Application of information theory methods to food web reconstruction

L.J. Moniz[a,*], E.G. Cooch[c], S.P. Ellner[d], J.D. Nichols[b], J.M. Nichols[e]

[a] US Geological Survey, Patuxent Wildlife Research Center, Laurel, MD, 20708 and Department of Mathematics, Trinity College, Washington, DC 20017, United States
[b] US Geological Survey, Patuxent Wildlife Research Center, Laurel, MD 20708, United States
[c] Department of Natural Resources, Cornell University, Ithaca, NY 14853, United States
[d] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, United States
[e] Naval Research Laboratory, 4555 Overlook Avenue, Washington, DC 20375, United States

ARTICLE INFO

ABSTRACT

In this paper we use information theory techniques on time series of abundances to determine the topology of a food web. At the outset, the food web participants (two consumers, two resources) are known; in addition we know that each consumer prefers one of the resources over the other. However, we do not know which consumer prefers which resource, and if this preference is absolute (i.e., whether or not the consumer will consume the non-preferred resource). Although the consumers and resources are identified at the beginning of the experiment, we also provide evidence that the consumers are not resources for each other, and the resources do not consume each other. We do show that there is significant mutual information between resources; the model is seasonally forced and some shared information between resources is expected. Similarly, because the model is seasonally forced, we expect shared information between consumers as they respond to the forcing of the resources. The model that we consider does include noise, and in an effort to demonstrate that these methods may be of some use in other than model data, we show the efficacy of our methods with decreasing time series size; in this particular case we obtain reasonably clear results with a time series length of 400 points. This approaches ecological time series lengths from real systems.

1. Introduction

In his classical text, "Animal Ecology," Elton (1927) made the following two observations: "Every animal is closely linked with a number of other animals living round it …" "These interrelations between animals appear fearfully complex at first sight". Interrelations in ecological communities frequently involve interactions between species and include such specific processes as predation, competition and mutualism.

Inference about the existence of interactions and estimation of interaction strength are important components of serious investigations of ecological system behavior (Paine, 1992; Wootton, 1997; Laska and Wootton, 1998; Abrams, 2001; Peacor and Werner, 2004). Such inference can proceed by any of three different methodological approaches. One approach involves detailed studies of the mechanisms involved in specific interactions (e.g., predation, Goldwasser and Roughgarden, 1993; Wootton, 1997. Recently, Scotti et al. (2006) have calculated

effective trophic position as an alternative to trophic level. Libralato et al. (2006) have identified keystone species in food webs, but this method is confined to models and not to time series data. Another approach utilizes species removal experiments with direct estimation of responses in abundances or densities of other species in the community (e.g., Paine, 1992; Berlow, 1999). The third approach involves analysis of time series of abundances of the potentially interacting species, even in the absence of experimental perturbations (e.g., Ives et al., 2003).

Here we consider the latter approach and attempt to draw inferences about food web interactions using time series data for species populations in the community of interest. Current methods for drawing inference about coupling and interaction strength in multi-species food web studies using time series data are based on methods that have evolved from Wright's path analysis (Wright, 1921, 1934) and include, as examples, approaches presented by Pfister (1995) and Ives et al. (1999, 2003). However,these methods are all based on approaches developed for linear models, and so typically assume that (on an appropriate scale, e.g., log-transformed) interspecific impacts can be described by linear equations with constant coefficients. The apparent prevalence of nonlinearity in ecological systems and processes (e.g., Ayala et al., 1973; Schaffer and Kot, 1986; Hastings et al., 1993; Constantino et al., 1997; Dennis et al., 1995; Cushing et al., 2003; Turchin, 2003; Boudjema and Chau, 1996) argues for consideration of methods that are applicable to nonlinear systems as well. Abrams (2001, page 209) argues: "because interactions are typically nonlinear, single numerical measures (of interaction strength) are generally poor characterizations".

In this paper, we address the fundamental question—can we derive a qualitative description of the topology of a food web from only time series data, given only a cursory description of the ecological setting (e.g., number of resources and consumers), typical of many empirical studies? In this study, we address this reverse inference question using statistical approaches based on mutual information and transfer entropy. *Mutual information*, first introduced by Shannon and Weaver (1949), later used in the investigation of chaotic dynamics by Fraser and Swinney (1986) and adapted to measure information transfer by Vastano and Swinney (1988), measures the amount of shared information between two time series. *Transfer entropy*, introduced by Schreiber in 2000(Schreiber, 2000), is less easy to interpret. Transfer entropy measures the extra information about dynamics of one system component provided by the knowledge of another component. While mutual information is a static, symmetric measure of shared information, transfer entropy is a dynamical, asymmetric measure of information transfer. We use both measures to reconstruct a food web topology based only on time series data. Ecological time series are generally short, and always noisy. However, some information statistics, based on the entropies of the time series in question, can be used when the time series are short, as long as they are stationary. There are methods to make density estimates relatively reliable with sparse data sets and we employ such methods here. Other approaches are available and might lead to further improvements.

The use of information theory to examine the stability of a food web is not a new idea in ecology. Beginning with MacArthur (1955) who examined stability via the Shannon Entropy, a long list of researchers has tackled the idea of describing stability (see, e.g., Ulanowicz, 2001 and references therein). Rutledge et al. (1976) discusses stability in three real systems using information theory methods and also provides an illuminating discussion of the effectiveness (or not) of only looking at complexity and species diversity to describe stability.

More recently, Ulanowicz (2001) and Ulanowicz and Abarca-Arenas (1997) has described *ascendency*, a measure that uses mutual information scaled by total system throughput. This idea was formalized by Pahl-Wostl (1992) and Latham and Scully (2004) have noted some recent discoveries with regard to its interpretation. One improvement of ascendency over mutual information is that the definition may be augmented to address spatially and temporally extended systems. Ulanowicz (2001) offers a lucid overview and history of information theory in ecology and an honest critique of the strengths and weaknesses of efforts to describe stability using information theory methods. Latham (2006) has contributed open-source algorithms and software to use in network flow analysis. Fath et al. (2003) recently used Fisher information to investigate regime changes in Network Environ analysis as well as direct and indirect effects (Fath and Patten, 1999). Fisher information has also been used by Mayer et al. (2006) to study dynamic regime changes. However, few ecological information theory applications have been concerned with the inverse problem of discovering the food web interactions from time series of abundances.

1.1. Information statistics—an overview

Information theoretical techniques first introduced by Shannon (1948) and Shannon and Weaver (1949) are widely used in nonlinear dynamics to analyze time series. Techniques for estimating both the extent of shared information and the direction of information flow can be useful in investigating food web topology.

Consider a time series as a discretization of a system state variable X. Note that $p(x)$ is the probability of landing in any one of the discrete states x that we observe in the time series. For the entire distribution X, we can compute the amount of information generated by each succesive observation. In the theoretical sense, "information" is defined as the measure of freedom-of-choice (or more accurately, indeterminacy—see, e.g., Ulanowicz, 1986, chapter 5) for system states (Shannon and Weaver, 1949). For example, in a periodic system, once we have completed one period, there is no additional information generated by successive observations. There are no more available choices for system states that differ from the ones already observed; all trajectories repeat. However, in a nonlinear chaotic system, there will be additional information (in the sense that any future generated by new observations that have not yet been observed) no matter how long the system is monitored.

Given the probability density $p(x)$ of landing in any one of the system states then, we quantify the information generated by $p(x)$ by calculating the Shannon entropy H of the system

component:

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x) \tag{1}$$

Looking more closely at the expression for entropy, we see that $\log_2 p(x)$ is the information content of a state $x \in X$. When we sum, we are averaging the information content over all the states. The entropy, then, is the average amount of information contained in X. The entropy is the foundation of the information statistics presented here.

Again considering the expression for entropy, Eq. (1), we see that it depends solely on the probability density function $p(x)$. Suppose we have an estimated or hypothesized density function for our system states, call it $\tilde{p}(x)$ that differs from the true density. Each would generate a (Shannon) entropy. If the entropies differ, and if $p(x)$ represents the better approximation of truth, the difference must depend on the deviation of $\tilde{p}(x)$ from $p(x)$. We can quantify this difference in entropy by introducing a variant of the Shannon entropy, the Kullback entropy. It is calculated as follows:

$$K_{\hat{p}} = \sum_{x \in X} p(x) \log_2 \frac{p(x)}{\tilde{p}(x)} \tag{2}$$

The Kullback entropy is a measure of the "error" that is obtained by using $\tilde{p}(x)$ in lieu of $p(x)$, or put another way, as the information gained per observation by using $p(x)$ in favor of $\tilde{p}(x)$ (Kullback, 1997). The reason we define the Kullback entropy as a separate entity from the Shannon entropy is that the Kullback entropy allows us to focus on a particular probability density $p(x)$ and then test the scenario that $\tilde{p}(x)$ describes the system as well as $p(x)$. Larger Kullback entropy reflects poorer performance of $\tilde{p}(x)$ as an approximation to $p(x)$.

