Title: `enaR`: An R package for Ecological Network Analysis

Running Title: R ecological network analysis package

Word Count: 3025

Authors: Stuart R. Borrett, Matthew K. Lau

Addresses:

- SRB: Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, NC, 28403
- MKL: Department of Biological Sciences and the Merriam-Powell Center for Environmental Research, Northern Arizona University, 617 S. Beaver St., Flagstaff, AZ, 86011

Contact Details:

- Email: borretts@uncw.edu
- Phone: 910.962.2411
- Fax: 910.962.4066

enaR: An R package for Ecological Network Analysis

Stuart R. Borrett[a,*] and Matthew K. Lau[b]

[a] Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, NC

[b] Department of Biological Sciences and the Merriam-Powell Center for Environmental Research, Northern Arizona University, 617 S

[*] Corresponding author, borretts@uncw.edu

February 13, 2014

Abstract

- Ecology at its core is a science that studies relationships. As technological developments are making the collection of complex, relational datasets more common, a network approach is becoming an essential analytical method. Ecological Network Analysis (ENA) is an approach rooted in ecosystem ecology with over 30 years of development that investigates the structure, function, and evolution of ecological systems, such as food-webs.

- Here, we introduce `enaR`, a set of R tools that enables ecologists to perform a broad set of ENA algorithms to analyze the structure and dynamics of ecosystem network models.

- In addition to introducing the primary functionality of the package, we highlight several value added features including a library of 100 empirical ecosystem models, the ability to analyze multiple models simultaneously, and connections to other network and ecological analysis tools in R.

- We have created this package to enable more ecologists to apply ecological network analysis and contribute to ENA software development.

KEYWORDS: network analysis, ecosystem, social network analysis, software, network environ

analysis, ascendency, input–output analysis, food web, Ecopath, WAND

1 Introduction

Network Ecology – the study of ecological systems using network models and analyses to characterize their structure, function, and evolution – is a large and rapidly growing area of ecology. Borrett *et al.* (submitted) found that more than 5% of the ecology and evolutionary biology papers published in 2012 and indexed by Web of Science could be classified as Network Ecology. Likewise, Ings *et al.* (2009) showed that a notable fraction of 2008 publications in 11 select journals were related to food webs (≈2.4%), mutualistic networks (≈0.9%), and host-parasitoid networks (≈0.055%). Network ecology is growing in part because ecology is fundamentally a relational science and network models are excellent tools for relational analyses. In addition, this rise of network ecology contributes to, mirrors, and builds on the more general development of network sciences (Barabási, 2012; Borgatti & Foster, 2003; Freeman, 2004; Newman, 2003; Wasserman & Faust, 1994)

Ecological Network Analysis (ENA) is a branch of network ecology that is rooted in ecosystem ecology (Borrett *et al.*, 2012). It works like a "macroscope" to investigate (1) whole system organization, (2) the direct and indirect effects among system components, and (3) the processes that create and sustain ecological systems. More specifically, ENA is a family of algorithms that are an ecological application and extension of the economic Input-Output Analysis developed by Leontief (1966). These algorithms are applied to network models of energy and matter exchange among ecosystem components with the iconic example of this being the food-web (Fath & Patten, 1999; Hannon, 1973; Patten *et al.*, 1976; Ulanowicz, 1986).

The development of ENA has contributed to a new theoretical understanding of ecosystems (Belgrano *et al.*, 2005; Higashi & Burns, 1991; Jørgensen *et al.*, 2007; Ulanowicz, 1986) and the techniques have been applied in a multiple ways. For example, Patten (1982) used a storage analysis to identify two nearly separate hydrologic subsystems in the Okefenokee Swamp, USA. Bondavalli & Ulanowicz (1999) showed that in the Florida Everglades the American alligator is an indirect mutualist with several of its prey, including frogs. Hines *et al.* (2012) used ENA to study the Cape Fear River estuary sediment nitrogen cycle, and applied the tools to quantify the coupling between biogeochemical processes (e.g., nitrification + anammox). Furthermore, several scientists have used ENA to investigate urban sustainability (Bodini & Bondavalli, 2002; Bodini *et al.*, 2012; Chen & Chen, 2012; Zhang *et al.*, 2010). Collectively, this work consistently shows the power of the interaction network to transform relationships among system components in non-obvious ways that require whole-systems analysis to elucidate (Fath *et al.*, 2007; Patten, 1991; Ulanowicz & Puccia, 1990).

We have created `enaR`to provide open-source access to ENA tools. We had three specific design objectives for this software. The first objective was to collect the major ENA analyses in a single software package. The second objective was to increase both the availability and extensibility of the software. Users can freely download a stable version of the package from the CRAN website (`http://cran.r-project.org/web/packages/enaR/`). Better access to the code is provided through GitHub, where users can obtain the source code, make modifications, and add new functionality as techniques develop (`https://github.com/TheSeeLab/enaR`). We chose to implement the software in R because of its increasing popularity as an analytical tool in the biological sciences (e.g., Dixon, 2003; Metcalf *et al.*, 2012; Revell, 2012). The third design objective was to let users connect to other analytical tools. To enable this, `enaR` was built specifically to connect to two existing R network analysis packages: `network` (Butts, 2008a) and `sna` (Butts, 2008b). In summary, the aim of the `enaR` package is to make ENA tools more available and easier to use, adapt, and extend. In this paper, we present `enaR`with a brief illustration of its functionality. For a more detailed user introduction, please see the package vignette: i.e., `vignette`('enaR').

