The Network Structure of Food Webs

Jennifer A. Dunne

Descriptions of food-web relationships first appeared more than a century ago, and the quantitative analysis of the network structure of food webs dates back several decades. Recent improvements in food-web data collection and analysis methods, coupled with a resurgence of interdisciplinary research on the topology of many kinds of “real-world” networks, have resulted in renewed interest in food-web structure. This chapter reviews the history of the search for generalities in the structure of complex food webs, and discusses current and future research trends. Analysis of food-web structure has used empirical and modeling approaches, and has been inspired both by questions from ecology such as “What factors promote stability of complex ecosystems given internal dynamics and external perturbations?” and questions from network research such as “Do food webs display universal structure similar to other types of networks?” Recent research has suggested that once variable diversity and connectance are taken into account, there are universal coarse-grained characteristics of how trophic links and species...
defined according to trophic function are distributed within food webs. In addition, aspects of food-web network structure have been shown to strongly influence the robust functioning and dynamical persistence of ecosystems.

1 INTRODUCTION

This chapter describes research that seeks to characterize and model the structure of food webs, complex networks of feeding (trophic) interactions among diverse species in communities or ecosystems. In particular, I discuss studies that search for generalities in the network structure, also referred to as topology, of food webs. Over the last fifteen years there has been dramatic growth in the field-based documentation of food webs, as well as a general increase in data quality in terms of diversity and resolution. This is a welcome trend, and provides fuel for further statistical assessment, model testing, and theory development. This chapter reviews the history of and current trends in the data, analyses, and models of food-web network structure. I attempt to be relatively comprehensive within this narrow domain, but as with all reviews, some papers and topics are not addressed. In the interest of focus and length, studies focused on food-web dynamics are not covered here, although they have been used to explore dynamic constraints on structure and vice versa (Pimm and Lawton 1978; McCann and Hastings 1997). Reviews of dynamical food-web models and their uses and limitations (Lawton 1989; Drossel and McKane 2003; Dunne et al. 2005; McKane and Drossel Chapter 9; Martinez and Bascompte Chapter 12), as well as “network analyses” of flow networks, primarily applied to particular marine or estuarine systems (Wulff et al. 1989; Christensen and Pauly 1993), can be found elsewhere.

Research on food-web structure is but one example in a very broad cross-disciplinary research agenda on the structure of all types of networks, both biotic and abiotic (Strogatz 2001; Albert and Barabási 2002). Network research is often couched in the framework and language of statistical mechanics and graph theory. Indeed, research on food-web structure has foundations in graph theory (the random graph theory of Erdős and Rényi 1960 as cited by Cohen 1990; see also Cohen 1977b; Sugihara 1982; Kenny and Loehle 1991) as well as natural history (the food web for Bear Island described by Summerhayes and Elton 1923). The chapter by Cartozo et al. (Chapter 3) goes into detail about how food-web studies fit into the broader network topology framework, and it introduces basic graph theory definitions and properties of interest. However, this review is primarily situated within the history, language, and theory of ecology, and given this context I generally use terms such as species rather than nodes and feeding links rather than edges. The biological terms evoke a rich conceptual history in ecology and connect food-web research to related areas of study, such
as evolutionary theory and conservation biology. However, more technical graph-theoretic terms do pop up, particularly in descriptions of more recent research.

The review is divided into three sections called “Early Phase: Pioneering Research,” “Middle Phase: Critique and Reassessment,” and “Current Phase: New Models, New Directions.” These phases correspond roughly to the late 1970s through the 1980s (Early Phase), the 1990s (Middle Phase), and the first half of the 2000s (Current Phase). In addition to many interesting questions, both basic and applied, that structural food-web research has been used to explore, its history provides a unique perspective on the interplay between theory, methods, and data, and on the role of critique. While the phases and topics emphasized within them are a subjective split of interrelated lines of research, they provide a useful way to organize a rich body of literature. The Early Phase and parts of the Middle Phase cover ground that is reviewed elsewhere (e.g., Lawton 1989; Cohen et al. 1990a; Pimm et al. 1991; Hall and Rafaelli 1993, 1997). However, a recent renewed interest in the network structure of food webs by ecologists as well as by physicists, mathematicians, biologists, and social scientists suggests that it would be wise to revisit early foundational work from a contemporary perspective. The final section, “Related Topics and Future Directions,” discusses relevant topics neglected by the three main sections and how those topics relate to future research directions.

2 EARLY PHASE: PIONEERING RESEARCH

Descriptions of feeding relationships among species go back at least to the late 1800s (studies by Forbes from 1876 on, reprinted in 1977; Camerano 1880 as cited by Cohen et al. 1990a). By the 1910s researchers began to produce images not unlike food-web figures seen today in textbooks, such as a network of insect predators and parasites on cotton-feeding weevils (Pierce et al. 1912 as cited by Pimm et al. 1991) and the hypothetical animal-oriented descriptions of Shelford (1913). By the 1920s, the first relatively detailed empirical descriptions of terrestrial (Summerhayes and Elton 1923, 1928) and marine (Hardy 1924) food webs appeared. Elton (1927) coined the term food chain, and termed all the food chains in a community a food cycle, which we now call a food web. Many descriptions of food webs, both hypothetical and empirical, followed those early efforts.

However, it was not until the late 1970s that quantitative, comparative research on potential generalities in the network structure of food webs arose. Cohen published the first collection of food webs in 1978, comprised of 30 webs with binary links (i.e., links showing only the presence and direction of a feeding relationship, with no weighting for flow or strength) compiled from the literature. He considered 14 to be community webs, webs that attempt to be reasonably inclusive of species in a particular system, and 16 were sink webs, selective webs that focus on one or more predator species, their prey, their preys’ prey, etc.
There are also source webs, selective webs that focus on one or more prey species, their predators, their predators' predators, etc. Both then and now, species in food webs, while often referring to actual biological species, in many cases represent other things: taxonomically related groups of species, all the way up to whole kingdoms (mites, arthropods, fungi); mixed groups of species (zooplankton); particular life-history stages (small-mouth bass young-of-year); parts of species (leaves, fruit); and non-living organic matter (detritus). Throughout this chapter, the term species will be used in the non-specific sense that includes all of these types of groupings within food webs, and the term taxa will sometimes be used in place of species.

In a strict sense, the trophic links in food webs are directed links, which means that a feeding relationship is directional (i.e., A eats B). However, effects of feeding move in both directions—A eating B has ecological and evolutionary implications for both species, at least when both “species” are living. Food-web interactions can be usefully represented as a matrix. The simplest way to build such a matrix is for rows to represent consumers/predators and columns to represent resources/prey. A “1” is assigned to the cell at the intersection of row $i$ and column $j$ if species $i$ feeds on species $j$, and a “0” is assigned if species $i$ does not feed on species $j$. Alternatively, rows can represent resources and columns consumers. Putting trophic relationships into a matrix format facilitates quantitative analysis of food-web structure. Larger datasets can be more efficiently represented using a two-column format, in which the first column lists the number of a consumer, and the second column lists the number of one of the resource species of that consumer.

2.1 COMPLEXITY-STABILITY AND FOOD-WEB STRUCTURE

Early collections of food webs provided, for better and for worse, empirical fodder for the complexity-stability debate. For several decades leading up to the 1970s, a dominant ecological paradigm was that complex communities are more stable than simple ones (Odum 1953; MacArthur 1955; Elton 1958; Hutchinson 1959). The argument in favor of complexity giving rise to stability in ecological communities was stated in a general way by MacArthur (1955), who hypothesized that “a large number of paths through each species is necessary to reduce the effects of overpopulation of one species.” He concluded that “stability increases as the number of links increases” and that stability is easier to achieve in more diverse assemblages of species, thus linking community stability with both increased trophic links and increased numbers of species. This convention was challenged by May in a seminal paper (1972) and book (1973) using dynamical mathematical modeling methods. May conducted local stability analyses of randomly assembled community matrices and demonstrated that network stability decreases with complexity, following Gardner and Ashby (1970). May found that simple, abstract communities of interacting species will tend to transition sharply from stable to unstable behavior as the complexity of the system in-
creases; in particular as the number of species \(S\), the connectance \(C\), or the average interaction strength \(i\) increase beyond critical values. May formalized this as a criterion that ecological communities near equilibrium will tend to be stable if \(i(SC)^{1/2} < 1\). In May’s framework, connectance refers to the probability that any two species will interact with each other; in a more technical sense, connectance refers to the percentage of non-zero elements in an interaction matrix (Gardner and Ashby 1970). The measure \(S\), the number of species in a food web, or more generally, the number of nodes in a network, will appear repeatedly throughout this review and will typically be referred to as species diversity or richness. “Diversity” and “richness” can also be quantified using more complex formulations not covered here.

Several papers since May (1972, 1973) have pointed out flaws and limitations in his analysis (e.g., Lawlor 1978; Cohen and Newman 1985a; Taylor 1988; Law and Blackford 1992; Haydon 1994). For example, patterns of species interactions are not random, and varying such patterns can have a significant impact on dynamics. However, May’s criterion and the general question of how diversity is maintained in natural ecosystems provided a framework on which to hang some readily accessible empirical data, namely the numbers of links and species in food webs. How does this work? Given a particular interest in species diversity \(S\), and assuming that average interaction strength \(i\) is constant, May’s criterion suggests that communities can be stable given increasing diversity, as long as connectance decreases. This can be empirically demonstrated using food-web data in three similar ways, by showing that (1) \(C\) hyperbolically declines as \(S\) increases, so that the product \(SC\) remains constant, (2) the ratio of links to species \(L/S\), also referred to as link or linkage density, remains constant as \(S\) increases, or (3) \(L\) plotted as a function of \(S\) on a log-log graph, producing a power-law relation of the form \(L = aS^b\), displays an exponent of \(b = 1\) (the slope of the regression) indicating a linear relationship between \(L\) and \(S\).

In an empirical framework, connectance is measured as the proportion of potential links among species that are actually realized. The simplest way of expressing this is \(C = L/S^2\) (directed connectance; Martinez 1991), where the numerator gives the number of observed links and the denominator includes all potential directed trophic links among \(S\) species, equal to \(S^2\). Food-web connectance has also been measured in more complex ways that exclude particular kinds of trophic links. For example, another measure is \(C = L/[S(S − 1)/2]\) which excludes all cycles, also called loops (Rejmánek and Starý 1979). Within food webs, cycles of length 1 (A eats A) are referred to as cannibalism, cycles of length 2 (A eats B eats A) are referred to as mutual predation, and even longer cycles are possible (e.g., A eats B eats C eats A). The \(S – 1\) part of the denominator excludes the main diagonal of an \(S\) by \(S\) matrix, thus eliminating cannibalism links, and dividing \(S(S − 1)/2\) by 2 excludes all other cycles. Thus, \(S(S − 1)/2\) constrains the zone of potential directed links in an \(S\) by \(S\) matrix to a triangle on one side of the main diagonal. It is the equivalent of counting the total number of possible undirected interactions (i.e., where “A interacts with B”
could mean that A eats B, B eats A, or both eat each other) in an S by S matrix, excluding intraspecific interactions, and thus is sometimes referred to as interactive connectance. However, if cycles do occur in food webs, \( C = L/[S(S - 1)/2] \) exaggerates connectance, since cycling links are counted in the numerator but are excluded from the denominator. This can be avoided by calculating interactive connectance based on undirected links, such that \( C = L/[S(S - 1)] \) (Warren 1990; Martinez 1991). How do these expressions relate to May’s criterion and the potential for demonstrating stability with increasing diversity? Given the form \( C = L/S^2 \), the hyperbolic decline of \( C \) with increasing \( S \) (so that \( SC \) is constant) is mathematically equivalent to constant \( L/S \), since \( SC = L/S \) is the same as \( C = L/S^2 \). Given the forms \( C = L/[S(S - 1)/2] \) or \( C = L/[S(S - 1)] \), the same equivalence between the hyperbolic decline of \( C \) and constant \( L/S \) occurs under the condition that \( S \) is large, that is, when \( (S - 1)/S \) approximates 1 (Macdonald 1979). This means that when \( S \) is small \( (S < 20) \), different forms of \( C \) can alter which hypothesis about variability of \( C \) is accepted or rejected (Martinez 1995).

Quickly following the publication of the first catalog of food webs (Cohen 1978), examples were published documenting all three ways of empirically corroborating the potential for stability with increasing diversity, assuming constant interaction strength (a big assumption that will be discussed briefly in Related Topics and Future Directions). MacDonald (1979) analyzed Cohen’s data and found that mean \( L/S \) of the 30 webs was 1.88 (SD = 0.27). Independently, Rejmánek and Starý (1979) compiled and analyzed 31 plant-aphid-parasitoid source webs and reported a hyperbolic relationship between \( S \) and \( C \) (using \( C = L/[S(S - 1)/2] \)) with a central tendency of \( C = 3/S \) (all data points fell between \( C = 2/S \) and \( C = 6/S \)) corresponding to \( L/S = 1.5 \) (MacDonald 1979). Briand (1983) analyzed 40 community webs including 13 from Cohen (1978), and using log-log regression analysis he found that trophic links increase as a nearly linear function of \( S \), with \( b = 1.10 \) (\( L = 1.3S^{1.10} \)). In these and other early studies (e.g., Pimm 1982; Auerbach 1984), the results seemed to indicate that connectance decreases with species richness and \( L/S \) is approximately constant with a value between 1 and 2. The ecological interpretation is that species tend to eat the same small number of prey (1 to 2 species on average) regardless of the diversity of the food web (Pimm 1982). Given May’s criterion, these results provided empirical support for the possibility of stable, diverse ecosystems.

E ensuing studies tested these findings with new data sets and introduced trophic aggregation. Expanding Briand’s 40 food webs, Cohen and Briand (1984) reported for 62 webs that \( L/S \) is “roughly independent of variation in \( S \)” with a value of 1.86 (\( SD = 0.07 \)). Unlike previous studies, they used trophic species aggregation (Briand and Cohen 1984; see also Sugihara 1982; Yodzis 1982), in which species that share the same set of predators and prey in a particular food web are lumped into a single trophic species. This was meant to reduce methodological artifacts due to researchers’ tendencies to resolve higher trophic level taxa more finely than lower trophic level taxa (Pimm 1982), which can add noise to trends and bias results (Briand and Cohen 1984). Sufficiently convinced
by the generality of the data that supported both scale invariance of $L/S$ and its value of $\sim 2$, Cohen and Newman (1985b) began to refer to the relationship as the link-species scaling law, and further corroborated it with an expanded set of 113 webs (Cohen et al. 1986) that will be referred to throughout this chapter as the 113 web catalog.

