Ecological network analysis: network construction

Brian D. Fath\textsuperscript{a,}\textsuperscript{*}, Ursula M. Scharler\textsuperscript{b,c}, Robert E. Ulanowicz\textsuperscript{d}, Bruce Hannon\textsuperscript{e}

\textsuperscript{a} Biology Department, Towson University, Towson, MD 21252, USA
\textsuperscript{b} School of Biological and Conservation Sciences, University of KwaZulu-Natal, Howard College Campus, George Campbell Building, 4041 Durban, South Africa
\textsuperscript{c} Department of Aquatic Ecology and Water Quality Management (AEW), Wageningen University and Research Centre, Ritzema Bosweg 32-A, 6703 AZ Wageningen, The Netherlands
\textsuperscript{d} Chesapeake Biological Laboratory, University of Maryland, Solomons, MD 20688, USA
\textsuperscript{e} Department of Geography, University of Illinois, Urbana, IL 61801, USA

\textbf{A R T I C L E  I N F O}

Article history:
Accepted 27 April 2007
Published on line 18 June 2007

Keywords:
Ascendency
Ecological network analysis
Food webs
Systems analysis

\textbf{A B S T R A C T}

Ecological network analysis (ENA) is a systems-oriented methodology to analyze within system interactions used to identify holistic properties that are otherwise not evident from the direct observations. Like any analysis technique, the accuracy of the results is as good as the data available, but the additional challenge is that the data need to characterize an entire ecosystem’s flows and storages. Thus, data requirements are substantial. As a result, there have, in fact, not been a significant number of network models constructed and development of the network analysis methodology has progressed largely within the purview of a few established models. In this paper, we outline the steps for one approach to construct network models. Lastly, we also provide a brief overview of the algorithmic methods used to construct food web typologies when empirical data are not available. It is our aim that such an effort aids other researchers to consider the construction of such models as well as encourages further refinement of this procedure.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Ecological network analysis (ENA) is a methodology to holistically analyze environmental interactions (see e.g., Hannon, 1973, 1985a,b, 1986, 1991, 2001; Hannon et al., 1986, 1991; Hannon and Joiris, 1989; Finn, 1976; Patten, 1978, 1981, 1982, 1985; Higashi and Patten, 1989; Fath and Patten, 1999; Ulanowicz, 1980, 1983, 1986, 1997, 2004; Ulanowicz and Kemp, 1979). As such, it is necessary that the network model be a partition of the environment being studied, i.e., be mutually exclusive and exhaustive. The latter criterion in particular is difficult to realize and most models such as Lotka–Volterra predator–prey or competition models represent only a small subset of the interactions occurring in the ecosystem, excluding both the majority of other species in the community and all abiotic processes. As a result of this limited perspective, it is impossible for such approaches to quantify the wholeness and consequent indirectness in the system, but this has been the trend of reductionist science for over a century. The reductionistic approach results in a self-fulfilling realization in that only the few species or processes in the model have influence and significance in the final interpretation, without considering the embedded nature of these activities within the larger ecological context. Ecosystems comprise a rich web of many interactions and it would be remiss to exclude, a priori, most of them or to rely on analysis techniques that do so. ENA, on the other hand, is capable of analyzing the structural and functional properties of this web of interactions without reducing the model to its presumed minimal constituents. Therefore, network models aim to include all ecological com-

\textsuperscript{*} Corresponding author. Tel.: +1 410 704 2535; fax: +1 410 704 2405.
E-mail address: bfath@towson.edu (B.D. Fath).
0304-3800/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2007.04.029
portions and interactions and the analysis determines the overall relationships and significance of each. The difficulty of course lies in obtaining the data necessary to quantify all the ecological compartments and interactions. When sufficient data sets are not available, simple algorithms, called community assembly rules, have been employed to construct realistic food webs to test various food web theories. Once the network is constructed, via data or algorithms, the ENA is quite straightforward and software is available to assist in this (Allesina and Bondavalli, 2004; Fath and Borrett, 2006). This paper outlines a possible scenario for developing network models.