2 Overview of `enaR`

ENA is applied to network models of energy or matter flow and storage in an ecosystem. After describing the data required as input to ENA, we highlight the primary ENA algorithms currently included in `enaR` . We then walk through an application of the `enaR` Flow analysis to an example model.

2.1 Data Requirements and Input

While many influences combined to create what we now call ENA (e.g., Golley, 1993; Hannon, 1973; Margalef, 1963; Patten, 1959; Pimm, 1982), since the 1970s two primary schools of thought have developed (Scharler & Fath, 2009). The first is based on the work of Dr. Robert E. Ulanowicz, which was centered at the University of Maryland (Ulanowicz, 1986, 1997, 2009). The Ulanowicz school of ENA is primarily focused on trophic ecology, and its starting point is a phenomenological map of the energy–matter exchanges among ecosystem components. A key contribution of this work is the use of information theory and the development of the ascendency concept that Ulanowicz (1986, 1997) used to characterize ecosystem growth and development. The second school is based on the work of Dr. Bernard C. Patten at the University of Georgia (Fath & Patten, 1999; Matis & Patten, 1981; Patten, 1982; Patten *et al.*, 1976). Its initial perspective was steeped in dynamic equations, simulations, and systems analysis. A key contribution of this work is the environ concept that formalizes the concept of environment for study inside the network models (Patten, 1978). The Patten School of work has often been referred to as "Network Environ Analysis". The Ulanowicz and Patten School's of ENA represent two distinct but interwoven developments. Together, they join information theory, environmental concepts, and network science to study ecosystems.

For ENA, the system is modeled as a set of compartments or network nodes that represent species, species-complexes (i.e., trophic guilds or functional groups), or non-living components of the system in which energy or matter is stored. These nodes are connected by a set of observed fluxes, termed directed edges or links. These models also have energy–matter inputs into the system and output losses from the system. While the Patten School treats all outputs the same, the Ulanowicz School partitions outputs into respiration and export to account for differences in energetic quality. Note that the more generic outputs can be the sum of the respiration and export values. Some analyses also need the amount of energy–matter stored in each node (e.g., biomass). The final required information is a categorization of each node as living or not, which is essential for algorithms from the Ulanowicz School. We specified this node attribute by creating a logical vector that indicates whether the node is living (TRUE) or not (FALSE). In summary, the full set of data required to perform ENA includes (1) internal flows, (2) boundary inputs, (3) boundary exports, (4) boundary respiration, (5) boundary outputs, which may be the sum of exports and respiration, (6) biomass or storage values, and (7) designation of living status of each node.

Most analytical functions in `enaR` assume the model data is presented as an R network data object defined in the `network` package. Given the data elements, the `pack` function can be used to manually combine the data elements to create the necessary R network data object. While there is no standard data format for an ENA model, there are two commonly used formats. First, there is the Scientific Committee for Ocean Research (SCOR) format that is the required input to NETWRK (Ulanowicz & Kay, 1991), and the second format is the Excel sheet formatted data that is the input to WAND (Allesina & Bondavalli, 2004). The `enaR` package includes a `read.scor` and a `read.wand` function to read in these common data formats.

2.2 Included Algorithms

While the long-term goal is for the `enaR` package to be comprehensive, this initial release is more limited, but provides a foundation for future expansion. The package currently includes many of the most commonly used algorithms (Table 1), along with a number of work flow tools (e.g., the read.x functions). `enaR` captures all of the Patten School algorithms previously implemented in NEA.m, along with some recent developments. Ulanowicz School algorithms are more limited, including the ascendency calculations (Ulanowicz, 1997) and mixed trophic impacts analyses (Ulanowicz & Puccia, 1990). We hope to grow the package in time and through collaboration with users.

2.3 Example Application

Given a network model, applying ENA algorithms with `enaR` is straight forward. Table 2 illustrates applying the ENA Flow analysis to the six compartment model of energy flow in a South Carolina oyster reef (Dame & Patten, 1981). After loading the `enaR` package, the first step is to enter the model data. In this example, we use the `read.scor` function to read the SCOR formatted data from a text file. We can then apply one of four automated balancing algorithms introduced by Allesina & Bondavalli (AVG, Input-Output, Output-Input, AVG2, 2003) to ensure that the model is at steady-state — one of the assumptions of the flow analysis. In this example we used the default AVG2 algorithm, which tends to cause the least distortion of flows while balancing the network (Allesina & Bondavalli, 2003). We then applied the `enaFlow` function to the model to perform the desired ENA flow analysis. This analysis returns 4 matrices ($\mathbf{G}$, $\mathbf{GP}$, $\mathbf{N}$, $\mathbf{NP}$) and two vectors (throughflow, T, and a vector of 20 whole-network statistics, ns). Guidance for how to interpret these results can be found in previously published literature (Fath & Borrett, 2006; Schramski *et al.*, 2011).

3 Value Added Features

Beyond the basic functionality of the `enaR` package, there are several features that add substantive value for users. We highlight four of these features here: visualization, model library, batch analysis, and connections to other network analysis tools.

