Ecological Network Metrics: Opportunities for Synthesis

Matthew K. Lau[1†], Stuart R. Borrett[2,3], Benjamin Baiser[4], Nicholas J. Gotelli[5], Aaron M. Ellison[1]

1 Harvard Forest, Harvard University, Petersham, MA 02138

2 Department of Biology and Marine Biology, University of North Carolina, Wilmington NC 28403

3 Duke Network Analysis Center, Social Science Research Institute, Duke University, NC 27708

4 Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611

5 Department of Biology, University of Vermont, Burlington, VT 05405

† Corresponding author: Matthew K. Lau (matthewklau@fas.harvard.edu)

Abstract

Network ecology provides a systems basis for approaching ecological questions, such as factors that influence biological diversity, the role of particular species or particular traits in structuring ecosystems, and long-term ecological dynamics (e.g., stability). Whereas the introduction of network theory has enabled ecologists to quantify not only the degree, but also the architecture of ecological complexity, these advances have come at the cost of introducing new challenges, including new theoretical concepts and metrics, and increased data complexity and computational intensity. Synthesizing recent developments in the network ecology literature, we point to several potential solutions to these issues: integrating network metrics and their terminology across sub-disciplines; benchmarking new network algorithms and models to increase mechanistic understanding; and improving tools for sharing ecological network research, in particular "model" data provenance, to increase the reproducibility of network models and analyses. We propose that applying these solutions will aid in synthesizing ecological subdisciplines and allied fields by improving the accessibility of network methods and models.

Keywords: benchmarking; computation; data provenance; metrics; network ecology; systems analysis

Introduction

Interactions are at the heart of ecology and drive many of its key questions. What are the roles of species interactions in ecological systems? When and why is biological diversity important? What factors influence the long-term dynamics of ecosystems? These are all questions with a long history in ecology (Cherrett 1989; Council 2003; Lubchenco et al. 1991; Sutherland et al. 2013) that are not addressed in isolation. Points of intersection include the relationship between diversity and stability (May 2001, 2006); the identity and role of species that are the main drivers of community structure (e.g., keystone species, Paine 1966), ecosystem engineers (Jones et al. 1994), or foundation species (Dayton 1972; Ellison et al. 2005); and the causes and consequences of introducing new species into existing assemblages (Baiser et al. 2008; Simberloff and Holle 1999). Furthermore, "systems thinking" has been a persistent thread throughout the history of ecology (Margalef 1963; Odum and Pinkerton 1955; Patten 1978; Patten and Auble 1981; Ulanowicz 1986), dating back at least to Darwin's *Origin of Species* in his famous pondering of an entangled bank (Bascompte and Jordano 2014; Golley 1993). The application of network theory has provided a formal, mathematical framework to approach systems (Bascompte and Jordano 2014; Proulx et al. 2005) and led to the development of network ecology (Borrett et al. 2014; Patten and Witkamp 1967; Poisot et al. 2016*b*).

Network ecology can be defined as the use of network models and analyses to investigate the structure, function, and evolution of ecological systems at many scales and levels of organization (Borrett et al. 2012; Eklöf et al. 2012). The influx of network thinking throughout ecology, and ecology's contribution to the development of network science highlights the assertion that "networks are everywhere" (Lima 2011). And, as one would expect, the field has grown rapidly, from 1% of the primary ecological literature in 1991 to over 6% in 2017 (Fig. 1A). Some examples include: applying network theory to population dynamics and spread of infectious diseases (May 2006); description and analysis of networks of proteins in adult organisms (Stumpf et al. 2007) or during development (Hollenberg 2007); expanding classical food webs to include parasites and non-trophic interactions (Ings et al. 2009; Kéfi et al. 2012); investigating animal movement patterns (Lédée et al. 2016) and the spatial structure of metapopulations (Dubois et al. 2016; Holstein et al. 2014); connecting biodiversity to ecosystem functioning (Creamer et al. 2016); identifying keystone species (Borrett 2013; Zhao et al. 2016); and using social network theory in studies of animal behavior (Croft et al. 2004; Fletcher et al. 2013; Krause et al. 2003; Sih et al. 2009). Further, ideas and concepts from network ecology are being applied to investigate the sustainability of urban and industrial systems (Fang et al. 2014;

Layton et al. 2016; Xia et al. 2016) and elements of the food-energy-water nexus (Wang and Chen 2016; Yang and Chen 2016).

Over the past 15 years, re-occurring themes for moving network ecology forward have emerged from reviews, perspectives, and syntheses (e.g., Bascompte 2010; Borrett et al. 2014; Poisot et al. 2015; Proulx et al. 2005). In this paper, we examine areas where the network approach is being applied to address important ecological questions and identify both challenges and opportunities for advancing the field. Among these are the need for shifting the focus toward mechanisms rather than observations, and increasing the resolution (e.g., individuals or traits as nodes and weighted edges of different interaction types) and replication of network models across different ecosystems and time (Ings et al. 2009; Poisot et al. 2016*b*; Woodward et al. 2010). After a brief primer of key concepts from network ecology, we discuss the following topics as they relate to these issues: the proliferation of terminology for ecological metrics with the increasing application of network methods; fully exploring the underlying assumptions of models of mechanistic processes for generating network structure; and the need for improved sharing and reproducibility of ecological network research and models. Although these topics are not new, the combination of the influx of metrics and theory and rapid increases in the computational intensity of ecology are creating novel challenges. With respect to these issues, we discuss recent advances that should be explored as tools to aid in a more effective integration of network methods for synthesis across ecological sub-disciplines.

A primer of ecological networks: models and metrics

Prior to the introduction of network methods in ecology, the primary way of studying interactions was limited to detailed studies of behaviors and traits of individual species important to interactions, or of relationships between tightly interacting pairs of species (Carmel et al. 2013). Some ecologists were advancing whole-system methods (Lindeman 1942; Odum 1957); however, quantifying interactions is costly, as compared to surveys of species abundances. This has created a significant barrier to studying interactions at the scale of entire communities, either at the scale of individuals or species pairs, because the number of interactions becomes intractable. For instance, even if one assumes that only pairwise interactions occur among S species, the number of possible pairs is $S(S-1)/2$. Local assemblages of macrobes often have $10^1 - 10^2$ species, and microbial diversity can easily exceed 10^3 OTUs (Operational Taxonomic Units).

This complexity of ecological systems is one reason there is a long tradition in community ecology of studying interactions within small subsets of closely-related species (e.g. trophic guilds) and using dimensionality reducing methods based on multivariate, correlative approaches (Legendre et al. 2012). While some approaches to studying subsets of species incorporate the underlying pattern of direct and indirect links (e. g., modules, (*sensu* Holt 1997; Holt and Hoopes 2005), the majority do not. Such limitations repeatedly have led to calls for the application of "network thinking" to ecological questions (e.g., Golubski et al. 2016; Ings et al. 2009; Jacoby and Freeman 2016; Patten and Witkamp 1967; Proulx et al. 2005; QUINTESSENCE Consortium et al. 2016; Urban and Keitt 2001). There are now many resources for learning about network ecology and network theory in general, and we point the reader in the direction of excellent reviews in this area (Bascompte and Jordano 2007; Borrett et al. 2012; Brandes et al. 2013; Ings et al. 2009; Proulx et al. 2005) and more comprehensive introductions (Brandes and Erlebach 2005; Estrada 2015; Newman 2010).

Network ecology employs network theory to quantify the structure of ecological interactions. All networks consist of sets of interacting nodes (e.g. species, non-living nutrient pools, habitat patches) whose relationships are represented by edges (e.g. nutrient or energy transfers, pollination, movement of individuals). Conceptually, a network is a set of things or objects with connections among them. Stated mathematically, a network is a generic relational-model comprised of a set of objects represented by nodes or vertices (N) and a set of edges (E) that map one or more relationships among the nodes, $G = (N, E)$. A canonical ecological example of a network is a food-web diagram, in which the nodes represent species, groups of species, or non-living resources, and the *edges* map the relationship who-*eats*-whom.

The analysis of networks is inherently hierarchical, ranging from the entire network down to individual nodes and edges. Depending on the characteristics and level of detail of the information provided for a given model, there is a large number of network analyses and metrics that can be used to characterize the system at multiple levels (similar to Hines and Borrett, 2014; Wasserman and Faust, 1994), including: (1) the whole network level (i.e., the entire network), (2) the sub-network level (i.e., groups of two or more nodes and their edges), and (3) the individual node or edge level (Fig. 2).

Network-level metrics integrate information over the entire set of nodes and edges. For example, the number of nodes (e.g., the species richness of a food web) and the density of connections or connectance are both network-level statistics used to describes the overall complexity of a network and have been investigated by ecologists for over 40 years (Allesina and Tang 2012; May 1972).

Sub-network level analyses focus on identifying specific subsets of nodes and edges. There are a variety of groups that have different names (e.g., module, motif, cluster, clique, environ) and different methods for measurement. Sub-networks often represent more tractable and meaningful units of study than individual nodes and edges on the one hand or entire networks on the other. For example, in landscape and population ecology, the preferential movement of individuals and genes (edges) between habitat patches (nodes) has implications for conservation of populations and the design of preserves (Calabrese and Fagan 2004; Fletcher et al. 2013; Holt and Hoopes 2005). Also, both nodes and edges can be divided into classes. An example of this is the bipartite graph, in which interactions occur primarily between, rather than within, each class or "part" of the community. A bipartite network has only two classes of nodes, such as in a pollination network in which the community is divided into plants being pollinated and insects that do the pollination (Petanidou et al. 2008). In this network, edges representing pollination visits can only map between two nodes in the different classes.