Using a separate set of 60 relatively well resolved insect-dominated food webs, drawn from a catalog of 95 such webs (Schoenly et al. 1991), Sugihara et al. (1989) reported $SC$ (equivalent to $L/S$) as “roughly independent of species number” for the subset of 41 webs with 10 or more species, ranging up to 87 species. While they did not use trophic species as the unit of analysis, they did follow methods introduced by Martinez (1988) and conducted an aggregation study of the 41 web subset, lumping species by trophic similarity until half of the original species remained. While the data showed that $SC$ “tended to fall slightly with increasing aggregation,” they nevertheless concluded that $SC$ was robust to aggregation. The aggregation study and its conclusions were meant to assuage concerns (Pimm 1982; Paine et al. 1988) that variable resolution in available data might create false patterns. However, the aggregation criterion used by Sugihara et al. (1989) resulted in about three-quarters of the insect webs being aggregated only to the trophic species level, so like Cohen and Briand they were primarily eliminating topological redundancy (Martinez 1993b). The 113 web catalog, the 60 insect webs, and other food webs were ultimately compiled in the ECOWeB database in a machine-readable format to facilitate analyses by other researchers (Cohen 1989).

### 2.2 EMPIRICAL REGULARITIES AND SCALE INVARIANCE

As the previous section suggests, a great deal of focus on constant linkage density ($L/S$) resulted from its connection to fundamental theory regarding ecosystem stability and diversity/complexity, a key ecological issue that continues to endure (McCann 2001). However, the first collection of food webs was compiled not to test the relationship between $L/S$ and $S$, but to look for other empirical regularities in the network structure of trophic interactions (Cohen 1978). Based on 14 community webs, Cohen (1977a) reported a ratio of prey to predators of $\sim 3/4$, constant across food webs with variable $S$. This ratio had been explored previously in other empirical datasets (e.g., Evans and Murdoch 1968; Arnold 1972; Cameron 1972). Cohen (1977b) also found that most of the 30 sink and community food webs are interval. This means that all of the species in a food web can be placed in a fixed order on a line such that each predator’s set of prey forms a single contiguous segment of that line. Intervality suggests that trophic niche space can be represented by a single dimension. Why this might be the case, or what the single dimension might represent, continues to be unclear (Williams and Martinez 2000). Other types of graph-theoretic properties quantifying how diets overlap in food webs such as triangulation have been explored (Sugihara 1982; Sugihara et al. 1989; Cattin et al. 2004).
A great deal of research following Cohen (1977a,b) focused on patterns of food-web network structure, and had more in common with Cohen’s initial search for empirical regularities than the search for constant \( L/S \) in the service of complexity-stability theory. From this viewpoint, \( L/S \) is just one of many properties of food-web structure, based on analysis of binary links, which may have a general value or central tendency across food webs with varying diversity. These types of food-web patterns, in which a property is found either to be constant, or under a weaker standard to not change systematically as the number of species across food webs changes, came to be referred to as *scale-invariant patterns* (Briand and Cohen 1984) or *scaling laws* (Cohen and Newman 1985b). Such scale-invariant patterns can be thought of as extremely general regularities that are theoretically valid from the smallest food webs to food webs that comprise the entire planet (Martinez and Lawton 1995).

Additional scale-invariant patterns found using early food-web data included: previously mentioned predator-prey ratios; constant proportions of top species (T, species with no predators), intermediate species (I, species with both predators and prey), and basal species (B, species with no prey), collectively called *species scaling laws* (Briand and Cohen 1984); and constant proportions of T-I, I-B, T-B, and I-I links between T, I, and B species, collectively called *link scaling laws* (Cohen and Briand 1984). Other general properties of food webs were thought to include: food chains are short (Elton 1927; Hutchinson 1959; Pimm 1982; Cohen et al. 1990a); cycling/looping is rare (Cohen and Newman 1985b); compartments, or subwebs with many internal links that have few links to other subwebs, are rare (Pimm and Lawton 1980); omnivory, or feeding at more than one trophic level, is uncommon (Pimm and Lawton 1978); and webs tend to be interval, with instances of intervality decreasing as \( S \) increases (Cohen 1977b; Yodzis 1984; Cohen and Palka 1990). Most of these patterns were reported for the 113 web catalog (Cohen et al. 1986, 1990). Select patterns, such as short food chains, constant predator-prey ratio, and scale-invariant fractions of T, I, and B species, were also documented in a subset of 41 of 60 insect webs with 10 or more species (Sugihara et al. 1989).

### 2.3 THE CASCADE MODEL

The many properties being proposed and explored threatened to become a kind of stamp collection of food-web patterns, with no particular rhyme, reason, or organizing principals. However, in a series of six papers published from 1985 to 1990 with the common title “A Stochastic Theory of Community Food Webs,” Cohen and colleagues sought to unify food-web patterns through a simple model called the *cascade model*. By ignoring dynamics and using a stochastic, binary link approach similar to that of random graph theory (Erdős and Rényi 1960), the cascade model sought to explain “the phenomenology of observed food web structure, using a minimum of hypotheses” (Cohen et al. 1990a). A number of other simple models for generating food-web structure were explored prior to
the cascade model (Cohen 1978; Pimm 1982; Sugihara 1982; Pimm 1984; Yodzis 1984) and several variations were explored concurrently with the cascade model (Cohen and Newman 1985b; Cohen 1990). Most performed poorly in predicting empirical trends, and only one appeared to perform as well as the cascade model (Cohen 1990).

The cascade model is based on two parameters, species richness $S$ and link density $L/S$. The model distributes species and feeding links stochastically, subject to two simple constraints: species are randomly placed in a one-dimensional feeding hierarchy, and species can only feed on species that are lower in the hierarchy than themselves (Box 1). This ensures a “triangular” matrix that prohibits cycles or loops, including cannibalism. The cascade model is simple enough that it is analytically tractable, although as a stochastic model it can also be explored through computational approaches such as numerical simulation. To explore whether the cascade model reproduced patterns of food-web structure in 62 community webs, Cohen and Newman (1985b) assumed constant $L/S$ and they tuned the value of $L/S$ to an average across the empirical data. They found that the cascade model successfully reproduced the qualitative patterns of both species and link scaling “laws” in all but the smallest webs. While it produced quantitatively similar values to those observed for species and link proportions, the cascade model was less successful in explaining variation in the data (Cohen and Newman 1985b).

Cohen et al. (1985) also tested whether the cascade model could reproduce the values of properties for particular food webs by tuning $S$ and $L/S$ to the values for each of the 62 webs. They found that the cascade model described the proportions of intermediate species (I) and B-I, I-I, and I-T links well. In the expanded 113 web catalog, Cohen et al. (1986) found that the cascade model made good predictions of numbers of food chains of each length (the frequency distribution) in most webs, while describing mean chain lengths adequately and chain-length variance less well. The cascade model also gave good qualitative support, and reasonable quantitative support, to the frequency of interval webs (Cohen and Palka 1990). In sum, the cascade model, a simple, stochastic, analytically tractable model, appeared remarkably successful at generating network topology and trends similar to those observed in empirical data at both “coarse-grained” and more detailed levels. It was suggested that the feeding hierarchy assumption of the cascade model might reflect natural processes such as body size constraints on feeding (Warren and Lawton 1987). Perhaps most importantly, the cascade model provided the first explicit and quantitative hypothesis that the network structure observed in early food-web data was not only non-random, but might be governed by simple rules.
3 MIDDLE PHASE: CRITIQUE AND REASSESSMENT

A rosy picture was thus painted by early practitioners. Empirical structural data appeared to corroborate aspects of dynamical complexity-stability theory; network structure appeared to be well-described by multiple empirical scaling laws; and a simple, stochastic graph-theoretic model based on the link-species scaling law and a seemingly ecologically reasonable principle of hierarchical feeding predicted the phenomenology of food-web patterns. However, the hounds were baying at the door. From the beginning, most of the purveyors of structural food-web research pointed out some of the limitations in the data they analyzed. The most obvious issue was that most of the food webs analyzed had very low diversity compared to the biodiversity known to be present in ecosystems. The webs of the 113 web catalog (Cohen et al. 1986) have 5 to 48 original species (mean = 17) and 3 to 48 trophic species (mean = 17), while the 60 insect webs in Sugihara et al. (1989) have 2 to 87 original species (mean = 22) and 2 to 54 trophic species (mean = 12). Another obvious issue was the highly uneven resolution and representation of “species” in most early food webs. Many types of organisms are aggregated, underrepresented, missing altogether, or misrepresented as basal species because no prey items were recorded for them. The webs of the 113 web catalog were culled from the literature, where they had been put together by a wide array of researchers, using a variety of methods, and for many different purposes, which did not include quantifying or testing structural food-web patterns. The 113 web catalog data were also explicitly purged of cannibalistic links. Since the cascade model excluded cannibalism, this increased the fit between the model and the data.

These and other methodological issues were taken up and amplified in several serious, and to many, devastating critiques of the adequacy of the data, casting doubt on the entire research program (Paine 1988; Polis 1991; Hall and Raffaelli 1993, 1997; Winemiller and Polis 1996). The most prominent early critique was by Paine (1988), who suggested that “future connectance-based development, even from sanitized webs, will not be enormously profitable,” due to the possibility of significant spatial and temporal variation in diets as well as idiosyncrasies in how researchers ascribe trophic links. He suggested that a tendency of researchers to describe trophic links more completely in small versus large webs, given the greater effort required to describe interactions in speciose systems in detail, could alone account for an apparent hyperbolic decline of connectance (C) with diversity (S).

Meanwhile, potential problems with the “conventional” view of C and L/S were emerging. Recall that the hyperbolic decline of C with S, constant L/S, and log L versus log S showing a power-law exponent of 1 all express the same thing: a linear increase of links with species in food webs. Thus, given May’s stability criterion, constant L/S associated with increasingly low C in increasingly diverse communities was thought to be a condition of stability. Empirical evidence began to point towards trends of higher connectance than expected in
apparently stable or persistent natural communities across gradients of diversity. Cohen and colleagues, the strongest proponents of a linear relationship between $L$ and $S$, noted that they could not exclude a nonlinear relationship like that hinted at by Briand (1983). When regressing log $L$ versus log $S$ for the 62 and 113 web catalogs, they found an exponent of $b = 1.36$. This makes $L^{3/4}$ rather than $L^1$ proportional to $S$, meaning links increase faster than species (Cohen and Briand 1984; Cohen et al. 1986; see also Schoener 1989). Based on relatively detailed trophic information for 24 species in an English stream, Hildrew et al. (1985) found higher $C$ than that reported for all but 2 of 40 webs analyzed by Briand (1983). Ten food webs based on detailed, highly resolved trophic interactions among 36 species in an English pond also displayed high $C$ (Warren 1989). Those were the first published data explicitly collected to test food-web patterns. A summary pond web including all species displayed higher $C$ than less diverse subwebs based on different habitats and sampling times. Detailed tropical fish sink webs (i.e., sampling focused on all fish present, their gut contents, prey lists for their prey, etc.) for swamps and streams in Costa Rica and Venezuela included 58 to 104 species and showed a trend of increasing $C$ with $S$ (Winemiller 1989, 1990).

There were also hints that other scale-invariant scaling “laws” had been overstated. A look at figures purporting to show scale-invariance of food-web properties provides little evidence supporting the strong version of scale invariance. In the classic scale invariance studies (Cohen and Briand 1984; Briand and Cohen 1984), a typical figure, such as percent of top species plotted as a function of $S$, displays a cloud of data points on which a line whose slope is constrained to be zero is superimposed. The height of this horizontal line is determined by the mean value of the property in question, calculated over all of the food webs. No regression is calculated or plotted; instead, the approximate visual “fit” of the line to the central tendency of the highly variable data is taken as evidence of scale invariance. In effect, a weak, non-statistical finding of “no relationship with $S$” was interpreted as evidence for a strong claim of the presence of scale-invariant scaling laws (Hall and Raffaelli 1993). In addition, exceptions to other food-web “generalities” started to appear. For example, detailed field-based food-web data suggested that food chains could be longer than previously claimed, and that omnivory and cannibalism might not be rare in some systems (Hildrew et al. 1985; Sprules and Bowerman 1988; Warren 1989).

By the early 1990s, most researchers readily acknowledged problems with the data and the potential impacts on food-web “laws,” particularly constant $L/S$ (Pimm et al. 1991). The question of where those problems might lead was summed up cogently in an excellent early review by Lawton (1989): “Confronted with limited data of highly variable quality, hardly any of which is really good, food web studies face either hand-wringing paralysis, or cautious efforts to see what can be discovered in the existing information. If nothing else, the latter course of action should serve as a spur to gather more and better data, particularly if published webs reveal evidence of interesting regularities and patterns
in nature. However, we have to accept that some of the patterns may eventually prove to be artifacts of poor information.” Indeed, a new wave of empirical food-web structure research was inspired by the inadequacy of the early data, the intriguing possibility of either corroborating or overthrowing previous theory or patterns, and the potential for describing new generalities. This research was fueled by improved data and more sophisticated methods, particularly concerning resolution of taxa, sampling effort, and sampling consistency.

3.1 IMPROVED DATA

An entirely new level of empirical detail characterizing diverse food webs was presented in 1991 in two seminal papers (Polis 1991; Martínez 1991). Polis (1991) published a dizzying array of trophic information, compiled over nearly two decades, for the Coachella Valley desert in California, whose biota include at least 174 species of vascular plants, 138 vertebrate species, 55 arachnid species, thousands of insect species including parasitoids, and unknown numbers of microorganisms, acari, and nematodes. Rather than trying to create a complete food web including all species, he compiled a number of detailed subwebs (a soil web, a gall web, a parasitoid web, a scorpion-focused web, a carnivore web, and a predaceous arthropod web) to demonstrate the enormous trophic diversity and complexity found in a type of ecosystem typically considered to be relatively simple and species-poor. Each subweb is more diverse than many of the community webs in the 113 web catalog and is also more complex in terms of number and density of feeding interactions. On the basis of the subwebs and a simplified, aggregated 30 taxa web of the whole community, Polis (1991) concluded, “most cataloged webs are oversimplified caricatures of actual communities” and are “grossly incomplete representations of communities in terms of both diversity and trophic connections.” Coachella Valley web properties include frequent omnivory, cannibalism, and looping; a high degree of interaction among species (i.e., $L/S$ close to 10); and a nearly complete lack of top species, since few species completely lack predators or parasites. These and other properties contradicted accepted food-web patterns, and Polis (1991) suggested that “theorists are trying to explain phenomena that do not exist.”

