Department of Environmental Sciences
University of Parma

SINGLE SPECIES FEATURES IN ECOLOGICAL NETWORKS

Author:
Marco Scotti

Thesis advisor:
Dr. Antonio Bodini

Submitted in fulfillment of the requirements
for the Degree of Doctor of Philosophy
January 2005 - December 2007

- Ecology -

Signed:

Author: Marco Scotti

Thesis advisor: Dr. Antonio Bodini

Thesis co-advisor: Dr. Cristina Bondavalli

PhD program director: Prof. Giulio De Leo

Thesis referees: Dr. Luis Gerardo Abarca-Arenas
Dr. Stefano Allesina
Dr. Maria Chiara Naldi

Acknowledgements

During the last three years I worked hard on my PhD program. Beside academic training and studying, I spent a lot of time travelling, meeting, knowing, and living with wonderful people.

First of all I would like to thank my research group at University of Parma. Antonio Bodini, Cristina Bondavalli, Stefano Allesina and Nicola Clerici. They are friends before than colleagues. Each one of us is totally different from the others, but this heterogeneity foster our friendship than creating distances.

Thanks to my referees, Luis Gerardo Abarca-Arenas, Stefano Allesina and Maria Chiara Naldi that helped me with comments and reviews. I am also grateful to Franco Sartore, Stefano Leonardi and Davide Menna for technical advices on computer science.

I can not forget all the people of the Department of Environmental Sciences in Parma: Silvia, Pio (with Patty and Anna), Andrea Mizio, Marta Marziana, Stefano and Maria Chiara, Valeria, Valentina, Joao, Danielino, Rossano, Enrica, Giulio, Matteo Parelli, Daniele Beva, Il Luca, Giulia, Benji, Max Pizzol, Sara Siussi, Ste Favi, Juliano, Stefanone, Matteo Manfre and Francis. Although is a supporter of the wrong football team, I would like to thank Alessandro Zaccagnini that helped me when I was a young student.

PhD also meant an exciting period abroad in Budapest. There I met Ferenc Jordán and worked with him at the Collegium Budapest. He is a "hub" of my scientific training and a dear friend with a wonderful family. I would like to thank the Rector of the Collegium Budapest, Prof. Imre Kondor that invited me as a junior fellow, Prof. Eörs Szathmáry and Prof. János Podani, the co-author of the paper I wrote with Ferenc.

In Budapest I lived at Eötvös József Collegium and my neighbours was Marie-Laetitia and Bertrand. With them I felt at home. In a perfect random order I also cite other people linked to my Magyar experience: Benjamin, Karolin, George, Katalin, Csilla, Kati Csepi, Noémi, Vera, Zsófia and Szilvia.

I can not forget "the Spolverini", my family in Parma: Marco Bego (with Chiara and Cristina too), Lorenzo and Patty with all the other friends: Lelo, Roby, Emiliano, Paolone, Emi Donati and Ale Bonatti. Great part of myself are Ale Ravera and Monica: they always tolerate my ups and downs and I can always trust them.

No one knows me better than my old and historical friends: Cecco, Dino, Lele and Ale, Marco and Miguel, Pello and Cristina, Pisco and Ale, Bargi, Paolino, Pletti, Sera, Buddy, Cocco and Manto. I can not forget the basketball, an important part of my life: Mario, Luciano, Teo, Cipe, Faro and Marco.

Thanks to my family, Andrea with Carlo and Ermanna, Tina and Beppe, Gino and Alice, Mario and Maria. If it was not for them, I would not be here writing this thesis. Thanks to Giulia that with her smiles and energy helped me flying and dreaming.

Parma, Italy,
December 2007,
Marco Scotti

Contents

1

Introduction

1.1 The long history of food webs

Within ecosystems, species interact in various ways (e.g. predator-prey, plant-seed disperser, host-parasite, plant-pollinator, plant-ant). Types and strengths of interaction change through time and space, varying between individuals that are subject to the rules of natural selection and genetic drift. In community ecology, food relations are certainly the most investigated ones. Dealing with the whole bunch of trophic interactions is possible by graph theory. So ecosystems become set of nodes (species or trophospecies) connected by a suite of edges (trophic relations). From this idea, two types of tools have come about: food webs and ecological networks. The former list feeding relations among species in a qualitative way (presence/absence); the latter include also the magnitude of the interactions, in terms of amount of matter (or energy) that is exchanged in a given time period.

Nowadays, food webs are widely applied but the idea of mapping trophic interactions with graphs is quite old. It traces back to the pioneering research of Camerano (1880), and its revisitation as the food chain description (Elton, 1927). The linear trophic chain depicted biological communities through feeding habits, according to the perceived flow of energy that was identified as a main path from producers to various consumers. This led to see the energy passage as composed by discrete steps called trophic levels. Although appealing, discrete trophic levels failed to aptly describe multitrophic and complex interactions. Then, the focus shifted from linear trophic chains to food webs, studying relations between complexity of flow structure and ecosystem stability (MacArthur, 1955; Hutchinson and MacArthur, 1959).

Until 1970s, ecologists accepted the generalization based on arguments by Elton (1958), stating that "*complex ecosystems are more stable*". They thought that richer networks of interactions, with many species, were better able to withstand disturbance, either human or natural created, but May (1972, 1973) demonstrated the shortcomings of such general proposition. However, real ecosystems, being the product of evolutionary processes, are clearly different from randomly assembled networks (Erdös and Rényi, 1960) and further studies on their regularities (Cohen and Newman, 1985; Pimm, 1980, 1982; Pimm and Lawton, 1978, 1980) were stimulated by the publication of the first collection of food webs (Cohen, 1978).

Beside emerging patterns revealed in the 1980s, increasing sample efforts provided better data that partially contradicted theoretical advances (Cohen and Newman, 1991; Cohen et al., 1990, 1993a,b; Pimm et al., 1991). The focus of food web theory switched from arithmetic properties such as number of connections, number of trophic levels and omnivory (Pimm, 1982; Cohen et al., 1990; Havens, 1992) to structural indices. This shift was promoted by advances in statistical physics and to the application of network theories to Internet, motorways, telephonic connections, social communities, sexual contacts and power grids (Watts and Strogatz, 1998; Liljeros et al., 2001; Girvan and Newman, 2002).

Ecologists have revisited complexity/stability question by exploring the dynamics of interactions, examining the properties of more realistic food web configurations and adding hierarchical structure to food webs (McCann et al., 1998). Still, a lot of interest towards structure and mechanisms of functioning is revealed by current works and publications (May, 2006; Pascual and Dunne, 2006; Allesina and Bodini, 2005; Garlaschelli et al., 2003).

In particular, many works on food webs and ecological networks focussed on models characterizing their structure. The basic statistic used to compute the structure of a large network is its degree distribution $p(i)$, the probability that a randomly chosen node is linked to i other nodes. The binomial distribution of Erdös-Rényi random graph is the oldest and best studied network model, with S nodes connected by randomly placed links between pairs of them (Newman, 2003). As asserted by May (1973), real food webs are not randomly assembled, but an active and still controversial area of theoretical ecology is dealing with patterns of variation of trophic links per node as regards to species in a web (May, 1983; Paine, 1988). Although it has been shown that connections increase with species (Cohen and Briand, 1988), the rate of increase is not clear (Martinez, 1992, 1993; Paine, 1988; Schoener, 1989; Pimm et al., 1991).

Recently, ecological networks have been compared to other types of complex networks that are "small-world" (Strogatz, 2001; Albert and Barabási, 2002). In "small-world" networks most nodes are not neighbors of one another, but most nodes can be reached from every other by a small number of steps. There are few nodes highly connected (hubs) and the remaining nodes with low connectivity (power-law distribution). Such networks are large and complex, exhibiting a highly clustered structure into sub-graphs (strongly connected components) and small diameter (maximum length of shortest paths connecting two nodes). Food webs partially act as "small world" graphs, showing short paths between species (Williams et al., 2002), but unlikely other networks have low clustering and high link densities (Dunne et al., 2002a). In general, food webs with a uniform degree distribution are associated to relatively high connectance, exponential distribution is correlated to middle connectance, while power-law or partial power-low are displayed by webs with very low connectance (Montoya and Solé, 2002, 2003; Dunne et al., 2002a; Camacho et al., 2002a). As well as identifying link distribution and structure, within food web theory the larger amount of studies are addressed at the whole system level (McCann, 2000): (a) estimating chain length (Post, 2002) and number of trophic levels (Pimm and Lawton, 1977); (b) measuring indices based on information theoretical analysis (Bondavalli et al., 2006); (c) detecting mechanisms linking structural complexity and ecosystem stability (Montoya et al., 2006; Yodzis, 1981; Polis, 1994); (d) simulating structural robustness to species loss (Dunne et al., 2002b, 2004; Allesina et al., 2006; Tilman and Downing, 1994); (e) defining simple models capturing food web patterns (Stouffer et al., 2006; Cattin et al., 2004; Williams and Martinez, 2000; Cohen et al., 1990).

Few publications contribute to the quantification of key players in ecosystem, introducing sociometrical methods in ecology (Jordán and Scheuring, 2002; Jordán et al., 1999, 2006; Vasas and Jordán, 2006).

To systematically carry out analysis and description of complicated networks (food webs with weighted trophic links), a collection of quantitative methods has been developed and widely applied: ecosystem network analysis (ENA; Ulanowicz, 1986). It has been introduced in ecology by Hannon (1973) that translated economic analysis (i.e. the structural analysis; Leontief, 1963) into ecological terms. In a few years the technique was extended to include more ecologically significant results thanks to the prominent contributions of Finn (1976), Ulanowicz (1980), Patten (1982) and many others. During the last decade, many papers on ENA have been published in top ecological journals (Bersier et al., 2002; Krause et al., 2003) and the importance of studying weighted food web data stimulates its application.

As suggested by Margalef (1968, 1991), multi-species trophic interactions can be analyzed using a network perspective. Then, natural systems are described as directed networks of species connected by weighted or binary trophic relations. The innovative feature of this work is represented by the analysis of the contribution of single compartments (species or trophospecies) in respect of food web properties. Then, in opposition to the system approach common to all the studies cited in the above review of food webs and network ecology, I explored the effects of each species as a conse-

quence of its position within the ecosystem. I adopted the methodological approach of ENA and reviewed fundamental aspects of food web theory and ecosystem network analysis. New techniques detecting peculiar aspects of ecosystems, and investigating regularities emerging from their topology and trophic structure, were proposed.
In the next paragraph a brief summary of the thesis is given.

1.2 Overview of the thesis

In Chapter 2, ecosystem network analysis procedures are briefly described. Indices and measures that take into account both structure of networks (link distribution) and weight of interactions (their intensities) were summarized. The following chapters illustrate my contributions to the development of the discipline. All these accomplishments are coupled with ad-hoc software ameliorations that can deeply impact the results and the way of doing research on ecological networks.

Trophic position is often used in ecology as a synthetic index characterizing species activity in flow networks. It measures average distance of each node from the external source of energy, defining species roles from primary producers to consumers. Within ENA, the suite of matrix manipulations called canonical trophic aggregation (CTA) apportions each species feeding activity to a series of discrete trophic levels sensu Lindeman, allowing the calculus of trophic positions. Despite some drawbacks, this procedure is widely applied. In Chapter 3 I propose an extended version of CTA, resolving its major ambiguities and defining the computation of effective trophic positions in ecological acyclic networks. In particular, mathematical details with examples are provided and addressed to: (a) including migratory import with trophic position far from zero (i.e. input to non-primary producer nodes); (b) making CTA scale-insensitive (avoiding biases such as different trophic position for the same node when calculated at different scale of resolution); (c) resolving inconsistencies when flows are measured using different currencies.

Chapters 4, 5 and 6 are strongly focussed on relations between whole network properties and relative contribution of single species. Although widely acknowledged, few studies explored and defined quantitative relationships linking flow structure and trophic hierarchy in food webs. In Chapters 5 and 6, weighted food webs were analyzed using species trophic position (TP), to characterize the trophic hierarchy, and single species average mutual information ($AMI_{.j}$) as a measure of the link density that combines the number of interactions pertaining each node and their magnitude (see Chapter 4).

The sign of the statistical correlation between the two indices is constantly positive across all the systems investigated. This reveals a characteristic pattern that, according to the meaning of the whole average mutual information (AMI), suggests how link density becomes rarer towards the top of the trophic hierarchy. Accordingly, trophic hierarchy seems to explain the topological configuration of links in food webs. Thus, my analysis indicates that species feeding at higher position tend to trophic specialization and a more generalist alimentary behaviour characterizes species at the bottom of the trophic hierarchy.

Linear correlation between TP and $AMI_{\cdot j}$ is preserved within taxonomic subgroups, pointing out how members of one taxon may occupy different roles in the same ecosystem. When coefficients are ranked, subgroups aggregate into three main clusters in accordance with phylogenetic and evolutionary schemes.

Analyzing weighted food webs I included functional information in addition to topological one, identifying trophic hierarchy as a key factor in energy delivery. The results allow to extend the discussion about certain key ecological issues such as the relation between weak and strong interactions in ecosystems, the role of omnivory and the connections characterizing functional and taxonomic features of species in ecological networks.

Moreover, the relationships between trophic position and link distribution observed with weighted and unweighted data add another evidence about the potential of niche and cascade models described in Chapter 4, emphasizing the potential of these extremely simple architectures that are able to capture the very essence link arrangements in food webs.

Finally, patterns between species trophic positions and topological centralities were detailed in Chapter 7, while effects of weighting links when identifying keystone species in different ecological networks are investigated in Chapter 8. Studying centralities and trophic positions opens new perspectives towards the identification of potential biases in finding high centrality nodes among basal, intermediate and top species. Observed features could lead to intriguing consequences on ecosystem functioning.

Estimating the importance of weighting links respect to centrality indices (direct or indirect), evolutionary stability of interaction types (e.g. predator-prey, plant-pollinator, plant-seed disperser, plant-ant and host-parasite) and network size were unveiled. I found that: (a) weighting affects node ordering very seriously; (b) food webs fundamentally differ from other network types in this respect, (c) direct and indirect indices provide fairly different results but indirect effects are similar if longer than two steps; (d) the effect of weighting depends on the number of network nodes in case of direct interactions only.

2

Ecosystem Network Analysis

2.1 Why ecosystem network analysis?

The ever increasing interest towards ecosystem status and performance, and the need to approach complex environmental problems, stimulate the application of tools for whole-system assessment (NSF, 1999). The most common method for quantifying system level events is simulation modeling that implies five main steps: (a) identifying relevant taxa; (b) defining the significant interactions among those taxa; (c) modeling such interactions; (d) calibrating and validating the model; (e) making predictions.