2. Data requirements and acquisition for developing network models

A network flow model is essentially an ecological food web (energy–matter flow of who eats whom), which also includes non-feeding pathways such as dissipative export out of the system and pathways to detritus. The first step is to identify the system of interest and place a boundary (real or conceptual) around it. Energy–matter transfers within the system boundary comprise the network; transfers crossing the boundary are either input or output to the network, and all transactions starting and ending outside the boundary without crossing it are external to the system and are not considered. Once the system boundary has been established, it is necessary to compartmentalize the system into the major groupings. The most aggregated model would have three compartments: producers, decomposers/detritus, and consumers; a slightly more disaggregated model could have producers, herbivores, carnivores, omnivores, decomposers, and detritivores (Fath, 2004); and the most disaggregated a different compartment for each species. Most models use some aggregation based on the functional groups of the ecosystem such that network models in the literature typically have between 6 and 60 compartments. However, this does not completely resolve the aggregation issue. It is likely that one is interested in greater detail for one group, but it is not entirely clear how disaggregation of one functional group and not others affects the analysis results. Identifying the major species or functional groups should be done by those knowledgeable about the system.

Once the compartments have been chosen, an energy–matter flow currency must be selected. Typically, the currency is biomass (e.g., grams of carbon) or energy (e.g., kilojoules) per area for terrestrial and aquatic ecosystems or volume for aquatic ecosystems per time. The flow dimensions then would be ML−2T−1 or ML−3T−1 where M = mass, L = length, and T = time. There is flexibility however in the biomass units chosen, which could also be grams of nitrogen, phosphorus, other nutrients, or even water per space dimension per time. Multiple currency network models using a combination of C, N, or P, etc. can also be constructed (Ulanowicz and Baird, 1999). In addition to the input, output, and within systems flow transfer values, it is also necessary to measure empirically as best as possible the mass density (biomass/area) of each compartment. Storage dimensions are ML−2 or ML−3, since they are not rates. Together the transfers and storages comprise the data requirements for ecological network analysis.

Once the currency has been chosen, we would arrange them in the columns and rows of an adjacency matrix to determine whether or not a resource flow of that currency occurs from each compartment to each other one. An adjacency matrix, is a representation of the graph structure such that \( a_{ij} = 1 \) if there is a flow from \( j \) to \( i \), else \( a_{ij} = 0 \), using a column to rows orientation (note that although we use a column to row orientation here, a row to column orientation is also used in the literature). This procedure forces one to ascertain the possible connectivity of each pair of compartments in the network, thus reducing the chances of over-looking certain connections. This exercise might also illuminate compartments that were excluded initially, thereby providing an iterative feedback in the network development.

The data required for ecological network analysis are as follows: For each compartment in the network, the biomass and physiological parameters, such as consumption (C), production (P), respiration (R) and egestion (E) must be quantified. It is possible to lump respiration and egestion into one outflow if necessary. Furthermore, the diet of each compartment must be apportioned amongst the inputs from other compartments (consumption) in the network. This apportionment of “who eats whom and by how much” can be depicted in a dietary matrix, where material flows from compartment \( j \) to compartment \( i \). For all compartments, inputs should balance outputs (C = P + R + E), in accordance with the conservation of matter and the laws of thermodynamics.

To quantify the network, flows of the chosen currency into and out of each compartment should be determined. Some of the flows could also be empirically gathered from primary field research regarding primary production, respiration, and feeding, but others could be assembled from various sources such as literature sources and simulation model results. Furthermore, two recently developed methods of assigning a flow value between compartments can be employed to estimate transfers (Ulanowicz and Scharler, in preparation). The first method, MATBLD, assigns the transfers according to the joint proportion of predator demand and prey availability. The second method, MATLOD, begins with assigning a very small flow to all designated links and keeps on doing so until either the demand is met or the source exhausted. The input data for both methods are the biomass, consumption, production, respiration, egestion, imports and exports of all compartments, and the topology of the networks (i.e., who eats whom). The networks originating from both methods are balanced using the algorithm developed by Allesina and Bondavalli (2003). A comparison of the two methods to networks constructed “by hand” revealed no statistical difference between the magnitudes of the compartmental transfers.