3.1 Visualization

Visualization of network models can be an essential analytical tool (Lima, 2011; Moody *et al.*, 2005). Because `enaR` is built on top of the `network` package and data type, it is possible to quickly create network plots of the model internal structure. Fig. 1a shows an example of the Oyster Reef ecosystem model. The `network` package includes three network layout algorithms: circle, Fruchterman-Reingold, and Kamada-Kawai. The Fruchterman-Reingold algorithm used here is the default.

3.2 Model Library

To facilitate new systems ecology and network science, we included a library of 100 previously published ecosystem network models with the `enaR` package. These models each trace a thermodynamically conserved unit (e.g., C, N, P) through a particular ecosystem. The models in this set are empirically-based in that the authors attempted to model a specific system and parameterized the model to some degree with empirical estimates. The library includes models used previously to test several systems ecology hypotheses (Borrett, 2013; Borrett & Salas, 2010; Borrett *et al.*, 2010; Salas & Borrett, 2011). This set has a 47% overlap with the set of models previously collected by Dr. Ulanowicz (`http://www.cbl.umces.edu/~ulan/ntwk/network.html`).

We have tentatively split these models into two classes. The most abundant class is the trophic network models. These models tend to have a food web at their core, but also include non-trophic fluxes generated by processes like death and excretion. The annual carbon flux model for the mesohaline region of the Chesapeake Bay is a typical example (Baird & Ulanowicz, 1989). The second class of models focuses on biogeochemical cycling. In contrast to the trophic networks, the biogeochemical cycling models tend to have more highly aggregated nodes (more species grouped into a compartment), include more abiotic nodes that could represent chemical species (e.g., ammonia in a nitrogen cycle), have a lower dissipation rate, and therefore they tend to have more recycling (Borrett *et al.*, 2010; Christian *et al.*, 1996). Christian & Thomas's (2003) models of nitrogen cycling in the Neuse River Estuary are good examples of the class. The package vignette has a full listing of the models included along with references to their original publications (Lau *et al.*, 2013).

3.3 Batch Analysis

Given a list of models like the model library, it is possible to efficiently batch apply one or more analyses to the models. This facilitates the kind of comparative network analysis often of interest to ecologists (Christian *et al.*, 2005; Monaco & Ulanowicz, 1997). For example, Christensen (1995) applied ENA to identify and compare the maturity of 41 ecosystem models, Baird *et al.* (2008) compared different nutrient dynamics in the Sylt-Rømø Bight ecosystem, and van Oevelen *et al.* (2011) compared the food webs and their organic matter processing in three sections of the Nazaré submarine canyon. The `enaR` tool simplifies the work flow for these types of comparison.

This batch analysis can be used in several additional ways. One application is for meta-analyses, such as tests of the generality of hypothesized ecosystem properties like network non-locality (Salas & Borrett, 2011), or to investigate how physical features might influence ENA results (Niquil *et al.*, 2012). Fig. 1b illustrates the rank-ordered network homogenization statistic for the 56 trophic-based ecosystem models in the library. Notice that the homogenization statistic is greater than one in all of these models indicating that the network of indirect interactions tend to more uniformly distribute the resources than is obvious from the direct interactions, which extends previous results of Borrett & Salas (2010) to include several new models. A second kind of application is the exploration of new

ENA inter-relationships. Given the collection of the Patten and Ulanowicz school algorithms and the library of models, the ENA community can investigate possible relationships among the ENA indicators from different schools (Fig. 1c). A third application of batch analysis is to investigate the previously unknown empirical ranges of ENA whole-network statistics, which may be useful for interpreting results from specific applications. Fig. 2 shows the observed distribution of values for selected network statistics from the 100 models in the library. The `enaR` package enables and simplifies these types of analysis.

3.4 New Connections

A fourth key feature of the `enaR` package design is that it enables network ecologists easier access to other network tools and analyses that might be useful. The `enaR` package uses the R network data structure defined in the `network` package (Butts, 2008a). This means that network ecologists using `enaR` can also use the network manipulation functions and visualization features of the `network` package. Further, the R Social Network Analysis (SNA) package, `sna`, (Butts, 2008b) also uses this network data object. This means that network ecologists can apply many of the SNA algorithms directly to their ecological network models. For example, Fig. 1d illustrates applying the betweenness centrality function to the Chesapeake Bay trophic model (Baird & Ulanowicz, 1989) and visualizing the results using the target centrality plot (Brandes *et al.*, 2003). This analysis highlights the central role of Sedimentary Particulate Carbon and bacteria in the Sediment Particulate Organic Carbon (POC) in the carbon flux of the estuary.

In addition, `enaR` can be a starting point for ecosystem network ecologists to use other R network tools. For example, the `iGraph` package provides functions to apply classic graph theory (Csardi & Nepusz, 2006). The `limSolve` package provides capabilities to infer network model fluxes from empirical data by linear inverse modeling (Soetaert *et al.*, 2009), which can also be used for uncertainty analyses of ENA (Kones *et al.*, 2009). There are a wealth of additional R package that network ecologists may find useful including `bipartite` (Dormann *et al.*, 2008), `vegan` (Dixon, 2003), `bioconductor` (Gentleman *et al.*, 2004), `Cheddar` (Hudson *et al.*, 2013), `Diversitree` (FitzJohn, 2012), and packages in the `statnet` family (Handcock *et al.*, 2008) beyond `network` and `sna`.

