

Ecological Network Structure: Data, Models, and Inference

Jennifer A. Dunne

Santa Fe Institute Pacific Ecoinformatics & Computational Ecology Lab

PEaCE Lab: <u>www.foodwebs.org</u>



Why is network anatomy so important to characterize? Because structure always affects function.



Strogatz (2001) Exploring complex networks. Nature



-predation -herbivory -detritivory -parasitism -cannibalism

Nodes = Species/Taxa Edges = Directed trophic links

1950's Paradigm: Complex communities MORE stable than simple communities



1970's Challenge: Complex communities LESS stable than simple communities







Current & Future Research: "Devious strategies" that promote stability and species coexistence

- 1. Food-Web Data
- 2. General Patterns Across Webs
- 3. Network Structure Models
- 4. Confronting Models with Data
- 5. Case Study: Ancient Food Webs

1. Food-Web Data

Earliest known graph of feeding relations

Lorenzo Camerano, 1880





1st generation data: 1920s-1980s:

Summerhayes & Elton 1923: Food web of Bear Island







S (# taxa) = 28 L (# links) = 59 L/S (links/species) = 2.1 C (connectance; L/S²) = 0.075 TL (mean trophic level) = 2.07

1 bacteria, 4 autotrophs, 13 invertebrates, 6 birds, 4 mammals

Directed Connectance (C): Proportion of possible links (S^2) that are realized (L)

2nd generation data: 1990s-present

Food Web of Little Rock Lake, Wisconsin



S = 92, L = 997, L/S = 11, C = 0.12, TL = 2.40



3rd generation" data: 2009 and beyond

Antarctic Weddell Sea Food Web



Highly & Evenly Resolved

Original species = 492

62 autotrophs 4 mixotrophs 345 invertebrates 48 ectotherm vertebrates 29 endotherm vertebrates 3 detritus 1 bacteria

S = 290 L = 7200 L/S = 24.8 C = 0.086 Mean TL = 3.79

Data compiled by Ute Jacob



2. General Patterns Across Webs



Beyond S and C: Link distributions



Apparent complexity

Raw data for 16 webs





Camacho et al. 2002, Dunne et al. 2002

Whence the exponential?

Principle of Maximum Entropy (MaxEnt, Jaynes 1957)

The probability distribution with the maximum information entropy is the least biased probability distribution which satisfies a set of information-containing constraints.



Percent of 51 food webs with degree distributions not significantly different from MaxEnt distributions constrained by # species, links, and top or basal spp.

Why should observed distributions tend toward MaxEnt distributions?

Observed large-scale patterns typically arise from aggregation of many small-scale processes, whose fluctuations tend to cancel each other out. A few constraints set the pattern in the aggregate, everything else tends to the greatest randomness (Frank, in press).

Williams 2009

Beyond degree distribution

Types of Organisms:

% Top spp.	= 1.1
% Intermediate spp.	= 85.9
% Basal spp.	= 13.0
% Cannibal spp.	= 14.1
% Herbivore spp.	= 37.0
% Omnivore sp.	= 39.1
% Species in loops	= 26.1

7.28 1.31 5.75 2.40 0.74 0.60 1.42 0.71 1.91 0.18

Linkage Metrics:

Mean food chain length	=
SD food chain length	=
Log number of chains	=
Mean trophic level	=
Mean max. trophic simil.	=
SD vulnerability (#pred.)	=
SD generality (#prey)	=
SD links (#total links)	=
Mean shortest path	=
Clustering coefficient	=



Scale dependence with 5 & C

2.8 2.4 2.2 2.0 1.8 1.8 1.8 1.6 1.2 0.04 0.06 0.08 0.10 0.30 Connectance

Data from 7 Food Webs

Williams et al. 2002, Vermaat et al. 2009

3. Network Structure Models

Empirical regularities \rightarrow modeling opportunities

Simple, stochastic, single-dimensional models of food-web structure

Explain "the phenomenology of observed food web structure, using a minimum of hypotheses"

- 1) Two Parameters: **S** (species richness) and **C** (connectance)
- 2) Assign each species i a uniform random "niche value" n_i of 0 to 1
- 3) Simple rules distribute links from consumers to resources



Cohen & Newman 1985, Williams & Martinez 2000

Cascade model



Link distribution rules:

→ Each species i has probability P = 2CS/(S-1) of consuming resource species j with lower niche values ($n_j < n_i$)

Effect of rules:

 \rightarrow Creates strict hierarchy of feeding (cannibalism & longer cycles prohibited)

Niche model



Link distribution rules:

- \rightarrow Species i is assigned a feeding range r_i (drawn from beta distribution)
- \rightarrow The center c_i of r_i is a uniform random number between r_i/2 and min (n_i, 1-r_i/2)
- \rightarrow Species i feeds on all species that fall within r_i

Effect of rules:

- \rightarrow Beta distribution generates exponential-type degree distributions
- \rightarrow The feeding hierarchy is slightly relaxed (cycles can occur)
- \rightarrow Webs are "interval" (species feed on contiguous sets of species)

Nested hierarchy model

Link distribution rules:

- \rightarrow Each consumer i's number of resource species j assigned using beta distribution
- → Resources j chosen randomly from species with n_j < n_i until all links are assigned or a j is obtained which already has at least one consumer
- \rightarrow Species i links to j and joins j's "consumer group"
- → Subsequent j chosen randomly from the set of j of this group until all of i's links are assigned or all j of the consumer group have been chosen
- \rightarrow Subsequent j chosen from remaining species with no consumers and $n_i < n_i$
- \rightarrow Subsequent j chosen randomly from species with $n_i \ge n_i$

Effect of rules:

- \rightarrow Rules meant to mimic phylogenetic effects (guilds of related consumers)
- \rightarrow Food webs are not "interval"
- \rightarrow Hierarchy relaxed in principle, in practice rarely violated

Generalized cascade model

Link distribution rules:

→ Species i consumes resources species j with $n_j \le n_i$ with a probability equal to a random number with mean 2C drawn from a beta distribution

Effect of rules:

 \rightarrow Create a simple, non-interval, beta-distributed hierarchical model that allows cannibalism

'Relaxed' niche models

Link distribution rules:

- → Same as niche model, but allow for gaps in a slightly expanded feeding range or for links external to feeding range
 - 1. Generalized niche model (Stouffer et al. 2005)
 - 2. Relaxed niche model (Williams & Martinez 2008)
 - 3. Minimum potential niche model (Allesina et al. 2008)

Effect of rules:

 \rightarrow Relax the intervality constraint of the niche model

Stouffer et al. 2005, Williams & Martinez 2008, Allesina et al. 2008

Random models

Link distribution rules:

 \rightarrow Distribute links randomly

- 1. Random model (Williams & Martinez 2000): P = C
- 2. Random beta model (Dunne et al. 2008): beta distribution

Effect of rules:

 \rightarrow Minimal constraints

- 1. Random: no hierarchy, no intervality, no beta distribution
- 2. Random beta: no hierarchy, no intervality

Summary of model constraints

			hierarchical feeding	
Model	beta distribution	intervality	hierarchy	exceptions
Random	no	no	no	
Random beta	yes	no	no	
Cascade	no	no	yes	no
Generalized cascade	yes	no	yes	$n_j = n_i$
Niche	yes	yes	yes	$n_j \ge n_i$
Relaxed niche	yes	no*	yes	$n_j \ge n_i$
Nested hierarchy	yes	no	yes	$n_j \ge n_i^*$

4. Confronting Models with Data

Inference methods



Degree distribution
 Suite of properties
 Likelihood

1) Degree distribution

Property	Expression
Distribution of number of prey	$p_{\rm prev}(k) = (1/2z)E_1(k/2z)$
Distribution of number of predators	$p_{\text{pred}}(m) = (1/2z)\gamma(m + 1, 2z)$
Distribution of number of links	$p_{\text{links}}(r) = \int_0^r p_{\text{prey}}(t) p_{\text{pred}}(r-t) dt = 1/(2z)^2 \int_0^r E_1(t/2z) \gamma(r-t+1, 2z) dt$
Fraction of top species	$T = (1 - e^{-2z})/2z$
Fraction of basal species	$B = \ln(1 + 2z)/2z$
Standard deviation of the vulnerability	$\sigma_V = \sqrt{1/3 + 1/z}$
Standard deviation of the generality	$\sigma_G = \sqrt{8/(3 + 6C) - 1}$

Data Normalization & Analytical Analyses



Camacho et al. 2002, Stouffer et al. 2005

'Degree distribution' summary

				hierarchical feeding	
	Model	beta distribution	intervality	hierarchy	exceptions
	Random	no	no	no	
		<u> </u>			
	Cascade	110	110	yes	110
J	Generalized cascade	yes	no	yes	$n_j = n_i$
J	Niche	yes	yes	yes	$n_j \ge n_i$
				Ş	
J	Nested hierarchy	yes	no	yes	$n_i \ge n_i^*$

2) Suite of properties

Beyond degree distribution ...

