Network ecology

Scientists studying diverse complex systems such as social communities, protein interactions, and **economies** use network models and network analysis to investigate the system's structure, function, and evolution. *Network ecology* is the investigation of ecological systems using these **tools**. Mathematically, a network model is a graph comprised of a set of *nodes* that represent the objects or actors in the system and *edges* that represent some relationship(s) that connects the nodes. For example, in a **food web**, nodes may be species populations and directed edges (i.e., having directional relationships) show the relationship of who eats whom. Network science in general has exploded in recent years [1], but the approach has deep roots in several disciplines including ecology [2]. Today, network models have a wide use in ecology including **metapopulation analysis** [3], **landscape ecology** [4], and mutualistic interactions [5]. This article focuses on ecosystem network ecology as an exemplar of the field as a whole.

In **ecosystem** *network ecology*, ecologists construct and analyze models to understand the intricacies of material and energy flows. In these networks, nodes are compartments that contain a designated form of material or energy; the directed edges represent flows of energy–matter between these compartments and with the external environment. The term most often associated with this discipline is *ecosystem* or *ecological network analysis* (*ENA*), though *network environ analysis* is also commonly used. The method is a collection of algorithms designed to aid the understanding of the flow structure of the network and, by inference, the corresponding ecosystem. These algorithms are particularly valuable to the ecologist in categorizing compartments and groups of compartments, defining indirect relationships between compartments, and indexing system-level attributes.

This article introduces the reader to network ecology and ENA in the following ways:

Based in part on the article "Network ecology" by Robert R. Christian and Robert E. Ulanowicz, which appeared in the *Encyclopedia of Environmetrics*.

1. It familiarizes the reader with the origins of network ecology and ENA and explains the construction of ecological energy–matter flow networks.
2. It provides an overview of the ENA algorithms and identifies software to implement them.
3. It describes a selected group of algorithms in more detail with special attention to system-level attributes.
4. It also highlights some of the applications of ENA and the new ecology it is generating.

Origins

Although the use of networks to describe the flow structure of ecosystems dates to the first half of the twentieth century [6], the use of formal algorithms in what is now called ENA did not begin to develop until the 1970s [7]. At that time it became obvious that mechanical simulation modeling was limited in its capacities to represent actual ecosystem dynamics [8]. Thus, ecologists began developing alternative tools like network analysis [9]. Responding to these developments, the Scientific Committee on Ocean Research (SCOR) of the United Nations formed Working Groups 59 (Mathematical Models in Biological Oceanography) and 73 (Ecosystem Theory in Relation to Biological Oceanography). These groups urged ecologists to shift attention from stocks and biomasses toward a greater emphasis upon the measurement and analysis of ecological processes in general and upon trophic transfers in particular [10]. Out of this movement arose the systematic analysis of ecological flow networks [11] and the protocols for network ecology.

Flow Networks

Ecosystem flow networks trace thermodynamically conserved currencies through the system, capturing the ecosystem architecture and illuminating the hidden connections between ecosystem **members**. Frequently used currencies include energy, carbon, nitrogen, and water. These models are like road maps for resource distribution in ecosystems. While the models can take several forms, two common types are trophic and **biogeochemical models**.