Martínez (1991) compiled a detailed community food web for Little Rock Lake, Wisconsin, in one of the earliest studies to explicitly test food-web theory and patterns (see also Martínez 1988; Warren 1988, 1989). By piecing together diversity and trophic information from multiple investigators actively studying the biota of Little Rock Lake, he was able to produce a relatively complete and highly resolved food web of 182 taxa, most identified to the genus, species, or ontogenetic life-stage level, including fishes, copepods, cladocera, rotifers, diptera and other insects, mollusks, worms, porifera, algae, and cyanobacteria. The resulting 182 original-species web (i.e., the web that includes whatever taxa and links the original investigator reported) and the 93 trophic-species web generated from it (i.e., the web that results from trophic species aggregation of an
original-species web) were by far the most diverse and complete depictions of the community food web of a complex ecosystem yet published. Previously, the largest community food webs described were an 87 original-species web of insects associated with a felled oak log (Sugihara et al. 1989) and a 48 trophic-species web of the Sonora Desert (Briand and Cohen 1987). In the 93 trophic-species Little Rock Lake web, \( L/S = 11 \), looping is common, food chains are long, and there are very few top species (see also Hildrew et al. 1985; Sprules and Bow-eraman 1988; Warren 1989; Polis 1991). This structure is quite different from patterns observed in prior web catalogs and patterns predicted by the cascade model (Cohen et al. 1990a).

While other detailed food webs were published at about the same time as the Little Rock Lake food web (Martinez 1991), they either were compiled for less diverse systems (Warren 1989); focused on subwebs based on particular sub-habitats, species, or substrates (Winemiller 1989, 1990; Polis 1991; Schoenly et al. 1991; Havens 1992), or had less even representation of taxa resulting from strict criteria for designating links (Hall and Rafaelli 1991). Because the Little Rock Lake web was based on expert knowledge developed over many years of known or probable feeding relationships, it has been suggested that this web overestimates feeding links compared to food webs based on a discrete set of observations, for example, one- or few-times sampling of species and their gut contents coupled with feeding trials (Hall and Raffaelli 1997). Indeed, “cumulative” webs like the Little Rock Lake web have been shown to alter \( S, L \), and other web properties compared to “snapshot” webs (Schoenly and Cohen 1991; and see Sampling Consistency section). For the purposes of discerning coarse-grained patterns of network structure in food webs, such cumulative webs may be more useful than snapshot webs. Approaches that narrowly constrain spatial and temporal boundaries of sampling may miss structurally and dynamically important species and links that are uncommon or rare (Martinez and Dunne 1998); for example, a little used food resource that becomes crucial during periods of scarcity of other resources. Regardless of the inconsistencies, flaws, and limitations still to be found among the second wave of data, all of the datasets mentioned, and others not mentioned, represent significant improvements over the 113 web catalog in terms of field-based observation, higher and more even resolution, and/or greater diversity. Most of these improved datasets were specifically collected to test different methodological aspects of analysis of food-web structure, as detailed below.

### 3.2 SPECIES AGGREGATION

In addition to setting a new standard for compiling community food-web data, Martinez (1988, 1991) was the first researcher to look systematically at the effects of variable species resolution and aggregation on the network structure of food webs. While trophic species aggregation (Cohen and Briand 1984) is based on 100% trophic equivalence of species, the threshold for aggregation can be re-
laxed to reduce the resolution of food-web data by progressively lumping taxa based on less and less trophic similarity. There are different indices that can be used to quantify similarity between objects, and the Jaccardian similarity index (Jaccard 1900) is probably the best known and most widely used in food-web research (Martinez 1988, 1991, 1993; Sugihara et al. 1989; Yodzis and Winemiller 1999). Within a food-web context, two basic choices are necessary for deciding how to aggregate taxa: whether to use an additive versus multiplicative Jaccardian index to define similarity between each pair of species, and how to define the similarity of two aggregates based on pairwise similarities of species in the aggregates (Yodzis and Winemiller 1999). Martinez (1988, 1991) used an additive Jaccard index to determine similarity of pairs of species in Little Rock Lake, and then aggregated taxa based on an average-linkage method that calculates similarity between two aggregates as the average of similarity indices across all possible pairs between the aggregates. Interestingly, a later detailed study of trophospecies aggregation methods concluded that out of 12 combinations of the two Jaccard similarity indices and six cluster linkage methods, the additive average-linkage method performed better than the other methods (Yodzis and Winemiller 1999).

Martinez (1991) explicitly tested the hypothesis that patterns observed in early food-web catalogs are an artifact of the low resolution and high aggregation of the data. Using the detailed Little Rock Lake food web, the aggregation methods discussed, and three criteria for designating links between aggregates (see Sampling Effort section), he created three sequences of increasingly aggregated versions of the original 182 species web. The sequences end when species in the aggregates share only 10% of their predators and prey, resulting in webs with nine highly aggregated taxa. In effect, trophic aggregation reduces the Little Rock Lake food web to levels of resolution and diversity comparable to those in the 113 web catalog. Martinez then analyzed how different food-web properties changed, or do not change, with increasing aggregation and thus decreasing $S$. For example, links per species, quite high in the 93 trophic-species web ($L/S = 11$), drops steadily as aggregation increases, so that the 9 to 42 species versions of the Little Rock Lake web display similar $L/S$ ($\sim 1$ to 4) and diversity to the webs in the 113 web catalog (Cohen et al. 1986). Chain-length statistics, trophic-level statistics, and several proportions of links and species are also very sensitive to aggregation (table 1), and display comparable values to earlier data once diversity is reduced to similar levels. Based on these results, Martinez (1991) argued that “most published food-web patterns appear to be artifacts of poorly resolved data.” The main property relatively robust to increasing aggregation and decreasing diversity is directed connectance, $C = L/S^2$, which displays values that hover near 0.10 until $S$ drops below $\sim 20$. A trophic resolution study performed on 11 of 60 insect webs (Sugihara et al. 1989; Schoenly et al. 1991) with 20 or more trophic species produced similar results: compared to poorly resolved versions of webs, higher resolved webs have higher $L/S$, $I$, $I-I$ and mean chain length, and lower $T$, $B$, and $T-B$ (Martinez 1993b). Directed connectance,
predator-prey ratio, and I-I link ratio are less sensitive to aggregation than other properties.

Hall and Raffaelli (1991) also conducted an aggregation study, in this case lumping species in a 92-species food web of the Ythan Estuary based on taxonomic rather than trophic similarity. They created four increasingly aggregated versions of the original web through qualitative decisions about how to lump taxonomically related taxa at each level. Similar to Martinez (1991) they found that percent basal species increases and mean chain length decreases with aggregation (lower S), but unlike Martinez (1991), most properties appear relatively insensitive to aggregation. Their web, while more speciose than webs in the 113 web catalog, was biased in similar ways, which may have impacted its sensitivity to lumping. There is a large number of top species (26 = 28%), and a small number of basal species (3 = 4%). Also, shorebirds are highly resolved, have no predators, and comprise more than a quarter of all taxa, while lower trophic level taxa are increasingly aggregated. Sugihara et al. (1997) conducted both taxonomic and trophic aggregation analyses on a set of 38 of 60 insect webs (Schoenly et al. 1991) and concluded that most properties except for $L/S$ and B are insensitive to changing $S$ due to aggregation. The sensitivity of $L/S$ in the insect webs was counter to Sugihara et al. (1989) but in line with Martinez (1993b). The lack of sensitivity in other properties is likely because the most aggregated versions of the insect webs usually retained more than 50% of the original $S$. Sugihara et al. (1997) acknowledged, “after a few food webs had been aggregated to 70–89% of their original size, even the sturdiest food web properties became sensitive to these coarser degrees of data resolution.” This corroborated results reported by Martinez (1991, 1993b). A potentially confounding aspect of both taxonomic aggregation analyses (Hall and Raffaelli 1991; Sugihara et al. 1997) is their use of a maximum linkage criterion for designating links between aggregates, which is a weak standard that tends to dampen changes in the value of structural properties with $S$ (see Sampling Effort section).

### 3.3 SAMPLING EFFORT

A question related to the effects of species aggregation on food-web structure is, How do different observation thresholds for links affect observed patterns? Winemiller (1989, 1990) was the first to explicitly quantify these sampling effort effects. In constructing a set of detailed tropical fish sink webs, he recorded not only the presence of feeding links but also the volumetric fraction of each prey species in a predator’s diet, usually based on gut content analysis. This volumetric fraction was used as an estimate of the relative strength of the predator-prey interaction, which then allowed the construction of webs that include all links versus webs that exclude “weak” trophic links. Winemiller (1989), who reported a strong positive relationship between $C$ and $S$ when all links are included, found that even when links representing $< 1\%$ of gut contents are excluded, there is still a slight positive relationship. Using a range of link thresholds from 0 to 4.5%,
Winemiller (1990) found that food-web properties are very sensitive to changing thresholds of link inclusion, with rapid changes between 0 and 1% thresholds that level off as thresholds approach 4.5%. Various food-web properties based on early catalogs look more like values observed for the detailed fish food webs at 4 or 4.5% link exclusion thresholds, suggesting that “food-web diagrams taken from the literature tend to depict only the strongest feeding interactions” (Winemiller 1990). Connectance increases as more links are included. Goldwasser and Roughgarden (1997) also found that $C$ increases as they include rarer diet items in a food web focused on Anolis lizards on the island of St. Martin. In general, $L/S$ decreases as its sensitivity to $S$ with increasing link thresholds, and the mean value of $L/S$ also decreases (Winemiller 1990; Tavares-Cromar and Williams 1996; Bersier et al. 1999).

Martinez looked at a slightly different aspect of link threshold effects on food-web structure by examining different criteria for assigning a feeding link between aggregates of species in resolution studies of the Little Rock Lake food web (Martinez 1991) and 11 insect webs (Schoenly et al. 1991) with 20 or more trophic species (Martinez 1993b). The strictest criterion is minimum linkage, which requires that every species in an aggregate be linked to every member of another aggregate for them to be linked. The weakest criterion is maximum linkage, which requires that only one member of each aggregate share a link for them to be linked. An average linkage criterion rests between the two; for example, two aggregates are linked if half their species share links. For all of the webs examined, many food-web properties are quite sensitive to linkage criteria, especially to minimum and average criteria (Martinez 1991, 1993b). Maximum linkage generally impacts the values of food-web properties the least across a range of aggregation. Some structural properties are more sensitive to linkage criterion choice than to aggregation. This means that webs with similar $S$ display a wider range of values across different linkage criteria than cases in which $S$ is allowed to vary for a particular linkage criterion (table 1).

Sampling effort potentially impacts the inclusion of both links and species in food webs. For example, source webs (one or more basal species, their consumers, their consumers’ consumers, etc.) are easier to compile than community webs, but, by definition, undersample species and links of the community within which they are embedded. About half of the 60 insect webs analyzed by Sugihara et al. (1989) and reanalyzed elsewhere have only a single basal “species” (often carrion or dung) and are, in effect, source webs. Detailed community food webs have been used to show that many properties are indeed sensitive to the inclusion of increasing numbers of source species (Hawkins et al. 1997); for example $L/S$ tends to increase. The study that perhaps best captures the impacts of sampling effort as it might actually manifest in field observations was conducted by Martinez et al. (1999). They analyzed a detailed grass source web (10 grass species plus 77 endophytic insect species, including herbivores and parasitoids, living inside the grass stems) that included data on the frequency with which each consumer and each feeding relationship were observed. They simulated increases
TABLE 1  Summary of sensitivity of selected food-web properties in the Little Rock Lake food web (Martinez 1991, table 2) to species aggregation and linkage criteria. For aggregation, High refers to a large and clear systematic change across variable $S$, Medium refers to a substantial change, and Low refers to a relatively small change. For linkage criteria, High means that the property is more sensitive to change in linkage criteria than to aggregation, Medium means that the sensitivity of the property is similar to both linkage criteria and aggregation, and Low means that the property is less sensitive to linkage criteria than to aggregation. The last seven properties are proportions relative to the total number of species in the web (Top, Intermediate, Basal) or the total number of links in the web (T-I, T-B, I-I, I-B).

<table>
<thead>
<tr>
<th>Property</th>
<th>Sensitivity to Aggregation</th>
<th>Sensitivity to Linkage Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interactive connectance</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Directed connectance</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>$L/S$</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>$SC$</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Average chain length</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Maximum chain length</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Predator/prey ratio</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Top species</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Intermediate species</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Basal species</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>T-I links</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>T-B links</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>I-I links</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>I-B links</td>
<td>Medium</td>
<td>Medium</td>
</tr>
</tbody>
</table>

in sampling effort in terms of inclusion of both links and species in different versions of the food web, and looked at the impacts on observed connectance. In this case, directed connectance initially decreases with lower thresholds for including species and links, but quickly reaches an asymptote beyond which $C$ changes very little with increasing sampling effort. This asymptote only appears among trophic-species versions of the food webs, but closely approximates $C$ of the original-species web.

3.4 SAMPLING CONSISTENCY

One of the critiques of the 113 web catalog was that sampling methodology varied widely. The first researcher to address this issue explicitly in the context of testing structural patterns was Havens (1992). He constructed food webs representing the pelagic communities of 50 small lakes and ponds in the Adirondacks,
New York, using consistent sampling methods and identical linkage criteria. The number of taxa ranges from 10 to 74, and they are resolved to the genus, species, or ontogenetic life-history stage level. Havens used a method that can be referred to as “subsampling from a metaweb.” Here, metaweb refers to a master web that includes all of the taxa found among multiple similar habitats and all possible feeding links if all of the species co-occurred in a single habitat. In this case, the metaweb includes 220 taxa found across all of the lakes and feeding links among them (see also Sprules and Bowerman 1988). The 50 individual webs were created using species lists for each lake and then deriving feeding interactions from the metaweb. The main drawback of this approach is that it assumes that predator selectivity does not vary among habitats (Havens 1992); that is, if species A eats species B in lake X, A will also eat B if they are both present in lake Y. While the strength or magnitude of interactions between A and B will almost certainly vary from lake to lake, it is less likely that the presence of particular feeding interactions will vary (Havens 1992), especially when the food webs are of the same habitat type and within the same climatic and geographic region.

