

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/309739386>

Species interaction networks

Chapter · January 2013

DOI: 10.1002/9780470057339.vnn142

CITATION

1

READS

7

1 author:



Grace S. Chiu

Australian National University

33 PUBLICATIONS 133 CITATIONS

SEE PROFILE

All content following this page was uploaded by [Grace S. Chiu](#) on 08 November 2016.

The user has requested enhancement of the downloaded file. All in-text references [underlined in blue](#) are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

Species interaction networks

Introduction

A trophic food web is a collection of identified species connected by trophic (feeding) links that conventionally point from the consumed to the consumer. To study an empirical food web at the individual or even species level is rare, as higher resolution naturally leads to more complexity. Rather, various species may be aggregated according to certain criteria [1], and represented as a single node or compartment in the directed network. Such a node is also referred to as a *trophic species* or *trophospecies*. Trophic interaction data appear in the form of a diet (adjacency) matrix, wherein the (i, j) th element y_{ij} – either binary for presence–absence or *weighted* for magnitude – quantifies the consumption of (transfer from) Node i by (to) Node j . A pictorial representation is $i \xrightarrow{y_{ij}} j$. For example, consider a hypothetical food web of size $n = 4$, comprising the nodes *Carnivores*, *Omnivores*, *Herbivores*, and *Plants*. Then, the 4×4 binary diet matrix $[y_{ij}]$ may be

$$\begin{matrix} i \setminus j & \text{Carn.} & \text{Omni.} & \text{Herb.} & \text{Plan.} \\ \text{Carn.} & \left(\begin{matrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{matrix} \right) \\ \text{Omni.} & & & & \\ \text{Herb.} & & & & \\ \text{Plan.} & & & & \end{matrix}$$

and the corresponding weighted matrix replaces each y_{ij} above with a magnitude measure for j 's consumption of i . Note that a typical diet matrix is asymmetric, that is, $y_{ij} \neq y_{ji}$ for some (i, j) . Diet data may be accompanied by nonconsumption characteristics of the trophic nodes, such as **biomass**, habitat type, taxonomy, and so on.

The reader may refer to **Food web modeling** for a primer on various perspectives of empirical food web research, some sources of uncertainty inherent in food web data, and implications of this uncertainty. The present article elaborates upon the quantitative practices in studying the network structure of species interactions in the form of predation or consumption in general. Emphasis is on illustrating how statistical modeling can exploit variability patterns based

empirically on trophic interaction data and theoretically on ecological understanding, so that uncertainty and power assessments are built into the estimation and prediction of feeding interaction and the inference for network structure.

Analysis of Network Structure

Trophic interaction patterns are associated with the food web's network topology. Many semiquantitative metrics developed for social network analysis (SNA), as well as information theoretic indices of entropy, have been directly adopted to describe food web structure. These include the connectance index, mean food chain length, and the Shannon index of diversity, to name a few. The Shannon index based on a diet matrix is $-\sum_i \sum_j y_{ij}^* \log y_{ij}^*$, where $y_{ij}^* = y_{ij} / \sum_a \sum_b y_{ab}$ [2]. It describes the diversity among consumer-and-consumed pairs; maximum diversity of pairs is attained if all y_{ij} s are equal and positive. Connectance is the number of observed links divided by the maximum possible number of links, and is therefore a link density measure. For the four-node example above, the connectance index is 0.5. The food chain length (FCL) associated with a given node is the number of intermediate links between it and the basal node that does not consume, such as *Plants* in the above example. Thus, the shortest FCL for *Omnivores* above is 1 (*Plan.* \rightarrow *Omni.*), and the longest is 3 (*Plan.* \rightarrow *Herb.* \rightarrow *Carn.* \rightarrow *Omni.*). Different versions of the FCL lead to different definitions of the mean FCL [3]. Related metrics such as mean trophic level and the omnivory index [3] are tailored to describing predator–prey networks.

