling time, describing how lagged in time the coordinates of the attractor will be (Abarbanel 1996). Pascual and Levin (1999) select τ as the value for which the autocorrelation function first crosses zero. Autocorrelation measures correlations between coordinates at different values of τ, and the point at which this function crosses zero indicates the time delay at which coordinates are independent but still dynamically linked.

However, Tong (1990) criticized use of this linear measure for a nonlinear analysis, and for this reason Abarbanel (1996), Kantz and Schreiber (1997) and Nichols and Nichols (2001) favor the use of Fraser and Swinney's (1986) mutual information function, a nonlinear equivalent of autocorrelation, to determine an optimal value for τ. Thus, in this research, τ was chosen using the first minimum point in the time delayed mutual information curve (Fraser and Swinney 1986, Abarbanel 1996, Kantz and Schreiber 1997, Nichols and Nichols 2001). This technique measures the amount of information shared between two measurements a and b. When the amount of information learned from a about b is at a minimum, the two time points are taken to be sufficiently independent (Abarbanel 1996).

The embedding dimension (d_e) is the minimum dimension in phase space needed to capture the system dynamics (Farmer 1982). To estimate d_e, the standard false nearest neighbors method is used (Kantz and Schreiber 1997), where the dimension is chosen when the percentage of false nearest neighbors falls within some small threshold of zero (Kennel et al. 1992). Each point on the attractor has some nearest neighbor in Euclidean space. False projections occur when the attractor is embedded in too few dimensions, and therefore is folded. The embedding dimension required to unfold the attractor is estimated to be where the number of false nearest neighbors drops below the level of noise and each point's nearest neighbors remain the same when the embedding dimension is increased (Liebert et al. 1991, Nichols and Nichols 2001).

Once τ and d_e are chosen, prediction of the attractor must be made. Nearby points on the attractor are followed to determine their location after some t time steps (Sugihara et al. 1990). To predict $x(t_i + \tau)$ from $x(t_i)$, a list of all the states of x previously visited is searched for those closest to $x(t_i)$. If the time series is long

enough, then some past states will be close to the present and the prediction will be close to the true state of the system (Kantz and Schreiber 1997). Some set of points (k) around the point trying to be predicted is used for the prediction. These points, or nearest neighbors, are chosen based on their proximity, and then are averaged to determine the prediction in phase space (Kaplan and Glass 1995, Kantz and Schreiber 1997). In this thesis, the approach was to choose a fixed number of k nearest neighbors (neighborhoods of *fixed mass*) and then weight the average of the neighbors by inverse distance (Schreiber 1995). The predicted value of each point is then the average of the observed values of its k nearest neighbors, and the average is weighted by inverse distance so that neighbors further away contribute less (Casdagli 1989).

The algorithm based on k-nearest neighbors (Kaplan and Glass 1995) is used to obtain a set of predicted values $\hat{X}_t$ from the reconstructed attractor. The value of X_t after a lag of h time steps is X_{t+h}. Predicted values of X_{t+h} are obtained by taking the weighted average of the trajectories of the k points closest to X_t in reconstructed phase space. Predictions are made for each point on the attractor, with h set equal to the time delay (τ).

Appendix B

The null case: no spatial pattern

The following proof shows that for a system with no spatial pattern, the expected values of prediction r^2 (Pascual and Levin 1999) and error X (Keeling et al. 1997) are constant across all window sizes. This theoretical result has been confirmed by analyzing a time series of randomly simulated landscapes.

Proof

Consider a time series of landscapes, L_{ij}^t, which are composed of independent, discrete valued pixels. The data series for a window of side length l is given by $N \approx Bin(l^2, \pi)$ with $E(N) = l^2 \pi$ and $Var(N) = l^2 \pi (1-\pi)$.

The binomial distribution is used in preference to the normal distribution because the binomial gives discrete valued data and so provides a null case comparable with individual-based spatially explicit models containing species $X_1 - X_n$. The value π can be thought of as the probability of observing a particular species X. The density of X in a window of side length l is

$$X_l^t = \frac{1}{l^2} \sum_{i=1}^{l} \sum_{j=1}^{l} L_{ij}^t = N^t \Big/ l^2 \text{ and } E_t(X_l^t) = \pi \text{ and}$$

$$Var(X_l^t) = \frac{1}{l^4} Var(N^t)$$

$$= \frac{1}{l^4} l^2 \pi (1-\pi)$$

$$= \frac{\pi(1-\pi)}{l^2} \tag{A1}$$

Because the landscapes L_{ij}^t are independent, the X_l^t are also independent. Thus, the *k*-nearest neighbors method averages *k* independent values of X_l^t to

determine $\hat{X}_l^t$.

Therefore $E_t(\hat{X}_l^t) = E_t(X_l^t) = \pi$ and $Var(\hat{X}_l^t) = Var\dfrac{1}{k}\displaystyle\sum_{i=1}^{k} X_l^{t_i}$

$$= \dfrac{1}{k^2} k \dfrac{\pi(1-\pi)}{l^2}$$

$$= \dfrac{\pi(1-\pi)}{kl^2} \qquad\qquad \text{(A2)}$$

$\hat{X}_l^t$ and X_l^t are independent, so using (A1) and (A2)

$$Var(X_l^t - \hat{X}_l^t) = \dfrac{\pi(1-\pi)}{l^2} + \dfrac{\pi(1-\pi)}{kl^2} = \left(\dfrac{\pi(1-\pi)}{l^2}\right)\left(\dfrac{k+1}{k}\right) \qquad \text{(A3)}$$

Since $E_t(X_l^t - \hat{X}_l^t) = 0$ then $Var(X_l^t - \hat{X}_l^t) = E_t[(X_l^t - \hat{X}_l^t) - E_t(X_l^t - \hat{X}_l^t)^2]$

$$= E_t[(X_l^t - \hat{X}_l^t)^2] - E_t(X_l^t - \hat{X}_l^t)^2$$

$$= E_t[(X_l^t - \hat{X}_l^t)^2] \qquad\qquad \text{(A4)}$$

Together (A3) and (A4) give $E_t[(X_l^t - \hat{X}_l^t)^2] = \left(\dfrac{\pi(1-\pi)}{l^2}\right)\left(\dfrac{k+1}{k}\right)$ which is used

to predict values of the CLS statistics:

for error X (Keeling et al. 1997):

$$\text{error X} = l\sqrt{E_t[(X_l^t - \hat{X}_l^t)^2]}$$

$$= l\dfrac{\sqrt{\pi(1-\pi)}}{l}\sqrt{\dfrac{k+1}{k}}$$

$$= \sqrt{\pi(1-\pi)\left(\dfrac{k+1}{k}\right)}$$

$$= \text{constant}$$

for prediction r^2 (Pascual and Levin 1999):

$$\text{prediction } r^2 = 1 - \frac{E_t[(X_l^t - \hat{X}_l^t)^2]}{Var(X_l^t)}$$

$$= 1 - \frac{\left(\frac{\pi(1-\pi)}{l^2}\right)\left(\frac{k+1}{k}\right)}{\left(\frac{\pi(1-\pi)}{l^2}\right)}$$

$$= -\frac{1}{k}$$

$$= \text{constant}$$

Expected values of prediction r^2 and error X for the null case were validated using a series of landscapes with $\pi = 0.3$ and $k = 10$. Expected values for prediction r^2

(i.e., $-\frac{1}{k} = -0.1$) and error X (i.e., $\sqrt{\pi(1-\pi)\left(\frac{k+1}{k}\right)} = \sqrt{0.21 \times 1.1} = 0.48$) were

produced in our simulations (results not shown).

Appendix C

DETERMINING NATURAL SCALES OF ECOLOGICAL SYSTEMS

Rebecca L. Habeeb[1], Jessica Trebilco[1], Simon Wotherspoon[2], Craig R. Johnson[1]*

[1]School of Zoology and Tasmanian Aquaculture and Fisheries Institute,

University of Tasmania, GPO Box 252-05, Hobart, Tasmania 7001 Australia

[2]School of Mathematics and Physics, University of Tasmania, GPO Box 252-37,

Hobart, Tasmania 7001 Australia

*craig.johnson@utas.edu.au

DETERMINING NATURAL SCALES OF ECOLOGICAL SYSTEMS

Rebecca L. Habeeb,[1] Jessica Trebilco,[1] Simon Wotherspoon,[2] and Craig R. Johnson[1,3]

[1]*School of Zoology and Tasmanian Aquaculture and Fisheries Institute, University of Tasmania, Private Bag 5, Hobart, Tasmania 7001 Australia*
[2]*School of Mathematics and Physics, University of Tasmania, Private Bag 37, Hobart, Tasmania 7001 Australia*

Abstract. A key issue in ecology is to identify the appropriate scale(s) at which to observe trends in ecosystem behavior. The characteristic length scale (CLS) is a natural scale of a system at which the underlying deterministic dynamics are most clearly observed. Any approach to estimating CLSs of a natural system must be able to accommodate complex nonlinear dynamics and must have realistic requirements for data. Here, we compare the robustness of two methods to estimate CLSs of dynamical systems, both of which use attractor reconstruction to account for the complex oscillatory dynamics of ecological systems. We apply these techniques to estimate CLSs of spatial multispecies systems of varying complexity, and show that both methods are robust for the simplest system, but as model complexity increases, the Pascual and Levin metric is more robust than that of Keeling et al. Both methods demonstrate some sensitivity to the choice of species used in the analysis, with closely connected species producing more similar CLSs than loosely connected species. In this context, connectivity is determined both by the topology of the interaction network and spatial organization in the system. Notably, systems showing complex spatial self-organization can yield multiple CLSs, with larger length scales indicating the emergent dynamics of interactions between patches. While the prediction r^2 metric of Pascual and Levin is suitable to estimate CLSs of complex systems, their method is not suitable to apply to most real ecosystems because of the requirement of long time series for attractor reconstruction. We offer two alternatives, both based on prediction r^2, but where repetition in space is largely (the "short time series" method) or wholly (the "sliding window" method) substituted for repetition in time in attractor reconstruction. Both methods, and in particular the short time series based on only three or four sequential observations of a system, are robust in detecting the primary length scale of complex systems. We conclude that the modified techniques are suitable for application to natural systems. Thus they offer, for the first time, an opportunity to estimate natural scales of real ecosystems, providing objectivity in important decisions about scaling in ecology.

Key words: attractor reconstruction; characteristic length scale; community dynamics; ecosystem; nonlinear dynamics; spatial and temporal dynamics; spatial scale; spatiotemporal models.

INTRODUCTION

A fundamental goal in ecology is to determine the dynamical processes underlying observed patterns. The single greatest difficulty confronting this important objective is that any pattern detected, and ultimately the understanding of the underlying dynamical processes, depends on the spatial scale at which we make our observations (Wiens 1989, Levin 1992, 2000, Schneider 1994). The challenge is to identify the appropriate scales of observation for ecological investigation (Levin 1992). Are there natural or "characteristic" scales of ecological systems that are optimal for observing a system's behavior that provide unambiguous information about the dynamic, and which minimize noise in the signal measured? If so, then the capacity to identify these scales may be useful to applied ecology, par-

ticularly to monitor and detect meaningful change in ecosystem state (Rand 1994, Bishop et al. 2002).

The search for a means to identify natural or "characteristic" length scales (hereafter CLSs) in ecological systems stems back at least to the 1950s (Grieg-Smith 1952, Kershaw 1957), but several more recent attempts have also addressed the problem (e.g., Carlile et al. 1989, De Roos et al. 1991, Schneider 1994). Most approaches have assumed either that ecological systems are stationary in space and time, that fluctuations are random around a stationary global average (e.g., Rand and Wilson 1995), or that any trends detected are linear (see Turner et al. 1991). These kinds of approaches fail to take into account the dynamical nature of ecological systems in space and time, which is arguably characterized by nonlinear oscillatory behaviors (e.g., Hastings et al. 1993, Ellner and Turchin 1995, Sole and Bascompte 1995, Little et al. 1996, Pascual and Ellner 2000).

In contrast, the relatively recent techniques of Keeling et al. (1997) and Pascual and Levin (1999) to iden-

tify CLSs are exceptional in that they both accommodate the complex nonlinear oscillations and nonuniform patterns in spatial variance that are pervasive in real ecosystems. The crucial development of these new approaches is the application of attractor reconstruction (Takens 1981) and prediction algorithms (Kaplan and Glass 1995) from nonlinear time series analysis to characterize dynamics at particular scales of observation. Takens (1981) shows that the attractor describing the dynamics of an entire system can be estimated from the time series of any one species. Both the Keeling et al. (1997) and Pascual and Levin (1999) techniques seek to identify the spatial scale that best distinguishes the deterministic dynamic, or trend, in a system from noise. The underlying tenet is that if the scale used to sample an ecosystem is too small, observations are swamped by noise due to strong correlations among individual samples (Durrett and Levin 2000, Wilson and Keeling 2000). If the sampling scale is too large, the nontrivial dynamics will be averaged out because distant parts of the landscape begin to act independently (Keeling et al. 1997, Pascual and Levin 1999, Wilson and Keeling 2000). The CLS is an intermediate scale which most clearly reflects the underlying deterministic signal (Pascual and Levin 1999, Wilson and Keeling 2000; for illustration, see Appendix A).

While the techniques of both Keeling et al. (1997) and Pascual and Levin (1999) show promise, thus far they have only been applied to simple model systems. Furthermore, these techniques require unrealistically long time series so that their application to most real ecosystems is impractical. Thus, if the goal of developing a technique to identify CLSs in real ecosystems is to be realized, there are two critical steps. First, these techniques must be able to indicate unambiguous length scales for dynamical systems more complex than those investigated to date. Second, the technique of attractor reconstruction needs to be modified to obviate the need for long time series.

These challenges define the two broad aims of our research: to assess the robustness of the approaches of Keeling et al. (1997) and Pascual and Levin (1999) so that we can be confident in applying the techniques to complex ecological systems, and to develop a method of attractor reconstruction that does not require an unrealistically long time series. We assess robustness by examining the capacity of the two techniques to identify unambiguous length scales for model systems spanning a spectrum of complexity, and by determining the sensitivity of CLS estimates to (1) the initial spatial arrangement of individuals in a system, (2) the choice of species used in attractor reconstruction, and (3) parameter choices in attractor reconstruction. We then develop two alternative methods which rely wholly or partially on substituting space for time in attractor reconstruction. One approach is based on using short time series obtained from multiple locations in space, while

in the other we slide "windows of observation" through space at a single point in time, entirely substituting repetition in space for repetition in time.

