

A stochastic theory of community food webs II. Individual webs

BY J. E. COHEN¹, C. M. NEWMAN² AND F. BRIAND³

¹ *Rockefeller University, 1230 York Avenue, New York, New York 10021–6399, U.S.A.*, ² *Department of Mathematics, University of Arizona, Tucson, Arizona 85721, U.S.A.*, and ³ *Department of Biology, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5*

(Communicated by J. H. Steele, F.R.S. – Received 4 July 1984)

The species scaling law and the link scaling law of community food webs can be derived from a simple mathematical model, called the cascade model, which incorporates the link-species scaling law. In the previous test of this model against data on 62 community food webs, the ratio of links to species is estimated from aggregated data on all webs taken together, on the assumption that the ratio is independent of the number of species in the web. This paper demonstrates that the ratio of links to species shows no pronounced increasing or decreasing trend, but varies substantially, over the observed range of variation in the number of species in a web. However, the ratio is higher for webs in constant environments than for webs in fluctuating environments. When the ratio of links to species is estimated separately for each web, the cascade model provides a good description of the numbers of intermediate species and of basal–intermediate, intermediate–intermediate, and intermediate–top links, aside from a single outlying web. The cascade model provides a fair description of the numbers of top and basal species, and a rather poor description of the number of basal–top links. The model describes the kinds of species and kinds of links of constant and fluctuating webs about equally well.

1. INTRODUCTION

A *food web* is a set of different kinds of organisms and a relation that shows the kinds of organisms, if any, that each kind of organism in the set eats. A *community food web* is a food web whose vertices are obtained by picking, within a habitat or set of habitats, a set of kinds of organisms (hereafter called *species*) on the basis of taxonomy, size, location, or other criteria, without prior regard to the eating relations (specified by trophic *links*) among the organisms (Cohen 1978, pp. 20–21).

In paper I (Cohen & Newman 1985) several models were proposed to describe the structure of community food webs. When models were tested against data on 62 community food webs in paper I, a crucial parameter in all the models, namely the ratio of links to species, was estimated from the aggregated data on all webs taken together. One model, the cascade model, successfully described, to a first approximation, the proportions of all species that are top, basal and intermediate, and the proportions of all links of each kind.

The purpose of this paper is to test how well the cascade model describes webs when the ratio of links to species is estimated separately for each web.

In §2 we describe the cascade model, show how to estimate the parameters of the model, and verify the correctness of the estimation procedure. In §3 we test the assumption, made in paper I, that the ratio of links to species is constant for all webs. We then test seven predictions of the cascade model, estimating this ratio separately for each web. In §4 we evaluate the results of this paper and relate them to the results of paper I.

We shall use a number of terms with special meanings that are given in §2 of paper I. These terms include: web, species, link, predator, prey, top, proper top, intermediate, basal, proper basal, adjacency matrix, isolated, triangular. We shall not repeat the definitions here.

Webs are classified as arising in 'fluctuating' or 'constant' environments. The environment is considered to be 'fluctuating' if the original report indicates temporal variations of substantial magnitude in temperature, salinity, water availability or any other major physical parameter. The magnitude, and not the predictability, of the fluctuations is the criterion of classification. Since the classification of an environment as constant or fluctuating is to some extent subjective, we point out that this task was carried out before we had analysed the webs and uncovered any pattern.

The 62 webs analysed here are drawn from published studies. They include the 40 webs assembled and described by Briand (1983). Of these, 13 are drawn and corrected from the 14 originally used by Cohen (1978). Details of the webs not yet presented will be published elsewhere (F. Briand, unpublished data). These data are used by Briand & Cohen (1984) and Cohen & Briand (1984).

2. THE CASCADE MODEL AND PARAMETER ESTIMATION

The cascade model assumes that the S species of a web may be labelled from 1 to S so that, for some finite positive real number $c \leq S$, the probability that species j feeds on species i is 0 if $j \leq i$. If $i < j$, then j feeds on i with probability $p = c/S$ and does not feed on i with probability $q = 1 - c/S$, independently for all $1 \leq i < j \leq S$.

All numerical predictions of the cascade model depend on the values of the model's two parameters c and S . These two parameters, in turn, depend only on the observed numbers of links and of species.

In the data we shall use to test this model, only proper top species (that is, those that eat at least one other species) and only proper basal species (that is, those that are eaten by at least one other species) are reported. Thus the total number of *observed* species in a web is not S but the number of *not isolated* species. The true number S of species in the web is not directly counted.

