DYNAMIC BASIS OF FOOD WEB ORGANIZATION

JOEL E. COHEN
The Rockefeller University, 1230 York Avenue, New York, New York 10021 USA

AND

CHARLES M. NEWMAN
Department of Mathematics, University of Arizona, Tucson, Arizona 85721 USA

Two well-established, and almost equivalent, empirical generalizations about community food webs are the link–species scaling law and the hyperbolic connectance law. The link–species scaling law says that, in community food webs with a moderate number of trophic species (≈ 5–50), the expected number of trophic links is directly proportional to the number of trophic species, with a coefficient of proportionality near 2; on average, community food webs have about twice as many trophic links as trophic species. The hyperbolic connectance law says that in community food webs with a moderate number of trophic species, the product of the number of species and the connectance (defined below) is approximately constant (with a value near 4), i.e., that the connectance is a hyperbolic function of the number of species. The link–species scaling law and the hyperbolic connectance law are mathematically equivalent when the number of trophic species is large compared with 1.

Here we explain these empirical laws quantitatively by combining a dynamic model of ecological communities with a model of the incompleteness of ecological observations. This conceptual link provides, for the first time, a partial dynamic basis for the cascade model, a static stochastic model that predicts important structural properties of community food webs. On the basis of parameter estimates obtained by fitting the models to the link–species scaling law, it is estimated that ecologists record in food webs 20% or fewer of the dynamic interactions among species in communities with 30 or more trophic species, and 10% or fewer of the dynamic interactions among species in communities with 50 or more trophic species.

A MODEL WITH A CHANGING COMMUNITY MATRIX

Paine (1988) pointed out that it would be highly desirable to find a dynamic basis for static empirical regularities and static stochastic models that describe observed community food webs. The same desire to find a dynamic basis has been expressed by others,

including Pimm (1982), Cohen and Newman (1985*b*: 442), Cohen et al. (1985:460), and Newman and Cohen (1986:376). Paine also pointed out, as have others, that ecological data generally, and especially data on trophic and dynamic interactions in communities, result from the selective and imperfect attention of ecologists.

Here we suggest that criteria for the probable stability or instability of an ecological community, in combination with a simple model of the incompleteness of ecological observations, can provide a partial dynamic basis for some static regularities (e.g., Briand 1983, Briand and Cohen 1984, Cohen and Briand 1984) and static models of food webs (e.g., Cohen and Newman 1985*b*, Cohen et al. 1985, 1986, Newman and Cohen 1986). The almost equivalent static empirical regularities that will be explained here are the link–species scaling law and the hyperbolic connectance law.

Others have proposed a dynamic derivation of the hyperbolic connectance law from May's criteria (1972, 1973) for the stability or instability of a model of ecosystems. Unfortunately, those criteria are erroneous for May's model (Cohen and Newman 1984, 1985*a*), and hence that proposed derivation fails.

Consider a simple, highly idealized community model (Cohen and Newman 1984) inspired by a model from May (1972, 1973). Suppose the state of a community is described by a real vector $x(t)$ of S_D elements. The positive integer S_D represents the number of species selected for a dynamic description of the community. These "dynamic" species might correspond to biological species, or to age groups within a species if age structure is important, or to genotypes within age groups if age structure and genetics are important. No matter what level of biological detail is selected, a purely biotic description of the state of a community represents a gross simplification that omits a large number of other potentially influential variables. Hence it is to be expected that a description of community dynamics in terms of S_D dynamic species will be incomplete and therefore subject to influences that may appear random within the framework of the biotic model.

The ith element $x_i(t)$ of $\mathbf{x}(t)$ is interpreted as the difference between the quantity (e.g., number of individuals or biomass) of dynamic species i at time t and the quantity of the same species at a hypothetical point of equilibrium, arbitrarily taken to be the origin $\mathbf{0}$. Suppose time t to be discrete, $t = 0, 1, 2, \ldots$. The generally nonlinear dynamics of the community are modeled by a linear approximation near the origin:

$$\mathbf{x}(t) = \mathrm{A}(t)\mathbf{x}(t - 1), \qquad t = 1, 2, \ldots, \qquad (1)$$

where $\mathbf{x}(0) \neq \mathbf{0}$ (i.e., the community is initially perturbed away from equilibrium) and each $\mathrm{A}(t)$ is a real $S_D \times S_D$ matrix with random elements $a_{ij}(t)$ independently and identically distributed for all $i = 1, \ldots, S_D$; all $j = 1, \ldots, S_D$; and all $t = 1, 2, \ldots$.

The matrix $\mathrm{A}(t)$ is called the community matrix. The element $a_{ij}(t)$ describes the effect of each incremental unit of $x_j(t - 1)$ on $x_i(t)$. The elements of $\mathrm{A}(t)$ represent, not only trophic relations, but all interactions that affect dynamics, including, e.g., competition, mutualism, and the effects of random environmental perturbations on these interactions. For this reason, we do not require that $a_{ij}(t) < 0$ whenever $a_{ji}(t) > 0$, as might be reasonable if we were modeling only trophic interactions, or that $a_{ii}(t) < 0$, as might be reasonable if density-dependent autoregulation were the only factor effective in controlling the size of the ith dynamic species. We assume that many factors affect dynamic interactions in the community and that trophic factors are not necessarily always dominant.

The assumption that the elements $a_{ij}(t)$ are random and independent for all i, j, and t reflects a lack of detailed knowledge of their actual values and is intended to approximate their behavior in an ensemble of communities. May (1972) assumed $\mathrm{A}(t)$ to be the same for all t. Allowing the community matrix to vary in time appears to be a step in the direction of realism.

