

The probabilistic niche model reveals substantial variation in the niche structure of empirical food webs

RICHARD J. WILLIAMS¹ AND DREW W. PURVES

Microsoft Research, 7 J J Thomson Avenue, Cambridge CB3 0FB United Kingdom

Abstract. The structure of food webs, complex networks of interspecies feeding interactions, plays a crucial role in ecosystem resilience and function, and understanding food web structure remains a central problem in ecology. Previous studies have shown that key features of empirical food webs can be reproduced by low-dimensional “niche” models. Here we examine the form and variability of food web niche structure by fitting a probabilistic niche model to 37 empirical food webs, a much larger number of food webs than used in previous studies. The model relaxes previous assumptions about parameter distributions and hierarchy and returns parameter estimates for each species in each web. The model significantly outperforms previous niche model variants and also performs well for several webs where a body-size-based niche model performs poorly, implying that traits other than body size are important in structuring these webs’ niche space. Parameter estimates frequently violate previous models’ assumptions: in 19 of 37 webs, parameter values are not significantly hierarchical, 32 of 37 webs have nonuniform niche value distributions, and 15 of 37 webs lack a correlation between niche width and niche position. Extending the model to a two-dimensional niche space yields networks with a mixture of one- and two-dimensional niches and provides a significantly better fit for webs with a large number of species and links. These results confirm that food webs are strongly niche-structured but reveal substantial variation in the form of the niche structuring, a result with fundamental implications for ecosystem resilience and function.

Key words: *ecological network; food web; hierarchy; likelihood; niche; probabilistic model; trophic structure.*

INTRODUCTION

Food webs, networks of feeding interactions between species in an ecosystem, have long been studied as a tractable yet informative simplification of the many feeding interactions between individual organisms in an ecosystem. The structure of food webs plays a crucial role in ecosystem stability (May 1972, Neutel et al. 2002, Emmerson and Raffaelli 2004, Beninca et al. 2008), robustness (Dunne et al. 2002b, Brose et al. 2006) and function (Petchey et al. 1999, Thébault and Loreau 2003). Understanding the structure of food webs remains a crucially important problem for both ecology and our understanding of other complex systems (May et al. 2008).

Early food web studies concentrated on finding structural regularities across empirical systems (Dunne 2006). A wide range of approaches have been used to study the mechanisms giving rise to these structural regularities. Valuable insights have been derived from studies of the stability of dynamical models (May 1972, Pimm and Lawton 1977, 1978, Yodzis 1981, Neutel et al. 2002) and empirical systems (Beninca et al. 2008),

models coupling evolutionary and population dynamics (Drossel et al. 2004, Loeuille and Loreau 2005), models of food web assembly (Morton et al. 1996), and studies of the effects of body size on the stability and persistence of species in dynamical food web models (Emmerson and Raffaelli 2004, Brose et al. 2006).

While many mechanisms and methods for explaining food web structure have been studied, simple stochastic models based on a few principles have provided surprisingly successful models of the presence and absence of links between species in complex food webs. The most successful models of this type are the niche model (Williams and Martinez 2000) and its variants (Stouffer et al. 2006, Allesina et al. 2008, Williams and Martinez 2008), and the groups model (Allesina and Pascual 2009). The allometric diet breath model (Petchey et al. 2008) grounds a body-size-based niche-structured model in individual feeding behavior, an important step toward a more mechanistic understanding of the diet selection processes which shape food webs.

The various niche-structured models assign each species a position (n_i), a feeding position (c_i), and a feeding range (r_i) along a one-dimensional niche axis (Hutchinson 1959, Cohen 1977), such that each species has high probability of eating only those species that fall within its feeding range. To generate niche-structured

Manuscript received 3 February 2011; revised 7 April 2011; accepted 13 April 2011. Corresponding Editor: J. Huisman.

¹ E-mail: ricw@microsoft.com

webs, the niche model and its variants are coupled with additional assumptions about the exact form of the niche structuring. These typically include a hierarchy such that all prey or all diet centers are above (or below) species niche positions (Cohen and Newman 1985, Cohen et al. 1990), uniform distributions of species positions along the niche axis and feeding positions within their possible values, and a beta distribution for normalized feeding ranges (r_i/n_i ; Williams and Martinez 2000). This last assumption results in generalist species (species with large feeding ranges r_i) tending to fall at high niche positions n_i .

Here we use a methodology based on maximum likelihood methods to parameterize a niche model directly against food web data. This probabilistic niche model (PNM) produces a maximum likelihood estimate of n_i , c_i , and r_i for each species in the web (Williams et al. 2010). In addition to testing overall model fit, the model allows us to test the various assumptions about parameter distributions and relationships between parameters made in previous niche-structured food web models. The model allows us to relax these assumptions and instead have the parameter distributions and degree of hierarchy follow directly from the data themselves. We test this niche-structured model against a much wider range of data than previously used when testing structural food web models, including data sets that are very poorly predicted by previous models.

The logic of the niche model is extended easily into multiple dimensions (Warren 1996, Allesina et al. 2008), but to date detailed studies have used only one dimension. In part, this may reflect the methodology used to date, which made it necessary to specify parameter distributions and make other assumptions. The total number of assumptions is substantial even for the one-dimensional model, and would at least double for a two-dimensional model. In contrast, a multi-dimensional probabilistic niche model simply has a higher-dimensional parameter space which has to be searched for optimal parameter values. This allows us to explore when the additional complexity of a two-dimensional niche space yields a significantly better model.