METHODS

Our overall approach was first to determine whether the metrics of Keeling et al. (1997) and Pascual and Levin (1999) indicate unambiguous length scales for systems of varying complexity across a range of species and choices of parameters for attractor reconstruction. We then derived two alternative approaches to estimate CLSs based on the most robust metric (i.e., that of Pascual and Levin 1999), where we substitute space for time to remove the need for long time series. Finally, we compared the estimates of CLSs derived using our two new methods with that of the original approach based on long time series.

Existing methods for estimating CLSs of dynamic oscillating systems

Both the methods of Keeling et al. (1997) and Pascual and Levin (1999) are based on a sequence of time series, each sampled at a different spatial scale. The time series for scale l is constructed by marking an $l \times l$ window on the landscape and recording the abundance or density of a single species over time as the system evolves. Nonlinear time series methods are used to make predictions for the time series, and the accuracy of these predictions is estimated. This process is repeated for a range of scales and prediction accuracy is compared across all scales.

Both Keeling et al. (1997) and Pascual and Levin (1999) use attractor reconstruction from nonlinear time series analysis as the means of prediction (Takens 1981, Kaplan and Glass 1995, Little et al. 1996). Assuming there is a deterministic system underlying the dynamics, observations of a single variable in the system over time are used as a proxy from which to predict the dynamics of the whole system. For example, in a predator–prey system, prey (or resource) density can be measured and used to predict the deterministic signal of the entire system (Rand and Wilson 1995). The strengths in using attractor reconstruction for prediction of ecological systems are that there are no assumptions of steady state, and that complex fluctuations can be accommodated. A more detailed description of the method of attractor reconstruction is given in Appendix B, while more comprehensive treatments can be found in several introductory texts on nonlinear dynamics and attractor reconstruction (Kaplan and Glass 1995, Abarbanel 1996). In summary, the phase space of the ecosystem dynamic is built from time delay coordinates of a single species, which act as surrogates for the unobserved variables of the system (Casdagli 1989, Abarbanel 1996, Kantz and Schreiber 1997). The trajectories in phase space are then used to reconstruct a set that is topologically equivalent to the attractor of

the full system (Farmer 1982, Sugihara et al. 1990, Kaplan and Glass 1995).

While both approaches for CLS estimation use attractor reconstruction as their basis, they differ in their measures of prediction error. Keeling et al. (1997) apply attractor prediction to calculate deviations around the underlying deterministic behavior and then to plot error variance (termed "error X") as a function of window size, producing a variance spectrum. For sufficiently large windows, the relative variance initially increases at a rate proportional to window size, and then plateaus. The window size at which error X reaches the asymptote is the length scale where correlations have decayed and values at different sites behave like independent random variables. This scale, where the windows become statistically independent, is identified as the CLS. It defines the "window size" in which the full spatial dynamics of the system can be observed.

Pascual and Levin's (1999) method is a variant of the approach of Keeling et al. (1997) and aims to extract the scale where the ratio of determinism to noise is maximized. This CLS will usually be slightly smaller than that required for the onset of independence defined by Keeling et al. (1997), because it is the minimum window size where the dynamics of the system can be accurately predicted (Pascual and Levin 1999). At each window size, the degree of determinism is evaluated from the prediction accuracy of the algorithm derived from attractor reconstruction. Pascual and Levin (1999) then examine how the prediction error of the trajectories changes with spatial scale (Kaplan and Glass 1995). They plot a statistic termed the prediction r^2 (or degree of determinism) for each window size to produce a variance spectrum similar to that of error X. The scale where the prediction r^2 spectrum first attains an asymptote with respect to window size identifies the scale of maximum determinism, or the CLS of the system.

By definition, sampling at a scale smaller than the CLS is suboptimal, and as window size increases towards the CLS, prediction accuracy should increase substantially. Sampling at scales larger than the CLS offers comparatively little gain in accuracy. Thus, the CLS is estimated as the scale or window size at which prediction accuracy plateaus.

Spatial models

Models of varying complexity, in terms of species richness, the network topology defining the number of direct and indirect interspecific interactions in the system (Johnson and Seinen 2002), and spatial pattern, were developed using the COMPETE software (see Supplement), which is a probabilistic individual-based system to model spatial competition between sessile colonial organisms. The models follow the fate of competing individuals in a two-dimensional landscape, and can demonstrate complex behaviors indicating nonlinear dynamics and spatial self-organization (Johnson

1997, Johnson and Seinen 2002). Any network topology among S species is possible, including intransitive loops (e.g., where $S_1 > S_2, S_2 > S_3, S_3 > S_1$; with $S_x > S_y$ indicating that species x outcompetes and displaces species y), which arise commonly in benthic marine systems (Johnson and Seinen 2002). We used models with intransitivities in their network topology because they enable persistence stability (sensu Johnson and Mann 1988) of the system without the need for elaborate model closures and forcing functions. We update the system synchronously, and use periodic (toroidal) boundary conditions.

Four model systems were implemented (in order of complexity of spatial pattern): symmetric networks of three, six, and 12 species, and a model of the dynamics of the benthos of a coral reef (Fig. 1). The three-species system is the simplest intransitive loop as described above; i.e., a circular network with binary interaction outcomes, so that each interaction has one unambiguous winner and one loser (Fig. 1A). The six-species system involves a symmetrical network in which each species overgrows and is overgrown by two species ($S_1 > (S_2, S_3), S_2 > (S_3, S_4), \ldots, S_6 > (S_1, S_2)$), with all other interactions as standoffs. In this system, the species spatially organize into two distinct groups of three species and, if the model is run for sufficient time, either group may eventually dominate (Fig. 1B). The 12-species system has a network structure of ($S_1 > (S_2, S_3, S_4), S_2 > (S_3, S_4, S_5), \ldots, S_{12} > (S_1, S_2, S_3)$) and organizes into three groups of four species, any of which may begin to dominate, as the emergent dynamic unfolds (Fig. 1C). In all three models the growth rates of all species are identical. For simplicity, there was no disturbance or mortality, and no recruitment of propagules to unoccupied sites. In the six- and 12-species models, two scales of self-organization are emergent: that of the colony and that of the patch, which is a distinctive group of colonies of several species. While the identity of species in particular patches is consistent, there is oscillation in the areal dominance of patches over long time series (1000x time steps). Patch dominance varies depending on the initial random configuration of "recruits" at time step zero and on stochasticity in execution of the rules of local interactions. This dominance arises as a result of finite landscape size, and becomes less obvious with very large landscapes (i.e., larger than ~500 × 500 cells). Colonies tend to become more aggregated over long time series, resulting in larger multispecies patches.

The coral reef model is more complex than the other three because (i) parts of colonies or whole colonies may die with subsequent recruitment to unoccupied space, and (ii) competitive outcomes are not binary so that for all pairwise interactions among species S_x and S_y, the probability of $S_x > S_y$ and $S_y > S_x$ is nonzero. There are 12 physiognomic life forms in the model (Fig. 1D). Neighbor-specific growth rates, interaction outcomes, and mortality and recruitment rates have

been parameterized from direct observations of communities on the Great Barrier Reef. The emergent community composition is very similar to that of real reefs after an appropriate period without a major disturbance event (notably, similar results are also evident from models of temperate marine benthic systems; see Dunstan and Johnson 2005). In this context, the coral reef model most closely simulates a natural system. We selected three physiognomic groups for analysis in our exploration of alternative methods to identify CLSs: (i) turf and coralline algae, (ii) corymbose and digitate Acroporidae (acroporid corals), and (iii) Faviidae (faviid corals). Acroporid and faviid corals occupy less cover than the turf and coralline algae, with faviids occurring in the smallest patches.

Determining the robustness of existing techniques to estimate CLS

We considered several properties in assessing the robustness of prediction r^2 (Pascual and Levin 1999) and error X (Keeling et al. 1997) spectra derived from long time series. First, by definition, neither measure should show any dependence on window size when spatial and temporal patterns are random, i.e., in the absence of any deterministic signal or oscillatory behavior. We show that this result holds true (Appendix C).

The other aspects of robustness considered are (1) the capacity to indicate an unambiguous length scale for spatial models of a range of complexity, (2) the capacity to identify similar length scales for different runs of the same system, (3) the capacity to identify similar length scales from different species in the same system, and (4) for any one model run, sensitivity of the CLS estimate to different arbitrary choices of parameters required for attractor reconstruction, namely τ (time delay), d_e (embedding dimension), and k (number of nearest neighbors; see Appendix B for details).

The overall approach was to run all four model systems for 10 000 time steps on landscapes of 500 × 500 cells, sampling landscapes from times 201 to 10 000 (the first 200 time steps are ignored while the system self-structures). Each model was begun with a random spatial arrangement of "recruits" (10% total cover) on the initial landscape, with identical amounts of each species. We sampled different window sizes l ($l = 5$ to 495 in steps of 5) within the 500 × 500 landscape, and observed those windows through time. At each time step there was a single window of each size. The density of each species for each window size $l \times l$ was determined, generating a separate time series for each l. For a selected species, the attractor of the system in d dimensional space was estimated for each time series (see Appendix B). For each model run, we derived estimates of CLSs using methods of both Keeling et al. (1997) and Pascual and Levin (1999).

In calculating the CLSs, abundances of the selected species were scaled by window area to convert each

time series of absolute counts to a time series of density. Attractor reconstruction was undertaken as described in Appendix B for particular choices of τ, d_e, and k ($k = 10$ unless otherwise specified). The CLS was estimated as the window size, l, at which the error X or prediction r^2 spectrum reached an asymptote. Monte Carlos of 100 independent runs of each model were used to calculate mean error X and prediction r^2 spectra with 95% confidence intervals. It should be noted that while we are interested in the sensitivity and robustness of both measures, prediction r^2 will yield a slightly smaller length scale than error X for the reasons outlined earlier, so we are not concerned with whether the two techniques provide similar absolute estimates of length scales.

Robustness to model complexity.—Because prediction r^2 and error X have thus far been calculated only for simple spatial model systems, it is necessary to ascertain whether the techniques produce meaningful results for more complex systems. Variance spectra of prediction r^2 and error X were produced for the three-, six-, and 12-species models, and for the coral reef model. Spectra from the two methods were compared for each of the four models.

Robustness to initial conditions.—For each model system, we assessed the variability of the CLS estimated from the dynamics of a given species over 100 runs of the model, each with a different initial random configuration.

Robustness to choice of species.—A system's attractor is built with observations on a single species. In theory, the choice of species to reconstruct the entire system's attractor is arbitrary, as every species reflects the same underlying attractor (Abarbanel 1996). Previously, CLSs have been generated from only a single species within the system, implicitly assuming that all species will indicate the same CLS (Keeling et al. 1997, Pascual and Levin 1999). We tested this assumption for each model by comparing the CLSs estimated from different species in the same system.

Robustness to parameters of attractor reconstruction.—The estimated CLS will depend in part on the accuracy of the attractor reconstruction, which itself depends on appropriate choices of the reconstruction parameters (τ, d_e, and k, see Appendix B; Buzug and Pfister 1992, Kantz and Schreiber 1997). No single unambiguous value exists for any of these parameters for a particular reconstruction, and indeed different techniques to estimate these parameters commonly yield dissimilar values (Buzug and Pfister 1992, Schreiber 1995, 1999, Kantz and Schreiber 1997). Thus, we examined the sensitivity of the estimated CLS to a range of reasonable potential choices of these parameters, as indicated below.

1. *Time delay (τ).*—From a mathematical perspective, the choice of delay is arbitrary because the data set is assumed to be infinitely long (Kantz and Schreiber 1997). However, for a finite set of data, the choice
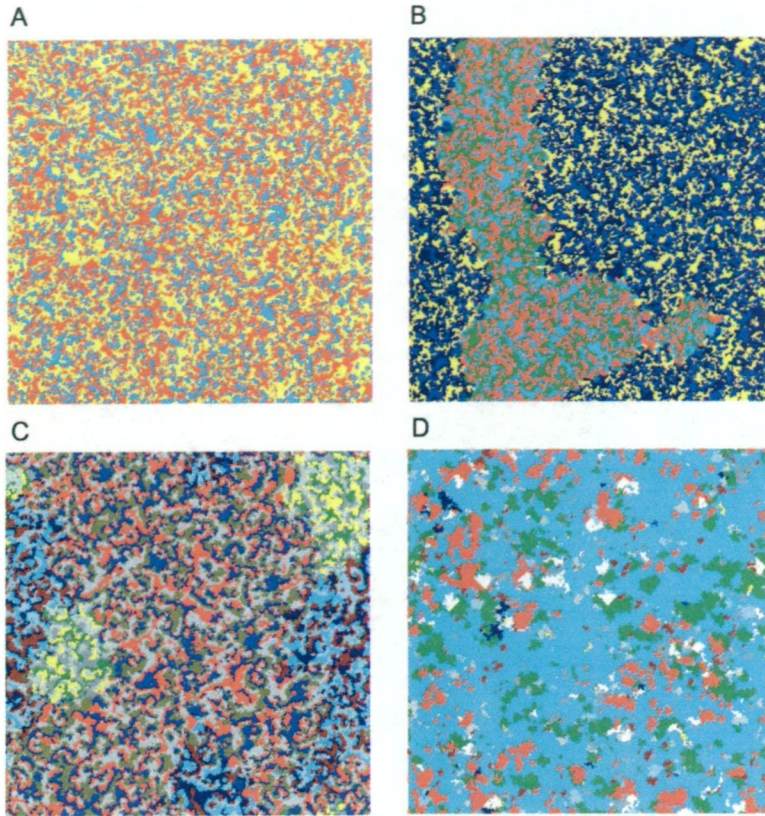
A   B

C   D

FIG. 1. Examples of 500 × 500 landscapes at time step 10 000 of the four model systems used to generate data: (A) three-species system, (B) six-species system, (C) 12-species system, and (D) coral reef system with 12 physiognomic groups. Each color represents a different species or group. Groups selected for analysis from the coral reef system are the dominant turf and coralline algae (light blue), digitate and corymbose Acroporidae (red patches), and Faviidae (dark gray patches). The different models represent a range of complexity of structure, dynamics, and spatial self-organizing. Note that some models self-organize at several spatial scales. All model runs used random initial configurations of recruits covering 10% of the landscape at time step zero.

of τ dictates the quality of the reconstructed trajectory (Liebert and Schuster 1989). If τ is too small, the co-ordinates $x(t_i)$ and $x(t_i + \tau)$ will be almost identical, offering redundant information about the state space. Alternatively, if τ is too large, the coordinates will be almost uncorrelated and their connection to one another is effectively random (Abarbanel 1996). The goal is to determine the delay where coordinates are independent while preserving their dynamical relationship (Nichols and Nichols 2001). We chose τ as the first minimum point of the mutual information (MI) function (Appendix B). However, because identifying the first minimum point from the plot of MI vs. τ requires subjective interpretation, we assessed the robustness of CLS estimates for a variety of choices of τ that might be considered reasonable using the MI approach.