Despite successful models describing the dynamics of one or a few subjects, mathematical modeling appears complicated when their number increases (Platt et al., 1981). As a consequence, to bypass the above mentioned inconsistencies in applications to system ecology, MacArthur (1955) and Platt et al. (1981) suggested an alternative approach paying more attention to processes (flows) than concentrating on what can be inferred from identification and analysis of single objects (stocks). In this conceptual framework, a prominent collection of quantitative methods consists of ecosystem network analysis (ENA; Ulanowicz, 1986; Wulff and Ulanowicz, 1989; Ulanowicz, 2004). The basic assumption behind ENA is that topology (statical graph configuration associated to trophic links between species) reveals much about history, current status and functioning of ecosystems.

2.2 Describing flow networks

An ecological network is a representation of the significant taxa or nodes comprising the ecosystem. It answers two questions: (a) who eats whom?, and (b) at what rate? Systems are depicted as directed graphs, or digraphs, with nodes (compartments) as boxes, vertices or points connected by arrowhead arcs portraying trophic relations (exiting the prey items and entering the predator). In addition, being ecological networks open systems exchanging material and energy with their surroundings, exogenous transfers are classified as: (a) input (e.g. primary production, immigration or inbound advection of material or energy), arrows that originate out of no visible taxon and terminate (with an arrowhead) into the receiving node; (b) export (e.g. emigration, harvesting by humans, and advection out of the system), arrows of useful matter or energy exiting from a node and pointing out to empty space; (c) respiration (e.g. energy dissipated into heat or material degraded into its lowest-energy form as denitrification to N_2), represented as "ground symbols" leaving compartments. Trophic links in these networks stand for energy flows (e.g. $kcal\ m^{-2}\ yr^{-1}$) or nutrient transfers of different currencies (e.g. carbon - $mgC\ m^{-2}\ yr^{-1}$; nitrogen - $mgN\ m^{-2}\ day^{-1}$; phosphorus - $gP\ m^{-2}\ yr^{-1}$). ENA allows only one medium per network, although in other works several are included (Hannon et al., 1991).

Setting ecosystem boundaries and the level of resolution are still open questions. Recent works evaluated effects of taxa aggregation in trophospecies or trophic roles (Yodzis and Winemiller, 1999; Abarca-Arenas and Ulanowicz, 2002; Luczkovic et al., 2003; Krause et al., 2003) and consequences of including non-living compartments (Allesina et al., 2005b). Generally, decisions about the degree of resolution are driven by the amount of available information and purpose of the study (Jordán, 2003).

As an example of ecological network, a schematic picture of the total suite of energy flows occurring in the Cone Spring ecosystem (Tilly, 1968) is drawn in Figure 2.1.

Beside the more intuitive graphical scheme, link topology can also be represented using linear algebra. The maximal amount of internal links in an ecosystem of S taxa is S^2 $(S \cdot S)$, to which add an upper limit of $3S$ flows across system bound (S inputs + S exports + S respirations). Whenever S exceeds about 20, the number of realized transfers usually falls well below the maximal limit of $S^2 + 3S$.

The flows depicted in Figure 2.1 can be expressed in matrix notation of order S as:

$$\begin{bmatrix} Plants \\ Bacteria \\ Detritus\ Feeders \\ Carnivores \\ Detritus \end{bmatrix}$$

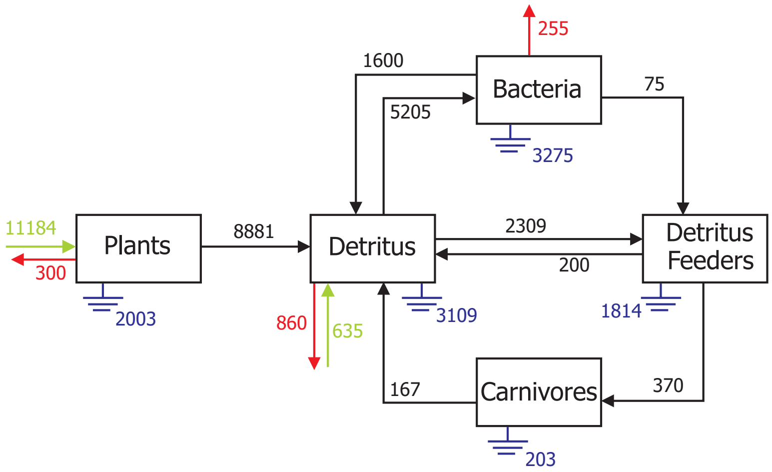


Figure 2.1: Cone Spring ecosystem is composed of 5 compartments (Plants, Bacteria, Detritus Feeders, Carnivores and Detritus) and it shows: 2 imports (green), 3 exports (red), 5 dissipations (blue) and 8 intercompartmental exchanges (black). Flows are expressed as $kcal\ m^{-2}\ yr^{-1}$.

$$\mathbf{T} = \begin{bmatrix} 0 & 0 & 0 & 0 & 8881 \\ 0 & 0 & 75 & 0 & 1600 \\ 0 & 0 & 0 & 370 & 200 \\ 0 & 0 & 0 & 0 & 167 \\ 0 & 5205 & 2309 & 0 & 0 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} 11184 \\ 0 \\ 0 \\ 0 \\ 635 \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} 300 \\ 255 \\ 0 \\ 0 \\ 860 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 2003 \\ 3275 \\ 1814 \\ 203 \\ 3109 \end{bmatrix}$$

$[T]$ represents the adjacency matrix with link strength at non-zero entries: the rate of the intercompartmental transfer from row (prey) taxon i to column (predator) node j is denoted as t_{ij}. Exogenous flows are classified as input (Z), export (E) and respiration (R) vectors, which elements z_i, e_i and r_i signify: (a) rate of external import to taxon i; (b) amount of loss of useful medium from taxon i to the outside world; (c) dissipation from node i.

ENA requires systems at steady state, that is, for each compartment i, the sum of all the inputs exactly balances the sum of all the outputs

$$\sum_{k=1}^{S} t_{ki} + z_i = \sum_{j=1}^{S} t_{ij} + e_i + r_i \tag{2.1}$$

or

$$T_{\cdot i} + z_i = T_{i\cdot} + e_i + r_i \tag{2.2}$$

since

$$\sum_{v=0}^{S} T_{iv} = T_{i\cdot} \tag{2.3}$$

$$\sum_{r=0}^{S} T_{rj} = T_{\cdot j} \tag{2.4}$$

Because of different methods used for quantifying flows in network construction (e.g. sampling and direct observation, literature screening of metabolic parameters and growth function of single compartments, contacting experts on specific taxa), many ecological networks appear, at first glance, unbalanced. Steady state condition can be achieved applying different balancing methods (Parker, 1977; Polovina, 1984; Allesina and Bondavalli, 2003).

Once we focus on ecological networks, matrix and vector manipulation adopted by ENA procedures allows to stress: (a) input-output analysis; (b) trophic structure analysis; (c) assessment of indirect effects using the mixed trophic impact; (d) cycling analysis and (e) estimation of whole-system and information-theoretic indices. The latter two aspects do not require that components be balanced.

2.2.1 Input-output analysis

Partial feeding matrix

While matrix of transfers $[T]$ describes absolute amount of matter or energy flowing between compartments, partial feeding matrix $[G]$ is estimated normalizing these connections by the total intake of each receiving node

$$g_{ij} = \frac{t_{ij}}{T_{\cdot j} + z_j} \tag{2.5}$$

Columns of $[G]$ ($G_{\cdot j}$, henceforth) sum to 1 in absence of exogenous input and to $0 \leq G_{\cdot j} < 1$ in all the other cases. Reading down j columns, a picture of relative percentages that each dietary taxa i constitutes of the full intake by j is supplied. Diet composition expressed by $[G]$, and its algebraic powers $[G]^m$, provide details on the fraction of flows moving from row-compartment to column taxa, through m steps (e.g. $[G]^2 = [G] \times [G]$

shows, in every non-zero element, relative contribution of i node to j diets, up to 2 steps).

Leontief structure matrix

In particular, being the $g_{ij} \leq 1$, Simon and Hawkins (1949) demonstrated how the series of matrices

$$[I] + [G] + [G]^2 + [G]^3 + [G]^4 \cdots \rightarrow [I - G]^{-1} \tag{2.6}$$

converges to the finite limit $[I - G]^{-1} = [S]$, the Leontief structure matrix (Leontief, 1951) where $[I] = [G]^0$ stands for the identity matrix with elements equal to 1 along the diagonal ($\delta_{ii} = 1$) and to 0 elsewhere ($\delta_{ij} = 0$, if $i \neq j$). Elements of $[S]$ summarize the fraction of matter/energy flowing from row-compartments to column-nodes over pathways of all lengths, per unit of final demand. Diagonal coefficients (s_{ii}) greater than 1 indicate compartments involved in cycles.

Total dependency matrix

Szyrmer and Ulanowicz (1987) used the Leontief structure matrix to derive the total dependency matrix $[D]$, which d_{ij} coefficients represent the fraction of diet of j that passes through i, scaled by their throughflows

$$d_{ij} = (s_{ij} - \delta_{ij}) \left(\frac{T_{i\cdot}}{s_{ii}\, T_{\cdot j}} \right) \tag{2.7}$$

where s_{ij} is the corresponding coefficient in $[S]$, s_{ii} is a diagonal element of $[S]$, δ_{ij} is the identity matrix element, $T_{i\cdot}$ and $T_{\cdot j}$ are throughflows of donor and receiving nodes, respectively.

Feeding and dependency analysis in Cone Spring

Using data of Cone Spring ecosystem (see Figure 2.1), the associated partial feeding $[G]$, Leontief structure $[S]$ and total dependency $[D]$ matrices are:

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.773 \\ 0 & 0 & 0.031 & 0 & 0.139 \\ 0 & 0 & 0 & 1 & 0.017 \\ 0 & 0 & 0 & 0 & 0.015 \\ 0 & 1 & 0.969 & 0 & 0 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} 1 & 0.933 & 0.933 & 0.933 & 0.933 \\ 0 & 1.169 & 0.201 & 0.201 & 0.169 \\ 0 & 0.039 & 1.039 & 1.039 & 0.039 \\ 0 & 0.018 & 0.018 & 1.018 & 0.018 \\ 0 & 1.207 & 1.207 & 1.207 & 1.207 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0.933 & 0.933 & 0.933 & 0.933 \\ 0 & 0.145 & 0.172 & 0.172 & 0.145 \\ 0 & 0.037 & 0.037 & 1.000 & 0.037 \\ 0 & 0.017 & 0.017 & 0.017 & 0.017 \\ 0 & 1 & 1 & 1 & 0.171 \end{bmatrix}$$

Partial host matrix

Analogously to the $[G]$ matrix obtained normalizing each t_{ij} element by the total input of the corresponding column node j, a matrix of host coefficients $[F]$ is computed as the ratio between each t_{ij} element and total flows exiting the row node i. Then, the fraction of the total activity of i that flows directly to taxa j is

$$f_{ij} = \frac{t_{ij}}{T_{i\cdot} + e_i + r_i} \tag{2.8}$$

Output structure matrix

Augustinovic (1970) formulated the output structure matrix $[\Sigma] = [I - F^T]^{-1}$ from which the total activity of each compartment j, generated by a unitary input to i, is estimated (σ_{ij}).

Total contribution matrix

Finally, the total contribution matrix $[C]$ is calculated weighting each σ_{ij} element by throughflows of corresponding row and column nodes

$$c_{ji} = (\sigma_{ij} - \delta_{ij}) \left(\frac{T_{\cdot j}}{\sigma_{ii}\, T_{i\cdot}} \right) \tag{2.9}$$

One may regard the elements of the total contribution matrix as the efficiencies with which medium flows from any one compartment to any given other.

Host and contribution analysis in Cone Spring

Partial host $[F]$, output structure (or Augustinovic) $[\Sigma]$ and total contribution $[C]$ matrices for the example of Cone Spring network (see Figure 2.1) would be:

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.794 \\ 0 & 0 & 0.014 & 0 & 0.307 \\ 0 & 0 & 0 & 0.155 & 0.084 \\ 0 & 0 & 0 & 0 & 0.451 \\ 0 & 0.453 & 0.201 & 0 & 0 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.434 & 1.169 & 0.084 & 0.247 & 0.547 \\ 0.199 & 0.092 & 1.039 & 0.113 & 0.251 \\ 0.031 & 0.014 & 0.161 & 1.018 & 0.039 \\ 0.958 & 0.374 & 0.186 & 0.545 & 1.207 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 0 & 0.371 & 0.192 & 0.030 & 0.794 \\ 0 & 0.145 & 0.089 & 0.014 & 0.310 \\ 0 & 0.072 & 0.037 & 0.158 & 0.154 \\ 0 & 0.211 & 0.109 & 0.017 & 0.451 \\ 0 & 0.468 & 0.241 & 0.038 & 0.171 \end{bmatrix}$$

Input and output analysis

Output (Augustinovic) and input (Leontief) structure matrices are starting tools for I/O analysis as well as for dependency and contribution studies. Assessing the proportions of any internal exchanges ascribed to a unitary input (or output) of matter/energy flows is the aim of I/O analysis. For example, to know the fate of 1 $kcal\ m^{-2}\ yr^{-1}$ input on plants, internal flows are defined multiplying the output structure matrix by a column vector with 1 in the first row (standing for plants) and 0 elsewhere

$$\mathbf{\Sigma}[,\mathbf{1}] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.434 & 1.169 & 0.084 & 0.247 & 0.547 \\ 0.199 & 0.092 & 1.039 & 0.113 & 0.251 \\ 0.031 & 0.014 & 0.161 & 1.018 & 0.039 \\ 0.958 & 0.374 & 0.186 & 0.545 & 1.207 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.434 \\ 0.199 \\ 0.031 \\ 0.958 \end{bmatrix}$$

Then, multiplying each row of the host coefficient matrix $[F]$ by the corresponding throughput in $\Sigma[,1]$, direct flows from inputs on plants $[T_{I1}]$ are found

$$\mathbf{T_{I1}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.794 \\ 0 & 0 & 0.006 & 0 & 0.134 \\ 0 & 0 & 0 & 0.031 & 0.017 \\ 0 & 0 & 0 & 0 & 0.014 \\ 0 & 0.434 & 0.193 & 0 & 0 \end{bmatrix}$$

The fate of a single arbitrary unit of energy entering in the plant compartment of Cone Spring network is depicted in Figure 2.2

The same procedure can be applied to input structure $[S]$ and partial feeding $[G]$ matrices, to estimate origins, and trace direct exchanges within the system, of a unitary output flow.