Model 1 shares with May's (1972) model a potential weakness pointed out by DeAngelis (1975): neither represents explicitly a material or energy balance of biomass flows. How important this potential weakness actually is depends on how variable assimilation efficiencies are over time and from species to species, and on how large and variable autotrophic flows are compared with heterotrophic flows. Some more general conditions for Model 1 considered by Cohen and Newman (1984) may incorporate a material or energy balance in biomass flows, but that remains to be determined.

The distance of the community's state vector $\mathbf{x}$ from the point of equilibrium $\mathbf{0}$ will be measured by $\|\mathbf{x}\| = \max_j |x_j|$ (but the following results apply to most other norms, such as Euclidean distance). For any nonrandom initial perturbation $\mathbf{x}(0)$, the asymptotic rate of growth of $\|\mathbf{x}\|$ is $\lambda[\mathbf{x}(0)] = \lim_{t \to \infty} \|\mathbf{x}(t)\|^{1/t}$, provided the limit exists with probability 1. In the cases that will be considered here, $\lambda[\mathbf{x}(0)]$ exists almost surely (i.e., with probability 1), is nonrandom, and does not depend on the initial state of the community $\mathbf{x}(0)$ [assuming always that $\mathbf{x}(0) \neq \mathbf{0}$]. Consequently, it will make sense to speak simply of the asymptotic growth rate λ of perturbations from equilibrium, keeping in mind that λ depends on both the number of dynamic species and the distribution assumed for the random elements $a_{ij}(t)$.

The community described by the time-series of community matrices $\{\mathrm{A}(t)\}_{t=1}^{\infty}$ is defined to be stable if and only if $\lambda < 1$, because the deviations $\mathbf{x}(t)$ from equilibrium $\mathbf{0}$ almost surely eventually decay toward $\mathbf{0}$ at an exponential rate. Similarly, the community is defined to be unstable if and only if $\lambda > 1$, because the deviations $\mathbf{x}(t)$ from equilibrium almost surely eventually increase in size at an exponential rate. When $\lambda = 1$, the community is defined to be at the frontier of stability. For a fixed number of dynamic species, there are some distributions (those in the domain of stability) of the random elements $a_{ij}(t)$ where $\lambda < 1$; some distributions (those in the domain of instability) where $\lambda > 1$; and some distributions (those at the frontier of stability) where $\lambda = 1$.

A community that is unstable, according to the model, is in theory less likely to persist as a community. However, a community could fail to satisfy the criterion of local linear stability used here and still satisfy other criteria of stability, such as having stable limit cycles. So it is not surprising, as Auerbach (1984) suggests, that communities that fail to satisfy a criterion of stability are observed in nature. The concept of stability used here is widely used by others (e.g., Pimm 1982, 1984) and may be useful, but clearly has its limits.

As did May (1972), consider a sequence of communities with increasing numbers S_D of dynamic species. For each S_D, the state vectors $\mathbf{x}(t)$ and matrices $\{\mathrm{A}(t)\}_{t=1}^{\infty}$ are of size S_D and $S_D \times S_D$, respectively. The sequence of model communities is defined to be asymptotically stable (or asymptotically unstable) if and only if, for large enough S_D, the community is always stable, $\lambda < 1$ (or unstable, $\lambda > 1$), i.e., if and only if there exists a positive integer S_0 such that, for all $S_D > S_0$, the community is stable (or unstable).

A Model of Imperfect Observation

Prompted by Paine's (1988) criticisms of ecological data, we suppose that a field ecologist has a threshold for noting the existence of an interaction between two species. Specifically, we suppose that there exists a positive number Δ such that, whenever $-\Delta \leq a_{ij}(t) \leq +\Delta$, the ecologist records $a_{ij}(t) = 0$; while whenever $a_{ij}(t) < -\Delta$ or $a_{ij}(t) > +\Delta$, the ecologists records $a_{ij}(t)$ as having some nonzero value. Define $P(S_D)$ to be the probability

that the ecologist records $a_{ij}(t)$ as having some nonzero value in a community matrix with S_D species, i.e., the probability that $a_{ij}(t) > +\Delta$ or $a_{ij}(t) < -\Delta$. If $F(x)$ is the cumulative distribution function of $a_{ij}(t)$, i.e., the probability that $a_{ij}(t) < x$, then $P(S_D) = F(-\Delta) + 1 - F(\Delta)$.

Let $L_D(t)$ be the number of *recorded* nonzero elements of $A(t)$, i.e., the number of *recorded* dynamic "links" in the community. Since the probability that any one link $a_{ij}(t)$ is recorded as nonzero is just $P(S_D)$, $L_D(t)$ is the number of "successes" (i.e., nonzero elements) in S_D^2 trials with probability of success $P(S_D)$. Thus the random variable $L_D(t)$ is binomially distributed with parameters S_D^2 and $P(S_D)$. Consequently, for $t = 1, 2, \ldots$, the mean of $L_D(t)$ is $E[L_D(t)] = P(S_D)S_D^2$. Since this expectation is independent of time, the indexing by time t will henceforth be dropped when referring to the expected number of dynamic links. According to the binomial distribution, the variance of $L_D(t)$ is $Var[L_D(t)] = P(S_D)[1 - P(S_D)]S_D^2$. The variance is independent of time. Thus:

$$E(L_D) = P(S_D)S_D^2, \qquad (2a)$$

$$Var(L_D) = P(S_D)[1 - P(S_D)]S_D^2. \qquad (2b)$$

The variables in Equations 1 and 2 are dynamical variables. The interacting groups of organisms, S_D in number, are defined as narrowly as is necessary to predict dynamics. By contrast, the variables recorded in food webs are trophic variables. A food web matrix records only feeding interactions. "Trophic species," S_T in number, are equivalence classes that consist of dynamic species with identical sets of predators and identical sets of prey (Sugihara 1982, Briand and Cohen 1984). Thus $S_T \leq S_D$.