4 Conclusion and Future Development

Several software tools have been created to previously to enable scientists to apply ENA. The first widely distributed tool was NETWRK (Ulanowicz & Kay, 1991). This program is a collection of analyses programmed in Fortran and distributed as a DOS executable file. Version 4.2 is available from `http://www.cbl.umces.edu/~ulan/ntwk/network.html`. WAND is a Microsoft Excel based re-implementation of many but not all of the algorithms in NETWRK (Allesina & Bondavalli, 2004). An explicit goal of WAND was to be more accessible for ecologists, who have tended to be more familiar with Excel than DOS. Fath & Borrett (2006) introduced a Matlab function, NEA.m, which collected algorithms largely developed for network environ analysis, hence NEA (Patten, 1991). One advantage of NEA.m is that the algorithms are transparent to the user and accessible for modification. While the NEA.m function is freely available (`http://www.mathworks.com/matlabcentral/fileexchange/5261-nea-m`) it requires Matlab, which is powerful but expensive proprietary software. With modification, the function can be run in Octave, an open source clone of Matlab, but it executes more slowly and doesn't have the same level of support provided by Matlab. EcoNet is a web-based tool that lets users apply ENA analyses similar to to NEA.m, but with some computational enhancements (Kazanci, 2007; Schramski *et al.*, 2011). Ecopath with Ecosim (Christensen & Pauly, 1992; Christensen & Walters, 2004) is used primarily for model construction and simulation, but it also includes a network analysis plug-in that implements several other ENA algorithms. Other tools have been created, but do not appear to have a large user base (Kones *et al.*, 2009; Latham II, 2006). A challenge for ENA users has been that no existing software covers all of the major analyses, which has lead to separate, over-lapping approaches to ENA and high variation in software availability, usability, and extensibility.

The `enaR` package addresses many of the limitations of the previously published set of ENA tools. The library joins analyses from both the currently separate schools of ENA into a single software package. The library is built in R so that the functions are transparent and adaptable by the community of users. It also lets users have access to other network and statistical analysis tools that are already part of R. In the future, we anticipate two initial lines of continued development for the `enaR` package. The first is to increase the connections between the `enaR` package and other modeling and analytical tools. For example, we are currently working with colleagues to enable users of Ecopath with Ecosim (Christensen & Walters, 2004) to apply the `enaR` tools in a seamless way. We are also developing functions to connect between `enaR` and the R limSolve package (Soetaert *et al.*, 2009) for creating models using Linear Inverse Modeling and to enable uncertainty analysis (Kones *et al.*, 2009).

The second line of development is to extend the package's capabilities. While it currently contains most of the many commonly used ENA algorithms used by ecologists, it is far from complete. For example, Ulanowicz's (1983) decomposition of cycles is not yet included nor is his construction for the Lindeman trophic spine (Ulanowicz & Kemp, 1979). The package could also include network model construction tools, such as least-inference methods for building models from empirical data (Ulanowicz & Scharler, 2008) and Fath's (2004) algorithm for constructing plausible ecosystems models. Looking to the future of ENA, we hope to facilitate the rapid developemnt of accessible network analysis tools for the ecological community. A major reason for our use of open source software is that we want to foster user driven development and extension of the package's functionality. It is our hope that `enaR` can serve as an organizing point for ENA methods with the hope that a by doing so not only produce relevant software, but also promote the feedback bewteen theory and application of network analytics. Toward this end, we have developed the GitHub development repository (https://github.com/MKLau/enaR_development) and project page (http://theseelab.github.io/enaR/), where researchers can find more information on how to contribute software. Together, the open-source tools for vasersion control and project management provided by Git and GitHub will increase the potential for collaborative software development. We look forward to working with the dynamic community of network analysts to promote the use of network tools in ecology.

5 Acknowledgments

We would like to acknowledge and thank David Hines for contributing to the initial code. We also thank several individuals who used the earlier versions of the software and provided helpful feedback for further development including Ursula Scharler, Shaoqing Chen, Emily Oxe, and John Mejaski. In addition, we thank the many ecosystem model authors who created, shared, and published their work. This work was funded in part by the US National Science Foundation (DEB1020944, DEB0425908), an NSF Integrative Graduate Education and Research Traineeship (MKL; DGE0549505) and a UNCW Cahill award (SRB).

References

Allesina, S. & Bondavalli, C. (2003) Steady state of ecosystem flow networks: A comparison between balancing procedures. *Ecol Model*, **165**, 221–229.

Allesina, S. & Bondavalli, C. (2004) Wand: An ecological network analysis user-friendly tool. *Environ Model Softw*, **19**, 337–340.

Baird, D., Asmus, H. & Asmus, R. (2008) Nutrient dynamics in the Sylt-Rømø Bight ecosystem, German Wadden Sea: An ecological network analysis approach. *Estuar Coast Shelf Sci*, **80**, 339–356.

Baird, D. & Ulanowicz, R.E. (1989) The seasonal dynamics of the Chesapeake Bay ecosystem. *Ecol Monogr*, **59**, 329–364.

Barabási, A.L. (2012) The network takeover. *Nature Physics*, **8**, 14–16.