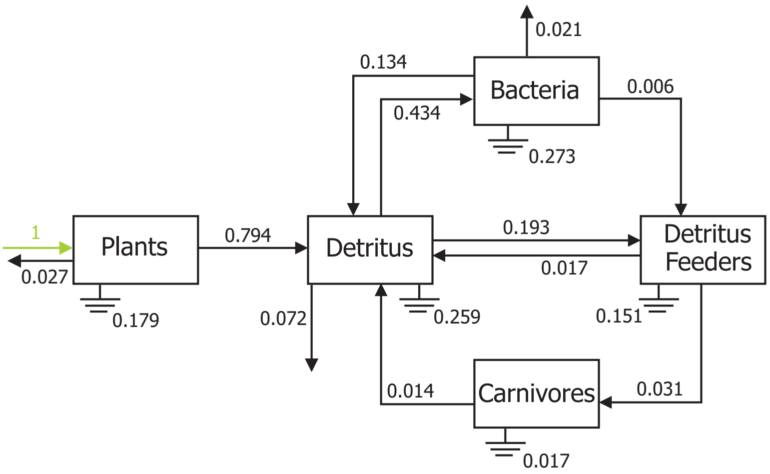


Figure 2.2: Description of internal direct Cone Spring flows, respirations and exports for a 1 $kcal\ m^{-2}\ yr^{-1}$ input set on plants (green arrow).

2.2.2 Trophic structure

The trophic status of ecological networks can be unveiled either by apportioning each species feeding activity to a series of discrete trophic levels (TLs) sensu Lindeman (1942) and by computing the effective (non-integer) trophic position (TP) of each compartment. In the former case, the intricacy of the whole ecosystem flows is transformed into a simple linear chain composed of discrete trophic levels. In the latter, effective trophic position is computed as the sum of the fractions of trophic activity that each species performs at different trophic levels.

In the absence of non-living nodes and cycles between living taxa, TPs of each i species could easily be defined by the i^{th} column sum of the Leontief structure matrix (Levine, 1983). Unfortunately cycles are the rule in nature and ecological networks often include non-living compartments. A widely applied alternative method to calculate TPs and dealing with this problem is called canonical trophic aggregations (CTA; Ulanowicz and Kemp, 1979; Ulanowicz, 1995). Basic tool of CTA is the trophic transformation matrix $[A]$, built up with a row by row procedure.

The method, based on trophic transformation matrix $[A]$, requires one to assign trophic positions to compartments according to diet composition. The first line of $[A]$ is the transposed normalized input vector $(N)^T$, where a_{1j} elements are computed as

$$a_{1j} = n_j = \frac{z_j}{T_{.j}} \tag{2.10}$$

Nodes receiving imports act as primary producers (with theoretical import $TL = 0$); if a compartment relies only on import, its TP will be 1.

Three main cases can be distinguished: (a) when $n_j = 0$, the compartment j does not receive input from the outside; (b) if $n_j = 1$, node j depends only on input; (c) with $0 < n_j < 1$, the node j receives medium both from the outside and the internal exchanges.
To compute the following rows of $[A]$, the matrix of partial feeding coefficient $[G]$ and the normalized import vector (N) are required:

$$A[i,] = (N)^T \times [G]^{i-1} \quad \text{with } i > 1 \tag{2.11}$$

where $A[i,]$ is the i^{th} row of the matrix, and powers of $[G]$ are defined as repeated product of $[G]$ by itself: $[G]^0$ stands for identity matrix (usually called $[I]$), $[G]^1 = [G]$, $[G]^2 = [G] \times [G]$, and so forth.
In $[A]$ each column sums to 1, satisfying the first law of thermodynamics (total input, throughput, export and respiration remain invariant under transformation), and all the resultant respirations are positive (in accordance with the second law of thermodynamics).

Besides computing TPs, CTA can also be used to ascertain the portions of all flows which are the same number of steps from any external input, defining the linear trophic chain of ecosystems. The composition of the i^{th} trophic level can be read along the i^{th} row of $[A]$, while the trophic behaviour of each j^{th} species can be inferred by the elements of the j column of $[A]$.
Non-integer TPs emerge as the result of a weighted mean of different trophic behaviours. In fact, each TP_j of the trophic position row vector (TP) is defined as

$$TP_j = \sum_{i=1}^{S-1} a_{ij} \cdot i \tag{2.12}$$

Cycles impose constraints to the transformation of ecological networks into linear trophic chains. Dealing with cyclic networks, powers of $[G]$ form an infinite sequence, and it is not clear how and where to interrupt the number of $[A]$ rows.
In presence of cycles, Ulanowicz (1995) suggested to split a system into two constitutive networks: one containing cyclic paths, and another depicting only once-through flow. Cycles are removed by a simple procedure (Ulanowicz, 1983): (a) listing all the simple directed cycles; (b) finding the weakest arc for every cycle (the one with the minimum flow); (c) grouping the cycles sharing the same weakest arc (also called *nexus*); (d) removing cycles according to their probabilities, making the *nexus*=0.
When cycles are widespread, a large amount of information about total system activity is lost during network decyclization, resulting in distorted values of calculated trophic efficiencies. However, as Pimm (1982) and May (1983) remark, feeding cycles between living compartment are rare (this is not true,

anymore, when considering detritus and nutrient pools), and networks often embed compartments as detritus and nutrient pools (non-living nodes). Non-living nodes have no trophic position, but they influence TPs of living compartments apportioning on them. In this case, submatrices $[T]_{living}$ and $[G]_{living}$, accounting only for exchanges between living compartments with scarcity of cycles, are isolated (Ulanowicz, 1995). As a consequence, applying the decyclization algorithm to living subsystems with few cycles does not distort CTA and the aggregation of the feeding web into discrete trophic levels.

Due to difficulties related to TP calculation in presence of migratory imports (with TP far from 0) and multiple non-living nodes (with $TP = 1$ or TP=0), the entire procedure is reviewed and amended, by an extended version of CTA, in Chapter 3.

As it is not possible to compute the matrix $[A]$ for the Cone Spring ecosystem (because of cycles and presence of detritus bridging "Plants" compartment and the other living nodes), CTA is now applied to the hypothetical network of Figure 2.3

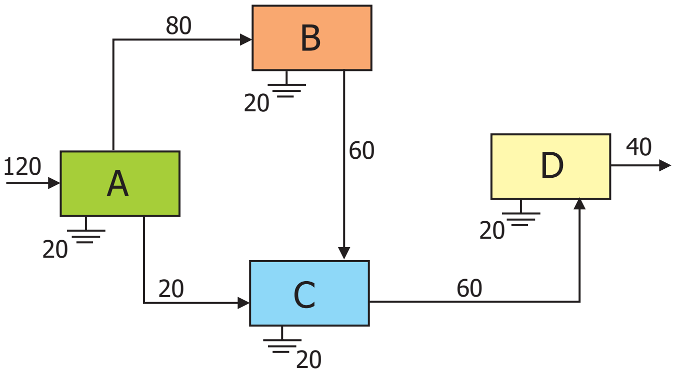


Figure 2.3: Hypothetical system made by 4 living compartments. Compartment A is a primary producer ($TP = 1$), relying only on imports, while B, feeding exclusively on A, is a herbivore ($TP = 2$). C shows an intermediate TP, since its diet is composed of A (primary producer) and B (herbivore). Non-integer TP is exhibited by D that prey upon C.

for which flow matrices are:

$$\mathbf{T} = \begin{bmatrix} 0 & 80 & 20 & 0 \\ 0 & 0 & 60 & 0 \\ 0 & 0 & 0 & 60 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} 120 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 40 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 20 \\ 20 \\ 20 \\ 20 \end{bmatrix}$$

The associated normalized import vector (N) and partial feeding matrix $[G]$ are

$$\mathbf{N} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{G} = \begin{bmatrix} 0 & 1 & 0.25 & 0 \\ 0 & 0 & 0.75 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

giving rise to the following trophic transformation matrix $[A]$ and trophic position vector (TP)

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0.25 & 0 \\ 0 & 0 & 0.75 & 0.25 \\ 0 & 0 & 0 & 0.75 \end{bmatrix} \mathbf{TP} = \begin{bmatrix} 1 \\ 2 \\ 2.75 \\ 3.75 \end{bmatrix}$$

Reading down the columns of $[A]$, A receives energy from outside and acts like a primary producer, with $TP = 1$ ($TP_1 = a_{11} \cdot 1 = 1 \cdot 1 = 1$); B feeds exclusively on A and is set to $TP = 2$ ($TP_2 = a_{22} \cdot 2 = 1 \cdot 2 = 2$); C shows a fractionary trophic position ($TP = 2.75$), as the weighted effects of two different trophic behaviours ($TP_3 = a_{23} \cdot 2 + a_{33} \cdot 3 = 0.25 \cdot 2 + 0.75 \cdot 3 = 2.75$); D, preying on C, has $TP = 3.75$ ($TP_4 = a_{34} \cdot 3 + a_{44} \cdot 4 = 0.25 \cdot 3 + 0.75 \cdot 4 = 3.75$).

Conversely, $[A]$ can be used to distribute each compartment trophic activity into integer TLs, obtaining a Lindeman spine for a complex network (see Figure 2.4). Then, A is the only primary producer (first line of $[A]$), herbivore behaviour (second line of $[A]$) is displayed by B (1/(1+0.25)=80%) and C (0.25/(1+0.25)=20%), primary consumer TL (third line of $[A]$) includes both C (75%) and D (25%) activities, while D is the only node with a partial secondary consumer diet composition (100%).

2.2.3 Mixed trophic impact and indirect effects

Conventional I/O and trophic analysis are often criticized because they pertain to steady-state or temporally-averaged configurations, dealing only

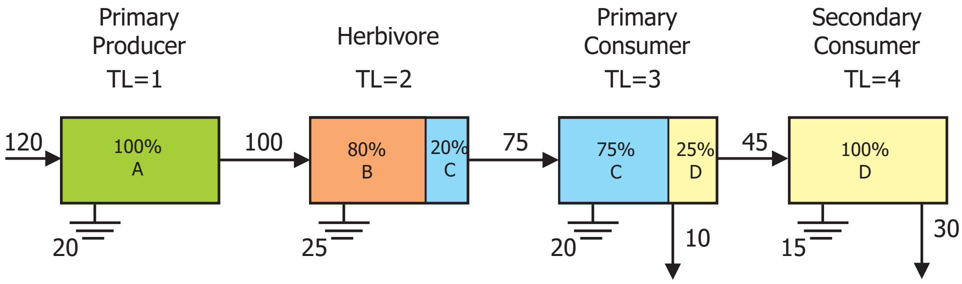


Figure 2.4: Flows along the trophic chain associated to the network depicted in Figure 2.3. While primary producer and secondary consumer include single species activity (A and D, respectively), herbivore and primary consumer behaviours are shared by multiple nodes. In the first case, both B (outside $\rightarrow A \rightarrow B$: $80/80 = 100\%$) and C (outside $\rightarrow A \rightarrow C$: $20/80 = 25\%$) feed at $TL = 2$; in the latter, C (outside $\rightarrow A \rightarrow B \rightarrow C$: $60/80 = 75\%$) and D (outside $\rightarrow A \rightarrow C \rightarrow D$: $15/60 = 25\%$), act as a primary consumer ($TL = 3$). Trophic efficiencies between TLs are estimated as usable energy fraction transferred in each step (i.e. 120 unit of energy entering primary producer and 100 unit flowing to herbivore $= 100/120 = 83.33\%$ of efficiency). Efficiencies measured in this trophic chain are: (a) 83.33% primary producer → herbivore; (b) 75.00% herbivore → primary consumer; (c) 60.00% primary consumer → secondary consumer.

with positive flows of medium. Ulanowicz and Puccia (1990) adopted an alternative approach for measuring direct and indirect trophic impacts from trophic flow data. Positive impact of prey i upon predator j is quantified by partial feeding coefficient (g_{ij}), defined as the proportion of i within the diet of j (see equation 2.5); the negative effect of predator j on its prey i (f_{ij}) is measured introducing an amended version of the equation 2.8, with t_{ij} normalized by the net output from i (flows to living members $= \sum_{z=1}^{living} t_{iz}$), excluding exports ($e_i$) and respiration ($r_i$)

$$f^*_{ij} = \frac{t_{ij}}{\sum_{z=1}^{living} t_{iz}} \tag{2.13}$$

The net impact of i upon j, accounting for one-step (direct) effect, is defined as

$$q_{ij} = g_{ij} - f^*_{ji} \tag{2.14}$$

where $-1 < q_{ij} < +1$ are constitutive elements of the net impact matrix $[Q]$. Net impacts upon x steps are measured by $[Q]^x$ matrix and, in almost all instances (with non-singular $[I - Q]$ matrix), the powers of $[Q]$ give rise to a convergent series (Hannon, 1973)

$$[I] + [Q] + [Q]^2 + [Q]^3 + [Q]^4 \cdots \rightarrow [I - Q]^{-1} \tag{2.15}$$

The total (direct and indirect) effects are calculated introducing the matrix $[M]$

$$[M] = \sum_{x=1}^{\infty} [Q]^x = [I - Q]^{-1} - [I] \tag{2.16}$$

By $[M]$ matrix one identifies species that have a disproportionate effect on the environment relative to their abundance (keystone species; Paine, 1995) and ranking the whole impacts exerted on a given taxa is possible (Bondavalli and Ulanowicz, 1999). Vasas and Jordán (2006) proposed to sum absolute values of m_{ij} elements of $[M]$, unveiling both positive and negative effects displayed by each species (otherwise strong negative and strong positive impacts might result masked in an effect around zero).

Similar analysis has been developed by Fath and Patten (1998). They demonstrated that the elements of $[M]$ tend to be more positive than the direct effects $[Q]$, with this trend strongly affected by the way they used to normalize t_{ij}. In fact, they adopted the equation (2.8) rather than (2.13), dividing each transfer from node i, by the full amount of its outflows ($T_{i\cdot}$).

2.2.4 Cycling analysis

Energy cycling is relevant to ecosystem properties: (a) residence time of nutrients (Herendeen, 1989); (b) buffering fluctuations in energy supply (Loreau, 1994); (c) augmenting stability (DeAngelis, 1980). Because of this, cycles have been deeply studied in ecological networks (Ulanowicz, 1983; Patten and Higashi, 1984; Christian and Thomas, 2003).

Within ENA, quantifying the fraction of total energy/matter recycled and describing paths involved in cycling (number of steps and compartments) are main targets. Finn (1976) dealt with the former issue and formulated a measure of it (Finn's cycling index, FCI). It accounts for the fraction of all the fluxes generated by cycling. Ulanowicz (1983) aimed to identify the structure of network cycles, characterizing the major pathways that recycle energy/matter .