Insight into food web topology can help to identify functional groups within a community [4], to compare interaction behavior, network stability, and so on, between communities [2, 5, 6], and these in turn help to address conservation concerns [7]. However, “[b]y condensing the biological information into a single number, the complexity of the ecological system has been greatly reduced” (quoted from **Ecological statistics**). Thus, descriptive metrics alone do not readily provide the necessary insight. Rather, a series of other intermediate steps is required to make sense of the suite of food web metrics. Many of these steps in the literature employ statistical modeling tools to correlate features associated with the metrics, but are typically unaccompanied by a discussion of the assumptions required by the statistical

2 Species interaction networks

model. A statistical model, through model assumptions, ought to reflect the nature of the uncertainty inherent in the food web data (due to feeding behavior responding to environmental influence, due to observation error in the field or laboratory, etc.). Thus, conclusions about the food webs could be skewed when model assumptions are violated, most commonly the independence assumption associated with linear regression (and hence, correlation significance tests). In fact, this type of independence contradicts the fundamental notion of food web *structure* being studied, as the hallmark of a network is the presence of complex dependencies due to interwoven linkages, forming network patterns. Irrespective of the validity of assumptions, the practice of applying stand-alone procedures sequentially poses an enormous challenge for a rigorous understanding of the intricate network of trophic interactions, as there is no obvious framework to consolidate the conclusions made from individual procedures or to assess their uncertainty as a whole.

While this practice continues at present, techniques that account for the uncertainty in trophic interactions have begun to emerge in food web literature. To measure the degree of specialization in consumption behavior, Blüthgen *et al.* [2] apply **permutation tests** to determine the statistical significance of their entropy-based indices against a null model that amounts to a patternless network. To identify trophic compartments, Krause *et al.* [4] apply a clustering algorithm developed in the 1990s for SNA via a rather complex statistical model [8]; they also use the permutation test for compartmentalization against a null patternless network. Alternatively, fully parametric hypothesis tests associated with the model in Ref. 8 can be used to yield direct inference for compartmental structure and its uncertainty. Yet, these approaches do not address the question: What biological or environmental factors may have contributed to the structure of specialization, compartmentalization, and other aspects of the food web?

To do so cohesively (avoiding the use of stand-alone post-analysis procedures), recently Chiu and Westveld [9,10] integrated trophic interactions and non-consumption information for a food web using a type of hierarchical **generalized linear mixed model** known as *latent space modeling* [11,12] developed for SNA in the social sciences. (Details appear in the section titled Latent Space Modeling of Trophic

Interactions.) This approach enables unified statistical inference for latent (unobservable) quantities $[s_i, r_i]'$, u_i , and v_i , respectively reflecting feeding activity, so-called “prey’s preference” when being consumed, and predator’s preference when consuming. On the basis of the inference for these quantities over $i = 1, \dots, n$, three network graphs labeled SR, U, and V can be generated to depict trophic compartmentalization from each of the perspectives of feeding activity, prey’s preference, and predator’s preference (Figure 1). These graphs allow visualization of trophic clustering and the associated uncertainty through a quantitative lens, with nonconsumption predictor variables (if used) and the uncertainty in the trophic data rigorously accounted for.

Note that the notion of prey’s preference originates from the generic context of sending and receiving within a network. For food web feeding interactions, Node i ’s sending preference refers to its tendency to be consumed by certain predators, or alternatively, the preference for i exhibited by i ’s predators in their consuming behavior. This perspective of prey’s preference, U, is different from predator’s overall consumption preference. The interpretation of these three perspectives of feeding behavior is summarized in Table 1 in the context of clustering with respect to SR, U, and V. Note that SR-clustering pertains to direct relationships, and is the model-based analogy of the definition of *trophic level*; the more novel U-clustering (nodes being similarly preferred as prey) and V-clustering (nodes having similar preference for prey) are complementary indirect relational features that, although related to specialization, may not have immediate counterparts from conventional food web analysis techniques. As illustrated by the Benguela food web graphs in Figure 1, compartments can differ drastically depending on the perspective from which inference for trophic relations is made. For example, in graph SR, the regions of low uncertainty (high posterior density for $[s_i, r_i]'$) show little overlap between $i = 29$ and 14 or 20, thus providing strong statistical distinction between sharks and goby/squid in their feeding activity (sharks are inactive as prey but active as predators, but goby or squid is average as either), after accounting for phylogenetic similarity as predictor variable for trophic interactions. In contrast, from the regions of low uncertainty in graph U (high posterior density for u_i), sharks and goby/squid are less statistically distinguishable with respect to their being