In most cases, field data, literature sources, or results from simulation models do not supply all the system-specific data necessary for the network construction. In those cases, it is recommended to perform a sensitivity analysis to assess a variation of the most inaccurate input data on network analysis results. Table 1 provides a step-by-step procedure for constructing ecological networks.
Table 1 – Check list for constructing ecological networks

1. Identify and demarcate ecosystem of interest
2. Make a list of the major species or functional groups in ecosystem
3. Select a unit of currency for the network
4. Construct the adjacency matrix to determine any possible flow interactions. Use this procedure to identify any possible holes in the initial classification
5. Empirically measure mass density of each compartment
6. Empirically measure input, output, and throughflows between compartments when possible
7. Use additional literature and models to quantify network flows not empirically determined
8. Employ flow-balancing algorithm to finalize flow matrix and storage, input, and outputs vectors
9. Apply ecological network analysis to network
10. Sensitivity analysis

3. Belize example

The following example illustrates the steps involved in network construction. The study ecosystem is a mangrove island called Twin Cays, situated in the Caribbean Sea, offshore Belize. The island features a tree height gradient from the fringe (tall trees) to the interior of the island (small trees). The tree height gradient has been attributed to nutrient limitations fringe (tall trees) to the interior of the island (small trees). The study involved a NSF Biocomplexity project, comprising 10 research groups. The project was part of a NSF Biocomplexity project, comprising 10 research groups. The study involved a NSF Biocomplexity project, comprising 10 research groups. The project was part of a NSF Biocomplexity project, comprising 10 research groups.

1. The ecosystems are the three zones on the island (fringe, transition, dwarf). The boundaries of the ecosystems are demarcated by the tree height gradient. The boundary to the sea is the mangrove prop roots hanging into the sea. Interior ponds marked the interior boundary of the dwarf zone.
2. The three zones have a different number of compartments, since not all species occur within all zones. The pattern of aggregation was the same over all three zones and flow currencies, except for the inorganic dissolved nutrients. The number of compartments ranges between 70 and 90. Some compartments comprise a single species only (e.g., mangrove trees), whereas others are taxonomic groups (e.g., lichen, bacteria, fungi, various macrozoobenthos groups) or functional groups (e.g., bird feeding guilds). Species, for which no information on density was available, were omitted from the networks. In most cases, these were species either living outside the ecosystem boundary and deriving part of their diet from within the ecosystem (e.g., fish), or they live within the ecosystem boundary and derive all their diet from outside the ecosystem boundary as an import (e.g., some of the submerged prop root fauna).
3. In all three zones, networks were constructed for three different flow currencies (carbon, nitrogen, phosphorus). The biomasses were expressed as gC/N/P m\(^{-2}\) and all flows (consumption, production, respiration, egestion, imports, exports, and intercompartmental transfers) as gC/N/P m\(^{-2}\) y\(^{-1}\). Flows were often inferred from metabolic ratios, such as P/B, P/R, or assimilation efficiencies. To assign certain respiration rates or assimilation efficiencies, criteria such as organism size, feeding guild, or taxonomy were used in the selection. In some cases (e.g., mangrove trees), models available in the literature helped to derive carbon balances.
4. Several of the flows in and out of the compartments were measured on site, e.g., for the tree species, macroalgalae, and microbial mats.
5. To derive non-measured flows going in and out of the compartments, literature sources were used. Data of similar mangrove systems (habitat wise and geographically) were preferred to others. Flows were often inferred from metabolic ratios, such as P/B, P/R, or assimilation efficiencies. To assign certain respiration rates or assimilation efficiencies, criteria such as organism size, feeding guild, or taxonomy were used in the selection. In some cases (e.g., mangrove trees), models available in the literature helped to derive carbon balances.
6. Several of the flows in and out of the compartments were measured on site, e.g., for the tree species, macroalgalae, and microbial mats.
7. After the biomasses, consumption, production, respiration, egestion, imports, exports, and the flow topology were determined, the actual transfers between compartments were estimated using the MATLAB procedures.