The amount of cycling in flow networks

After Hannon (1973) introduced I/O analysis into ecology, attempts to quantify cycling were developed (Patten et al., 1976; Finn, 1976). Because in the Leontief structure matrix $[S]$ each diagonal element s_{ii} relates to the probability that a quantum of medium visits the designated compartment more than once, the amount of throughflow due to cycling is

$$T_{c,i} = T_{\cdot i} \left(\frac{s_{ii} - 1}{s_{ii}} \right) \tag{2.17}$$

where $T_{\cdot i}$ is the throughflow of node i. FCI is then obtained summing the contribution to cycling of each taxa and dividing this value by the total system throughput (TST).

$$FCI = \frac{1}{TST} \sum_{i=1}^{S} T_{c,i} \tag{2.18}$$

being S the number of compartments (both living and non-living) and TST the sum of all the flows (Ulanowicz, 1986; Patten and Higashi, 1984), as detailed below:

$$TST = \sum_{i=1}^{S} \sum_{j=1}^{S} t_{ij} + \sum_{i=1}^{S} (z_i + e_i + r_i) \tag{2.19}$$

Szyrmer and Ulanowicz (1987) proposed a revised form of FCI, replacing the use of Leontief structure matrix $[S]$ with total dependency $[D]$ or total contribution $[C]$ matrices. In fact, they noticed how diagonal elements of $[D]$ and $[C]$ are identical, representing a more detailed probability that a given quantum of currency leaves a particular taxon and returns to it. Han (1997), independently, achieved the same results. Finally, Allesina and Ulanowicz (2004) implemented FCI and its improved versions, describing how cycling flows in ecosystems were underestimated. Simple cycles (simple paths in which the starting and the ending node coincide) and compound cycles (repeated cycles) are accounted by FCI, but the contribution of compound paths (paths with repeated compartments) remains unexplored (see Figure 2.5). Allesina and Ulanowicz (2004) overcame this drawback with a comprehensive cycling index (CCI). The strong linear relationship between FCI and CCI ($CCI = FCI \cdot 1.142$) and the enormous amount of time required to compute the latter suggested to approximate the effective amount of cycling (CCI) simply by inflating the FCI by some 14.2%.

To calculate the cycling index for Cone Spring network (see Figure 2.1), Leontief structure matrix $[S]$, effective amount of flows attributable to cycling (T_c) and TST have been employed

$$\mathbf{S} = \begin{bmatrix} 1 & 0.933 & 0.933 & 0.933 & 0.933 \\ 0 & 1.169 & 0.201 & 0.201 & 0.169 \\ 0 & 0.039 & 1.039 & 1.039 & 0.039 \\ 0 & 0.018 & 0.018 & 1.018 & 0.018 \\ 0 & 1.207 & 1.207 & 1.207 & 1.207 \end{bmatrix}$$

$$\begin{aligned}
T_c &= \sum_{i=1}^{5} T_{\cdot i}\left(\frac{s_{ii}-1}{s_{ii}}\right) \\
&= 11184\left(\frac{1-1}{1}\right)+5205\left(\frac{1.169-1}{1.169}\right)+2384\left(\frac{1.039-1}{1.039}\right) \\
&+ 370\left(\frac{1.018-1}{1.018}\right)+11483\left(\frac{1.207-1}{1.207}\right) \\
&= 2817\ kcal\ m^{-2}\ yr^{-1} \qquad (2.20)
\end{aligned}$$

$$\begin{aligned}
TST &= \sum_{i=1}^{5}\sum_{j=1}^{5} t_{ij}+\sum_{i=1}^{5}(z_i+e_i+r_i) \\
&= 18807+11819+1415+10404 \\
&= 42445\ kcal\ m^{-2}\ yr^{-1} \qquad (2.21)
\end{aligned}$$

yielding to $FCI = 6.63\%$

$$FCI = \frac{T_c \cdot 100}{TST} = 6.63\% \qquad (2.22)$$

This value is different from the 9.20% calculated by Finn (1980) because, in that original formulation, TST^* stood for the sum of internal transfers plus imports

$$\begin{aligned}
TST^* &= \sum_{i=1}^{5}\sum_{j=1}^{5} t_{ij}+\sum_{i=1}^{5} z_i \qquad (2.23) \\
&= 30626\ kcal\ m^{-2}\ yr^{-1} \qquad (2.24)
\end{aligned}$$

$$FCI^* = \frac{T_c \cdot 100}{TST^*} = 9.20\% \qquad (2.25)$$

The structure of network cycles

Ulanowicz (1983, 1986) devised a procedure to extract cycles from the network and compare their activities with the remaining unidirectional flows. He adopted a backtracking algorithm to find all the simple cycles (Mateti and Deo, 1976), increasing its efficiency with a suitable pruning method (Knuth, 1973). The whole process can be summarized as follows: (a) a preliminary depth-first search to count the number of cycle arcs incident to each node; (b) ordering nodes by decreasing number of incoming cycle

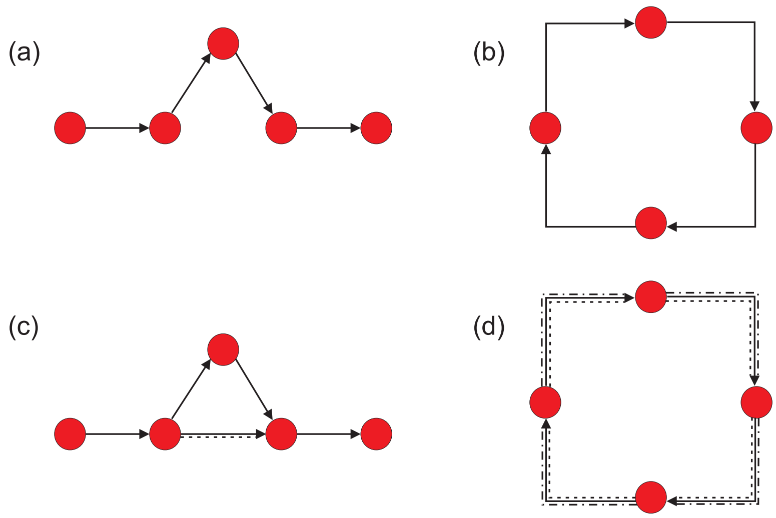


Figure 2.5: Pathways can be classified into four categories (Allesina and Ulanowicz, 2004): (a) simple paths; (b) simple cycles; (c) compound paths; (d) compound cycles. Excluding the first, remaining structures constituted cycled flows. While FCI exclusively consider (b) and (d), its amended version (CCI) also includes (c). Dotted lines stand for repeated transfers.

arcs (compartments with no cycle arcs are eliminated from further analysis); (c) scrutinizing each cycle in search for the smallest cycle arc; (d) identifying if this cycle link is common to other cycles and grouping cycles sharing it (*nexus*); (e) starting backtracking algorithm to delete the *nexus* and distribute its magnitude among member cycles (in proportion to circuit probabilities and magnitudes of links). Finally, the acyclic residual once-through flow web (tree, in graph theory) and the aggregated network of cycled medium result as two separated parts of the whole network.

The separation of cyclical transfers from the complementary web of once-trough flow are depicted for Cone Spring in Figure 2.6

2.2.5 Information-theoretic indices

[1]The ever increasing interest towards ecosystem status and performance stimulates the application of tools for whole-system assessment. The apparatus of ENA comprises indices, derived from information theory, that quantify global attributes. Total system throughput (TST) and average mutual information (AMI) are basal indices advocating a systemic approach (Ulanowicz, 2004). Ascendency (A), development capacity (C), overhead (Φ) and redundancy (Φ_R) are measures combining the total activity, or

---

[1] Published section: Scotti, M., in press. Ecological Indicators: Development Capacity and Overhead. For Jørgensen, S. E., Encyclopedia of Ecology - Elsevier.

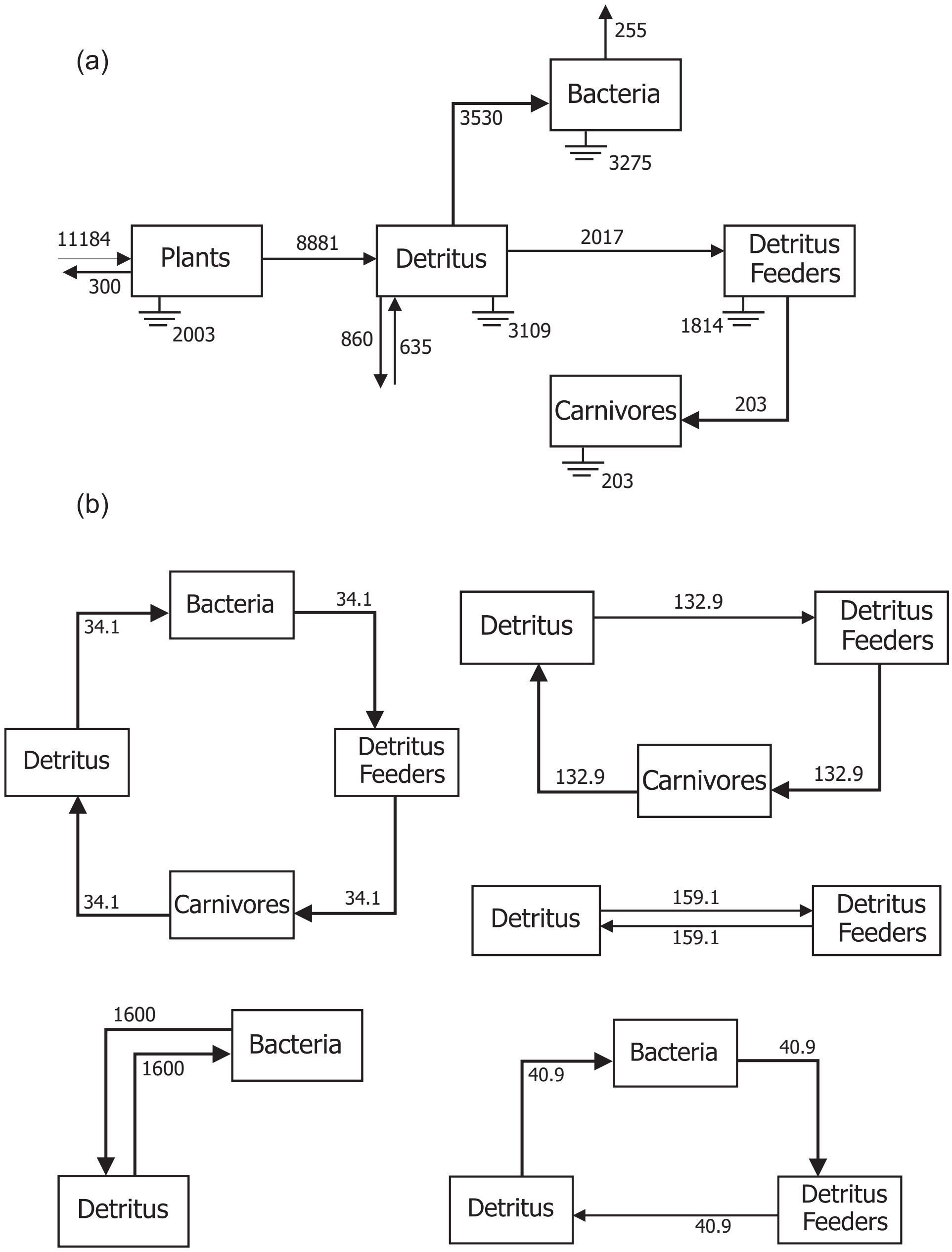


Figure 2.6: Cone Spring (see Figure 2.1) is decomposed into two constitutive components (Ulanowicz, 1983): (a) the residual acyclic network; (b) five simple directed cycles. Summing these complementary topologies yields the whole system structure. Flows are measured as $kcal\ m^{-2}\ yr^{-1}$.

power of the system, with the organization by which the component processes are linked (Latham and Scully, 2002). To facilitate their computation, a graphical scheme based on a $(S + 3) \times (S + 3)$ extended transfer matrix

$[T]^*$ can be introduced (Figure 2.7). Each t^*_{ij} element describes an outflow from row-compartment i to column-node j. First row stands for imports from outside to the column compartments, and the last two columns list exports and respirations from any row compartment. The internal $S \times S$ sub-matrix reports internal exchanges.

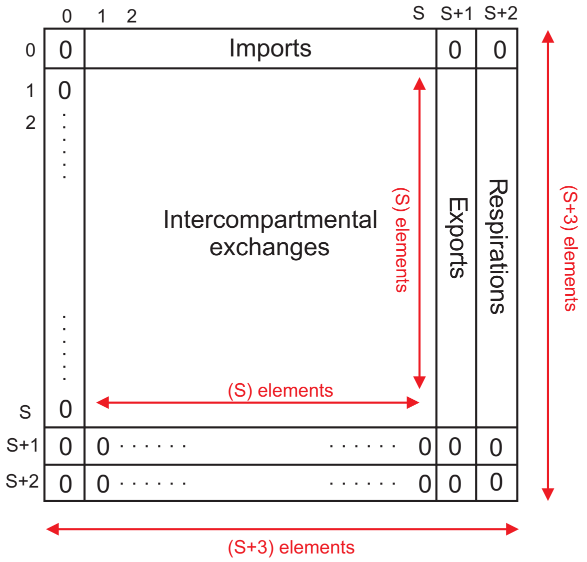


Figure 2.7: Graphical example of $[T]^*$ matrix for a generic S-compartment ecosystem, with size $= (S+3) \times (S+3)$. Imports to system nodes are summarized in the first row, exports and dissipations are set in the last two columns and the internal $S \times S$ matrix shows transfers between system compartments. This model of $[T]^*$ matrix is adopted, below, for Cone Spring.

$$\mathbf{T}^* = \begin{bmatrix} 0 & 11184 & 0 & 0 & 0 & 635 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8881 & 300 & 2003 \\ 0 & 0 & 0 & 75 & 0 & 1600 & 255 & 3275 \\ 0 & 0 & 0 & 0 & 370 & 200 & 0 & 1814 \\ 0 & 0 & 0 & 0 & 0 & 167 & 0 & 203 \\ 0 & 0 & 5205 & 2309 & 0 & 0 & 860 & 3109 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Total system throughput

The total amount of flows occurring in the system is called total system throughput (TST), and for Cone Spring is equal to 42445 $kcal\ m^{-2}\ year^{-1}$.

$$TST = \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} t^*_{ij} = T_{..} \tag{2.26}$$

TST defines ecosystem activity (size) and it conceptually corresponds to what in economy is known as gross national product (GDP), an indicator of economic community size.