2. *Embedding dimension* (d_e).—If an attractor is projected in too few dimensions, the observed orbits will overlap and distinct segments on the attractor become confused (Abarbanel 1996). The appropriate d_e allows the attractor to be sufficiently unfolded in space such that this overlap no longer occurs. Over-embedding (embedding in too many dimensions) requires larger numbers of co-ordinates, and increases the likelihood of noise in the dimensions of the embedded space where no dynamics are operating (Kennel et al. 1994). The d_e is chosen as the smallest dimension required to sufficiently unfold the attractor, and is indicated as the first minimum of the false nearest neighbors vs. dimension curve (see Appendix B). We assessed the robustness of CLS estimates over a range of embedding dimensions around the value suggested by the false nearest neighbor method.

REBECCA L. HABEEB ET AL.

3. *Number of nearest neighbors (k).*—The number of nearest neighbors (k) used in the prediction is a tunable parameter that influences the quality of the prediction (see Appendix B). If too few neighbors are picked, then important nonrandom information may be missed. If too many are picked, the points may be widely spread in space, decreasing the accuracy of the prediction (Kantz and Schreiber 1997). We assessed the sensitivity of CLS estimates over a range of reasonable choices of k.

Alternative methods of CLS detection not dependent on long time series

To address the problem of unrealistic data requirements of the existing approaches based on analysis of long time series, we derive two alternative methods which we term the "short time series" and "sliding window" approaches. Both new methods use prediction r^2 (because this metric proved most robust), and substitute repetition in space for repetition in time. In the "short time series" method, time series are obtained over only three to four time steps, but at many locations in space. In the "sliding window" method, space is wholly substituted for time by sliding a window of observation across space at a single point in time. The logic underpinning these developments is that distant locations in space within the same dynamical system are likely to be at different points on the system's attractor. If so, then sampling sequentially in space should produce a reconstruction sufficient to replace long time series sampling. Substituting space for time has been successful in other contexts (see, for example, Allain and Cloitre 1991, Marcos-Nikolaus et al. 2002).

Short time series analysis.—In this approach, short time series from multiple locations in space are embedded to reconstruct a system's attractor piece-wise. Our data are from landscapes consisting of a large but finite array of contiguous cells where each cell can be occupied by a single species. We position windows so that they overlap spatially, with successive windows displaced by the width of a single cell, either horizontally or vertically (Fig. 2A). A short time series is obtained for each window. Thus, the short time series approach samples the whole landscape at every time step over a short time. The embedding parameters τ and d_e for attractor reconstruction are not determined in the same way as for long time series analysis. In principle, any value of τ should suffice for an embedding (Takens 1981), but in practice some delays are more effective (Kantz and Schreiber 1997). Here, we use a time delay of one and we ensure that the spacing between landscapes allows the system to evolve from one time step to the next. The number of embedding dimensions d_e is the number of time steps sampled minus one, as the final time step provides a data set against which the accuracy of predictions is assessed.
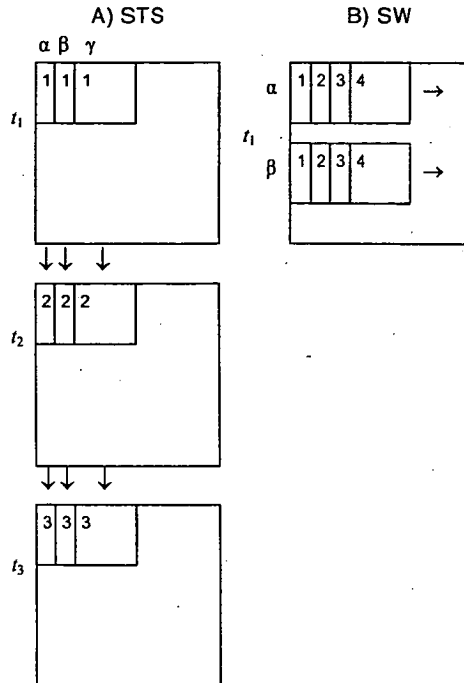

FIG. 2. Diagrams of (A) the short time series (STS) and (B) sliding window (SW) methods of attractor reconstruction. The large square represents an entire landscape that is pixelated at some resolution such that each pixel is occupied by a single species. The small squares represent windows of observation on the landscape. (A) The short time series approach uses a temporal sequence of only three or four consecutive time steps (e.g., $t_1 - t_3$). At each time step we sample over the entire landscape, with successive windows separated by a single pixel, horizontally or vertically. In this particular representation, there are three time steps t_1, t_2, and t_3, and three data series at positions α, β, and γ. Three short time series are generated, and the system's attractor is reconstructed piecewise using the delayed coordinates. (B) In the sliding window approach, multiple short data series are generated for a single time step t_1 by sliding windows horizontally across the landscape. The vertical position of each series differs by one pixel, so the entire landscape is sampled. The sequence in each data series as the window of observation slides across the landscape is indicated as 1, 2, 3, 4.... In this example, two data series are shown at positions α and β.

Embeddings for different values of d_e are evaluated in the results.

A modified prediction algorithm is required for attractors reconstructed from short time series. When using long time series, changing the window size does not change the number of points used to predict the attractor, because replication occurs through time. However, in short time series embeddings where replication is partially spatial, larger windows will gen-

erate fewer points in the reconstructed attractor than smaller windows, which could bias prediction accuracy. Thus, predictions for the short time series are based on a random sample of 1000 windows, which compensates for the discrepancy between small and large windows. Random samples are also used for k nearest neighbor predictions in the sliding window approach described below.

Sliding window analysis.—The sliding window sampling method also uses multiple short data series to reconstruct an attractor in sections, but there is no temporal component. A single landscape is sampled by sliding windows of observation cell by cell (or pixel by pixel) horizontally across the landscape, in a manner similar to the "gliding-box" described by Allain and Cloitre (1991) and Plotnick et al. (1996). Each data series begins one cell lower than the previous series so that the whole landscape is sampled (Fig. 2B). Attractor reconstruction from a spatial series assumes that spatial data capture the general shape of the system's dynamics in a similar way to a time series. The assumption is reasonable if distant parts of the landscape are out of phase, that is, if they are on different parts of the system's attractor. Because of the large model landscapes used in the current study, this assumption is met for all analyses.

There are two sensible options in selecting the spatial delay, τ, in the sliding window attractor reconstruction. One approach is to use the same delay for every window size, as for the time series embeddings. However, for spatial sampling the units of embedding delay are cells (or pixels) rather than time steps. Thus, a potential problem with using delays of a fixed number of cells is that successive embedding dimensions will have a greater overlap for large window sizes than for small window sizes, which could bias CLS estimates. An alternative is to use delays set as some proportion of the linear dimensions of the window, $\tau = \alpha \times$ window length, provided that this proportion is some whole number of cells. We evaluated fixed and proportional delays by assessing their performance for the "null case," that is, landscapes composed of independent, discrete valued pixels (results of comparison are presented in Appendix D). The results indicated that proportional delays were more appropriate, and so we adopted proportional delays in our sliding window analyses.

RESULTS

Robustness of existing techniques based on long time series

For each model system, we compare spectra based on error X and prediction r^2. Length scales determined using error X (after Keeling et al. 1997) are designated as CLS_k, while those based on prediction r^2 spectra (after Pascual and Levin 1999) are designated as CLS_p.

Robustness of CLS estimates to initial conditions

For the simplest model system (three species), spectra of both error X and prediction r^2 were of the expected shape, increasing to an asymptote as a function of window length, and demonstrating a single length scale (Fig. 3A). As expected, CLS_k (50–60 cells) was slightly larger than CLS_p (30–40 cells; Pascual and Levin 1999). There was little variation between runs with different initial conditions. However, with increasing complexity of the model system, CLS_k became more difficult to determine and confidence intervals around the error X spectrum broadened considerably (Fig. 3B, C, D). The error X spectrum for the 12-species system was not the expected positive asymptotic shape, and was not readily interpretable (Fig. 3C). Similarly, based on error X, no length scale could be determined for the coral reef model because the curve deviated dramatically from the expected shape with no inflection or asymptote (Fig. 3D). This was true of individual spectra as well as the mean spectrum depicted in Fig. 3D. Conversely, CLS_p was not highly sensitive to changes in initial conditions for any of the models, and produced an interpretable curve for the full range of models we examined (Fig. 3).

Robustness of CLS estimates to choice of species

Estimates of CLS_k and CLS_p in the three-species model system were not dependent on species identity. All three species are ecologically equivalent in the system (Johnson and Seinen 2000), and displayed curves identical to those in Fig. 3A. In contrast, curves for the more complex model systems were often highly sensitive to the choice of species (for example, Figs. 4 and 5). CLS_k depended heavily on the species used in the attractor reconstruction for all three complex model systems. For example, in the 12-species system, CLS_k ranged from 50–60 cells for one species to 240–250 cells for another (Fig. 5A). For spectra based on error X, species in the same self-organized patch demonstrated similar curves (Figs. 4A, 5A, C). However, not only were error X spectra dissimilar for species from different patches in the same run (Figs. 4A, 5A, C), but the overall shape of the curve, and therefore the CLS_k indicated, changed markedly among runs for the same species. For example, one species in the 12-species system had a CLS_k of approximately 80–90 cells in one run (Fig. 5A), but 290–300 cells in another (Fig. 5C). The error X spectra from the coral reef model displayed little variation from run to run, but only spectra for some species were interpretable (not shown).

As model complexity increased, prediction r^2 spectra also usually separated into groups that corresponded to species within the same spatially self-organized patch. However, CLS_p estimates of the different groups of species were more similar than indicated by spectra based on error X (e.g., Fig. 4A vs. 4B). In the two most complex systems, CLS_p was more sensitive to species,
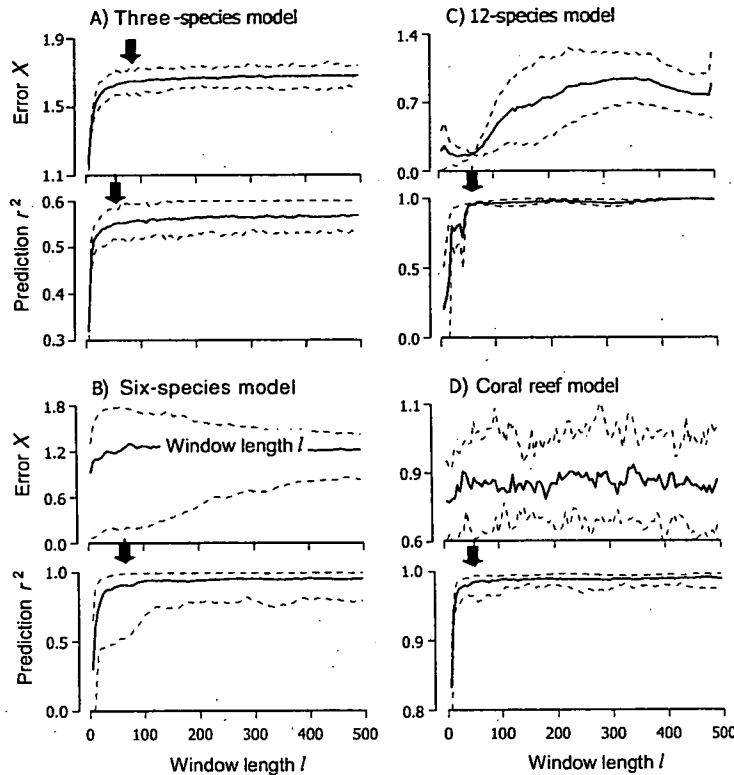
FIG. 3. Average spectra (solid lines) based on error X (Keeling et al. 1997) and prediction r^2 (Pascual and Levin 1999) and the 95% confidence intervals (dotted lines) for a selected species, calculated from a Monte Carlo of 100 independent runs for the (A) three-species model, (B) six-species model, (C) 12-species model, and (D) coral reef model. For the three-, six-, and 12-species models, the CLS was determined for a randomly selected species. The CLS for the coral reef model is based on the spatially dominant group, comprising turf and coralline algae. Estimated CLSs are indicated by arrows where curves are interpretable.

with the scale of the first asymptote differing by up to 50 cells between species groups (Fig. 5B, D). For the six- and 12-species systems, the curves of prediction r^2 for a given species sometimes demonstrated multiple peaks (Figs. 4B, 5B), indicating the potential for more than one length scale in a single system.

Robustness to choices of time delay in attractor reconstruction.—Changes in time delay of the three-species system shifted the CLS curves on the y-axis, but did not change the magnitude of the CLS estimate for either method (Fig. 6A, B). Similarly, time delay had little effect on estimates of the CLS interpreted from curves of the six-species system. However, in the two most complex model systems, CLS_k was notably less robust, with the shape of the curve changing with delay. For the coral reef model in particular, the error X curves were ambiguous at best, and varied with the delays (Fig. 6C). Conversely, CLS_p was robust to changes in

delay in all four model systems we examined (e.g., Fig. 6D).

Robustness to choices of embedding dimension in attractor reconstruction.—Changes in the embedding dimension of the three-species and six-species systems shifted the curves on the y-axis, but had no effect on estimates of CLS_k or CLS_p. In the 12-species system, the shape of the error X curve remained robust to increasing dimension, but interpreting the curves to define the CLS was difficult (Fig. 7A). For the coral reef system, the error X spectra indicated an interpretable CLS_k only with an overly large embedding dimension (Fig. 7C). For both of these more complex systems, the estimates of CLS_p were reasonably robust to changing dimension (Fig. 7B, D).

