Network Models of Food Webs

Jennifer A. Dunne

Santa Fe Institute
Pacific Ecoinformatics & Computational Ecology Lab
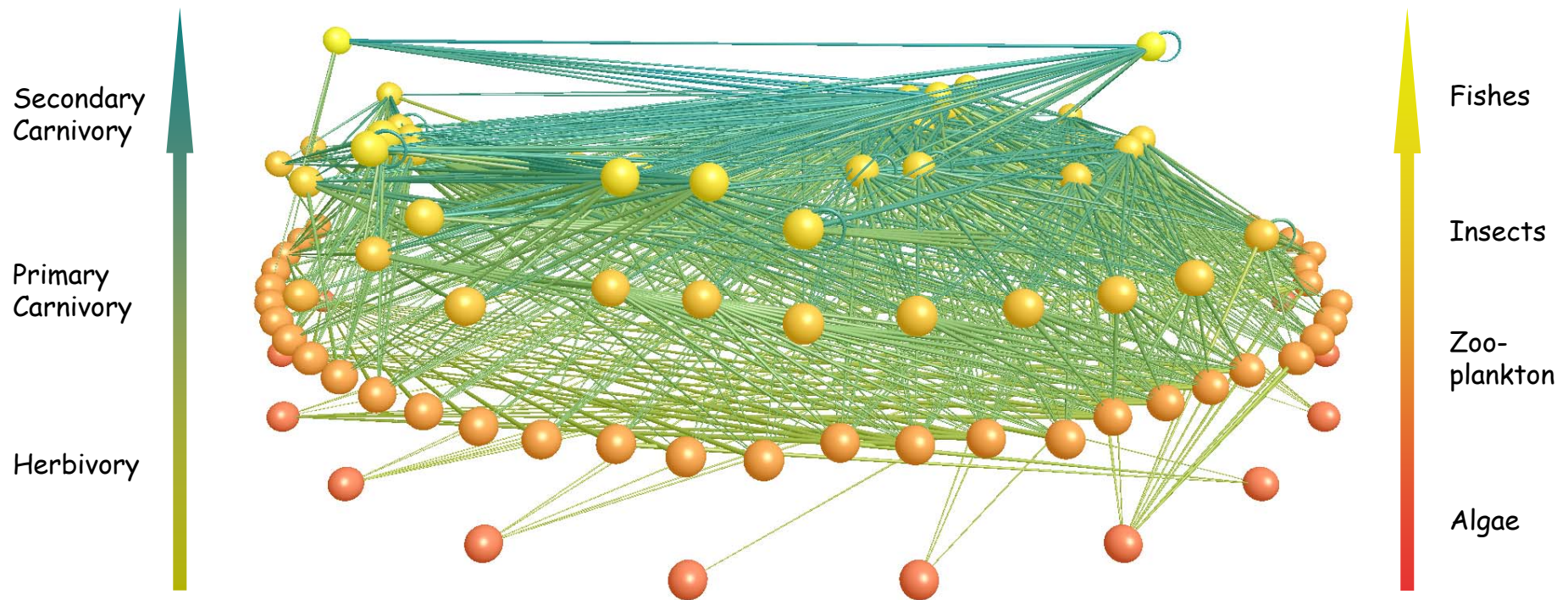
www.foodwebs.org

1) **Data**
2) Models
3) Models $\leftarrow\rightarrow$ Data
4) Thoughts

Food Web of Little Rock Lake, Wisconsin

181 taxa in original network: 11 fishes, 110 invertebrates, 59 autotrophs, 1 detrital category

Secondary Carnivory

Primary Carnivory

Herbivory

Fishes

Insects

Zoo-plankton

Algae

Nodes → Trophic Species (S)
Edges → Directed Feeding Links (L)
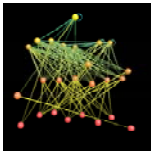Cycles → 1 (cannibalism), 2 (mutual predation), etc.

S = 92; L = 997
L/S = 11 (average degree)
C (L/S²) = 0.12 (connectance)
Mean Trophic Level = 2.40

Examples of currently used datasets

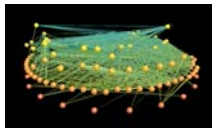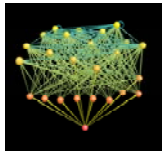S ~ 25 to 180, C ~ 0.03 to 0.3

Stream Webs

Canton Creek   Stony Stream

Lake & Pond Webs

Bridge Brook Lake

Little Rock Lake

Skipwith Pond
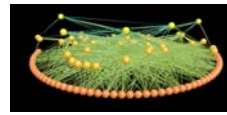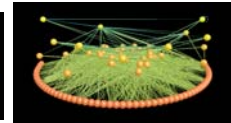
Estuary Webs

St. Marks Seagrass
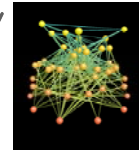
Chesapeake Bay

Ythan Estuary

Terrestrial Webs

El Verde Rainforest

St. Martin Island

Coachella Valley

Grassland   Scotch Broom
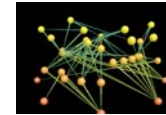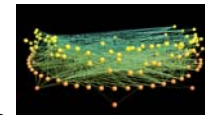
Marine Webs
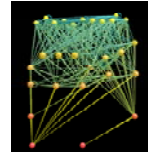
Benguela   Caribbean Reef   NE US Shelf
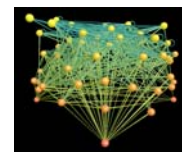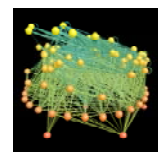
1950s Paradigm:

Complex communities MORE
stable than simple communities
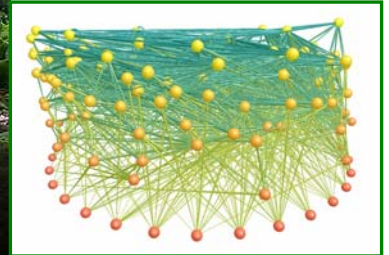
1970s Challenge:

Complex communities LESS
stable than simple communities
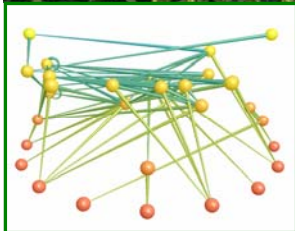
Current & Future Research:

"Devious strategies" that promote
stability and species coexistence

Marine

Estuary

Lake

Rain-forest

Desert
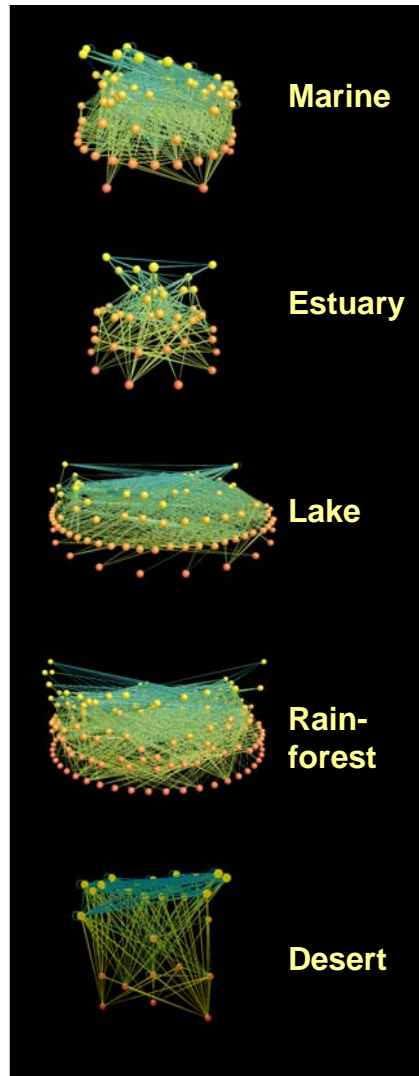
Apparent Complexity

Degree distributions

Apparent complexity $\longrightarrow$ **Underlying simplicity**

Raw data for 16 webs (log-linear)   Normalized data for 16 webs (log-log)
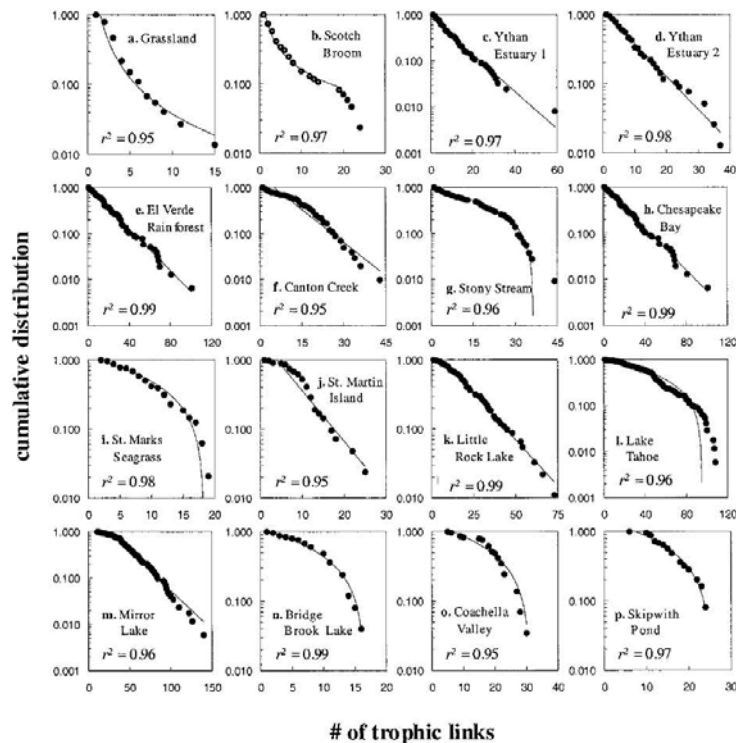
Dunne et al. (2002) Food-web structure and network theory. *PNAS* 99:12917-12922
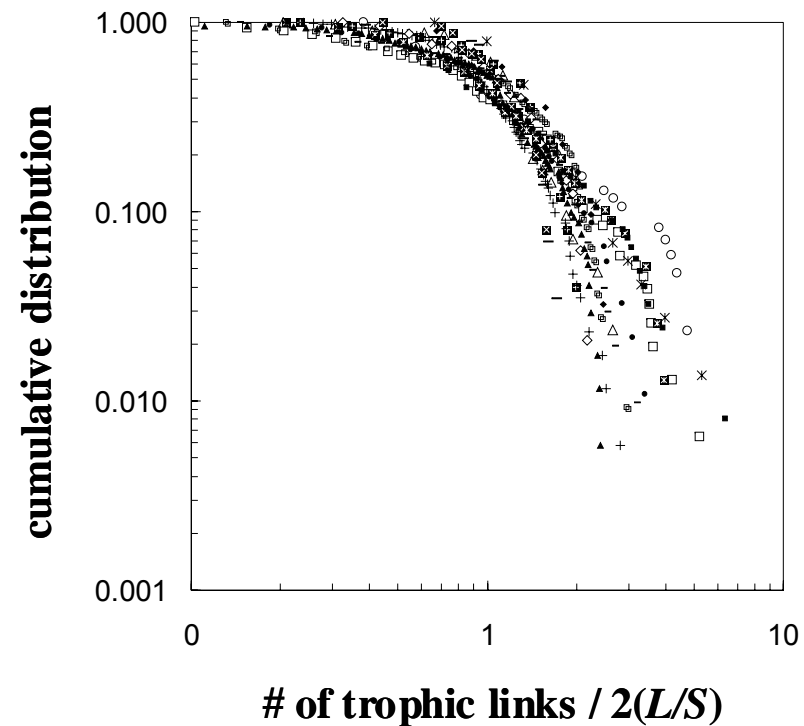
What about other properties?
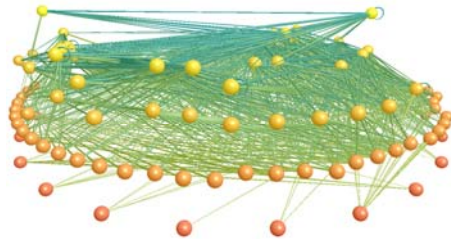
Types of Organisms:

% 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

Linkage Metrics:

Mean food chain length	= 7.28
SD food chain length	= 1.31
Log number of chains	= 5.75
Mean trophic level	= 2.40
Mean max. trophic sim.	= 0.74
SD vulnerability (#pred.)	= 0.60
SD generality (#prey)	= 1.42
SD links (#total links)	= 0.71
Mean shortest path	= 1.91
Clustering coefficient	= 0.18

Properties scale with C and/or S

Data from 7 Food Webs

Williams et al. (2002) Two degrees of separation in complex food webs. *PNAS* 99:12913-12916
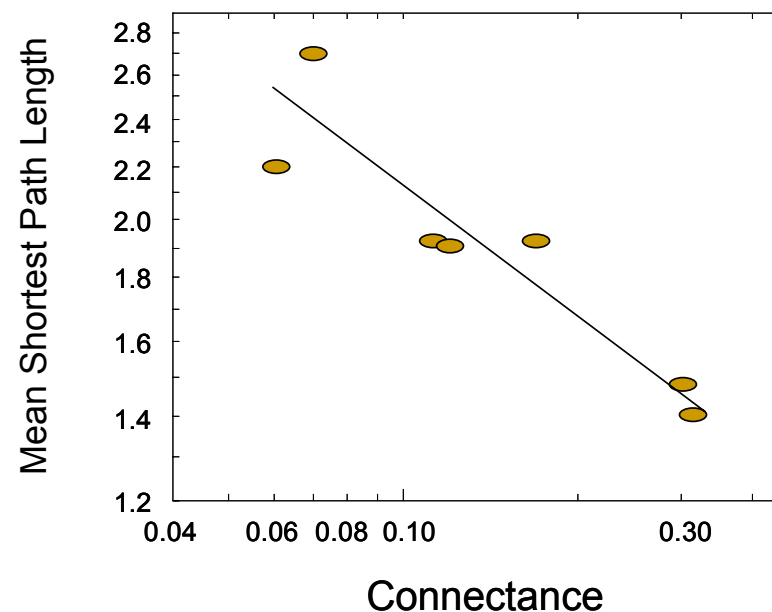
1) Data
2) Models
3) Models $\leftarrow\rightarrow$ Data
4) Thoughts

Empirical regularities provide modeling opportunities

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

Explain *"the phenomenology of observed food web structure, using a minimum of hypotheses"* (Cohen & Newman 1985)

→Two Parameters: S (species richness) and C (connectance)

→Assign each species i a uniform random "niche value" n_i along a "niche dimension" of 0 to 1 (i.e., $0 \leq n_i \leq 1$)

→Simple rules distribute links from consumers (predators) to resources (prey)

Cascade model (Cohen & Newman 1985)

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 link distribution rules:

→Creates strict hierarchy of feeding (no cannibalism or longer cycles possible)

Cohen, Newman (1985) A stochastic theory of community food webs: I. Models and aggregated data. *PRSLB* 224:421-448

Niche model (Williams & Martinez 2000)

Link distribution rules:

→ Species i is assigned a feeding range r_i
  - drawn from beta distribution

→ The center c_i of the feeding range r_i is a uniform random number between $r_i/2$ and $\min(n_i, 1-r_i/2)$
  - $c_i < n_i$
  - r_i placed entirely on the niche dimension
  - consumers' diets biased towards resources with lower n_i