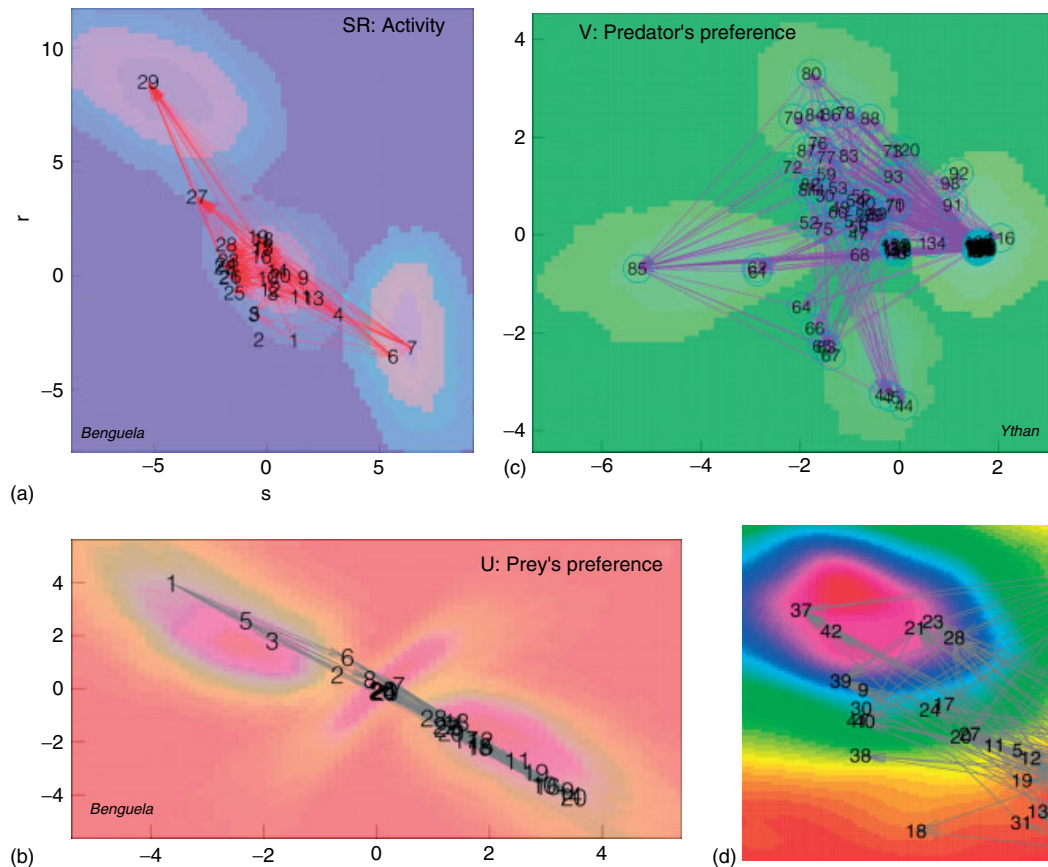


Figure 1 Visualizing the statistical inference for trophic compartmentalization using Chiu and Westveld's approach for binary trophic links [10]. Edges in each food web graph in panels (a) to (d) point from prey to predator. Nodes are arranged according to the "estimated values" of feeding activity (graph SR in panels (a) and (d)), so-called prey's preference when being consumed (graph U in panel (b)), and predator's preference when consuming (graph V in panel (c)). (An "estimate" here is the posterior mean for the $[s_i, r_i]'$, u_i , or v_i vector, from fitting the model made up of (1)–(6).) All node labels follow the original literature cited in Ref. 10. Close neighbors in graph SR/U/V are similar with respect to their feeding activity/preference as prey/preference as predator. (a) Feeding activity for the Benguela web ($n = 29$) with phylogenetic similarity as predictor; heat map is based on the uncertainty for $i = 7$ (macrozooplankton), 14 (goby), and 29 (sharks) – hotter colors represent less uncertainty (i.e., higher posterior density). (b) Same as (a) but for prey's preference, with heat map based on the uncertainty for $i = 1$ (phytoplankton), 7, and 20 (squid); near 7 is a tight cluster of five nodes (geelbek, kob, tunas, whales and dolphins, and yellowtail). (c) Predator's preference for the Ythan web ($n = 92$) without predictors; heat map is based on the uncertainty for $i = 43$ (river otter), 80 (common goby), 85 (flounder), and 116 (mud snail), which is next to a tight cluster of 30 nodes near $v = (2, 0)$. (d) Close-up view of the Goose Creek Bay web's graph SR, showing only nodes that are the most active consumers (large r) and least preyed upon (small s) after accounting for phylogenetic similarity (full graph in Ref. 10); heat map is based on the uncertainty for $i = 37$ (benthos-eating birds).

preferred as prey, since the estimated location (posterior mean) for sharks is near where there is least uncertainty for squid while the estimated locations for squid and goby virtually coincide. Figure 1 also shows graph V for the Ythan web and part of graph

SR for the Goose Creek Bay web; more webs are analyzed this way in Refs 9,10.