The expected number $E(N)$ of not isolated species depends on both c and S according to (6.3c) in paper I. Similarly, the expected number $E(L)$ of links in a web depends on c and S according to (6.5) in paper I. To test the predictions of the cascade model with individual webs, we estimate c and S by the method of moments. That is, if S' is the observed number of species (that is, S' is the observed

value of the random variable N , the number of not isolated species), and L' is the observed number of links (that is, L' is the observed value of the random variable L , the number of links in a web), we replace $E(N)$ on the left of (6.3c) in paper I by S' and $E(L)$ on the left of (6.5) in paper I by L' . The resulting equations are restated as (A 1) in the appendix. We then solve this system of two nonlinear equations for the two unknowns c and S by using Newton's method, as described in the appendix, except for the one web with $S' = 3$. For this web, we take $S = 3$ and then compute c by solving (A 1a).

As a check on the correctness of the numerical solutions c and S , we used the numerical values of c and S to compute $E(L)$ from (6.5) in paper I and $E(N)$ from (6.3c) in paper I. In figures not shown, we plotted L' as a function of the calculated $E(L)$ and S' as a function of the calculated $E(N)$. A line of slope one through the origin passed through all the plotted characters except, as expected, the web with $S' = 3$, verifying that the computed numerical solutions for c and S in fact satisfy (A 1) adequately.

The computed values of S are not in general integers. We could force them to be integers by replacing S with the integer closest to S and then solving (A 1a) for a new value of c . A simpler alternative, which we adopt here, is to interpret the equations of the cascade model derived in paper I as applying whether S is integral or positive real.

From (6.3c) in paper I or (A 1b), it follows that if $S \gg 1$ and $e^{-c} \ll 1$, then S is approximately S' . In the data plotted in figure 1, S does not greatly exceed the observed values of S' . In fact $S - S' < 2.1$ for all webs but one. For the one exceptional web (Paviour-Smith 1956), $S - S' = 5.5$, where $S = 37.5$. This exceptional case is visible as the outlying fluctuating web in the lower right corner of figure 1. Briand & Cohen (1984) also noted that this web was an outlier on a plot of prey against predators based on unlumped data. This web appears to be unusual in both the relation between links and total species and the relation between predators and prey.

3. TESTING THE PREDICTIONS OF THE CASCADE MODEL

The tests of the cascade model in paper I use a single value of c for all webs. If this procedure is correct, then a plot of c against S , estimated individually for each web, should display no increasing or decreasing trend. Substantial variability in c as a function of S is expected because the realized number of not isolated species need not exactly equal the mean $E(N)$ and the realized number of links need not exactly equal the mean $E(L)$.

Figure 1 shows that there is no evidence of a pronounced trend in the estimated c as a function of the estimated S . Because S and $S - 1$ are close to the observed number S' of not isolated species, the observation that $c = 2L'/(S - 1)$ has no significant trend as a function of S follows from the link-species scaling law (paper I) that L'/S' has no increasing or decreasing trend as a function of S' .

The observation of a slightly positive slope in figure 1 is consistent with two earlier observations. First, by using multiple versions of the unlumped community webs of Cohen (1978), Yodzis (1980) observed that with increasing S' , the observed

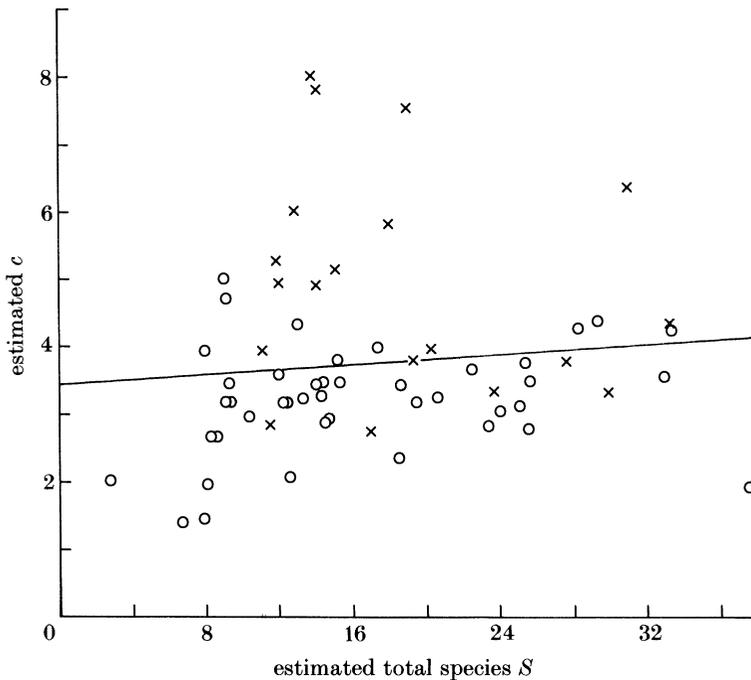


FIGURE 1. The estimated value of c as a function of the total number of species S . The straight line $c = 3.438 + 0.017S$ is an ordinary least-squares regression line fitted under the assumption that the variance of the residuals is independent of S , and without constraints on the slope or intercept. The slope 0.017 has a standard deviation of 0.022. There is no evidence of a rise or fall in c with increasing S . In this and all subsequent figures, the plotted symbols have been perturbed by a small random amount from their exact positions to indicate when several symbols coincide. \times , constant web; \circ , fluctuating web. Only 'lumped' webs are used (Briand & Cohen 1984).