Average mutual information

Ecosystem development is connected to flow organization and it increases when uncertainty is diminishing. Mathematically, uncertainty (H) is related to a distribution of probability over n categories (with scalar constant K) and is equivalent to

$$H = -K \sum_{i=1}^{n} p_i \; log \; p_i \tag{2.27}$$

System flow disorganization is measured by uncertainty and its amount can be distinguished into output $H(a)$ and input $H(b)$ contributions.

$$H(a) = -K \sum_{i=0}^{S+2} p(a_i) \; log \; p(a_i) \tag{2.28}$$

$$H(b) = -K \sum_{j=0}^{S+2} p(a_j) \; log \; p(a_j) \tag{2.29}$$

With completely independent events, total system uncertainty becomes $H(a) + H(b)$. However, inputs and outputs in ecosystems are not always independent and the associated uncertainty can be computed adopting joint probabilities $p(a_i, b_j)$

$$H(a,b) = -K \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} p(a_i, b_j) \; log \; p(a_i, b_j) \tag{2.30}$$

Therefore, with inputs and outputs that are not completely independent

$$H(a,b) < H(a) + H(b) \tag{2.31}$$

and, in this case, the degree of system organization is defined as

$$A(a;b) = H(a) + H(b) - H(a,b) \tag{2.32}$$

yielding to

$$A(a;b) = K \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} p(a_i, b_j) \; log \; \frac{p(a_i, b_j)}{p(a_i)p(b_j)} \tag{2.33}$$

because

$$p(a_i) = \sum_{j=0}^{S+2} p(a_i, b_j) \tag{2.34}$$

$$p(b_j) = \sum_{i=0}^{S+2} p(a_i, b_j) \tag{2.35}$$

To measure transfer uncertainty with network analysis notation, the probability that a quantum of matter (or energy) would flow from compartment i to j becomes

$$p(a_i, b_j) \cong \frac{t^*_{ij}}{T_{..}} \tag{2.36}$$

with output and input probabilities that can be written as marginal sums of joint probabilities

$$p(a_i) \cong \sum_{j=0}^{S+2} \frac{t^*_{ij}}{T_{..}} \tag{2.37}$$

$$p(b_j) \cong \sum_{i=0}^{S+2} \frac{t^*_{ij}}{T_{..}} \tag{2.38}$$

The average mutual information (AMI) is the index measuring system organization (ecosystem development intended as flow articulation). Using the expressions provided by formulas (2.36), (2.37) and (2.38) into (2.33), one obtains, after some algebraic manipulations, the following:

$$AMI = \frac{K}{TST} \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} t^*_{ij} \; log \left[\frac{t^*_{ij} \; TST}{\sum_{r=0}^{S+2} t^*_{rj} \; \sum_{v=0}^{S+2} t^*_{iv}} \right] \tag{2.39}$$

Ascendency and development capacity

Ascendency, being the product of TST by AMI, takes the following mathematical form

$$A = AMI \cdot TST = \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} t^*_{ij} \; log \left[\frac{t^*_{ij} \; TST}{\sum_{r=0}^{S+2} t^*_{rj} \; \sum_{v=0}^{S+2} t^*_{iv}} \right] \tag{2.40}$$

Assessing ecosystem growth and development can be done by comparing ascendency with its maximum and minimum limits. Ascendency (A), as obtained from TST and AMI, shows a minimum value of 0 (when output and input flow probabilities are completely independent, see Figure 2.8a), whereas the upper boundary is defined as development capacity (C).

For each ecosystem, the development capacity depends on the constraints established by real network topology. When the number of compartments (S) and TST are assigned, the highest development capacity is associated to a wholly connected and balanced network, decreasing when flows become more articulated (its minimum value is achieved with closed linear chain topology, when it corresponds to system ascendency - see Figure 2.8 and Table 2.1).

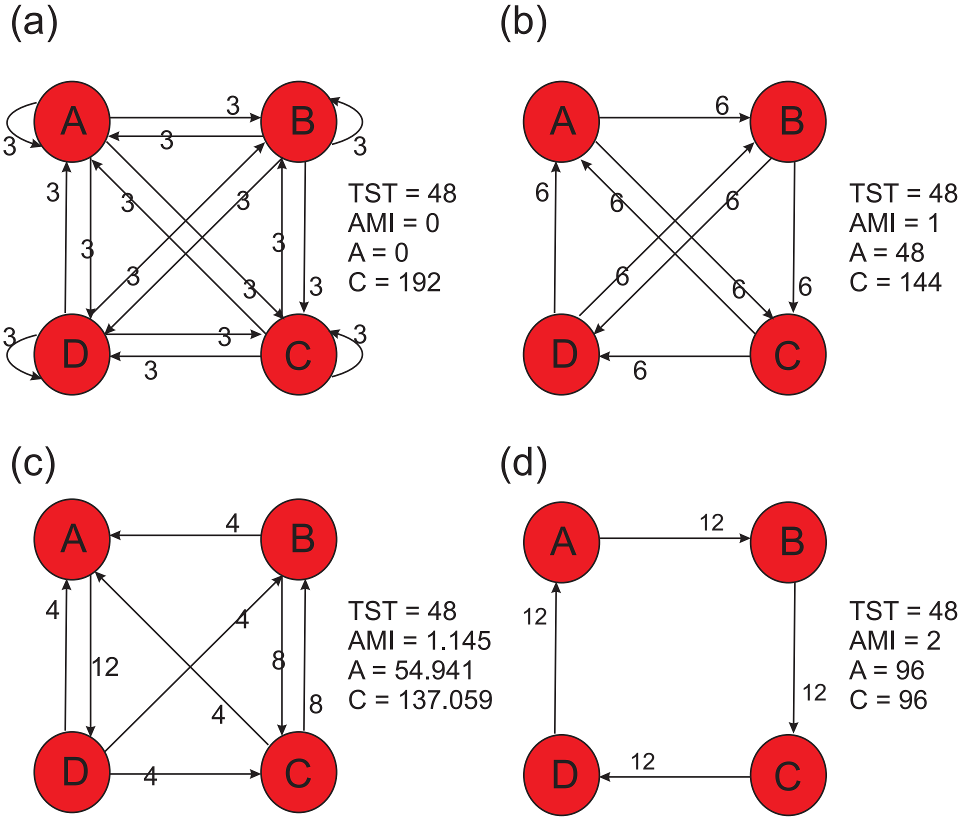


Figure 2.8: Hypothetical networks with four compartments and $TST = 48$ energy units: (a) is the more unarticulated topology with minimum AMI and ascendency values (both equal to 0) and maximum development capacity (192); (b) and (c) are intermediate configurations showing increasing AMI and A; C is lower than that observed for the first network; in (d) is depicted a closed and linear chain (maximally articulated flows) with highest AMI (2) and ascendency equal to development capacity (96).

In what follows, minimum and maximum ascendency are explained through a probabilistic approach. The lowest ascendency value is associated to fully connected topology, that is when input and output flow probabilities are completely independent:

$$p(a_i, b_j) = p(a_i)p(b_j) \tag{2.41}$$

Substituting this relation into equation (2.33) yields

Network topology	C
Fully connected (a)	192
Intermediate I (b)	144
Intermediate II (c)	137
Linear and closed chain (d)	96

Table 2.1: Development capacity (C) calculated for the hypothetical networks depicted in Figure 2.8; the values of C are decreasing when the network topology becomes more articulated.

$$\begin{aligned} A(a;b) &= K \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} p(a_i)p(b_j)\ log\ \frac{p(a_i)p(b_j)}{p(a_i)p(b_j)} \\ &= K \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} p(a_i)p(b_j)\ log(1) = 0 \end{aligned} \tag{2.42}$$

Conversely, under minimum uncertainty conditions, ascendency can be inferred setting each inflow and outflow probabilities as mutually determined (when the output a_i coincides, exclusively, with the input b_j)

$$p(a_i, b_j) = p(a_i) = p(b_j) \tag{2.43}$$

and the consequent development capacity is

$$C = K \sum_{i=0}^{S+2} p(a_i)\ log \frac{p(a_i)}{p(a_i)p(a_i)} = -K \sum_{i=0}^{S+2} p(a_i)\ log\ p(a_i) \tag{2.44}$$

$$C = K \sum_{j=0}^{S+2} p(b_j)\ log \frac{p(b_j)}{p(b_j)p(b_j)} = -K \sum_{j=0}^{S+2} p(b_j)\ log\ p(b_j) \tag{2.45}$$

or, in terms of energy (or matter) transfers (with $K = TST$)

$$C = -\sum_{i=0}^{S+2} \sum_{j=0}^{S+2} t^*_{ij}\ log \frac{t^*_{ij}}{TST} \tag{2.46}$$

The more articulated topology (Figure 2.8d), with receiving node always determined by the knowledge of the donor compartment, implies that $C = A$. In general

$$C \geq A \geq 0 \tag{2.47}$$

System overhead (Φ) measures the degree of freedom in flow organization preserved by an ecosystem. It is estimated from the not negative difference between development capacity and ascendency.

$$\Phi = C - A \tag{2.48}$$

Ascendency and system overhead are usually scaled with development capacity to define them as percentage of the theoretical upper bound on organization.

$$A(\%) = \frac{100 \cdot A}{C} \tag{2.49}$$

$$\Phi(\%) = \frac{100 \cdot \Phi}{C} \tag{2.50}$$

System overhead and its constitutive terms

Path multiplicity and a low level of flow organization, giving rise to overhead (Φ), can be interpreted as system inefficiency in processing material and energy but, in case of stress and perturbations, they represent an advantage in terms of system adaptability to new threats. The system overhead can be divided into four separate contributions, each related to a certain form of multiplicity of pathways: input from outside (overhead on imports, Φ_I), exports to other systems (overhead on exports, Φ_E), respirations (dissipative overhead, Φ_D) and internal transfers (redundancy, Φ_R).

$$\Phi = \Phi_I + \Phi_E + \Phi_D + \Phi_R \tag{2.51}$$

$$\Phi_I = -\sum_{j=1}^{S} t^*_{0j}\; log\left[\frac{t^{*2}_{0j}}{\sum_{r=0}^{S} t^*_{rj} \sum_{v=1}^{S} t^*_{0v}}\right] \tag{2.52}$$

$$\Phi_E = -\sum_{i=1}^{S} t^*_{i,S+1}\; log\left[\frac{t^{*2}_{i,S+1}}{\sum_{r=1}^{S} t^*_{r,S+1} \sum_{v=1}^{S+2} t^*_{iv}}\right] \tag{2.53}$$

$$\Phi_D = -\sum_{i=1}^{S} t^*_{i,S+2}\; log\left[\frac{t^{*2}_{i,S+2}}{\sum_{r=1}^{S} t^*_{r,S+2} \sum_{v=1}^{S+2} t^*_{iv}}\right] \tag{2.54}$$

$$\Phi_R = -\sum_{i=1}^{S}\sum_{j=1}^{S} t^*_{i,S+2}\; log\left[\frac{t^{*2}_{ij}}{\sum_{r=0}^{S} t^*_{rj} \sum_{v=1}^{S+2} t^*_{iv}}\right] \tag{2.55}$$

where t_{ij} stands for a transfer from compartment i to j; t_{0j} depicts imports to j; $t_{i,S+1}$ and $t_{i,S+2}$ denote, respectively, export and respiration flows from node i.

Overhead on inputs summarizes the fraction of "flow inefficiency" related to number and intensities of flows coming from the outside environment. When the number of inputs increases and becomes more evenly distributed, overhead on imports increases, signifying a higher inefficiency in getting usable matter beyond the system boundaries. Nevertheless, the maintenance of an adequate portion of overhead on imports become essential for the system to survive. In fact, a system depending only on one input would be extremely efficient (in this case overhead on imports is minimized and equal to 0, regardless of the magnitude of the flow) but too fragile, showing risks of catastrophic extinctions in case of collapse of the external source. Overhead on exports quantifies the multiplicity of pathways that medium takes to exit the system in a usable form. Like the overhead on inputs, it ranges from a minimum of 0, when all the matter (or energy) leaving the system is concentrated on a single node, to a maximum value in case of exports evenly distributed among all the compartments. When the topology of exports and their relative importance are assigned, this overhead component tends to increase with higher amount of matter exported. The dissipative overhead is related to the fraction of medium that is modified by internal processes (i.e. respiration in ecosystems) and exiting the system in an unusable form (flows that do not connect boxes). It increases with dissipation intensity and because of thermodynamic and ecological constraints its value must be greater than 0. The fourth component of overhead is related to the redundancy of pathways within the system. It is a contribution to disorder (inefficiency - disorganization) because sending medium over diverse routes costs more in terms of dissipation than channeling it over few efficient pathways; nevertheless it becomes absolutely essential to system survival whenever an unexpected perturbation occurs. Under these circumstances, redundancy reflects "strength in reserve" from which the system can draw to adapt to the new conditions. The lower limit for redundancy is 0 when no uncertainty is preserved by internal flow structure (this is the unlikely case of a linear chain - see Figure 2.8d), while the value of its upper bound depends on TST and is associated to a completely connected topology (maximum of flow uncertainty - see Figure 2.8a).

Table 2.2 summarizes, for 14 real ecosystems, the values of development capacity, ascendency and overhead, with this latter separated into its four constitutive components. Values are also given as percentage of the development capacity.

Additionally, one can also compute an internal development capacity (IC), considering only intercompartmental exchanges. This form finds its counterparts in other internal indices such as internal ascendency (IA) and internal redundancy (IR). Final results will be measured as percentage of the maximum upper bound (internal capacity - see Table 2.3).