To translate dynamical variables into trophic variables, suppose that, when the number of species is sufficiently large, there is a direct proportionality between S_D and S_T; say $S_D = c_1 S_T$. Suppose also that there is a direct proportionality between L_D, the number of dynamic links or elements of $A(t)$ recorded as nonzero, and L_T, the number of recorded trophic links in the community's food web; say $L_D = c_2 L_T$. Then Equations 2a and 2b may be expressed in terms of variables measurable in trophically lumped food webs as

$$E(L_T) = P(c_1 S_T)(c_1^2/c_2)S_T^2, \qquad (3a)$$

$$Var(L_T) = P(c_1 S_T)[1 - P(c_1 S_T)](c_1^2/c_2^2)S_T^2. \qquad (3b)$$

Eq. 3a (but not 3b) remains valid even if the proportionality between L_D and L_T is not valid for the variables themselves but only for their means.

What is c_1? If dynamic species interact trophically according to the cascade model (Cohen and Newman 1985b) and if (as observed) there are twice as many links as species, the expected fraction of nonisolated

(i.e., trophically connected) dynamic species lost due to trophic lumping approaches $16e^{-8}/3 \approx 0.002$ as S_D becomes large (Cohen and Newman 1985b:440). Thus c_1 is very close to 1, and may reasonably be approximated as 1.

What is c_2? If dynamic species j eats dynamic species i, the trophic link from i to j contributes a negative term to the sum of terms that compose $a_{ij}(t)$ and contributes a positive term to the sum of terms that compose $a_{ji}(t)$; assuming the nontrophic terms do not cancel the trophic term in each element, one trophic link contributes to two nonzero dynamic links in the community matrix. Interactions between species other than through feeding could cause additional nonzero elements in the community matrix. As a first approximation, we shall take $c_2 = 2$.

This analysis of the proportionality between L_D and L_T may seem to contradict our justification for independent and identically distributed elements in the dynamic community matrix. Although we make no claim of a complete argument, that apparent contradiction can be greatly lessened by the remark following Eq. 3b.

Cohen and Briand (1984) and Cohen et al. (1986) observed empirically that, to good approximation,

$$E(L_T) = cS_T, \qquad (4)$$

with $c \approx 2$. This relation is the link–species scaling law. The question is to what extent the empirical linear relation (4) between species and links can be reconciled with the theoretical relations (2a or 3a). Evidently what is required is that $P(S_D)$ be inversely proportional to S_D.

MAJOR ASSUMPTIONS: LIFE AT THE FRONTIER OF STABILITY

We have described a model of community dynamics and a model of imperfect observation.

In spite of the limitations on the concept of stability used here, we suppose further that a community evolves to the frontier of stability, either by evolution in the number of species S_D or by evolution in the distribution of interaction coefficients or by both.

For each S_D, we also suppose that each element $a_{ij}(t)$ has a density function sufficiently smooth to satisfy the general conditions of Cohen and Newman (1984). Within these general assumptions, we shall consider two distinct possibilities: first, that the distribution of each element has a finite variance $\sigma^2(S_D)$ or, equivalently, a finite standard deviation $\sigma(S_D)$, and, second, that the distribution has an infinite variance. The reason for considering distributions with infinite variance will become clear. When $a_{ij}(t)$ has a mean, it will be 0. In all our specific examples, the distribution of $a_{ij}(t)$ will be symmetric.

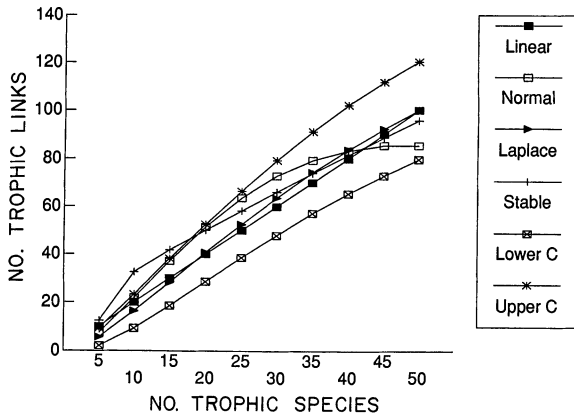We shall show that being at the frontier of stability

FIG. 1. Number of trophic links as a function of the number of trophic species. Linear: a straight line $E(L_D) = 2S_D$, summarizing the empirical link–species scaling law. Normal: the elements of the community matrix are assumed to be normally distributed, and $\Delta = 0.26$. Laplace: the elements of the community matrix are assumed to be Laplace distributed, and $\Delta = 0.2525$. Stable: the elements of the community matrix are assumed to be symmetric stably distributed, and $\kappa = \Delta^{-\alpha} = 15$. The upper and lower confidence intervals for the Laplace distribution span ± 3 standard deviations.

constrains the parameters of any distribution of $a_{ij}(t)$, and that this constraint in combination with the model of imperfect observation explains the link–species scaling law and the hyperbolic connectance law.

INTERACTION COEFFICIENTS HAVE FINITE VARIANCE

Suppose now that each element $a_{ij}(t)$ has a finite variance. We consider specifically the normal distribution and the Laplace (or two-tailed exponential) distribution (Johnson and Kotz 1970).