Robustness to choices of k nearest neighbors in attractor reconstruction.—The number of k nearest neighbors used to reconstruct the attractor of the three-
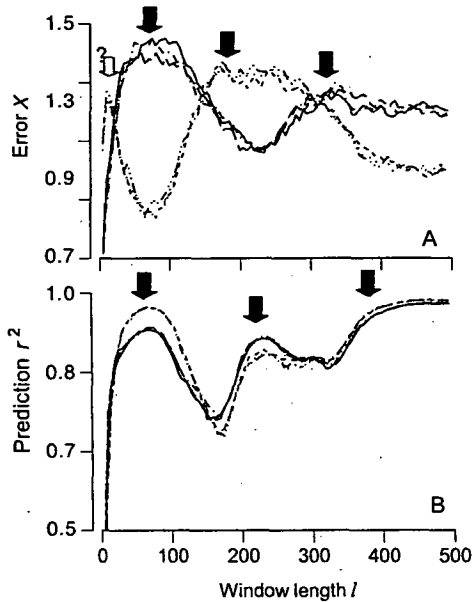
FIG. 4. Variance spectra for each species of the six-species model system based on (A) error X and (B) prediction r^2. In this system, species spatially self-organize into two groups of three (Fig. 1B), which is reflected as two groups of curves on the graphs. Data shown are for a single run of the model. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. Note the potential of multiple CLSs, i.e., length scales larger than the primary (smallest) CLS.

species system had no effect on estimates of CLS_k or CLS_p. However, with the more complex systems, the shape of the error X spectra shifted from an inverted shape with low numbers of neighbors to a curve of the expected shape with excessively high numbers of neighbors (Fig. 8A). In contrast, spectra based on prediction r^2 for the same system were robust to varying numbers of neighbors in attractor reconstruction, and indicated a consistent length scale (Fig. 8B).

Alternative methods of CLS detection

Our results indicate clearly that the prediction r^2 measure (Pascual and Levin 1999) is robust in providing an unambiguous estimate of CLS_p across a range of model complexities, initial conditions, choice of species, and choice of parameters in attractor reconstruction. Because of this robustness we chose to use prediction r^2 in our alternative methods, which differ from the original approach of Pascual and Levin (1999) only in their approach to attractor reconstruction. We refer to the original method of CLS estimation based on prediction r^2 (Pascual and Levin 1999) as a "long time

series" method, which distinguishes it from the new methods we derive based on very short time series (the "short time series" method) and on a single spatially resolved landscape (the "sliding window" method).

Short time series analysis.—Our important overall result is that prediction r^2 spectra derived from attractor reconstruction based on the "short time series" method indicated unambiguous length scales for all four models we examined (Figs. 9B, C, 10C, 11B, 12D). Moreover, for any one model, the smallest CLS (i.e., primary CLS) indicated from spectra produced using both long and short time series were similar (Table 1). It is also encouraging that CLS estimates from prediction r^2 curves derived using the short time series approach are robust to the number of embedding dimensions used for attractor reconstruction (Fig. 9B, C). While results are presented for embeddings of only two and three dimensions for the three-species system, analyses using larger numbers of embedding dimensions (eight, nine, or 10 dimensions) also indicated similar CLSs (not shown). CLS estimates for the six- and 12-species systems and for the model coral reef system were also robust to the number of embedding dimensions, and spectra were consistent when the length of the time step was varied (results not presented).

Prediction r^2 curves were consistent between model runs for the three-, six- and 12-species systems indicating CLSs of 20–40 cells for the three-species system, 30–50 cells for the six-species system and 40–60 cells for the 12-species system (Figs. 9B, C, 10C, and 11B respectively). The CLS for turf and coralline algae in the coral reef model was clearly defined between 20 and 40 cells (Fig. 12D) but spectra derived from corymbose and digitate acroporid and faviid corals were more variable and not readily interpretable (Fig. 12E, F). Changing the length of the time step or changing the number of time steps used in the analysis did not affect variability in results among different runs for these coral groups which occur at relatively low cover.

Two notable differences between CLSs estimated using the long time series and short time series methods arose. First, prediction r^2 curves developed from the long time series analyses were not consistent between species, while prediction r^2 curves from short time series analyses were (for example, cf. Fig. 10A, C). Second, curves of species using long time series often indicate several peaks, suggesting the possibility of multiple length scales (for example, Fig. 4B), while only a single CLS is usually evident in spectra from the short time series. Both approaches yield a similar primary CLS, which is the smallest length scale and is indicative of local dynamics among species. We interpret additional larger length scales to indicate emergent dynamics.

A characteristic of the six-species system on small landscapes (~200 × 200 cells) is that over many time steps and without disturbance, the system's attractor is nonstationary. A shifting attractor is evidenced by a
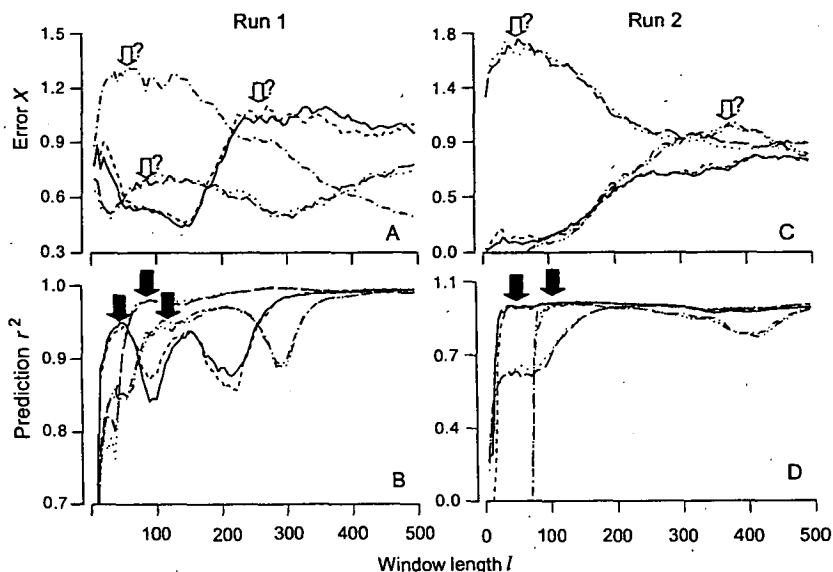
FIG. 5. Variance spectra for six species of the 12-species model system shown from two separate runs. For run 1, spectra are based on (A) error X and (B) prediction r^2, and for run 2, spectra are given for (C) error X and (D) prediction r^2. Species organize into three groups of four on the landscape, and species in the same patch on the landscape show similar CLS curves. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. For clarity, we have identified only the primary length scales for each group of species; however, multiple peaks in the curves (B) may be evidence of multiple length scales.

particular patch type expanding to eventually dominate the landscape (e.g., Fig. 13A, B), and/or coalescence of many small patches of a particular type into a single large patch of that type. The identity of the dominant patch type and the rapidity with which it realizes dominance, or the rate at which small patches of a particular type coalesce to form large patches, may vary from one run of the model to the next. In the example in Fig. 13, the community composition of the 6-species system changed gradually over 5000 time steps, so that one patch type became more abundant at the expense of the other (cf. Fig. 13A, B). Interestingly, using short time series analyses, this shift was reflected as a change in the shape of spectra, and therefore as a change in the CLS estimates derived at the beginning and at the end of the run (compare Fig. 13C, D). This kind of nonstationary behavior in the six-species system can be eliminated and the system made stationary by introducing a low level of disturbance, and allowing open recruitment of all species to disturbed areas with equal probability. As was the case for the short time series analysis (not shown), the long time series analysis of a stationary six-species time series produced almost identical prediction r^2 curves and CLS estimates for all six species (Fig. 13F). This is quite different from the non-stationary case where the spectra and CLS estimates for the two groups of three species based on

analysis of long time series were distinctly different (Fig. 13E). Unlike the six- and 12-species systems, the coral reef model is essentially stationary (after a self-structuring period of ~200 time steps) and so variability in CLS estimates between runs was minimal.

Sliding window analysis.—The most obvious difference between the prediction r^2 spectra produced by this method and those derived using the short and long time series approaches is that the spectrum is inverted relative to the shapes based on short and long time series analysis. Thus, the spectra yield interpretable estimates of length scales as the first minimum in the curve. Likely reasons for the inversion are addressed in *Discussion*.

In determining parameters for this method of attractor reconstruction, we have shown that delays proportional to the window size are more appropriate than are fixed delays (Appendix D). However, the question of how to select a suitable proportional delay arises. For all four model systems, the first minimum on curves of scaled mutual information vs. τ was at $\tau = 0.8 \times$ window length, suggesting that this is the most appropriate embedding delay for attractor reconstruction. However, sliding window analyses for the three-species system indicated that CLS estimates were sensitive to the choice of proportional delay (cf. Fig. 9D, E). Estimated CLS ranges increased from 20–30 cells to 50–
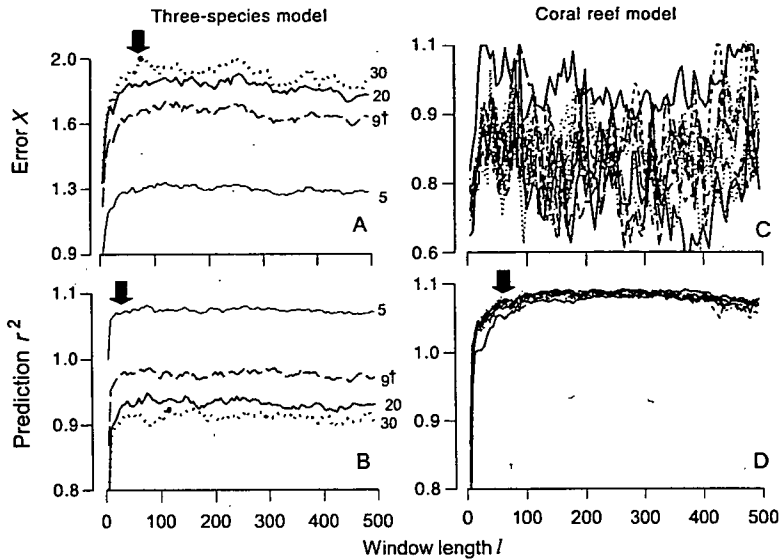
FIG. 6. The effect of choice of time delay (τ) on spectra constructed from one species of the three-species system (A and B) and the dominant species group (filamentous turf and coralline algae) of the coral reef system (C and D). In (A) and (B), the time delays are shown to the right of the relevant spectrum based on (A) error X and (B) prediction r^2. A dagger (†) denotes the delay indicated from the mutual information method. In (C) and (D), different lines denote spectra calculated from delays between 20 and 200 (in steps of 20) based on (C) error X and (D) prediction r^2. Arrows indicate the estimated CLSs of these spectra.

60 cells when the delay changed from τ = 0.8 × window length to τ = 0.2 × window length, respectively. Despite this sensitivity, the most easily interpretable curve was provided by using the delay indicated by the scaled mutual information (here τ = 0.8 × window length; Fig. 9E). Moreover, using this delay indicated a CLS virtually identical to that based on analysis of a long time series from the same system (cf. Fig. 9A, E).

Indeed, CLSs estimated with τ = 0.8 × window length using the sliding window method corresponded to those estimated from analyses of long time series for all four model systems. CLS estimates for the six- and 12-species systems were between 30–40 cells and 20–30 cells respectively (Figs. 10D, 11C). All physiognomic groups in the coral reef model indicated a similar CLS range of 20–40 cells (Fig. 12G–I). However, just as was the case for the 3-species system, CLS estimates for the more complex systems were also sensitive to the proportional delay chosen. Notably, for a given value of τ, the prediction r^2 spectra were generally consistent between model runs and across species within a system.

Comparison of CLS estimates derived by all methods

For each of the four model systems that we analyzed, the three different techniques yielded similar estimates

of the smallest CLS (Table 1). CLS ranges for the six- and 12-species systems were alike across all the different methods, and were consistently larger than those for the 3-species system. Our short time series and sliding window approaches provided more precise CLS estimates for the six- and 12-species system than did the long time series method. However, prediction r^2 spectra produced using long time series analysis more often suggested multiple CLSs for a given species within the six- and 12-species systems than did the alternative analyses (Table 1).

DISCUSSION

The issue of spatial scale is a central theme in ecology. The scale at which a system is observed affects relationships between pattern and process, and between space, time, and organizational complexity (Levin 1992, 2000, Levin et al. 1997, Tilman and Kareiva 1997, Tyre et al. 1997, Dieckmann et al. 2000, Wilson and Keeling 2000, Molofsky et al. 2002). While investigators have acknowledged the need to address ecological questions at appropriate scales (Carlile et al. 1989, Wiens 1989, De Roos et al. 1991), how these "appropriate scales" are identified is often ambiguous. The application of methods from nonlinear time series analysis (Rand and Wilson 1995, Keeling et al. 1997, Pascual and Levin 1999) has refined a crucial aspect
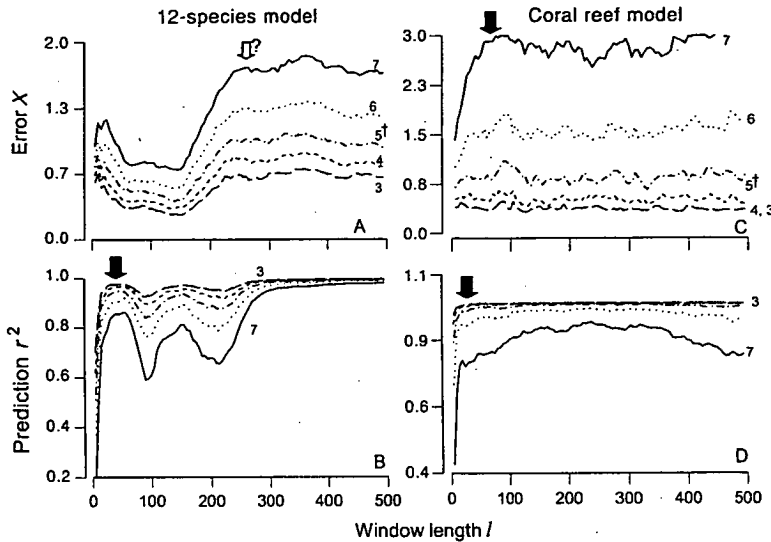
FIG. 7. The effect of choice of embedding dimension (d_e) for a single species in the 12-species system (A and B), and for the dominant species group (filamentous turf and coralline algae) of the coral reef system (C and D). Different lines denote particular embedding dimensions (shown to the right of curves) used in attractor reconstruction based on (A and C) error X and (B and D) prediction r^2. A dagger (†) denotes the dimension indicated using the false nearest neighbors method. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear.

of the study of scale in ecology, allowing a shift in focus from observing mean behaviors to extracting the deterministic signal from dynamical systems. Towards the ultimate goal of estimating CLSs of natural systems, here we evaluated the robustness of these relatively recent methods, and then developed modifications to reduce the data requirements necessary to estimate CLSs. We first discuss the robustness of the original long time series methods, and then evaluate the general behaviors of our alternative methods in detecting the smallest or primary length scale of spatial systems. We examine the interpretation of multiple length scales, as indicated by different species groups or by a single species in the system, and finally, we comment briefly on the possible future of CLSs in ecology.