	C	**A**	$\rho_{\mathbf{I}}$	$\rho_{\mathbf{E}}$	$\rho_{\mathbf{D}}$	$\rho_{\mathbf{R}}$
Charca de Masp.	39886000	16871000	2755700	906760	5538700	13814000
		42.30%	**6.91%**	**2.27%**	**13.89%**	**34.63%**
Chesapeake Bay Meso.	19655000	8593800	1702300	79705	3565200	5714500
		43.72%	**8.66%**	**0.41%**	**18.14%**	**29.07%**
Crystal River C. (c)	70712	28340	3205	6193	18408	14566
		40.08%	**4.53%**	**8.76%**	**26.03%**	**20.60%**
Crystal River C. (dT)	56315	22434	2588	3892	15030	12372
		39.84%	**4.60%**	**6.91%**	**26.69%**	**21.97%**
Everglades Gram. (ws)	79572	38643	11391	675	10181	18682
		48.56%	**14.32%**	**0.85%**	**12.79%**	**23.48%**
Florida Bay (ws)	18540	7004	2064	53	2629	6791
		37.78%	**11.13%**	**0.29%**	**14.18%**	**36.63%**
L. Chesapeake Bay	7713700	2966500	633140	81527	1271200	2761400
		38.46%	**8.21%**	**1.06%**	**16.48%**	**35.80%**
M. Chesapeake Bay	9328300	3872600	634340	37609	1548700	3235000
		41.51%	**6.80%**	**0.40%**	**16.60%**	**34.68%**
U. Chesapeake Bay	4583700	1822300	387190	15984	791020	1567200
		39.76%	**8.45%**	**0.35%**	**17.26%**	**34.19%**
St. Marks River	11264	3726	1488	353	2267	3432
		33.08%	**13.21%**	**3.13%**	**20.12%**	**30.47%**
Lake Michigan	140690	65649	10409	1814	12013	50805
		46.66%	**7.40%**	**1.29%**	**8.54%**	**36.11%**
Mondego Estuary	39126	16547	4799	500	6932	10347
		42.29%	**12.27%**	**1.28%**	**17.72%**	**26.45%**
Final Narragansett	20464000	7506700	586940	360110	2742100	9268300
		36.68%	**2.87%**	**1.76%**	**13.40%**	**45.29%**
Ythan Estuary	23397	8663	1845	1363	4158	7368
		37.02%	**7.89%**	**5.82%**	**17.77%**	**31.49%**

Table 2.2: 14 ecosystem networks are listed with their values of development capacity (C), ascendency (A) and overhead ($\rho_I, \rho_E, \rho_D, \rho_R$). Data are obtained from Prof. Ulanowicz database (datall.dat) and processed with NETWRK 4.2b software. Flows are measured in $mgC\ m^{-2}\ d^{-1}$ (Charca de Maspalomas, Crystal River Creek control and delta temp., St. Marks River and Lake Michigan), $mgC\ m^{-2}\ sum^{-1}$ (Chesapeake Bay Mesohaline, Lower, Middle and Upper Chesapeake Bay), $gC\ m^{-2}\ yr^{-1}$ (Everglades Graminoids - wet season, Florida Bay - wet season, and Ythan Estuary), $mgC\ m^{-2}\ yr^{-1}$ (Final Narragansett) and $gAFDW\ m^{-2}\ yr^{-1}$ (Mondego Estuary).

$$IC = -\sum_{i=1}^{S}\sum_{j=1}^{S} t^*_{ij}\ log\frac{t^*_{ij}}{TST} \tag{2.56}$$

$$IA = -\sum_{i=1}^{S}\sum_{j=1}^{S} t^*_{ij}\ log\frac{t^*_{ij}\ TST}{\sum_{r=0}^{S+2} t^*_{rj} \sum_{v=0}^{S+2} t^*_{iv}} \tag{2.57}$$

$$IR = IC - IA \tag{2.58}$$

$$IA(\%) = \frac{100 \cdot IA}{IC} \tag{2.59}$$

$$IR(\%) = \frac{100 \cdot IR}{IC} \tag{2.60}$$

	IC	**IA**	**IR**	**IA(%)**	**IR(%)**
Charca de Masp.	25147000	11333000	13814000	45.07	54.93
Chesapeake Bay Meso.	11584000	5869700	5714500	50.67	49.33
Crystal River C. (c)	26223	11657	14566	44.45	55.55
Crystal River C. (dT)	21267	8895	12372	41.83	58.17
Everglades Gram. (ws)	34090	15407	18682	45.20	54.80
Florida Bay (ws)	11291	4500	6791	39.85	60.14
L. Chesapeake Bay	4782000	2020500	2761400	42.25	57.75
M. Chesapeake Bay	5867500	2632500	3235000	44.87	55.13
U. Chesapeake Bay	2862600	1295400	1567200	45.25	54.75
St. Marks River	5507	2075	3432	37.68	62.32
Lake Michigan	71540	20735	50805	28.98	71.02
Mondego Estuary	14285	3938	10347	27.57	72.43
Final Narragansett	13929000	4661100	9268300	33.46	66.54
Ythan Estuary	12805	5437	7368	42.46	57.54

Table 2.3: Internal capacity (IC), internal ascendency (IA) and internal redundancy (IR) for the 14 ecosystems extracted from Prof. Ulanowicz database. The last two columns show the percentage of internal ascendency and internal redundancy respect to internal capacity.

Internal capacity aims to define the upper limit to intercompartmental flow organization. While internal redundancy (IR) and redundancy (Φ_R) coincides, we get a further detail estimating internal ascendency (IA), that is the fraction of rigidly linked flows between system nodes respect to the whole ascendency (A).

3

Trophic positions

3.1 Effective trophic position in ecology

[1]The trophic-level ideal of a simple linear chain of energy passages had great appeal as an easy and intuitive description of complex energy-based ecosystem processes. Further, this approach has inspired several applications in ecology and management such as the cascade trophic interaction theory (Carpenter et al., 1986; Carpenter and Kitchell, 1993) and the associated biomanipulation idea (Gophen, 1990; Shapiro, 1990). On the other hand, the structural intricacy (e.g. richness and topology of connections between species) of food webs gives rise to a vast array of functional behaviours that do not easily accommodate into the framework of the "green world" theories (Polis and Strong, 1996). Donor-controlled diffuse omnivory, for example, shunts the flow of matter and energy away from adjacent trophic components thus challenging the idea that populations aggregate into discrete homogeneous trophic levels each of which receives energy solely from its adjacent level nearer the ultimate source of energy (the outside or abiotic environment) and passes it to the next (Vadas Jr., 1990; Winemiller, 1990; Polis, 1991, 1994). The *"...hawk that feeds at five trophic levels..."* (Cousins, 1985, 1987) embarrasses less the ecologists now that the trophic-dynamic description of the ecosystem needs not exclude the reticulate connections between the diversity of consumers and resources. Mapping energy movements according to feeding relation in complex ecological communities

---

[1] Published Chapter: Scotti, M., Allesina, S., Bondavalli, C., Bodini, A., Abarca-Arenas, L. G., 2006. Effective trophic positions in ecological acyclic networks. Ecological Modelling, 198(3-4):495-505.

yields to ecological flow networks (Ulanowicz, 1986), from which one can appreciate the continuum of trophic positions of the species in the ecosystem. Matrix manipulation on flow coefficients allows one to calculate trophic positions: while some species do behave as obligate autotrophs or herbivores, many heterotrophs assume fractional trophic position, as the result of their feeding at multiple levels (Christian and Luczkovich, 1999). This scenario remains problematic to many, essentially because the recognition of trophic levels implies the acceptance of the tropho-dynamic viewpoint of ecosystem organization (Cousins, 1987; Oksanen, 1991). The more realistic notion of trophic position or trophic role has gained ground among ecologists because of its potential both for conceptual developments and practical applications (Burns, 1989; Pauly et al., 1998; Luczkovic et al., 2003). In particular, investigating trophic position allows insights into trophic transfer efficiencies (Christensen and Pauly, 1993), the overall energy budget in ecosystems (Wulff and Ulanowicz, 1989; Burns et al., 1991; Gaedke and Straile, 1997) and ecosystem response to stress (Ulanowicz, 1996; Bondavalli et al., 2006). The emphasis in all of these studies has been at the ecosystem level but effective trophic position can be used also to infer controlling factors in ecological communities (Christian and Luczkovich, 1999), showing the potential of the concept for community-level investigations. At present, the bulk of ecological trophic analysis conducted in the framework of ecosystem network analysis (Ulanowicz and Kay, 1991; Christensen and Pauly, 1993; Christian and Luczkovich, 1999) makes use of the canonical trophic aggregation (CTA; Ulanowicz and Kemp, 1979; Ulanowicz, 1995) a matrix-based algorithm that defines trophic positions as the weighted average distance of the compartment from the ultimate source of energy (Imports - outside environment). When analyzed in detail, however, we found that this method can give rise to some inconsistencies. In particular, we noted that: (a) there are ambiguities for the role of migratory imports (inflows to the system coming from another ecosystem); (b) there is scale dependency (e.g. a species trophic position varies when one considers just a subset of the original network); (c) the computation can lead to ecologically unrealistic values for trophic position when performed on networks that use currencies others than carbon (e.g. phosphorous and nitrogen), or from multiple non-living nodes (nutrient pools, detritus, etc.). In this work we present a natural extension of the original formulation of the CTA algorithm (Ulanowicz and Kemp, 1979; Ulanowicz, 1995) for computing trophic positions, showing how all the inconsistencies vanish after its implementation. In what follows, we briefly describe the method proposed by Ulanowicz and Kemp (1979) (detailed information can be found in the Appendix TPs) and the possible drawbacks. Finally, we explain in detail how this procedure can be generalized, and contrast the two approaches.

3.2 Canonical trophic aggregation

Ecosystem network analysis (Ulanowicz, 1986; Baird and Ulanowicz, 1989; Fath and Patten, 1999; Christensen and Pauly, 1992) is a technique that depicts ecosystems as composed of compartments (that represent species or aggregates of species, nutrient pools, etc.) exchanging flows (that can stand for energy or matter nutrients) with each others. One can discriminate between several types of fluxes: inter-compartmental fluxes denote internal exchanges (t_{ij} that stands for a flux from compartment i to compartment j), while exchanges with the external world can be divided into imports from the outside (z_j indicates an external input flow to compartment j), exports to the outside (e_j indicates an outflow from compartment j), and respirations (symbolized, for the compartment j, as r_j). Such scheme is given in matrix form with three column vectors, namely (Z) - Import, (E) - Export and (R) - Respiration and the matrix $[T]$ of internal exchanges. All vectors will have S coefficients, where S is the number of compartments, and the matrix $[T]$ will have dimension $S \cdot S$. As an example, consider the network depicted in Figure 3.1.

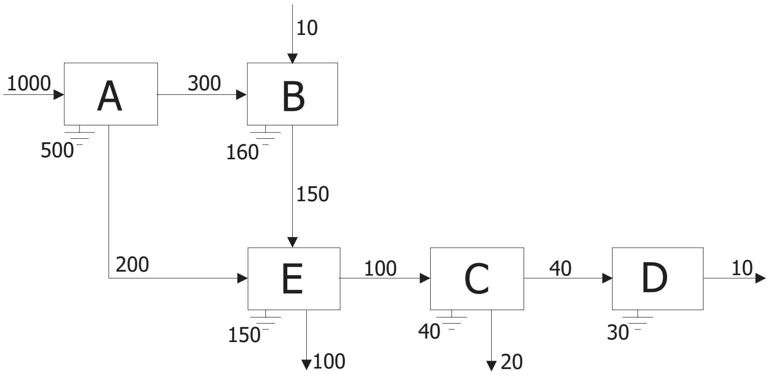


Figure 3.1: Hypothetical 5 species network showing: 5 inter-compartmental exchanges, 2 imports (on species A and B), 3 exports (on species C, D and E) and 5 respiration flows (one for each compartment). Link strength is quantified by numbers next to arrows and "ground symbols" (respirations).

The 5 compartment network of Figure 3.1 can be represented in matrix form as:

$$\mathbf{T} = \begin{bmatrix} 0 & 300 & 0 & 0 & 200 \\ 0 & 0 & 0 & 0 & 150 \\ 0 & 0 & 0 & 40 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 \end{bmatrix} \mathbf{Z} = \begin{bmatrix} 1000 \\ 10 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 0 \\ 0 \\ 20 \\ 10 \\ 100 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 500 \\ 160 \\ 40 \\ 30 \\ 150 \end{bmatrix}$$

Effective trophic position is defined as the weighted average length of all the pathways that originate from outside the system and reach a given compartment. Autotrophs will have distance (and therefore TP - trophic position) equal to 1, herbivores 2 and so on. Usually TPs are fractionary: for a species that bases half of its diet on primary producers and the other half on herbivores, the TP will be $2 \cdot 0.5 + 3 \cdot 0.5 = 2.5$ (one pathway of length 2: outside → primary producers and one of length 3: outside → primary producers → herbivores). Organisms feeding on this latter species would have $TP = 3.5$ and so forth. If one considers the network in Figure 3.1, it is clear that species A receives flows just from the outside, having therefore $TP = 1$. Computing the other TPs is more difficult, as there are multiple pathways connecting the external environment to each compartment. After some computation, we see that species B is almost herbivore ($TP = 1.968$), as the energy (or matter) it receives from the outside (10) is much smaller than the amount it gets from species A (300). So it acts as a primary producer by a 3.22% (10/(10+300)) and for the remaining 96.78% as a primary consumer. Computation yields to $0.0322 \cdot 1 + 0.9678 \cdot 2 = 1.968$. Similarly, one can compute the trophic positions of all species in the network. When the network under examination comprises non-living compartments, such as detritus and nutrient pools, the classical analysis considers only the subsystem formed by the ensemble of living compartments. Non-living nodes are then assigned to trophic level 1, equiparating their outflows as imports to the living subsystem. Even though this is the most common procedure (Cousins, 1985; Baird and Ulanowicz, 1989), other studies calculate trophic positions of non-living compartments as the number of transfers required to reach the given node starting from the outside (imports) (Burns et al., 1991; Higashi et al., 1989, 1991, 1992; Whipple and Patten, 1993; Whipple, 1998). Cycles embedded in the network would therefore imply an infinity of pathways connecting any two nodes in the same strongly connected component (Allesina et al., 2005a), making the actual trophic level computation more complicated. The vast majority of cycles, however, involves non-living compartments as an intermediate step (Pimm, 1982; May, 1983; Ulanowicz, 1983, 1995): removing the non-living nodes would therefore leave just a few residual cycles. Nevertheless, recent studies highlight the presence of cycles involving only living compartments (Dunne et al., 2002a, 2004; Williams and Martinez, 2000). According to these evidences, while we leave to a more detailed study the problem of assigning trophic positions in the presence of cycles, in what follows we discuss the question in the framework of CTA by analyzing simple networks, both hypothetical and extracted from real ecosystems, in which living compartments do not form cycles.

3.2.1 Drawbacks of CTA

Having sketched the basic idea behind CTA, we can discuss some general drawbacks:

Migratory Imports

In CTA, as it was initially conceived, all imports, corresponding to inputs from outside the system, are assigned to virtual trophic position 0, thus considering them as nutrients received by primary producers. But this is not always true. When individuals of a prey species migrate from another area (e.g. think about stream - forest ecosystems) there is an import with effective trophic position different from 0. Therefore, setting these inflows to trophic level 0 will affect the final outcome of trophic aggregation and the trophic positions of the various compartments are likely to be underestimated. The principal software packages for network analysis are NETWRK (Ulanowicz and Kay, 1991), WAND (Allesina and Bondavalli, 2004) and ECOPATH (Pauly et al., 2000; Christensen et al., 2005). The first version of NETWRK set all imports to trophic level 0, while a more recent one (Ulanowicz, 2002) assigns migratory inputs to heterotrophs to the same trophic position of the receiving node, whereas imports to primary producers are set to $TP = 0$. WAND and ECOPATH use an approach similar to the older NETWRK version.

Scale dependency

Assigning external input to trophic level 0, trophic positions of network compartments are sensible to scale. If a compartment is computed to have a trophic position equal to, say, x, considering a sub-network, "external environment" boundary changes and some intercompartmental flows in the original network appear now as import flows, that will be assigned to trophic level 0. Recomputing the trophic position of the compartment yields to a value $y \leq x$. The lack of a clear procedure for handling migratory imports makes the computation of trophic positions scale-dependent.