In addition to trophic clustering, inference from this unified statistical SNA approach simultaneously addresses various food web features of ecological

4 Species interaction networks

Table 1 Interpretation of nearby nodes i and j in a food web graph that results from a statistical SNA based on (1)–(6).

Graph	Interpretation ^a
SR	Nodes i and j are trophically similar (similarly active as prey and as predator)
U	Nodes i and j are similarly preferred as prey by their predators (but their predators do not necessarily prefer the same prey in general). More technically, i and j are similar in their sending behavior aside from their activity as prey
V	Nodes i and j have similar consumption preference. More technically, i and j are similar in their receiving behavior aside from their activity as predator

^aThe interpretation is for the mean-zero random effects, that is, for trophic behavior after accounting for regressors.

interest. As can be seen from the modeling framework in the section titled Latent Space Modeling of Trophic Interactions, these features include the tendency for an (in)active predator to be an (in)active prey (through the model parameter ρ_{sr}), the tendency for predator–prey reciprocity (through ρ), the statistical relevance of the predictor variables \mathbf{x}_{ij} to feeding behavior (through β and model goodness-of-fit), and the uncertainty in any descriptive metric (through posterior predictive inference for y_{ij} , in the context of **missing data** imputation or otherwise). Chiu and Westveld provide details in Refs 9,10.

Latent Space Modeling of Trophic Interactions

Given diet data y_{ij} s for a food web of size n , a latent space model is

$$E(y_{ij}|\theta_{ij}) = \theta_{ij}$$

$$\text{link}(\theta_{ij}) = \beta_0 + \mathbf{x}'_{ij}\beta + s_i + r_j + \mathbf{u}'_i\mathbf{v}_j + \varepsilon_{ij}$$

$$\text{for } i \neq j \quad (1)$$

$$\mathbf{u}_i = [u_{i1}, u_{i2}]' \quad \mathbf{v}_i = [v_{i1}, v_{i2}]' \quad (2)$$

$$[s_i, r_i]' | \rho_{sr}, \sigma_s^2, \sigma_r^2 \sim N\left(\mathbf{0}, \begin{bmatrix} \sigma_s^2 & \rho_{sr}\sigma_s\sigma_r \\ \rho_{sr}\sigma_s\sigma_r & \sigma_r^2 \end{bmatrix}\right) \quad (3)$$

$$u_{iq} | \sigma_{uq}^2 \sim N(0, \sigma_{uq}^2) \quad \text{for } q = 1, 2 \quad (4)$$

$$v_{iq} | \sigma_{vq}^2 \sim N(0, \sigma_{vq}^2) \quad \text{for } q = 1, 2 \quad (5)$$