If $a_{ij}(t)$ is a normal random variable $N[0, \sigma^2(S_D)]$ with mean 0 and variance $\sigma^2(S_D)$, then the model community is at the frontier of stability ($\lambda = 1$) if and only if $\sigma^2(S_D) = (1/2)\exp[-\psi(S_D/2)]$, where ψ is the digamma function (Cohen and Newman 1984:287, Eq. 2.9). Let $\Phi(x) = P[N(0,1) < x]$ be the cumulative distribution function of the *standard* normal distribution with mean 0 and variance 1. The threshold of observation Δ is $\Delta/\sigma(S_D)$ standard deviations from the mean of $N[0, \sigma^2(S_D)]$. Since $\Phi(-x) = 1 - \Phi(x)$,

$$
\begin{aligned}
P(S_D) &= F(-\Delta) + 1 - F(\Delta) \\
&= \Phi[-\Delta/\sigma(S_D)] + 1 - \Phi[+\Delta/\sigma(S_D)] \\
&= 2[1 - \Phi(+\Delta\{(1/2)\exp[-\psi(S_D/2)]\}^{-1/2})].
\end{aligned}
$$

Using this formula in Eq. 3a, with $S_D = c_1 S_T$, gives an exact formula for the expected number of trophic links $E(L_T)$ as a function of the number of trophic species. That formula can be evaluated numerically using rational approximations for Φ (e.g., Abramowitz and Stegun 1965: 932) and finite sums for ψ at integral and

half-integral arguments (e.g., Abramowitz and Stegun 1965:258). We fixed the parameter values $c_1 = 1$ and $c_2 = 2$ and selected $\Delta = 0.26$ by numerical experimentation and visual inspection.

Fig. 1 shows that the predictions from the normal distribution (open rectangles) are not far from the empirical link–species scaling law (solid rectangles), but the normal predictions are curved concavely.

For distributions of $a_{ij}(t)$ other than normal, the exact variance at the frontier of stability is unknown when S_D is finite. More is known about the *asymptotic* frontier of stability in the limit as S_D becomes large. We shall use what is known by making the additional important assumption that the theory for large S_D can be applied when S_D is finite, as in data.

Under quite general conditions, which are satisified by the Laplace distribution among others, Cohen and Newman (1984) proved that if

$$\lim_{S_D \uparrow \infty} S_D \sigma^2(S_D) < 1, \tag{5}$$

then the sequence of model communities is asymptotically stable, while if

$$\lim_{S_D \uparrow \infty} S_D \sigma^2(S_D) > 1, \tag{6}$$

the sequence of model communities is asymptotically unstable.

Inequalities 5 and 6 are similar in form to May's (1972) criteria for stability and instability, respectively, but there is a crucial difference. Here the elements of the community matrix are assumed to have probability 0 of being equal to 0. Every dynamic species interacts with every other. If each element $a_{ij}(t)$ were equal to 0 with some positive probability, then the community modeled by Eq. 1 would be asymptotically stable with probability 1, regardless of $\sigma^2(S_D)$, because there would be a positive probability for each t that all the elements of $A(t)$ would be zero simultaneously, and hence it would be inevitable that for some t all the elements of $A(t)$ would be zero simultaneously.

Separating the domain of asymptotic stability (Inequality 5) from the domain of asymptotic instability (Inequality 6) is the asymptotic frontier of stability, which is close to the solution of $S_D \sigma^2(S_D) = 1$ for large S_D. Thus for large S_D, $\sigma^2(S_D) \approx 1/S_D$ at the asymptotic frontier of stability. As a first approximation, we shall take this approximate equality as an equality: $\sigma^2(S_D) = 1/S_D$. Let $u = \Delta/\sigma(S_D)$. Then $F(\Delta) = F([\Delta/\sigma(S_D)]\sigma(S_D)) = F[u\sigma(S_D)]$. At the asymptotic frontier of stability, we have, for large S_D,

$$
\begin{aligned}
P(S_D) &= F(-\Delta) + 1 - F(\Delta) \\
&= F[-u\sigma(S_D)] + 1 - F[+u\sigma(S_D)] \\
&\leq u^{-2} = \sigma^2(S_D)/\Delta^2 \approx 1/(S_D\Delta^2). \tag{7}
\end{aligned}
$$

The inequality in (7) is just the Bienaymé-Chebyshev inequality. This simple model of imperfect observation

implies that, for large S_D, the probability of observing a link must fall at least as fast as a constant times $1/S_D$. The fatter the tail of the distribution F, the closer the inequality in (7) will be to an equality. However, the inequality in (7) cannot be replaced by equality for all large S_D if the variance of $a_{ij}(t)$ is finite, as assumed so far.

The Laplace distribution has tails that are fatter and easier to describe analytically than those of the normal distribution. The Laplace distribution with density function $f(x) = (2\phi)^{-1}\exp(-|x|/\phi)$ has mean 0 and variance $2\phi^2$; the parameter ϕ is assumed to be positive. At the asymptotic frontier of stability, for the Laplace distribution, $2\phi^2 = \sigma^2(S_D) \approx 1/S_D$, and hence $P(S_D) = \exp(-\Delta[2S_D]^{1/2})$. Taking this approximation as exact for the purpose of calculation, we fixed the parameter values $c_1 = 1$ and $c_2 = 2$ and selected $\Delta = 0.2525$ by numerical experimentation. For $5 \leq S_T \leq 50$, the resulting agreement between the predicted $E(L_T)$ (Fig. 1, solid triangles) and the empirical linear function $2S_T$ (Fig. 1, solid rectangles) is remarkable.

The confidence intervals for the Laplace distribution given by ± 3 standard deviations around the predicted mean numbers of links (Fig. 1) include the curve predicted from the normal distribution. Confidence intervals are shown in Fig. 1 only for the Laplace distribution to avoid cluttering the graph.

The class of distribution functions of the interaction coefficients $a_{ij}(t)$ that will closely approximate the linear relation (Eq. 4) over the observed range of trophic species remains to be determined. The ease here of finding one such distribution function, the Laplace, suggests that this class of distribution functions is not small.

INTERACTION COEFFICIENTS HAVE INFINITE VARIANCE

To reconcile Prediction 3a from the model of imperfect observation with the empirical link–species scaling law (Eq. 4) requires distributions such that, as nearly as possible, $P(S_D)$ is proportional to $1/S_D$. We therefore give up the assumption that the distribution F has a finite variance in order to get a distribution with fatter tails. We again resort to the assumption that the theory for large S_D can be applied when S_D is finite.