Havens (1992) found that $L/S$ increases fourfold over the range of diversity of the webs, providing more evidence that constant $L/S$ is related to deficiencies in the earlier data and “does not reflect a real ecological trend.” He also reported that fractions of species and links show no significant trends with $S$, and concluded that those properties are scale invariant, consistent with Cohen and Briand (1984). However, compared to empirical values (Briand and Cohen 1984) and values predicted by the cascade model (Cohen and Newman 1985b), the lake food webs display central tendencies much lower for $T$ and $T-I$ (similar to Warren 1989; Martinez 1991) and higher for $B$ and $I-B$.

Havens’s approach is only one way to create multiple food webs in a consistent manner. Another common method is simply to chose a type of habitat, and to create food webs for particular instances of that habitat (e.g., multiple streams within a region) using consistent field sampling and trophic link attribution methods at each site (Winemiller 1990; Townsend et al. 1998; Schmid-Araya et al. 2002). This approach has also been used to create multiple time-specific webs at particular sites in order to look at temporal resolution issues (Baird and Ulanowicz 1989; Schoenly and Cohen 1991; Tavares-Cromar and Williams 1996). Composite webs that integrate over time will necessarily have greater $S$ and $L$ than the time-specific webs on which they are based. They have also been used to demonstrate the sensitivity of other structural food-web properties to temporal variation and integration (Schoenly and Cohen 1991; Deb 1995; Tavares-Cromar and Williams 1996; Thompson and Townsend 1999).

3.5 ABANDON SHIP?

Improved data, as well as studies that looked at species aggregation, sampling effort, and sampling consistency indicated that many conventionally accepted patterns of food-web structure, especially constant $L/S$, were almost certainly
artifacts of weak data or methodology. For many researchers, the array of issues being brought to light was enough for them to wash their hands of the entire research program. This attitude was reflected in the change in focus between the first and second major international symposia on food-web research. The first symposium, sponsored by the U.S. Department of Energy and organized by Oak Ridge National Laboratory, occurred in October 1982 and was followed a year later by a proceedings (De Angelis et al. 1983). This landmark meeting was dominated by talks relating to theory, dynamical and structural models, and topological patterns. The second major symposium on food webs, convened in September 1993 at Colorado State University’s Pingree Park Conference Center, with a proceedings following a few years later (Polis and Winemiller 1996), was dominated by talks on empirical and experimental research. While several of the talks and resulting book chapters discussed food-web structure in the context of particular systems, there was almost no treatment of general properties or patterns, apart from a review of criticisms of such research in the introduction (Winemiller and Polis 1996). As Paine notes in the preface, “web metrics...are pleasantly inconspicuous.” Even dynamical modeling was sidelined—of 37 main chapters, only 4 focused primarily on dynamical models (Abrams 1996; Arditi and Michalski 1996; Hastings 1996; Yodzis 1996). The overwhelming attention on experimentation and describing the impact of variability (e.g., spatial, temporal, environmental, habitat) on the dynamics of particular food webs or sets of trophic relationships came as a relief to many dissatisfied with the earlier work, which was often viewed as a gross oversimplification of the natural world. The pendulum in food-web research had swung sharply in favor of the WIWAC school—the “world is infinitely wonderful and complex” (Lawton 1995), and away from the search for generalities.

Nevertheless, over the course of the 1990s, ecologists continued to put together improved, detailed, field-based data on food-web network structure. An article co-authored by 24 top food-web researchers suggested a variety of ways for “improving food webs” with a focus on better data collection (Cohen et al. 1993). However, there is no universal correct way to compile a food web: all food webs will reflect the focus of researchers on particular methods, taxa, habitats, questions, and spatial and temporal boundaries and scales. The promise that structural food-web research holds is that by stepping far enough back from noisy details of particular systems, coarse-grained attributes and generalities will emerge, despite inherent variability and noise associated both with particular systems and with particular research approaches and agendas. In effect, this is a statistical mechanics approach to ecology. In this spirit, a few researchers continued to explore food-web patterns to call into question or corroborate old “laws” and theory, and also to offer alternative hypotheses about the network structure of food webs based on improved data and/or statistical analyses.
3.6 SCALE DEPENDENCE

One alternative hypothesis to scale-invariant patterns is “scale dependence.” In his Ecological Society of America MacArthur Award Lecture article, Schoener (1989) suggested that “S-independence” of food-web properties including $L/S$, prey-predator ratios, and fractions of basal and top species was unlikely in both principle and practice. In other words, rather than being scale-invariant, many food-web patterns are likely to be scale-dependent. Schoener explored a simple conceptual model based on an extension of Pimm’s (1982) notion that species should be limited in the number of prey species they can consume. Schoener suggested that not only may the number of prey species (generality) be constrained, but the number of predator species against which a species can defend (vulnerability) may also be constrained. He suggested that as food-web $S$ increases, the vulnerability of a species also increases. The consequences of these basic assumptions include scale-dependence of many food-web properties. Schoener (1989) tested those predictions with a statistical analysis, which Cohen and colleagues had avoided, of 98 webs drawn from the 113 web catalog and from source papers for particular webs in the catalog. He found the data generally agreed with the $S$-dependent predictions, not with scale-invariance. The data also supported the basic assumptions of his conceptual model: generality does not increase with $S$, while vulnerability does.

Several studies previously mentioned utilized improved data as well as novel methodological analyses and found that many purported scale-invariant properties of food webs appeared to be attributes of poorly resolved webs, with quite different values observed in highly resolved or sampled webs (Warren 1989, 1990; Winemiller 1990, 1991; Martinez 1991, 1993b; Hall and Rafaelli 1991; Polis 1991; Hawkins et al. 1997). This provided some empirical support for “scale-dependence” of most food-web properties. Additional, more explicit support came in the form of reanalysis of Havens’ 50 lake webs (Martinez 1993a) and subsets of Schoenly et al.’s (1991) insect webs (Martinez 1994). The reanalyses, which used more appropriate statistical tests and/or trophic species aggregation, reported significant scale-dependent trends of species and link proportions across food webs, similar to what Schoener (1989) found in his reanalysis of the 113 web catalog. The scale-dependent trends reported by Martinez (1993a, 1994) were actually present in data originally used to support scale invariance, but either the trends were weak and lacked significance, or significance tests were not conducted properly, if at all (Havens 1992; Sugihara et al. 1989). More sophisticated statistical methods based on generalized linear models also suggested that fractions of species are sensitive to $S$, based on a subset of 61 of the insect webs (Murtaugh and Kollath 1997).

An analysis of observed patterns of species fractions ($T$, $I$, $B$) at local scales and expected patterns of species fractions at regional and global scales based on biodiversity estimates suggests a limit to food-web scale dependence (Martinez and Lawton 1995). Empirical webs with less than a hundred species appear to
display strong scale dependence. However, hypothetical webs with more than a thousand species appear to display scale invariance due to unchanging fractions of species: no top species (due to cannibalism, parasitism, mutual predation, and longer loops), a small percentage of basal species, and a large majority of intermediate species. In essence, given that researchers are unlikely to compile detailed food webs with more than a thousand species, the food webs they can describe are likely to exhibit scale dependence. Using a similar concept, Bersier and Sugihara (1997) attempted to rehabilitate scale invariance for webs with \( S \) of \( \sim 10 \) to 100, arguing that species and link fractions in 60 insect food webs (Sugihara et al. 1989) display scaling regions: one region of scale dependence and one region of scale invariance. However, in this case, webs containing 12 or fewer original species or 7 or fewer trophic species are suggested to exhibit scale dependence, while webs with more species are suggested to exhibit scale invariance. These cut-off points are one to two orders of magnitude lower than those suggested by Martinez and Lawton (1995). At this point, the massaging and remassing of the Schoenly et al. (1991) insect web catalog is more confusing than convincing. Regardless of what is thought of the various statistical approaches, the fact that the insect webs are not comprehensive community webs, but are mostly source/substrate webs with mean original \( S \) of 22 and mean trophic \( S \) of 12, should lessen their importance in the overall assessment. The balance of evidence suggests that early patterns of scale invariance are artifacts of poorly resolved data, and that scale dependence of most properties is likely to be observed across higher quality datasets, at least within the range of diversity (i.e., \( S < 1000 \)) that ecologists are likely to be able to sample in detail.

3.7 CONSTANT CONNECTANCE

Just as improved data and methodological analyses highlighted problems with scale invariance, they also showed problems with constant \( L/S \) and hyperbolic decline of \( C \) with \( S \), as discussed previously. \( L/S \) was shown to be much greater than 2 for a variety of improved food-web datasets (e.g., Polis 1991; Martinez 1991; Hall and Raffaelli 1991). A set of fish sink webs with \( \sim 20 - 120 \) \( S \) displayed strongly increasing \( C \) with \( S \), although when links representing 1% or less of diet were excluded, \( C \) increased only slightly with \( S \) (Winemiller 1989). Several aggregation and sampling effort analyses suggested that \( C \) is relatively robust to changes in \( S \) (Warren 1989; Martinez 1991, 1993b; Martinez et al. 1997) while \( L/S \) is not, except possibly at low levels of link sampling (Bersier et al. 1999). Consonant with the former studies, an alternative hypothesis of “constant connectance” emerged out of new data and analyses (Martinez 1992). The mathematical difference between constant \( C \) and constant \( L/S \) can be simply stated using a log-log graph of links as a function of species (fig. 1). As discussed previously, if a regression can be reasonably fit to the data, it produces a power law of the form \( L = aS^b \). In the case of the link-species scaling law, \( b = 1 \), which means that \( L = aS, L/S = a \), and thus \( L/S \) (directed connectance) is constant. In
the case of constant connectance, $b = 2$, which means that $L = aS^2, L/S^2 = a$, and thus $L/S^2$ (directed connectance) is constant. Instead of $L/S$ being constant, $L/S$ increases as a fixed proportion of $S$. One ecological interpretation of constant connectance is that consumers are likely to exploit an approximately constant fraction of available prey species, so as diversity increases, the number of links per species increases (Warren 1990).

Do data support constant connectance? For a set of 15 webs derived from an English pond (11 habitat webs, 4 arbitrary subwebs, and a summary composite web), Warren (1990) reported $b = 2$, and more specifically $L = 0.24S^2$. He suggested that within a community, “increasing $S$ should produce a curvilinear relationship whereby $L$ is roughly proportional to $S^2$.” He hypothesized that the value of that proportionality will be greater for subwebs dominated by generalist feeders and lower for those dominated by specialists. Warren (1990) suggested that looking at the $L - S$ relationship across communities would result in a great degree of scatter and only a slight upward trend due to sampling effects, a hypothesis that appeared consistent with previous food-web data. However, another study hypothesized that the proportionality of $L$ with $S^2$ does hold across communities (Martinez 1992). Trophic species versions of 175 webs display $b = 1.54$ ($R^2 = 0.93$), smaller than $b = 2$ expected for constant $C$ (Martinez 1992). The 175 webs, which include the 113 web catalog, the 60 insect webs, the Little Rock Lake food web, and a food web of the island of St. Martin (Goldwasser and Roughgarden 1993), are dominated by the earlier, more poorly resolved data. To address this, Martinez (1992) conducted an analysis of a subset of the 12 most “credible” datasets in terms of resolution and completeness, representing the whole range of $S$ (2 to 93), and found $b = 1.73$ ($R^2 = 0.98$). Further distillation of the data by eliminating food webs with less than 10 species resulted in 5 webs with $b = 2.04$ ($R^2 = 0.95$).

Based on these and other analyses, Martinez (1992) argued that the true value of $b$ in highly resolved webs across a wide range of $S$ is likely closer to 2 than to 1, suggesting “roughly constant connectance within relatively homogeneous environments.” If connectance is roughly constant ($b \approx 2$), mean $C$ is quantified by the value of $a$, which is 0.014 ($SD = 0.06$) for all 175 webs and 0.11 ($SD = 0.03$) for the five web subset, suggesting that 0.11 is the best estimate of mean $C$. This indicates that approximately 10% of all possible feeding links, including cannibalism and other loops, are actually realized in food webs compiled “within” habitats where species are likely to be relatively well mixed. Food webs compiled across obvious environmental boundaries (e.g., a lake and its surroundings) are likely to have lower $C$, since there will be species that never encounter one another and thus have no chance of a feeding relationship. Also, connectance will vary to the degree that specialists, generalists, or omnivores are prevalent (Warren 1990).

The hypothesis of constant connectance, or $b = 2$, has been called into question by a few later studies, although in some cases a second look suggests otherwise. Havens (1992) reported $b = 1.4$ for 50 pelagic lake food webs. Re-
analysis without forcing the regression through 0 showed that the data display $b = 1.9$ ($R^2 = 0.92$) (Martinez 1993a) and mean $C$ of 0.10 ($SD = 0.02$). These results corroborate prior predictions of the likely values of both $a$ and $b$ (Martinez 1992). Using a set of 22 stream food webs with 22 to 212 species, Schmid-Araya et al. (2002) reported $b = 1.30$, lower than the value of $b = 1.36$ for the 113 web catalog, as well as relatively low $C$ that ranges from 0.03 to 0.12. The webs focus on algae, micro- and in some cases meiofauna, and macroinvertebrates, with vertebrates excluded. It is unclear how many and what types of vertebrates might be a part of any of the food webs. The exclusion of higher trophic-level taxa, such as generalist opportunistic feeders on invertebrates, would tend to decrease $C$ and could also lower the value of $b$. However, it may be that low mean $C$ and different scaling of $C$ are characteristic of stream webs.

Figure 1 shows the relationship between links and species for 19 relatively diverse food webs from a variety of habitats (e.g., pond, lake, stream, desert, grassland, rainforest, coral reef, marine shelf), with trophic species of 25 to 172 (refer to Dunne et al. 2002b, 2004 for information on individual webs). For this set of data, $b = 1.5$, but there is a large amount of scatter ($R^2 = 0.68$) (see also Montoya and Solé 2003). Directed connectance ranges from 0.03 to 0.3, with a mean of 0.13. The regression line for the data does not deviate strongly from the predicted line for constant connectance, but does obviously deviate from the line predicted by the link-species scaling law (fig. 1). In any case, these data are clearly too variable to convincingly demonstrate a particular $L - S$ relationship, other than not supporting constant $L/S$, perhaps because they represent such widely different ecosystems and methods (Warren 1990). Other studies have questioned the fit of constant connectance to empirical data, for example by conducting more appropriate statistical analysis of existing data (Murtaugh et al. 1998) or by doing detailed studies of consumer guild diets (Winemiller et al. 2001). The latter approach may not be a good way to test for constant connectance since it focuses on the trophic breadth of particular kinds of organisms, and $C$ is a global property of a food web calculated across many types of species. Whether values of $b$ are closer to 1.5 or 2, directed connectance for resolved community webs appears constrained to $\sim 0.03$ to 0.3 out of a possible range of 0 to 1. The central tendency across communities appears to be $\sim 0.10$ to 0.15, much lower than a null expectation of 0.5 (Kenny and Loehle 1991).