With the Kullback entropy in hand, we will describe the following two statistics that we will use to draw inferences about the nature and degree of coupling and dynamical effects between species in an unknown food web. All of the statistics that we use will be built in some way from the Kullback entropy (Eq. (2)) and/or the Shannon entropy (Eq. (1)).

1.2. Mutual information statistic

If we have two processes, we often want to measure the amount of shared information between them, in other words, their deviation from independence. This would provide inferences about the coupling between the two processes. If we start with X and Y, with probability distributions $p(x)$ and $p(y)$, we can compute the degree to which $p(x, y)$ deviates from independence. This is

$$\log_2 \frac{1}{p(y)p(x)} - \log_2 \frac{1}{p(x, y)} \tag{3}$$

If X and Y are independent, note that we get zero. If we average over all points, we can compute the *mutual information* between X and Y as a Kullback entropy:

$$MI(X, Y) = \sum_X \sum_Y p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \tag{4}$$

In a different form, we can represent the mutual information as a sum of Shannon entropies:

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) \tag{5}$$

where H is the entropy from Eq. (1). The mutual information $MI(X, Y)$ is therefore the average loss of indeterminancy that results from assuming that X and Y are independent, so $MI(X, Y) = 0$ if X and Y actually are independent. This measure is symmetric in X and Y: $MI(X, Y) = MI(Y, X)$. The mutual information then measures the amount of information gained about one of the system components through the measurement of the other. If the two processes are independent, no information about the first system component can be extracted from the second, and vice versa. If the food web were seasonally forced, i.e., if one or more species were seasonally forced, we would expect to see positive mutual information between all the species: there is some shared information from the forcing. However, we expect to see less mutual information between consumers than between, e.g., a resource and a consumer of that resource; shared information between consumers could be a direct or indirect result of the forcing or an indirect result of their both consuming the same resource or it could be an indirect effect of consuming different resources if the resources competed with each other for some other unseen resource. In a system that is not externally forced, positive mutual information between *any* two system components could indicate that all the components interact.

Although mutual information will show that two system components are coupled, it gives no indication of the dynamical *effect* of one system component on the other. One way to glean some knowledge of about dynamical effects is to focus on information transfer by modifying the mutual information to compare coupling at different times in the time series. The *time-delayed mutual information*, introduced by Vastano and Swinney (1988), measures the information gained about one system component by measuring another at a past or future time. Although time-delayed mutual information does not take the dynamics of the two time series into account explicitly, the change in time-delayed mutual information over a given time window can indicate that one system affects another with a particular response time. For example, in a food web, we might expect to see a flat time-delayed mutual information for coupling induced by an outside forcing, while we would expect to see a well-defined maximum in the mutual information for true transfer of dynamics—for example a consumer affecting a resource because of predation (Nichols, 2005).

The time-delayed mutual information (TDMI) can be defined as follows. To find out how much information about a process $X(t)$ is gained by observing a future or past value of another process $Y(t)$, we compute

$$TDMI(X, Y + \tau) = H(X) + H(Y + \tau) - H(X, Y + \tau) \tag{6}$$

This statistic can be computed for both positive and negative values of τ. If the maximum of $TDMI(X, Y + \tau)$ as a function of τ occurs for $\tau > 0$ this indicates a flow of information from X to Y; X "now" tells us about Y "later". Similarly a maximum for $\tau < 0$ indicates information flow from Y to X. Although this

statistic is not symmetric, it is reflected through zero, that is $\text{TDMI}(X, Y + \tau) = \text{TDMI}(X - \tau, Y)$.

1.3. Transfer entropy statistic

This statistic, introduced by Schreiber in 2000 (Schreiber, 2000), measures the amount of dynamical information flow between two processes. We can, at least in the discrete case, compute transition probabilities for a trajectory of a process X. This is of course related to the entropy. Now, we consider both the transition probability of going from some state x_0 to a state in the next time step, x_1, and ask whether there is any additional information about that probability provided by another process Y, that is, $p(x_1|x_0, y)$. Considering this, we can see if the process Y actually affects the dynamics of X; does knowing y give us any information about the transition probabilities of x? If it does, then we can quantify this additional information with the transfer entropy.

To define the transfer entropy, we start by looking at the transition probability from one state to another: $p(x_{i+1}|x_i)$. We can form a conditional entropy from this (see, e.g., Kaiser and Schreiber, 2002) and then consider the hypothesis that the process Y affects the transition probabilities in some way. From this we form another Kullback entropy to see if there is any additional gain in information from considering Y in the transition probability:

$$\text{TE}(X_{i+1}|X_i, Y_i) = \sum_{x_{i+1}, x_i, x_i} p(x_{i+1}, x_i, y_i) \log_2 \left[\frac{p(x_{i+1}|x_i, y_i)}{p(x_{i+1}|x_i)} \right] \quad (7)$$

where p is the density. Thus the transfer entropy measures the change in entropy rate of process X when affected by process Y. Note that the denominator of Eq. (7) depicts a first-order Markov process in which the transition probability from time i to time $i+1$ depends only on the system state at time i, whereas a process of order > 1 can be considered as well. The numerator of Eq. (7) focuses on the influence of y_i; this state variable can also be considered at other times (Schreiber, 2000).

We note most importantly that transfer entropy is not symmetric; entropy can be transferred in one direction and not necessarily in another. Because the transfer entropy yields information about the direction of dynamical information transfer, it is a powerful tool for analysis of coupling relationships. We will discuss the interpretation of transfer entropy in Section 1.4. Finally, because a shorter time series contains less information than a long time series. the rate of information transfer is expected to go down as the time series length is shortened.

1.4. Interpretation of information statistics

The mutual information statistic, in both static and time-delayed forms, is relatively easy to interpret. However, there are some characteristics of the transfer entropy, described below, which can cause confusion in its analysis. Marschinski and Kantz (2002) noted these difficulties and offered an alternative formulation of transfer entropy. Here, we instead explore the characteristics of transfer entropy in a very simple

case in order to provide interpretation guidelines of transfer entropy in its original form.

We stress that transfer entropy is ambiguous in that if two processes are completely independent, transfer entropy can be zero and if two processes are identical, transfer entropy can also be zero. We also note that transfer entropy can be non-zero between two processes if they are driven by a third, unseen system. We offer the following "toy" example, two discrete, stochastic systems driven by a third system, to illustrate this concept.

Let F be a deterministic, discrete driving system that cycles between states 0 and 1 with probability 1 (e.g., a sequence for F with initial condition 0 is 0, 1, 0, 1,...). The only non-deterministic quantity for F, then is the choice of initial condition (0 or 1). We then consider two systems, G and H which are stochastic and depend on the state of F (F drives both G and H). The rule for update of G is: $g_n = (F_n + 1) \bmod 3$ with probability $(1 - c)/2$ and $g_n = (F_n + 2) \bmod 3$ with probability $(1 + c)/2$. This is somewhat similar to the discrete model described in Kaiser and Schreiber (2002). We then let H be a similar, but not necessarily identical, system to G. Then $h_n = (F_n + 2) \bmod 3$ with probability $(1 - d)/2$ and $h_n = (F_n + 3) \bmod 3$ with probability $(1 + d)/2$. We can compute the transfer entropies between H and G directly:

$$\text{TE}_{G \to H} = \frac{1}{16}[(28 - 4d^2) \log 2 - (8 + 8d + (1 - d^2)(1 - c)) \log(1 + d)$$
$$- (8 - 8d + (1 - d^2)(1 + c)) \log(1 - d)]$$

and

$$\text{TE}_{H \to G} = \frac{1}{16}[(28 - 4c^2) \log 2 - (8 + 8c + (1 - c^2)(1 - d)) \log(1 + c)$$
$$- (8 - 8c + (1 - c^2)(1 + d)) \log(1 - c)]$$

Thus we see that even if c and d are equal, the transfer entropies are non zero. If we assign quite different values to c and d, e.g., $c = 0.2$ and $d = 0.8$, we get $\text{TE}_{G \to H} \approx 0.84207$ and $\text{TE}_{H \to G} \approx 1.5853$. Taking c and d to be close to each other for example $c = 0.4$ and $d = 0.5$ we find $\text{TE}_{H \to G} \approx 1.2652$ and $\text{TE}_{G \to H} \approx 1.3829$.

The mutual information for this model can also be calculated directly:

$$\text{MI}(G, H) = \frac{1}{8}((1 + c)(1 - d)[\log(1 - d) + \log(1 + c)]$$
$$+ (1 - c)(1 + d)[\log(1 + d) + \log(1 - c)]$$
$$+ (2 + 2cd) \log 2 + (4 + 4cd) \log(1 + cd))$$

For the first example $c = 0.2$, $d = 0.8$ we find $\text{MI}(G, H) \approx 0.4471$ and for the second example, $c = 0.4$, $d = 0.5$ we find $\text{MI}(G, H) \approx .0.395695$.