Belgrano, A., Scharler, U.M., Dunne, J. & Ulanowicz, R.E. (2005) *Aquatic Food Webs: An Ecosystem Approach*. Oxford University Press, New York, NY.

Bodini, A. & Bondavalli, C. (2002) Towards a sustainable use of water resources: a whole-ecosystem approach using network analysis. *Int J Environmental Pollution*, **18**, 463–485.

Bodini, A., Bondavalli, C. & Allesina, S. (2012) Cities as ecosystems: Growth, development and implications for sustainability. *Ecol Model*, **245**, 185–198.

Bondavalli, C. & Ulanowicz, R.E. (1999) Unexpected effects of predators upon their prey: The case of the American alligator. *Ecosystems*, **2**, 49–63.

Borgatti, S.P. & Foster, P.C. (2003) The network paradigm in organizational research: A review and typology. *J Manage*, **29**, 991–1013.

Borrett, S.R. (2013) Throughflow centrality is a global indicator of the functional importance of species in ecosystems. *Ecol Indic*, **32**, 182–196.

Borrett, S.R., Christian, R.R. & Ulanowicz, R.E. (2012) Network ecology. A.H. El-Shaarawi & W.W. Piegorsch, eds., *Encyclopedia of Environmetrics*, pp. 1767–1772. John Wiley & Sons, 2nd edition.

Borrett, S.R., Moody, J. & Edelmann, A. (submitted) The rise of network ecology: maps of the topic diversity and scientific collaboration. *Ecol Model*.

Borrett, S.R. & Salas, A.K. (2010) Evidence for resource homogenization in 50 trophic ecosystem networks. *Ecol Model*, **221**, 1710–1716.

Borrett, S.R., Whipple, S.J. & Patten, B.C. (2010) Rapid development of indirect effects in ecological networks. *Oikos*, **119**, 1136–1148.

Brandes, U., Kenis, P. & Wagner, D. (2003) Communicating centrality in policy network drawings. *IEEE Transactions on Visualization and Computer Graphics*, **9**, 241–253.

Butts, C. (2008a) network: A package for managing relational data in R. *J Stat Softw*, **24**.

Butts, C. (2008b) Social network analysis with sna. *J Stat Softw*, **24**, 1–51.

Chen, S. & Chen, B. (2012) Network environ perspective for urban metabolism and carbon emissions: A case study of Vienna, Austria. *Environ Sci Tech*, **46**, 4498–4506.

Christensen, V. (1995) Ecosystem maturity—towards quantification. *Ecol Model*, **77**, 3–32.

Christensen, V. & Pauly, D. (1992) Ecopath-II—a software for balancing steady-state ecosystem models and calculating network characteristics. *Ecol Model*, **61**, 169–185.

Christensen, V. & Walters, C.J. (2004) Ecopath with Ecosim: Methods, capabilities and limitations. *Ecol Model*, **172**, 109–139.

Christian, R.R., Baird, D., Luczkovich, J., Johnson, J.C., Scharler, U.M. & Ulanowicz, R.E. (2005) Role of network analysis in comparative ecosystem ecology of estuaries. A. Belgrano, J. Scharler U. M. Dunne & R. Ulanowicz, eds., *Aquatic Food Webs: An Ecosystem Approach*, pp. 25–40. Oxford University Press, New York, NY.

Christian, R.R., Fores, E., Comin, F., Viaroli, P., Naldi, M. & Ferrari, I. (1996) Nitrogen cycling networks of coastal ecosystems: influence of trophic status and primary producer form. *Ecol Model*, **87**, 111–129.

Christian, R.R. & Thomas, C.R. (2003) Network analysis of nitrogen inputs and cycling in the Neuse River Estuary, North Carolina, USA. *Estuaries*, **26**, 815–828.

Csardi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *Inter-Journal*, **Complex Systems**, 1695.

Dame, R.F. & Patten, B.C. (1981) Analysis of energy flows in an intertidal oyster reef. *Mar Ecol Prog Ser*, **5**, 115–124.

Dixon, P. (2003) VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, **14**, 927–930.

Dormann, C.F., Gruber, B. & Fründ, J. (2008) Introducing the bipartite package: analysing ecological networks. *R News*, **8**, 8–11.

Fath, B.D. (2004) Network analysis applied to large-scale cyber-ecosystems. *Ecol Model*, **171**, 329–337.

Fath, B.D. & Borrett, S.R. (2006) A Matlab© function for network environ analysis. *Environ Model Softw*, **21**, 375–405.

Fath, B.D. & Patten, B.C. (1999) Review of the foundations of network environ analysis. *Ecosystems*, **2**, 167–179.

Fath, B.D., Scharler, U.M., Ulanowicz, R.E. & Hannon, B. (2007) Ecological network analysis: network construction. *Ecol Model*, **208**, 49–55.

FitzJohn, R.G. (2012) Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol Evol*, **3**, 1084–1092.

Freeman, L.C. (2004) *The development of social network analysis: A study in the sociology of science*. Empirical Press Vancouver.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*, **5**, R80.

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

Handcock, M., Hunter, D., Butts, C., Goodreau, S. & Morris, M. (2008) statnet: Software tools for the representation, visualization, analysis and simulation of network data. *J Stat Softw*, **24**, 1548.