4 CURRENT PHASE: NEW MODELS, NEW DIRECTIONS

The cascade model was largely neglected during the 1990s as researchers either focused on methodological issues or veered entirely away from structural food-web analyses, especially research on the relationship of dynamical stability to diversity and connectance. Given that the cascade model was developed with data that turned out to be poorly resolved, it was unlikely to survive testing with improved data that differed dramatically from earlier data. Indeed, its fun-
The fundamental assumption of no looping, particularly cannibalism, was shot down repeatedly in improved datasets. It was also unclear whether the cascade model even described early data well (Schoener 1989). A graph theoretic analysis suggested that the relationship between connectance and species in early empirical webs that appeared to support the cascade model could not be distinguished from the relationship generated by a random model that included sampling effects, although other aspects of potentially non-random network structure were not considered (Kenny and Loehle 1991). Another analysis looked at the goodness of fit of the cascade model to the early ECOWeB data and found that, while it did seem to capture the central tendency of the data, it failed to characterize variability. Namely, the data were over-dispersed in relation to model predictions, leading to rejection of the model and a modified form of the model at “essentially any significance level” (Solow 1996). Furthermore, the random distribution of links by the cascade model, albeit within a constrained portion of an interaction matrix, resulted in less species lumpiness than found in the Little Rock Lake and larger ECOWeB food webs. In other words, there is more overlap of predators and prey for species in diverse empirical webs than in simulated webs based on a random distribution of links (Solow and Beet 1998).
4.1 NICHE MODEL

A new food-web structure model, in the tradition of the phenomenological, graph-theoretic, stochastic cascade model, was proposed by Williams and Martinez (2000) (Box 1). This niche model addresses several limitations of the cascade model, particularly the assumption of link-species scaling, the exclusion of looping, and the lack of trophic overlap among species. However, the niche model retains much of the simplicity and tractability embodied by the earlier model. As in the cascade model, the niche model has two tunable parameters that determine the number of species and links. In the case of links, instead of assuming constant $L/S$ (Cohen and Newman 1985b), Williams and Martinez used directed connectance $C$ ($L/S^2$) as the link parameter, allowing it to vary (Cohen et al. 1985). Thus, Williams and Martinez (2000) made no assumption about link-species scaling, and did not replace it with any other hypothesis about the relationship between $L$ and $S$. However, their choice of $C$ reflected the notion that $C$ is independent of $S$ (Martinez 1992, 1993a) just as Cohen and Newman’s (1985b) choice of $L/S$ reflected the notion that $L/S$ is independent of $S$ (Cohen and Briand 1984). Beyond the two input parameters $S$ and $C$, the niche model, like the cascade model, orders species along a single dimension. However, instead of the cascade model’s simple rank-ordering of species, the niche model assigned each species a uniformly random niche value along a line, and that niche value corresponds to the position of each species on the line. To distribute links, each species is assigned a feeding range that represents an interval of the line whose midpoint is a uniformly random number less than the niche value of the species possessing the range. All species that fall in this range are eaten by the consumer species (Box 1). Feeding range sizes are drawn randomly from a beta distribution to produce a $C$ close to the target $C$ (Williams and Martinez 2000).

The distribution of feeding links in the niche model leads to several outcomes: (1) cannibalism and feeding on species with higher niche values can occur, (2) there is higher trophic overlap than in purely stochastic link distribution schemes, since species with similar niche values are more likely to share consumers, and (3) food webs are rendered interval due to contiguous feeding by each consumer on resources within a single range or segment of the line. The first two outcomes, along with using variable $C$ as a parameter, address the three main limitations of the cascade model. The third outcome represents an acknowledged limitation of the niche model (Williams and Martinez 2000), since diverse empirical food webs are known to not be interval (Cohen and Palka 1990). Williams and Martinez (2000) argued that intervality is a delicate property that is easily broken, often by the loss of just one link in a web. They hypothesized that quantitative measures of intervality will show that the degree of intervality, rather than just the presence of intervality, is quite high in empirical food webs.

In addition to these model innovations, Williams and Martinez (2000) also introduced the use of numerical simulations to compare statistically the ability of the niche model and alternate network models to fit empirical food-web
data. They standardized the form of multiple models by creating versions of the cascade and a random model that use variable $S$ and $C$ as input parameters. This ensures that the three models vary only in how they distribute links among species (Box 1). Because of stochastic variation in aspects of how species and links are distributed in any particular model web, analysis begins with the generation of hundreds to thousands of model webs with the same $S$ and similar $C$ as an empirical food web of interest. Model webs that fall within 3% of the target $C$ are retained. Species that are either disconnected (i.e., species that lack links) or trophically identical to other species, or webs that contain such species, are eliminated and replaced. The exclusion of trophically identical species means that model webs are most fruitfully compared to trophic-species versions of empirical food webs, rather than original-species webs, unless most or all of their species are already trophically distinct. Additionally, in the cascade and niche models, no prey are assigned to the species with the lowest niche value, ensuring there is at least one basal species per web. Once a set of model webs is generated, model means and standard deviations are calculated for each food-web property of interest, which can then be compared to empirical values. Raw error, the difference between the value of an empirical property and a model mean for that property, is normalized by dividing it by the standard deviation of the property’s simulated distribution. This approach allows assessment not only of whether a model over- or underestimates empirical properties as indicated by the raw error, but also to what degree a model’s mean deviates from the empirical value. Normalized errors within $\pm 2$ are considered to indicate a good fit between the model prediction and the empirical value (Williams and Martinez 2000).

All three models were evaluated by how well they fit up to 12 structural properties of seven empirical community food webs drawn from the expanding set of improved food-web datasets, with 25 to 92 trophic species and 0.061 to 0.32 connectance (Williams and Martinez 2000). The properties include old standbys such as $T$, $I$, and $B$ as well as other properties of interest in food-web research (proportions of omnivores, cannibals, and species in loops; chain length properties; trophic similarity; variation of generality and vulnerability). Link proportions were excluded due to their strong correlation with $T$, $I$, and $B$. In summary, the random model performed poorly, with an average normalized error (ANE) of 27.1 ($SD = 202$), the cascade model performed much better, with an ANE of $-3.0$ ($SD = 14.1$), and the niche model performed an order of magnitude better than that, with an ANE of 0.22 ($SD = 1.8$). Only the niche model falls within $\pm 2$ ANE and is considered to show a good fit to the data. Not surprisingly, there is variability in how all three models fit different food webs and properties. For example, the niche model generally overestimates food-chain length. Specific mismatches are generally attributable either to limitations of the models or biases in the data. A separate test of the niche and cascade models with three marine food webs, a type of habitat not included in the original analysis, obtained similar results (Dunne et al. 2004). That test added four properties to the analysis: proportion of herbivores, mean trophic level, clus-
tering coefficient, and characteristic path length. The latter two properties are borrowed from “small-world” network research and will be discussed more below (see Small-World Properties and Degree Distribution section). The niche model’s performance on the new properties is similar to previous properties. For example, the niche model tends to systematically underestimate the proportion of herbivores, which partly explains its overestimation of food-chain length.

Like the cascade model, but taking into account improved food-web data, these niche model studies demonstrate that the structure of food webs is far from random, and that simple link distribution rules can yield apparently complex network structure, comparable to that observed in empirical data. In addition, like the cascade model, the niche model is simple enough that it is analytically solvable, leading to theoretical predictions similar to trends in numerical simulations and empirical data (Camacho et al. 2002a,b). The hypothesis of scale dependence suggests that there are not simple generalities or constant values of network properties that hold across food webs with varying S. However, the concordance of the niche model with empirical data suggests that once variable S and C are taken into account, there appear to be universal coarse-grained characteristics of how trophic links and species (defined according to trophic function) are distributed within food webs across a wide array of habitats. Of course, the current success of the niche model needs to be taken in context, since, like the cascade model it orients itself to data that are flawed and limited. New data and better understanding about key processes, properties, and interactions in food webs, or more broadly, ecological networks, may lead to more efficacious approaches or to the rejection of any general model. Martinez et al. (Chapter 6) and others have already proposed some variations on the niche model as discussed below.

4.2 NESTED-HIERARCHY MODEL

Another simple topological model similar to the niche and cascade models, the nested-hierarchy model, addresses the intervality limitation of the niche model (Box 1, Cattin et al. 2004). Following methods introduced by Williams and Martinez (2000), Cattin et al. use S and C as input parameters, distribute species niche values randomly along a line, and stochastically assign the number of prey items for each species using a beta distribution. They then distribute links, starting with the species with the lowest niche value, in a way that (1) avoids creating interval webs, (2) generates webs with trophic overlap (Solow and Beet 1998; but see Stouffer et al. 2005)), and (3) allows a low probability of looping. First, a link is randomly assigned from a consumer species i to another species j with a lower niche value. If that resource species j is also fed upon by other species, the consumer species i’s next feeding link is randomly selected from the pool of resource species of a set of consumer species defined as follows: they share at least one prey species, and at least one of them feeds on resource species j. If more feeding links are required, links are randomly assigned to species without
predators and with lower niche values. If yet more feeding links are needed, links are randomly assigned to species with equal or higher niche values.

Cattin et al. (2004) compared the fit of this model and the niche model to the 7 webs and 12 properties analyzed by Williams and Martinez (2000) plus two additional properties that reflect aspects of intervality. They compared model means and standard deviations with empirical values, but did not calculate normalized error. They concluded that the two models perform comparably on the original 12 properties, but the nested-hierarchy model does better for the two intervality properties. This is not surprising since the niche model necessarily returns 0 values for the intervality properties. A re-analysis of Cattin et al.’s results illuminates some overstated or incorrect claims in their abstract, namely that the nested-hierarchy model “better reflects the complexity and multidimensionality of most natural systems,” and that the niche model fails “to describe adequately recent and high-quality data.” Martinez and Cushing (Box A) suggest that Cattin et al. base their claims on selective favoring of intervality, and gloss over other details of fit. Indeed, these types of analyses will always be influenced by how much importance is ascribed to particular properties (Cattin et al. Box B). However, any slight difference in the performance of the niche and nested-hierarchy models, which could be assessed more rigorously using normalized error, is not comparable to the order of magnitude improvement of the niche model over the cascade model, the quantitative assessment of which did not even include the cascade model’s failure to generate cannibalism or other looping (Williams and Martinez 2000). Indeed, the fit of the nested-hierarchy model to intervality properties may come at the expense of reduced fit to other properties (Martinez and Cushing Box A). Another way the niche model could be modified to break intervality is by generating feeding ranges that are slightly larger, and by making the probability that the consumer species eat species within their feeding ranges slightly less than 100% (R. J. Williams, personal communication). However, all previous attempts to modify the niche model based on ecological understanding to improve its overall fit to data failed (unpublished data, Williams and Martinez).

4.3 GENERALIZED ANALYTICAL MODEL

An analytical study building on Camacho et al. (2002a,b) provides a compelling reason why the nested-hierarchy model does not generally improve on the niche model. Although the nested-hierarchy model “appears to be quite different in its description, it nevertheless generates webs characterized by the same universal distributions of numbers of prey, predators, and links” (Stouffer et al. 2005, original emphasis). Stouffer et al. found that only two conditions must be met for network models to reproduce several central properties of currently available improved food-web data: (1) species niche values form a totally ordered set, and (2) each species has a specific probability, drawn from an approximately exponential distribution, of preying on species with lower niche values. The first
Box 1: Models of Food-Web Structure

Four simple, stochastic models that have been proposed to generate and predict the network structure of empirical food webs are described. The models share two empirically quantifiable input parameters: (1) $S$, the number of species or taxa in a web, and (2) $C$, connectance, a metric determined by the number of links and species in a food web. There are $S^2$ possible and $L$ actual links in a particular food web, and directed connectance $C$ is defined as $L/S^2$, or the proportion of possible links that are actually realized. The models differ in the rules they use to distribute links among species, as follows:

- **Random Model** (inspired by Erdős and Rényi 1960; see also Cohen 1977b)
  Any link among $S$ species occurs with the same probability $P$ equal to $C$. This creates food webs as free as possible from biological structuring.

- **Cascade Model** (modified from Cohen and Newman 1985b)
  Each species is assigned a random value drawn uniformly from the interval $[0, 1]$. Each species has the probability $P = 2CS/(S - 1)$ of consuming species with values less than its own. This creates a feeding hierarchy and disallows cannibalism and feeding on species higher in the hierarchy, as illustrated in the following diagram.

  ![Diagram of Cascade Model]

  This formulation is a modified version (Williams and Martinez 2000) of the original cascade model (Cohen and Newman 1985b), which tuned $L/S$ (link density) to the data by using an average across empirical webs to look at scaling patterns generated by the model. This assumed constant $L/S$ and ensured that $C$ declined hyperbolically with $S$, as suggested by early theory and data. Recent studies use variable $C$ as the parameter that constrains the number of links in model webs, employing a similar approach to Cohen et al. (1985) who tuned $L/S$ to values for particular webs.

- **Niche Model** (Williams and Martinez 2000)
  As in the modified cascade model, each species is assigned a random value drawn uniformly from the interval $[0, 1]$, referred to as the species’ niche value, $n_i$. Each species consumes all species within a range of niche values $r_i$. The size of $r_i$ is randomly assigned using a beta function, producing a $C$ close or identical to the target $C$. The center of the range $c_i$ is drawn uniformly from the interval $[r_i/2, n_i]$ or $[r_i/2, 1 - r_i/2]$ if $n_i > 1 - r_i/2$. This keeps all of the feeding range within $[1, 0]$ and places the center of a species’ range lower than its niche value.

continued on next page
The Network Structure of Food Webs

Box 1 continued

As shown in the diagram below, the $c_i$ rule relaxes the strict feeding hierarchy of the cascade model by allowing up to half of the feeding range to include species with niche values $\geq n_i$, thus permitting cannibalism and feeding on species with higher niche values.

Species $i$ feeds on all four species, shown in grey, within its feeding range $r_i$. This includes a cannibalistic link to itself and a link to a species with a higher niche value.