(lower) connectance C' , defined by $C' = 2L'/[S'(S' - 1)]$, decreases nearly but not quite as fast as $1/S'$. Now $C' = c(1/S')[(S - 1)/(S' - 1)]$ and the last factor $[(S - 1)/(S' - 1)]$ approximates 1. Therefore if c has no trend as a function of S , C' would be expected to decline approximately as $1/S'$. Secondly, by using 40 unlumped webs, including 13 of Cohen's, Briand (1983) observed that the number L' of links was proportional to $S'^{1.1}$ rather than to S' . Because of the overlaps among the sets of data used by Yodzis (1980), Briand (1983) and here, the findings of Yodzis (1980) and Briand (1983) are by no means independent of ours. Moreover, the webs analysed by Yodzis and Briand were unlumped while ours are lumped. Thus there is no persuasive evidence against the natural null hypothesis that for the lumped webs studied here, c is effectively constant as S increases.

Of the 19 constant webs, 14 fall above the regression line in figure 1. Of the 43 fluctuating webs, 34 fall below the regression line. The difference in the proportions of webs above the regression line (74% for constant webs, 21% for fluctuating webs) is too large to be attributed to chance ($\chi^2 = 15.7$ with one degree of freedom, a value with extremely low probability if one chooses to believe the underlying but doubtful assumption of independence among webs). For a given number of species, constant webs have more links than fluctuating webs (Briand 1983).

This difference demonstrates at the level of individual webs the aggregate difference in the ratio of links to species between constant and fluctuating webs. For constant webs, the ratio of links to species is $811/351 = 2.31$, while for fluctuating webs the ratio is $1108/683 = 1.62$.

The use of a single value for c in paper I overlooks differences in the typical values of c of two distinguishable kinds of webs, the constant and the fluctuating, making it all the more surprising that the aggregated predictions of the cascade model in paper I are not worse. In paper II, since c and S are estimated separately for each web, we are testing how the cascade model applies to individual webs, both constant and fluctuating.

We now test seven predictions of the cascade model. In figures 2–8 the abscissa is the expected value of some feature of a web, according to the cascade model, and the ordinate is the observed value of that feature. If the estimated values of c and S corresponded exactly to the true values of c and S and if the observed value of each feature in each web corresponded to the expected value, then all data points would fall along a line of slope one through the origin. The cascade model is a stochastic model, however, so the data points are expected to deviate from such a line, but not systematically. Since the scales of the abscissa and ordinate vary, a line of slope one through the origin is drawn in figures 2–8 for comparison.

There is no reason to assume that half of the data points should fall above, and half below, the line of slope one, because we have not proved that, according to the cascade model, the variables of interest are symmetrically distributed about their mean. However, as the number of species in a web increases, it seems reasonable to suppose that the distributions of these variables approach normality. In this limit of large S , it seems reasonable to anticipate roughly half of the data points above and half below the line of slope one if the cascade model is correct.

As might be expected, in figures 2–8 the variance of the observed number, plotted on the ordinate, increases as the expected number, plotted on the abscissa, increases. Since all the abscissae are increasing functions of the number of species in a web, the variances of observed numbers also increase with increasing size of web.

Figure 2 plots the observed (proper) basal species against the expected proper basal species, computed from (6.3a) in paper I. There appear to be ‘rows’ of data points in figure 2 because the observed numbers of basal species are constrained to be integers, while the expected numbers can vary continuously. The bulk of the data points, though by no means all, fall below the line of slope one. This finding is consistent with the fact that fewer basal than top species are observed and with the observation in paper I that fewer basal species are observed than expected using an aggregate estimate of c . No difference between constant and fluctuating webs in the success of the cascade model is immediately evident from figure 2. This absence of apparent difference is consistent with the finding of Briand & Cohen (1984) that the proportions of (proper) basal, intermediate, and (proper) top species are homogeneous between constant and fluctuating webs, within statistical fluctuations.

Figure 3 plots the observed intermediate species against the expected intermediate species, computed from (6.3b) in paper I. The constant webs fall nearly evenly above and below the line of slope one (9 fall above, 10 fall below). The bulk