Trophic flow networks have a food web at their core and are most common. Trophic processes are
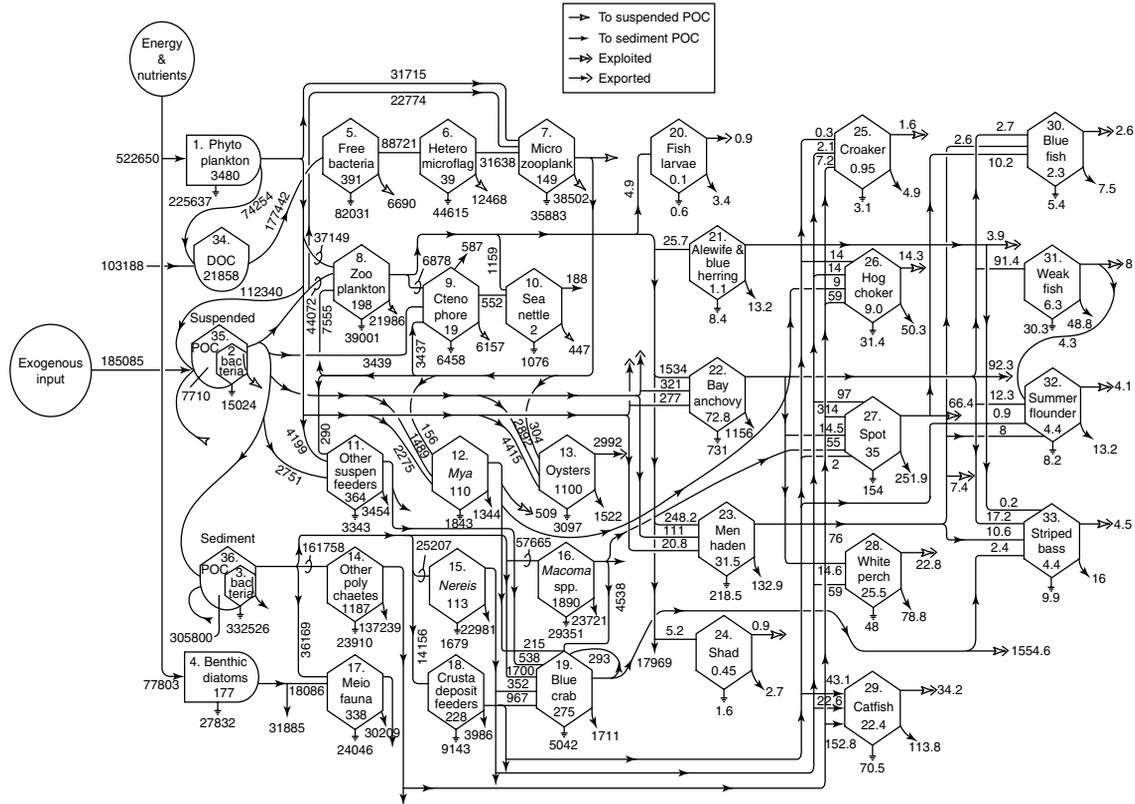
Figure 1 Schematic representation of the annual carbons flows among the 36 principal components of the Chesapeake Bay mesohaline ecosystem [12]. Carbon standing crops are indicated within the compartments in mg m^{-2} and the indicated carbon flows are in mg C m^{-2} y^{-1}. *Source*: Reproduced by permission of Ecological Society of America [12].

often the focus because consumption is an important way of transferring energy–matter from one compartment to another. Trophic flow networks focus on two questions: "Who eats whom?" and "At what rates?". An important difference between trophic networks and classic food **webs**, however, is that food webs restrict their relationship to who eats whom. Trophic ecosystem network models include all physical, chemical, and biological processes for energy–matter movement including respiration, excretion, and death. Typically these models explicitly map energy–matter movement to and from living compartments to nonliving pools of dead organic matter (detritus). Also depicted in these models are transfers of the energy–matter to or from the external world, for example, primary productions, respirations, and other exports and losses. Figure 1 is a diagram of the Chesapeake Bay trophic ecosystem network [12] that shows these flows.

Biogeochemical ecosystem network models usually trace a nutrient currency such as nitrogen or **phosphorus**. These models may have trophic components, but they also tend to have more nonliving compartments representing different states of the nutrient (e.g., nitrate and ammonia) and have more nontrophic biogeochemical transactions like nitrification [13]. When compared to trophic networks, the models tend to have more aggregated biological nodes, higher connectance, and to be functionally less dissipative, and to have more recycling [14].