- Nested-Hierarchy Model (Cattin et al. 2004)

Following the cascade and niche models, each species is assigned a random niche value $[0, 1]$. Like the niche model, the number of prey items for each species is drawn randomly from a beta distribution that constrains $C$ to be close to the target. Feeding links are assigned in a multistep process. First, a link is randomly assigned from species $i$ to a species $j$ with a lower niche value. If that prey species $j$ is also fed upon by other species, the next feeding link for species $i$ is selected randomly from the pool of resource species fed on by a set of consumer species defined as follows: they share at least one prey species, and at least one of them feeds on species $j$. If more feeding links are required, links are randomly assigned to species without predators and with niche values $< n_i$. If more feeding links are required, links are randomly assigned to species with niche values $\geq n_i$. Thus, the model relaxes the contiguous feeding, and thus intervality, of the niche model.

The food web for St. Marks Estuary (Christian and Luczkovich 1999), and an example of four types of model webs with the same $S$ and $C$ as the empirical web. Images produced with FoodWeb3D, written by R. J. Williams and available at www.foodwebs.org.

condition was met by the cascade model (Cohen and Newman 1985b) and is equivalent to their rule that species are ordered along a single dimension. The niche model adopted that rule by assigning each species a unique niche value drawn randomly from the interval $[0, 1]$. The second condition was newly em-
bodied in the niche model through its use of a beta distribution as the means by which the size of feeding ranges is distributed (Williams and Martinez 2000). Stouffer et al. (2005) note that the beta distribution is a type of exponential distribution. They show that the specific form of the exponential distribution does not alter predictions about the distributions of numbers of prey and predators, nor does it alter a number of network structure properties derivable from those two distributions. Their focus on numbers of prey and predators as two fundamental, non-identical distributions relates back to Schoener’s (1989) work discussing the importance of understanding differences in vulnerability (number of predators) and generality (number of prey) for developing simple models of food-web structure, differences reflected in improved empirical data and the niche model (Williams and Martinez 2000).

Ordered niche values and the beta distribution (Williams and Martinez 2000) were adopted by the nested-hierarchy model (Cattin et al. 2004). The nested-hierarchy model has different link distribution rules that are meant to mimic phylogenetic constraints. However, Stouffer et al. (2005) point out that Cattin et al.’s distribution rules ensure that a species is assigned prey essentially randomly from the set of species with lower niche values, as constrained by the beta distribution. As a result, the nested-hierarchy model returns the same distributions of numbers of predators, prey, and links among species as does the niche model. Stouffer et al. (2005) test their hypothesis about the centrality of the two conditions by modifying the cascade model to meet the second condition. This modified cascade model also produces the same general analytical expressions, or universal functional forms, for distributions as the niche and nested-hierarchy models. Based on their analysis, Stouffer et al. consider the nested-hierarchy and modified cascade models to be “randomized” versions of the niche model, which was the first model to embody the two fundamental conditions they identify as central to model representations of empirical food-web network structure.

Several common food-web properties previously used to assess particular models (Williams and Martinez 2000; Cattin et al. 2004) can be derived from the analytical expressions for the distributions of numbers of prey and predators, including fractions of top, basal, and cannibalistic species, and standard deviations of vulnerability and generality (Stouffer et al. 2005). Other derivable properties of potential interest include the correlation coefficient between each species’ number of prey and predators, and assortativity, the probability with which nodes with high degree (many links) link to other high-degree nodes (Newman 2002). Food webs tend to be negatively assortative, or disassortative (Newman 2002; Stouffer et al. 2005). With reference to this set of properties, and drawing on 15 of 19 improved food webs described in Dunne et al. (2002ab, 2004), Stouffer et al. (2005) find that empirical values for 11 of the 15 food webs are well described by the model’s analytical expressions and numerical simulations. Stouffer et al. (2005) also added the potential for cannibalism (i.e., feeding on species with equal niche values) to the modified cascade model to make it even more comparable to the other two models. However, allowing for a low
probability of cannibalism, or a low probability of feeding on species with higher niche values as in the niche and nested-hierarchy models, does not alter analytical predictions of predator and prey distributions and thus is not included in condition two (Stouffer, personal communication).

What Stouffer et al. (2005) do not address are other aspects of the models and data, particularly with regard to looping, intervality, omnivory, herbivory, food-chain statistics, trophic level, and other properties of ecological interest (Williams and Martinez 2000, 2004b; Dunne et al. 2004a,b; Cattin et al. 2004). While they appear to have identified fundamental aspects of species and link distributions that underlie and emerge from currently successful models of food-web network structure, not all properties of ecological interest are derivable from distributions of numbers of predators and prey. Differences and similarities among particular models with respect to such properties are not addressed by the general analytical model (Stouffer et al. 2005). For example, the nested-hierarchy and modified cascade models allow for non-interval webs, and the niche and nested-hierarchy models allow for feeding on species with higher niche values. These particular aspects of the models may not significantly impact the overall distribution of numbers of predators and prey, but will affect how well the models capture other quantifiable and ecologically interesting variability in food-web network structure. Also, while Stouffer et al. (2005) have demonstrated core conditions and universal functional forms for some aspects of food-web structure that emerge out of the niche model, how well those conditions and functional forms continue to fit the data depends on evolving data availability and standards. Access to more comprehensive data that is more highly and evenly resolved, or a switch in focus to other more ecologically compelling ways to slice and dice data, may necessitate the development of some other approach to modeling ecological network structure. Whether and how the simple rules that appear to generate food-web-like topology actually connect back to ecological, evolutionary, thermodynamic or other principles remains a wide-open area of inquiry. This, more than anything, is likely to impact how ecological network data is modeled in the future.

4.4 SMALL-WORLD PROPERTIES AND DEGREE DISTRIBUTION

The resurgence of interest in the late 1990s across disciplines in describing general properties of the network structure of everything from social groups to WWW page links to power grids to transportation systems to metabolic pathways to scientific citations, brought to light a few topological properties that had not previously been explicitly evaluated for food webs. It also brought a new question to the table: do food webs have similar topology to other biotic, social, and abiotic networks? A paper published by Watts and Strogatz (1998) led the resurgence by bringing the notion of “small-world” network structure to the foreground. They suggested that most real-world networks look neither like randomly connected graphs (Erdős and Rényi 1960) nor regularly connected lattices in which every
node has the same number and pattern of links. Instead, real-world networks appear to combine aspects of both—they have high clustering, like regular lattices, but short paths between nodes, like random graphs. These features are typically expressed by two properties: clustering coefficient, or the average fraction of pairs of nodes connected to the same node that are also connected to each other, and characteristic path length, or the average shortest distance between pairs of nodes. Work initiated by Barabási and colleagues (as reviewed in Albert and Barabási 2002) suggested that most real-world networks also display power-law degree distributions, which refers to the distribution of the number of links per node. Regular lattices display a constant distribution of links among nodes, while random graphs display a Poisson distribution of links among nodes. Most empirical networks appear to display a highly uneven power-law or scale-free distribution of links among nodes, with most nodes having few links and a few nodes having a very large number of links.

Several papers published in 2002 considered the question of whether empirical food webs display small-world, scale-free network structure similar to many other real-world networks (also reviewed in Cartozo et al. Chapter 3). Using original species versions of three relatively diverse and well-resolved webs, Montoya and Solé (2002) suggested that food webs do tend to display small-world, scale-free structure, although the Little Rock Lake food web did not fit those patterns well. Looking across trophic-species versions of the seven community webs analyzed by Williams and Martinez (2000), Camacho et al. (2002b) contradicted Montoya and Solé (2002) by concluding that clustering coefficients of food webs appear similar to random expectations, less than the clustering observed in small-world networks. Using cumulative rather than density distributions due to the noisiness of the data (Amaral et al. 2000), Camacho et al. (2002b) also concluded that food webs do not display scale-free distributions of links, irrespective of whether total links, number of predators, or number of prey are considered. However, when the distributions are normalized for linkage density by dividing the number of links by $2L/S$, empirical food-web link distribution data appears to display universal functional forms (Camacho et al. 2002b). For example, cumulative degree distribution, assessed using data pooled across six of seven food webs, shows a systematic exponential decay in its tail. This type of distribution, less skewed than a power-law distribution, has been observed in a few other “real-world” networks (Amaral et al. 2000). Clustering coefficient and path length also appear to follow universal functional forms that scale with linkage density (Camacho et al. 2002b). Numerical and analytical predictions of the niche model (Williams and Martinez 2000) fit empirical data quite closely for clustering coefficient, characteristic path length, and distributions of numbers of predators, prey, and links (Camacho et al. 2002b; Williams et al. 2002; Dunne et al. 2004; Stouffer et al. 2005).

Dunne et al. (2002a) attempted to resolve differences between Montoya and Solé (2002) and Camacho et al. (2002b) by examining a larger array of 16 trophic-species food webs, including those analyzed in the other two studies. Corroborat-
The Network Structure of Food Webs

ing aspects of Camacho et al. (2002b), Dunne et al. found that most food webs display low clustering coefficients and non-power-law degree distributions, in particular less skewed exponential and uniform distributions. However, they also found that webs with very low connectance (e.g., the Scotch broom source web with $C = 0.03$) were more likely to display both higher-than-random clustering and power-law degree distributions, consistent with the small-world, scale-free structure of many other types of networks, and as reported for food webs by Montoya and Solé (2002). Using linkage density normalization to overlie cumulative degree distributions of the 16 webs, Dunne et al. (2002a) concluded there was too much variation in the data to support the notion of a universal functional form (Camacho et al. 2002b; Stouffer et al. 2005), and pointed out that such variation can be masked by pooling the data and excluding datasets that don’t fit the pattern well. However, the data are obviously constrained within a region that is not power-law in its form. Both the tendency for much, but not all, improved food-web data to converge on universal functional forms that scale with linkage density (Stouffer et al. 2005), as well as potentially systematic deviations from or variability around those central tendencies (Dunne et al. 2002a), are important research issues that can reveal interesting insights at different levels of analysis. For example, while the niche model produces network structures that have exponential degree distributions (Camacho et al. 2002b), individual empirical webs can show other distributions, particularly uniform distributions (Dunne et al. 2002a). A simple model that starts with a randomly linked “regional” pool of species, and then creates “local” food webs via random immigration from the regional pool coupled with random extinctions from the local web, produces significant percentages of webs with exponential as well as uniform degree distributions (Arii and Parrott 2004). This study highlights simple assembly mechanisms that can produce variable degree distributions, mechanisms that may be relevant for empirical webs.

All studies looking at small-world structure in food webs (Montoya and Solé 2002; Camacho et al. 2002b; Dunne et al. 2002a, 2004; Williams et al. 2002) have reported short path lengths similar to random expectations (i.e., “two degrees of separation”; Williams et al. 2002), consistent with one aspect of small-world structure. However, apart from path length, most currently available food-web data clearly deviate from the small-world, scale-free topology observed for other biotic and abiotic networks (Camacho et al. 2002b; Dunne et al. 2002a). Compared to other networks, food webs have low diversity and high connectance, which appear to be drivers of clustering coefficient and degree distribution patterns. The ratio of observed clustering coefficient to the random expectation, which is 1 for many food webs, scales approximately linearly with network size among a wide range of biological, social, and technological networks (Dunne et al. 2002a). Degree distribution is related to connectance, with networks that show power-law distributions being much more sparsely connected (i.e., lower $C$) than most food webs. While many empirical food webs display exponential
degree distributions, higher connectance webs often display less skewed uniform distributions (Dunne et al. 2002a).

4.5 UNIVERSAL PATTERNS?

In addition to the question of whether food webs display small-world, scale-free topology, researchers have considered other ways of identifying general topological patterns in binary-link food webs, often borrowing methods from other types of network research. Research on scale-dependence of empirical food-web properties, the niche model and its offshoots, and analyses of small-world structure and degree distribution all suggest that food webs generally do not have scale-invariant patterns conventionally understood to be “universal.” Instead, food webs tend to display scale-dependent patterns that can be characterized, depending on how much variability is ignored, as universal functional forms (Camacho et al. 2002b). These scale-dependent patterns or functional forms emerge once data from different webs are normalized for link density, \(2L/S\), the relationship between the number of links \((L)\) and species \((S)\) in a web (Camacho et al. 2002b; Dunne et al. 2002a; Stouffer et al. 2005).

However, other studies suggest that there are universal topological patterns in empirical food webs that hold regardless of \(S\), \(L\), or their relationship. Milo et al. (2002) developed an algorithm for detecting network motifs based on the statistical approach of Williams and Martinez (2000). They define network motifs as “recurring, significant patterns of interconnections.” For a variety of biotic and abiotic networks, they identified and counted all possible configurations of three- and four-node subgraphs, and then compared the frequency of different subgraphs in empirical networks to their frequency in comparable randomized networks. Particular subgraph types are considered motifs when they occur significantly more often than expected for random webs. Milo et al. (2002) analyzed the seven food webs from Williams and Martinez (2000) and found that five of seven share a three-node motif referred to as three-chain or a three-species food chain. Additionally, all seven share a four-node motif referred to as bi-parallel in which two species share a common predator and prey. The food webs do not share the three-node motif with other types of networks, but do share their four-node motif with the \(C.\) elegans neuronal network and all five electronic circuit networks considered. The presence of significant motifs in food webs appears to be independent of network size \((S)\), but the frequency of motifs in food web and other networks appears to grow linearly with size, unlike the frequency in randomized networks. Milo et al. (2002) speculate that motifs may be interpreted as “structures that arise because of the special constraints under which the network has evolved.”

Food webs have also been investigated as a type of transportation network, using methods applied to river basins and vascular systems (see Cartozo et al. Chapter 3 for detailed discussion). Garlaschelli et al. (2003) decompose seven food webs with 42 to 123 trophic species into minimal spanning trees. A minimal
spanning tree is a simplified version of a network that is created by removing links to minimize the distance between nodes and some destination. One way to create a minimal spanning tree of a food web is to add an additional environment node to which all basal species link, and then to trace the shortest food chain from each species to the environment node. Links that are not a necessary part of any of these shortest food chains (e.g., cannibalism, and other links in loops) are excluded from the tree. Garlaschelli et al. (2003) analyzed the allometric scaling of these trees, or how their branching properties change with network size. They found that all seven food webs display a power-law, scale-free relationship with an exponent of 1.13, although the three smallest webs display marginally larger exponents. This suggests that minimal spanning trees may successfully characterize a universal core structure in food-web networks. However, the “universality” of the exponent for food-web minimal spanning trees has been called into question by an analysis of a broader set of 17 food webs, which display exponents ranging from 1.09 to 1.26 (Camacho and Arenas 2005). The short range of the exponent values is attributed to the relatively small mean trophic level of most food webs. The particular exponents for food-web minimal spanning trees suggests that they can transport resources more efficiently across the whole network than river or vascular systems, which display higher exponents. Within a reasonable range of connectance (0.05 to 0.3), the niche model (Williams and Martinez 2000) underestimates empirically observed exponents, returning values of 1.06 to 1.08 (Garlaschelli et al. 2003).