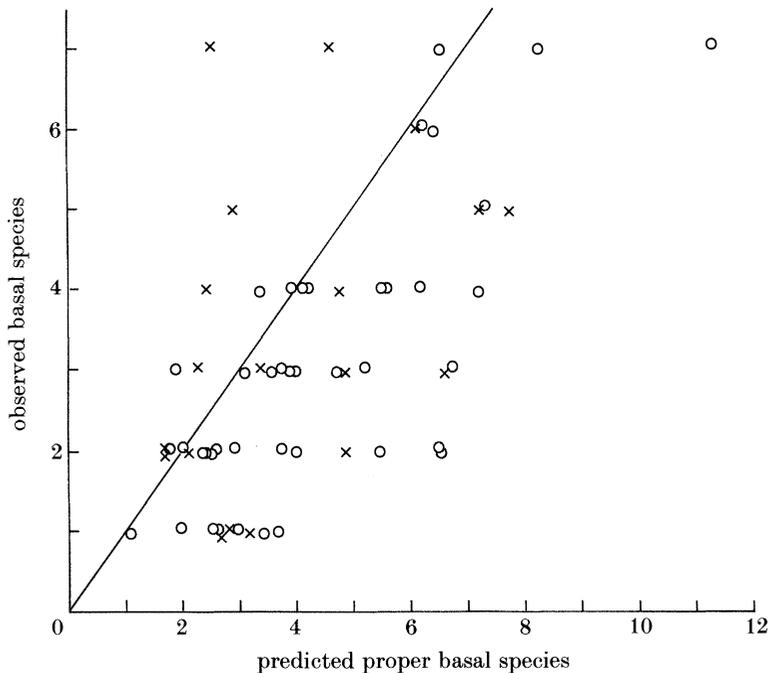


FIGURE 2. The observed number of (proper) basal species as a function of the expected number of proper basal species according to the cascade model. In this and all the following figures, the solid straight line passes through the origin with slope 1. This is not a regression line, but should describe the trend of the data if the predictions of the cascade model are approximately correct.

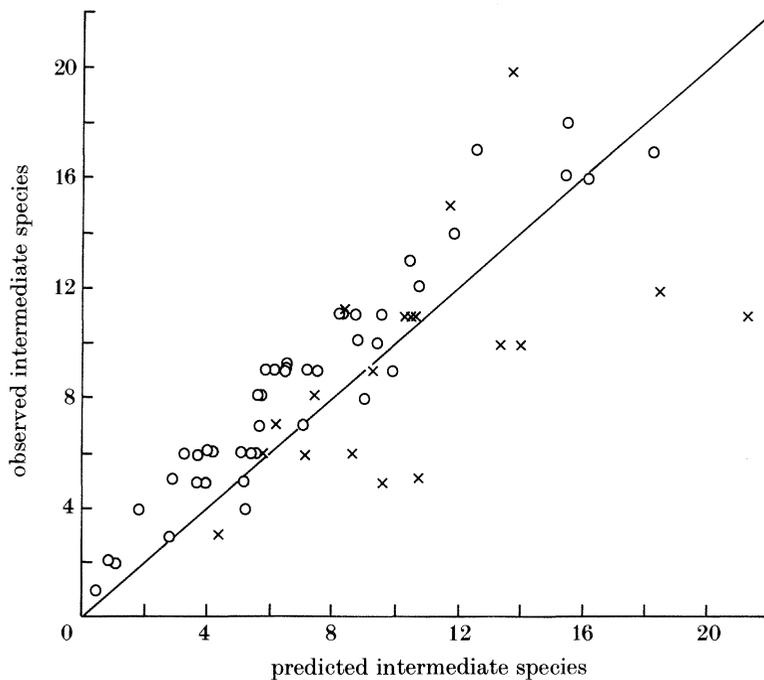


FIGURE 3. The observed number of intermediate species as a function of the expected number of intermediate species according to the cascade model.

of the fluctuating webs fall slightly above the line. This small difference is consistent with the insignificantly greater aggregate proportion of intermediate species in fluctuating webs than in constant webs (54 % versus 50 %). Overall, the agreement between observed and expected species is good.

Figure 4 plots the observed (proper) top species against the expected proper top species, computed from (6.3a) in paper I. The constant web with 17 top species appears to be an outlier. This same web, which describes the rocky shore of Lake Nyasa (Fryer 1959) also appeared as an outlier in a plot, with unlumped data, of prey against predators (Briand & Cohen 1984). For both constant and fluctuating webs, there is a suggestion that the remaining points may rise convexly. At least in the middle range of expected values, however, the agreement between observation and expectation is good.