Non-living compartment trophic level

In carbon based networks primary producers receive their requisite medium from outside the system in form of atmospheric carbon dioxide and their TP is 1. Considering different currencies, such as nitrogen or phosphorous, or also multiple non-living nodes (e.g. nutrient pool and detritus) in a carbon based-network, the calculation is not that straightforward. To include the contribution of non-living components to the budget of the other nodes in the network one could either decide to assimilate inflows coming from these nodes to imports (that is to say assigning a virtual trophic position 0 to non-living compartments), or treat those inflows, as suggested by Ulanowicz (1995), as primary production (assigning virtually detritus and nutrient pools to level 1). The latter approach is axiomatically utilized in NETWRK (Ulanowicz and Kay, 1991) and ECOPATH (Pauly et al., 2000) software, because: "*...it's sometime difficult to separate living from dead plant tissue...*"

(Kay et al., 1989). It is clear that assimilating inflows from non-living nodes as of level 1 does not always hold. In the case of nitrogen, in fact, this would lead to an unfeasible $TP = 2$ for primary producers. The same problem may arise in carbon based network with multiple non-living nodes (e.g. nutrient pool, suspended POC, sediment POC - see Table 3.4). This problem has been analyzed in detail by Gaedke and Straile (1997). They compared four different definitions for the trophic position of dead autochthonous organic material, from a logical and descriptive point of view, suggesting to allocate all the dead organic material to the "zeroth" trophic level. In what follows, we will describe a procedure able to cope with migratory imports. The outflows of non-living compartments will be considered to be imports with: (a) $TP = 0$ when non-living nodes apportion flows to, at least, a primary producer; (b) $TP = 1$ when non-living nodes act as autotrophic compartments and exclusively show outflows to heterotrophs.

3.3 Extending CTA

To deal with all the above inconsistencies we extended canonical trophic aggregation introducing multiple imports of different trophic position. The new algorithm naturally extends the one presented by Ulanowicz (1995). We define an import vector $(Z)_0$ (energy or nutrients to plants) of trophic level 0, and an import vector $(Z)_\alpha$ for any trophic position α different from 0. Also we define a $[T]_{living}$ sub-matrix that accounts only for exchanges between living compartments (species or group of species), while non-living nodes become imports aggregated in a non-living import vector $(K)_\alpha$. In the next paragraphs we present the building blocks of the algorithm and in the last paragraph of this section we assemble them into a general framework.

3.3.1 Partial feeding matrix

If we divide every non-zero coefficient in the $[T]_{living}$ matrix by the column sum plus all the imports to that column compartment we obtain a partial feeding matrix $[G]_{living}$ that specifies the fractionary diet of each compartment. The columns of $[G]_{living}$ will sum to 1 in case the compartment does not receive imports, and to less than 1 elsewhere:

$$g_{ij} = \frac{t_{ij}}{\sum_{v=1}^{living} t_{vj} + z_j + k_j} \tag{3.1}$$

where (Z) is an undifferentiated vector with z_j coefficient that sums all the external imports $(z_{\alpha,j})$, and (K) is a vector with k_j coefficient that sums all the flows from non-living nodes $(k_{\alpha,j})$.

For the network in Figure 3.1, $[G]_{living}$ becomes:

$$\mathbf{G_{living}} = \begin{bmatrix} 0 & 0.968 & 0 & 0 & 0.571 \\ 0 & 0 & 0 & 0 & 0.429 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

The powers of $[G]_{living}$ are defined as repeated product of $[G]$ by itself: $[G]^0_{living}$ stands for the identity matrix $[I]$, $[G]^1_{living} = [G]_{living}$, $[G]^2_{living} = [G]_{living} \times [G]_{living}$, and so forth. These powers account for the fraction of matter flowing from the row compartment to the column compartment in exactly x steps, being x the exponent of the $[G]_{living}$ matrix.

3.3.2 Normalized import vectors

These vectors $(N)_\alpha$ quantify what amount of the flows to a given compartment comes from external subsidies. They are expressed similarly to $[G]_{living}$ coefficients. If a compartment j bases its diet only on external inflows of trophic position α, then $n_{\alpha,j}$ will be 1. Conversely, a value of 0 means that the compartment is not directly connected to the surrounding system; if $0 < n_{\alpha,j} < 1$ the node j receives matter, or energy, both from outside and from internal exchanges. Then we may calculate normalized import vectors $(N)_\alpha$, using

$$n_{\alpha,j} = \frac{z_{\alpha,j}}{\sum_{v=1}^{living} t_{vj} + z_j + k_j} \tag{3.2}$$

In the example network there is an import to plants (node A) which can be classified as of trophic level 0 ($\alpha = 0$). Plus, an additional import to node B (omnivores) exists. For the sake of the methodological explanation this matter/energy is assumed to be at trophic position 2.825 ($\alpha = 2.825$). We get now

$$\mathbf{N_0} = \begin{bmatrix} 1.000 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{N_{2.825}} = \begin{bmatrix} 0 \\ 0.032 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

3.3.3 Normalized non-living import vectors

These vectors $(W)_\alpha$ are obtained with the same procedure used for normalized import vectors. The coefficients of normalized non-living import vectors $(W)_\alpha$ are computed as:

$$w_{\alpha,j} = \frac{k_{\alpha,j}}{\sum_{v=1}^{living} t_{vj} + z_j + k_j} \tag{3.3}$$

where the trophic position α is set to 0 when non-living compartments apportion to primary producers, and 1 elsewhere. The Figure 3.1 network does not have non-living imports, but the calculation is straightforward once one know how to compute normalized vector of living matter.

3.3.4 Trophic Positions

Two types of trophic transformation matrices (Ulanowicz, 1995) can be computed: $[A]_\alpha$, that considers imports of living matter with various trophic positions, and $[B]_\alpha$ that refers to non-living compartments. Computation can be conducted one row at a time, using the formulas

$$a_{\alpha,i} = (N_\alpha)^T \times G_{living}^{i-1} \tag{3.4}$$

$$b_{\alpha,i} = (W_\alpha)^T \times G_{living}^{i-1} \tag{3.5}$$

where i stands for the row number and the superscript ^{T} stands for transpose form. Putting the rows together we define the final form of matrices $[A]_0$ and $[A]_{2.825}$ for the example network of Figure 3.1:

$$\mathbf{A_0} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.968 & 0 & 0 & 0.571 \\ 0 & 0 & 0.571 & 0 & 0.415 \\ 0 & 0 & 0.415 & 0.571 & 0 \\ 0 & 0 & 0 & 0.415 & 0 \end{bmatrix}$$

$$\mathbf{A_{2.825}} = \begin{bmatrix} 0 & 0.032 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.014 \\ 0 & 0 & 0.014 & 0 & 0 \\ 0 & 0 & 0 & 0.014 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore there are two vectors of trophic positions $(TP)_{0,j}$ and $(TP)_{2.825,j}$ whose coefficients (e.g. trophic positions) are computed as:

$$TP_{0,j} = \sum_{i=1}^{living-1} a_{0,ij} \cdot i \tag{3.6}$$

$$TP_{2.825,j} = \sum_{i=1}^{living-1} a_{2.825,ij} \cdot (i + 2.825) \tag{3.7}$$

yielding to:

$$\mathbf{TP_0} = \begin{bmatrix} 1 \\ 1.935 \\ 3.373 \\ 4.359 \\ 2.387 \end{bmatrix} \mathbf{TP_{2.825}} = \begin{bmatrix} 0 \\ 0.123 \\ 0.081 \\ 0.095 \\ 0.067 \end{bmatrix}$$

The final vector of trophic positions results from the sum of $(TP)_0$ and $(TP)_{2.825}$:

$$\mathbf{TP_{Final}} = \begin{bmatrix} 1 \\ 2.058 \\ 3.454 \\ 4.454 \\ 2.454 \end{bmatrix}$$

This procedure can be generalized into

$$TP_{Final} = \left(\sum_{i=1}^{q} (\alpha_i \cdot \vec{1} + [\ 1 \quad 2 \quad 3 \quad 4 \quad 5\ ]) \times [M]_{\alpha,i} \right)^T \tag{3.8}$$

This relation allows computing the trophic position in the system with different types of import flows. In particular, α_i represents the scalar TP of the i^{th} import and q the total number of imports (including non-living compartments). $[M]_{\alpha_i}$ is the trophic transformation matrix associated to the input with TP α_i. In our example $q = 2$, in fact we have $\alpha_1 = 0$, and $\alpha_2 = 2.825$. The number of nodes in the system is 5, and the trophic transformation matrices are: $[M]_{\alpha_1} = [A]_0$, and $[M]_{\alpha_2} = [A]_{2.825}$. So, for the 5-species system of Figure 3.1 the compact representation becomes explicit as follows:

$$\begin{aligned} TP_{Final} &= \Big((0 \cdot \vec{1} + [\ 1 \quad 2 \quad 3 \quad 4 \quad 5\ ]) \times [A]_0 \Big)^T + \\ &+ \Big((2.825 \cdot \vec{1} + [\ 1 \quad 2 \quad 3 \quad 4 \quad 5\ ]) \times [A]_{2.825} \Big)^T \end{aligned}$$

3.4 Applying "extended" CTA

3.4.1 Crystal River network with multiple migratory imports

We now apply the "extended" CTA to the Crystal River Creek ecosystem (Homer and Kemp, unpublished ms; see also Ulanowicz, 1983; Ulanowicz, 1986), to highlight differences in trophic positions of each compartment respect to the outcomes obtained with "classical" CTA. In this example

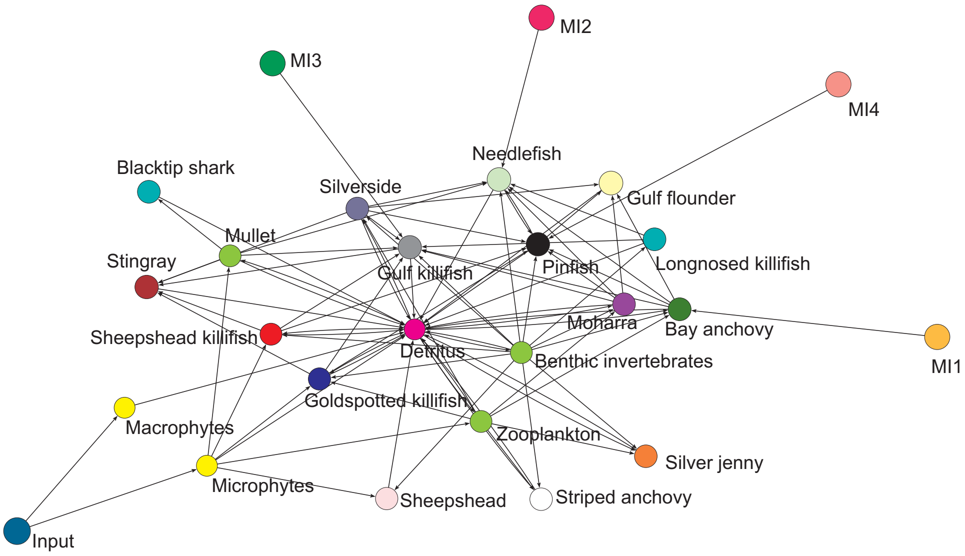


Figure 3.2: Crystal River Creek network. Nodes exhibiting different TPs are drawn varying colors (e.g. microphytes and macrophytes with $TP = 1$ are yellow while sheepshead killifish, in red, is associated to $TP = 2.109$).

(Figure 3.2) flows are measured in $mgC\ m^{-2}\ day^{-1}$, and there are 21 compartments: 20 living nodes (mainly vertebrate fishes) and 1 non-living compartment (detritus).

The network consists of 6 imports (to compartment microphytes, macrophytes, bay anchovy, needlefish, gulf killifish and pinfish), 20 exports (one for each compartment, excluding microphytes), 21 respiratory flows and 82 internal exchanges. Although much currency cycles between detritus and various living compartments, there is no need for decyclization because our trophic analysis treats non-living nodes (detritus) like exogenous inputs. Therefore we carry out "extended" CTA on 20 living compartments. The new subsystem (Figure 3.2) is comprised of: 2 imports of light or energy (to microphytes and macrophytes), 4 migratory imports (to bay anchovy, needlefish, gulf killifish and pinfish), 20 exports, 21 respiratory flows, 52 internal exchanges, 10 non-living imports from detritus to living compartments (zooplankton, benthic invertebrates, striped anchovy, bay anchovy, sheepshead killifish, goldspotted killifish, silverside, moharra, silver jenny and mullet) and 20 exports from living nodes to detritus. We consider imports to primary producers (microphytes and macrophytes) as light or energy with trophic level equal to 0, while other imports are migratory income with trophic position far from 0 because they are directed to fish. There is no logic to think of these imports to fishes as light or energy. To fishes components receiving migratory imports in Crystal River Creek network, we assign the following trophic positions, according to the Florida Bay network dataset (Ulanowicz et al., 1998):

- TP import to bay anchovy = 2.55
- TP import to needlefish = 2.9
- TP import to gulf killifish = 2.24
- TP import to pinfish = 2.07

The trophic position of migrating prey is estimated by subtracting 1 from the average trophic position of the receiving compartment. This method is what we recommend to apply. The average trophic position of a compartment should be obtained from existing datasets and literature (e.g. FishBase database available on the web site *www.fishbase.org*, Florida Bay network dataset), or should be estimated diminishing by one unit the trophic position of the receiving node without the contribution of the imported prey. The trophic level of imports from detritus, in Crystal River Creek ecosystem, is fixed to 1 because this compartment does not show flows directed to primary producers.

3.4.2 Comparison with "classical" CTA

We stress now the differences between the trophic positions of species in Crystal River Creek determined using our "extended" CTA with respect to those computed with the "classical" CTA (see Table 3.1).

Species	**CTA**	**"extended" CTA**	**Variation (%)**
Microphites	1.000	1.000	0.000
Macrophites	1.000	1.000	0.000
Zooplankton	2.000	2.000	0.000
Benthic invertebrates	2.000	2.000	0.000
Blacktip shark	3.000	3.000	0.000
Stingray	3.833	3.847	0.364
Striped anchovy	2.667	2.667	0.000
Bay anchovy	2.020	3.138	55.373
Needlefish	3.438	3.949	14.864
Sheepshead killifish	2.109	2.109	0.000
Goldspotted killifish	2.446	2.446	0.000
Gulf killifish	3.392	3.531	4.105
Longnosed killifish	3.000	3.000	0.000
Silverside	2.937	2.937	0.000
Moharra	2.859	2.859	0.000
Silver jenny	2.846	2.846	0.000
Sheepshead	2.500	2.500	0.000
Pinfish	3.225	3.706	14.930
Mullet	2.000	2.000	0.000
Gulf flounder	3.820	4.155	8.759

Table 3.1: Trophic positions for living compartments in Crystal River Creek, calculated according to "classical" CTA (CTA) and its extended version ("extended" CTA).