Robustness of estimates

Both the methods of Keeling et al. (1997) and Pascual and Levin (1999) maintained high robustness and low sensitivity to parameter choices in the simplest model system. As expected (Pascual and Levin 1999), CLS_k was consistently larger than CLS_p, reflecting the need for larger "windows of observation" to observe the full spatial dynamic than to accurately predict trends in dynamics. However, as the complexity of the model systems increased, the error X spectra of Keeling et al. (1997) became more difficult to interpret, and

spectra for the 12-species and coral reef model systems were not interpretable (Fig. 3C, D). Also, with increasing model complexity, error X spectra, and therefore estimates of CLS_k, were increasingly sensitive to changes in initial conditions, to the species on which the analysis focused, and to values of parameters used in attractor reconstruction. The reason for the ambiguity and, in some cases, failure (Fig. 6C) of error X spectra to indicate a CLS is unclear. Error X may be more susceptible to random noise than prediction r^2 or, alternatively, may be more sensitive to a system's complexity and particular dynamics than is prediction r^2. However, even if the observed sensitivity of these spectra are accurate reflections of subtle features of a particular dynamic, this level of sensitivity is not helpful towards the overall objective, which is to identify the CLS of a system irrespective of the particular state of that system over a finite period of observation. Further research may demonstrate that the sensitivity of error X reveals useful information about a system, such as identifying features of spatial pattern. However, while this may be useful for other applications, it is unsuitable for estimating CLSs of complex systems.

In contrast, the Pascual and Levin (1999) method provided interpretable spectra and thus clearer estimates of the CLS with increasing model complexity. The prediction r^2 spectra from all model systems were consistently interpretable and the CLS estimates of a
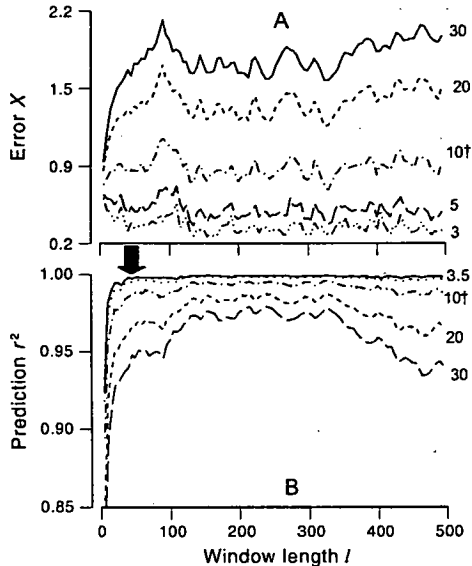
FIG. 8. The effect of increasing numbers of k nearest neighbors (k is given to the right of the curves) for the dominant species group (filamentous turf and coralline algae) of the coral reef model system, based on (A) error X and (B) prediction r^2. A dagger (†) denotes the k chosen for use in all other model runs. The solid arrow indicates the estimated CLS.

particular system were not sensitive to changes in parameters for attractor reconstruction, or to varying initial conditions. Thus, prediction r^2 spectra using the approach of Pascual and Levin (1999) appears to be a robust means of estimating CLSs of complex ecological systems. We suggest that this metric is likely to be useful and reliable in estimating CLSs of natural ecosystems.

Estimating CLSs: alternative methods without long time series

While Pascual and Levin's (1999) approach to detecting CLSs accommodates oscillatory behavior in dynamic systems and proves reasonably robust over a range of complexities of model ecosystems, the technique requires long time series of data which are not attainable for most natural systems. Our two alternative approaches have considerable potential for overcoming this problem in estimating the CLSs of real ecosystems. These methods detect primary CLSs similar to those determined with the original long time series method (Table 1), but require a maximum of only three to four spatially resolved landscapes instead of thousands.

For the remainder of the discussion, reference to CLSs estimated from the long time series method refers

to those derived using the original method of Pascual and Levin (1999).

CLSs from short time series.—The short time series method uses variability through time as the basis to derive CLS estimates, but employs a unique approach of sampling over the entire landscape at every time step. This generates sufficient data points for nonlinear time series analysis, but removes the requirement for a long series of sequential observations. The underlying motivation of this work is to develop robust techniques for application to real ecosystems, and the requirement of only three or four sampling occasions is realistic for many applications in ecology. Results for the three-species model indicate that the technique generates sensible CLS estimates and is robust to the number of time steps considered. Our more general investigations of the three-, six- and 12-species and coral reef models also indicate that CLS estimates are robust to both the length of the time step and to the number of time steps sampled (detailed results not provided here).

Another advantage of using information from only a small number of time steps is that ecological systems are likely to be essentially stationary in time over the sampling period, whereas the same may not be true of a system's attractor over a longer time period. Furthermore, the short time series approach can be used to detect temporal nonstationarity as a change in the CLS over time (e.g., Fig. 13C, D). Unlike simpler measures such as changes in species' abundances (which occur without any shift in the attractor of a system), a change in the CLS through time indicated from short time series analysis reflects a shift in the underlying dynamics (i.e., the attractor). Note that, in the six-species system for which we demonstrated the shift in the attractor using this approach, the network interaction topology remained constant over time. Thus, the short time series approach could potentially be used to detect fundamental change in natural systems, where a shifting attractor is reflected as a changing CLS. Applied examples of shifting attractors may include systems affected by pollutants or impacted by foreign species.

Similar to the assumption of temporal stationarity that underpins the long time series method, is the assumption of spatial stationarity which applies to our short time series method. This assumption, that species sampled on different parts of the spatial landscape are on the same attractor, cannot be tested by simply examining whether community composition is similar in all areas of space. Different points on the same attractor can, of course, reflect states of vastly dissimilar community structure, for example as arises in simple stable limit cycles. However, in the same way that we detected potential temporal non-stationarity in the attractor by determining the CLS at two different points in time using the short time series method, so could one determine the CLS at two regions in space. If the CLSs are different in different regions of the landscape, then
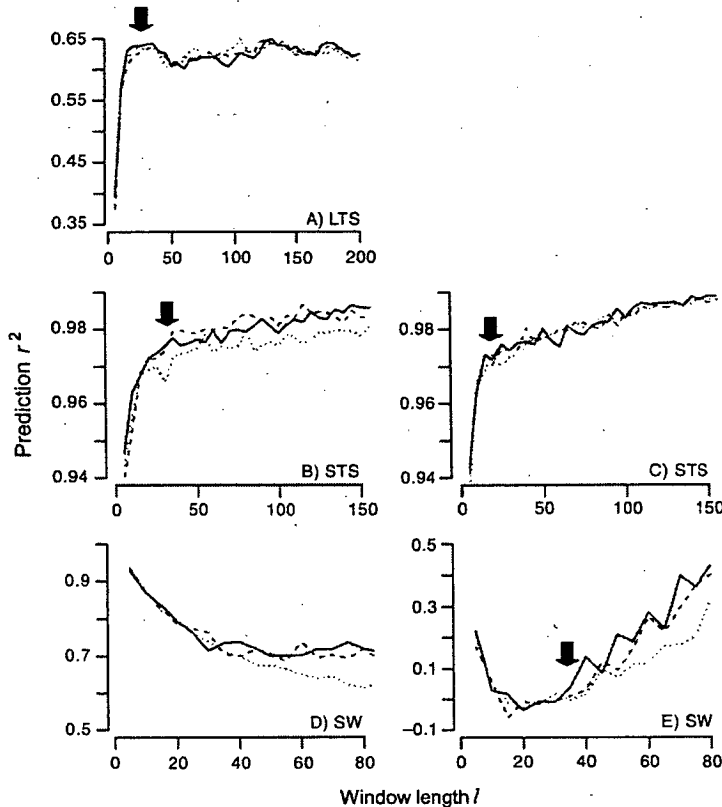
FIG. 9. Prediction r^2 as a function of window length l for the three-species system derived using the three methods. Curves for the three species in the model are shown as different lines. Solid arrows indicate estimated CLSs where curves are interpretable. (A) Long time series analysis of 10000 time steps with time delay $\tau = 9$ and embedding dimension $d_e = 6$; the first 200 time steps are discarded in the analysis. (B and C) Short time series analysis with $\tau = 1$: (B) spectra derived from analysis of three consecutive time steps (498–500; $d_e = 2$); (C) spectra derived from analysis of four consecutive time steps (497–500; $d_e = 3$). (D and E) Analysis using sliding windows: (D) $\tau = 0.2 \times$ window length l, $d_e = 5$; (E) $\tau = 0.8 \times$ window length l, $d_e = 5$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 500 × 500 cells, and the number of nearest neighbors $k = 10$. Key to abbreviations: LTS, long time series analysis; STS, short time series analysis; SW, sliding window analysis.

the attractor cannot be assumed to be spatially stationary across that space. Nevertheless, given the relatively small magnitude of the CLSs and that it is not necessary to sample over large spatial areas to estimate them, we expect that the assumption should hold true provided that a landscape is not sampled across strong environmental gradients or environmental discontinuities likely to realize dissimilar community dynamics.

CLSs could not be identified consistently for the acroporid and faviid corals in the coral reef model based on analysis of short time series. Changing the number of time steps and/or the length of time steps used for attractor reconstruction did not significantly change the shape of the curves. Since both of these coral groups

occur at relatively low densities in this system, the dynamic signal for these groups over short time series may be inadequate for attractor reconstruction and for the determinism test.

CLSs from sliding windows.—Our sliding window method derives CLS estimates from spatial data obtained at a single sampling occasion, using spatial sampling similar to that of others (Allain and Cloitre 1991, Plotnick et al. 1996, Marcos-Nikolaus et al. 2002). The approach replaces temporal variability with spatial variability in reconstructing the system's attractor, but this introduces several complications. In particular, it is clear that "delays" of a fixed number of cells (or pixels) are inappropriate for this method. Consideration
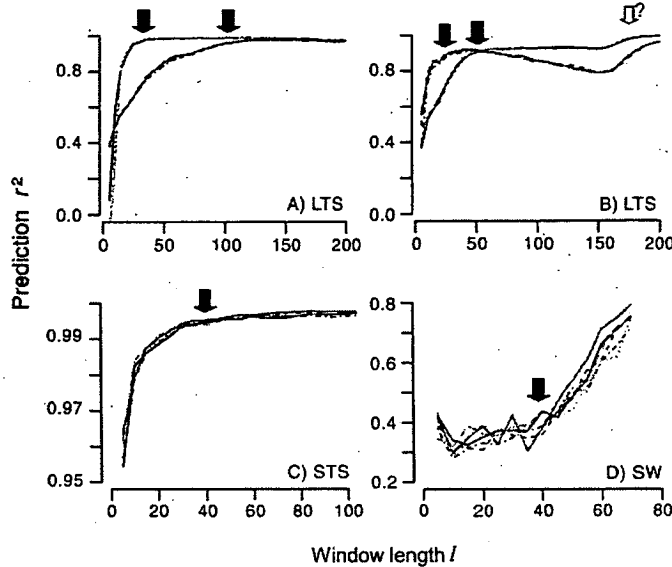
FIG. 10. Prediction r^2 as a function of window length l for the six-species system derived using the three methods. Curves for the six species in the model are shown as different lines. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. (A and B) Long time series analysis of two independent model runs with different initial random configurations of recruits with $\tau = 12$, $d_e = 6$. The analyzed time series contains 10 000 time steps, with the first 200 time steps discarded. (C) Short time series analysis with $\tau = 1$. This analysis is based on three consecutive time steps (498–500; $d_e = 2$). (D) Sliding window analysis of the landscape with $\tau = 0.8 \times$ window length l, and $d_e = 5$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 500 $\times$ 500 cells and $k = 10$. Key to abbreviations: LTS, long time series analysis; STS, short time series analysis; SW, sliding window analysis.

of prediction r^2 spectra produced from landscapes of independent random pixels (i.e., the null case of no spatial pattern) suggests that delays that are proportional to window size are most appropriate. Proportional delays ensure that the overlap of coordinates is constant for all window sizes (see Appendix D).

For all four model systems, when a delay of 0.8 $\times$ window length l was used (i.e., the delay suggested by the scaled mutual information plot), we obtained CLSs similar to those estimated from the long time series analyses. For sliding window analysis, the CLS was taken as the point of increase of the (inverted) prediction r^2 curve. The inverted shape of prediction r^2 curves using this approach indicates that prediction is good at very small spatial scales, but then declines with increasing window size. The CLS is the point where prediction r^2 then begins to improve again, thereafter increasing with window size. This relationship can be explained in terms of spatial patterns that arise on the model landscapes. At small scales of observation below the size of individual colonies, prediction is good because there is a high likelihood that successive windows in any one "sliding series" will be the same species. At slightly larger scales prediction will be

poorer, because more than one species will occur in sampling windows but the scale of observation will be insufficient to capture the community dynamic among colonies. As window size continues to increase, a point is reached where the spatial pattern becomes more predictable, reflecting the deterministic dynamic among colonies (Johnson 1997, Johnson and Seinen 2002). At this point there is a rise in prediction r^2, indicating the CLS.