We see that in the first example, although the transfer entropy from G to H is relatively high, the asymmetry is also large. In this case, the dynamics are dissimilar. However, each provides information about the driver that the other does not. We see in the second example that transfer entropy is more symmetric. In this case, the dynamics of both G and H are similar. There is some transfer entropy; one is generating information that the other does not. In both cases mutual

information was positive and did not differ greatly. This shows that there is shared information. However, we see that the rate of information transfer in both examples is greater than the shared information.

In this very elementary example, we see that the driven systems can exhibit positive transfer entropy with no direct information transfer. Thus it is likely that in a driven system with actual information exchange, the transfer entropy from the driver can either mask or enhance effects of the two system variables on each other. This characteristic of transfer entropy, however, rather than being a problem, can be used to glean some interpretation guidelines from this simple system. In computer science terms, this is a "feature" and not a "bug". We see primarily that *disparate transfer entropies signal that the dynamics are not necessarily shared* and we may be seeing transfer of indirect information about the forcing from the other system. Thus the transfer entropy can give us information about direct and indirect coupling. This information is not accessible with the mutual information statistic alone.

This demonstrates that the results of transfer entropy need to be interpreted *in toto*. transfer entropy is a one-sided statistic; it is possible to have entropy transferred in one direction and not another. However, as we see from our food web model, the more likely scenario for real systems and real data is that some entropy will be transferred in a driven or interacting system because of indirect effects. As we saw in this case, the asymmetry can be exploited in the interpretation of transfer entropy in order to discern stronger or weaker relationships between one time series and another. More asymmetric transfer entropies signal less coupling of dynamics, and less asymmetric transfer entropies signal more coupling of dynamics.

2. The model

One subset of the authors – the "modeling team" – generated artificial time series data from a food web model which included two resource (prey) and two consumer (predator) species. These data were provided to another subset of the authors—the "analysis team". The analysis team knew the numbers of resource and consumer species and which time series represented resource and consumer species, but did not know the form of the underlying model equations, or the values of model parameters. After the food web structure was determined by the analysis team, as described in Section 5, below, the modeling team revealed the underlying equations and parameters.

2.1. *True structure of the underlying food web model*

The food web used to generate the time-series data consisted of two resource species and two consumer species. The resource species, η_1 and η_2, exhibit density-dependent growth and are environmentally forced. The consumer species, η_3 and η_4, have growth dependence only on the resource species within the framework of the community matrix $\mathbf{A}$, but mortality based on a constant factor m, where $m = 0.1$ in this simulation. The coupled equations for the model, based on generalized Lotka–Volterra expressions (assuming a Type I

predator–prey interaction), were:

$$\frac{\partial \eta_1}{dt} = r_1 z_t \eta_1 (1 - 0.1\eta_1) - \alpha_{1,3}\eta_3\eta_1 - \alpha_{1,4}\eta_4\eta_1$$

$$\frac{\partial \eta_2}{dt} = r_2 z_t \eta_2 (1 - 0.1\eta_2) - \alpha_{2,3}\eta_3\eta_2 - \alpha_{2,4}\eta_4\eta_2$$

$$\frac{\partial \eta_3}{dt} = \alpha_{3,1}\eta_3\eta_1 + \alpha_{3,2}\eta_3\eta_2 - m\eta_3$$

$$\frac{\partial \eta_4}{dt} = \alpha_{4,1}\eta_4\eta_1 + \alpha_{4,2}\eta_4\eta_2 - m\eta_4$$

where (i) $\alpha_{i,j}$ is the direct effect of species j on species i, (ii) m_i is the intrinsic mortality rate of the predator/consumer species in the absence of food, and (iii) z_t is an exogenous environmental forcing variable with both seasonal and random components,

$$z_t = 1 + 0.25 \cos\left(\frac{2\pi t}{52}\right) + 0.1Q(t)$$

Here t is time in weeks (i.e., the model uses a year consisting of 52 7-day weeks) and the random component $Q(t)$ was computed by generating $Q(0), Q(52), Q(104), \ldots$ as independent Gaussian random numbers with mean = 0, variance = 1, and interpolating these values by a cubic spline. Each run of the model used a different randomly generated $Q(t)$ function.

Because this was a blind study (i.e., the model structure was unknown by the analysis team), we used a resource/consumer labeling system to facilitate topology identification. In the analysis, η_1 is labeled R1 (or resource1), η_2 is labeled R2 (resource 2), η_3 is labeled C1 (consumer1), and η_4 is labeled C2 (consumer2). For this study, data were generated using the following parameter values: $m = 0.1$, $r_1 = 1.25$, $r_2 = 1.5$, $\alpha_{3,1} = 1$, $\alpha_{4,1} = 0.3$, $\alpha_{3,2} = 0.3$, and $\alpha_{4,2} = 1$. Thus, in the underlying model used to generate the time-series data there was (i) no *direct* competition between species for a given trophic level (i.e., the resource species neither directly interact, nor do the two consumer species interact), and (ii) consumer species 1 interacted most strongly with resource species 1, whereas consumer species 2 interacted most strongly with resource species 2—with symmetry in the strength of the interactions (i.e., effect of consumer species i on resource species j is equivalent to effect of resource species i on consumer species j).

It is important to note that the results that we present were obtained without knowledge of the model, model parameters, or community matrix. Although we hoped to be able to discover the possibility of preference of one resource over another for each consumer, we did not assume this information at the outset.

3. Time series: the "Data"

A plot of the first 1000 points of the time series, with transients removed, appears in Fig. 1. From inspection of the time series we see that is is quite likely the model is seasonally forced; this was verified by the modeling team before the analysis was begun. It was also revealed by the modeling team that there is some noise in the model. However, no other information about the model was known before the analysis.

All of the statistics mentioned here require the computation of an estimated density from a time series. These time
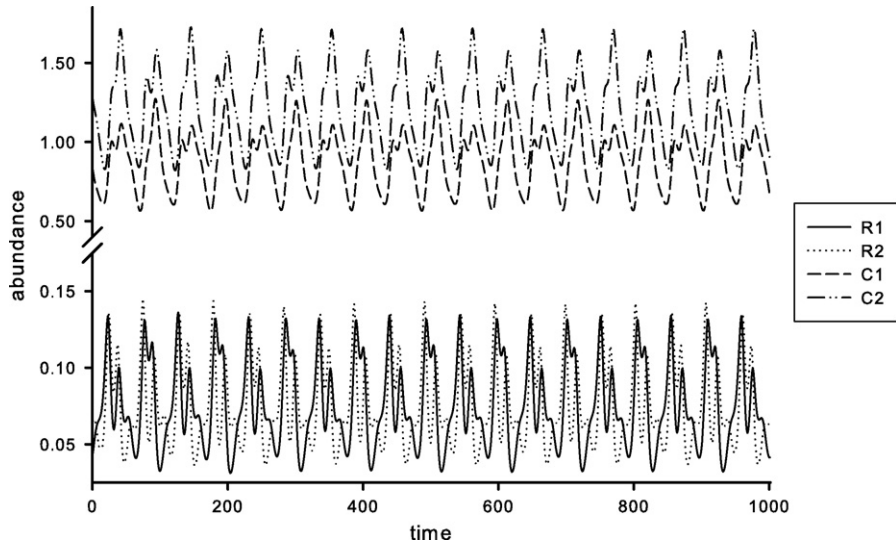
Fig. 1 – First 1000 points of the consumer–resource time series (without transients). C1, consumer1; C2, consumer2; R1, resource1; R2, resource2. N refers to the number of near neighbors used in the density estimation ($\hat{p}(x)$) used to calculate mutual information and transfer entropy. All information measures are in bits.

series are discretizations of continuous processes and in particular these processes are nonlinear. Thus, the usual method of binning or aggregating a discrete distribution in order to find the density is unlikely to provide accurate estimates of density. One reason for this is that the points are not uniformly and evenly distributed in state space; some regions are likely to be sparser than others. The usual method of using a Heaviside Step Kernel (Liebert and Shuster, 1989; Prichard and Theiler, 1995) may not work for very sparse densities—which we hope to be able to investigate in order to investigate the viability of the method for ecologically realistic time series lengths. The Heaviside Step Kernel estimates density at a point x_i by finding all time series points x_j within a radius R of x_i, then counts the number of such points x_j. However, if the data are not *uniformly* sparse, the fixed radius R (spherical binning) can yield a poor estimate. Thus we will, for these experiments, use a fixed-mass approach (see, e.g., Kaiser and Schreiber, 2002) along with a naive kernel density estimator (Silverman, 1986)(see also Marschinski and Kantz, 2002 for another estimate of kernel density for transfer entropy with a relatively short time series).

We begin, for our density estimate, with a kd-tree (see Bentley, 1979), implemented with a fast near neighbor search (see Hjaltason and Samet, 1995). The kd-tree is a computational tool for creating an easy-to-search multidimensional adaptive data structure. This structure is especially suited to the highly non-uniform distribution that is commonly found in time series from nonlinear systems.