Metrics at the individual node or edge level quantify differences in relative importance. Whether we are interested in an individual or species that transmits disease, species whose removal will result in secondary extinctions, or key habitat patches that connect fragmented landscapes, identifying important nodes is a critical component of network analysis. Another type of node or edge-level metric classifies nodes or edges according to their roles within a network. This classification can use information from differing levels. Additionally, nodes and edges can have variable characteristics. Edges can be weighted and they can map a directed relationship (as opposed to a symmetric or undirected relationship). For example, in ecosystem networks, the edges show the directed movement of energy or nutrients from one node to another by some process like feeding, and the edge weight can indicate the amount of energy or mass in the transaction (Baird and Ulanowicz 1989; Dame and Patten 1981). Nodes also can be weighted (e.g., size of individual, population size, biomass of a given species). Lastly, network models are flexible enough to accommodate variation in edge types and relationships among edges (e.g., hypergraphs), but analysis of these more complicated models is challenging and has only begun to be applied in ecology (e.g., Golubski et al. 2016).

Resolving network metrics

The application of network theory defines an explicit mathematical formalism that provides a potentially unifying set of terms for ecology and its inter-disciplinary applications (QUINTESSENCE Consortium et al. 2016). Ironically, the development of ecological network metrics has had an opposing affect. One reason for this is that introductions have occurred in multiple sub-disciplinary branches (Fig. 1B) (Blüthgen 2010; Borrett et al. 2014; Carmel et al. 2013). Having separate research trajectories can facilitate rapid development of ideas and the process of integration can lead to novel insights (Hodges 2008). At the same time, these innovations in network ecology have come at the cost of the "rediscovery" of the same network metrics and subsequent description of them with new terms. This has led to different metrics with similar purposes existing in separate areas of ecology (Table 1).

Ecological studies using network approaches draw from a deep well of general network theory (Newman 2003, 2006; Strogatz 2001). Ecologists broadly use network concepts, techniques, and tools to: (1) characterize the system organization (Borrett 2013; Croft et al. 2004; Ulanowicz 1986); (2) investigate the consequences of the network organization (Borrett et al. 2006; Dunne et al. 2002; Grilli et al. 2016); and (3) identify the processes or mechanisms that might generate the observed patterns (Allesina and Pascual 2008; Fath et al. 2007; Guimarães et al. 2007; Poisot et al. 2016*b*; Ulanowicz et al. 2014; Williams and Martinez 2000). The unnecessary proliferation of network metrics is exemplified by "connectance" (C), which is used by food-web ecologists to mean the ratio of the number of edges in the network divided by the total number of possible edges. Elsewhere in the network science literature, this measurement is referred to as network density (Newman et al. 2001). As another example, what ecosystem ecologists have described as "average path length" (total system through-flow divided by the total system input) (Finn 1976) also has been called network aggradation (Jørgensen et al. 2000). In economics, average path length is known as the multiplier effect (Samuelson 1948).

Another kind of redundancy is the creation and use of multiple statistics that measure the same or very similar network aspects. A clear example of this is inherent in the proliferation of centrality measures to indicate node or edge importance. Network scientists have shown that many centrality metrics are correlated (Jordán et al. 2007; Newman 2006; Valente et al. 2008). Likewise, Borrett and Osidele (2007) found that nine commonly reported ecosystem network analysis metrics covaried in 90 plausible parameterizations of a model of phosphorus biogeochemical cycling for Lake Lanier, GA, but that all these metrics were associated strongly with only two underlying factors. However, even a perfect correlation does not mean that two metrics have identical properties, and they still may diverge in different models. Therefore, it is important to have mathematically based comparisons of metrics (Borgatti and Everett 2006; Borrett 2013; Kazanci and Ma 2015; Ludovisi and Scharler 2017). It is incumbent on network ecologists to establish clearly the independence and uniqueness of the descriptive metrics used.

From the perspective of the broader field of ecology, the proliferation of concepts, terms, and metrics is not a new issue (e.g., Ellison et al. 2005; Tansley 1935). Ecologists have a long history of using network concepts and related models in multiple subdomains (e.g., metapopulations, matrix population models, community co-occurrence models, ecosystems) without fully recognizing or capitalizing on the similarities of the underlying models. Each subdomain has constructed its own concepts and methods (occasionally borrowing from other areas), and established its own jargon that impedes scientific development. Previous suggestions for solving this issue have focused on maintaining an historical perspective of ecology (Graham and Dayton 2002); Blüthgen et al. (2008) is an excellent example of how this can be done through peer-reviewed literature.

One possible approach that would go beyond such a diffuse, literature-centered approach would be to develop a formal ontology of concepts and metrics. An ontology is a a set of related terms that are formally defined and supported by assertions (Bard and Rhee 2004). An ontology therefore provides a framework for developing concepts within a discipline and presents the opportunity for more efficient synthesis across disciplinary boundaries. The concept of an ontology is not new, but more rapid sharing of ontologies and their collaborative development have been enabled by the Internet. For example, the Open Biological Ontologies (OBO, `http://www.obofoundry.org`) supports the creation and sharing of ontologies over the web. Currently, there is no OBO for a "network ecology metric" ontology, and as far as we are aware, ontologies have yet to be explored or developed for network metrics.

The OBO could provide a platform for harmonizing ecological network metrics, terms, and concepts. Key obstacles to such harmonization include a requirement that network ecologists work within a common framework, and the need for an individual or leadership team to periodically curate the ontology based on new developments in the field. In determining the best course of action, network ecologists could follow the example of how similar OBO projects have been managed in the past. The *FOODON* food role ontology project (`http://www.obofoundry.org/ontology/foodon.html`) contains information about "materials in natural ecosystems and food webs as well as human-centric categorization and handling of food." It could serve as an example or even the basis of a ecological network metric ontology.

Benchmarking: Trusting our models of mechanisms

Inferences about processes in ecological systems have relied in part on the application of simulation models that generate matrices with predictable properties. As discussed in the previous section, the proliferation of network metrics points to the need for the investigation and comparison of how these metrics will behave in the context of different modeling algorithms. Once a metric or algorithm has been chosen, it is tempting apply them widely to empirical systems to detect patterns, but before research proceeds, a process of "benchmarking" with artificial matrices that have predefined amounts of structure and randomness should be used to examine the behavior of the algorithms and the metrics that are applied to them.

Benchmarking of ecological models developed from null model analysis in community ecology (Atmar and Patterson 1993; Connor and Simberloff 1979; Gotelli and Ulrich 2012). Null models are specific examples of randomization or Monte Carlo tests (Manly 2007) that estimate a frequentist P value, the tail probability of obtaining the value of some metric if the null hypothesis were true (Gotelli and Graves 1996). The aim of a null model is to determine if the structure of an observed ecological pattern in space or time is incongruous with what would be expected given the absence of a causal mechanism. A metric of structure calculated for a single empirical data set is compared to the distribution of the same metric calculated for a collection of a large number of randomizations of the empirical data set. The data are typically randomized by reshuffling some elements while holding other elements constant to incorporate realistic constraints. Comparison with a suite of null models in which different constraints are systematically imposed or relaxed may provide a better understanding of the factors that contribute most to patterns in the network (see Box 1). However, the devil remains in the details and there are also a variety of ways to randomize data and impose constraints to construct a useful null model. If the null model is too simplistic, such as a model in which edges and nodes are sampled with uniform probability, it will always be rejected and provide little insight into important ecological patterns, regardless of what metric is used. At the other extreme, if the null model incorporates too many constraints from the data, it will be difficult or impossible to reject the null hypothesis, even though the network may contain interesting structure.

In network theory, the Erdos-Renyi (ER, (Erdös and Rényi 1959)) model is a now-classic example of a model used to generate networks via a random process for creating matrix structure. The ER model is a random graph that starts with an $N \times N$ adjacency matrix of nodes and assigns to it K edges between randomly chosen pairs of nodes. The ER model has been applied in ecology to address questions about the relationship between stability and complexity (May 1972) and the structure of genetic networks (Kauffman et al. 2003). For example, randomized networks have been used to link motifs (Milo et al. 2002) to network assembly (Baiser et al. 2016), stability (Allesina and Pascual 2008; Borrelli et al. 2015), and persistence in food webs (Stouffer and Bascompte 2010).

In addition to the random matrix approaches of null and ER models, there are other, more complex algorithms that are used to generate structured matrices. Perhaps one of the best known in network theory is the Barabasi-Albert (BA, Barabási and Albert 1999) model, which adds nodes and edges to a growing network with a greater probability of adding edges to nodes with a higher degree. The BA algorithm is similar to ecological network algorithms that generate non-random structure, because of either direct influence or similar processes operating in systems of interest. Some of these models include processes of "preferential attachment" in which organisms tend to interact with the same, common species. Food-web modeling algorithms also have been developed that use a trait-based approach (e.g., Allesina and Pascual 2009), consumer-resource models (Yodzis and Innes 1992), niches (Williams and Martinez 2000), cyber-ecosystem algorithms (Fath 2004), and cascade models (Allesina and Pascual 2009; Allesina and Tang 2012; Cohen and Łuczak 1992).