Several placeholder networks were constructed throughout the duration of the project. The networks were continuously updated as more system-specific data became available from within the project research group.

4. Application to existing models

In situations in which the model has already been developed, for example, in STELLA or other software, the network construction follows directly from the simulation results. For example, Spieles and Mitsch (2003) have provided a complete model of the macroinvertebrate trophic structure for the Olenyong River Wetland Research Park in Columbus, Ohio. The model consists of nine compartments: Metaphyton, Macrophytes, Periphyton, Collectors, Shredders, Scrapers, Predators, Coarse Detritus and Fine Detritus (Fig. 1). The model, therefore,
contains all the flow and storage connections needed for the network analysis matrices. These values can be taken directly from the simulation results, which in STELLA are available in tabular form ready for input to Excel. This particular model has a seasonal forcing function, and reaches a dynamic steady state by the end of the first year. Once the steady state is reached, one can use the average of any 1-year period as representative of the model values to create the flow, storage, input and output network data. For this model, the following data are obtained:

\[
F = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.028 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.010 & 0 & 0 & 0 & 0 & 0 & 0.005 & 0 \\
0 & 0 & 0.005 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2.34E-7 & 5.54E-7 & 8.97E-10 & 0 & 0 & 0 \\
0 & 219.99 & 0 & 0.007 & 0.002 & 0.003 & 4.94E-6 & 0 & 0 \\
34.97 & 0 & 9.995 & 0 & 0 & 0 & 0 & 220.00 & 0
\end{bmatrix}
\]

\[
x = \begin{bmatrix}
349.6751 & 2199.896 & 99.95435 & 0.0732 & 0.01718 & 0.0304 & 4.12E-5 & 366.6695 & 533.2744
\end{bmatrix}
\]

\[
z = \begin{bmatrix}
34.9955 & 220 & 10 & 0 & 0 & 0 & 2.47E-5 & 0 & 55
\end{bmatrix}^T
\]

\[
y = \begin{bmatrix}
0 & 0 & 0 & 0.0490 & 0.0137 & 0.00152 & 2.06E-5 & 0 & 319.9647
\end{bmatrix}
\]

In this manner, any simulation model has the necessary information for network construction. The obstacle is that many simulation models only represent a portion of the overall ecosystem network, such as a predator–prey relation, and therefore are mostly unsuitable for a full network analysis which is most revealing when the entire ecological network pathways are captured.
5. Additional methods

Other approaches have been used to construct flow networks from ecological data. In particular, the efforts by (1992; Pauly et al., 2000) in developing Ecopath have gained wide usage in fisheries. Activity is centered at the University of British Columbia’s Fishery Centre, but Ecopath has more than 2000 registered users in over 120 countries (see www.ecopath.org). The following description of Ecopath is based on information found at their Web site. Ecopath is publicly available software to construct and analyze mass-balanced flow networks. The ecosystem interactions in the networks represent trophic links at the species or functional level. It allows users to input known data regarding their ecosystem such as total mortality estimates, consumption estimates, diet compositions, and fishery catchers, however, data requirements are kept to a minimum because databases and balancing procedures are used to fill in missing aspects. The model employs two main equations, one regarding production (production = catch + predation + net migration + biomass accumulation + other mortality) and the other consumption (consumption = production + respiration + unassimilated food). In addition to examining the models using network analysis, Ecopath highlights the benefits from the model construction process in bringing together scientists, government managers, and public interest groups in collaborative, interdisciplinary team building exercise. To date, many different network models have been constructed using this approach (e.g., Okey and Pauly, 1999; Pauly et al., 1998; Haggan and Beattie, 1999). Recently, a dynamic time simulation module, Ecosim, has been added, which further benefits applications for environmental management. Ecopath provides a good platform for introducing network flow models to new users by removing many of the technical barriers.