Regardless of the model type, constructing these networks synthesizes existing data and knowledge and often identifies missing information. While the magnitudes of some fluxes are available for most ecosystems, direct measurements of all transfers are rare. For trophic process, one may combine the biomass densities (e.g., g C m^{-2}) of the various

components with tabulated data or allometric equations for physiological ratios, such as consumption: biomass, respiration: biomass, and assimilation efficiency. Then with knowledge of the diet composition, one may estimate the trophic flows. Several software packages can assist with model construction and balancing including MATLOD (cf. NETWRK below), WAND, WAND Balance, and Ecopath (see URLs in the next section).

Ecological Network Analysis

Given an ecosystem network model, ecologists apply ENA to investigate the organization and function of these complex systems. Algorithms for many analyses have been developed, including analyses of network structure, flow, storage, utility, environ, and information characteristics. For example, structural analysis focuses on the topology of the network. It can enumerate pathways between nodes and can identify the characteristic cycles within the system. Flow analysis provides matrices and vectors designed to quantify both direct and indirect relationships of nodes and the origin and fates of material flowing through the system. Selected analyses are described in more detail in the next section.

Software to perform ENA calculations is available. Ulanowicz developed one of the first packages called *NETWRK* (http://www.cbl.cees.edu/~ulan/ntwk/network.html). *WAND* (http://www.dsa.unipr.it/netanalysis/?Software) is a Microsoft Windows implementation of many but not all of the algorithms in NETWRK. *Ecopath*, developed by Christensen and Pauly (http://www.ecopath.org/), focuses on model construction, but implements some of the network analyses. It was intended for application to fisheries, which is reflected by some of its terminology. Fath and Borrett introduced *NEA.m*, a Matlab® function to perform a broad spectrum of ENA from the environ perspective (http://people.uncw.edu/borretts/research.html). EcoNet® (http://eco.engr.uga.edu/) is a newer online software for ENA that adds new capabilities. Each package contains somewhat different algorithms, assumptions, benefits, and limitations. To date there exists no comprehensive software that offers a unified framework.

Selected ENA Algorithms

ENA is based on the premise that analysis of network representations may aid understanding of complex, natural systems. With mathematical tools, one can glean useful insights at several levels of resolution into the functioning of the ecosystem. Some of the analyses provide insight into how components interrelate and some refer to system-level attributes. Here, we describe selected ENA algorithms to highlight a range of analyses and insight available. These are by no means the only analyses that can be employed, nor the only ways they can be used.

Structural Analysis

Structural network analysis concentrates on the unweighted topological arrangements of the network. It begins with counts of the number of network nodes and edges as a way of characterizing the size of the network and its internal connectivity. As indirect influences can be important in systems, network scientists use network analysis to count and occasionally to identify pathways between nodes that may cross several edges. For example, in the Chesapeake Bay network (Figure 1) the pathway from free bacteria (#5) → heteromicroflagellates (#6) → microzooplankton (#7) → oysters (#13) has a length of 3. Because of cycling, the number of pathways in ecosystems tends to increase geometrically as path length increases (Figure 2(a)). The rate at which pathways proliferate is a whole-system indicator of the system connectivity.