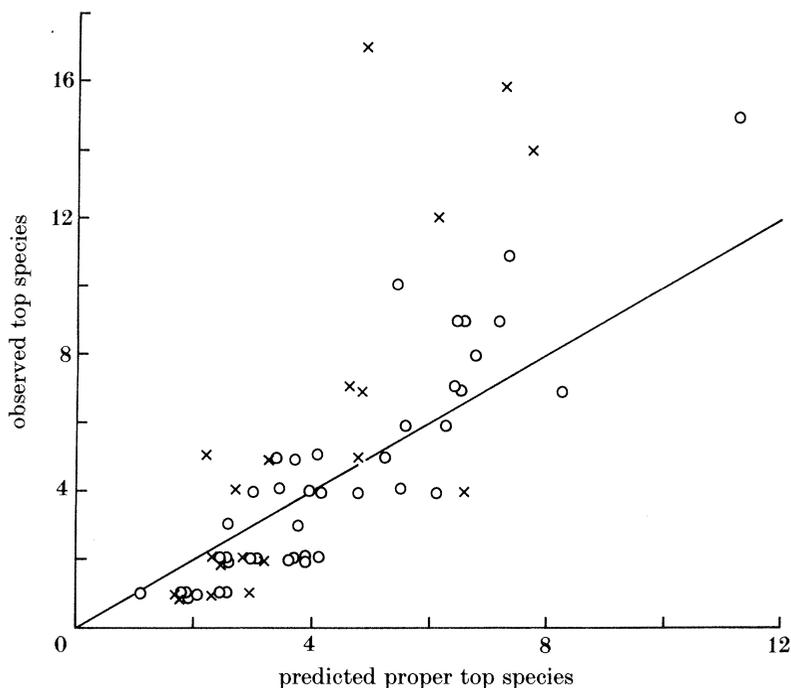


FIGURE 4. The observed number of (proper) top species as a function of the expected number of proper top species according to the cascade model.

In summary, when the expected numbers of species of each kind are compared with the observed, the agreement is best for intermediate species and is fair for proper top and proper basal species. The cascade model describes the kinds of species in constant and fluctuating webs about equally well.

Figure 5 plots the observed basal-intermediate links against the expected basal-intermediate links, computed from (6.6a) in paper I. There is no sign of systematic deviation between the points and the line of slope one, for the constant and fluctuating webs considered separately or together.

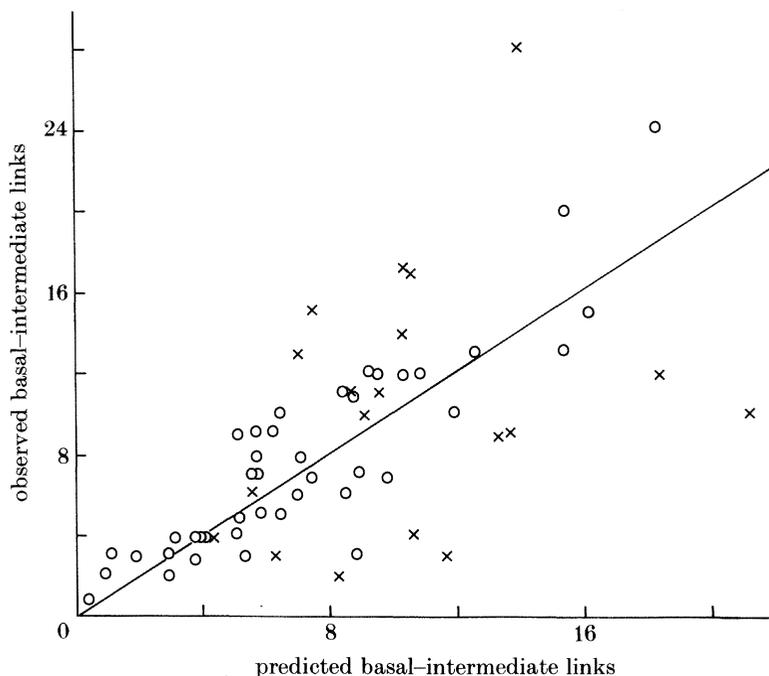


FIGURE 5. The observed number of basal-intermediate links as a function of the expected number of basal-intermediate links according to the cascade model.

Figure 6 plots the observed basal-top links against the expected basal-top links, computed from (6.6*b*) in paper I. Contrary to expectation, there are many webs with no basal-top links or one only. The line through the origin with slope one passes through the mass of the remaining points, but even for these the scatter about the line is large, compared with that in figures 5, 7 and 8.

Figure 7 plots the observed intermediate-intermediate links against the expected intermediate-intermediate links, computed from (6.6*c*) in paper I. The apparent outlier with six observed intermediate-intermediate links in the lower right corner of figure 7 is the same web that appears above as the potential outlier in figure 4. This same web appears again as the outlier with 59 observed intermediate-top links in the top right corner of figure 8. Clearly this web is exceptional in several respects, when compared with other webs. Aside from this outlier, the remaining webs are scattered more or less symmetrically about the line of slope one, and no systematic deviations are evident.

Figure 8 plots the observed intermediate-top links against the expected intermediate-top links, computed from (6.6*a*) in paper I. As in figure 5 (apart from the single outlier), there is no sign of systematic deviation of the points from the line of slope one, for the constant and fluctuating webs separately or together.

In summary, the cascade model provides a good description of the numbers of basal-intermediate, intermediate-intermediate, and intermediate-top links, aside from one outlying constant web, and a rather poor description of the number of basal-top links. The cascade model describes the links of constant and fluctuating webs about equally well.