In summary, we use a likelihood-based probabilistic niche model to ask four questions: (1) how well does an optimally parameterized niche model perform; (2) are previously assumed parameter distributions and the previous assumption of hierarchy correct for most webs; (3) how much web-to-web variation is there in the parameter distributions and the degree of hierarchy; and (4) how much is niche-structured model performance improved by adding a second dimension to the niche space?

METHODS

The probabilistic niche model (PNM; Williams et al. 2010) assumes that a predator i 's probability of consuming a prey item j is a continuous function of

the mismatch between the prey species position in (potentially multidimensional) niche space n_j , and the predator's diet location c_i in the same space. The D -dimensional PNM (where D is the total number of dimensions) uses an exponential function with a parameterized cutoff rate for the probability that species i eats species j :

$$P(i, j | \theta) = \alpha \prod_{d=1}^D \exp\left(-\left|\frac{n_{d,j} - c_{d,i}}{r_{d,i}/2}\right|^e\right) \quad (1)$$

where $P(i, j | \theta)$ is the probability that species i eats species j given a particular parameter set θ where $\theta = \{n_{1,1} \dots n_{D,S}, c_{1,1} \dots c_{D,S}, r_{1,1} \dots r_{D,S}, e\}$; the parameter $n_{d,j}$ is the niche position of species j in dimension d ; the parameter $c_{d,i}$ is the optimal feeding position of species i in dimension d ; the parameter $r_{d,i}$ is the feeding range of species i in dimension d ; the parameter e varies the cutoff rate of the niche probability function (larger values of the exponent mean that the niche probability function is flatter at the center of the niche and cuts off more quickly at the edges of the niche), and the parameter α is the probability that i eats j when j is exactly on i 's feeding optimum (i.e., when $n_{1,j} \dots n_{D,j} = c_{1,i} \dots c_{D,i}$).

There is always a nonzero probability that any predator eats any prey (which is necessary in order to employ likelihood-based statistics). For each species, high feeding probabilities occur in a single contiguous region in the niche space. Each species i has D parameters defining its location in the D -dimensional niche space, $n_{d,i}$ where d is the dimension. Each species has $2D$ additional parameters defining its D -dimensional diet location $c_{d,i}$, and the feeding range $r_{d,i}$ which determines how quickly the feeding probability declines as a function of the distance between the prey location and predator diet location. In this way, the PNM naturally includes some realities present in real food webs which were absent from the original niche model, but which have been introduced recently to create variants of the original. For example, the generalized niche model (Stouffer et al. 2006), minimum potential niche model (Allesina et al. 2008) and relaxed niche model (Williams and Martinez 2008) allow predators to have a non-interval set of prey (Cohen 1977, 1978, Williams and Martinez 2000, Cattin et al. 2004, Stouffer et al. 2006). Non-intervality occurs naturally in the PNM because it is probabilistic, such that there are random "holes" in predators' diets; but these holes occur more often toward the edge of an otherwise interval diet. In addition, because in the PNM parameters follow entirely from food web data, it allows for the species to clump into functional groups that consume and/or are consumed by the same set of species, as stressed by the groups model (Allesina and Pascual 2009).

We parameterized the PNM for each of 37 food webs ranging in size S from 18 to 155 trophic species (Cohen et al. 1990) and with 48 to 1509 links. Data sets were restricted to food webs that are non-interval (all links in

interval food webs are trivially predicted by the probabilistic niche model) and to webs with 25 or more trophic species except for three smaller webs that were used in the study of the allometric diet breadth model (Petchey et al. 2008). The original niche model analyzed seven food webs and the minimum potential niche model analyzed 10 food webs. We have included a number of food webs not analyzed in these earlier studies, which have similar levels of taxonomic resolution as the previously studied data sets. Details of the data sets are given in Appendix A: Table A.1.

For each food web, we estimated the maximum likelihood parameter set, from which we computed the probability of each link according to the model. The model was initialized by placing the $n_{d,i}$, $c_{d,i}$, and $r_{d,i}$ of the species randomly in the niche space between 0 and 1. We then use simulated annealing (Kirkpatrick et al. 1983) to find the maximum likelihood parameter set given the observed feeding relationships. The values of $n_{d,i}$ and $c_{d,i}$ were constrained to remain between 0 and 1 while the $r_{d,i}$ were constrained to be between 10^{-8} and 10. Defining \mathbf{X} as the connection matrix for an S -species food web containing an observation X_{ij} for each link i, j ($X_{ij} = 1$ means i eats j ; $X_{ij} = 0$ means i does not eat j), the log-likelihood is defined as

$$\ell(\mathbf{X}|\theta) = \sum_i \sum_j \ln \left\{ \begin{array}{ll} P(i, j|\theta) & \text{if } X_{ij} = 1 \\ 1 - P(i, j|\theta) & \text{if } X_{ij} = 0 \end{array} \right\}.$$

In each food web, the expected total number of links given by the maximum likelihood estimate (MLE) parameter set was very close to the observed number (mean fraction = 0.993, SD = 0.0076), and so the expected fraction of links predicted correctly ($f_L = \text{expected number of links predicted correctly}/\text{total number of links in empirical food web}$) provides an easily understood, scale-independent measure of the overall performance of the model (Petchey et al. 2008). We also report model likelihoods. The performance of models with different numbers of parameters was evaluated by the Akaike information criterion (AIC; Akaike 1974), recognizing that AIC values might suffer from as yet unsolved problems with overfitting (Allesina 2010). These problems are partially corrected by AIC_c (Sugiura 1978, Hurvich and Tsai 1989), which we use when comparing the one and two dimensional PNM to see whether the improved fit to data afforded by the extra dimension justifies the additional parameters.