→ Species i feeds on all species that fall within the feeding range r_i
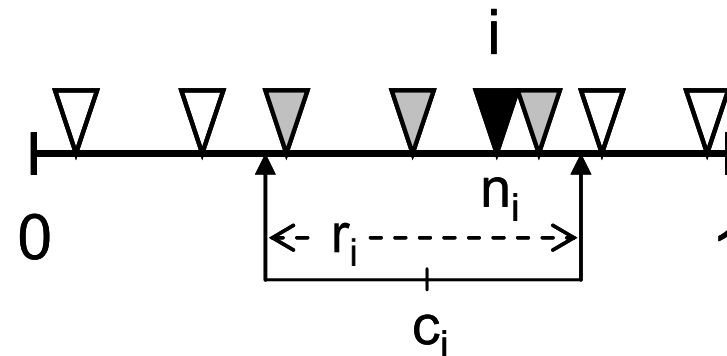
Effect of link distribution rules:

→ The feeding hierarchy is slightly relaxed (cycles can occur)

→ Food webs are "interval" (species feed on contiguous sets of species along a single dimension)

→ The beta distribution generates exponential-type degree distributions

The beta distribution
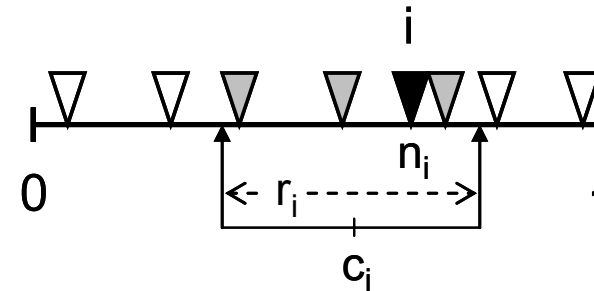
Beta distributions for various C

probability density

value chosen

0.10*
0.15* 0.2*
0.3*
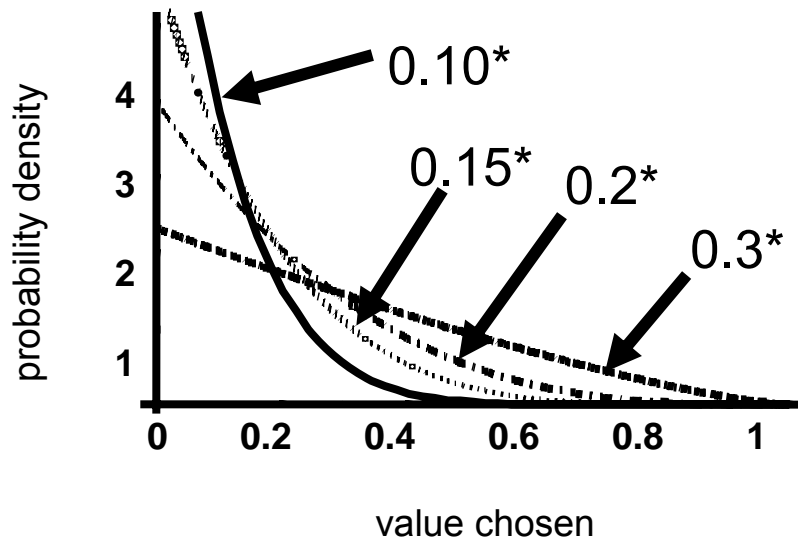
→ The niche range $r_i = xn_i$, where x is a random variable between 0 and 1 with a beta-distributed probability density function $p(x) = \beta(1-x)^{(\beta-1)}$ with $\beta = (1/2C)-1$

→ By paramaterizing the beta distribution with $2C$ and multiplying by n_i, the target C is achieved: mean $n_i = 0.5$, thus, mean $n_i(2C) = C$

→ Species' generality $\propto n_i$

→ Beta distribution ~exponential for $C < 0.15$

Williams, Martinez (2000) Simple rules yield complex food webs. *Nature* 404:180-183

Nested hierarchy model (Cattin et al. 2004)

Link distribution rules:

→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

→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

→Subsequent j chosen from remaining species with no consumers and $n_j < n_i$

→Subsequent j chosen randomly from species with $n_j \geq n_i$

Effect of link distribution rules:

→Rules meant to mimic phylogenetic effects

→Food webs are not "interval"

→Hierarchy relaxed in principle, in practice rarely violated

Cattin et al. (2004) Phylogenetic constraints and adaptation explain food-web structure. *Nature* 427:835-839

Generalized cascade model (Stouffer et al. 2005)

Link distribution rules:

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

Effect of link distribution rules:

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

Stouffer et al. (2005) Quantitative patterns in the structure of model and empirical food webs. *Ecology* 86:1301-1311

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. 2006)
    2. Relaxed niche model (Williams & Martinez 2008)
    3. Minimum potential niche model (Allesina et al. 2008)

Effect of link distribution rules:

→Relax the intervality constraint of the niche model

Stouffer et al. (2006) A robust measure of food web intervality. *PNAS* 103:19015-19020
Williams, Martinez (2008) Success and its limits among structural models of complex food webs. *JAE* 77:512-519
Allesina et al. (2008) A general model for food web structure. *Science* 320:658-661

Random models

Link distribution rules:

→Distribute links randomly

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

Effect of link distribution rules:

→Minimal constraints

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

Williams, Martinez (2000) Simple rules yield complex food webs. *Nature* 404:180-183
Dunne et al. (2008) Compilation and network analyses of Cambrian food webs. *PLoS Biology* 6:e102