The statistical behavior of some models and metrics can be understood analytically. For example, the networks generated by the BA algorithm display degree distributions that approximate a power-law distribution, like many real-world "scale-free" networks (Albert et al. 2002). If the network is sparse (i.e., ($K \ll N^2$)), the degree distribution of the network should follow a Poisson distribution. However, as new models and metrics are introduced, new benchmarking should be done and compared to previous results. Newman et al. (2016) is one example of how benchmarking can be used for investigating processes operating on ecological networks. Ludovisi and Scharler (2017) advocate the same approach for the analysis of network models in general. The `benchmark` (Eugster and Leisch 2008) package in R (R Core Team 2017) is a general algorithm-testing software package that provides a useful starting point.

Reproducibility: Open-data, Open-source and Provenance

As analyses of network models increase in computational intensity, there is a concomitant increase in the need for new tools to track and share key computational details. This need is compounded when models incorporate data from multiple sources or analyses involve random processes. The combination of the volume of data and computational intensity of studies of ecological networks further increases the burden on ecologists to provide information needed to adequately reproduce datasets, analyses, and results. As the sharing and reproducibility of scientific studies are both essential for advances to have lasting impact, finding easier, faster, and generally more convenient ways to record and report relevant information for ecological network studies is imperative for advancing the field.

Sharing data and open-source code have become established in ecology, and network ecologists are now producing more network models and data (e.g., Fig. 1A). These include not only ecological interaction networks, but also an influx of other relevant networks, including ecological genomic networks generated by next-generation, high-throughput sequencing technologies (Langfelder and Horvath 2008; Zinkgraf et al. 2017). There are now multiple web-accessible scientific databases (e.g., NCBI, Data Dryad, Dataverse) and at least four databases have been constructed specifically to curate ecological network data: including "Kelpforest" (Beas-Luna et al. 2014), "The Web of Life" (Fortuna et al. 2014), "Mangal" ecological network database (Poisot et al. 2015) and the "Interaction Web Database" (`https://www.nceas.ucsb.edu/interactionweb/resources.html`).

The increase in ecological network data is linked to an increasing rate of shared analytical code and other open-source software. It is now commonplace for ecologists to have a working knowledge of one or more programming languages, such as R, Python, SAS, MatLab, Mathematica, or SPSS. Multiple software packages exist for doing ecological analyses, including ecological network analyses. In addition to the general network analysis packages available in R, there are at least two packages aimed specifically at ecological network analysis: `bipartite` and `enaR`. The former provides functions drawn largely from community ecology (Dormann et al. 2009), whereas the latter provides a suite of algorithms developed in the ecosystem network analysis literature (Borrett and Lau 2014; Lau et al. 2015).

Although, ecology has long had a culture of keeping records of important research details, such as field and lab notebooks, these practices put all of the burden of recording "metadata" on the researcher. Manual record-keeping methods, even when conforming to metadata standards (e.g., EML, see Boose et al. 2007), do not take advantage of the power of the computational environment. Data-provenance methods aim to provide a means to collect formalized information about computational processes, ideally in a way that aids the reproducibility of studies with minimal impact on the day-to-day activities of researchers (Boose et al. 2007). These techniques have been applied in other areas of research and could provide an effective means for documenting the source and processing of data from the raw state into a model (Boose and Lerner 2017).

The reproducibility of scientific studies is imperative for advances to have lasting impact through the independent verification of results. Although this has been an ongoing topic of discussion in ecology (Ellison 2010; Parker et al. 2016), the need was highlighted by a recent survey finding issues with reproduction of studies across many scientific disciplines (Baker 2016). There is significant motivation from within the ecological community to move toward providing detailed information about computational workflows for both repeatability and reproducibility, which includes repetition by the original investigator (Lowndes et al. 2017). It is also important in network ecology for data sources and methods for model construction be standardized and transparent, and that models be curated and shared (McNutt et al. 2016).

Collecting details, such as those enabled by data-provenance capture software, is one innovative way forward. These tools have been developing in the computer-science domain for decades; however, only recently have they gained a foothold in ecology (Boose et al. 2007; Ellison 2010) or the broader scientific community. Although there are many challenges in the development and application of data-provenance principles, multiple software packages do exist for collecting data provenance in the context of scientific investigations. Two provenance capture packages exist in R, the `recordr` package associated with the DataOne repository (Cao et al. 2016) and `RDataTracker` (Lerner and Boose 2014). In addition, although they do not collect formal data provenance, there are methods developed for "literate computing" that help to collect code along with details about the code and the intention of the analyses (e.g., the Jupyter notebook project: (Shen and Barabasi 2014)).

For ecological networks, there is software that captures the "data pedigree" of food-web models, but it does not capture data provenance. Data pedigree was initially implemented in the EcoPath food-web modeling package (Guesnet et al. 2015; Heymans et al. 2016) to define confidence intervals and precision estimates for network edges. It has been developed further to allow for the use of informative priors in Bayesian modeling of ecological networks. This is done by linking models to the literature sources from which estimates were derived, an approach that is similar to incorporating metadata information within databases of ecological networks. Although this approach focuses only on a subcomponent of provenance, this still is a promising way to address the issue that networks, network metrics, and simulation models used to analyze them commonly assume a lack of uncertainty (e.g., Borrett and Osidele 2007; Kauffman et al. 2003; Kones et al. 2009), and typically ignore inaccuracy in the empirical data (Ascough et al. 2008; Gregr and Chan 2014).

Moving Forward

Development and application of new technologies (e.g., sequencing methods and computational, data-driven approaches) have the potential to increase both the abundance and quality of ecological networks. For the future development of network ecology, there is a pressing need not only to share data and code, but also to integrate and use the large amounts of information enabled by technological advances. For example, synthetic networks (i.e., merging network models from different studies Poisot et al. 2016*a*) are a promising new direction; however, the structural properties of synthetic networks and the behavior of network metrics applied to them will require careful investigation, including the application of systematic benchmarking. Multi-trophic networks provide a precedence for these studies to move forward, but synthesizing models from across many different sources produces new challenges for developing and benchmarking metrics, as well as an opportunity for new technologies, like data provenance, to help establish better connections among studies and researchers.

The burgeoning of "open" culture in the sciences (Hampton et al. 2014) also has the potential to serve as a resource for models and a clearinghouse for resolving the validity of metrics, models, and algorithms. First, because code is openly shared, functions used to calculate metrics are open for inspection and, if coded and documented clearly using software "best-practices" (e.g., Noble 2009; Visser et al. 2015), the code provides a transparent documentation of how a metric is implemented and its computational similarity to other metrics. Second, enabled by the ability to write their own functions and code, researchers can do numerical investigations of the similarities among metrics. Through comparison of metrics calculated on the same or similar network models, a researcher could at least argue, for a given set of models, that two or more metrics produce similar results. Third, data provenance provides a useful tool to aide in the dissemination and synthesis of network models and increases the reproducibility of ecological network studies, including those documenting new metrics and benchmarking those metrics and associated algorithms for generating or analyzing empirical models. Last, as with data provenance, formalizing ecological network metrics and concepts requires a mathematically rigorous foundation that is developed by the community of researchers working along parallel lines of inquiry. Whether this is done through an ontological approach or some other formalized "clearing-house," an open process of exchange that integrates multiple perspectives is essential to prevent the rapid dilution of concepts in ecological network research as these concepts continue to proliferate, develop and evolve.

Over half a century ago, Robert MacArthur published his first paper on the relationship between diversity and stability, initiating multiple research trajectories that have now become the mainstay of many ecological research programs (MacArthur 1955). The theory that MacArthur applied was based on flows of energy through networks of interacting species. Thus, network theory is at the roots of one of the most widely studied topics in ecology and is now a part of the broader context of integration across many scientific disciplines that is aimed at consilience of theory (Wilson 1999). The synthesis of ecological concepts through the mathematically rigorous "lingua franca" of network terminology has the potential to unify theories across disciplines. As with previous concepts (e.g., keystone species, foundation species, ecosystem engineer), greater clarity and less redundancy will come about as network methods are used more commonly and researchers compare the mathematical and computational underpinnings of the metrics that they are using. With the increased use of these approaches, the network concept has and will continue to serve as a common model that transcends disciplines and has the potential to serve as an inroad for new approaches. With thoughtful dialogue across sub-disciplines and among research groups, further infusion of network theory and methods will continue to advance ecology.

Acknowledgments

This work was supported by the US National Science Foundation under grant SSI-1450277 End-to-End Provenance.

Author contributions statement

All authors contributed to the conception, writing and review of the manuscript.

References

Albert, R., A. L. Barabasi, and A.-L. Barabási. 2002. Statistical mechanics of complex networks. Reviews of Modern Physics **74**:47–97.

Allesina, S., and M. Pascual. 2008. Network structure, predatorprey modules, and stability in large food webs. Theoretical Ecology **1**:55–64.

Allesina, S., and M. Pascual. 2009. Food web models: A plea for groups. Ecology Letters **12**:652–662.

Allesina, S., and S. Tang. 2012. Stability criteria for complex ecosystems. Nature **483**:205–8.

Ascough, J., H. Maier, J. Ravalico, and M. Strudley. 2008. Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. Ecological Modelling **219**:383–399.

Atmar, W., and B. D. Patterson. 1993. The measure of order and disorder in the distribution of species in fragmented habitat. Oecologia **96**:373–382.