In multidimensional models, the niche width in one of the dimensions sometimes becomes very large. This means that in that dimension, the niche is very broad and so no information about a prey's niche value $n_{d,j}$ is needed to determine the predator's diet. Knowing this, the predator's $r_{d,i}$ and $c_{d,i}$ can be set to a large value and the center of the niche range respectively, and that dimension's contribution to the probability is approximately 1 for all values of $n_{d,i}$. This then reduces the number of free parameters in the model by 2.

RESULTS

Across the 37 food webs, the expected fraction of links predicted correctly by the one-dimensional PNM ranged from $f_L = 0.502$ to $f_L = 0.971$ with an average of 0.678. For the two-dimensional PNM, f_L ranged from 0.590 to 0.999, with an average of 0.804. Fig. 1 plots empirical links and link probabilities of the one-dimensional PNM for two of the food webs to illustrate some patterns found generally across the set of food webs studied here. (Similar plots of all 37 food webs are in Appendix B.) The Skipwith Pond web (Fig. 1a) is a small ($S = 25$), high connectance ($C = 0.315$) web that is quite well predicted by the one-dimensional PNM ($f_L = 0.86$). The St. Martin Island web (Fig. 1b) is a larger web ($S = 42$, $C = 0.116$) with $f_L = 0.67$. In these empirical webs, the niches of generalists have high-probability, near-contiguous "cores" and lower-probability edges; empirically observed links with low probability of prediction tend to be in these edge regions rather than randomly placed either within, or a long way from, the contiguous core of the diet. These patterns recur throughout the empirical data sets studied (see Appendix B). Maximum likelihood parameter estimates and link probabilities for all webs are in the Supplement.

In Fig. 2a, we compare the performance of the PNM against the deterministic allometric diet breadth model (ADBM) based on the fraction of links predicted correctly. In Fig. 2b, we compare the performance of the PNM against the minimum potential niche model (MPNM) and groups model using our AIC estimates of the PNM vs. previously published AIC values of the other models (Allesina et al. 2008, Allesina and Pascual 2009). Data are shown in Appendix A: Tables A.2 and A.3; Table A.2 also includes AIC values from a probabilistic version of the allometric diet breadth model (Allesina 2010). In all cases, the PNM consistently performs better than any of these recent models.

Fig. 3 shows that the fraction of links predicted correctly (f_L) for the one-dimensional PNM is correlated significantly with both the number of species (S ; Fig. 3a) and number of links (L) in the network (Fig. 3b; $r^2 = 0.69$, $P = 1.8 \times 10^{-6}$ and $r^2 = 0.51$, $P = 0.0012$, respectively). These is also a suggestion in Fig. 3b that goodness of fit does not continue to decline as rapidly with number of links once networks have about 400 links, although the number of data sets with a large number of feeding links is small. This overall pattern may result in part from the fact that the ratio of parameters to data decreases with increasing S , but we also find substantial variation in f_L for a given S . For example, circled in Fig. 3a are two relatively species-rich webs ($S = 61$ and 85) that are very well fit ($f_L = 0.90$ and 0.89 of links predicted correctly, respectively). These webs are the UK Grassland and Broom webs, two low connectance ($C = 0.026$ and 0.031, respectively) parasitoid webs.

Rather than assuming hierarchical ordering, we measured the degree of hierarchical ordering of each