Other approaches for identifying universal network structure have yet to be applied to food webs. For example, Song et al. (2005) found that the network structure of a variety of real-world networks obeys power-law scaling, as if they are fractal shapes. This self-similarity was found to apply to the relationship between the number of boxes (i.e., sub-groups of connected nodes) needed to cover a network and the size of the box. Song et al. (2005) also used a renormalization procedure to coarse-grain the networks by sequentially collapsing boxes into single nodes and then creating new boxes. These aggregated networks fit the same power-law scaling as individual-node networks. Given that food webs differ in some basic ways from many other empirical networks, it will be interesting to see if this self-similarity applies to ecological networks. While analysis of food-web data may be inconclusive, given how small they are relative to most other networks studied, food-web models such as the niche model could be used to explore higher diversity webs as well as any sensitivity to changes in diversity and connectance.

4.6 NETWORK STRUCTURE AND ROBUSTNESS OF FOOD WEBS

As the previous sections highlight, research on general or universal aspects of food-web network structure underwent a renaissance at the beginning of the 2000s as a result of improved data, new topological models, and concepts and approaches borrowed from general network theory. Network approaches also
brought back into focus the question of how food-web structure might relate to issues of core ecological interest, such as ecosystem stability. Within ecology during the 1990s, questions about stability were increasingly transformed into questions about ecosystem responses to perturbations and the relationship between ecosystem complexity, especially diversity, and ecosystem function (McCann 2000). Classic research into connectance and how it relates to May's stability criterion (May 1973) was largely abandoned, as foreseen by Paine (1988). However, in a new introduction to the 2001 Princeton Landmarks in Biology Edition of his 1973 book, May wrote, “...the theme of the relationship between the network structure of food webs and their ability to handle perturbation is central in ecology, as in many other subjects. ...The reorientation of this question to what kinds of connectance patterns are likely to be most resistant to specific kinds of disturbance is of continuing relevance in ecology, as elsewhere.”

May’s (2001) comments were inspired partly by emerging research in the broader arena of network theory. About that time, a series of papers examined the response of a variety of networks including the Internet and WWW pages (Albert et al. 2000) and metabolic and protein networks (Jeong et al. 2000, 2001) to the simulated loss of nodes. In each case, the networks, all of which display highly skewed power-law degree distributions, appear very sensitive to the targeted loss of highly connected nodes but relatively robust to random loss of nodes. When highly connected nodes are removed from scale-free networks, the average path length tends to increase rapidly, and the networks also quickly fragment into isolated clusters. In essence, paths of information-flow in highly skewed networks are easily disrupted by the loss of nodes that are directly connected to an unusually large number of other nodes. In contrast, random networks with much less skewed Poisson degree distributions display similar responses to the targeted loss of highly connected nodes versus random node loss (Strogatz 2001).

Within ecology, species deletions on small (S < 14) hypothetical food-web networks as well as a subset of the 113 web catalog were used to examine the reliability of network flow, or the probability that sources (producers) are connected to sinks (consumers) in food webs (Jordán and Molnár 1999). They concluded that the structure of the empirical webs appeared to conform to reliable flow patterns identified using the hypothetical webs, but that result was based on the early poorly resolved data. Following Albert et al. (2000), Solé and Montoya (2001) used three improved, diverse food webs to conduct species knockout simulations. Instead of path length, Solé and Montoya looked at the level of secondary extinctions potentially triggered by different patterns of primary species loss. This is easily measured in binary food-web networks using the simple algorithm that if primary extinctions cause a consumer to lose all of its resources, it too goes extinct. In all three food webs, removal of highly connected species resulted in much higher rates of secondary extinctions than random loss of species, and also fragmented the webs more rapidly, similar to results seen for other types of networks (e.g., Albert et al. 2000). Solé and Montoya (2001) attributed this to highly skewed power-law degree distributions (Montoya and Solé 2002). How-
ever, most food webs do not have power-law link distributions (Camacho et al. 2002b; Dunne et al. 2002a), so the generality of those results was unclear. Also, the web that showed extreme fragility to the loss of highly connected species was a detailed source web based on only one basal species, Scotch broom (Memmott et al. 2000). In this web, Scotch broom has a large number of species linking to it and, as a result, is removed very early in the non-random deletion sequence, leading to the necessary collapse of the entire web. That collapse is attributable to the particular and peculiar characteristics of the Scotch broom dataset, and is not good evidence for a general trend in community food webs (Jordán 2002; Dunne et al. 2002b).

To address these issues, a set of 16 non-marine food webs (Dunne et al. 2002b) and 3 marine food webs (Dunne et al. 2004) was used for similar biodiversity loss simulations. It was found that even without highly skewed degree distributions, food webs are much more robust to random loss of species than to loss of highly connected species. These results suggest that any substantial skewness in degree distribution will tend to alter the response of a network to different kinds of node loss. Similarly, the order of pollinator loss was found to have an effect on potential plant extinction patterns in two detailed, speciose plant-pollinator networks (Memmott et al. 2004). Loss of plant diversity associated with targeted removal of highly connected pollinators was not as extreme as comparable secondary extinctions in food webs, which Memmott et al. (2004) attribute to pollinator redundancy and the nested topology of the networks.

The previous studies all point to the trend that sequential loss of highly connected species has a greater impact than random losses. However, the “knock-out highly connected species” approach is not necessarily useful for identifying particular species likely to have the greatest impact: loss of a particular highly connected species may or may not result in a large number of secondary extinctions. To address this, Allesina and Bodini (2004) used a dominator tree approach to reduce the topological structure of 13 empirical food webs into linear pathways that define the essential chains of energy delivery in the network. A particular node dominates another node if it passes energy to it along a chain in the dominator tree. In addition to corroborating prior findings of higher secondary extinctions with targeted loss of species (in this case, the loss of species that dominate many other species) versus random losses, Allesina and Bodini (2004) showed that the higher the number of species that a particular species dominates, the greater the secondary extinctions that result from its removal.

Dunne et al. (2002b, 2004) provide yet another answer to May’s (2001) question about patterns of connectance and ecosystem responses to perturbation. They found that food-web robustness, defined as the fraction of primary species loss that induces at least 50% total species loss (primary + secondary extinctions) for a particular trophic-species web, increases with increasing connectance, $L/S^2$. This holds both for random species loss and targeted removal of highly connected species. While systematic removal of species with few links generally leads to low levels of secondary extinctions, there are exceptions, and there is no
obvious correlation of those exceptional responses with global food-web properties such as $S$ or $C$ (Dunne et al. 2002b). Such apparently idiosyncratic effects of the loss of least-connected species probably have to do with finer-level patterns of link distribution within particular webs. One possibility is that such species “dominate” many other species within a dominator-tree context (Allesina and Bodini 2004), even though they have few direct links to other species. Or such effects may simply be hard to assess a priori using standard measures of network topology. Unpredictability of species’ likely importance for extinction dynamics was reported in a dynamical food-web modeling study on the resistance of communities to non-random extinctions, based on species’ sensitivity to a theoretical stressor (Ives and Cardinale 2004).

While the previous species removal studies (Jordán and Molnár 1999; Solé and Montoya 2001; Dunne et al. 2002b, 2004; Memmott et al. 2004; Allesina and Bodini 2004) are informed by a purely network-structure perspective that ignores dynamics, it has been demonstrated repeatedly that constraints imposed by structure can have a significant role in the outcomes of dynamics (e.g., Pimm and Lawton 1978; McCann and Hastings 1997; Jordán et al. 2002; Martinez et al. Chapter 6; see also the review by Jordán and Scheuring 2004). Ideally, such in silico biodiversity loss and related species invasion experiments will be conducted using approaches that integrate structure and dynamics (see Integrating Structure and Dynamics section). Even without explicit dynamical analysis, food-web topology research shows that more than 95% of species are within three links of each other, and that species draw ever closer as connectance and species richness increase (Williams et al. 2002). This suggests that the dynamics of species in complex ecosystems are more tightly connected than conventionally thought, which has profound implications for the impact and spread of perturbations.

5 RELATED TOPICS AND FUTURE DIRECTIONS

As the previous sections highlight, a great deal of interesting research on general aspects of network structure of complex food webs has occurred since the first 30-web catalog was analyzed and published (Cohen 1977a,b, 1978). Such research is now thriving and has found a broader context in interdisciplinary network research, following a lull during the 1990s when much of the focus was on systematically exploring the impacts of methodology on patterns of food-web properties. By limiting this review to research explicitly focused on potential generalities in complex food-web network structure, I have neglected a number of exciting topics that relate to food-web topology, many of which provide more of an ecological context for such research. It is from these additional areas that many of the future directions of structural food-web research are likely to emerge. I wrap up the chapter with a brief discussion of a few of many possible related topics.
66  The Network Structure of Food Webs

5.1 VARIABLE STRUCTURE ACROSS ECOSYSTEMS AND/OR ENVIRONMENT

Since the beginning of food-web structure research, some studies have looked for systematic differences among types of webs. For example, early data suggested that food webs from fluctuating environments have lower connectance than those from more constant environments (Briand 1983), and that three-dimensional habitats have longer food chains than more two-dimensional habitats such as forest canopy or grassland (Briand and Cohen 1987). Differences in food webs across broad categories of habitat have also been considered, for example aquatic versus terrestrial webs (Chase 2000), and marine versus “continental” webs (Cohen 1994; Link 2002; Dunne et al. 2004). Unfortunately, there is currently not enough high-quality data, representing multiple webs from a variety of ecosystems, to be more than suggestive in any assessment of systematic changes in structure across environment or habitat (Dunne et al. 2004).

In addition, any comparisons of the details of food-web structure need to be mediated by the understanding that species richness and connectance vary across webs in ways that systematically impact structure. Unless food webs are constructed with the same methodology, which is almost impossible to standardize across different habitats, such methodological variation may result in inconsistent levels of sampling effort and thresholds for link inclusion. This can make it difficult to discern systematic differences in food-web structure attributable to ecological or environmental processes. The niche model or other models that use $S$ and $C$ as inputs can help with this problem by acting as “benchmarks” for assessing similarities and differences between food webs with varying $S$ and $C$ (Dunne et al. 2004). The benchmark approach works by helping to factor out variation due to species and/or links being differentially observable among habitat types. $S$ and $C$ may also vary due to ecological or other non-methodological reasons. How and why they vary beyond methodology has important implications for food-web structure and dynamics and is a basic question for all of ecology. For example, a source of variation in the relationship between $S$ and $C$ has been suggested by the model of Solé et al. (2002), where the immigration probability of new species into the system appears to be critical. This suggests that the degree of openness of the system will influence the relationship in systematic ways (Pascual et al. Chapter 15).

5.2 COMPARTMENTS

The presence or lack of compartmentalization may have implications for the transmission of both beneficial and harmful effects throughout food webs, and thus can affect the stability and robustness of ecosystems. Early food-web research suggested that there is little detectible compartmentalization in food webs (Pimm and Lawton 1980) although other studies using early data suggested otherwise (Yodzis 1982; Raffaelli and Hall 1992). High connectance in food webs, compared to many other types of networks studied, should tend to obscure
compartmentalization, as indicated by low clustering coefficients (Amaral et al. 2002b; Dunne et al. 2002a) and short path lengths (Williams et al. 2002). However, some recent studies suggest that there are ways to identify compartments in detailed food-web data. For example, Girvan and Newman (2002) borrowed a “centrality” or “betweenness” index common to social network research, applied it to links, and successfully detected pelagic and benthic subcommunities of species within the Chesapeake Bay food web (Baird and Ulanowicz 1989) using only network structure data. Also borrowing from social network analysis, Krause et al. (2003) used a clustering algorithm (“KliqueFinder”) which identifies subgroups with concentrated interactions (Frank 1995). They found compartmentalization in three of five complex food webs, but the ability to find compartments depended on inclusion of flow data. They generally did not detect compartments in less complex food webs with or without flow data. Allesina et al. (in press) identified multiple strongly connected components (SCCs) in 17 food webs, where an SSC is a set of species interconnected by cycles or loops. However, they found that the number of SSCs identified is very sensitive to removal of weak links (Winemiller 1990).

A promising algorithm for systematically identifying a type of food-web compartment in any size web using only structural data was developed by Mélián and Bascompte (2004). A \( k \) subweb is defined as a subset of species that are connected to at least \( k \) predator and/or prey species within that subset. Each species belongs to only one subweb, the subweb where each species has the highest \( k \) value. In effect, \( k \) is the measure of links per species, or linkage density, within the context of the subweb. Any particular species can have a higher absolute number of links if connected to species in other subwebs. Mélián and Bascompte (2004) assessed the frequency distribution of subweb size in five speciose webs with \( S \) of 134 to 237 and found that all follow a power-law distribution, with zero or few links per species in most subwebs and a unique subweb with the most links per species. The scaling exponents increase with web size and vary from \( -1.87 \) to \( -0.65 \). Mélián and Bascompte (2004) also calculated the connectance of the most dense subweb for each web. They compared those values to means for the most dense subweb of networks (with the same levels of \( S \)) generated by various topological models including the niche model (Williams and Martinez 2000). No model was highly successful at generating webs with most dense subweb connectance closely comparable to that of the five empirical webs. However, to be a more appropriate test of the niche model, the analysis should be redone for trophic species versions of the food webs.

An issue for all studies considering compartmentalization or sub-web structure is whether such structures play any functional role within ecosystems. There are likely innumerable ways for chopping up networks into clusters, but that does not mean that such clusters are necessarily meaningful for ecological function or dynamics. Ideally, future compartmentalization studies will be more directly linked to issues such as ecosystem function, robustness, or resilience to perturbations.
5.3 CHAIN LENGTHS AND TROPHIC LEVELS

While from one perspective, food-chain length and trophic-level properties are just a few of many possible metrics of food-web structure (Williams and Martinez 2000), there is a very rich body of research specifically on chain lengths and trophic levels that spans theoretical, observational, experimental, and applied approaches dating back at least to Elton (1927). This research reflects the understanding that food-chain length, or “the number of transfers of energy or nutrients from the base to the top of a food web,” is a central characteristic of ecological communities due to its impact on ecosystem functioning, such as nutrient and carbon cycling, contaminant concentration, and trophic cascades (Post 2002). A great deal of research has gone into characterizing food-chain length, exploring its implications for ecosystem functioning, and determining whether and how it is constrained (Post 2002). Similar questions have been explored for the closely related concept of trophic level, or “the number of times chemical energy is transformed from a consumer’s diet into a consumer’s biomass along the food chains that lead to the species” (Williams and Martinez 2004b), although some researchers have claimed that the concept lacks scientific utility (Polis and Strong 1996).