To compute the density estimate for our M-point time series **X**, for each representative point x_0 (which, as we shall see, is not necessarily one-dimensional), we use the near-neighbor search to determine the N (we use $N = 5$ and $N = 10$) nearest neighbors to the representative point. We then compute the largest side length of the rectangle, centered at x_0, that exactly encompasses the nearest neighbors. Using the naive kernel estimator (correction factor) 1/2, we compute the

estimate for the density $\hat{p}$ (**X**), using the procedure outlined in Kaiser and Schreiber (2002).

To insure that the effects that we saw for short time series lengths were not an anomaly or peculiar to the particular time series fragment we used, we also front-truncated the original time series both by 1000 points and by 2000 points, chose the remaining shorter time series fragments from these truncated time series and computed the statistics in the same fashion. We discuss these results along with the original time series results, in Section 5.1.

4. Parametric analysis

To motivate our analyses using information statistics, a team member who generated the data and therefore had knowledge of the structure of the generating model performed a conventional parametric analysis. The details of the analysis appear in the appendix. We will summarize the results here.

The analysis yielded perfect results (coefficients for the consumers are estimated nearly perfectly)) for the consumer because the fitted statistical model for their dynamics is exactly right. Those for the resources (species 1 and 2) are not: the consumer impact coefficients are all underestimated. For the resource species the fitted model is not quite the truth, because the data-generating model for resource species i actually has the time-dependent coefficient z_t multiplying $(1 - \eta_i)$. That is, the fitted model (in the appendix, see (A.2)) has time dependence only in the intercept of the linear regression, while the data-generating model also has time-dependence in one of the coefficients. This error in model specification leads to biased and imprecise estimates of interaction parameters, even for interactions that are specified correctly in the fitted model.

Moreover, the parametric analysis is so powerful that it provides clear evidence for some things that are not actually true.

For each artificial data series we also fitted the model

$$R_i = s_i^{(1)}(t \bmod 52) + s_i^{(2)}(t) + s_{i1}(\eta_1) + s_{i2}(\eta_2) + s_{i3}(\eta_3) + s_{i4}(\eta_4) \quad (8)$$

where s_{ij} are the fitted spline functions representing the effect of species j on the per-capita growth rate of species i. This means that the species interactions were not assumed to be linear, even though they really are. In (8) the fitted model "tries" to capture the missing interaction term by putting non-linearities somewhere else, and it succeeds because there are consistent patterns of covariation among state variables. Model (8) really does predict better than (A.2)– given the structure of the data – but it does so for the wrong reason.

As this example illustrates, the potential power of a parametric analysis is only realized when the fitted parametric model is exactly right. A subtle error in the fitted model – which could easily occur when working with real data – markedly degrades the accuracy of the numerical estimates of food web parameters, and produces incorrect conclusions (at very high statistical confidence levels) about the qualitative nature of the dynamics. By using information statistics, we are able to dispense with all assumptions about the form of the data-generating process. The unavoidable cost is loss of precision relative to the ideal situation of fitting a correct parametric model. The potential gain is robustness against our incomplete knowledge of the data-generating mechanisms.

5.   Results—verification of resources and consumers

With the limitations of parametric methods in mind, we turn to the analysis using information theory methods. Although the analysis in the previous section assumed knowledge of the model, we stress that the following was undertaken with only the time series data of abundances.

We consider Fig. 2, the plots of mutual information for all combinations of resources and consumers for the food web model. Recall that mutual information is symmetric; e.g., MI(R1,C1)=MI(C1,R1); only one direction is given in the plot key. We note first of all that mutual information between consumers is the lowest of all the time series pairs. This is expected; the two consumer's dynamics are not identical, but because of seasonal forcing the resources on which they prey are coupled. Thus, the consumers are indirectly linked through the resources and show some information overlap. As mentioned, the mutual information between resources is relatively high. We see also that mutual information for some resource/consumer pairs is larger than for others. As expected, mutual information decreases as time series information declines (reduction in time series length).

We turn to the plot of transfer entropy, Fig. 3. This plot shows very low transfer of information between consumers; this was something of a sanity test and also provides evidence that this is not a tritrophic food web; the consumers are not resources for each other. This is bolstered by the even stronger result for mutual information: the consumer–consumer mutual information is lower than that for the other relationships. We note however, that consumer-to-consumer transfer entropy is not zero; this reflects the indirect transfer of entropy through the resources. We also see that as in other predator/prey systems, the predator–prey transfer entropy (colloquially, the amount of dynamical information about the prey that the predator exhibits) is larger than the prey–predator transfer entropy. A resource/consumer system can be thought of, very loosely, as a drive/response system where the resource can be thought of as the drive and the consumer, the response. Although in a food web the consumer (response) does affect the resource (driver), the resource can exist without the consumer, but the consumer cannot exist without the resource. Hence, the resource corresponds somewhat to the drive. Since the response contains the information in the drive, we expect that the response will carry more information about the drive than vice versa. Thus, we also have a sanity check that the resources and consumers are correctly identified. We note, however, the analogy to the
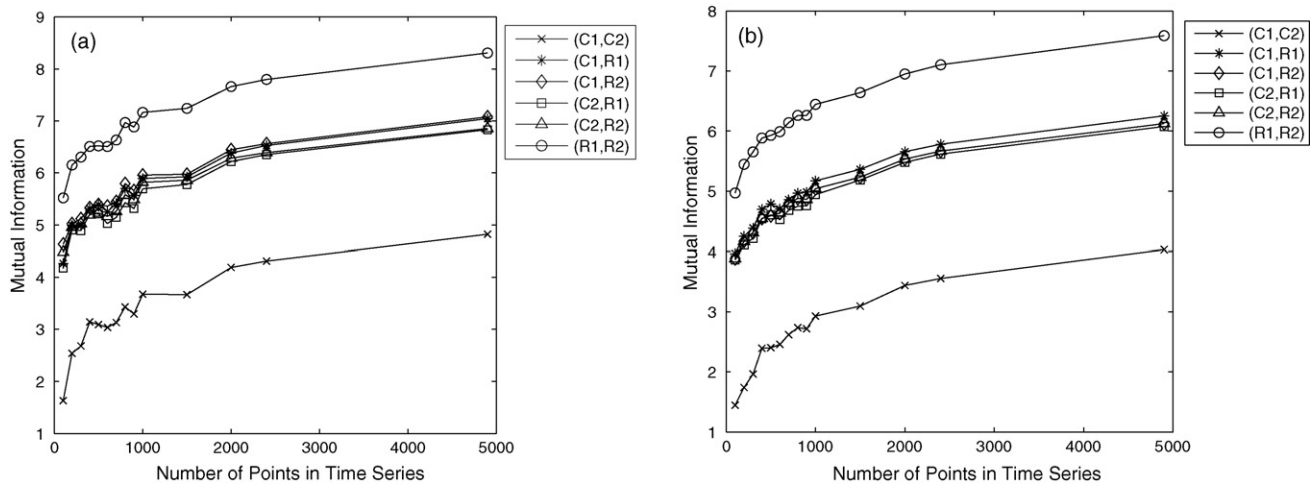

Fig. 2 – Mutual information – food web model – between consumers and resources for (a) N = 5 and (b) N = 10 near neighbors used in the density estimation. C1, consumer1; C2, consumer2; R1, resource1; R2, resource2. N refers to the number of near neighbors used in the density estimation ($\hat{p}(x)$) used to calculate mutual information and transfer entropy. All information measures are in bits.
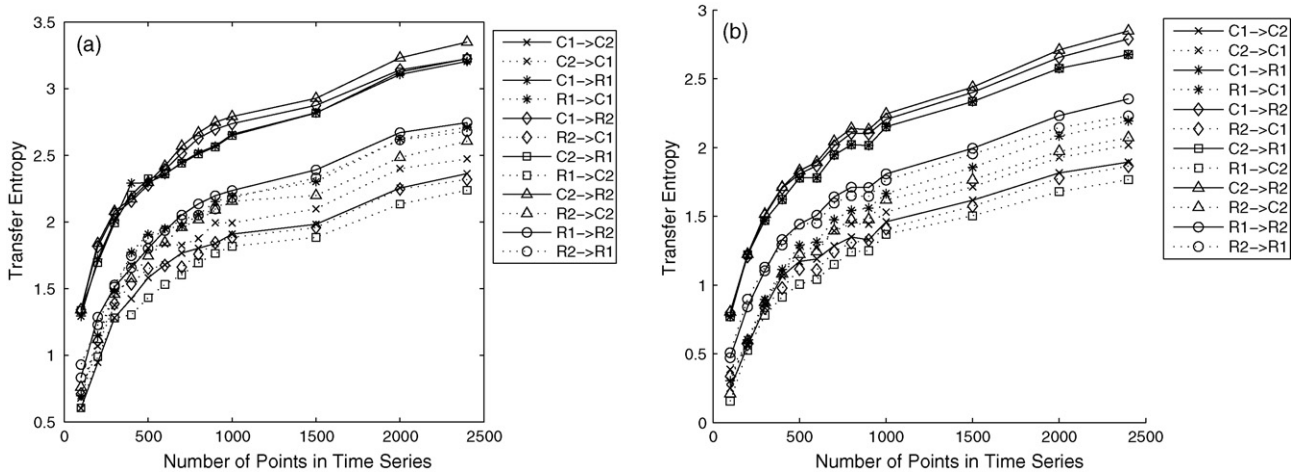
Fig. 3 – Transfer entropy – food web model – from resources to consumers and consumers to resources for (a) N = 5 or (b) N = 10 near neighbors in the density estimation. C1, consumer1; C2, consumer2; R1, resource1; R2, resource2. N refers to the number of near neighbors used in the density estimation ($\hat{p}(x)$) used to calculate mutual information and transfer entropy. All information measures are in bits.

drive–response system is imperfect. The consumer does affect the resource, and more so if a particular resource is the main prey.