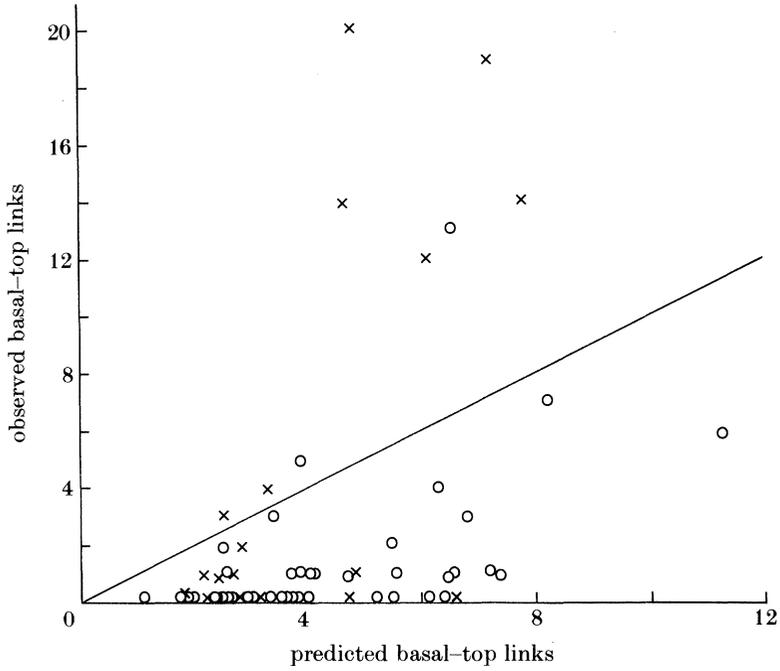


FIGURE 6. The observed number of basal-top links as a function of the expected number of basal-top links according to the cascade model.

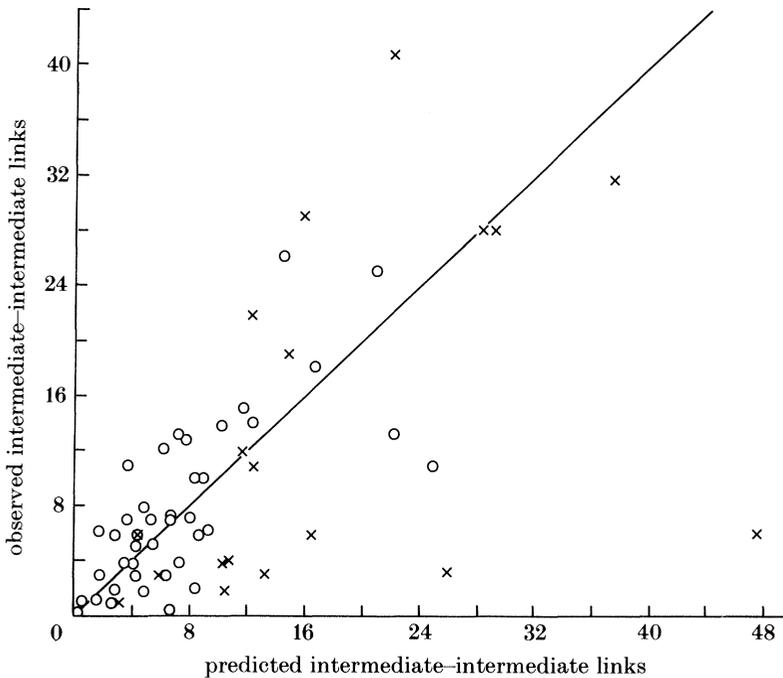


FIGURE 7. The observed number of intermediate-intermediate links as a function of the expected number of intermediate-intermediate links according to the cascade model.

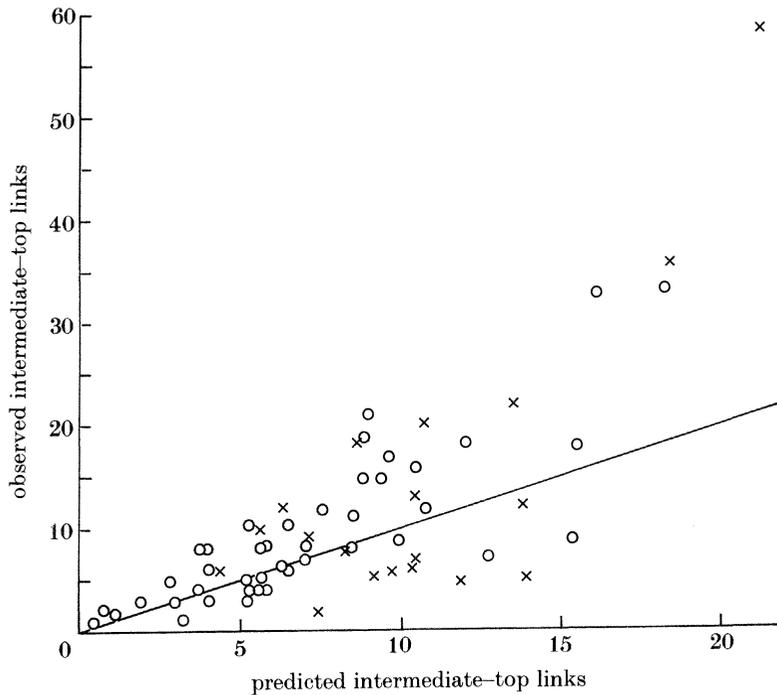


FIGURE 8. The observed number of intermediate-top links as a function of the expected number of intermediate-top links according to the cascade model.