Baird, D., and R. E. Ulanowicz. 1989. The seasonal dynamics of the Chesapeake Bay ecosystem. Ecological Monographs **59**:329–364.

Baiser, B., R. Elhesha, and T. Kahveci. 2016. Motifs in the assembly of food web networks. Oikos **125**:480–491.

Baiser, B., J. L. Lockwood, D. La Puma, and M. F. J. Aronson. 2008. A perfect storm: two ecosystem engineers interact to degrade deciduous forests of New Jersey. Biological Invasions **10**:785–795.

Baker, M. 2016. 1,500 scientists lift the lid on reproducibility. Nature **533**:452–454.

Barabási, A.-L., and R. Albert. 1999. Emergence of scaling in random networks. Science **286**:509–512.

Barabási, A.-L., R. Albert, and H. Jeong. 2000. Scale-free characteristics of random networks: the topology of the world-wide web. Physica A: statistical mechanics and its applications **281**:69–77.

Bard, J. B. L., and S. Y. Rhee. 2004. Ontologies in biology: design, applications and future challenges. Nature Reviews Genetics **5**:213–222.

Bascompte, J. 2010. Structure and dynamics of ecological networks. Science **329**:765–6.

Bascompte, J., and P. Jordano. 2007. Plant-Animal Mutualistic Networks: The Architecture of Biodiversity. Annual Review of Ecology, Evolution, and Systematics **38**:567–593.

Bascompte, J., and P. Jordano. 2014. Mutualistic networks. Princeton University Press.

Beas-Luna, R., M. Novak, M. H. Carr, M. T. Tinker, A. Black, J. E. Caselle, M. Hoban, D. Malone, and A. Iles. 2014. An online database for informing ecological network models: http://kelpforest.ucsc.edu. PLoS One **9**:e109356.

Blüthgen, N. 2010. Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. Basic and Applied Ecology **11**:185–195.

Blüthgen, N., J. Fründ, D. P. Vázquez, and F. Menzel. 2008. What do interaction network metrics tell us about specialization and biological traits? Ecology **89**:3387–3399.

Bonacich, P. 1987. Power and Centrality: A Family of Measures. American Journal of Sociology **92**:1170.

Boose, E. R., A. M. Ellison, L. J. Osterweil, L. A. Clarke, R. Podorozhny, J. L. Hadley, A. Wise, and D. R. Foster. 2007. Ensuring reliable datasets for environmental models and forecasts. Ecological Informatics **2**:237–247.

Boose, E. R., and B. S. Lerner, 2017. Replication of data analyses: Provenance in R. Pages 195–212 *in* A. Shavit and A. M. Ellison, editors. Stepping in the Same River Twice: Replication in Biological Research. Yale University Press, New Haven, Connecticut, USA.

Borgatti, S. P., and M. G. Everett. 2006. A Graph-theoretic perspective on centrality. Social Networks **28**:466–484.

Borrelli, J. J., S. Allesina, P. Amarasekare, R. Arditi, I. Chase, J. Damuth, R. D. Holt, D. O. Logofet, M. Novak, R. P. Rohr, A. G. Rossberg, M. Spencer, J. K. Tran, and L. R. Ginzburg. 2015. Selection on stability across ecological scales. Trends in Ecology & Evolution **30**:417–425.

Borrett, S. R. 2013. Throughflow centrality is a global indicator of the functional importance of species in ecosystems. Ecological Indicators **32**:182–196.

Borrett, S. R., W. Bridewell, P. Langely, and K. R. Arrigo. 2006. A method for representing and developing process models. Ecological Complexity **4**:28.

Borrett, S. R., R. R. Christian, and R. E. Ulanowicz, 2012. Network Ecology (Revised). Pages 1767–1772 *in* A. El-Shaarawi and W. Piegorsch, editors. Encyclopedia of Environmetrics (2nd edition). John Wiley and Sons, Chinchester, second edition.

Borrett, S. R., B. D. Fath, and B. C. Patten. 2007. Functional integration of ecological networks through pathway proliferation. J. Theor. Biol. **245**:98–111.

Borrett, S. R., and M. K. Lau. 2014. enaR: An R package for Ecosystem Network Analysis. Methods in Ecology and Evolution **11**:1206–1213.

Borrett, S. R., J. Moody, and A. Edelmann. 2014. The rise of Network Ecology: Maps of the topic diversity and scientific collaboration. Ecological Modelling page 18.

Borrett, S. R., and O. O. Osidele. 2007. Environ indicator sensitivity to flux uncertainty in a phosphorus model of Lake Sidney Lanier, USA. Ecological Modelling **200**:371–383.

Borrvall, C., B. Ebenman, and T. Jonsson. 2000. Biodiversity lessens the risk of cascading extinction in model food webs. Ecology Letters **3**:131–136.

Brandes, U., and T. Erlebach. 2005. Network analysis: methodological foundations. Springer.

Brandes, U., G. Robins, A. McCranie, and S. Wasserman. 2013. What is network science? Network Science **1**:1–15.

Calabrese, J. M., and W. F. Fagan. 2004. A comparison-shopper's guide to connectivity metrics. Frontiers in Ecology and the Environment **2**:529–536.

Cao, Y., C. Jones, V. Cuevas-Vicenttín, M. B. Jones, B. Ludäscher, T. McPhillips, P. Missier, C. Schwalm, P. Slaughter, D. Vieglais, L. Walker, and Y. Wei, 2016. DataONE: A Data Federation with Provenance Support. Pages 230–234 . Springer International Publishing.

Carmel, Y., R. Kent, A. Bar-Massada, L. Blank, J. Liberzon, O. Nezer, G. Sapir, and R. Federman. 2013. Trends in ecological research during the last three decades–a systematic review. PLoS One **8**:e59813.

Cherrett, J. M., 1989. Key concepts: The results of a survey of our members' opinions. Pages 1–16 *in* J. M. Cherrett, A. D. Bradshaw, F. B. Goldsmith, P. G. Grubb, and J. R. Krebs, editors. Ecological concepts: The contribution of ecology to an understanding of the natural world. Blackwell Scientific Publications, Oxford, UK.

Cohen, J. E., and T. Łuczak. 1992. Trophic levels in community food webs. Evolutionary Ecology **6**:73–89.

Colwell, R. K., and D. W. Winkler. 1984. A null model for null models in biogeography. Princeton University Press.

Connor, E. F., and D. Simberloff. 1979. The Assembly of Species Communities: Chance or Competition? Ecology **60**:1132.

Council, N. R. 2003. Neon. National Academies Press, Washington, D.C.

Creamer, R., S. Hannula, J. Van Leeuwen, D. Stone, M. Rutgers, R. Schmelz, P. De Ruiter, N. B. Hendriksen, T. Bolger, M.-L. Bouffaud, et al. 2016. Ecological network analysis reveals the inter-connection between soil biodiversity and ecosystem function as affected by land use across Europe. Applied Soil Ecology **97**:112–124.

Croft, D. P., J. Krause, and R. James. 2004. Social networks in the guppy *Poecilia reticulata*. Proc. Royal Soc. Lond. B. **271**:S516–S519.

Dame, R. F., and B. C. Patten. 1981. Analysis of energy flows in an intertidal oyster reef. Marine Ecology Progress Series **5**:115–124.

Dayton, P. K. 1972. Toward an understanding of community resilience and the potential effects of enrichment to the benthos at McMurdo Sound, Antarctica. Proceedings of the Colloquium on Conservation Problems in Antarctica pages 81–96.

Dormann, C. F., C. F. Dormann, J. Fründ, N. Blüthgen, and B. Gruber. 2009. Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. Open Ecology Journal **2**:7–24.

Dubois, M., V. Rossi, E. Ser-Giacomi, S. Arnaud-Haond, C. López, and E. Hernández-García. 2016. Linking basin-scale connectivity, oceanography and population dynamics for the conservation and management of marine ecosystems. Global Ecology and Biogeography **25**:503–515.

Dunne, J. a., R. J. Williams, and N. D. Martinez. 2002. Network structure and biodiversity loss in food webs: Robustness increases with connectance. Ecology Letters **5**:558–567.

Eklöf, A., and B. Ebenman. 2006. Species loss and secondary extinctions in simple and complex model communities. Journal of Animal Ecology **75**:239–246.

Eklöf, A., M. R. Helmus, M. Moore, and S. Allesina. 2012. Relevance of evolutionary history for food web structure. Proc. Royal Soc. B **279**:1588–96.

Ellison, A. M. 2010. Repeatability and transparency in ecological research. Ecology **91**:2536–2539.

Ellison, A. M., M. S. Bank, B. D. Clinton, E. A. Colburn, K. Elliott, C. R. Ford, D. R. Foster, B. D. Kloeppel, J. D. Knoepp, G. M. Lovett, J. Mohan, D. A. Orwig, N. L. Rodenhouse, W. V. Sobczak, K. A. Stinson, J. K. Stone, C. M. Swan, J. Thompson, B. Von Holle, and J. R. Webster. 2005. Loss of foundation species: consequences for the structure and dynamics of forested ecosystems. Frontiers in Ecology and the Environment **3**:479–486.

Erdös, P., and a. Rényi. 1959. On random graphs. Publicationes Mathematicae **6**:290–297.

Estrada, E., 2015. Introduction to Complex Networks: Structure and Dynamics. Pages 93–131 . Springer International Publishing.