There are some constraints to the general application of the sliding window technique. Although the approach dramatically reduces data requirements compared with the original method based on long time series, the amount of spatial data necessary for the method can be large. Depending on the magnitude of the delay, the maximum possible window size for sliding window analysis may be much smaller than the landscape size. It is therefore necessary to select a landscape size that ensures that the maximum possible window size is larger than the anticipated CLS. Because replication occurs in space, landscapes must generally be larger than those used in the long time series method, where replication occurs by sampling through time. Similarly, landscapes must also be large when using
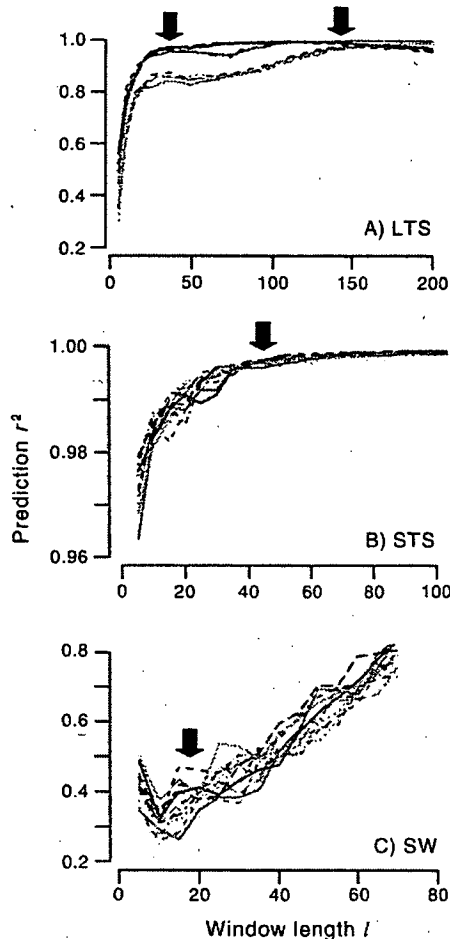
REBECCA L. HABEEB ET AL.

consideration is that, while the shape of the prediction r^2 curve in the null case of no spatial pattern is acceptable to allow interpretation of CLSs where they do arise, we cannot verify mathematically that this is the expected shape (Appendix C). The reason for sensitivity of CLS estimates to the value selected for the proportional delay is also unclear. However, despite these caveats, the important finding is that using the proportional delay indicated by the minimum mutual information produces a clearly interpretable prediction r^2 spectrum which indicates a CLS similar to that estimated from analysis of a long time series of the same system (Table 1).

Multiple length scales

While different scales will undoubtedly exist that reflect other properties of ecosystems, here we sought the natural scales at which to optimally observe a system's dynamics. It was not surprising that we detected multiple scales in the more spatially complex systems. Several length scales were detected in the complex model systems, either because different species groups indicated different CLSs or a single species indicated multiple CLSs. The mechanisms underpinning these two phenomena are different. Different species in the same system can indicate different scales when the connectivity between groups of species in patches is low. The prediction r^2 spectrum of a single species can also display several length scales, but with the larger scales indicative of that species' emergent dynamics. We discuss each scenario.

Multiple length scales indicated by different species.—In theory, any species in a system can be observed to indicate the length scale of that system, irrespective of its life history or dynamical behavior. However, biologists might intuitively suggest that species with dissimilar life history parameters could be expected to yield dissimilar length scales for the same system. Indeed, our results for the more complex models suggest that estimates of CLS_p show some dependence on the species used for attractor reconstruction. Analyses based on both short and long time series could indicate different scales for different species, but the phenomenon arose more readily in analyses based on long time series.

In long time series analyses, scale differences among species occurred in systems that were strongly spatially self-organizing at several spatial scales, as in the six- and 12-species systems. Notably, these differences were not related to abundances. For example, in the 12-species model the curves for the species form three distinct groups, each with a different estimate of CLS_p (Fig. 5). While all species in this model are topologically equivalent with respect to network structure (each is able to overgrow three others in a symmetrical network), the system spatially self-organizes into three distinct patch types, with each patch type containing four species. Because species within patches are more
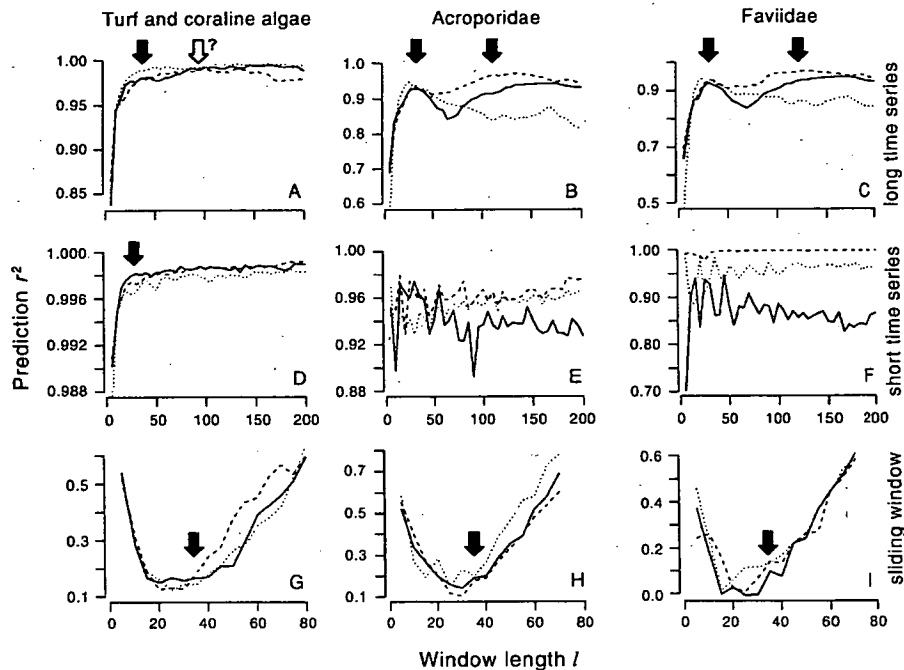

FIG. 11. Prediction r^2 as a function of window length l for the 12-species system derived using the three methods. Curves for the 12 species in the model are shown as different lines. Solid arrows indicate estimated CLSs. (A) Long time series analysis with $\tau = 12$ and $d_e = 6$. The analyzed time series contains 10 000 time steps, with the first 200 time steps discarded. (B) Short time series analysis with $\tau = 1$. This analysis is based on three consecutive time steps (498–500; $d_e = 2$). (C) Sliding window analysis of the landscape with $\tau = 0.8 \times$ window length l, $d_e = 6$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 500 × 500 cells and $k = 10$. Key to abbreviations: LTS, long time series analysis; STS, short time series analysis; SW, sliding window analysis.

the short time series method. Thus, when applying these spatial techniques to natural systems, some prior understanding of the general scales of dynamical processes is required to ensure that the maximum sampling scale is large enough to include its CLSs. A second

Fig. 12. Prediction r^2 as a function of window length l derived using the three methods applied to cover of three physiognomic groups in the coral reef system: (A, D, G) turf and coralline algae, (B, E, H) corymbose and digitate Acroporidae, and (C, F, I) Faviidae. Curves for three model runs with different initial random configurations of recruits are represented as different lines. Solid arrows indicate estimated CLSs where curves are interpretable, while open arrows with question marks indicate ambiguous CLSs where interpretation is unclear. (A, B, C) Long time series analysis with (A) $\tau = 100$, $d_e = 5$; (B) $\tau = 75$, $d_e = 6$; and (C) $\tau = 100$, $d_e = 6$. The analyzed time series contains 10 000 time steps, with the first 200 time steps discarded. Landscape size is 300 × 300 cells. (D, E, F) Short time series analyses with $\tau = 1$. These analyses are based on four consecutive time steps (498–500; $d_e = 3$). (G, H, I) Sliding window analysis of the landscape with $\tau = 0.8 \times$ window length l. For (G) $d_e = 5$, and for (H and I) $d_e = 6$. Parameter choices were based on mutual information and false nearest neighbors techniques. For D–I, landscape size is 500 × 500 cells. For all analyses, $k = 10$.

TABLE 1. Summary of CLS (Pascual and Levin 1999) estimates from long time series (LTS), short time series (STS) and sliding window (SW) analyses for the three-species, six-species, 12-species, and coral reef systems.

System	Analysis	Scale									
		20	40	60	80	100	120	140	160	180	200
Three species	LTS	▨									
	STS	▨									
	SW	▨									
Six species†	LTS		▨						▨		
	STS		▨								
	SW	▨									
12 species	LTS		▨				▨				
	STS		▨								
	SW	▨									
Coral‡	LTS		▨								
	STS	▨									
	SW	▨									

† CLS ranges indicated are for the nonstationary case.
‡ CLS range estimated using short time series analysis is taken for turf and coralline algae only.

REBECCA L. HABEEB ET AL.

likely to interact with each other than with species from other patches given their spatial proximity, the realized connectivity on a spatially organized landscape is higher among species within patches than it is among species between patches. This occurs even though, in this example, the interaction network defines that each species interacts equivalently with others both inside and outside a patch type.

Not surprisingly, highly connected species with tightly coupled dynamics (i.e., species within patches) indicate similar length scales, while species whose dynamics are more weakly linked (because they are spatially separated in different patches) can manifest dissimilar length scales, even though they have identical "life history" attributes. In our systems where different length scales arise for different species, differential connectivity among species arises through spatial self-organizing. We anticipate that other factors which influence the connectivity among species, such as the topology of food webs in which some groups of species are tightly coupled trophically while others are poorly connected (e.g., O'Neill et al. 1986, Johnson et al. 1995), will have a similar effect. Because the approach is based on reconstruction of deterministic dynamics, it should not be surprising that species that are weakly linked dynamically can indicate different length scales for their different behaviors. Conversely, species that are highly connected in a system, as in our three-species model, all indicate very similar estimates of CLS_p.

In the coral reef model, the different physiognomic groups provide similar CLS_p estimates despite notable differences among groups in their life history parameters such as rates of recruitment, growth, and mortality. Marked differences in CLS_p estimates among "species" (in reality, guilds) in this model do not arise because they do not self-organize into distinct patches. In one sense, this system is maximally connected in that each species competes with all others for space and there are no standoffs (Johnson and Seinen 2002). However, the species are not identically connected because they have different interaction strengths and neighbor-specific growth rates.

In short time series analyses using both the six- and 12-species models, the spectra of species from different patches also separated, revealing different length scales among species. With this method, the different CLSs likely reflect the connectivity of spatial patches captured in the set of three to four landscapes that are used to reconstruct the attractor. However, in general, spectra produced from long time series tended to more clearly differentiate among weakly connected species.

In summary, because CLS estimates reflect the strength of dynamical connectivity among species, we predict that for complex real systems different species or functional groups that are loosely connected, either as a result of spatial separation or weak direct interactions, may indicate dissimilar length scales. This will

arise whether the long time series method or the short time series method is used.

Multiple length scales indicated by a single species.—For complex model systems, analyses based on long time series from a single species may detect several different length scales, reflecting the different dynamics within the system. For example, the prediction r^2 spectrum of a single group in the six-species system sometimes displays several critical points (Fig. 4B), which we interpret as multiple length scales. The smallest length scale (the first peak in the curve) is consistent among runs, and is also consistently indicated by our alternative short time series and sliding window methods. This "primary scale" is the scale at which the local dynamic is best predicted, and it reflects the scale of interaction between colonies of different species within patches. Thus, for any one species, we expect at least one CLS, which is influenced by colony size, to reflect the nature of local interactions between species within patches. In the models we examined in which both distinct colonies and clearly differentiated groups of colonies (i.e., patches) formed, the primary CLS is larger than the mean colony size but smaller than the patch size. In real ecological systems, we might expect the primary CLS to indicate the most appropriate scale for monitoring system dynamics, and therefore the scale that most efficiently identifies meaningful trends in species abundances.

Scales larger than the primary scale likely reflect the system's emergent dynamics, which may include the emergence of and interactions among patches, and nonstationary attractors. Because emergent dynamics are highly variable among runs (e.g., the size and shape of patches vary), length scales larger than the primary scale tend to be more variable among runs (e.g., Fig. 5B, D). Thus, while any single run may demonstrate several length scales, the average of several prediction r^2 curves for a given species is likely to indicate only a single CLS (e.g., Fig. 3B). The widening of the 95% confidence intervals around the average curve reflects that the variable secondary peaks are absorbed as noise (Fig. 3B).

While the short time series method occasionally identified secondary CLSs, it is clear that for analyses based on a single species, the short time series and sliding window methods are less likely to identify additional CLSs larger than the primary CLS than is the method based on long time series. This result may indicate that emergent dynamics are more difficult to detect from spatial pattern than from time series. When using the short time series or sliding window method, if it is important to detect secondary length scales reflecting emergent dynamics, it may be necessary to sample many windows that collectively cover areas much larger than the primary length scale to more comprehensively sample the attractor. The tradeoff in this approach is that, as the distance over which a landscape is sampled increases, more care is likely to be required
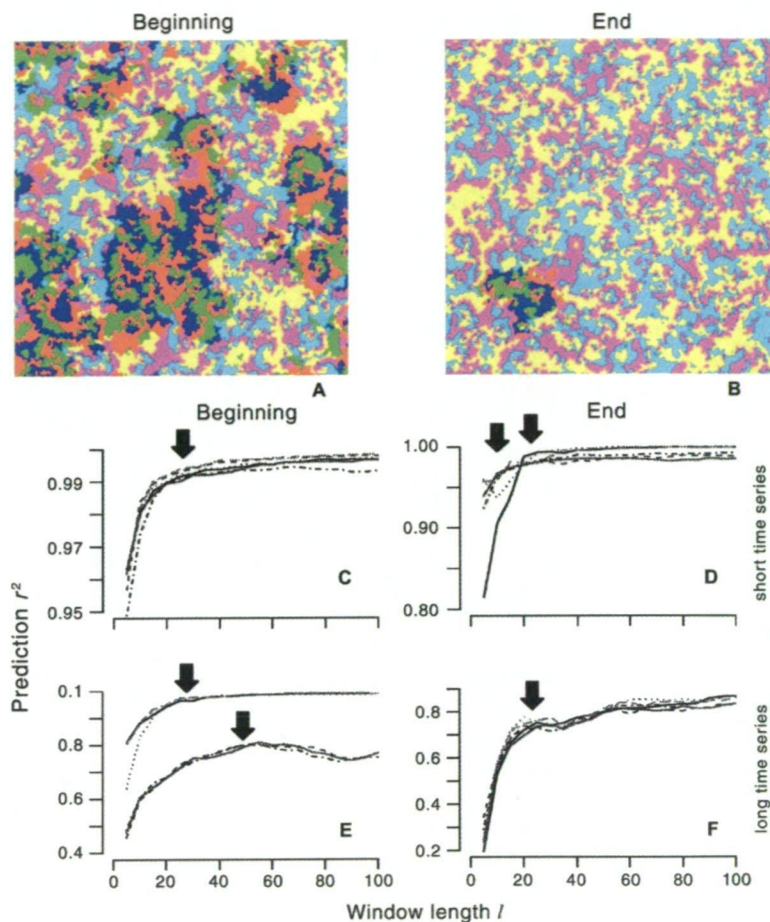
FIG. 13. Comparison of prediction r^2 spectra derived from short and long time series approaches for the six-species system. Arrows identify estimated CLSs. In (C) and (D), short time series analysis uses time steps from the start (A) and end (B) of a 5000-time step time series. These analyses are based on four consecutive time steps ($d_e = 3$) with $\tau = 1$: (C) time steps 201–204; (D) time steps 4997–5000. Divergence in species abundances by the end of the time series is reflected in the two distinct groupings of prediction r^2 curves in (D). (E and F) Long time series analysis for a nonstationary and a stationary system, respectively (see *Results: Short time series analysis* for explanation). The analyzed time series in (E) and (F) contains 5000 data points, sampled at each time step. For (E), $\tau = 12$, $d_e = 6$; for (F), $\tau = 19$, $d_e = 6$. Parameter choices were based on mutual information and false nearest neighbors techniques. For all analyses, landscape size is 200 × 200 cells, and $k = 10$.

to ensure that all samples are represented by the same attractor.