Summary of model constraints

Model	beta distribution	intervality	hierarchical feeding 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	$n_j \geq n_i$
Relaxed niche	yes	no*	yes	$n_j \geq n_i$
Nested hierarchy	yes	no	yes	$n_j \geq n_i$*

1) Data
2) Models
3) **Models ←→ Data**
4) Thoughts
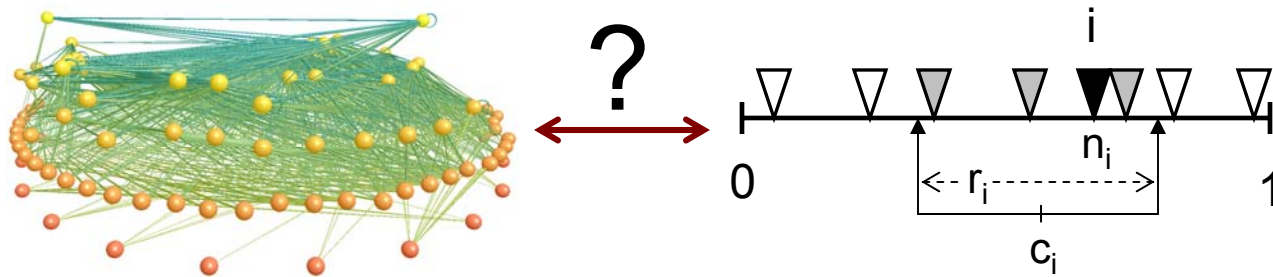
1. Degree distribution
2. Suite of properties
3. Likelihood

Degree distribution

Niche Model Simulation Results
(generality: links to prey)
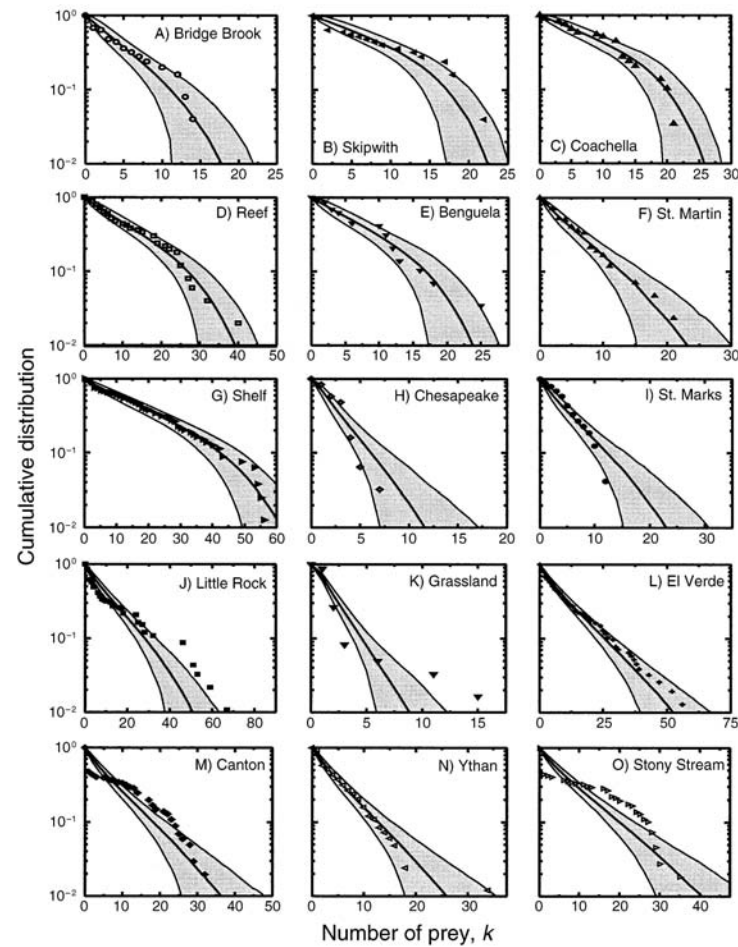
Niche Model Analytical Results

TABLE 2. Summary of the analytical expressions obtained for the niche model in the limits $S \gg 1$ and $C \ll 1$. These properties also hold for the nested-hierarchy and generalized cascade models in the same limits.

Property	Expression
Distribution of number of prey	$p_{prey}(k) = (1/2z)E_1(k/2z)$
Distribution of number of predators	$p_{pred}(m) = (1/2z)\gamma(m + 1, 2z)$
Distribution of number of links	$p_{links}(r) = \int_0^r p_{prey}(t)p_{pred}(r - t)dt$
	$= 1/(2z)^2 \int_0^r E_1(t/2z)\gamma(r - t + 1, 2z)dt$
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	$\sigma_G = \sqrt{8/(3 + 6C) - 1}$