Distributions of $a_{ij}(t)$ with infinite variance lead to criteria for asymptotic stability and asymptotic instability that are different from Inequalities 5 and 6 (Cohen and Newman 1984). While Inequalities 5 and 6 are valid for a robust family of models, they are not completely insensitive to the probability distribution of the interaction coefficients $a_{ij}(t)$.

We now consider symmetric stable distributions as possible distributions for the independently and identically distributed random elements $a_{ij}(t)$. Feller (1971: 169–176) gives a clear, elementary description of sym-

metric stable distributions and of their applications in astronomy, Brownian motion, and economics. The use of the term "stable" for these distributions is too well established in probability theory to change here in order to avoid conflict with the use of "stable" to describe ecological communities. The reader is warned that there is no direct connection between the two uses of the term.

Specifically, we suppose that each element $a_{ij}(t)$ has the distribution $\rho(S_B)W$, where W is a symmetric stable random variable of exponent α, and $0 < \alpha \leq 2$. Each element $a_{ij}(t)$ has a "spread" $\rho(S_D)$ that depends on S_D so as to keep the model community at the asymptotic frontier of stability. When $\alpha = 2$, W is a standard normal random variable; the standard deviation of $\rho(S_D)W$ is $\rho(S_D)$. When $\alpha < 2$, W has infinite variance, but $\rho(S_D)$ is still a natural measure of how spread out the distribution is. We know already that distributions with finite variance cannot make $P(S_D)$ be proportional to $1/S_D$ asymptotically for large S_D, though such distributions certainly can approximate this behavior for finite S_D. So we henceforth confine attention to $\alpha < 2$.

Let $\lambda(S_D, \alpha)$ denote the growth rate λ for a model community with S_D dynamic species when the elements $a_{ij}(t)$ have the distribution $\rho(S_D)W$. The assumption that the community evolves to the frontier of stability means that $\rho(S_D)$ is chosen so that $\lambda(S_D, \alpha) = 1$. We have proved that for any fixed α, $0 < \alpha < 2$, if $\lambda(S_D, \alpha) = 1$, then for any $\Delta > 0$,

$$P(S_D) = \frac{1}{\Delta^\alpha S_D \log S_D} + o\left(\frac{1}{S_D \log S_D}\right) \text{ as } S_D \uparrow \infty. \quad (8)$$

The proof is given in the Appendix.

Since $\log S_D$ is a slowly increasing function of S_D, $1/(S_D \log S_D)$ should be very close to $1/S_D$ as desired. Asymptotically, $P(S_D)$ does not depend on Δ and α separately, but only through the combination $\kappa = 1/\Delta^\alpha$.

Fig. 1 shows that for any combination of α and Δ such that $\kappa = 15$ (a value obtained by numerical experimentation), the expected number of trophic links (+ marks) given by Eq. 3a can approximate the empirical linear relation (Eq. 4). The predicted expected numbers of links display notable concavity between 5 and 15 species. Above 10 trophic species, the predictions are very nearly linear, but with a slope <2.

Altogether, Fig. 1 shows that if every $a_{ij}(t)$ is assumed to have a Laplace distribution and if an appropriate threshold of observation Δ is assumed, then the expected number of trophic links increases nearly linearly, with slope near 2, as the number S_T of trophic species increases from 5 to 50. The confidence intervals around the predictions of the Laplace distribution include most of the predictions that follow if every $a_{ij}(t)$
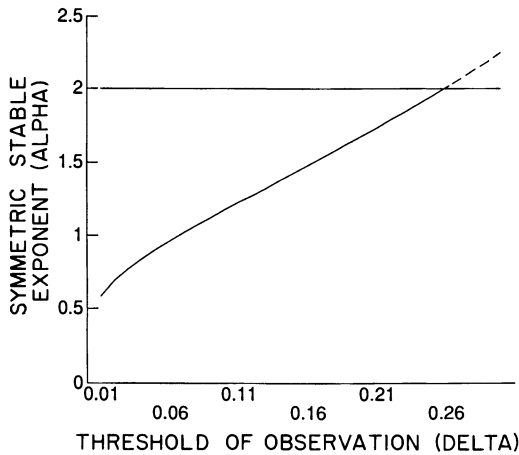
FIG. 2. Combinations of parameters α and Δ such that 15 = $\Delta^{-\alpha}$. The horizontal line marks $\alpha = 2$.

is assumed to have a normal or symmetric stable distribution. Given broad freedom in the choice of a distribution for $a_{ij}(t)$, this model of dynamics and imperfect observation can reproduce well the empirical link–species scaling law.

Given that $\kappa = 15$ approximately for an assumed symmetric stable distribution of $a_{ij}(t)$, it is possible to compute α as a function of Δ. Since the theory requires that α cannot exceed 2, not all values of Δ are permitted. Fig. 2 shows that for $\Delta > 0.26$ approximately, α does exceed 2. Symmetric stable distributions with infinite variance require that $\Delta < 0.26$. It is interesting that the best value of Δ, obtained independently of Fig. 2, for normally distributed $a_{ij}(t)$ is $\Delta = 0.26$, and for Laplace distributed $a_{ij}(t)$ is $\Delta = 0.2525$. In the present framework, these values of Δ correspond to a value of α just at or under 2.

CONNECTANCE

Connectance C is usually, and here, defined as $L_T/[S_T(S_T - 1)/2]$. The mean and the standard deviation of connectance are given by the corresponding quantities (Eq. 3a and square root of Eq. 3b) for L_T, each divided by $[S_T(S_T - 1)/2]$. The standard deviation of connectance and the weak law of large numbers give asymptotically normal confidence intervals for connectance. To our knowledge, confidence intervals for connectance have not been derived previously.