4. DISCUSSION AND CONCLUSIONS

We have tested a model, called the cascade model, which assumes that species in a community are arranged in a hierarchy or cascade of potential feeding relations. This model assumes that whether a potential feeding relation becomes an actual feeding relation is determined randomly, independently of all other potential feeding relations. The probability that a potential feeding relation becomes actual is assumed to be the same for every potential feeding relation within a community, and to vary inversely as the number of species in the community.

Consequently, according to the model, for a randomly chosen species in a community, the mean number of other species that prey on it or that are prey to it is independent of the total number of species in the community. Thus the model is consistent with the hypothesis, suggested by Pimm (1982, p. 89), that 'each species in a community feeds on a number of species of prey that is independent of the total number of species in the community', provided that the term 'each species' is replaced by the term 'a randomly chosen species'.

In paper I we showed that the cascade model describes several important properties of webs, to a first approximation, when a fixed probability parameter, estimated from aggregated data, is applied to all webs. There is no logical necessity for the cascade model to describe individual webs, given that it succeeds reasonably in the macroscopic analysis. Tests of the cascade model using data on individual webs are logically and empirically independent of tests using aggregated data.

Indeed, it would be surprising to find that ecological 'assembly rules' as simple as the cascade model apply to communities that arise in diverse environments.

To test the cascade model's ability to describe individual webs, we used two numbers, the observed number of not isolated species and the observed number of links, to estimate the two parameters of the cascade model: the unknown true number S of species and the unknown constant c to which the probability of a feeding relation is proportional. We then computed the expected values of seven characteristics of webs and compared them with the observed.

In 62 webs, with the exception of an occasional outlier, the cascade model describes well the numbers of intermediate species (figure 3), basal–intermediate links (figure 5), intermediate–intermediate links (figure 7), and intermediate–top links (figure 8). It describes fairly the numbers of proper basal (figure 2) and proper top (figure 4) species. It describes poorly the numbers of basal–top links (figure 6).

For a given value of S , the probability c/S that a species will prey on another species, when their positions in the trophic hierarchy permit, is higher in constant webs than in fluctuating webs. Given c and S , the cascade model describes the numbers of kinds of species and kinds of links in constant and fluctuating webs about equally well, according to our examination of the data.

Cohen & Briand (1984) noted that the proportions of each kind of link appear to differ between constant and fluctuating webs. Since the cascade model describes the numbers of each kind of link about equally well in constant and fluctuating webs, the difference in proportions may be explained by the difference in the typical values of c for constant and fluctuating webs, rather than by some deeper structural difference between constant and fluctuating webs. The difference in the typical values of c between constant and fluctuating webs is not explained by the cascade model.

In testing the model, we present graphical comparisons of the observations and predictions so that the reader can make his or her own verbal summaries of how good or bad the fit is. We avoid formal statistical measures of goodness of fit because the data may not be independent and because we are interested in simultaneous inference about the model as a whole. The assumption of independence among webs appears doubtful, since some authors contributed more than one web and the proclivities of authors do appear to influence the structure of webs. We are less concerned to test hypotheses about any portion of the cascade model than to see how well, on the whole, it describes simultaneously several major features of webs, some of which may not be independent. (For example, given the total number of not isolated species, the observed numbers of proper top, intermediate, and proper basal species are not independent.)

How should one evaluate the discrepancies between theory and observation most evident in figure 6, and strongly suggested by figures 2 and 4? One can be sceptical about the model, but not the data; or one can be sceptical about the data, but not the model; or one can be sceptical about both. We are sceptical about both.

As for the data, the earlier papers of Briand and Cohen, jointly and separately, indicate that there is great variability among observers in the operational definitions of species and links and in the detail of published reports of field work. Often, these operational definitions are not even described in the published reports.

A first step that field ecologists could make toward improving the data would be to describe in detail how the species and links are operationally defined. A second step would be to work toward some uniform definitions.

Nevertheless, the data analysed here are the best available at present. The regularities in these data merit theoretical attention.

As for the model, numerous assumptions underlying it are unrealistic. Is it plausible to assume that the species at the top of the hierarchy or cascade is equally likely to prey on all other species in the community? Is it plausible to assume that the prey species a predator eats are chosen independently of the abundance of the prey species and stochastically, once and for all, as the model implicitly assumes? We think not.