Do CLSs have a future in ecology?

Both the short time series and sliding window techniques we describe offer workable alternatives to the existing approach based on analyses of long time series (Pascual and Levin 1999) for estimating the primary CLS of dynamical systems. The distinct advantage of these new approaches is that their data requirements can be met for natural ecosystems. This is an important step towards the goal of estimating CLSs of real ecosystems. Of our two alternatives, the short time series method seems to be the most robust and interpretable, combining both temporal and spatial data to reconstruct the attractor.

Several questions arise with regards to the application of the new CLS methods to natural ecosystems.

Do these new methods unambiguously identify the primary CLS of the system? We have shown that the short time series method and sliding windows method can be used to readily detect primary CLSs virtually identical to those identified by the original long time series method, in some cases with less ambiguity (Table 1).

Another important issue is whether a system has only a single length scale at which it is optimal to observe dynamics, as suggested by the current theory, or whether there can be many. Most ecologists would argue that any ecological system is likely to manifest several length scales (Levin 2000, Bishop et al. 2002), and it is therefore important for the techniques to have the potential to identify them. Using the long time series method, different length scales are indicated by different species within the same system if their dynamics are weakly linked (e.g., Figs. 4B, 5B, D). Using our new method based on a short series of three to four time steps, species in different patches with weakly coupled dynamics also indicate different CLSs in the six- (Fig. 13D) and 12-species (not shown) model systems. Thus, when applied to a natural system in which subcomponents are weakly linked, this technique should have the ability to indicate that the partially decoupled dynamics among subsets of weakly linked species are best observed at different spatial scales. Similarly, these approaches to estimating length scales can identify secondary scales that indicate emergent dynamics. Finally, we note that detection of natural scales for optimal observation of ecosystem dynamics does not preclude that there may be other natural scales that reveal other ecologically meaningful properties of real ecosystems.

Our overall conclusion is that the development of new techniques that can be realistically applied in ecology to produce prediction r^2 spectra for complex oscillating dynamical systems is useful progress towards objectively defining appropriate scales for observing natural ecological systems. The next challenge is to use the modified techniques to assess what CLSs inform us about natural systems, and to evaluate their utility in providing objective estimates for scaling issues in applied ecology.

LITERATURE CITED

Abarbanel, H. 1996. Analysis of observed chaotic data. Springer-Verlag, New York, New York, USA.

Allain, C., and M. Cloitre. 1991. Characterizing the lacunarity of random and deterministic fractal sets. Physical Review A 44:3552–3558.

Bishop, M. J., A. J. Underwood, and P. Archambault. 2002. Sewage and environmental impacts on rocky shores: necessity of identifying relevant spatial scales. Marine Ecology Progress Series 236:121–128.

Buzug, T., and G. Pfister. 1992. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. Physica D 58:127–137.

Carlile, D., J. Skalski, J. Batker, J. Thomas, and V. Cullinan. 1989. Determination of ecological scale. Landscape Ecology 2:203–213.

Casdagli, M. 1989. Nonlinear prediction of chaotic time series. Physica D 35:335–356.

De Roos, A., E. McCauley, and W. Wilson. 1991. Mobility versus density-limited predator-prey dynamics on different spatial scales. Proceedings of the Royal Society of London B 246:117–122.

Dieckmann, U., R. Law, and J. A. J. Metz, editors. 2000. The geometry of ecological interactions: simplifying spatial complexity. Cambridge University Press, Cambridge, UK.

Dunstan, P. K., and C. R. Johnson. 2005. Predicting global dynamics from local interactions: individual-based models predict complex features of marine epibenthic communities. Ecological Modelling, 186:221–233.

Durrett, R., and S. A. Levin. 2000. Lessons on pattern formation from planet WATOR. Journal of Theoretical Biology 205:201–214.

Ellner, S., and P. Turchin. 1995. Chaos in a noisy world: new methods and evidence from time-series analysis. American Naturalist 145:343–375.

Farmer, J. 1982. Chaotic attractors of an infinite-dimensional dynamical system. Physica D 4:366–393.

Greig-Smith, P. 1952. The use of random and contiguous quadrats in the study of the structure of plant communities. Annals of Botany 16:293–316.

Hastings, A., C. L. Hom, S. Ellner, P. Turchin, and H. C. J. Godfray. 1993. Chaos in ecology: is mother nature a strange attractor? Annual Review of Ecology and Systematics 24:1–33.

Johnson, C. R. 1997. Self-organising in spatial competition systems. Pages 245–263 in N. Klomp and 1. Lunt, editors. Frontiers in ecology: building the links. Elsevier, Oxford, UK.

Johnson, C. R., D. Klumpp, J. Field, and R. Bradbury. 1995. Carbon flux on coral reefs: effects of large shifts in community structure. Marine Ecology Progress Series 126:123–143.

Johnson, C. R., and K. H. Mann. 1988. Diversity, patterns of adaptation, and stability of Nova Scotian kelp beds. Ecological Monographs 58:129–154.

Johnson, C. R., and I. Seinen. 2002. Selection for restraint in competitive ability in spatial competition systems. Proceedings of the Royal Society of London B 269:655–663.

Kantz, H., and T. Schreiber. 1997. Nonlinear time series analysis. Cambridge University Press, Cambridge, UK.

Kaplan, D., and L. Glass. 1995. Understanding nonlinear dynamics. Springer-Verlag, New York, New York, USA.

Keeling, M. J., I. Mezic, R. Hendry, J. Mcglade, and D. Rand. 1997. Characteristic length scales of spatial models in ecology via fluctuation analysis. Philisophical Transactions of the Royal Society of London B 352:1589–1601.

Kennel, M. B., R. Brown, and H. D. I. Abarbanel. 1994. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Physics Review A 45:3403–3411.

Kershaw, K. A. 1957. The use of cover and frequency in the detection of pattern in plant communities. Ecology 38:291–299.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Levin, S. A. 2000. Multiple scales and the maintenance of biodiversity. Ecosystems 3:498–506.

Levin, S. A., B. Grenfell, A. Hastings, and A. S. Perelson. 1997. Mathematical and computational challenges in population biology and ecosystems science. Science 275:334–343.

Liebert, W., and H. Schuster. 1989. Proper choice of the time delay for the analysis of chaotic time series. Physics Letters A 142:107–111.

Little, S., S. Ellner, M. Pascual, M. Neubert, D. Kaplan, T. Sauer, H. Caswell, and A. Solow. 1996. Detecting nonlinear dynamics in spatio-temporal systems, examples from ecological models. Physica D 96:321–333.

Marcos-Nikolaus, P., J. M. Martin-Gonzalez, and R. V. Sole. 2002. Spatial forecasting: detecting determinism from single snapshots. International Journal of Bifurcation and Chaos 12:369–376.

Molofsky, J., J. Bever, J. Antonovics, and T. Newmaan. 2002. Negative frequency dependence and the importance of spatial scale. Ecology 83:21–27.

Nichols, J. M., and J. D. Nichols. 2001. Attractor reconstruction for non-linear systems: a methodological note. Mathematical Biosciences 171:21–32.

O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen. 1986. A hierarchical concept of ecosystems. Princeton University Press, Princeton, New Jersey, USA.

Pascual, M., and S. Ellner. 2000. Linking ecological patterns to environmental forcing via nonlinear time series models. Ecology 81:2767–2780.

Pascual, M., and S. A. Levin. 1999. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. Ecology 80:2225–2236.

Plotnick, R. E., R. H. Gardner, W. W. Hargrove, K. Prestegaard, and M. Perlmutter. 1996. Lacunarity analysis: a general technique for the analysis of spatial patterns. Physical Review E 53:5461–5468.

Rand, D. 1994. Measuring and characterizing spatial patterns, dynamics and chaos in spatially extended dynamical systems and ecologies. Philisophical Transactions of the Royal Society of London A 348:497–514.

Rand, D., and H. Wilson. 1995. Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially extended ecosystems. Proceedings of the Royal Society of London B 259:111–117.

Schneider, D. C. 1994. Quantitative ecology: spatial and temporal scaling. Academic Press, San Diego.

Schreiber, T. 1995. Efficient neighbor searching in nonlinear time series analysis. International Journal of Bifurcation and Chaos 5:349–358.

Schreiber, T. 1999. Interdisciplinary application of nonlinear time series methods. Physics Reports 308:2–64.

Sole, R. V., and J. Bascompte. 1995. Measuring chaos from spatial information. Journal of Theoretical Biology 175:139–147.

Sugihara, G., B. Grenfell, and R. M. May. 1990. Distinguishing error from chaos in ecological time series. Philisophical Transactions of the Royal Society of London B 330:235–251.

Takens, F. 1981. Detecting strange attractors in turbulence. Pages 366–381 in D. Rand and L. Young, editors. Dynamical systems and turbulence, Warwick 1980. Lecture notes in mathematics. Springer-Verlag, New York, New York, USA.

Tilman, D., and P. Kareiva, editors. 1997. Spatial ecology: the role of space in population dynamics and interspecific interactions. Princeton University Press, Princeton, New Jersey, USA.

Turner, S., R. V. O'Neill, W. Conley, M. Conley, and H. Humphries. 1991. Pattern and scale: Statistics for landscape ecology. Pages 17–47 in S. J. Turner and R. H. Gardner, editors. Quantitative methods in landscape ecology. Springer Verlag, New York, New York, USA.

Tyre, A. J., H. P. Possingham, and C. M. Bull. 1997. Characteristic scales in ecology: fact, fiction or futility. Pages 233–243 in N. Klomp and I. Lunt, editors. Frontiers in ecology. Elsevier Science, New York, New York, USA.

Wiens, J. 1989. Spatial scaling in ecology. Functional Ecology 3:385–397.

Wilson, H. B., and M. J. Keeling. 2000. Spatial scales and low dimensional deterministic dynamics. Pages 209–226 in U. Dieckmann, R. Law, and J. A. J. Metz, editors. The geometry of ecological interactions: simplifying spatial complexity. Cambridge University Press, Cambridge, UK.

APPENDIX A

A figure showing species' trajectories using different sized windows of observation is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A1.

APPENDIX B

A discussion of attractor reconstruction using nonlinear time series analysis is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A2.