The empirical link–species scaling law (Eq. 4) is mathematically equivalent, as Cohen and Briand (1984) pointed out, to a hyperbolically decreasing relation between expected connectance and the number of trophic species, when the number of trophic species is large compared with 1, i.e., when $S_T/(S_T - 1)$ approximates 1. In observed food webs, the product of connectance

times the number of species is roughly constant at 4 (see Rejmánek and Starý 1979, Pimm 1982, and Auerbach 1984:418). This is equivalent to the numbers of trophic links being twice the number of trophic species, or $c \approx 2$ in Eq. 4.

If $P(S_D)$ in Eq. 3a is inversely proportional to S_D, as required by the empirical link–species scaling law (Eq. 4), then expected connectance is inversely proportional to the number of trophic species, as observed. In this case, for large numbers of trophic species, the standard deviation of connectance approaches zero.

Fig. 3 shows the expected connectance C computed from the hyperbolic curve $4/S_T$ (solid rectangles) and from assumed normal, Laplace, and symmetric stable distributions for each element $a_{ij}(t)$, in the range of S_T from 5 to 50 species. Confidence intervals for connectance derived from the Laplace distribution are also shown in Fig. 3. As might be expected, the expected connectance from the Laplace distribution agrees best with the hyperbolic law. Above 25 trophic species, there is little possibility of distinguishing among the models empirically.

PROBABILITY OF RECORDING A LINK

Under all the preceding assumptions, we can estimate the probability that ecologists record any given dynamic link as a function of the number of trophic species. This probability is just $P(c_1 S_T)$. Fig. 4 shows the estimates of this probability using the normal, Laplace, and symmetric stable distributions. The three hypothetical distributions give divergent estimates of the probability of recording a dynamic link when the number of trophic species is below 20 or so, but remarkably similar estimates for larger numbers of trophic
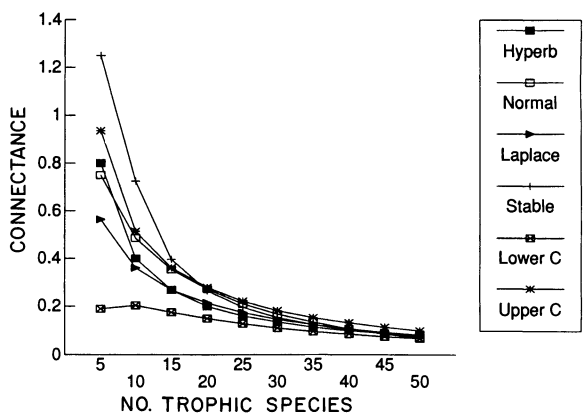

FIG. 3. Connectance as a function of the number of trophic species. Hyperbolic: $C = 4/S_D$, summarizing the empirical hyperbolic connectance law. The normal, Laplace, and symmetric stable distributions are as in Fig. 1.

species. According to the model developed here, the probability of recording a dynamic link when the number of trophic species is large is startlingly low. It is estimated that in communities with 30 or more trophic species, 20% or fewer of dynamic interactions among species are recorded, while in communities with 50 or more trophic species, 10% or fewer of dynamic interactions among species are recorded.

ASSUMPTIONS AND ALTERNATIVES

We have made many assumptions in deriving the link–species scaling law and in estimating the fraction of recorded dynamic links in community food webs. Ecologists will receive these assumptions with the skepticism they deserve. To reinforce that skepticism, we review here the major assumptions and their biological interpretations. We consider first the model of community dynamics, then the model of imperfect observation.

The dynamic model assumes that it is useful to describe the biological components only, omitting physical components, of community dynamics; that an equilibrium point exists for the abundance or biomasses of the dynamic species; that the community is always observed near enough to its equilibrium point so that the dynamics can usefully be approximated as linear; and that in this linear representation of the biotic dynamics, it is not necessary to impose any structure whatsoever on the interaction coefficients. Balances of materials and energy are ignored. The elements of the community matrix are *not* assumed to be negative on the diagonal, as they might be if such elements represented autoregulation of a population, and are *not* assumed to be antisymmetric, as they might be if symmetric pairs of elements represented only feeding relations; instead, each element of the community matrix is assumed to combine so many different terms, representing so many different kinds of biological and environmental interactions, that it is reasonable to treat all elements of the community matrix as independently and identically distributed random variables.

The community matrix elements are assumed to have mean zero, so that on average no species consistently influences any other, or itself, favorably (positively) or unfavorably (negatively). The matrix elements are assumed to change randomly in time, e.g., from positive in one time period to negative in the next, through the influence of the nonlinear interactions of all the dynamic variables that are omitted from explicit representation in the model. The distribution of community matrix elements is assumed to be smooth; all gradations of interaction strength between any two species are possible. The distribution of community matrix elements is assumed to have no mass concentrated at 0; this means that every species is supposed to affect
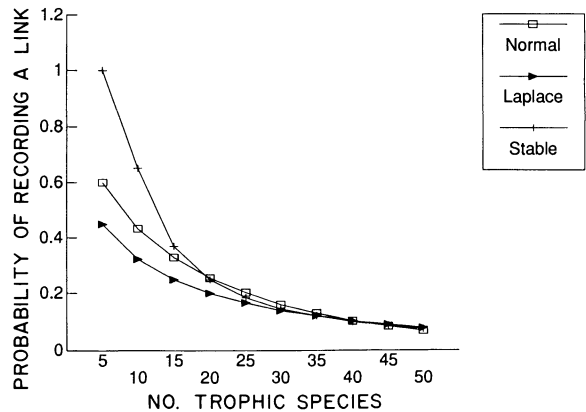

FIG. 4. Estimated probability of recording a dynamic link as a function of the number of trophic species, according to the model developed here. The normal, Laplace, and symmetric stable distributions are as in Fig. 1.

directly every other. This last assumption contrasts with that of May (1972), who assumed that any pair of species has a positive probability of not interacting directly at all.