We also consider the fact that the resources may be in competition with each other (see, e.g., Ulanowicz and Puccia, 1990). However, the transfer entropy between the two resource species is nearly the same in both directions. This leads us to believe that neither has much of a negative impact on the other and, as with the consumers, do not consume each other.

5.1.   Identification of arrows in the food web topology

The existence of significant information transfer and information sharing between all combinations of resources and consumers indicate that while there may be a preferred resource for the consumers, the consumers will take

advantage of both resources. Thus there is very little difference in the transfer entropies between, for example, resource1/consumer1 and resource2/consumer1. This particular statistic, then, transfer entropy, allows us to "draw the arrows" in the food web, i.e., both consumers prey on both resources, but the small differences in transfer entropy do not allow us to determine the preferences, if any.

We do find some weak evidence for drawing preferences in the food web, from the time-delayed mutual information (Fig. 4). In these plots, we see that the time-delayed mutual information statistic for the consumer1/resource1 and consumer2/resource2 pairs have maxima near zero. The other pairs (consumer1/resource2 and consumer2/resource1) show much flatter overall plots than the others. Because we expect constant flow of information for relationships based on coupling from forcing, the flatter plots are more indicative of
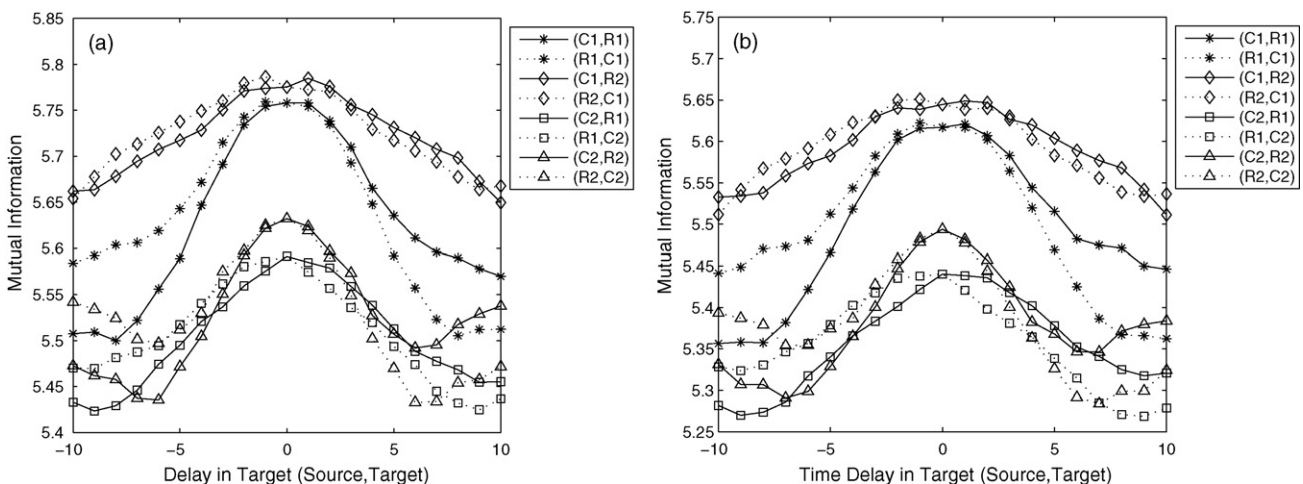

Fig. 4 – Time-delayed mutual information – food web model – between consumers and resources, for time series length (a) 2400 points and (b) 2000 points. There were N = 10 near neighbors used in the density estimation. C1, consumer1; C2, consumer2; R1, resource1; R2, resource2. N refers to the number of near neighbors used in the density estimation ($\hat{p}(x)$) used to calculate mutual information and transfer entropy. All information measures are in bits.
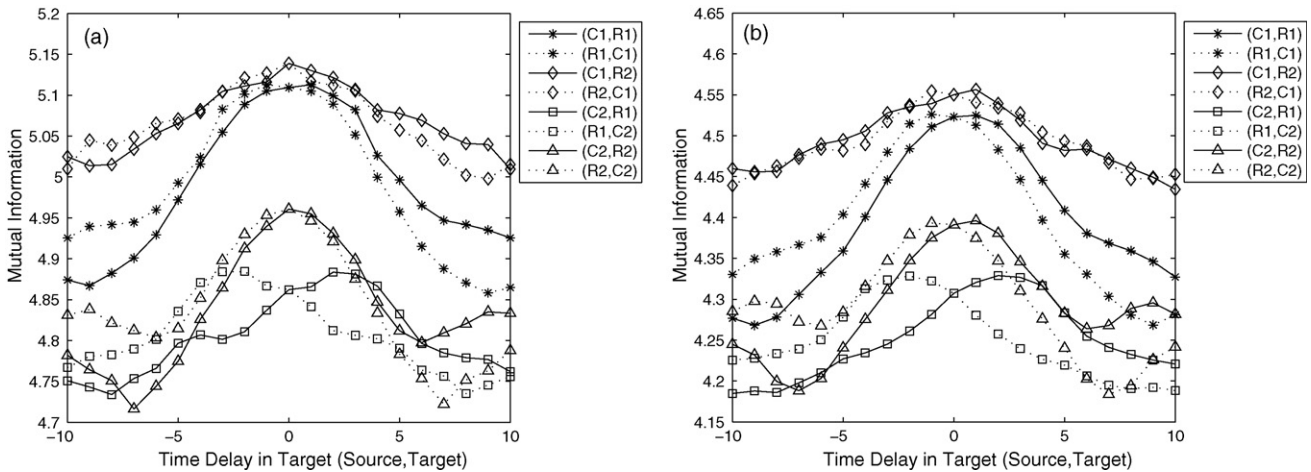
Fig. 5 – Time-delayed mutual information – food web model – for time series lengths (a) 1000 and (b) 500. There were N = 10 nearest neighbors were used in the density estimation. C1, consumer1; C2, consumer2; R1, resource1; R2, resource2. N refers to the number of near neighbors used in the density estimation ($\hat{p}(x)$) used to calculate mutual information and transfer entropy. All information measures are in bits.

coupling based on forcing rather than direct coupling. However, we see that this relationship is not as persistent as some of the other statistics; it disappears when the time series have fewer than about 1000 points (Fig. 5). Thus, the time-delayed mutual information provides only weak evidence that there is a preference for resources by particular consumers.

Now we turn to measurement of asymmetry in the transfer entropy to identify preferences for either resource by a consumer. The asymmetry in transfer entropy is computed for arbitary X and Y by taking:

$$|TE_{X \to Y} - TE_{Y \to X}| \tag{9}$$

Recall from the results in Section 1.4 that if two systems are strongly and directly coupled, we expect that mutual information will be relatively high and there will be significant transfer entropy as well. We also expect that if the coupling is direct, the transfer entropy between the two systems should be similar, but not identical. If the coupling is either indirect or weak, we expect more significant asymmetry between transfer entropies in the two directions. The forcing yields positive transfer entropy, but dissimilar dynamics (weaker coupling) will exhibit asymmetry. In our resource/consumer system we expect that if the consumer does not eat as much of the resource, the resource does not affect the consumer's population as highly. Conversely, if the resource is the preferred prey, then the growth of the consumer population is more highly dependent on the resource. However, this inference about the relationship must be tempered by the mutual information and the values of the transfer entropies themselves. If the transfer entropy and mutual information are low (as one would expect between resources or between consumers), the asymmetry carries less meaning.

We refer to the plot of asymmetry in transfer entropy to assign the preferences (see Fig. 6). Asymmetry between resource1/consumer2 and resource2/consumer1 is substantially higher than the other relationships. This higher asymmetry persists for time series to around 400 points. For time series shorter than 300 points (N = 10) or 400 points

(N = 5), the relationship is not discernable from the plot. The asymmetry in transfer entropy was lowest between consumer/consumer and resource/resource pairs, indicating little transfer of entropy. However, mutual information was also lowest for these pairs, providing little indication of shared dynamics between resources or between consumers. This agrees with the other results indicating that the food web only has one trophic level. The TE asymmetries between resource1/consumer1 and resource2/consumer2, however, suggest shared dynamics. The transfer entropy was relatively high and showed more symmetry than the other resource/consumer pairs. This, coupled with relatively high MI, indicates that these pairs influenced each other's dynamics more closely. Thus, we conclude that consumer1 prefers resource1 and consumer2 prefers resource2. However, because there is no direct formula that links transfer entropy to coupling strength, we were not able to assign a meaningful ratio to the preferences of the consumer/resource pairs. In the absence of knowledge of the model parameters, then, we were simply able to indicate the preferences and not the strength of the preferences with these statistics.