APPENDIX C

A discussion of prediction r^2 and error X in the null case of no spatial pattern is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A3.
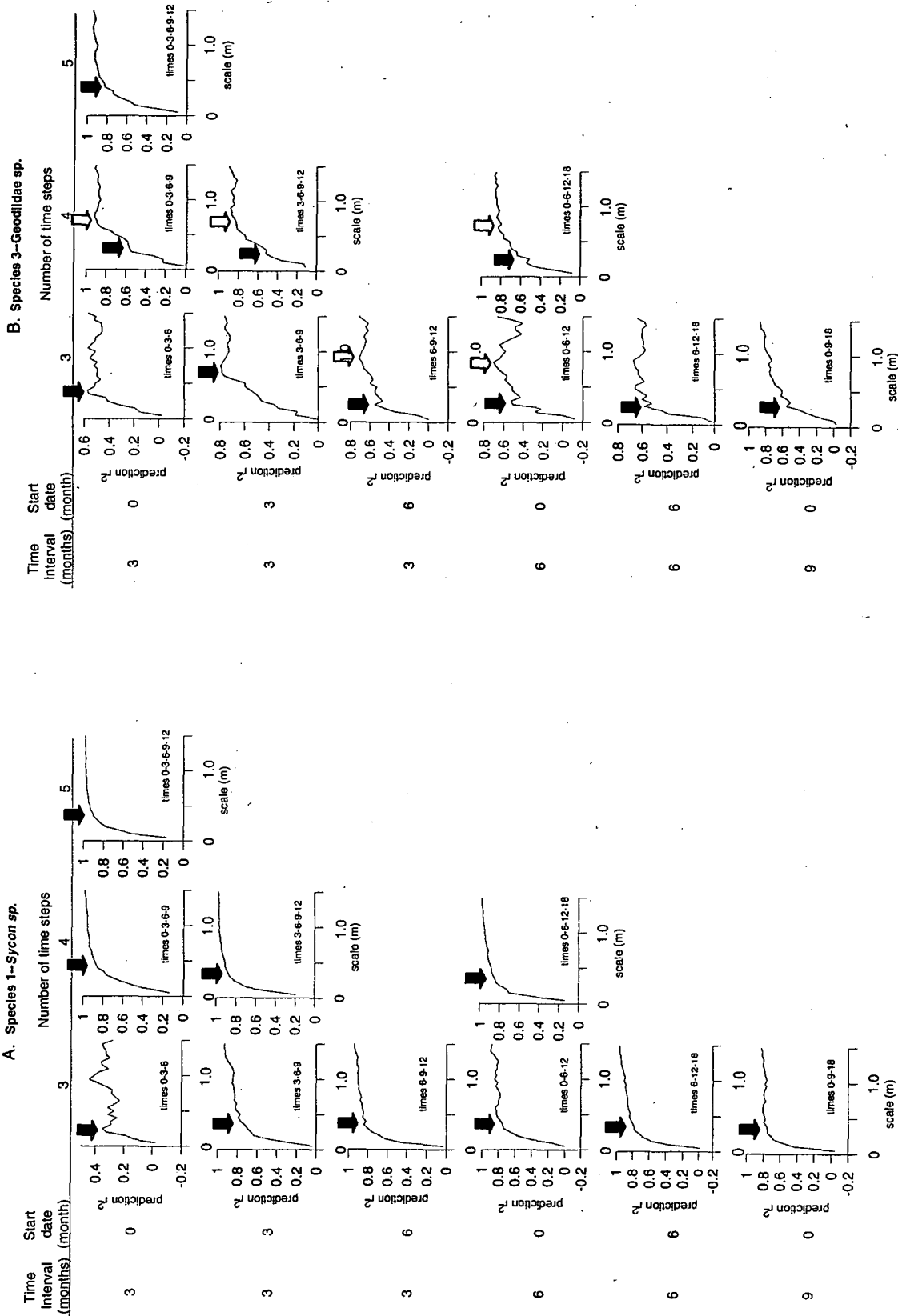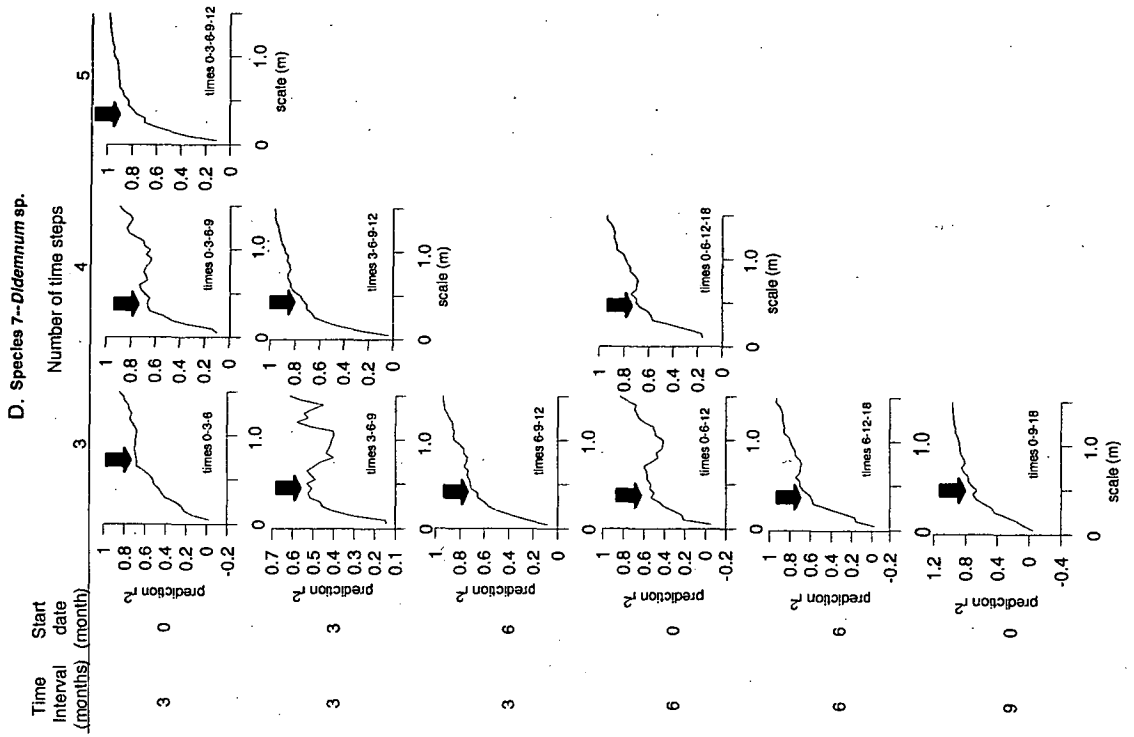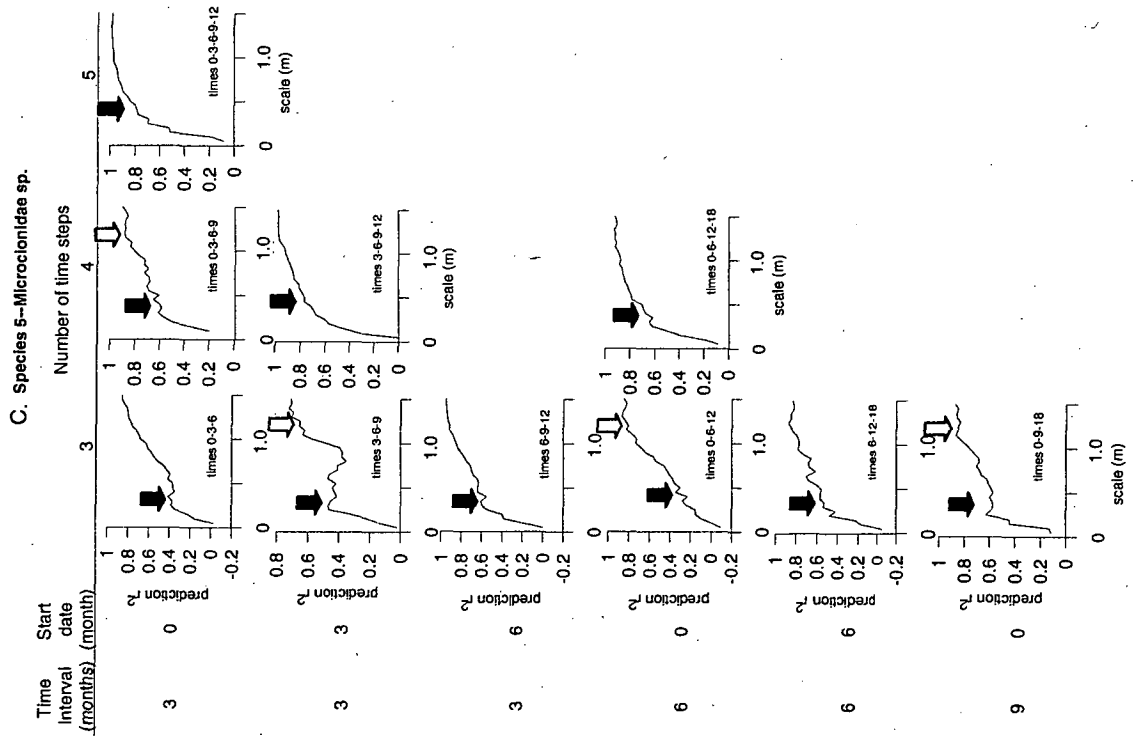
APPENDIX D

A discussion of the evaluation of fixed and proportional delays is available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-A4.

SUPPLEMENT

The COMPETE software (a menu-driven individual-based spatial modeling program) and user manual are available in ESA's Electronic Data Archive: *Ecological Archives* M075-018-S1.

Appendix D. Same as Figure 4.1, for (A) Sycon sp. — species 1, (B) Geodiidae sp. — species 3, (C) Microcionidae sp. — species 5, (D) *Didemnum* sp.— species 7, (E) Geniculate coralline algae – species 9, and (F) Foliose red algae – species 10.

C. Species 5--Microclonidae sp.

D. Species 7--Didemnum sp.

E. Species 9—Geniculate coralline algae

F. Species 10—Foliose red algae
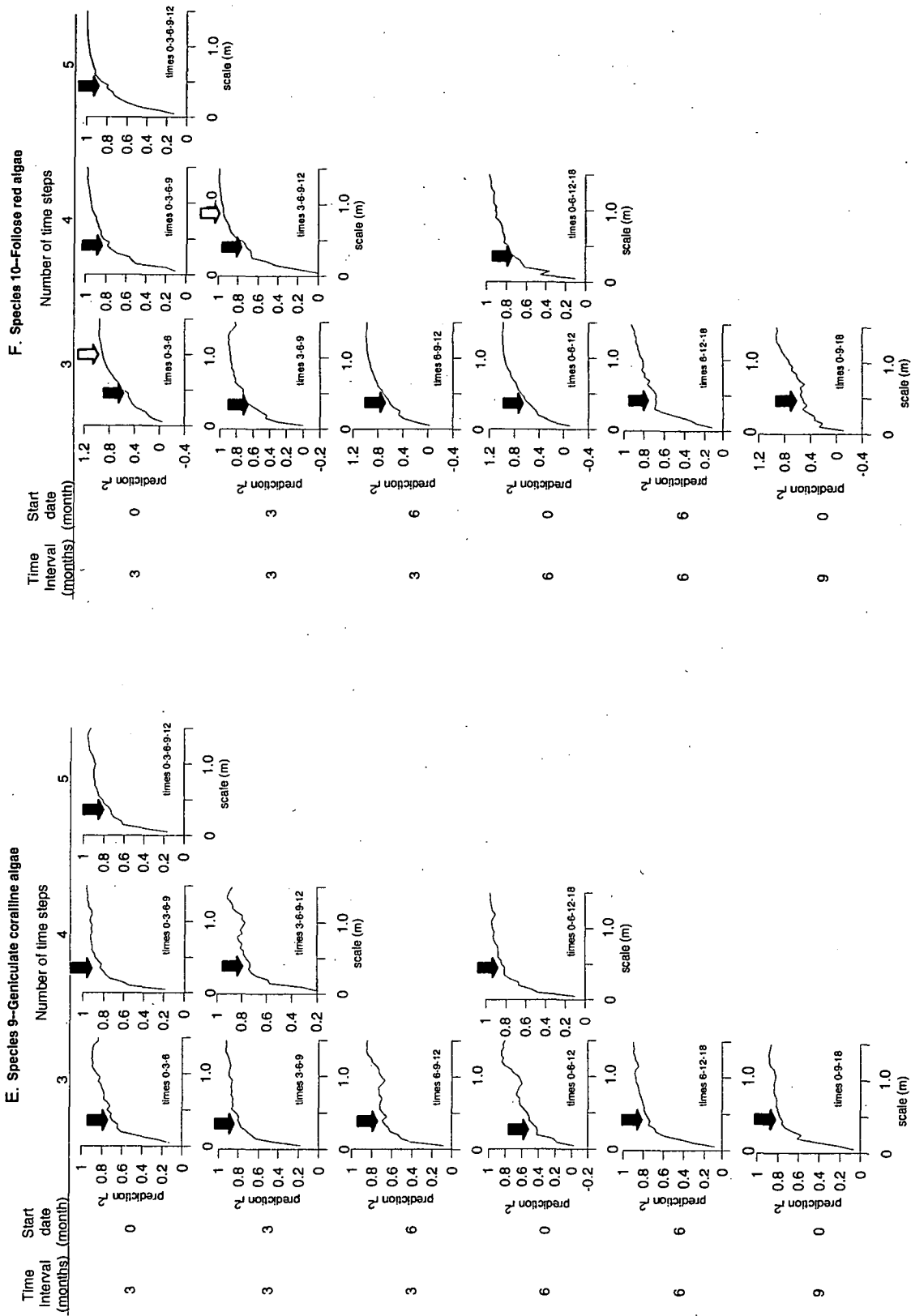
Appendix E

Evaluation of fixed and proportional delays

Landscapes composed of independent, discrete valued pixels provide a null case for evaluating fixed and proportional delays. We can generate data series for windows of side length l using the binomial distribution $N \sim \text{Bin}\,(l^2, \pi)$, which gives independent discrete valued data and so provides a null case comparable with our individual-based spatially explicit models of n-species systems. The value π can be thought of as the probability of observing a particular species, so that $\pi = 0.33$ is the randomized equivalent of a 3-species system. Since, by definition, the determinism to noise ratio does not change with window size l in our null model, we should not expect a typical increasing asymptotic in the prediction r^2 spectrum, or indeed any interpretable prediction r^2 curve. Note, however, that we do not expect a constant slope in the prediction r^2 spectrum for the null model generated using the sliding window approach (Fig D1), unlike the situation for the long time series method (see Appendix B).

It is clear that using a fixed embedding delay in the sliding window analysis is inappropriate because it yields an interpretable prediction r^2 spectrum in the null case for which no CLS exists (Fig D1). The shape of the prediction r^2 curve in our example with $\tau = 10$ suggests a CLS of approximately 60 cells where none can exist (Fig D1, A). Qualitatively similar results were obtained from other integer values of τ ($1 \leq \tau \leq 10$). Thus, using a fixed delay for sliding window attractor reconstruction would confound CLS estimation. By comparison, trends in prediction r^2 curves are reduced in the null case when τ is proportional to window length ($\tau = 0.8 \times$ window length l) and, most importantly, the prediction r^2 spectrum does not indicate a CLS (Fig D1, B). The trend that is evident arises

because the overlap of successive spatial series introduces correlations in k-nearest neighbor predictions (see Appendix B). Thus, proportional delays are preferable to fixed delays for sliding window embeddings because the resultant prediction r^2 spectra do not indicate CLSs in the null case.
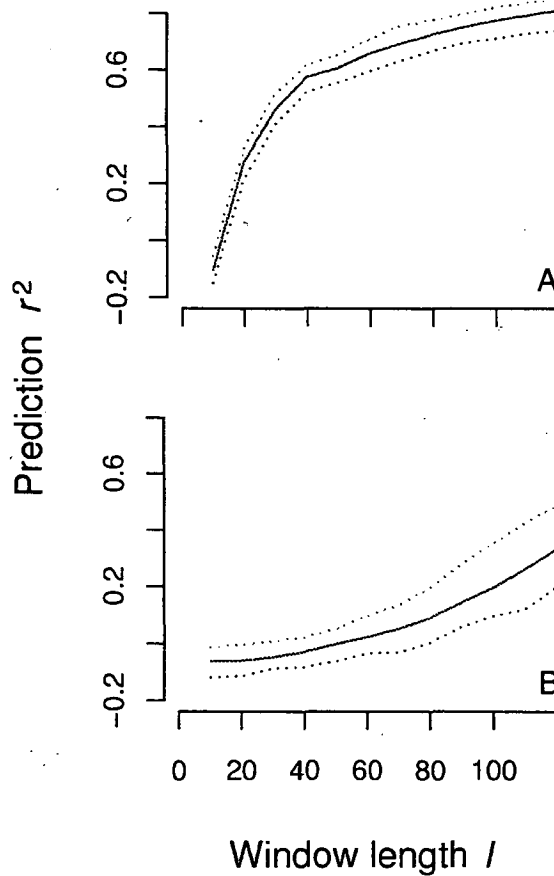

Figure D1. Prediction r^2 as a function of window length l for sliding window analysis using a 3-species null model of independent, discrete valued pixels. In simulated frames, the probability of observing pixels with a particular discrete value is $\pi = 0.33$. In (A) $\tau = 10$ (fixed delay), while in (B) $\tau = 0.8 \times$ window length l (proportional delay). For both analyses $d_E = 5$, $k = 10$. Landscape size is 700×700 pixels. Solid lines indicate average curves for 100 replications and dotted lines indicate 95% confidence intervals. Note that the fixed delay yields an interpretable prediction r^2 curve which is misleading because no CLS exists in the null case. Although the spectrum for the proportional delay is not flat, it does not indicate an interpretable CLS, which is desirable for the null case, and the r^2 values are very low.