The analysis of this dynamic model assumes that a community must be stable in the technical sense of linear local stability analysis. Volumes have been written about the biological interpretation of this assumption (see Pimm 1982, 1984).

The analysis assumes further that species diversity (i.e., number of species) and the variability of the community matrix elements are adjusted so that a community is at the frontier of linear local stability. By what biological mechanisms could this adjustment occur? Communities in the domain of instability could suffer local extinctions of species, while communities within the domain of stability could support speciation or immigration of species. By these means, the number of species could move toward the frontier of stability even when the variance of the community matrix elements is constant. For a fixed number of species, behavioral and ecological mechanisms might permit changes in the variance of community matrix elements that would move the community toward the frontier of stability. Such changes might be expected to be more rapid than evolutionary changes in the number of species.

When the community matrix elements are Laplace or symmetric stably distributed, we do not know exactly where the frontier of stability is for finite numbers of species. So we resorted to an asymptotic analysis for large numbers of species. In using that asymptotic analysis, we assumed that the exact frontier of stability for finite numbers of species is close to that derived from the asymptotic theory. For the normal distribution, we

know where the frontier of stability is exactly for finite numbers of species and asymptotically for large numbers of species; although we have not given details here, the asymptotic frontier is not far from the exact. We therefore think that the asymptotic frontier may be a good approximation to the exact frontier for nice smooth distributions in addition to the normal distribution.

The numerical calculations included three possible forms for the distribution of community matrix elements: normal, Laplace, and symmetric stable. If the community matrix elements are the sum of many, small, independent terms that reflect various forms of interaction between two species, they will be (under certain assumptions) approximately normal by the central limit theorem. Models could also be constructed to lead to the Laplace and symmetric stable distributions. In fact those distributions, like the normal, were chosen for mathematical convenience, and should not be taken seriously in advance of data about the real distribution of community matrix elements. What is remarkable is that, for model communities with >20 or 30 species, the choice of a hypothetical distribution for community matrix elements appears to make very little difference to the predicted link–species scaling law or to the estimated probability of recording a dynamic interaction.

So much for the dynamic model. We turn now to the model of imperfect observation. The model assumes that an ecologist who observes a given community has a single, fixed threshold of observation such that any interaction below that threshold is overlooked altogether. The model assumes that the threshold is the same for every pair of species: the ecologist is hypothetically as interested in granivorous birds as in insectivorous reptiles. The model assumes the threshold is the same for all time: the ecologist never gets tired and never improves field techniques or hires more field assistants. The model assumes the threshold is the same regardless of the number of species in the community: smaller numbers of species are assumed not to permit more sensitive observations. The model assumes the threshold of observation is the same for all ecologists; this assumption seems even less likely to be true than the other assumptions. A possible justification for these assumptions about the threshold is that the assumed constant threshold actually refers to the average of various fluctuating thresholds, and that the fluctuations around that threshold have insigificant effects compared with other sources of randomness.

In translating the results of the calculations from hypothetical dynamical species and community matrix elements into potentially observable trophic species and trophic links, we assumed simple proportionality. We assumed the constants of proportionality to be the

same for all communities, regardless of the number of species or links. We offered a justification for the assumption that there are about as many dynamic species as trophic species, and about twice as many recorded (i.e., above-threshold) community matrix elements as trophic links. It would be desirable to investigate empirically what contribution trophic links make to elements of a community matrix.

These many assumptions make it possible to link a dynamic model of community stability and a model of the imperfection of ecological observations with the empirical link–species scaling law. We believe it is unlikely that any of the detailed assumptions faithfully mirrors reality, but we would not be surprised if the qualitative behavior derived from the collection of assumptions corresponds well to what would be predicted by vastly more realistic models. It would be desirable in the future to test the assumptions directly. For now, only a desire for intellectual coherence justifies them.

The derivation of the link–species scaling law given here is not the only possible explanation of that law, nor even necessarily the most attractive one. Pimm (1982: 189) observed that if the average number of prey species per predator species is independent of the total number of species in a community, then the product of the number of species and the connectance is independent of the number of species. Limitations on the behavioral repertoire or on the trophic apparatus of a consumer might limit the number of its prey species regardless of the total number of species in a community. Thus the link–species scaling law could be derived from behavioral or anatomical hypotheses as well as from dynamical and observational hypotheses. Of course, the average recorded number of prey species per predator species could be independent of the total number of species because the observing ecologist simply stops recording prey species when their contribution to a predator's diet becomes too small. Thus observational imperfections may act jointly with behavioral and anatomical mechanisms.

WHERE DOES THE OTHER ASSUMPTION OF THE CASCADE MODEL COME FROM?

The cascade model of Cohen and Newman (1985b) rests on two key assumptions: first, that connectance is inversely proportional to the number of trophic species (equivalently, the link–species scaling law); and second, that trophic species are arranged in a linear "pecking order" of possible predator–prey relations. We have just suggested models that lead to the first assumption. The second assumption was introduced (Cohen and Newman 1985b) to eliminate an excess of cycles in model food webs that lacked this assumption.

One might hope to relate the assumed linear order of possible predator–prey relations in the cascade model to the empirical finding that the overlaps in the diets of predators can very often be represented by intervals of a one-dimensional continuum (Cohen 1978, Sugihara 1982). However, it is not difficult to construct a hypothetical web whose trophic overlaps can be represented in intervals but that has no linear order as assumed in the cascade model, as well as a hypothetical web consistent with the cascade model whose trophic overlaps can not be represented by intervals. Indeed, J. Cohen and Z. Palka (*personal observation*) have shown that, according to the cascade model, the probability of intervality is near 1 for very small webs, but approaches 0 for very large webs. Intervality cannot explain the ordering assumed in the cascade model.