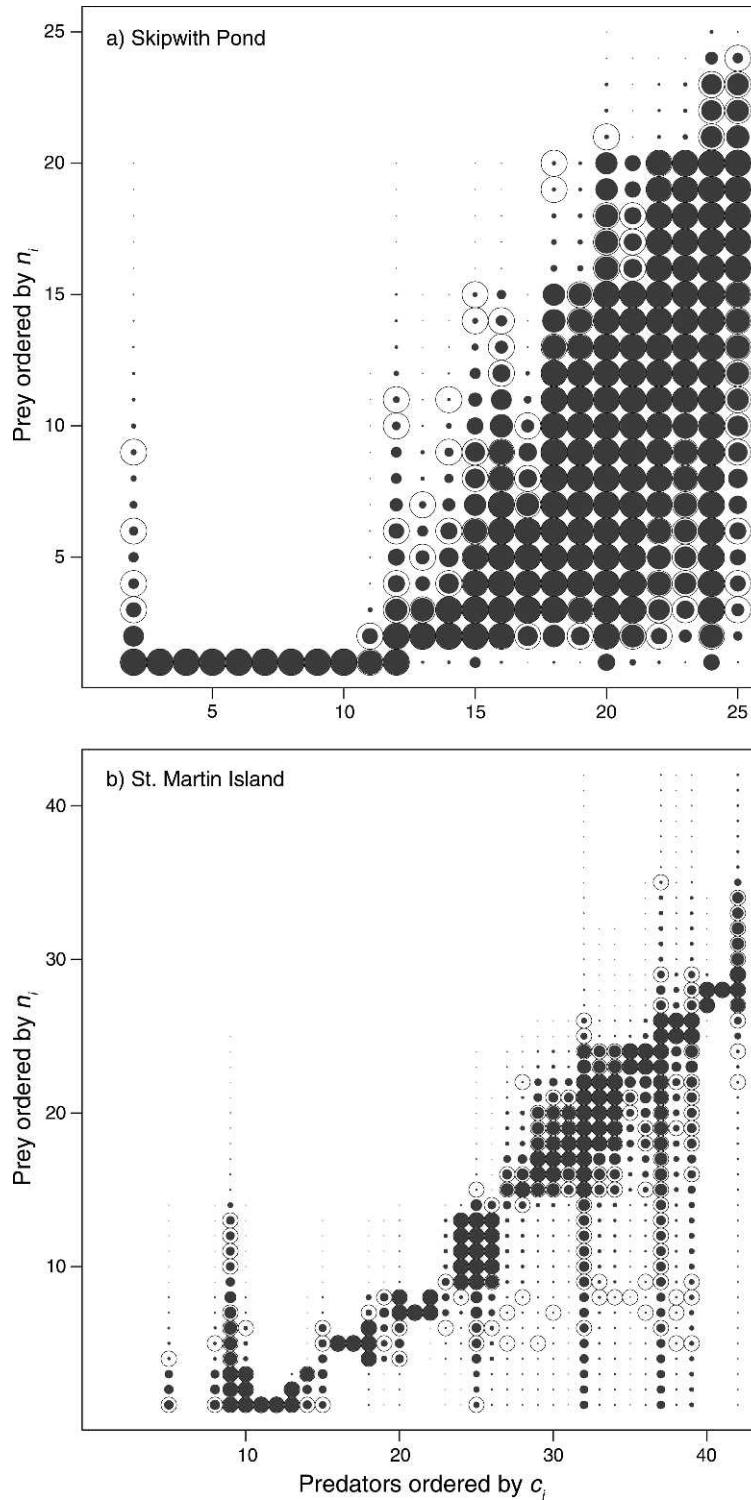


FIG. 1. Link probabilities predicted by the probabilistic niche model (solid gray circles, diameter is proportional to probability) and empirically observed links (rings around circles) for the (a) Skipwith Pond and (b) St. Martin Island food webs. Each prey species has a position (n_i), and each predator species has a feeding position (c_i).

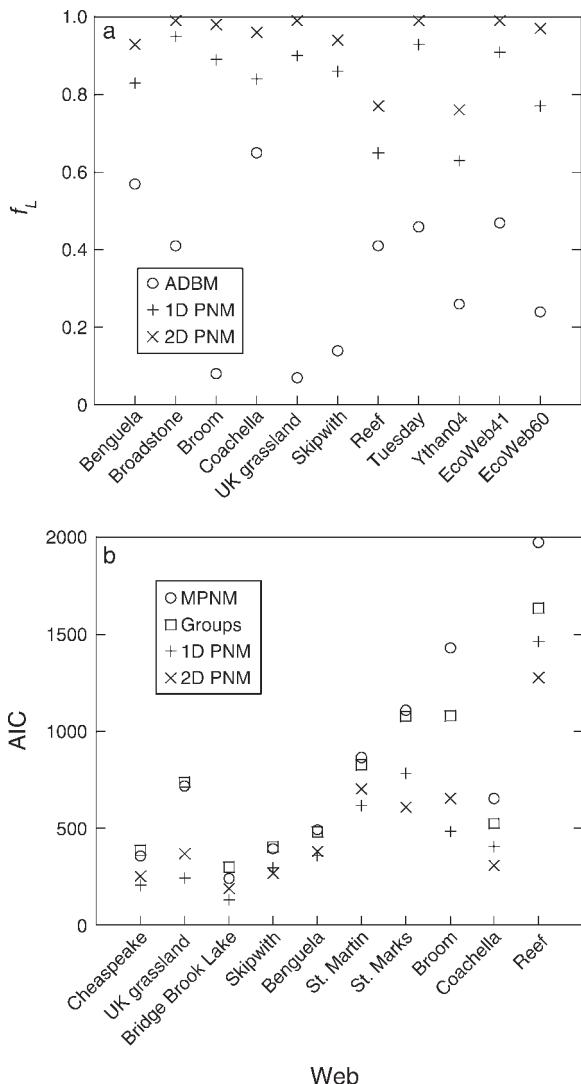


FIG. 2. Comparison of the performance of the one-dimensional probabilistic niche model (1D PNM) and two-dimensional PNM (2D PNM) against (a) the allometric diet breadth model (ADBM; Petchey et al. 2008), model performance measured by the fraction of links predicted correctly ($f_L = [\text{expected number of links predicted correctly}]/[\text{total number of links in empirical food web}]$), and (b) the minimum potential niche model (MPNM; Allesina et al. 2008) and groups model (Allesina and Pascual 2009), model performance measured by Aikake information criteria (AIC; better performing models have lower values). $AIC_{MPNM} = 2(-\ell + S + 3)$; $AIC_{Groups} = 2(-\ell + S + k^2)$ where ℓ is model likelihood and k is the number of groups.

web as h , the fraction of consumers with $c_i < n_i$ in the MLE parameter set (if this fraction is less than 0.5, the hierarchy is reversed and h is replaced with $1 - h$). We found that 10 webs (27%) are completely hierarchically ordered, while 27 webs (73%) deviate from the previous assumption of hierarchical ordering and have at least one species that is not hierarchically ordered. Of these 27 webs, some had no significant hierarchy compared to

random: 7 of 37 (19%) have $P > 0.05$ and 12 of 37 (32%) have $P > 0.01$ that they are unordered (tested against the null expectation of a binomial distribution with $k = hS_C$, $n = S_C$ and $P = 0.5$, where S_C is the number of consumers). The seven strongly nonhierarchical food webs range widely in size from 29 to 106 species but all are fairly low connectance ($C_{max} = 0.073$). The nonhierarchical webs also range widely in how well the nonhierarchical model predicts the links in the food web, from $f_L = 0.53$ for the Troy food web to $f_L = 0.89$ for the Broom food web.