Hannon, B. (1973) The structure of ecosystems. *J Theor Biol*, **41**, 535–546.

Higashi, M. & Burns, T.P. (1991) *Theoretical studies of ecosystems: The network perspective.* Cambridge University Press, Cambridge.

Hines, D.E., Lisa, J.A., Song, B., Tobias, C.R. & Borrett, S.R. (2012) A network model shows the importance of coupled processes in the microbial N cycle in the Cape Fear River estuary. *Estuar Coast Shelf Sci*, **106**, 45–57.

Hudson, L.N., Emerson, R., Jenkins, G.B., Layer, K., Ledger, M.E., Pichler, D.E., Thompson, M.S.A., O'Gorman, E.J., Woodward, G. & Reuman, D.C. (2013) Cheddar: analysis and visualisation of ecological communities in R. *Methods Ecol Evol*, **4**, 99–104.

Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., van Veen, F.J.F. & Warren, P. H. nad Woodward, G. (2009) Review: Ecological networks–beyond food webs. *J Anim Ecol*, **78**, 253–269.

Jørgensen, S.E., Fath, B.D., Bastianoni, S., Marques, J.C., Müller, F., Nielsen, S., Patten, B.C., Tiezzi, E. & Ulanowicz, R.E. (2007) *A new ecology: Systems perspective.* Elsevier, Amsterdam.

Kazanci, C. (2007) EcoNet: A new software for ecological modeling, simulation and network analysis. *Ecol Model*, **208**, 3–8.

Kones, J.K., Soetaert, K., van Oevelen, D. & Owino, J.O. (2009) Are network indices robust indicators of food web functioning? a Monte Carlo approach. *Ecol Model*, **220**, 370–382.

Latham II, L.G. (2006) Network flow analysis algorithms. *Ecol Model*, **192**, 586–600.

Lau, M.K., Borrett, S.R. & Hines, D.E. (2013) *enaR: Tools for ecological network analysis in R.* R package version 2.5.

Leontief, W.W. (1966) *Input–Output Economics.* Oxford University Press, New York.

Lima, M. (2011) *Visual complexity: mapping patterns of information.* Princeton Architectural Press.

Margalef, R. (1963) Certain unifying principles in ecology. *Am Nat*, **97**, 357–374.

Matis, J.H. & Patten, B.C. (1981) Environ analysis of linear compartmental systems: the static, time invariant case. *Bull Int Stat Inst*, **48**, 527–565.

Metcalf, C.J.E., McMahon, S.M., Salguero-Gómez, R. & Jongejans, E. (2012) IPMpack: an R package for integral projection models. *Methods Ecol Evol*, **4**, 195–200.

Monaco, M.E. & Ulanowicz, R.E. (1997) Comparative ecosystem trophic structure of three us mid-Atlantic estuaries. *Mar Ecol Prog Ser*, **161**, 239–254.

Moody, J., McFarland, D. & Bender-deMoll, S. (2005) Dynamic network visualization. *American Journal of Sociology*, **110**, 1206–1241.

Newman, M. (2003) The structure and function of complex networks. *SIAM review*, **45**, 167–256.

Niquil, N., Chaumillon, E., Johnson, G., Bertin, X., Grami, B., David, V., Bacher, C., Asmus, H., Baird, D. & Asmus, R. (2012) The effect of physical drivers on ecosystem indices derived from ecological network analysis: Comparison across estuarine ecosystems. *Estuar Coast Shelf Sci*, **108**, 132–143.

Patten, B.C. (1959) An introduction to the cybernetics of the ecosystem: The trophic dynamic aspect. *Ecology*, **40**, 221—231.

Patten, B.C. (1978) Systems approach to the concept of environment. *Ohio J Sci*, **78**, 206–222.

Patten, B.C. (1982) Environs: Relativistic elementary particles for ecology. *Am Nat*, **119**, 179–219.

Patten, B.C. (1991) Network ecology: Indirect determination of the life–environment relationship in ecosystems. M. Higashi & T. Burns, eds., *Theoretical Studies of Ecosystems: The Network Perspective*, pp. 288–351. Cambridge University Press, New York.

Patten, B.C., Bosserman, R.W., Finn, J.T. & Cale, W.G. (1976) Propagation of cause in ecosystems. B.C. Patten, ed., *Systems Analysis and Simulation in Ecology, Vol. IV*, pp. 457–579. Academic Press, New York.

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

Revell, L.J. (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol*, **3**, 217–223.

Salas, A.K. & Borrett, S.R. (2011) Evidence for dominance of indirect effects in 50 trophic ecosystem networks. *Ecol Model*, **222**, 1192–1204.

Scharler, U. & Fath, B. (2009) Comparing network analysis methodologies for consumer–resource relations at species and ecosystems scales. *Ecol Model*, **220**, 3210–3218.

Schramski, J.R., Kazanci, C. & Tollner, E.W. (2011) Network environ theory, simulation and EcoNet© 2.0. *Environ Model Softw*, **26**, 419–428.

Soetaert, K., Van den Meersche, K. & van Oevelen, D. (2009) *limSolve: Solving Linear Inverse Models*. R package version 1.5.1.

Ulanowicz, R.E. (1983) Identifying the structure of cycling in ecosystems. *Math Biosci*, **65**, 219–237.