- <u>Test:</u> against the structure of empirical food webs
- <u>Assess</u>: a suite of structural properties
- Generate: a set of 1000 model webs for each empirical web
- Evaluate: how well does the model perform?
 - → Normalized Model Error = (empirical value model mean) / (model median value value at upper or lower 95% boundary of model distribution)
 - \rightarrow MEs \leq |1| show 'good' fit of model mean to empirical value



Williams & Martinez 2008



Summary Stats

	ME	ME	% ME
	mean	SD	> 1
Gen Cas	-0.57	2.37	46%
Gen Nic	-0.50	1.40	39%
Nes Hier	-0.53	1.45	26%
Niche	-0.10	1.32	25%
Rel Nic	-0.40	1.58	33%

Mean ME ≤ |1| for all models: effect of hierarchy + beta distribution constraints
 Niche: lowest ME mean & SD, most properties closest to 0, fewest properties >|1|
 All models drastically underestimate herbivory/detritivory

'Suite of properties' summary

				hierarchical feeding	
	Model	beta distribution	intervality	hierarchy	exceptions
	Random	no	no	no	
	Random beta	yes	110	110	
	Cascade	110	110	yes	110
J	Generalized cascade	yes	no	yes	$n_j = n_i$
J	Niche	yes	yes	yes	$n_j \ge n_i$
J	Relaxed niche	yes	no*	yes	$n_j \ge n_i$
J	Nested hierarchy	yes	no	yes	$n_j \ge n_i^*$

3) Likelihood: topology as a whole



- 1) 3 models (Cascade, Niche, Nested hierarchy) and 10 datasets analyzed. All webs have links that violate assumptions of each model.
- 2) Use a genetic algorithm to order species to minimize violating links (Matrix $A \rightarrow A^*$).
- 3) Split datasets into compatible links (Matrix N) and incompatible links (Matrix K).
- 4) Calculate probability of obtaining Matrix N with the model and Matrix K with a random graph.
- 5) Product of those P's gives a "total likelihood" (Tot L) of that model for that dataset.

Alternate model: The minimum potential niche model

Link distribution rules:

→ Same as niche model, but define a feeding range where the consumer has a probability of <1 of feeding on species in that range.

Effect of link distribution rules:

- \rightarrow Relax the intervality constraint of the niche model
- \rightarrow No empirical links are incompatible with the models
- \rightarrow Introduces an extra model parameter



Minimum potential (relaxed) niche model performs best:

- \rightarrow no irreproducible links (Niche model has most)
- \rightarrow marginally better Tot \angle than the Niche model on every dataset
- \rightarrow much better Tot \angle than Nested hierarchy or Cascade models

'Likelihood' summary

				hierarchical feeding	
	Model	beta distribution	intervality	hierarchy	exceptions
	Cascade	no	110	yes	<u>no</u>
				÷	
J.	Niche	yes	yes	yes	$n_j \ge n_i$
J	Relaxed niche	yes	no*	yes	$n_j \ge n_i$
	Nested hierarchy	yes	no	yes	$n_j \ge n_i^*$

Pros and cons of inference approaches

1) Degree Distributions

Pros: Characterizes a central tendency of structure Cons: Very minimal notion of "structure"

2) Suite of Properties

Pros: Allows assessment of details of how/why structure differs Cons: Properties are not independent, making overall evaluation problematic

3) Likelihood

Pros: Based on full structure of network
Cons: How to understand details of how/why structure differs?
How to interpret magnitude of differences in Tot L?





- 'Complex' food webs aren't intractably complex: underlying shared scale-dependent structure.
- MaxEnt provides a simple null model for a central feature of data and models, the degree distribution.
- The Niche model and its variants, but not the Cascade model, do a good job of predicting many aspects of empirical web structure.

 \rightarrow Hierarchical Feeding + Beta/Exponential Distribution

- The Niche and Relaxed Niche models fit data better than other variants.
 - \rightarrow Intervality + Cycles
- Evaluation of models largely indifferent to inference method.

5. Case Study: Ancient Food Webs

Are species interactions structured differently in ancient versus modern ecosystems?

(can we even put convincing data together?)

- Have food webs become more complex since the beginning of the Phanerozoic?
- What do differences/similarities in ecological network structure suggest about fundamental constraints on species interactions?
 - Does food-web complexity or structure change across extinction boundaries?
 - Do major evolutionary innovations ramify throughout food webs?
 - How does community structure respond to major environmental perturbations?



Lagerstätten: Fossil assemblages with exceptional soft-tissue preservation



Lines of evidence for feeding interactions

Every link is a hypothesis based on inferences

- Gut contents
- Body size
- By analogy with associated taxa
- Damage patterns
- Environmental deposition
- Functional morphology
- Stable isotopes
- Trace fossils
- Coprolites
- The occasional smoking gun...