We also examined the distribution of the estimates of niche position n_j against the assumption used in the various niche-structured models that these are uniformly distributed over the niche axis. We found that niche values n_i were only occasionally uniformly distributed. Using a K-S test of uniform distribution, the n_i distributions of 32 of 37 webs (86%) deviated from the

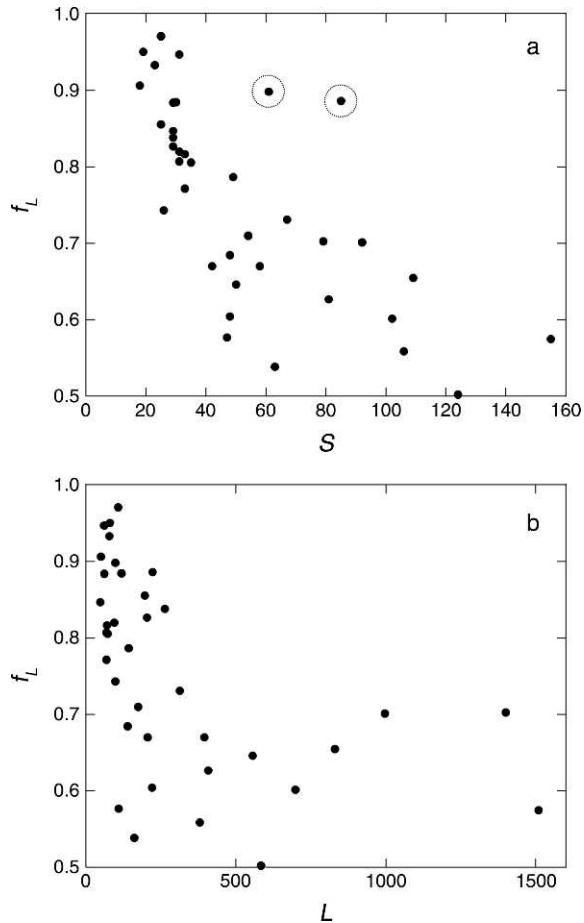


FIG. 3. Expected fraction of links correctly predicted by the one-dimensional probabilistic niche model (f_L) vs. (a) species richness (S) and (b) number of links (L). Each data point represents a different food web (see Table S.1 in the Supplement). The circled points in panel (a) are the U.K. Grassland and Broom webs, two low-connectance ($C = 0.026$ and 0.031 , respectively) parasitoid webs.

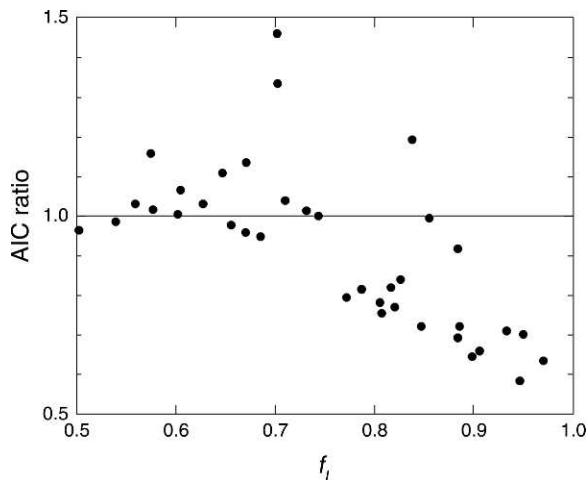


FIG. 4. AIC_{c1D}/AIC_{c2D} (AIC ratio) vs. expected fraction of links correctly predicted by the 1D model (f_L). AIC ratio > 1 means that the 2D model performed better even when taking into account its increased number of parameters. Each data point represents a different food web (see Table S.1 in the Supplement).

assumptions of the niche model and were significantly nonuniform at $P < 0.05$, and 27 (73%) were nonuniform at $P < 0.01$. It is possible to constrain the PNM to force the niche values to be approximately uniformly distributed, but this gave consistently worse (lower likelihood) solutions (not shown). This shows that in the majority of food webs, nonuniform distributions of n_i are necessary for finding an optimal solution.

The various niche-structured models assume that generality, as expressed by diet widths r_i , is strongly correlated with niche position n_i , with species with higher n_i tending to be generalists. We tested whether n_i and r_i are correlated in the PNM maximum likelihood parameter sets and did not find a significant correlation (Spearman rank correlation; $P > 0.05$), and therefore deviation from the assumptions of the niche model, in 15 of 37 webs (41%).

Following earlier work, we quantified the taxonomic resolution of each food web as the fraction of taxa resolved to the level of genus or species (Dunne et al. 2002a). We did not find any significant relationship between taxonomic resolution and fraction of links predicted correctly by the PNM. There was also no significant relationship between taxonomic resolution and the level of hierarchy h . We also tested whether the fit of the PNM was dependent on habitat type, coarsely defined as terrestrial, aquatic, or marine, but again found no significant relationship.

We found that those networks that were better predicted by the one-dimensional model rarely benefited from the addition of a second niche dimension, while the worse-predicted networks, those with $f_L < 0.75$ for the one-dimensional model, almost all benefited significantly from the addition of a second niche dimension (Fig. 4). The networks that were best predicted by the one-

dimensional model tended to have smaller S and L (Fig. 3), and the benefit of an additional dimension increases with both S and L . Further, the AIC ratio for the two-dimensional vs. one-dimensional model was more strongly correlated with the number of links L ($r^2 = 0.64$, $P = 1.7 \times 10^{-5}$), than with the number of species S ($r^2 = 0.40$, $P = 0.014$).