Flow Analysis

Flow analysis focuses on the weighted flows in the system and is the core of many ENA algorithms. The first step in flow analysis is to determine the total amount of energy−matter flowing into or out of each node, which is called *node throughflow*. In the Chesapeake Bay model, the croaker compartment (#25) has $0.3 + 2.1 + 7.2 = 9.6 \, \text{mg C m}^{-2} \, \text{y}^{-1}$ flowing into the node and $1.6 + 4.9 + 3.1 = 9.6 \, \text{mg C m}^{-2} \, \text{y}^{-1}$ flowing out of the node. In this case the total input and output throughflow is equal because the model is at steady state. The sum of the node throughflows is termed *total system throughflow*, which is a measure of the total system activity. If we sum all of the model inputs, internal flows, and outputs we obtain a
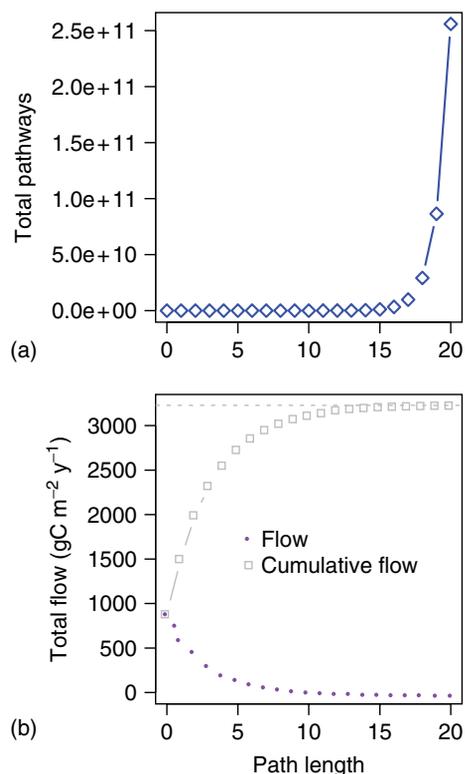
Figure 2 The (a) total number of pathways and (b) total and cumulative flow as path length increases in the Chesapeake Bay ecosystem model.

related whole-system indicator called the *total system throughput*, which is larger than total system through*flow*. Another useful whole-system indicator of the system function is the total system throughflow divided by the total system inputs. This indicates the average amount of activity that a unit input generates and is called the *average path length* or *network aggradation*.

Another ENA flow algorithm determines the *total contribution matrix*. It is used to evaluate the fraction that any compartment's throughflow contributes to any other compartment's activity. This contribution is derived from the matrix of exchanges among the compartments. The matrix coefficients represent connections between compartments that may be either direct (path length of one) or indirect (path length greater than one). Thus, no direct connection is necessary for a contribution to occur. For example, zooplankton directly eats phytoplankton in Figure 1.

However, the relationship between fish larvae and phytoplankton is indirect as it transfers through zooplankton. The total contribution matrix includes the relative contributions of both the direct connections (e.g., the fraction of production from phytoplankton to zooplankton), as well as the indirect connections (e.g., the fraction of production from phytoplankton to fish larvae). Indirect effects tend to be relatively large and develop rapidly in ecosystems, even though the total amount of flow falls off significantly as pathway length increases [14] (Figure 2(b)).

The *total dependency matrix* evaluates the fraction of a compartment's throughflow that resided at some point in another compartment. As such, this analysis is the obverse of total contribution and can be used to assess the extended diet of consumers.

Biogeochemical cycle analysis evaluates the nature of flows associated with cycles within the network [15]. A *cycle* is a series of transfers that, in combination, pass material from a compartment, through one or more other compartments, and returns material to the original one. This creates a positive feedback or autocatalytic loop. The transfer within a cycle that has the smallest flux is called the *weak arc*, and a group of cycles sharing the same weak arc is called a *nexus*. The weak arc is potentially the controlling flow within a cycle, and as such, all cycles in a nexus share a common control. In the example as shown in Figure 1, a cycle exists as the following: carbon is passed from particulate organic carbon (POC) (#35) to attached bacteria (#2) to zooplankton (#8) and back to POC. The smallest flux (weak arc) is the consumption of bacteria by zooplankton. There is also a cycle from 35 to 2 to 8 to ctenophores (#9) and back again to 35, and its smallest flow is from 2 to 8. The two cycles form a nexus controlled by their common weak arc. If all flows within the cycles were to be reduced by the amount of flow from bacteria to zooplankton, then both cycles would be broken. If zooplankton feeding were to increase, then one might also infer that both cycles would have greater flow.