Eugster, M. J. A., and F. Leisch, 2008. Bench Plot and Mixed Effects Models: First Steps toward a Comprehensive Benchmark Analysis Toolbox. Pages 299–306 *in* P. Brito, editor. Compstat 2008—Proceedings in Computational Statistics. Physica Verlag, Heidelberg, Germany.

Fang, D., B. D. Fath, B. Chen, and U. M. Scharler. 2014. Network environ analysis for socio-economic water system. Ecological Indicators **47**:80–88.

Fath, B. D. 2004. Network analysis applied to large-scale cyber-ecosystems. Ecological Modeling **171**:329–337.

Fath, B. D., U. M. Scharler, R. E. Ulanowicz, and B. Hannon. 2007. Ecological network analysis: network construction. Ecological Modelling **208**:49–55.

Finn, J. T. 1976. Measures of ecosystem structure and function derived from analysis of flows. Journal of theoretical biology **56**:363–380.

Finn, J. T. 1980. Flow analysis of models of the Hubbard Brook ecosystem. Ecology **61**:562–571.

Fletcher, R. J., A. Revell, B. E. Reichert, W. M. Kitchens, J. D. Dixon, and J. D. Austin. 2013. Network modularity reveals critical scales for connectivity in ecology and evolution. Nature Communications **4**:2572–2576.

Fortuna, M. A., R. Ortega, and J. Bascompte. 2014. The Web of Life. www.web-of-life.es .

Freeman, L. C. 1979. Centrality in networks. I. Conceptual clarificaiton. Social Networks **1**:215–239.

Golley, F. 1993. A history of the ecosystem concept in ecology: More than the sum of the parts. Yale University Press, New Haven, CT.

Golubski, A. J., E. E. Westlund, J. Vandermeer, and M. Pascual. 2016. Ecological Networks over the Edge: Hypergraph Trait-Mediated Indirect Interaction (TMII) Structure. Trends in Ecology & Evolution .

Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. Ecology **81**:2606–2621.

Gotelli, N. J., R. M. Dorazio, A. M. Ellison, and G. D. Grossman. 2010. Detecting temporal trends in species assemblages with bootstrapping procedures and hierarchical models. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **365**:3621–3631.

Gotelli, N. J., A. M. Ellison, and B. A. Ballif. 2012. Environmental proteomics, biodiversity statistics and food-web structure. Trends in Ecology & Evolution **27**:436–42.

Gotelli, N. J., and G. R. Graves. 1996. Null models in ecology. Ecology **14**:368.

Gotelli, N. J., and W. Ulrich, 2012. Statistical challenges in null model analysis.

Graham, M. H., and P. K. Dayton. 2002. On the evolution of ecological ideas: paradigms and scientific progress. Ecology **83**:1481–1489.

Gregr, E. J., and K. M. A. Chan. 2014. Leaps of Faith: How Implicit Assumptions Compromise the Utility of Ecosystem Models for Decision-making. BioScience **65**:43–54.

Grilli, J., T. Rogers, and S. Allesina. 2016. Modularity and stability in ecological communities. Nature Communications **7**:12031–12041.

Guesnet, V., G. Lassalle, A. Chaalali, K. Kearney, B. Saint-Béat, B. Karimi, B. Grami, S. Tecchio, N. Niquil, and J. Lobry. 2015. Incorporating food-web parameter uncertainty into Ecopath-derived ecological network indicators. Ecological Modelling **313**:29–40.

Guimarães, P. R., G. Machado, M. A. M. de Aguiar, P. Jordano, J. Bascompte, A. Pinheiro, and S. F. Dos Reis. 2007. Build-up mechanisms determining the topology of mutualistic networks. Journal of theoretical biology **249**:181–9.

Hampton, S. E., S. Anderson, S. C. Bagby, and C. Gries. 2014. The Tao of Open Science for Ecology. PeerJ pages 1–30.

Heymans, J. J., M. Coll, J. S. Link, S. Mackinson, J. Steenbeek, C. Walters, and V. Christensen. 2016. Best practice in Ecopath with Ecosim food-web models for ecosystem-based management. Ecological Modelling **331**:173–184.

Hines, D. E., and S. R. Borrett. 2014. A comparison of network, neighborhood, and node levels of analyses in two models of nitrogen cycling in the Cape Fear River Estuary. Ecological Modelling **293**:210–220.

Hodges, K. E. 2008. Defining the problem: terminology and progress in ecology. Frontiers in Ecology and the Environment **6**:35–42.

Hollenberg, D. 2007. On the evolution and dynamics of biological networks. Rivista di biologia **100**:93–118.

Holstein, D. M., C. B. Paris, and P. J. Mumby. 2014. Consistency and inconsistency in multispecies population network dynamics of coral reef ecosystems. Marine Ecology Progress Series **499**:1–18.

Holt, R., 1997. Community modules. Pages 333–349 *in* Multitrophic interactions in terrestrial ecosystems, 36th Symposium of the British Ecological Society. Blackwell Science Oxford.

Holt, R. D., and M. F. Hoopes. 2005. Food web dynamics in a metacommunity context. Metacommunities: Spatial dynamics and ecological communities pages 68–93.

Ings, T. C., J. M. Montoya, J. Bascompte, N. Blüthgen, L. Brown, C. F. Dormann, F. Edwards, D. Figueroa, U. Jacob, J. I. Jones, R. B. Lauridsen, M. E. Ledger, H. M. Lewis, J. M. Olesen, F. J. F. van Veen, P. H. Warren, and G. Woodward. 2009. Ecological networks–beyond food webs. The Journal of animal ecology **78**:253–69.

Jacoby, D. M. P., and R. Freeman. 2016. Emerging network-based tools in movement ecology. Trends in Ecology & Evolution **31**:301–314.

Jones, C. G., J. H. Lawton, and M. Shachak. 1994. Organisms as Ecosystem Engineers. Oikos **69**:373–386.

Jordán, F., Z. Benedek, and J. Podani. 2007. Quantifying positional importance in food webs: A comparison of centrality indices. Ecological Modelling **205**:270–275.

Jørgensen, S. E., B. C. Patten, and M. Straškraba. 2000. Ecosystems emerging: 4. Growth. Ecological Modelling **126**:249–284.

Kauffman, S., C. Peterson, B. R. Samuelsson, C. Troein, and P. W. Anderson. 2003. Random Boolean network models and the yeast transcriptional network. PNAS **100**:14796–14799.

Kazanci, C., and Q. Ma, 2015. Chapter 3 System-wide measures in ecological network analysis. Pages 45–68 *in* Y.-S. Park, S. Lek, C. Baehr, and S. E. Jorgensen, editors. Advanced Modelling Techniques Studying Global Changes in Environmental Sciences, volume 27. Elsever.

Kéfi, S., E. L. Berlow, E. A. Wieters, S. A. Navarrete, O. L. Petchey, S. A. Wood, A. Boit, L. N. Joppa, K. D. Lafferty, R. J. Williams, N. D. Martinez, B. A. Menge, C. A. Blanchette, A. C. Iles, and U. Brose. 2012. More than a meal{. . . } integrating non-feeding interactions into food webs. Ecology Letters **15**:291–300.

Kones, J. K., K. Soetaert, D. van Oevelen, and J. O. Owino. 2009. Are network indices robust indicators of food web functioning? A Monte Carlo approach. Ecological Modelling **220**:370–382.

Krause, A. E., K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. W. Taylor. 2003. Compartments revealed in food-web structure. Nature **426**:282–5.

Langfelder, P., and S. Horvath. 2008. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics **9**:559.

Lau, M. K., S. R. Borrett, D. E. Hines, and P. Singh, 2015. enaR: Tools for Ecological Network Analysis.

Layton, A., B. Bras, and M. Weissburg. 2016. Ecological Principles and Metrics for Improving Material Cycling Structures in Manufacturing Networks. Journal of Manufacturing Science and Engineering **138**:101002.

Lédée, E. J. I., M. R. Heupel, A. J. Tobin, A. Mapleston, and C. A. Simpfendorfer. 2016. Movement patterns of two carangid species in inshore habitats characterised using network analysis. Marine Ecology Progress Series **553**:219–232.

Legendre, P., L. Legendre, L. Legendre, and P. Legendre. 2012. Numerical ecology. Elsevier.

Lerner, B., and E. Boose, 2014. RDataTracker: Collecting Provenance in an Interactive Scripting Environment. Pages 1–4 *in* 6th USENIX Workshop on the Theory and Practice of Provenance (TaPP 2014). USENIX Association, Cologne.

Lima, M. 2011. Visual Complexity: Mapping Patterns of Information. Princeton Architectural Press.

Lindeman, R. L. 1942. The trophic-dynamic aspect of ecology. Ecology **23**:399–418.

Lowndes, J. S. S., B. D. Best, C. Scarborough, J. C. Afflerbach, M. R. Frazier, C. C. O'Hara, N. Jiang, and B. S. Halpern. 2017. Our path to better science in less time using open data science tools. Nature Ecology & Evolution **1**:0160.

Lubchenco, J., A. M. Olson, L. B. Brubaker, S. R. Carpenter, M. M. Holland, S. P. Hubbell, S. A. Levin, J. A. MacMahon, P. A. Matson, J. M. Melillo, H. A. Mooney, C. H. Peterson, and H. Ronald Pulliam. 1991. The Sustainable Biosphere Initiative: An Ecological Research Agenda: A Report from the Ecological Society of America. Risser Source: Ecology **72**:371–412.