In order to be certain that these results were not an accident of the particular portion of the time series that we used, we extracted fragments of the time series from other portions of the entire time series and computed the statistics. The results from these alternate fragments were nearly identical to those of the original time series fragments. The asymmetry in transfer entropy for several such fragments, taken from the original time series with the first 1000 points truncated, appears in Fig. 7 for comparison.

We are now in a position to identify the food web topology. From the evidence presented above, we conclude that consumer1 prefers resource1 but will also consume resource2. consumer2 prefers resource2 but will also consume resource1. The consumers do not eat each other and the resources do not eat each other. Thus, the topology is represented by Fig. 8. After the analysis was complete, the modeling team revealed the actual food web topology. This is given in Fig. 9. The actual
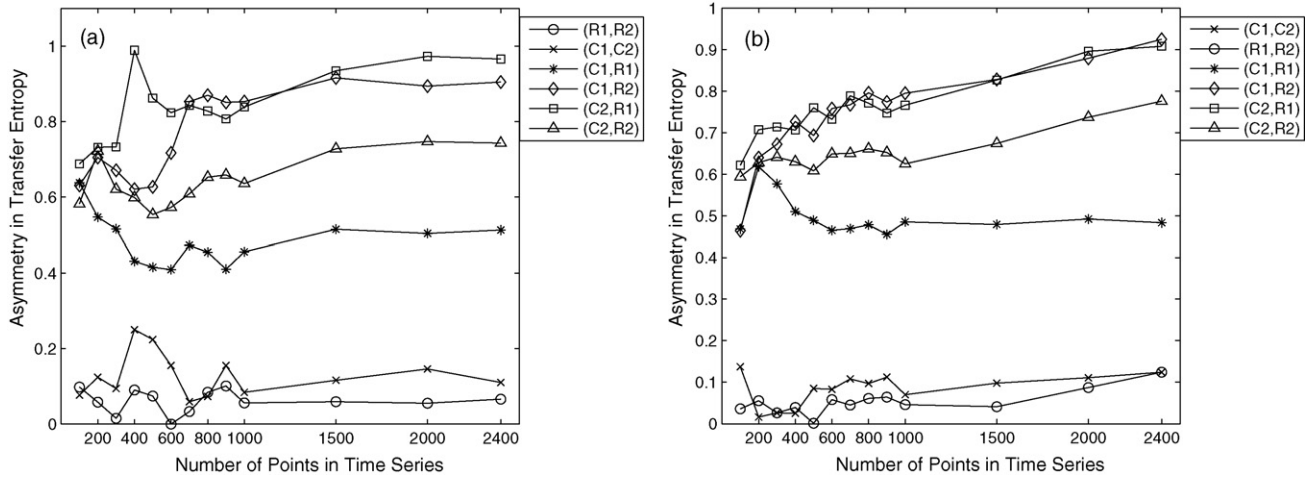
Fig. 6 – Asymmetry in transfer entropy – food web model – (e.g., (C1,C2) = |$TE_{C1 \to C2}$ − $TE_{C2 \to C1}$|) for (a) N = 5 and (b) N = 10 near neighbors in the density estimation. C1, consumer1; C2, consumer2; R1, resource1; R2, resource2. N refers to the number of near neighbors used in the density estimation ($\hat{p}(x)$) used to calculate mutual information and transfer entropy. All information measures are in bits.

topology and that obtained by using the information statistics are operationally identical.

5.2. Caveats

In order to compute densities for the time series, the time series must be stationary. For example, if there were a catastrophic event during collection of the data in which the availability of a resource changes completely (or a consumer appears or disappears), the density estimates will not be valid. It is possible to compute information statistics on stationary pieces of such data (to determine stationarity in a nonlinear time series, see, e.g., Kennel, 1997, Kantz and Schreiber, 1997
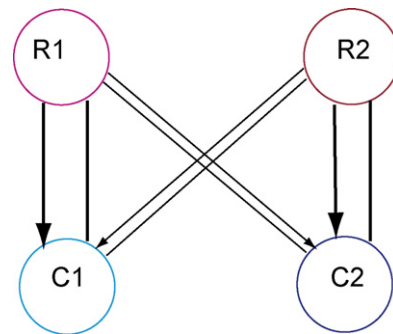

Fig. 8 – Reconstructed food web based only on analysis of time series data. The arrows indicate the direction of positive interaction. Thickness of the line indicates differences in the strengths of the interactions; thicker lines indicate stronger interactions. C1, consumer1; C2, consumer2; R1, resource1; R2, resource2.
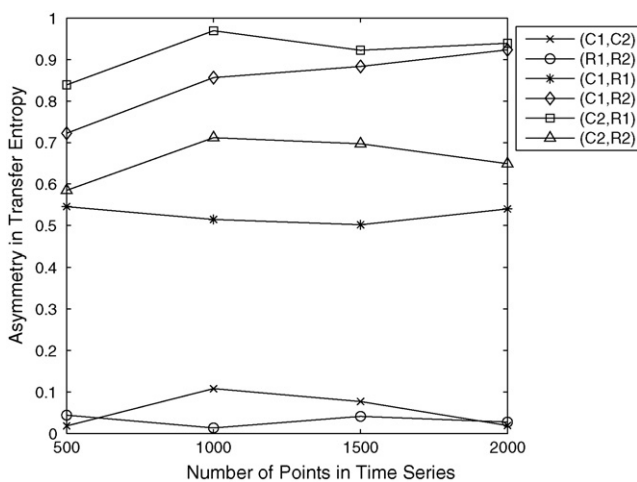

Fig. 7 – Asymmetry in transfer entropy – truncation 1k – using alternate fragment from the time series using N = 5 near neighbors in the density estimation. C1, consumer1; C2, consumer2; R1, resource1; R2, resource2. N refers to the number of near neighbors used in the density estimation ($\hat{p}(x)$) used to calculate mutual information and transfer entropy. All information measures are in bits.

and references therein) but accurate determination of the food web topology (using information statistics) during transient periods or using a non-stationary time series is not possible with the information statistics alone. However, we do not believe that this should necessarily be viewed as a methodological shortcoming. It may not be reasonable to think that a single topology may characterize a system undergoing change. The use of system topology to make predictions typically requires stationarity as well, so the stationarity assumption is not unusually restrictive. We note that conventional ecological wisdom indicates that transients are more information-rich than stationary dynamics. Although this may be the case, the information is not accessible via the transfer entropy.

In this analysis, the transients that appeared at the beginning of the time series (about 500 data points) were truncated in order to assure that the time series were stationary. Although it would be quite interesting to study the transient behavior of the food web before it settles into a steady state, it
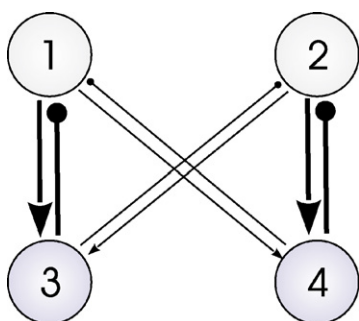
Fig. 9 – True food web topology, based on model parameters. Again, arrows indicate the direction of interaction and thicker lines indicate stronger interactions.

is not possible to do that with the statistics used here because of the stationarity requirement. Therefore, we wistfully leave the study of the transient behavior to another day and another metric.

We should mention that this model was not a complex one. While it was seasonally driven and did include noise, it included only one compartment and did not include any non-relevant participants. However, some of the authors have used transfer entropy and an attractor-based method, continuity, to investigate a spatially extended predator–prey model (Nichols et al. (2005); Moniz et al. (in press)) and real engineering systems (Nichols et al. (2006)) with notably more complexity. We have found that these attractor-based and signal-processing methods are able to distinguish very subtle effects. Thus, we are eager to work on the problem with a more complex model and possibly real time series data, if available, to determine interactions when they cannot be discerned by other methods.

5.3. *Consistency of results for short time series*

The importance of these results probably does not lie in the fact that we were able to uncover the relationships in the food web, but that the conclusions that we can reach using these statistics were robust to a 300–400 point time series. In this case, there is no hope of using attractor reconstruction order to investigate the dynamics of the system; it is generally conjectured (see, e.g., Tsonis, 1992) that in order to reach reasonable conclusions with attractor reconstruction methods it is necessary to have at least 10,000 points per dimension in a reconstruction. Although this requirement has not been theoretically proven and there have been studies that show a shorter time series can still yield meaningful results (Nichols et al., 2005) a *long* time series from ecological field data would tend to be fewer than 500 observations. Thus, the fact that the relationships between information statistics persisted in the 300–400 point time series is perhaps more important than the food web topology itself. The information statistics used here are as fully nonparametric as classical attractor reconstruction approaches (i.e., they make no assumptions whatsoever about the form of the underlying dynamical system). The difference is that the information statistics can exploit the fact that the complete state vector is observed. Having the attractor in hand rather than needing to reconstruct it from a single state variable, we could use information statistics to directly

probe the interrelationships among state variables and get meaningful results without requiring implausibly long data series for ecological systems.