Several other approaches have been suggested to explain the rarity of cycles and the possibility of a linear ordering of species. Based on numerical simulations, Sugihara (1982) argued that webs with "triangulated niche overlap graphs" (a slightly weaker property than intervality) are more likely to be stable than webs without this property. Pimm and Rice (1987) found, in numerical simulations of Lotka-Volterra equations, that trophic cycles decrease stability; like Sugihara (1982), they argued that the requirement of stability explains the observed rarity of cycles in real food webs. Another approach is based on the observation that, in food chains that exclude parasites and decomposers, consumers are usually larger than their prey. Cohen (1988) and independently Warren and Lawton (1987) suggested that an ordering of consumers by body size may explain the ordering assumed in the cascade model, and Warren and Lawton offered one food web as supporting evidence.

LITERATURE CITED

Abramowitz, M., and I. A. Stegun. 1965. Handbook of mathematical functions with formulas, graphs, and mathematical tables. Dover, New York, New York, USA.

Auerbach, M. J. 1984. Stability, probability and the topology of food webs. Pages 412–436 *in* D. R. Strong, Jr., D. Simberloff, L. G. Abele, and A. B. Thistle, editors, Ecological communities: conceptual issues and the evidence. Princeton University Press, Princeton, New Jersey, USA.

Briand, F. 1983. Environmental control of food web structure. Ecology 64:253–263.

Briand, F., and J. E. Cohen. 1984. Community food webs have scale-invariant structure. Nature 307:264–266.

Cohen, J. E. 1978. Food webs and niche space. Princeton University Press, Princeton, New Jersey, USA.

———. 1988. Food webs and community structure. *In* J. Roughgarden, R. M. May, and S. Levin, editors. Perspectives on ecological theory. Princeton University Press, Princeton, New Jersey, USA.

Cohen, J. E., and F. Briand. 1984. Trophic links of community food webs. Proceedings of the National Academy of Sciences (USA) 81:4105–4109.

Cohen, J. E., F. Briand, and C. M. Newman. 1986. A stochastic theory of community food webs. III. Predicted and observed lengths of food chains. Proceedings of the Royal Society of London B 228:317–353.

Cohen, J. E., and C. M. Newman. 1984. The stability of large random matrices and their products. Annals of Probability 12(2):283–310.

Cohen, J. E., and C. M. Newman. 1985a. When will a large complex system be stable? Journal of Theoretical Biology 113:153–156.

Cohen, J. E., and C. M. Newman. 1985b. A stochastic theory of community food webs. I. Models and aggregated data. Proceedings of the Royal Society of London B 224:421–448.

Cohen, J. E., C. M. Newman, and F. Briand. 1985. A stochastic theory of community food webs. II. Individual webs. Proceedings of the Royal Society of London B 224:449–461.

DeAngelis, D. L. 1975. Stability and connectance in food web models. Ecology 56:238–243.

Feller, W. 1971. An introduction to probability theory and its applications. Second edition. John Wiley and Sons, New York, New York, USA.

Johnson, N. L., and S. Kotz. 1970. Continuous univariate distributions—2. Houghton Mifflin, Boston, Massachusetts, USA.

May, R. M. 1972. Will a large complex system be stable? Nature 238:413–424.

———. 1973. Stability and complexity in model ecosystems. Princeton University Press, Princeton, New Jersey, USA.

Newman, C. M., and J. E. Cohen. 1986. A stochastic theory of community food webs. IV. Theory of food chain lengths in large webs. Proceedings of the Royal Society of London B 228:355–377.

Paine, R. T. 1988. Food webs: road maps of interactions or grist for theoretical development? Ecology 69:1648–1654.

Pimm, S. L. 1982. Food webs. Chapman and Hall, London, England.

———. 1984. The complexity and stability of ecosystems. Nature 307:321–326.

Pimm, S. L., and J. C. Rice. 1987. The dynamics of multispecies, multi-life-stage models of aquatic food webs. Theoretical Population Biology 32:303–325.

Rejmánek, M., and P. Starý. 1979. Connectance in real biotic communities and critical values for stability of model ecosystems. Nature 280:311–313.

Sugihara, G. 1982. Niche hierarchy: structure, organization and assembly in natural communities. Dissertation. Princeton University, Princeton, New Jersey, USA.

Warren, P. H., and J. H. Lawton. 1987. Invertebrate predator–prey body size relationships: an explanation for upper triangular food webs and patterns in food web structure? Oecologia (Berlin) 74:231–235.

APPENDIX

Proof of Eq. 8

According to Cohen and Newman (1984:289–290, [2.17], [2.20], [2.22] and [2.25]), when $\lambda(S_D, \alpha) = 1$,

$$(S_D \log S_D)[\rho(S_D)]^\alpha \to 1/K(\alpha) \text{ as } S_D \uparrow \infty, \qquad (9)$$

where $K(\alpha) = 2\Gamma(\alpha)\sin(\alpha\pi/2)/(\alpha\pi)$. Now using Eq. 2.26 of Cohen and Newman (1984:290),

$$P(S_D) = \text{Prob}[|\rho(S_D)W| \geq \Delta] = \text{Prob}\{|W|^\alpha \geq [\Delta/\rho(S_D)]^\alpha\}$$

$$= K(\alpha) \int_{[\Delta/\rho(S_D)]^\alpha}^{\infty} \nu^{-2} \, d\nu + O\left(\int_{[\Delta/\rho(S_D)]^\alpha}^{\infty} \nu^{-3} \, d\nu\right)$$

$$= K(\alpha)[\rho(S_D)/\Delta]^\alpha + O([\rho(S_D)/\Delta]^{2\alpha}).$$

Multiplying both sides by $S_D \log S_D$ and using Expression 9 several times gives

$$(S_D \log S_D)P(S_D) = 1/\Delta^a + o(1),$$

which leads to Eq. 8.