Ludovisi, A., and U. M. Scharler. 2017. Towards a sounder interpretation of entropy-based indicators in ecological network analysis. Ecological Indicators **72**:726–737.

MacArthur, R. 1955. Fluctuations of Animal Populations and a Measure of Community Stability. Ecology **36**:533.

Manly, B. F. J. 2007. Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall.

Margalef, R. 1963. Certain unifying principles in ecology. The American Naturalist **97**:357–374.

Martinez, N. D., 1992. Constant Connectance in Community Food Webs.

Maslov, S., and K. Sneppen. 2002. Specificity and stability in topology of protein networks. Science **296**:910–3.

May, R. M. 1972. Will a Large Complex System be Stable? Nature **238**:413–414.

May, R. M. 2001. Stability and Complexity in Model Ecosystems. Princeton University Press.

May, R. M. 2006. Network structure and the biology of populations. Trends in Ecology & Evolution **21**:394–399.

McNutt, M., K. Lehnert, B. Hanson, B. A. Nosek, A. M. Ellison, and J. L. King. 2016. Liberating field science samples and data. Science **351**:1024–1026.

Milo, R., S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, S. H. Strogatz, D. Watts, S. H. Strogatz, A.-L. Barabási, R. Albert, M. Newman, H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A. L. Barabasi, R. F. Cancho, C. Janssen, R. V. Sole, R. F. Cancho, R. V. Sole, L. Amaral, A. Scala, M. Barthelemy, H. Stanley, B. Huberman, L. Adamic, S. Shen-Orr, R. Milo, S. Mangan, U. Alon, N. Guelzim, S. Bottani, P. Bourgine, F. Kepes, M. Newman, S. H. Strogatz, D. Watts, S. Maslov, K. Sneppen, D. Thieffry, A. M. Huerta, E. Perez-Rueda, J. Collado-Vides, M. C. Costanzo, R. Williams, N. Martinez, S. Pimm, J. Lawton, J. Cohen, J. White, E. Southgate, J. Thomson, S. Brenner, D. Callaway, J. Hopcroft, J. Kleinberg, M. Newman, and S. H. Strogatz. 2002. Network motifs: simple building blocks of complex networks. Science **298**:824–7.

Newman, M. 2010. Networks an Introduction. Oxford University Press.

Newman, M. E. J., 2003. The Structure and Function of Complex Networks.

Newman, M. E. J. 2006. Modularity and community structure in networks. PNAS **103**:8577–82.

Newman, M. E. J., A. Clauset, C. Aicher, A. Z. Jacobs, A. Clauset, S. Fortunato, P. Holme, M. Huss, H. Jeong, R. Guimerà, L. A. N. Amaral, D. Hric, R. K. Darst, S. Fortunato, M. Barthélemy, A. Z. Jacobs, A. Clauset, K. Zuev, G. B. M. Boguñá, D. Krioukov, B. H. Good, Y.-A. d. Montjoye, A. Clauset, C. Bothorel, J. D. Cruz, M. Magnani, B. Micenková, N. Binkiewicz, J. T. Vogelstein, K. Rohe, E. Galbrun, A. Gionis, N. Tatti, T. J. Hansen, M. W. Mahoney, Y. Zhang, E. Levina, J. Zhu, P. Expert, T. S. Evans, V. D. Blondel, R. Lambiotte, L. Peel, P. Zhang, C. Moore, L. Zdeborová, X. Maa, L. Gaoa, X. Yongb, L. Fua, Z.-Y. Zhang, P. W. Holland, K. B. Laskey, S. Leinhardt, B. Karrer, M. E. J. Newman, A. Decelle, F. Krzakala, C. Moore, L. Zdeborová, E. Mossel, J. Neeman, A. Sly, J. Moody, L. Danon, J. Duch, A. Diaz-Guilera, A. Arenas, A. F. McDaid, D. Greene, N. Hurley, U. Brose, G. Woodward, A. L. Traud, P. J. Mucha, M. A. Porter, P. C. Bull, D. B. Larremore, A. Clauset, C. Z. Buckee, D. B. Larremore, A. Clauset, A. Z. Jacobs, G. M. Warimwe, P. C. Bull, A. Clauset, C. Moore, M. E. J. Newman, R. Milo, S. P. Borgatti, M. G. Everett, B. Ball, M. E. J. Newman, P. D. Hoff,

A. E. Raferty, and M. S. Handcock. 2016. Structure and inference in annotated networks. Nature Communications **7**:11863.

Newman, M. E. J., S. H. Strogatz, and D. J. Watts. 2001. Random graphs with arbitrary degree distributions and their applications. Physical Review E **64**:026118.

Noble, W. S. 2009. A quick guide to organizing computational biology projects. PLoS computational biology **5**:e1000424.

Odum, H. T. 1957. Trophic structure and productivity of Silver Springs, Florida. Ecol. Mono. **27**:55–112.

Odum, H. T., and R. C. Pinkerton. 1955. Time's speed regulator: the optimum efficiency for maximum power output in physical and biological systems. American Scientist **43**:331–343.

Paine, R. T. 1966. Food Web Complexity and Species Diversity. The American Naturalist **100**:65.

Parker, T. H., S. Nakagawa, and J. Gurevitch. 2016. Promoting transparency in evolutionary biology and ecology. Ecology Letters **19**:726–728.

Patten, B. C. 1978. Systems approach to the concept of environment. Ohio Journal of Science **78**:206–222.

Patten, B. C., and G. T. Auble. 1981. System theory of the ecological niche. American Naturalist **117**:893–922.

Patten, B. C., and M. Witkamp. 1967. Systems analysis of ˆ134cesium kinetics in terrestrial microcosms. Ecology **48**:813–824.

Petanidou, T., A. S. Kallimanis, J. Tzanopoulos, S. P. Sgardelis, and J. D. Pantis. 2008. Long-term observation of a pollination network: Fluctuation in species and interactions, relative invariance of network structure and implications for estimates of specialization. Ecology Letters **11**:564–575.

Pimm, S. L. 1982. Food webs. Chapman and Hall, London; New York.

Poisot, T., B. Baiser, J. A. Dunne, S. Kéfi, F. Massol, N. Mouquet, T. N. Romanuk, D. B. Stouffer, S. A. Wood, and D. Gravel. 2015. mangal making ecological network analysis simple. Ecography **39**:384–390.

Poisot, T., D. Gravel, S. Leroux, S. A. Wood, M.-J. Fortin, B. Baiser, A. R. Cirtwill, M. B. Araújo, and D. B. Stouffer. 2016*a*. Synthetic datasets and community tools for the rapid testing of ecological hypotheses. Ecography **39**:402–408.

Poisot, T., D. B. Stouffer, and S. Kéfi. 2016*b*. Describe, understand and predict: why do we need networks in ecology? Functional Ecology **30**:1878–1882.

Post, D. M., M. L. Pace, and N. G. Hairston. 2000. Ecosystem size determines food-chain length in lakes. Nature **405**:1047–1049.

Proulx, S. R., D. E. L. Promislow, and P. C. Phillips. 2005. Network thinking in ecology and evolution. Trends in Ecology & Evolution **20**:345–53.

QUINTESSENCE Consortium, M. E., S. Carpenter, e. al., G. Daily, e. al., C. Raudsepp-Hearne, e. al., I. P. o. B. Services, Ecosystem, J. Silvertown, M. Schröter, e. al., S. Naeem, e. al., N. Biggs, e. al., A. Mashaghi, e. al., T. Ideker, R. Sharan, J. Gardy, e. al., T. Valente, D. Acemoglu, e. al., M. Jackson, M. Janssen, e. al., G. Woodward, e. al., C. Mulder, J. Elser, R. Thompson, e. al., J. Montoya, e. al., O. Petchey, e. al., J. Reiss, e. al., A. Barabasi, R. Albert, F. Harary, R. Haines-Young, S. Macfadyen, e. al., K. Rathwell, G. Peterson, R. Costanza, I. Kubiszewski, C. Mulder, e. al., H. Tallis, e. al., A. Tavoni, S. Levin, A. Ma, R. Mondragón, M. Pascual, J. Dunne, F. Jordán, e. al., H. Ernstson, e. al., U. Narloch, e. al., M. Hagen, e. al., R. Stewart, e. al., A. Shmida, M. Wilson, M. Palmer, e. al., G. Mace, e. al., R. d. Groot, e. al., H. Zimmermann, R. Albert, e. al., M. Pocock, e. al., K. Chan, e. al., L. Dicks, e. al., G. McInerny, e. al., M. Pocock, e. al., A. Vespignani, D. Bohan, e. al., P. Anderson, J. Cohen, e. al., J. Montoya, e. al., S. Pimm, C. Holling, A. Trichard, e. al., D. Bohan, and e. al. 2016. Networking Our Way to Better Ecosystem Service Provision. Trends in Ecology & Evolution **31**:105–15.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Samuelson, P. A. 1948. Economics: An Introductory Analysis. McGraw–Hill Book Co., New York,.

Shen, H.-W., and A.-L. Barabasi. 2014. Collective credit allocation in science. PNAS **111**:12325–12330.

Sih, A., S. F. Hanser, and K. a. McHugh. 2009. Social network theory: New insights and issues for behavioral ecologists. Behavioral Ecology and Sociobiology **63**:975–988.