Ulanowicz, R.E. (1986) *Growth and Development: Ecosystems Phenomenology*. Springer–Verlag, New York.

Ulanowicz, R.E. (1997) *Ecology, the Ascendent Perspective*. Columbia University Press, New York.

Ulanowicz, R.E. (2009) *A third window, Natural life beyond Newton and Darwin*. Templeton Foundation Press, West Conshohocken, PA.

Ulanowicz, R.E. & Kay, J. (1991) A package for the analysis of ecosystem flow networks. *Environmental Software*, **6**, 131–142.

Ulanowicz, R.E. & Kemp, W.M. (1979) Toward canonical trophic aggregations. *Am Nat*, **114**, 871–883.

Ulanowicz, R.E. & Puccia, C.J. (1990) Mixed trophic impacts in ecosystems. *Coenoses*, **5**, 7–16.

Ulanowicz, R.E. & Scharler, U.M. (2008) Least-inference methods for constructing networks of trophic flows. *Ecol Model*, **210**, 278–286.

van Oevelen, D., Soetaert, K., García, R., de Stigter, H.C., Cunha, M.R., Pusceddu, A. & Danovaro, R. (2011) Canyon conditions impact carbon flows in food webs of three sections of the nazaré canyon. *Deep-Sea Res Pt II*, **58**, 2461–2476.

Wasserman, S. & Faust, K. (1994) *Social network analysis: Methods and applications*. Cambridge University Press, Cambridge; New York.

Zhang, Y., Yang, Z.F., Fath, B.D. & Li, S.S. (2010) Ecological network analysis of an urban energy metabolic system: Model development, and a case study of four Chinese cities. *Ecol Model*, **221**, 1865–1879.

6 Tables

Table 1: Primary Ecological Network Analysis algorithms in `enaR`.

Analysis	**Function Name**	**School**
Structure	`enaStructure`	foundational, Patten
Flow	`enaFlow`	foundational, Patten
Ascendency	`enaAscendency`	Ulanowicz
Storage	`enaStorage`	Patten
Utility	`enaUtility`	Patten
Mixed Trophic Impacts	`enaMTI`	Ulanowicz
Control	`enaControl`	Patten
Environ	`enaEnviron`	Patten

Table 2: Example code for applying `enaR` Flow analysis to Dame & Patten's (1981) oyster reef model.

```
> library(enaR)                   # load package
> m <- read.scor("oyster.dat")  # read model data from SCOR formatted file
> m <- balance(m)                 # balance model using AVG2 algorithm
[1] BALANCED
> u <- unpack(m)                  # unpack model data to illustrate components
> attributes(u)
$names
[1] "F"      "z"      "r"      "e"      "y"      "X"      "Living"

> F <- enaFlow(m)                 # perform ENA flow analysis
> attributes(F)                    # show analysis objects created
$names
[1] "T"  "G"  "GP" "N"  "NP" "ns"

> F$ns                            # show flow analysis network statistics
     Boundary     TST     TSTp      APL       FCI       BFI       DFI       IFI
[1,]    41.47 83.5833 125.0533 2.015512 0.1101686 0.4961517 0.1950689 0.3087794
         ID.F   ID.F.I   ID.F.O    HMG.I    HMG.O AMP.I AMP.O mode0.F  mode1.F
[1,] 1.582925 1.716607 1.534181 2.051826 1.891638     3     1   41.47 32.90504
      mode2.F  mode3.F mode4.F
[1,] 9.208256 32.90504   41.47
>
```