The total flow associated with these cycles is termed *recycled flow*. The *Finn Cycling Index* is the fraction of total system throughput that is recycled [16]. This is a whole-system indicator of the functional connectivity of the system components.

Environ Analysis

Environ analysis builds on the core flow analysis, but it further evaluates the origin and fate of

energy–matter exiting or entering the system [17]. It generates a partition set of node-specific subnetworks. Input environs reveal the origins of material exiting one of the compartments, while the output analysis shows the fate of input to a node. This analysis can be used, for example, to determine the fate of primary production from either benthic diatoms or phytoplankton as it passes through and out of the ecosystem.

Information Analysis

Information analysis consists of a set of whole-system indicators that describe the organization and development of ecological systems [18]. These indicators are based on information theory and reflect the agencies potentially responsible for changes in network structures [19]. They stem from the role of chance and historical contingency in effecting change. Ulanowicz [19] argues that *autocatalytic* configurations (i.e., positive feedback loops or cycles) affected by chance disturbances can generate a nonmechanical system response. This response is reflected in whole-system attributes including selection, growth enhancement, symmetry breaking, centripetality of resources, inducement of competition between autocatalytic clusters, and partial autonomy.

Three indicators are central to this analysis: capacity, ascendency, and overhead. Each is a combination of an extensive (size dependent) and intensive (size independent) element. Total system through*put* (TSTp) is the extensive scalar in all three indicators. The capacity of the system is a measure of the diversity of energy–matter movement through the system, which is indicated by the familiar Shannon–Wiener index [20], H, applied to the flows in the network. *Capacity* (C) is then $H \times$ TSTp. In ecological systems, only a fraction of this capacity is achieved because ecological processes are not randomly construed; the flow structure is more ordered. *Average mutual information* (AMI) is the amount by which the constraints encumber the potential complexity, H, and *Ascendency* (A) is then AMI scaled by TSTp, such that $C \geq A \geq 0$. The remaining potential complexity, $C - A$ (>0), is called *system overhead*, ϕ. In contrast to the ascendency, which measures the amount of complexity that is expressed as constrained flow structure, the overhead assigns

a number to the residual flexibility that the system retains.

Ecologists have used these information analyses in several ways. Their origination was to characterize the maturity of ecosystems in terms of development [18]. They have also been used to try to quantify the concepts of ecosystem health and integrity [19]. These measures are built into Ecopath and are frequently used to characterize the organization of ecosystems associated with fisheries.

Limitations

While ENA is a powerful approach for studying complex ecosystems, it has limitations. An important constraint is that network models are data intensive. Each compartment requires an estimate of standing stock, and estimates of the magnitude of energy–matter flux between the compartments and between the compartments and the system environment. There is both a qualitative concern that each connection be adequately identified and a quantitative concern that the values for standing stocks and flows are appropriate. Direct measurement of each stock, flow or parameter is rare, and the consequences of using information from such sources are rarely evaluated. Furthermore, little evaluation is usually made of the ramifications of **stochasticity** or **uncertainty** upon the network analysis. A second limitation is that many analyses assume the model is at steady-state and thus do not capture the effects of system dynamics. Further, most of the analyses are based on **linear algebra** and, therefore, are not prone to the complexities associated with nonlinearities. Third, the analyses typically only trace a single conserved currency at a time, which may not reflect all aspects of energy–matter transactions and derived relationships. Multiple models might be needed to build a more comprehensive understanding. Finally, the analysis only captures the relationships mediated by or captured in energy–matter fluxes. Many important indirect, behavioral, and semiotic effects will not be captured in this analysis.