Simberloff, D., and B. V. Holle. 1999. Positive interactions of nonindigenous species: invasional meltdown? Biological Invasions pages 21–32.

Solé, R. V., and J. M. Montoya. 2001. Complexity and fragility in ecological networks. Proc. Royal Soc. B **268**:2039–2045.

Stouffer, D. B., and J. Bascompte. 2010. Understanding food-web persistence from local to global scales. Ecology Letters **13**:154–161.

Strogatz, S. H. 2001. Exploring complex networks. Nature **410**:268–76.

Stumpf, M. P. H., W. P. Kelly, T. Thorne, and C. Wiuf. 2007. Evolution at the system level: the natural history of protein interaction networks. Trends in Ecology & Evolution **22**:366–373.

Sutherland, W. J., R. P. Freckleton, H. C. J. Godfray, S. R. Beissinger, T. Benton, D. D. Cameron, Y. Carmel, D. A. Coomes, T. Coulson, M. C. Emmerson, R. S. Hails, G. C. Hays, D. J. Hodgson, M. J. Hutchings, D. Johnson, J. P. G. Jones, M. J. Keeling, H. Kokko, W. E. Kunin, X. Lambin, O. T. Lewis, Y. Malhi, N. Mieszkowska, E. J. Milner-Gulland, K. Norris, A. B. Phillimore, D. W. Purves, J. M. Reid, D. C. Reuman, K. Thompson, J. M. J. Travis, L. A. Turnbull, D. A. Wardle, and T. Wiegand. 2013. Identification of 100 fundamental ecological questions. Journal of Ecology **101**:58–67.

Tansley, A. G. 1935. The Use and Abuse of Vegetational Concepts and Terms. Ecology **16**:284–307.

Ulanowicz, R. E., 1986. Introduction. Pages 1–8 *in* Growth and Development. Springer New York, New York, NY.

Ulanowicz, R. E., R. D. Holt, and M. Barfield. 2014. Limits on ecosystem trophic complexity: insights from ecological network analysis. Ecology letters **17**:127–36.

Ulrich, W., and N. J. Gotelli. 2007. Null model analysis of species nestedness patterns. Ecology **88**:1824–1831.

Ulrich, W., and N. J. Gotelli. 2010. Null model analysis of species associations using abundance data. Ecology **91**:3384–97.

Urban, D., and T. Keitt. 2001. Landscape connectivity: A graph-theoretic perspective. Ecology **82**:1205–1218.

Valente, T. W., K. Coronges, C. Lakon, and E. Costenbader. 2008. How Correlated Are Network Centrality Measures? Connections **28**:16–26.

Vermaat, J. E., J. a. Dunne, and A. J. Gilbert. 2009. Major dimensions in food-web structure properties. Ecology **90**:278–282.

Visser, M. D., S. M. McMahon, C. Merow, P. M. Dixon, S. Record, E. Jongejans, W. Michener, M. Jones, S. Petrovskii, N. Petrovskaya, A. Ellison, B. Dennis, G. Hager, G. Wellein, P. Zuidema, E. Jongejans, P. Chien, H. During, F. Schieving, B. V. Putten, M. Visser, H. Muller-Landau, P. Jansen, L. Comita, H. Muller-Landau, S. Aguilar, S. Hubbell, C. Merow, N. LaFleur, J. Silander, A. Wilson, M. Rubega, M. Visser, E. Jongejans, M. V. Breugel, P. Zuidema, Y.-Y. Chen, A. Kassim, J. Chambers, B. Kernighan, P. Plauger, G. Amdahl, A. Porter, R. Selby, R. Bryant, D. OHallaron, M. Schmidberger, M. Morgan, D. Eddelbuettel, H. Yu, L. Tierney, U. Mansmann, A. Grama, G. Karypis, V. Kumar, A. Gupta, P. LEcuyer, D. Eddelbuettel, R. François, D. Eddelbuettel, C. Sanderson, H. Robbins, S. Monro, A. Mantoglou, J. Wilson, A. Finley, Z. Merali, A. Guisan, A. Lehmann, S. Ferrier, M. Austin, J. Overton, B. Brook, J. OGrady, A. Chapman, M. Burgman, H. Akçakaya, R. Frankham, F. Isbell, V. Calcagno, A. Hector, J. Connolly, W. Harpole, P. Reich, E. V. Moran, J. Clark, G. Bohrer, G. Katul, R. Walko, and R. Avissar. 2015. Speeding Up Ecological and Evolutionary Computations in R; Essentials of High Performance Computing for Biologists. PLoS Computational Biology **11**:e1004140.

Wang, S., and B. Chen. 2016. Energy-water nexus of urban agglomeration based on multiregional input-output tables and ecological network analysis: A case study of the Beijing-Tianjin-Hebei region. Applied Energy **178**:773–783.

Wasserman, S., and K. Faust. 1994. Advances in Social Network Analysis: Research in the Social and Behavioral Sciences. SAGE Publications.

Williams, R. J., E. L. Berlow, J. a. Dunne, A.-L. Barabási, and N. D. Martinez. 2002. Two degrees of separation in complex food webs. PNAS **99**:12913–12916.

Williams, R. J., and N. D. Martinez. 2000. Simple rules yield complex food webs. Nature **404**:180–183.

Wilson, E. O. 1999. Consilience: The unity of knowledge. Vintage.

Woodward, G. U. Y., J. P. Benstead, O. S. Beveridge, J. Blanchard, T. Brey, L. E. E. E. Brown, W. F. Cross, N. Friberg, C. Ings, U. T. E. Jacob, S. Jennings, M. E.

Ledger, A. M. Milner, J. M. Montoya, E. O. Gorman, J. M. Olesen, O. L. Petchey, E. Pichler, D. C. Reuman, M. S. A. Thompson, F. J. F. V. a. N. Veen, G. Yvon-Durocher, T. C. Ings, U. T. E. Jacob, S. Jennings, M. E. Ledger, A. M. Milner, J. M. Montoya, E. O 'gorman, J. M. Olesen, O. L. Petchey, D. E. Pichler, D. C. Reuman, M. S. A. Thompson, F. J. F. Van Veen, and G. Yvon-Durocher. 2010. Ecological Networks in a Changing Climate. Advances In Ecological Research 42. Ecological Networks. **42**:72–120.

Xia, L., B. D. Fath, U. M. Scharler, and Y. Zhang. 2016. Spatial variation in the ecological relationships among the components of Beijing's carbon metabolic system. Science of the Total Environment **544**:103–113.

Yang, J., and B. Chen. 2016. Energy–water nexus of wind power generation systems. Applied Energy **169**:1–13.

Yodzis, P., and S. Innes. 1992. Body Size and Consumer-Resource Dynamics. The American Naturalist **139**:1151–1175.

Zhao, L., H. Zhang, E. J. O'Gorman, W. Tian, A. Ma, J. C. Moore, S. R. Borrett, and G. Woodward. 2016. Weighting and indirect effects identify keystone species in food webs. Ecology Letters **19**:1032–1040.

Zinkgraf, M., L. Liu, A. Groover, and V. Filkov. 2017. Identifying gene coexpression networks underlying the dynamic regulation of wood-forming tissues in *Populus* under diverse environmental conditions. New Phytologist **214**:1464–1478.

Boxes

Box 1. Benchmarking Ecological Models The most basic test is to feed the algorithm a set of "random" matrices to make sure that the frequency of statistically significant results is no greater than 5%. Otherwise, the algorithm is vulnerable to a Type I statistical error (incorrectly rejecting a true null hypothesis). However, specifying a matrix produced by random sampling errors is not so easy. By definition, if a null model algorithm is used to generate the random matrices, then no more than 5% of them should be statistically significant (unless there were programming errors). For binary matrices, two log-normal distributions can be used to generate realistic heterogeneity in row and column totals, while still maintaining additive effects for cell occurrence probabilities (Ulrich and Gotelli 2010). "Structured" matrices are needed to test for Type II errors (incorrectly accepting a false null hypothesis), and these require a careful consideration of exactly what sort of pattern or mechanism the test is designed to reveal. One approach is to begin with a perfectly structured matrix, such as one derived from a mechanistic model for generating network structure, contaminate it with increasing amounts of stochastic noise, and test for the statistical pattern at each step (Gotelli 2000). A plot of the P value versus the added noise should reveal an increasing curve, and will indicate the signal-to-noise ratio below which the test cannot distinguish the pattern from randomness. Alternatively, one can begin with a purely random matrix but embed in it a non-random substructure, such as a matrix clique or a node with extreme centrality. The size, density, and other attributes of this matrix can be manipulated to see whether the test can still detect the presence of the embedded structure (Gotelli et al. 2010). Because all null model tests (and all frequentist statistics) are affected by sample size and data structure, these benchmark tests can be tailored to the attributes of the empirical data structures for better focus and improved inference.