---

6. Conclusions and recommendations

The information statistics correctly identified the food web topology in this simple model, including verifying the preference of one resource over another for a given consumer. These results were valid even for relatively short time series lengths. Thus, there is reason for further investigation of the use of information statistics in this context.

Although the blind analysis of the food web topology was able to correctly identify the relationships between the resources and consumers and the preference for one resource over another, the large (3-to-1) ratio of the preference of resource1 over resource2 for consumer1 and the corresponding preference for resource2 was not apparent from the statistics. However, this model did include some noise and although the community matrix could not be recovered from the data, the information statistics, when taken as a whole, were able to discern the preferences clearly for ecologically realistic time series lengths. More generally, by combining evidence from mutual information and transfer entropy statistics, we were able to distinguish between direct interactions (A eats B or competes with B) and indirect interactions (A and B are both eaten by C), which is essential for correctly recovering food web topology.

This was a very simple food web topology. More study is needed on more complicated systems, especially those with multiple trophic levels, to see if the relationships persist in the information statistics. Although the method looks promising in light of the results for short time series lengths, those results also need to be demonstrated to be consistent for different models or known relationships in real data before the method can be verified.

Our information-statistical analyses represent an extreme alternative to a conventional parametric analysis such as (A.2), avoiding any presumption that functional form of the data-generating process is either known *a priori* or can be inferred from the data. There is quite a bit of "model space" in between these extremes, raising the the possibility of extracting sharper information from observational data on food webs without having to make questionable assumptions about unstudied processes.

For example, semiparametric approaches replace (A.2) by a general regression model $R_i(t) = F_i(N(t), t)$ and the functions F_i are estimated by nonparametric regression. But given limited data, completely nonparametric estimation of the high-dimensional functions F_i is problematic at best. Success typically depends on being able to impose *a priori* qualitative constraints. Semiparametric analyses have recently been applied successfully to data on laboratory populations (e.g., Wood, 2001; Lingjaerde et al., 2001; Moe et al., 2002, 2005; Ellner et al., 2002), using constraints based on knowledge of how the study species progresses through different stages of the life cycle, and about which developmental steps were likely targets for density-dependence. Would it also be possible for an information-statistical analysis to be structured so as to gain

precision by exploiting solid basic knowledge about species' life cycles?

Finally, recall that inference based on multiple time series of abundance is the least direct of the three approaches listed for investigating species interactions. Direct investigation of interaction mechanisms and use of species removal experiments should yield stronger inferences. However, these more direct approaches typically require substantial effort and have not been carried out in most existing food webs. Approaches based on time series of abundances have potential utility for use with existing data obtained from the various animal monitoring programs established around the world (e.g., Peterjohn and Sauer, 1993; Gregory, 2000; Yoccoz et al., 2001; Webber et al., 2004). Even if a time series from one location has fewer than 300 observations, use of the data concatenation technique proposed by Banbrook et al. (1997) on time series from adjacent locations with similar habitats may yield a long enough time series to employ the techniques described in this paper successfully. Thus we believe that efforts to use information-theoretic statistics in food web investigations using such data hold promise.

Acknowledgements

Appendix A. Detailed description of the parametric analysis

The following analysis was done by a team member who had knowledge of the the structure of the data-generating model. A short summary is included in the body of the paper; details appear here. The following calculation describes a conventional parametric analysis using approaches like those in Laska and Wootton (1998), Ives et al. (1999, 2003) and Ellner et al. (2002).

Replacing (for the moment) the exogenous forcing z_t by its expected value $E[z_t] = 1$, the model has the form:

$$R(t) = \mathbf{A}N(t) + b \tag{A.1}$$

where $N = (\eta_4, \eta_4, \eta_4, \eta_4)$, R is the vector whose ith component is $(1/\eta_i)(d\eta_i/dt) = (d \log \eta_i/dt)$, $\mathbf{A}$ is a constant matrix of coefficients and b is a vector of coefficients (the constant terms in the expressions for R_i). The interpretation of (A.1) is that population change results from net per-capita birth and death rates, and each of those rates is (by assumption) a linear function of the vector of population densities. The goal of the data analysis is to estimate the entries of $\mathbf{A}$, which specify the pattern and strength of interactions among the species.

Given the weekly time-series "data", the values of R can be estimated well by first interpolating the values of $\log \eta_i(t)$ with a cubic spline, and then numerically differentiating the fitted spline. This simple method works here because the "data" are error-free and the data-generating dynamics are smooth. With real data, much more work would have to go into estimating the derivative. Using the estimated $R_i(t)$ values, the matrix $\mathbf{A}$ can be estimated by linear regression. As an upper bound on what could be achieved this way, while taking account of the exogenous forcing z_t, we fit a functional form that is very close to the truth:

$$R_i(t) = s_i^{(1)}(t \bmod 52) + s_i^{(2)}(t) + a_{i1}\eta_1 + a_{i2}\eta_2 + a_{i3}\eta_3 + a_{i4}\eta_4 \tag{A.2}$$

Here $s_i^{(j)}$ denotes a spline function of its argument. The two spline terms on the right-hand side of (A.2) are therefore a regular seasonal component (a function of week mod 52) and a smooth but otherwise irregularly varying function of time, which corresponds to the actual structure of the forcing z_t in the data-generating model. In the data-generating model both resource species experience the same forcing function, but that constraint was not imposed in the regression analysis.

The matrix $A = (a_{ij})$ used in generating the "data" was

$$\begin{bmatrix} -0.125 & 0.00 & -1.0 & -0.3 \\ 0.000 & -0.15 & -0.3 & -1.0 \\ 1.000 & 0.30 & -0.1 & 0.0 \\ 0.300 & 1.00 & 0.0 & -0.1 \end{bmatrix}$$

We ran the model to generate 15 years of monthly values, and discarded the first 5 years. In 100 replicates of this process, the elementwise average estimates were

$$\begin{bmatrix} -0.12 & 0.02 & -0.87 & -0.18 \\ 0.01 & -0.12 & -0.14 & -0.85 \\ 1.00 & 0.30 & 0.00 & 0.00 \\ 0.30 & 1.00 & 0.00 & 0.00 \end{bmatrix}$$

with standard deviations

$$\begin{bmatrix} 0.06310 & 0.04401 & 0.11067 & 0.10454 \\ 0.07661 & 0.05158 & 0.13281 & 0.12555 \\ 0.00011 & 0.00017 & 0.00005 & 0.00004 \\ 0.00017 & 0.00027 & 0.00008 & 0.00006 \end{bmatrix}$$

We note again here that the coefficients for the consumers (species 3 and 4) are estimated almost perfectly and that those for the resources (species 1 and 2) are not: the consumer impact coefficients are all underestimated. The sampling variability is high enough to frequently reverse the sign of the smaller consumption coefficient having true value -0.3.

For the second parametric analysis (8), both models were fitted using the *gam* function in *mgcv* package of R (Wood, 2004, 2005; R Core Development Team, 2005). The significance of the nonlinearity ((8) versus (A.2)) can be assessed by an F-test using *anova.gam*. For the four species, the (true) hypothesis of linearity was rejected at the $\alpha = 0.05$ level in (100, 100, 90, 90)

out of 100 replicates, and rejected at the $\alpha = 0.01$ level in (100, 100, 81, 80) out of 100 replicates.

REFERENCES

Abrams, P., 2001. Describing and quantifying interspecific interactions: a commentary on recent approaches. Oikos 94, 209–218.

Ayala, F., Gilpin, M., Ehrenfeld, J., 1973. Competition between species: theoretical models and experimental tests. Theor. Popul. Biol. 4, 331–356.

Banbrook, M., Ushaw, G., MacLaughlin, S., 1997. How to extract Lyapunov exponents from short and noisy time series. IEEE Trans. Signal Process. 45 (5), 1378–1382.

Bentley, J., 1979. Multidimensional binary search trees used for associative searching. Commun. ACM 18 (9), 509–517.

Berlow, E., 1999. Strong effects of weak interactions in ecological communities. Nature 398, 330–334.

Boudjema, G., Chau, N., 1996. Revealing dynamics of ecological systems from natural recordings. Ecol. Modell. 91, 15–23.

Constantino, R., Desharnais, R., Cushing, J., Dennis, B., 1997. Chaotic dynamics in an insect population. Science 275, 389–391.

Cushing, J., Constantino, R., Dennis, B., Desharnais, R., Henson, S., 2003. Chaos in Ecology: Experimental Nonlinear Dynamics. Academic Press, San Diego, CA USA.