Significance

Applications of network ecology are revealing the hidden relationships among components of many types of ecological systems. For example, Bondavalli

and Ulanowicz [21] showed that although American alligators are direct predators of frogs in the Everglades swamp, their net relationship is mutualistic because the alligators provide a net benefit by eating snakes, which are major predators on the frogs. Baird and Ulanowicz [12] found that the Chesapeake Bay ecosystem was divided into two separate subsystems – the pelagic and the detrital. The inclusion of ENA methods in Ecopath has insured their continued contribution to fisheries ecology. Interestingly, new applications of ENA are being used to evaluate the **sustainability** of **water** use in such places as Sarmato, Italy [22] and the energy metabolism of cities in China [23]. ENA is helping scientists perceive and manage specific systems, and it is building a new theoretical understanding of ecological systems in general [18, 24].

Possibly ENA's most important contribution to science has been the required change in philosophical perspective. In this age of complexity, there is no more salient metaphor than the network, which captures the entanglement of chance and necessity essential for the growth, development, and persistence of ecological systems. In short, ENA is in the process of catapulting ecology from deep among the pack of sciences to the very forefront of our understanding of the natural world.

References

[1] Newman, M.E.J., Barabasi, A.-L., & Watts, D.J. (2006). *The Structure and Dynamics of Networks*, Princeton University Press, Princeton.

[2] Proulx, S.R., Promislow, D.E.L., & Phillips, P.C. (2005). Network thinking in ecology and evolution, *Trends in Ecology and Evolution* **20**, 345–353.

[3] Hanski, I. & Ovaskainen, O. (2000). The metapopulation capacity of a fragmented landscape, *Nature* **404**, 755–758.

[4] Urban, D. & Keitt, T. (2001). Landscape connectivity: a graph–theoretic perspective, *Ecology* **82**, 1205–1218.

[5] Bascompte, J. & Jordano, P. (2007). Plant-animal mutualistic networks: the architecture of biodiversity, *Annual Reviews of Ecology and Systematics* **38**, 567–593.

[6] Golley, F.B. (1993). *A History of the Ecosystem Concept in Ecology*, Yale University Press, New Haven.

[7] Hannon, B. (1973). The structure of ecosystems, *Journal of Theoretical Biology* **41**, 535–546.

[8] Mann, K.H., Platt, T.C., & Ulanowicz, R.E., eds, (1981). *Mathematical Models in Biological Oceanography*, Vol. 7, UNESCO Monographs on Oceanographic Methodology UNESCO Press, Paris.

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

[10] Fasham, M.J.R., ed. (1984) *Flows of Energy and Materials in Marine Ecosystems: Theory and Practice*, Plenum Press, New York.

[11] Wulff, F., Field, J.G., & Mann, K.H., eds (1989). *Network Analysis in Marine Ecology*, Coastal and Estuarine Studies Series, Springer-Verlag, Berlin.

[12] Baird, D. & Ulanowicz, R.E. (1989). The seasonal dynamics of the Chesapeake Bay ecosystem, *Ecological Monographs* **59**, 329–364.

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

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

[15] Ulanowicz, R.E. (1983). Identifying the structure of cycling in ecosystems, *Mathematical Biosciences* **65**, 219–237.

[16] Finn, J.T. (1980). Flow analysis of models of the Hubbard Brook ecosystem, *Ecology* **61**, 562–571.

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

[18] Ulanowicz, R.E. (1986). *Growth & Development: Ecosystems Phenomenology*, Springer-Verlag, New York.

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

[20] Shannon, C.E. (1948). A mathematical theory of communication, *Bell System Technical Journal* **27**, 379–423.

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

[22] Bodini, A. & Bondavalli, C. (2002). Towards a sustainable use of water resources: a whole-ecosystem approach using network analysis, *International Journal of Environment and Pollution* **18**, 463–485.

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

[24] Jorgensen, S.E., Fath, B.D., Bastianoni, S., Marques, J.C., Muller, F., Neilsen, S.N., Patten, B.C., Tiezzi, E., & Ulanowicz, R.E. (2007). *A New Ecology: Systems Perspective*, Elsevier, Amsterdam.

Stuart R. Borrett, Robert R. Christian & Robert E. Ulanowicz