The two-dimensional niche space allows for variation in the dimensionality of the species diets; when $r_{d,i}$ is large the link probability does not vary significantly with $n_{d,i}$ and so that dimension is not constraining the diet. This means that each species' diet is free to be constrained by both, either, or neither of the two dimensions. Considering only predators with two or more prey (predators with 1 prey require no information about feeding range and so all have very small values for both $r_{d,i}$) and considering all 37 food webs, we found that 70% of species with more than one prey have strongly two-dimensional diets, while the remaining species' diets are effectively one-dimensional. The fraction of species with two-dimensional niches is strongly correlated with both S and L , with a much higher fraction of two-dimensional niches in webs with high S and L (Fig. 5).

DISCUSSION

The overall fit of the one and two-dimensional PNMs is better than that of any of the recent simple structural food web models (Fig. 2). In part, this improved fit reflects the fact that within the maximum likelihood framework we have removed previous constraints on parameter values and are free to actively select the particular set of n , c , and r parameters that provide the best fit to a given set of data. It also reflects the fact that niches in the PNM are structured to have high feeding probabilities in the center of the feeding range and lower probabilities toward the edge of the feeding range. Overall, the results confirm the original niche model's proposition that the majority of observed links can be reproduced by a model assuming that species both occupy a position in a low-dimensional niche space and preferentially consume species from a certain region of that same space.

The results do not support the long-used assumption of a feeding hierarchy, or the assumed distributions of n , c , and r used previously in the niche model and its variants. We tested whether the maximum likelihood parameter sets found here followed three constraints on parameter values used in the various niche-structured models. Across the 37 food webs, only one had uniformly distributed n_i , completely hierarchical feeding relationships and feeding range (r_i) significantly correlated with n_i . This food web was the 19 trophic species Broadstone Stream food web, the smallest food web tested here. These results show that there is much more variability in the structure of food web niche space than assumed by previous models. The significant improvement of the PNM over previous models is in large part

because it is not encumbered with these unnecessary constraints.

Our results confirm that many species diets are non-contiguous, as stressed by the minimum potential niche, generalized niche, and relaxed niche models, but not the idea that the holes are randomly strewn throughout the diet, as assumed in earlier models (Allesina et al. 2008, Williams and Martinez 2008). Neither are prey outside a central contiguous niche randomly distributed among other potential prey (Stouffer et al. 2006). The PNM instead assumes that holes in the niche occur mostly toward the edge of a narrowly constrained diet, and comparison with the distribution of links in the empirical data support this assumption (Fig. 1).

There has been frequent speculation that the niche axis is closely related to or corresponds exactly to species' body sizes (Williams and Martinez 2000, Woodward et al. 2005, Stouffer et al. 2007, Allesina et al. 2008). The niche-structured allometric diet breadth model (ADBM; Petchey et al. 2008) positions species on the niche axis according to their body size and determines contiguous diet ranges and positions using optimal foraging theory. The model's performance is quite variable, indicating that traits other than body size are important in structuring some food webs. Comparison of the PNM's performance with food webs previously analyzed using the ADBM (Fig. 5a) shows that food webs that are poorly predicted by the body-size-structured model can still be well predicted by a low-dimensional niche-structured model, and even those webs best predicted by the ADBM can be better predicted by an unconstrained probabilistic niche model. In addition, a body-size-structured PNM has been shown to perform significantly worse than an unconstrained PNM on a marine food web (Williams et al. 2010). Together, these results suggest that it is likely that species positions and feeding traits in the one- or two-dimensional niche spaces generally correspond to a complex variety of physical and behavioral traits, with body size often playing an important role, and other traits generally needing to be considered to understand species' diet choices. It is likely that different traits are more or less important in different systems. For example, while body size is generally important in producing non-random levels of intervality in feeding niches (Zook et al. 2011), it is probably more important in structuring the niche space of aquatic systems than in terrestrial systems (Shurin et al. 2006).

We have analyzed a number of food webs not previously used in studies of simple structural food web models, including several with a very high fraction of basal species. The constraints on parameter values in the original niche model cause it to have very poor performance on these food webs, while the freedom of parameter choice in the PNM allows it to find parameter sets for these food webs that give similar performance to the other food webs tested. This shows that these food webs are strongly niche structured and the species and