Camacho et al. (2002) Robust patterns in food web structure. *Phys Rev Lett* 88:228102
Stouffer et al. (2005) Quantitative patterns in the structure of model and empirical food webs. *Ecology* 86:1301-1311
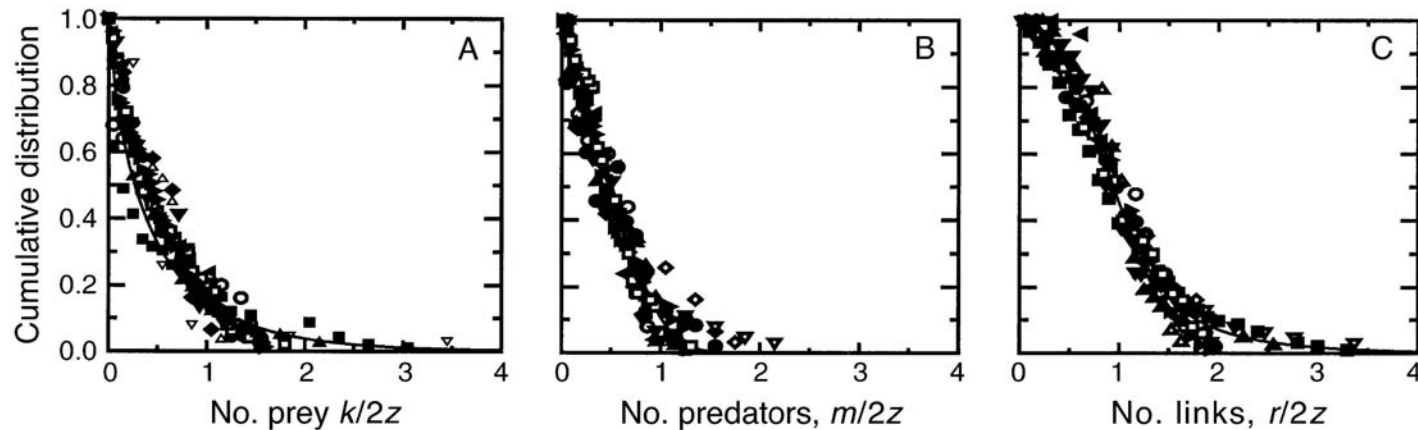
Degree distribution

Model	beta distribution	intervality	hierarchical feeding	
			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	$n_j \geq n_i$
~~Relaxed niche~~	yes	no*	yes	$n_j \geq n_i$
✓ Nested hierarchy	yes	no	yes	$n_j \geq n_i$*

Suite of properties

Beyond degree distribution...

- **Assess:** a suite of single-number structural properties

- **Generate:** sets of 1000 model webs with same S & C as empirical webs

- **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 distr.) [for one-tailed distributions]

  →MEs > |1| indicates that the empirical property value is not within the most likely 95% model property values and is significantly different from the range of property values produced by the model

---

Types of Organisms:

% Top spp.
% Intermediate spp.
% Basal spp.
% Cannibal spp.
% Herbivore spp.
% Omnivore spp.
% Species in loops

Linkage Metrics:

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

Results: Original test (7 webs, 3 models, 10 properties)

Histogram of NEs across properties, webs
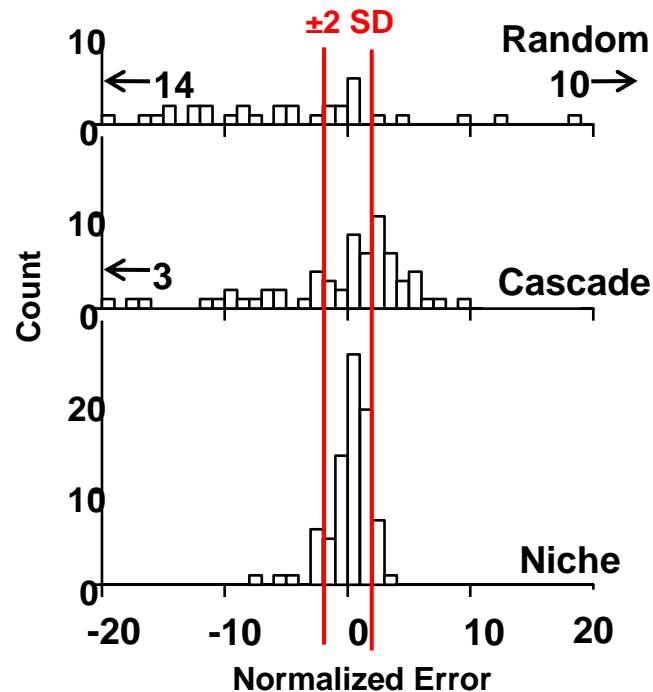
Old normalized error*

= (empirical value – model mean) / model SD

*assumes normal distribution of model values, |2| is cutoff

% of NEs ≤ |2|

Niche model: 79%
Cascade model: 27%
Random model: 16%

Similar results for 3 marine webs

Williams, Martinez (2000) Simple rules yield complex food webs. *Nature* 404:180-183

Dunne et al. (2004) Network structure and robustness of marine food webs. *Mar Ecol Prog Ser* 273:291-302

Results: Recent test (10 webs, 5 models, 15 properties)