Dennis, B., Desharnais, R., Cushing, J., Constantino, R.F., 1995. Nonlinear demographic dynamics: mathematical models, statistical methods, and biological experiments. Ecol. Monogr. 65, 261–281.

Ellner, S., Seifu, Y., Smith, R., 2002. Fitting population models to time series data by gradient matching. Ecology 83, 2256–2270.

Elton, C., 1927. Animal Ecology. Sidgwick and Jackson, London.

Fath, B.D., Patten, B.C., 1999. Review of the foundations of network environ analysis. Ecosystems 2, 167–179.

Fath, B.D., Cabezas, H., Pawlowski, C.W., 2003. Regime changes in ecological systems: an information theory approach. J. Theor. Biol. 222, 517–530.

Fraser, A., Swinney, H., 1986. Independent coordinates for strange attractors from mutual information. Phys. Rev. A 33, 1134–1140.

Goldwasser, L., Roughgarden, J., 1993. Construction and analysis of a large Caribbean food web. Ecology 74, 1216–1233.

Gregory, R., 2000. Development of breeding bird monitoring in the united kingdom and adopting its principles elsewhere. Ring 22, 35–44.

Hastings, A., Hom, C., Ellner, S., Turchin, P., Godfray, H., 1993. Chaos in ecology: is mother nature a strange attractor? Annu. Rev. Ecol. Syst. 24, 1–33.

Hjaltason, G.R., Samet, H., 1995. Ranking in spatial databases. In: Engenhoffer, Herring (Eds.), Proceedings of the 4th Symposium on Large Spatial Databases. Vol. 951 of Lecture Notes in Computer Science. Springer-Verlag, Berlin, pp. 83–95.

Ives, A., Carpenter, S., Dennis, B., 1999. Community interaction webs and the response of a zooplankton to experimental manipulations of planktivory. Ecology 80, 1405–1421.

Ives, A., Dennis, B., Cottingham, K., Carpenter, S., 2003. Estimating community stability and ecological interactions from time-series data. Ecol. Monogr. 73, 301–330.

Kaiser, A., Schreiber, T., 2002. Information transfer in continuous processes. Physica D 166, 43–62.

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

Kennel, M., 1997. Statistical test for dynamical non-stationarity in observed time-series data. Phys. Rev. E, 56.

Kullback, S., 1997. Information Theory and Statistics. Dover, Mineola, New York.

Laska, M., Wootton, J., 1998. Theoretical concepts and empirical approaches to measuring interaction strength. Ecology 79, 461–476.

Latham, L., 2006. Network flow analysis algorithms. Ecol. Modell. 192, 586–600.

Latham, L., Scully, E., 2004. Network optimization model implies strength of average mutual information in ascendancy. Ecol. Modell. 179, 373–392.

Libralato, S., Christensen, V., Pauly, D., 2006. A method for identifying keystone species in food web models. Ecol. Modell. 195, 153–171.

Liebert, W., Shuster, H., 1989. Proper choice of the time delay for the analysis of chaotic time series. Phys. Lett. A 142, 107–111.

Lingjaerde, O., Stenseth, N., Kristofferson, A., Smith, R., Moe, S., Read, J., Daniels, S., Simkiss, K., 2001. Exploring the density-dependent structure of blowfly populations by nonparametric additive modeling. Ecology 82, 2645–2658.

MacArthur, R., 1955. Fluctuations of animal populations and a measure of community stability. Ecology 36 (3), 533–536.

Marschinski, R., Kantz, H., 2002. Analysing the information flow between financial time series: an improved estimator for transfer entropy. Eur. Phys. J. B 30, 275–281.

Mayer, A., Pawlowski, C., Cabezas, H., 2006. Fisher information and dynamic regime changes in ecological systems. Ecol. Modell. 195, 72–82.

Moe, S., Kristofferson, A., Smith, R., Stenseth, N., 2005. From patterns to processes and back: analysing density-dependent responses to an abiotic stressor by statistical and mechanistic modeling. Proc. R. Soc.: B: Biol. Sci. 272, 2133–2142.

Moe, S., Stenseth, N., Smith, R., 2002. Density dependence in blowfly populations: experimental evaluation of non-parametric time-series modeling. Oikos 98, 523–533.

Moniz, L., Nichols, J., Nichols, J., in press. Mapping the information landscape: discerning peaks and valleys for ecological monitoring. J. Biol. Phys.

Nichols, J., 2005. Inferences about information flow and dispersal for spatially extended population systems using time series data. Proc. R. Soc. B 272 (1565), 871–876.

Nichols, J., Moniz, L., Nichols, J., Pecora, L., Cooch, E., 2005. Assessing spatial coupling in complex population dynamics using mutual prediction and continuity statistics. Theor. Popul. Biol. 67, 9–21.

Nichols, J., Seaver, M., Trickey, S., Salvino, L., Pecora, D., 2006. Detecting impact damage in experimental composite structures: and information theoretic approach. Smart Mater. Struct. 15, 424–434.

Pahl-Wostl, C., 1992. Information theoretical analysis of functional temporal and spatial organization in flow networks. Math. Comput. Modell. 16 (3), 35–52.

Paine, R., 1992. Food web analysis through filed measurement of per capita interaction strength. Nature 355, 73–75.

Peacor, S., Werner, E., 2004. How dependent are species-pair interaction strengths on other species in the food web? Ecology 85, 2754–2763.

Peterjohn, B., Sauer, J., 1993. North american breeding bird survey annual summary 1990–91. Bird Popul. 1, 1–15.

Pfister, C., 1995. Estimating competitive coefficients from census data: a test with field manipuations of tide pool fishes. Am. Nat. 146, 271–291.

Prichard, D., Theiler, J., 1995. Generalized redundancies for time series analysis. Physica D 84, 476–493.

R Core Development Team, 2005. R: A Language Environment for Statistical Computing. ISBN 3-900051-07-0, URL: http://www.r-project.org.

Rutledge, R.W., Basore, B.L., Mulholland, R.J., 1976. Ecological stability: an information theory viewpoint. J. Theor. Biol. 57, 355–371.

Schaffer, W., Kot, M., 1986. Chaos in ecological systems: the coals that newcastle forgot. Trends Ecol. Evol. 1, 58–63.

Schreiber, T., 2000. Measuring information transfer. Phys. Rev. Lett. 85, 462–464.

Scotti, M., Allesina, S., Bondavalli, C., Bodini, A., Abarca-Arenas, L., 2006. Effective trophic positions in ecological acyclic networks. Ecol. Modell. 198, 495–505.

Shannon, C., 1948. The mathematical theory of communication. Tech. Rep. 27, Bell System Technical Journal, pp. 379–423:623–656.

Shannon, C., Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press, Chicago, IL.

Silverman, B., 1986. Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability. Chapman and Hall, London, available online at: http://ece.ut.ac.ir/dbrg/seminars/Chehreghani/SpecialTopicInDAtabase/ReferencePapers/Density Estimation for Statistics and Data Analysis.pdf.

Tsonis, A., 1992. Chaos: From Theory to Applications. Plenum Press, New York.

Turchin, P., 2003. Complex Population Dynamics. Princeton University Press, Princeton, NJ, USA.

Ulanowicz, R.E., 1986. Growth and Development: Ecosystems Phenomenology. Springer Verlag, New York.

Ulanowicz, R.E., 2001. Information theory in ecology. Comput. Chem. 25, 393–399.

Ulanowicz, R.E., Abarca-Arenas, L.G., 1997. An informational synthesis of ecosystem structure and function. Ecol. Modell. 95, 1–10.

Ulanowicz, R.E., Puccia, C., 1990. Mixed trophic impacts in ecosystems. Coenoses 5, 7–16.

Vastano, J., Swinney, H., 1988. Information transfer in spatiotemporal systems. Phys. Rev. Lett. 60 (18), 1773–1776.

Webber, D., Hintermann, U., Zangger, A., 2004. Scale and trends in species richness; considerations for monitoring biological diversity for political purposes. Global Ecol. Biogeogr. 13, 97–104.

Wood, S., 2001. Partially specified ecological models. Ecol. Monogr. 71, 1–25.

Wood, S., 2004. Stable and efficeint multiple smoothing parameter estimation for generalized additive models. J. Am. Stat. Assoc. 99, 673–686.

Wood, S., 2005. mgcv:GAMs with GCV smoothness estimation and GAMMs by REML/PQL. URL: http://www.cran.R-project.org, version 1.3.8.

Wootton, J., 1997. Estimates and tests of per captia interaction strength: diet, abundance and impact of intertidally foraging birds. Ecol. Monogr. 67, 45–64.

Wright, S., 1921. Correlation and causation. J. Agric. Res. 20, 557–585.

Wright, S., 1934. The method of path coefficients. Ann. Math. Stat. 5, 161–215.

Yoccoz, N., Nichols, J., Boulinier, T., 2001. Monitoring of biological diversity in space and time. Trends Ecol. Evol. 16, 446–453.