Nevertheless, the cascade model provides a very simple unifying perspective, quantitatively testable and open to improvement, that explains for the first time several empirical regularities in the structure of webs. The ecological generalizations explained by the cascade model still need to be derived from a persuasive and testable theory of behaviour, population dynamics, and trophic interactions.

The cascade model also needs to be tested further against macroscopic data. Can the cascade model explain the observed frequency of intervality (Cohen 1978) in food webs? Can the cascade model explain the observed frequency distributions of length of food chains?

A referee made helpful comments. J. E. C. acknowledges with thanks the John D. and Catherine T. MacArthur Foundation Fellows Program, U.S. National Science Foundation grant BSR84-07461, and the hospitality of Mr and Mrs William T. Golden. C. M. N. acknowledges the support of U.S. National Science Foundation grant MCS80-19384. F. B. acknowledges the support of grant A9868 from the Natural Science and Engineering Research Council of Canada.

APPENDIX: ESTIMATION OF c AND S BY NEWTON'S METHOD

Given observed numbers of species S' and observed numbers of links L' , the parameter c and the true number of species S satisfy, according to model 3 (the cascade model):

$$0 = c(S-1)/2 - L' \equiv v_1, \quad (\text{A } 1 \text{ } a)$$

$$0 = (1 - [1 - c/S]^{S-1})S - S' \equiv v_2. \quad (\text{A } 1 \text{ } b)$$

We find c and S numerically by Newton's method (for example, Rektorys 1969, pp. 1180-1181).

Step 1. Let $c_0 = 2L'/(S' - 1)$, $S_0 = S'$. (These are initial estimates.)

Step 2. Compute

$$J_{11} = (S_0 - 1)/2,$$

$$J_{12} = c_0/2,$$

$$J_{21} = (S_0 - 1)(1 - c_0/S_0)^{S_0-2},$$

$$J_{22} = 1 - (1 - c_0/S_0)^{S_0-1} [1 + c_0(S_0 - 1)/(S_0 - c_0) + S_0 \lg(1 - c_0/S_0)].$$

(This is the Jacobian of the nonlinear system (A 1), that is, $J_{11} = \partial v_1 / \partial c$, etc., evaluated at (c_0, S_0) .)

Step 3. Compute $A = J^{-1}$, that is,

$$A = \begin{pmatrix} J_{22} & -J_{12} \\ -J_{21} & J_{11} \end{pmatrix} / (J_{11}J_{22} - J_{12}J_{21})$$

and v_1, v_2 from (A 1) with c replaced by c_0 , S by S_0 , then

$$c_1 = c_0 - a_{11}v_1 - a_{12}v_2, \quad S_1 = S_0 - a_{21}v_1 - a_{22}v_2.$$

(The new estimates of c and S are c_1, S_1 .)

Step 4. Stop if $|c_1 - c_0| + |S_0 - S_1| < \delta = 0.01$. (Stop when the procedure quasi-converges. The value of $\delta = 0.01$ was chosen to require the final estimates c_1, S_1 to be changing far less than the uncertainty in the data.)

Step 5. Otherwise, replace the value of c_0 by the value of c_1 and replace the value of S_0 by the value of S_1 . Then go to step 2. (Iterate with improved estimates.)

When applied to the 62 pairs (S', L') from the webs assembled by Briand, this procedure stopped after at most 3, and generally 2, iterations, except for the pair (S', L') = (3, 2) from one web (code number 10). For this web, the procedure diverged, and we used the initial estimate.

REFERENCES

- Briand, F. 1983 Environmental control of food web structure. *Ecology* **64**, 253–263.
- Briand, F. & Cohen, J. E. 1984 Community food webs have scale-invariant structure. *Nature, Lond.* **307**, 264–266.
- Cohen, J. E. 1978 *Food webs and niche space*. Princeton: Princeton University Press.
- Cohen, J. E. & Briand, F. 1984 Trophic links of community food webs. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4105–4109.
- Cohen, J. E. & Newman, C. M. 1985 A stochastic theory of community food webs. I. Models and aggregated data. *Proc. R. Soc. Lond. B* **224**, 421–448.
- Fryer, G. 1959 The trophic interrelationships and ecology of some littoral communities of Lake Nyasa with especial reference to the fishes, and a discussion of the evolution of a group of rock-frequenting Cichlidae. *Proc. Zool. Soc. Lond.* **132**, 153–281.
- Paviour-Smith, K. 1956 The biotic community of a salt meadow in New Zealand. *Trans. R. Soc. N.Z.* **83**, 525–554.
- Pimm, S. 1982 *Food webs*. London: Chapman and Hall.
- Rektorys, K. 1969 *Survey of applicable mathematics*. Cambridge: M.I.T. Press.
- Yodzis, P. 1980 The connectance of real ecosystems. *Nature, Lond.* **284**, 544–545.