7 Figures

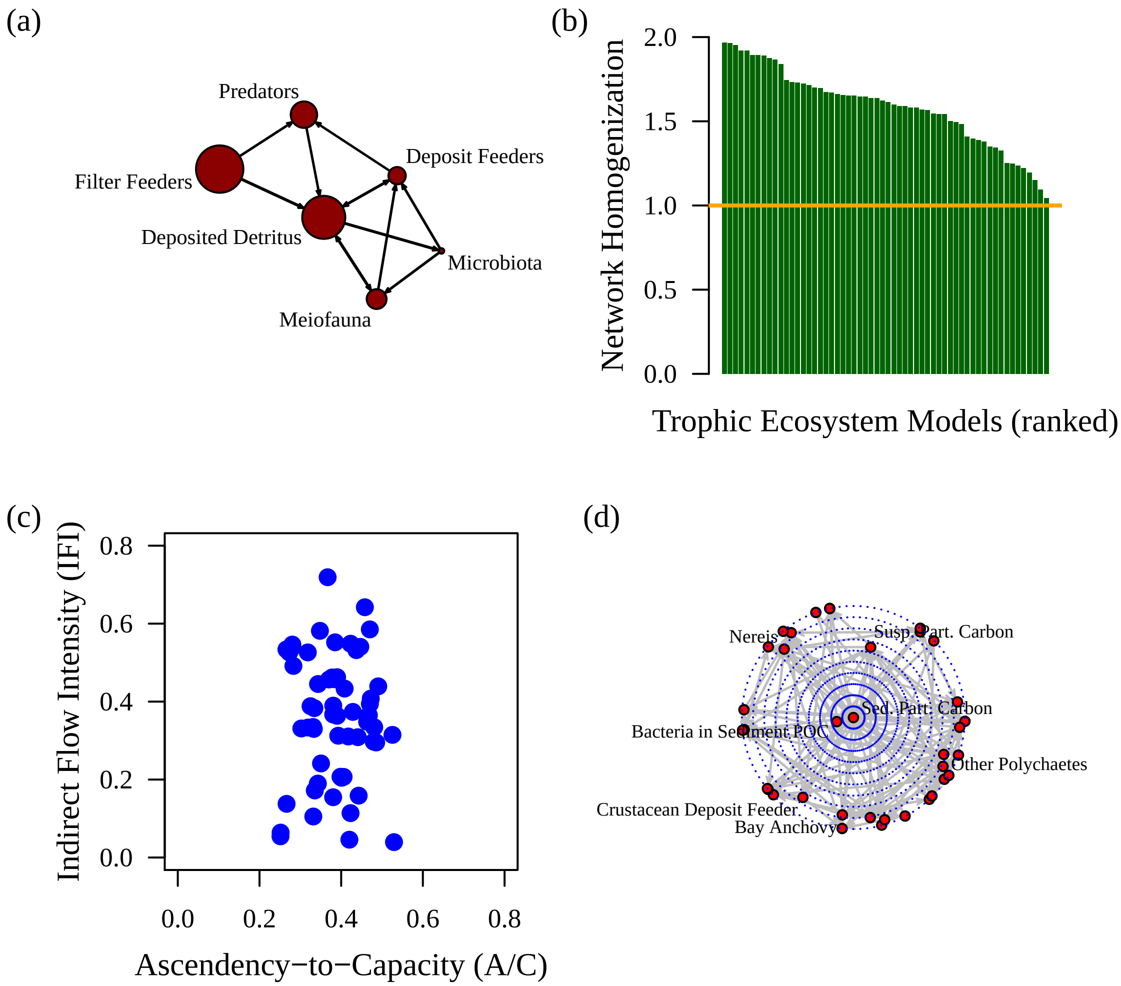
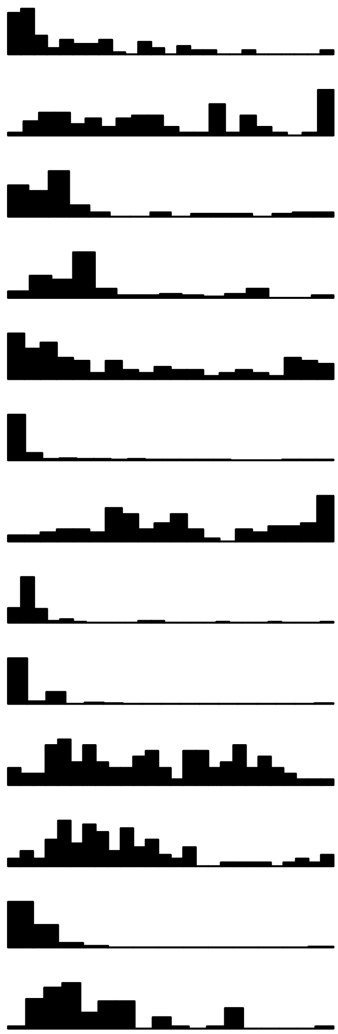


Figure 1: Example of analysis and visualizations created with `enaR` (a) network digraph of the internal flows of an oyster reef ecosystem model (Dame & Patten, 1981), (b) network homogenization statistic for 56 trophic ecosystem models (rank-ordered), (c) scatter plot showing the relationship between the ascendency-to-capacity ratio and the indirect flow index for the 56 trophic ecosystem models (Table xx), and (d) target plot of the betweenness centrality from social network analysis calculated for the xx nodes of the Chesapeake Bay ecosystem model (Baird & Ulanowicz, 1989).



Statistic	Min	Distribution	Max	Median	Mean	CV
n	4		125	15	26.66	1.02
C	0.05		0.45	0.22	0.25	0.51
LD	1		16.91	3.14	4.58	0.89
lam1A	0		14.17	3.27	4.27	0.76
FCI	0		0.98	0.26	0.38	0.86
APL	1.37		186.25	3.67	20	1.91
IFI	0.04		0.99	0.53	0.56	0.52
HMG.O	1.04		13.07	1.78	2.3	0.83
AMP.O	0		323	6.5	19.77	1.91
AMI	1		2.25	1.57	1.58	0.21
ASC.CAP	0.25		0.75	0.39	0.42	0.28
synergism.F	2.41		60.51	3.95	5.69	1.1
mutualism.F	0.6		4	1.16	1.43	0.5

Figure 2: Distributions of selected ENA network statistics from to the 100 empirically-based ecosystem models included in `enaR` 2.0. The results are summarized using a histogram showing the distribution of the values of each network statistic between the observed minimum and maximum values. The median, mean, and coefficient of variation (ratio of standard deviation and mean) values are also reported. The network statistics are the number of nodes (n), the connectance ($C = L/n^2$), link density ($LD = L/n$), pathway proliferation rate (lam1A), Finn cycling index (FCI), average path length (APL), indirect flow intensity (IFI), output oriented network homogenization ratio (HMG.O), output-oriented network amplification ratio (AMP.O), average mutual information (AMI), the ascendency-to-capacity ratio (ASC.CAP), flow-based network synergism (synergism.F) and mutualism (mutualism.F).