Another approach for constructing ecological flow networks has been inverse modelling (Vezina and Platt, 1988). In this approach, a structural food web is constructed for the ecosystem based on available feeding and flow patterns. The compartmental biomasses are assumed to be at steady state such that the total flows entering a compartment are equal to the total amount exiting it. The flows are then back calculated for this particular structure using mass balance equations and basic biological constraints. Recent work has looked at the limitations of the steady-state assumption (Vézina and Pahlow, 2003) on inverse modelling as well as modifications to relax it (Richardson et al., 2003). This approach, although not as well known as Ecopath, has been employed in many instances. A recent citation index search revealed 41 citations for the original Vezina and Platt (1988) article, including many recent citations (e.g., Breed et al., 2004; Leguerrier et al., 2004; Richardson et al., 2004; Savenkoff et al., 2004).

6. Community assembly rules

One approach that has been used to account for the lack of empirically derived data is the development of simple algorithms to construct hypothetical, but ecologically realistic networks. However, there have been two distinct approaches marked by the initial assumptions one makes. The first group, based on population/community ecology, focuses strictly on “who eats whom”, producing structures involving primary producers, grazers, and predators, but explicitly lacks decomposers and detritus. As a result, these networks do not typically contain cycling. The two main algorithms in this category are the cascade model (Cohen and Newman, 1985) and niche model (Williams and Martinez, 2000). The second group, based on ecosystem ecology focuses on energy flow in the system and includes all functional groups including detritus and decomposers. The algorithms here are a modified niche model (Halnes et al., 2007), a cyber-ecosystem model (Fath, 2004), and structured food webs of realistic trophic relationships with transfer coefficients drawn from uniform and lognormal distributions (Morris et al., 2005). In the cyber-ecosystem approach, Fath (2004) used a meta-structure with six classifications – (1) primary producers, (2) herbivores, (3) omnivores, (4) carnivores, (5) detrital feeders, and (6) detritus – and linked each class based on known ecological relationships, such that grazers can feed only on primary producers, omnivores can feed on primary producers, carnivores, detrital feeders, and other omnivores, etc.; all compartments deposit into detritus; and detrital feeders feed on detritus. Resulting in the following generalized adjacency matrix:

\[
A = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

The matrix size is expandable, since the algorithm allows for the selection of any number of each of the six classifications while retaining the generalized connectance pattern. The number and placement of connections between individual nodes in targeted classifications (such as from primary producers to grazers) are randomly assigned. Note also that in the constructed matrices cannibalism (a_ii = 1) was not allowed although it is shown in the generalized structure. This classification as currently defined guarantees that cycling will occur since energy must pass out of each compartment and into detritus, some of which is taken up by detrital feeders. The spectral radius of Eq. (1), as given by the maximum eigenvalue of A, \( \lambda_{\text{max}} = 2.62 \), whereas without the diagonal \( \lambda_{\text{max}} = 1.94 \). This value gives an indication of the cyclic pathways inherent in the structure.

The cyber-ecosystem algorithm was used initially to create large-scale networks with 600 compartments (100 of each classification) to test the values of certain network properties. A structural measure of the network is connectance, the ratio of number of links (L) to total possible number of connections (n^2). The average connectance for the 600 compartment networks was around 0.18. Fath and Killian (2007) have looked at other combinations with different number of nodes within each classification such as at a classic ecological numbers pyramid in which there are the largest amount of primary producers and subsequently less at each higher level and inverse pyramids as well as a uniform distribution.

The inclusion of six classifications although attempting to represent known ecological categories may be more compli-
Ecological network analysis is an important tool to understand whole-system interactions, and the lack of quantified network models, and the difficulty in constructing them is one of the main impediments to further application of this methodology. There is no one correct way to construct a network model, but here we try to offer some assistance for doing so, which hopefully will increase the number of networks that are developed. Having network construction guidelines will provide some consistency in both the procedure and product making it easier to compare and contrast various ecological networks. This is only a first step and we expect that the more experience that is gained with constructing networks the more refined the procedure will become.

7. Conclusions

Ecological network analysis is an important tool to understand whole-system interactions, and the lack of quantified network models, and the difficulty in constructing them is one of the