Functional morphology + damage patterns



Image by Ken Dowd (via NSF): Anomalocaris canadensis hunting trilobites

Burgess Shale Food Web



S = 85, **L** = 559, **L/S** = 6.6, **C** = 0.08, **TL** = 2.99







St. Martin (terrestrial)

Coachella Valley (terrestrial)



St. Marks (estuary)



Chesapeake Bay (estuary)

?



Bridge Brook (lake)



Chengjiang (marine)

Burgess (marine)





Caribbean Reef (marine)



Benguela (marine)

Normalized link distributions



Niche model errors

Model Errors for 17 properties for 10 food webs



Results generally robust to removal of uncertain links

A few potentially meaningful differences?



C = Chengjiang shale B = Burgess shale M = 8 modern Webs

- → Higher LinkSD in both Cambrian webs Reflects higher vulnerability to predation
- → Longer Path lengths in Chengjiang web Reflects lower integration among taxa
- → More taxa in Loops in Chengjiang web Reflects less hierarchical trophic organization

- Detailed species interaction data compiled for ancient ecosystems from the early Phanerozoic (> 500 MA).
- The structure of Cambrian & modern webs is very similar and well-predicted by the niche model.
- Results are robust to removal of uncertain or random links.
- Differences in Cambrian structure may reflect a rapid transition to more stable, constrained, hierarchical, integrated, trophic organization following the Cambrian "explosion" of diversity, body plans, and trophic roles.
- Shared architecture across habitats and deep time is suggestive of constraints on trophic organization.
 - \rightarrow null: MaxEnt
 - \rightarrow thermodynamic constraints
 - \rightarrow dynamical stability of complex systems
 - \rightarrow evolutionary processes



Dunne et al. 2008



Allesina, S., D. Alonso, & M. Pascual (2008) A general model for food web structure. Science 320: 658-661.

Camacho, J., R. Guimera, & L.A.N. Amaral (2002) Robust patterns in food web structure. *Physical Review Letters* 88: 228102.

Camerano, L. (1880) On the equilibrium of living beings by means of reciprocal destruction. *Atti della Reale Accademia delle Scienze di Torino* 15: 393-414.

Cattin, M.-F., L.-F. Bersier, C. Banašek-Richter, M. Baltensperger, & J.-P. Gabriel (2004) Phylogenetic constraints and adaptation explain food-web structure. *Nature* 427: 835-839.

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

Dunne, J.A., R.J. Williams, & N.D. Martinez (2002) Food-web structure and network theory. *Proceedings of the National Academy of Sciences USA* 99: 12917-12922.

Dunne, J.A., R.J. Williams, & N.D. Martinez (2004) Network structure and robustness of marine food webs. *Marine Ecology Progress Series* 273:291-302

Dunne, J.A., R.J. Williams, N.D. Martinez, R.A. Wood, & D.E. Erwin (2008) Compilation and network analyses of Cambrian food webs. *PLoS Biology* 6: 693-708.

Martinez, N.D. (1991) Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. *Ecological Monographs* 61: 367-392.

Stouffer, D.B., J. Camacho, R. Guimerà, C.A. Ng, & L.A.N. Amaral (2005) Quantitative patterns in the structure of model and empirical food webs. *Ecology* 86:1301–1311.

Summerhayes, V.S., & C.S. Elton (1923) Contributions to the ecology of Spitzbergen and Bear Island. *Journal of Ecology* 11:214–286.

Williams, R.J. (2009) Simple MaxEnt models explain food web degree distributions. *Theoretical Ecology*.

Williams, R.J., E.L. Berlow, J.A. Dunne, A.-L. Barabási, & N.D. Martinez (2002) Two degrees of separation in complex food webs. *Proceedings of the National Academy of Sciences USA* 99: 12913–12916.

Williams, R.J., & N.D. Martinez (2000) Simple rules yield complex food webs. Nature 404:180-183.

Williams, R.J., & N.D. Martinez (2008) Success and its limits among structural models of complex food webs. *Journal of Animal Ecology* 77:512–519.

Reviews:

Dunne, J.A. (2006) The network structure of food webs. pp. 27-86 in <u>Ecological</u> <u>Networks: Linking Structure to Dynamics in Food Webs</u>. M. Pascual & J.A. Dunne, eds. Oxford University Press.

Dunne, J.A. (In press) Food webs. <u>Encyclopedia of Complexity and Systems Science</u>. Springer.



Im Lavrakas / Anchorage Daily News