Even simple randomization algorithms may require further filters to ensure that random matrices retain a number of desirable network properties. For example, Dunne et al. (2002) created random food-web matrices with constant species richness and connectance, but they discarded webs with unconnected nodes and subwebs because these topologies were not observed in the empirical webs. A "stub reconstruction" algorithm builds a topology that is constrained to the observed number of edges per node (Newman et al. 2001). Each node is assigned the correct number of edges, and then nodes are successively and randomly paired to create a growing network. However, this algorithm also generates multiple edges between the same two nodes, which must be discarded or otherwise accounted for. Maslov and Sneppen (2002) use a "local re-wiring algorithm" that preserves the number of connections for every node by swapping edges randomly between different pairs of nodes. This algorithm is closely analogous to the swap algorithm used in species co-occurrence analyses that preserves the row and column totals of the original matrix (Connor and Simberloff 1979). The more constraints that are added to the algorithm, the less likely it is that simple sampling processes can account for patterns in the data. However, some constraints, such as connectivity or matrix density, may inadvertently "smuggle in" the very processes they are designed to detect. This can lead to the so-called "Narcissus" effect (Colwell and Winkler 1984). Finding the correct balance between realistic constraints and statistical power is not easy (Gotelli et al. 2012), and there are many potential algorithms that reasonably could be used, even for simple binary matrices (Gotelli 2000).

966 # Tables

Sub.discipline	Level	Metric	Concept	Reference
General	W	Density	The proportion of possible edges that are actually associated with nodes; called Connectance in Food Web ecology.	
General	N	Centrality	Multiple ways to characterize the relative importance of nodes.	Wasserman and Faust (1994)
General	N	Degree	Number of edges connected to a given node, which is a type of local centrality.	
General	N	Eigenvector Centrality	Global centrality metric based on number of walks that travel through a node	Bonacich (1987)
General	W	Centrality Distribution	Shape of the frequency distribution of edges among nodes.	Barabási and Albert (1999); Dunne et al. (2002)
General	W	Centralization	The concentration (versus evenness) of centrality among the nodes.	Freeman (1979)
General	W	Graph diameter	The longest path between any two nodes in a graph.	Barabási et al. (2000); Urban and Keitt (2001)
General	W	Modularity	Degree to which edges are distributed within rather than between distinct sets of nodes.	Newman (2010)
General	G	Motifs	Small sets of nudes with similar distributions of edges.	Milo et al. (2002)
General	W	Link density	Average number of edges per node.	Martinez (1992)
Community	N	Temperature	Measures the nestedness of a bipartite network.	Ulrich and Gotelli (2007)
Community	W	Co-occurrence	Degree of overlapping spatial or temporal distributions of species relative to a null model.	Gotelli (2000)
Community	N	Indicator Species	The degree to which the abundance of a taxonomic group responds to an environmental gradient.	
Community	W	Nestedness	The degree to which interactions can be arranged into subsets of the larger community	
Community	W	Evenness	Deviation of the distribution of observed abundances relative to an even distribution among taxonomic groups in a community	
Community	W	Diversity	Distribution of abundances among taxonomic groups in an observed community	
Community	W	Richness	The number of taxonomic groups in a community	
Community	W	Stability	The change in the abundances of taxonomic groups across a set of observations	
Food-Web	N	Removal Importance	The degree to which removal of a compartment or species produces subsequent removals in the ecosystem.	Borrvall et al. (2000); Dunne et al. (2002); Eklöf and Ebenman (2006); Solé and Montoya (2001)
General	N	Connectance	Proportion of realized out of possible edges	Pimm (1982); Vermaat et al. (2009)
Food-Web	G	Food-chain length	The number of feeding relationships among a set of compartments in a food-web.	Post et al. (2000); Ulanowicz et al. (2014)
Ecosystem	W	Finn cycling index	Degree to which matter or energy passes through the same set of compartments.	Finn (1980)
Ecosystem	G	Environ	The sub-network of the probability of movement of energy or matter among compartments generated by a single unit of input (output) into a selected node.	Patten (1978); Patten and Auble (1981)
Ecosystem	N	Throughflow	Amount of energy or matter passing into or out of a node	Finn (1976)
Ecosystem	N	Throughflow Centrality	The proportion of energy or matter that passes through a given compartment in an ecosystem.	Borrett (2013)
General	G	Chain Length	Number of edges between two nodes in a group	
Food-Web	G	Average Path Length	The average number of times a unit of matter or energy travels from one compartment to another before exiting the ecosystem	Finn (1976)
Ecosystem	W	Pathway Proliferation	Rate of increase in the number of edges between nodes with increasing path length	Borrett et al. (2007)
Ecosystem	W	Ascendency	Measures the average similarity in matter or energy flows among compartments in an ecosystem.	Ulanowicz (1986)
Food-Web	N	Trophic Level	Ordinal classification of a compartment or taxonomic group based on the relative position in the ecosystem.	Allesina and Pascual (2009); Fath (2004); Williams et al. (2002)

Table 1: Ecological network metric summary and classification. Level indicates the hierarchy of the metric (W = Whole network, G = Group or sub-network, N = Node). The Sub-disciplines include 'General' network theory, 'Community' ecology, 'Food-web' and 'Ecosystem' ecology. Also available at `https://figshare.com/s/1bf1a7e0a6ee3ac97a4b`.

967 # Figures

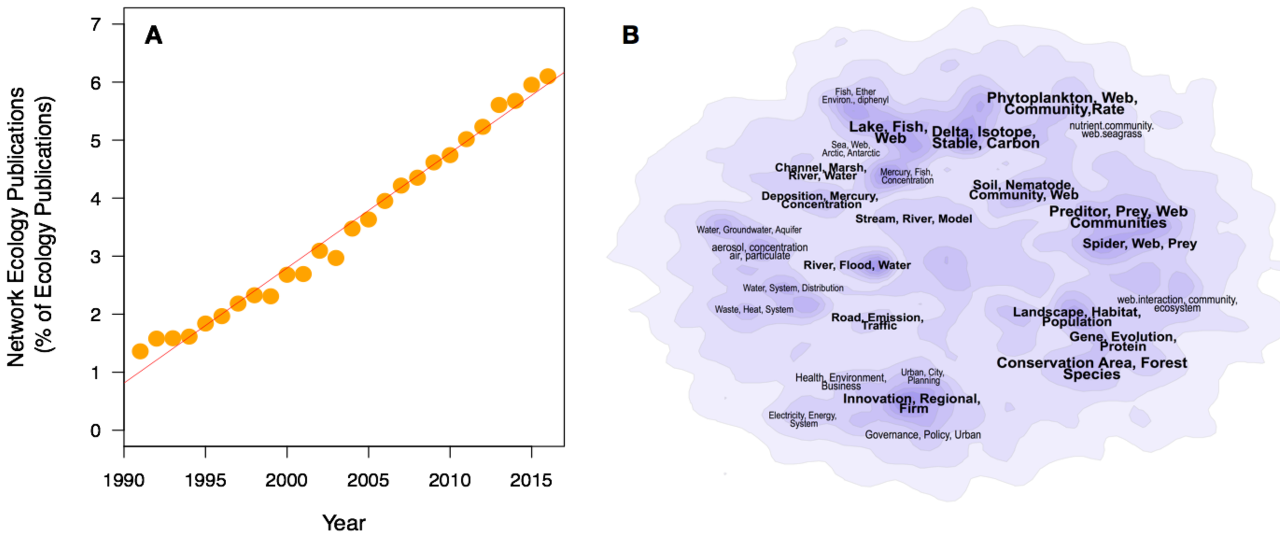

Figure 1: Although systems thinking has been a part of ecology since at least the work of Darwin, network ecology has grown rapidly since the turn of the last century but has been developing in isolated sub-fields. (A) Plot showing the increase in "network ecology" keywords in the literature from 1991 to current (updated using search developed by Borrett et al. (2014)). (B) Contour plot of common topics in network ecology with peaks indicating clusters of related topics. The regions are labeled with the most common terms found in the clusters. From Borrett et al. (2014), reproduced with permission.

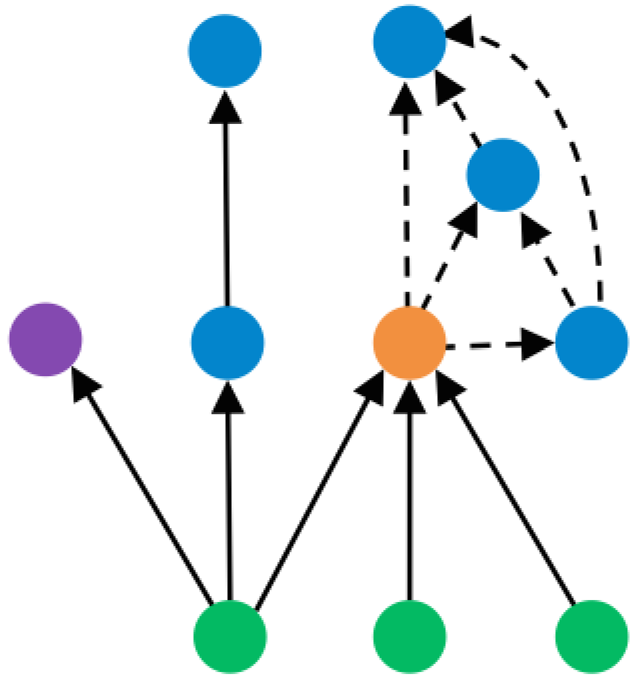

Figure 2: Hypothetical unweighted, directed network showing examples of the four classes of network metrics. *Node Level:* the purple node exhibits low centrality while the orange node exhibits high centrality. *Group or Sub-Network Level:* the blue nodes connected with dashed edges shows a module. *Global or Whole Network Level:* using the edges of all nodes we can measure the connectance of the entire network ($c = edges/nodes^2 = 0.12$).