Model	Mean of model error	SD of model error	Fraction outside ± 1 model error
Generalized Cascade (GCM)	−0·57	2·37	0·46
Generalized Niche (GNM)	−0·50	1·40	0·39
Relaxed Niche (RNM)	−0·40	1·58	0·33
Nested Hierarchy (NHM)	−0·53	1·45	0·26
Original Niche (ONM)	−0·10	1·32	0·25
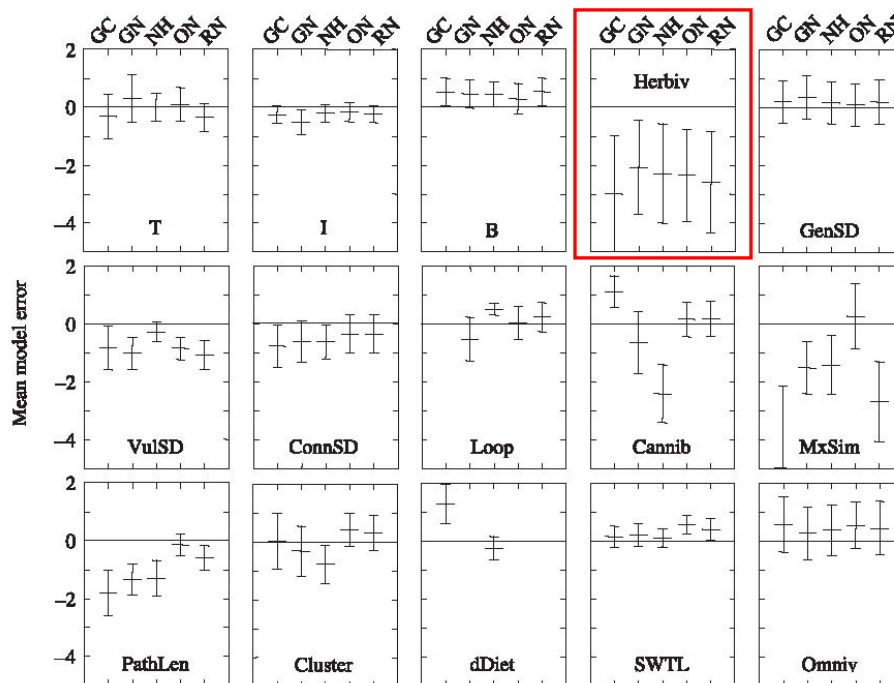
Summary:

Mean ME ≤ |1| for all models: effect of hierarchy + beta distribution constraints.

Niche has lowest ME mean & SD, most properties closest to 0, fewest properties outside |1|.

All models drastically underestimate herbivory.

Williams, Martinez (2008)Success and its limits among structural models of complex food webs. *J Animal Ecology* 77:512-519

Suite of properties

Model	beta distribution	intervality	hierarchical feeding	
			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	$n_j \geq n_i$
√ Relaxed niche	yes	no*	yes	$n_j \geq n_i$
√ Nested hierarchy	yes	no	yes	$n_j \geq n_i$*

Likelihood: Assessing topology as a whole

1) 3 models (Cascade, Niche, Nested hierarchy) and 10 datasets considered.

2) All empirical webs have links that violate assumptions of each model.

3) Use GA to order species in datasets to minimize violating links for each model (Matrix $A \rightarrow A*$). Split datasets into a set of links compatible with the model of interest (Matrix N), and a set of links incompatible with the model (Matrix K).
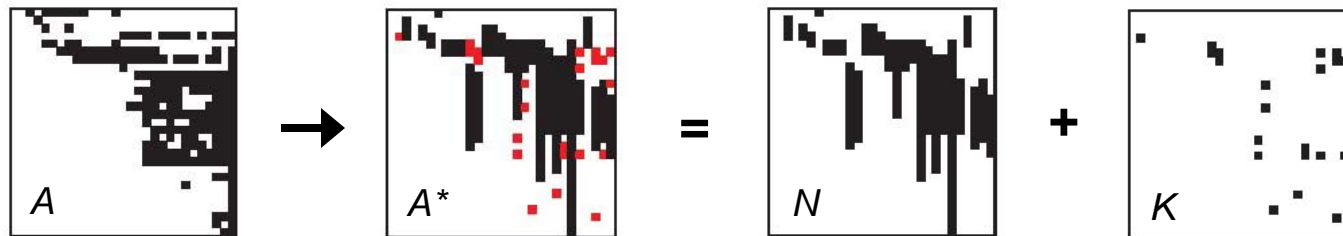
4) Calculate probability of obtaining Matrix N with the model and Matrix K with a random graph. Product gives a "total likelihood" (Tot L) of that model for that dataset.

5) 4th model: The **Minimum potential niche model** defines a feeding range where the consumer has a probability <1 of feeding on each species in that range. It is general: no incompatible links. While it introduces an extra parameter, its Tot L is comparable to other models, which include an extra parameter to reflect the random Matrix K.

Allesina et al. (2008) A general model for food web structure. *Science* 320:658-661

S = # of taxa (nodes)
L = # of links (edges)
I = # of irreproducible links
$L(K)$ = log-likelihood of obtaining I with random graphs
Tot L = total log-likelihood for the model