Food-chain length and trophic level have been explored in detail using structural food-web data and models (e.g., Cohen and Newman 1991), but much of the research on these properties falls outside strictly structural approaches (Post 2002). Constraints on food-chain length or trophic levels in ecosystems have often been attributed to dynamical stability or resource availability, although this is increasingly questioned (e.g., Sterner et al. 1997) in favor of ecosystem size (Post et al. 2000). Post (2002) suggests that the debate is shifting from the search for singular explanations to “a complex and contingent framework of interacting constraints that includes the history of community organization, resource availability, the type of predator-prey interactions, disturbance and ecosystem size.” Although the nuances regarding such interacting constraints would seem to require the inclusion of data on relative flows along links to accurately characterize food-chain lengths and trophic levels, complex binary food webs without such flow data appear to provide a successful and simple quantitative framework for analysis. For example, a binary link-based measure called short-weighted trophic level yields surprisingly accurate estimates of species’ trophic and omnivory levels as compared to flow-weighted data (Williams and Martinez 2004b).

5.4 INTERACTION STRENGTH

Most early structural food-web research concerning May’s stability criterion (1973) generally focused on diversity ($S$) and connectance ($C$), the two parameters readily computed from structural data. Those analyses were based on the implicit assumption that the third parameter, interaction strength ($i$), was roughly constant. May’s analysis actually assigned interaction strengths ran-
domly. From very early on, researchers noted that these assumptions are flawed since interaction strengths are likely to vary non-randomly in real communities, with ramifications for ecosystem structure and stability (e.g., Paine 1969, 1980). One of the earliest analyses of this issue using structural food-web data was by Yodzis (1981), who used 40 food webs (Briand 1983) to construct community matrices based on structured interaction patterns. He found that stability, in May’s sense, was far more likely to occur if interaction strengths are assigned non-randomly based on ecological understanding. More recent research suggests that “non-random patterning of strong and weak links can be critical for the stability or persistence of theoretical and empirically observed complex communities” (Berlow et al. 2004; de Ruiter et al. 1995; Kokkoris et al. 1999; Neute et al. 2002). There are many opportunities and pitfalls in interaction strength research, as discussed in an excellent review by Berlow et al. (2004). Done thoughtfully, such research can provide linkages between ecological network structure and dynamics, and can facilitate future links between theoretical work and experimental and other empirical research.

5.5 INTEGRATING STRUCTURE AND DYNAMICS

Most models of food-web dynamics have focused on small modules with only a few species (see review by Dunne et al. 2005). While topology is generally thought to constrain ecological dynamics, most studies have explored this issue using very simplified network structures that can be imposed on species-poor dynamical models (see review by Jordán and Scheuring 2004). However, researchers are increasingly integrating dynamics with complex food-web structure in modeling studies that move beyond modules. The Lotka-Volterra cascade model (Cohen et al. 1990b; Chen and Cohen 2001a,b) was probably the first incarnation of this type of integration. As its name suggests, the Lotka-Volterra cascade model runs classic L-V dynamics, including a non-saturating linear Type I functional response, on sets of species interactions structured according to the cascade model (Cohen and Newman 1985b). The cascade model was also used to generate the structural framework for a recent dynamical food-web model with a Type I functional response (Kondoh 2003) that was used to study the effects of prey-switching on ecosystem stability. Improving on aspects of biological realism of both dynamics and structure, Yodzis used a bioenergetic dynamical model with nonlinear functional responses (Yodzis and Innes 1992), in conjunction with empirically-defined trophic network structure among 29 species, to simulate the biomass dynamics of a marine fisheries food web (Yodzis 1998, 2000). This bioenergetic dynamical modeling approach has been integrated with cascade and niche model network structure by other researchers (Brose et al. 2003; Williams and Martinez 2004a; Martinez et al. Chapter 6).

All of the approaches mentioned, as well as others, have been used to examine a variety of aspects of food-web complexity and stability including persistence, and are proving to be a valuable way to explore structural constraints
The Network Structure of Food Webs

on dynamics in complex ecosystems (Martinez et al. Chapter 6). An alternative approach is to develop assembly models that include ecological and/or evolutionary dynamics and see if they generate plausible ecological diversity and network structure (see review by McKane and Drossel Chapter 9), or to use methods such as genetic algorithms to explore the space of dynamically possible or probable structures (Ruiz-Moreno et al. Chapter 7). Integration of plausible ecological and evolutionary dynamics and network structure is a grand challenge for ecological modeling, with potential benefits for conservation and management of ecosystems (Yodzis 1998, 2000), as well as for fundamental scientific understanding of complex adaptive systems.

5.6 QUANTIFYING SPECIES AND LINKS

There have been many calls to go beyond species presence/absence data and binary link designations to include quantitative aspects of species and/or links during documentation and analysis of food webs (Cohen et al. 1993; Borer et al. 2002). Obvious ways to quantify “species” include characterizing their population abundance, population biomass, average individual biomass, and average body size. Links can be quantified through the amount of flow, usually biomass, attributable to the link (e.g., Baird and Ulanowicz 1989; Winemiller 1990), or the frequency of the occurrence of the links (Martinez et al. 1999). While researchers have compiled “quantitative” data for food-web components for as long as predator-prey relationships and food webs have been objects of study, it is only recently that some studies are attempting to integrate quantitative estimates with comprehensive descriptions of complex food-web network structure.

For example, Bersier et al. (2002) introduced a set of quantitative descriptors corresponding to several commonly studied food-web structure metrics (e.g., link and species proportions, L/S, C, chain properties, omnivory, generality, vulnerability). They used indices based on incoming and outgoing biomass flow for each species, as inspired by prior ecological network analysis applications of information theory (Ulanowicz 1986; Ulanowicz and Wulff 1991). Binary link metrics and quantitative link metrics do not return the same values for the Chesapeake Bay web (Baird and Ulanowicz 1989), and Bersier et al. (2002) suggest that their combined use will provide the most insight into food-web structure and function. A follow-up study looking at ~10 empirical webs with flow data suggests that quantitative indices are more robust to variable sampling effort than metrics based on binary links, although precision decreased (Banašek-Richter et al. 2004). A different approach focusing on species traits was introduced by Cohen et al. (2003), who document numerical abundance (varying across 10 orders of magnitude) and average body size (varying across ~ 12 orders of magnitude) for each species in the detailed pelagic food web of Tuesday Lake, Michigan (Jonsson et al. in press). They characterize this approach as a new way to describe ecological communities and they identify new community patterns for analysis. The approach of Cohen et al. (2003) was used to describe the detrital soil food-
Jennifer A. Dunne

While these and other “quantitative” approaches are intriguing, it remains to be seen whether they are primarily a tool for richer description of particular ecological communities, or whether they also give rise to generalities, novel models and predictions, or theory. Such approaches have much greater data requirements than just trying to characterize network structure. In many cases it will be difficult to document the necessary link and species characteristics given time and monetary constraints. Also, many of the concerns about “binary” data apply to “quantitative” data, and may be even more complicated to resolve. For example, biomass flow along trophic links and levels of species abundances vary spatiotemporally. A snapshot of a set of flows or abundances at a particular time and in a particular space ignores variability that may result in a very misleading picture of typical flow or abundance levels. Indeed, it may be difficult to characterize “typical” or average levels that are ecologically meaningful. Additional quantitative detail may be unnecessary for some kinds of investigations, since indices based on binary link data can be a good proxy for, or improvement upon, those based on richer quantitative data. For example, as previously mentioned, calculation of mean trophic level for food webs using detailed flow data is well approximated by a binary link-based structural index (Williams and Martinez 2004b). Also, species similarity measures based on binary link data outperform flow-based similarity measures (Yodzis and Winemiller 1999).

5.7 ECOLOGICAL NETWORKS

The food webs that are the object of study of most of the research reviewed in this chapter generally have as their focus classic predator-herbivore-primary producer feeding interactions. However, the basic concept of food webs can be extended to a broader framework of ecological networks that is more inclusive of different components of ecosystem biomass flow, and that takes into consideration different kinds of species interactions that are not strictly trophic. I give three of many possible examples here. First, parasites have typically been given short shrift in traditional food webs, although exceptions exist (Huxham et al. 1996; Memmott et al. 1999). Almost a decade after “a plea for parasites” in food webs (Marcogliese and Cone 1997), there are still few food-web studies that systematically incorporate or focus on parasites. Dobson et al. (Chapter 4) take up this issue and explore it further. A second issue that has yet to be resolved adequately for structural or dynamical food-web studies is the role of detritus, or dead organic matter, in ecosystems. Detritus has been explicitly included as one or several separate nodes in many binary-link and flow-weighted food webs. In some cases, it is treated as an additional primary producer, while in other cases both primary producers and detritivores connect to it. Researchers must think much more carefully about how to include detritus in all kinds of ecological studies (Moore et al. 2004), given that it plays a fundamental role in most
ecosystems and has particular characteristics that differ from other food-web nodes: it is dead and not living organic matter, it is heterogeneous, and it has an ambiguous trophic role. The third example concerns the analysis of ecological networks focused on other interactions besides strictly predator-prey relationships. Plant-animal mutualistic networks, particularly pollination and seed dispersal networks, have received the most attention thus far, and their network structure is discussed by Bascompte and Jordano (Chapter 5). While pollination and seed dispersal each involve a trophic interaction and can be portrayed as two trophic level networks; unlike in a classic predator-prey relationship, a strong positive benefit is conferred upon both partners in the interaction. The evolutionary and ecological dynamics of such mutualistic relationships may place unique constraints on the network structure of such interactions and the dynamical stability of such networks (Jordano 1987; Memmott 1999; Jordano et al. 2003; Bascompte et al. 2003; Vásquez and Aizen 2004).

5.8 ECOINFORMATICS

Future food-web research will depend on the collection of and access to increasingly high-quality data from systems spanning the globe. To address these needs, and to help identify gaps in knowledge that can be filled with strategic sampling, recent and historical food-web datasets are being compiled for a WWW-based publicly accessible trophic interaction database. This database will be integrated with structural and dynamical modeling tools as well as three-dimensional food-web graphics and animation tools. This project, called Webs on the Web (WoW), in effect updates and expands Cohen’s (1989) Ecologists Co-operative Web Bank (ECOWeB), a set of over 200 “machine readable” food-web datasets, and adds many layers of additional functionality and database capabilities. Initially, the WoW knowledgebase will include hundreds of food-web datasets, thousands of instances of consumer-resource relationships, and associated quantitative species and link information where available. Database tools will facilitate the decentralized addition of data to the knowledgebase, as well as annotation of existing data. Ideally, WoW will increase the ability of scientists, managers, and students to exchange and analyze information regarding the structure, function, and dynamics of ecological networks and the species within them.

Webs on the Web is just one example of an ever-expanding array of biodiversity-related databases that are, or will be, available on the WWW (Graham et al. 2004). In turn, these types of databases are just one of many kinds of ecoinformatics tools, which include technologies and practices for gathering, analyzing, visualizing, storing, retrieving, and otherwise managing ecological knowledge and information. There is a great deal of ecological information potentially available on the WWW, but it is widely dispersed and comes in a large variety of formats. Current search tools are very limited in their ability to effectively mine the data. To address these issues, efforts are underway to develop languages and tools for a Semantic Web that will allow for more sophisticated, content-based
access to dispersed data and information on the WWW (Hendler 2003). Some of these tools are being developed in an ecological network research context through the Semantic Prototypes in Research Ecoinformatics project (spire.umbc.edu) in conjunction with Webs on the Web (www.foodwebs.org). The emerging Semantic Web technologies, if properly developed and widely implemented, have the potential to transform the scope, effectiveness, and efficiency of ecological research across spatial and temporal scales (Green et al. 2005).

6 FINAL THOUGHTS

It is a vibrant time for research on the structure and dynamics of complex systems. Food webs, and more broadly, ecological networks, are a paradigmatic example of such systems. Research on complex food webs can not only benefit from research on other types of networks, but can provide novel insights that have implications for network research beyond ecology. However, as a cautionary note, what Lawton stated in 1989 is completely relevant today. Our data are still limited and of highly variable quality, and we may come to realize that some of the apparent generalities or universalities we see in that data are merely artifacts of poor information. Fortunately, researchers have not succumbed to “hand-wringing paralysis” given these limitations. However, the emergence of novel ways of analyzing and modeling ecological network data need to be accompanied by meticulous collection of detailed, comprehensive field data that seeks to address or overcome the previous empirical limitations. In addition, scientists must make careful decisions about what data to use, and perhaps more importantly, what data not to use in analyses. Beyond the data issues, researchers need to think continually about how to connect food-web properties, both structural and dynamical, back to issues of fundamental ecological interest, to avoid chasing research down dead-end alleys of little scientific interest. While some hand-wringers may continue to assert that research on food-web structure is one of those unfortunate alleys, the rich history outlined here combined with the recent tonic of interdisciplinary network research where it is well understood that structure always affects function (Strogatz 2001), shows that ecological network research is an important avenue for future advances. Such advances are likely to occur both with regard to more applied concerns such as the response of ecosystems to perturbations including biodiversity loss and species invasions (Memmott et al. Chapter 14), as well as more basic research questions such as identifying the processes or constraints that give rise to general structural patterns. Thoughtful and careful development of ecological network data, analyses, and models can provide the backbone for robust and general ecological theory on the complex networks of interactions among species across many scales, especially if integrated with other macroecological theory on allometric scaling and species-area relationships (Brose et al. 2004).
ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation (DBI-0234980, ITR-0326460). I thank the Santa Fe Institute and the Rocky Mountain Biological Laboratory for hospitality and administrative support, and appreciation goes to Brooke Ray Smith and Neo Martinez for helpful edits and comments made on a draft of the manuscript.

REFERENCES


The Network Structure of Food Webs


The Network Structure of Food Webs


The Network Structure of Food Webs


The Network Structure of Food Webs


The Network Structure of Food Webs