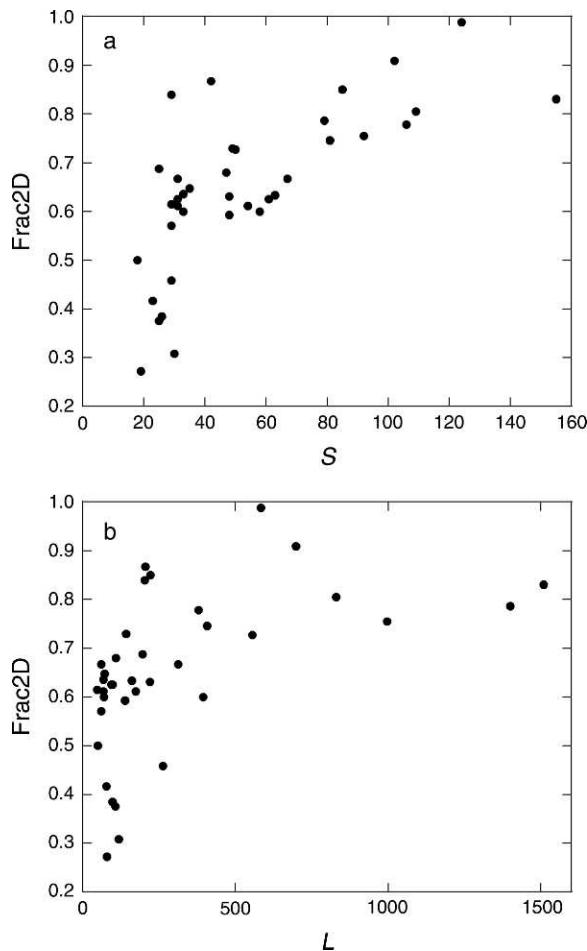


FIG. 5. Fraction of species with fully two-dimensional niches (Frac2D) vs. (a) species richness S and (b) number of links L . Each data point represents a different food web (see Table S.1 in the Supplement).

diets can be arranged in a low-dimensional niche space such that the observed high fraction of basal species is reproduced.

Strongly niche-structured webs, with overall fit to the PNM controlled by network scale, were found irrespective of taxonomic resolution or habitat type. The uneven taxonomic resolution and aggregation of food web data sets is a well-known problem (Cohen et al. 1993), but similar to previous studies (Stouffer et al. 2007), our results are robust to differences in resolution and aggregation across the data sets we studied. Similarly, we don't find any consistent trends in the performance of the PNM on food webs from different habitats, confirming earlier work that found a lack of empirical evidence for hypothesized differences between terrestrial and aquatic food webs (Chase 2000) and that found differences between marine and other food webs were largely due to variations in their size and connectance (Dunne et al. 2004).

Previous work has noted considerable scale dependence in the performance of niche-structured food web models (Williams and Martinez 2008), and our results also clearly show that models with a single niche dimension perform better on less species rich and less highly connected food webs. In addition, as webs become more diverse and more complexly connected, the overall benefit of adding a second niche dimension becomes greater. We also found that while the number of additional parameters in the two-dimensional PNM scales with the number of species, the improved performance of the two-dimensional PNM was more highly correlated with the number of links than the number of species.

The estimates of n , c , and r for each species could form the basis of additional studies. As noted above, many studies have highlighted the important role of body size in structuring food webs and in mediating the dynamics of energy flow through food webs and it has frequently been suggested that the niche axis in the various niche-structured food web models corresponds closely to species body size. Similarly, phylogenetic similarity has been proposed as being important in structuring feeding interactions (Cattin et al. 2004). Future work will consider the relationship between these and other traits and the abstract niche traits used here.

Although the results confirm that smaller food webs can be explained in terms of single-dimensional niche spaces, they show that in larger and more complex food webs, a two-dimensional niche space, with species having a mix of one- and two-dimensional diets, gives a significantly better model. It is becoming clear that food web structure and niche partitioning are critically important in establishing ecosystem functioning (Cardinale et al. 2009). The web-to-web variation in food web niche structure revealed here likely reflects important variation in community dynamics and biogeochemical functioning, despite the essential niche structuring that all webs share. Determining the species traits that control niche structure remains an open and important question, and developing such a trait-based understanding of food web structure will facilitate developing new understanding of the connections between species traits and ecosystem function in complex ecological communities (Loreau 2009).

ACKNOWLEDGMENTS

Thanks to Jen Dunne and Ross Thompson for generously sharing their data sets.

LITERATURE CITED

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:716–723.
- Allesina, S. 2010. Predicting trophic relations in ecological networks: a test of the Allometric Diet Breadth Model. *Journal of Theoretical Biology*. [doi: 10.1016/j.jtbi.2010.06.04]
- Allesina, S., D. Alonso, and M. Pascual. 2008. A general model for food web structure. *Science* 320:658–661.
- Allesina, S., and M. Pascual. 2009. Food web models: a plea for groups. *Ecology Letters* 12:652–662.
- Benincà, E., J. Huisman, R. Heerkloss, K. D. Johnk, P. Branco, E. H. Van Nes, M. Scheffer, and S. P. Ellner. 2008. Chaos in a long-term experiment with a plankton community. *Nature* 451:822–825.
- Brose, U., R. J. Williams, and N. D. Martinez. 2006. Allometric scaling enhances stability in complex food webs. *Ecology Letters* 9:1228–1236.
- Cardinale, B. J., J. E. Duffy, D. S. Srivastava, M. Loreau, M. Thomas, and M. Emmerson. 2009. Towards a food web perspective on biodiversity and ecosystem functioning. Pages 105–120 in S. Naeem, D. Bunker, A. Hector, M. Loreau, and C. Perrings, editors. *Biodiversity, ecosystem functioning and human wellbeing*. Oxford University Press, Oxford, UK.
- Cattin, M.-F., L.-F. Bersier, C. Banasek-Richter, R. Baltensperger, and J.-P. Gabriel. 2004. Phylogenetic constraints and adaptation explain food-web structure. *Nature* 427:835–839.
- Chase, J. M. 2000. Are there real differences among aquatic and terrestrial food webs? *Trends in Ecology and Evolution* 15:408–412.
- Cohen, J. E. 1977. Food webs and dimensionality of trophic niche space. *Proceedings of the National Academy of Sciences USA* 74:4533–4536.
- Cohen, J. E. 1978. *Food webs and niche space*. Princeton University Press, Princeton, New Jersey, USA.
- Cohen, J. E., et al., editors. 1993. *Improving food webs*. *Ecology* 74:252–258.
- Cohen, J. E., F. Briand, and C. M. Newman. 1990. *Community food webs: data and theory*. Springer, Berlin, Germany.
- Cohen, J. E., and C. M. Newman. 1985. A stochastic theory of community food webs: I. models and aggregated data. *Proceedings of the Royal Society B* 224:421–448.
- Drossel, B., A. J. McKane, and C. Quince. 2004. The impact of nonlinear functional responses on the long-term evolution of food web structure. *Journal of Theoretical Biology* 229:539–548.
- Dunne, J. A. 2006. The network structure of food webs. Pages 27–86 in M. Pascual and J. A. Dunne, editors. *Ecological networks: linking structure to dynamics in food webs*. Oxford University Press, New York, New York, USA.
- Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002a. Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences USA* 99:12917–12922.
- Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002b. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters* 5:558–567.
- Dunne, J. A., R. J. Williams, and N. D. Martinez. 2004. Network structure and robustness of marine food webs. *Marine Ecology Progress Series* 273:291–302.
- Emmerson, M. C., and D. Raffaelli. 2004. Predator–prey body size, interaction strength and the stability of a real food web. *Journal of Animal Ecology* 73:399–409.
- Hurvich, C. M., and C.-L. Tsai. 1989. Regression and time series model selection in small samples. *Biometrika* 76:297–307.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? *American Naturalist* 93:145–159.
- Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science* 220:671–680.
- Loeuille, N., and M. Loreau. 2005. Evolutionary emergence of size-structured food webs. *Proceedings of the National Academy of Science USA* 102:5761–5766.
- Loreau, M. 2009. Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philosophical Transactions of the Royal Society B* 365:49–60.
- May, R. M. 1972. Will a large complex system be stable? *Nature* 238:413–414.