Food web	S	L		Cascade			Niche			Nested hierarchy			Min. potential	
			I	$L(K)$	Tot L	I	$L(K)$	Tot L	I	$L(K)$	Tot L	Tot L	f
Benguela	29	203	12	−62.91	−343.62	23	−105.46	−234.22	1	−7.73	−349.39	−213.52	0.170
Bridge	25	107	4	−24.19	−217.16	1	−7.44	−94.42	1	−7.44	−162.32	−92.18	0.013
Broom	85	223	4	−33.99	−857.42	36	−226.77	−737.56				−626.54	0.336
Chesapeake	31	68	1	−7.87	−199.59	10	−55.60	−166.84	3	−20.30	−200.15	−145.11	0.314
Coach	29	262	41	−163.85	−443.67	37	−151.75	−296.76	7	−40.49	−381.57	−296.10	0.240
Grass	61	97	0	0	−379.31	10	−69.18	−327.08	13	−86.52	−437.81	−294.94	0.243
Reef	50	556	59	−279.34	−1106.54	196	−687.11	−970.28	22	−126.03	−1053.50	−934.71	0.416
Skip	25	197	12	−59.32	−259.02	22	−95.24	−191.11	5	−29.12	−254.74	−169.67	0.142
St. Marks	48	221	3	−22.93	−576.69	72	−320.40	−546.48	18	−105.27	−634.04	−504.49	0.554
St. Martin	42	205	0	0	−472.58	52	−234.48	−421.53	10	−61.70	−531.55	−388.06	0.443

Minimum potential (relaxed) niche model performs best:

→ no irreproducible links (the Niche model has the most)

→ slightly better Tot L than the Niche model on every dataset

→ much better Tot L than Nested hierarchy or Cascade models

Allesina et al. (2008) A general model for food web structure. *Science* 320:658-661

Likelihood

Model	beta distribution	intervality	hierarchical feeding	
			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	$n_j \geq n_i$
Relaxed niche	yes	no*	yes	$n_j \geq n_i$
~~Nested hierarchy~~	~~yes~~	~~no~~	~~yes~~	~~$n_j \geq n_i$*~~

Pros and Cons

1) Degree Distributions
Pros: Characterizes a central tendency of structure
Cons: Very limited/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 assessment suspect

3) Likelihood
Pros: Based on full structure of network
Cons: Doesn't allow one to understand details of how/why structure differs;
   Had to add parameter to models to calculate likelihood;
   Not clear how to interpret the magnitude of differences in Tot L

Together, the 3 approaches suggest the following:

→ The Niche and Relaxed niche models fit data much better than Random or Cascade models, somewhat better than other beta-distributed models.

→ Thus, the combination of beta distribution, hierarchical feeding, and intervality or near-intervality constraints performs best.

1) Data
2) Models
3) Models $\leftarrow\rightarrow$ Data
4) Thoughts

Summary

- 'Complex' food webs aren't so complex: underlying common scale-dependent structure.

- The Niche model and its recent spin-offs (but not Random or Cascade models) do a good job of predicting many aspects of fine-grained structure of empirical food webs.
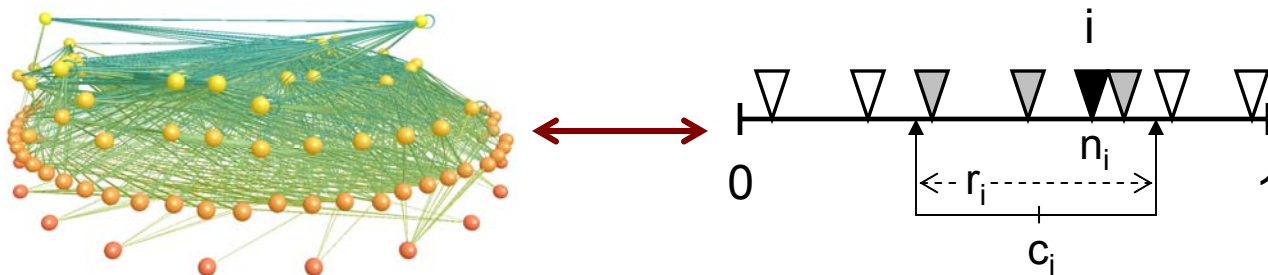
  →**Hierarchical Feeding + Beta Distribution**

- The Niche and Relaxed niche models fit data slightly better than non-interval variants (Nested hierarchy, Generalized cascade).

  →**Intervality + Cycles**

- Common structure across habitat and deep time suggests strong constraints on the organization of species interactions in communities.
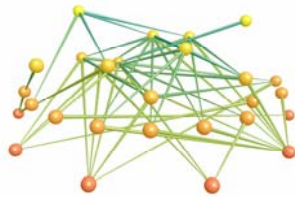
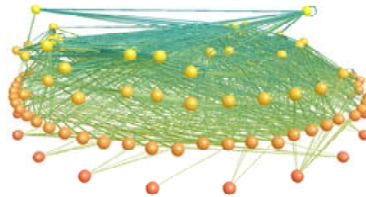 →**Ecology, Evolution, Energetics**

A few questions...

- Are there better ways of assessing the fit of slightly different models to data?

- Are there better ways of understanding differences/similarities across datasets?

- What happens when we move to 3rd generation data:

Gen 3: ?????
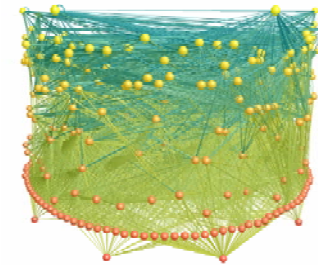
Gen 1: Cascade

Gen 2: Niche

1923

1991

2008

- Data & models of food-web assembly and disassembly?

- How does structure affect dynamics and vice-versa? Robustness?

- Mechanisms that give rise to shared, scale-dependent network structure?
  Are the model 'constraints' pointing us in useful directions?

- Compartments/structure within a food web?