$$[\varepsilon_{ij}, \varepsilon_{ji}]' | \rho, \sigma^2 \sim N\left(\mathbf{0}, \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right) \quad \text{for } i \neq j \quad (6)$$

where (1)–(6) are for $i, j = 1, \dots, n$, and \mathbf{x}_{ij} is a vector of optional nonconsumption regressors associated with the trophic quantity y_{ij} ; each regressor may describe the i th or j th node alone (e.g., node biomass), or the (i, j) th pair of nodes (e.g., phylogenetic similarity between the pair). A logit link (see **Logistic regression**) is used for binary y_{ij} , and identity link, for nonnegative real y_{ij} after a suitable transformation. Note that the equations altogether constitute a single statistical model; cannibalism is not considered by the model (y_{ii} is excluded for all i). Specifically, (1) decomposes the total conditional random error into the following mean-zero random effects: s_i , sender (prey) i 's activity level; r_j , receiver (predator) j 's activity level; $\mathbf{u}'_i\mathbf{v}_j$, the interaction between i and j , where \mathbf{u}_i and \mathbf{v}_i are the respective mean-zero effects corresponding to prey i 's and predator j 's preference; and ε_{ij} , the remaining component unattributable to the former three.

Bayesian inference for this model yields the trio of food web graphs SR, U, and V (Figure 1), in which Node i is located according to the **posterior distribution** of the latent **random effects** $[s_i, r_i]'$ for graph SR, \mathbf{u}_i for graph U, or \mathbf{v}_i for graph V; uncertainty in node location (conditioned on the observed data) is fully quantified by the posterior. Inference for this model can lead to additional insight into food web structure (see the end of the the section titled Analysis of Network Structure).

Acknowledgment

The writing of this article was sponsored by the CSIRO Water for a Health Country Flagship under the Healthy Water Ecosystems theme. The author thanks the Section Editor, the Reviewer, Dr. G.R. Hosack (CSIRO), and Dr. A.H. Westveld (University of Arizona) for their valuable suggestions.

References

- [1] Yodzis, P. & Winemiller, K.O. (1999). In search of operational trophospecies in a tropical aquatic food web, *Oikos* **87**, 324–340.
- [2] Blüthgen, N., Menzel, F. & Blüthgen, N. (2006). Measuring specialization in species interaction networks, *BMC Ecology* **6**, 9. DOI: 10.1186/1472-6785-6-9.
- [3] Williams, R.J. & Martinez, N.D. (2004). Limits to trophic levels and omnivory in complex food webs: theory and data, *The American Naturalist* **163**, 458–468.
- [4] Krause, A.E., Frank, K.A., Mason, D.M., Ulanowicz, R.E. & Taylor, W.W. (2003). Compartments revealed in food-web structure, *Nature* **426**, 282–285.
- [5] Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance, *Ecology Letters* **5**, 558–567.
- [6] Thébault, E. & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks, *Science* **329**, 853–856.
- [7] Tylianakis, J.M., Laliberté, E., Nielsen, A. & Bascompte, J. (2010). Conservation of species interaction networks, *Biological Conservation* **143**, 2270–2279.
- [8] Frank, K.A. (1995). Identifying cohesive subgroups, *Social Networks* **17**, 27–56.
- [9] Chiu, G.S. & Westveld, A.H. (2010). A statistical social network model for consumption data in food webs, [arXiv:1006.4432 v2](https://arxiv.org/abs/1006.4432v2).
- [10] Chiu, G.S. & Westveld, A.H. (2011). A unifying approach for food webs, phylogeny, social networks, and statistics, *Proceedings of the National Academy of Sciences of the USA*, **108**, 15881–15886.
- [11] Hoff, P.D., Raftery, A.E. & Handcock, M.S. (2002). Latent space approaches to social network analysis, *Journal of the American Statistical Association* **97**, 1090–1098.
- [12] Ward, M.D. & Hoff, P.D. (2007). Persistent patterns of international commerce, *Journal of Peace Research* **44**, 157–175.

(See also **Species competition; Bayesian methods and modeling; Hierarchical model**)

GRACE S. CHIU