- May, R. M., S. A. Levin, and G. Sugihara. 2008. Complex systems: ecology for bankers. *Nature* 451:893–895.
- Morton, R. D., R. Law, S. L. Pimm, and J. A. Drake. 1996. On models for assembling ecological communities. *Oikos* 75:493–499.
- Neutel, A. M., J. A. P. Heesterbeek, and P. C. de Ruiter. 2002. Stability in real food webs: weak links in long loops. *Science* 296:1120–1123.
- Petchey, O. L., A. P. Beckerman, J. O. Riede, and P. H. Warren. 2008. Size, foraging, and food web structure. *Proceedings of the National Academy of Science USA* 105:4191–4196.
- Petchey, O. L., P. T. McPhearson, T. M. Casey, and P. J. Morin. 1999. Environmental warming alters food-web structure and ecosystem function. *Nature* 402:69–72.
- Pimm, S. L., and J. H. Lawton. 1977. Number of trophic levels in ecological communities. *Nature* 268:329–331.
- Pimm, S. L., and J. H. Lawton. 1978. Feeding on more than one trophic level. *Nature* 275:542–544.
- Shurin, J. B., D. S. Gruner, and H. Hillebrand. 2006. All wet or dried up? Real differences between aquatic and terrestrial food webs. *Proceedings of the Royal Society B* 273:1–9.
- Stouffer, D. B., J. Camacho, and L. A. N. Amaral. 2006. A robust measure of food web intervality. *Proceedings of the National Academy of Science USA* 103:19015–19020.
- Stouffer, D. B., J. Camacho, W. Jiang, and L. A. N. Amaral. 2007. Evidence for the existence of a robust pattern of prey selection in food webs. *Proceedings of the Royal Society B* 274:1931–1940.
- Sugiura, N. 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics, Theory and Methods* A7:13–26.
- Thébault, E., and M. Loreau. 2003. Food-web constraints on biodiversity-ecosystem functioning relationships. *Proceedings of the National Academy of Sciences USA* 100:14949–14954.
- Warren, P. H. 1996. Structural constraints on food web assembly. Pages 142–161 in M. E. Hochberg, J. Clobert, and R. Barbault, editors. *Aspects of the genesis and maintenance of biological diversity*. Oxford University Press, New York, New York, USA.
- Williams, R. J., A. Anandanadesan, and D. Purves. 2010. The probabilistic niche model reveals the niche structure and role of body size in a complex food web. *PLoS ONE* 5:e12092.
- Williams, R. J., and N. D. Martinez. 2000. Simple rules yield complex food webs. *Nature* 404:180–183.
- Williams, R. J., and N. D. Martinez. 2008. Success and its limits among structural models of complex food webs. *Journal of Animal Ecology* 77:512–519.
- Woodward, G., B. Ebenman, M. Emmerson, J. M. Montoya, J. M. Olesen, A. Valido, and P. H. Warren. 2005. Body size in ecological networks. *Trends in Ecology and Evolution* 20:402–409.
- Yodzis, P. 1981. The stability of real ecosystems. *Nature* 289:674–676.
- Zook, A. E., A. Eklof, U. Jacob, and S. Allesina. 2011. Food webs: ordering species according to body size yields high degree of intervality. *Journal of Theoretical Biology* 271:106–113.

APPENDIX A

Tables listing the data sets used and comparisons of the probabilistic niche model (PNM) with the allometric diet breadth model, minimum potential niche model, and groups model (*Ecological Archives* E092-155-A1).

APPENDIX B

Link probabilities predicted by the probabilistic niche model and empirically observed links for all 37 food webs (*Ecological Archives* E092-155-A2).

SUPPLEMENT

Data files containing overall model performance, maximum-likelihood parameters, and link probabilities for each of the 37 food webs (*Ecological Archives* E092-155-S1).