It is evident that microphytes and macrophytes conserve the same trophic level of 1 both in "classical" and in "extended" CTA because their inflows are exclusively external inputs that maintain trophic level 0 in both the methods. Also zooplankton, benthic invertebrates and mullet maintain their trophic level (2) since they are strictly herbivores depending on microphytes (primary producers) and detritus (placed to trophic level 1 as default in both CTA versions in this case). Longnosed killifish and blacktip shark maintained their trophic level (3), being placed in a hypothetical linear chain without other inflows (detritus → benthic invertebrates → longnosed killifish, and microphytes → mullet → blacktip shark), playing the role of primary carnivores. Migratory imports that change their trophic values from 0 to a scalar number α, in the "extended" CTA, affect directly bay anchovy, needlefish, gulf killifish and pinfish, and indirectly gulf flounder (that feeds on bay anchovy and on pinfish), and stingray (that feeds on gulf killifish). It is interesting to notice that with the exception of the bay anchovy the others compartments show both direct and indirect dependence on migratory imports: needlefish that feeds on bay anchovy and pinfish; gulf killifish and pinfish that feed on bay anchovy. The last column of the Table 3.1 presents the percentage of variation produced by considering the trophic position of imports with a scalar number α far from 0 respect to what happens in "classical" CTA. The percentage of variation is calculated as the difference between the trophic position obtained considering migratory import trophic positions different from 0 and the trophic position calculated setting all the imports to 0, divided by the latter and multiplied by 100.

3.4.3 Sensitivity analysis of migratory imports

Prey migration counts for about 49% (1.57/(1.57+0.64+1)) of bay anchovy diet, while other species rely less on external inputs: needlefish = 1.3%; gulf killifish = 2.4%; pinfish = 3.4%. Bay anchovy distributes its outflows in some intra-compartmental exchanges (to needlefish, gulf killifish, pinfish and gulf flounder), and it is more connected than needlefish, gulf killifish and pinfish. Therefore we decided to perform a sensitivity analysis on migratory imports to bay anchovy. We carry out 10,000 random samplings of migratory import amounts to bay anchovy, into a range that is ±50% its original value, obtaining the following interval: $0.755\ mgC\ m^{-2}\ day^{-1} \leq 1.57\ mgC\ m^{-2}\ day^{-1} \leq 2.355\ mgC\ m^{-2}\ day^{-1}$. The relative importance of prey migration flow with respect to total inflows into bay anchovy, in this simulation, varies between 32% (0.785/(0.785+1+0.64) and 59% (2.355/2.355+0.64+1). For each value sampled, considered as a possible amount of prey migration, we apply the "extended" CTA to calculate trophic positions of different compartments and their confidence interval at 95% and 99% (see Table 3.2).

Only compartments that are directly (bay anchovy) or indirectly (stingray, striped anchovy, needlefish, gulf killifish, silverside, silver jenny, pinfish and gulf flounder) affected by the prey migration to bay anchovy, show a con-

Species			**95%**			**99%**
Microphytes	1.000	±	0.000e+00	1.000	±	0.000e+00
Macrophytes	1.000	±	0.000e+00	1.000	±	0.000e+00
Zooplankton	2.000	±	0.000e+00	2.000	±	0.000e+00
Benthic invertebrates	2.000	±	0.000e+00	2.000	±	0.000e+00
Blacktip shark	3.000	±	0.000e+00	3.000	±	0.000e+00
Stingray	3.844	±	3.028e-05	3.844	±	4.542e-05
Striped anchovy	2.667	±	8.882e-18	2.667	±	1.332e-17
Bay anchovy	2.708	±	3.990e-03	2.708	±	5.986e-03
Needlefish	3.767	±	1.681e-03	3.767	±	2.522e-03
Sheepshead killifish	2.109	±	0.000e+00	2.109	±	0.000e+00
Goldspotted killifish	2.446	±	0.000e+00	2.446	±	0.000e+00
Gulf killifish	3.498	±	3.028e-04	3.498	±	4.542e-04
Longnosed killifish	3.000	±	0.000e+00	3.000	±	0.000e+00
Silverside	2.937	±	8.882e-18	2.937	±	1.332e-17
Moharra	2.859	±	0.000e+00	2.859	±	0.000e+00
Silver jenny	2.846	±	8.882e-18	2.846	±	1.332e-17
Sheepshead	2.500	±	0.000e+00	2.500	±	0.000e+00
Pinfish	3.548	±	1.462e-03	3.548	±	2.193e-03
Mullet	2.000	±	0.000e+00	2.000	±	0.000e+00
Gulf flounder	4.034	±	1.124e-03	4.034	±	1.686e-03

Table 3.2: Trophic position sensitivity analysis for migratory imports to bay anchovy (Crystal River Creek ecosystem): confidence interval at 95% and 99%.

fidence interval around their mean values. Confidence intervals are bigger for bay anchovy, needlefish, gulf killifish, pinfish and gulf flounder, because these nodes are closer to prey migration than others.

3.4.4 Scale independence

Trophic positions of compartments depicted in Figure 3.2 are summarized in the column labeled *"extended" CTA* of Table 3.1. If we selected a subsystem (Figure 3.3) composed by gulf killifish and longnosed killifish, CTA would treat inflows to these nodes from benthic invertebrates, bay anchovy, sheepshead killifish, goldspotted killifish, silverside, moharra and mullet as energy fluxes with trophic level equal to 0, while our "extended" CTA allows to maintain the same trophic positions they had in the whole network.

In Table 3.3 results of "extended" CTA are compared with those of "classical" CTA. The trophic position of gulf killifish and longnosed killifish remains the same in the whole network and in the selected subsystem when using "extended" CTA.

The role of inorganic nutrients

"Extended" CTA permits to consider flows from non-living nodes of a system as inputs from outside. We may set non-living nodes trophic position to 0 when primary producers (plants, phytoplankton, etc.) receive flows

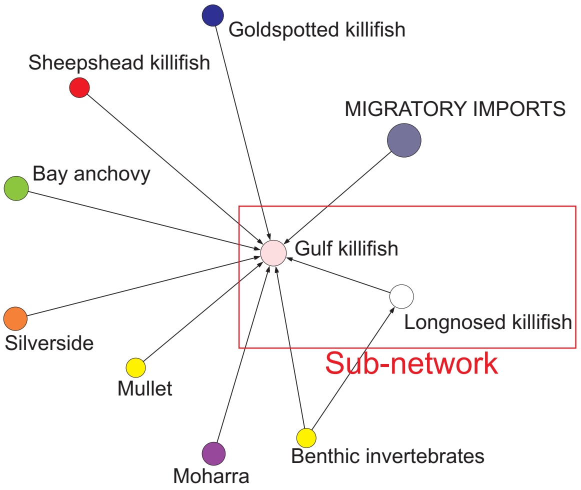


Figure 3.3: Crystal River Creek: a sub-network. Flows from benthic invertebrates, bay anchovy, sheepshead killifish, goldspotted killifish, silverside, moharra and mullet are migratory imports with TPs far from 0.

	Gulf killifish	**Longnosed killifish**
whole network "classical" CTA	3.392	3.000
whole network "extended" CTA	3.531	3.000
subsystem "classical" CTA	1.005	1.000
subsystem "extended" CTA	3.531	3.000

Table 3.3: Gulf killifish and longnosed killifish trophic positions investigated with "classical" and "extended" CTA, at different scale levels (Crystal River Creek ecosystem).

from them, or 1 elsewhere. Here follows an application to Ythan estuary ecosystem (Baird and Milne, 1981), whose scheme is depicted in Figure 3.4.

Flows are measured in $gC\ m^{-2}\ yr^{-1}$. There are 13 compartments: 10 living (benthic macrophytes, phytoplankton, benthic microflora, herbivorous birds, zooplankton, invertebrate suspension feeders, meiofauna, invertebrate deposit feeders, carnivorous birds and carnivorous fish) and 3 non-living (nutrient pool, suspended POC, and sediment POC). Non-living nodes apportion to living ones as follows:

- Nutrient pool → benthic macrophytes, phytoplankton and benthic microflora (three primary producers).
- Suspended POC → zooplankton and invertebrate suspension feeders.
- Sediment POC → meiofauna and invertebrate deposit feeders.

The network consists of 2 imports into nutrient pool and suspended POC, 10 exports (from benthic macrophytes, phytoplankton, zooplankton, invertebrate suspension feeders, meiofauna, invertebrate deposit feeders, carnivorous birds, carnivorous fish, nutrient pool and suspended POC), 12 respiration flows (one for each node, excluding nutrient pool) and 40 intercompartmental exchanges. We isolated a sub-network that includes only living compartments, using flows from non-living nodes as imports. We set to 1 the trophic level of suspended POC and sediment POC, while to 0 the trophic level of the nutrient pool flowing to primary producers (benthic macrophytes, phytoplankton and benthic microflora). The "living subsystem" that we isolated from the whole network is comprised of: 7 imports (formerly flows from non-living to living nodes), 8 exports, 10 respiratory flows, 12 internal exchanges and 17 "non-living exports" (formerly flows to non-living compartments). Results after applying CTA and "extended" CTA are summarized in Table 3.4.

"Classical" CTA yields to strange results: primary producers appear at trophic level 2 (benthic macrophytes, phytoplankton and benthic microflora), with a consequent shift in the trophic position of herbivorous birds, that now feed at trophic level 2. These incongruities arise because in the "classical" trophic analysis all the non-living nodes are positioned at trophic level 1, including those that contribute to primary producers (nutrient pool). In "extended" CTA we distinguish the trophic level of non-living nodes (0 if the compartment outflows flux to primary producers, and 1 elsewhere). Trophic positions of the Ythan estuary sub-network are ecologically consistent when computed using "extended" CTA:

- Benthic macrophytes, phytoplankton and benthic microflora (primary producers) = TP 1.

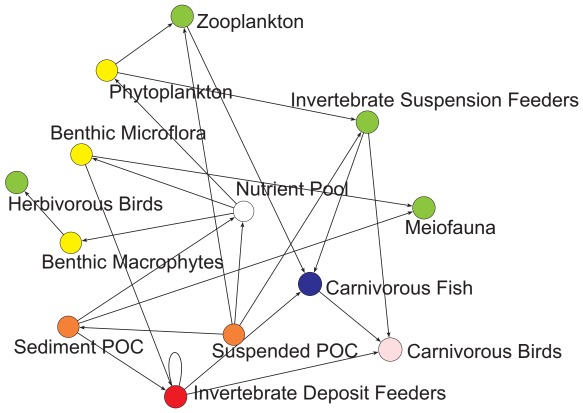


Figure 3.4: Ythan estuary ecosystem. Primary producers ($TP = 1$) are drawn in yellow and herbivores ($TP = 2$) in green. Invertebrate deposit feeders ($TP = 2.004$), carnivorous fish ($TP = 3.003$) and carnivorous birds ($TP = 3.030$), with non-integer TPs, are colored in red, blue and pink, respectively. Non-living nodes are set to $TP = 1$ (suspended POC and sediment POC, in orange) and $TP = 0$ (nutrient pool, in white).

Species	**CTA**	**"extended" CTA**	**Variation (%)**
Benthic macrophytes	2.000	1.000	50.000
Phytoplankton	2.000	1.000	50.000
Benthic microflora	2.000	1.000	50.000
Herbivorous birds	3.000	2.000	33.333
Zooplankton	2.143	2.000	6.667
Invertebrate susp. feed.	2.151	2.000	7.006
Meiofauna	2.049	2.000	2.411
Invertebrate dep. feed.	2.053	2.004	2.411
Carnivorous birds	3.125	3.030	3.047
Carnivorous fish	3.062	3.003	1.918

Table 3.4: Trophic positions in the Ythan estuary living sub-network, calculated according to "classical" CTA (CTA) and its extended version ("extended" CTA).

- Herbivorous birds = TP 2.

Applying the "extended" CTA instead of the "classical" approach, these nodes largely change their trophic position. The other compartments (zooplankton, invertebrate suspension feeders, meiofauna, invertebrate deposit feeders, carnivorous birds and carnivorous fish) present a lower variation.

3.5 Trophic position in acyclic networks

Canonical trophic aggregation is a matrix-based device that has been developed for ecosystem trophic analysis. One of its outcomes is the trophic position that species occupy in an ecosystem (Kercher and Shugart, 1975; Levine, 1983; Higashi et al., 1989; Christian and Luczkovich, 1999; Heymans et al., 2002). In the algorithm for calculus, inputs from outside the system are all allocated to the "zeroth" trophic level. The consequences of this assumption is that organisms first assimilating this energy are allocated to trophic level 1. In doing so, however, the procedure completely excludes from the analysis the question of ecosystem subsidies, that is, imports of energy and nutrients from other ecosystems (Polis and Hurd, 1996). There is a vast array of publications that highlight the importance of ecosystem subsidies (Polis et al., 1997; Polis et al., 2004 and citations therein) and others address the problem of the import of nutrients or non-living organic matter (Gaedke and Straile, 1997). Whereas non-living imports can easily be treated in the "classical" canonical trophic aggregation, the many other forms of living material exchanges between ecosystems, such as immigration, prey migration and so forth, still need to be accommodated in the CTA. Living imports and their effect on the trophic position of species has not been treated extensively in the literature. Gaedke and Straile (1997) consider this as a scale problem, stating that "*...If physical ecosystem boundaries are* *chosen in such a way that species immigrating from other systems constitute major prey items, we suggest that they maintain the trophic position they had in their previous food web, and that this number is included into the calculations...*". However, they do not treat the question in detail and restricted their analysis to dead material. In this paper, while we generalize the CTA procedure by including any kind of import (i.e. living and non-living), we also show how this approach can overcome the difficulties posed by ecosystem boundaries and related scale problems. The "extended" CTA allows, in fact, to maintain the same trophic position when sub-models of the original food web are analyzed. Often, incidentally, the trophic position of an immigrating prey is not known. In those cases it is not possible to calculate exactly the trophic position of the species that receive this incoming energy (detailed information on simulation approaches can be found in the Appendix TPs). However, in this respect a reasonable guess can be made considering the type of organism and its ecology, by taking into account that no organism can efficiently combine different ways of energy intake (Oksanen, 1991). Although the main object of investigation of this work are import flows, the "extended" CTA provides insight into the problem of trophic position as related to flows from inorganic nutrients and non-living organic matter present in the system. Assigning all non-living organic matter to the first trophic level like autotrophs (Pimm, 1988; Wulff and Ulanowicz, 1989) raises the problem of having plants feeding at trophic level 2 when the currency is inorganic nutrients. Gaedke and Straile (1997) criticized this approach and emphasized that the trophic position of the various species is largely independent from the currency (organic or inorganic) when all the non-living organic material is considered as system input. Accordingly, primary producers and bacteria would be allocated to trophic level 1; we think that this assumption underestimates the trophic role played by bacteria that can perform either as autotrophs and heterotrophs (Azam et al., 1983; Pomeroy, 1984; Bratbak, 1987; Sherr et al., 2003; Bennett et al., 1990). The "extended" CTA results show that bacteria can be allocated either at trophic level 1 or higher according to their function but this requires that non-living material be allocated: (a) to trophic level 1 if bacteria act as heterotrophs; (b) to trophic level 0 when bacteria are autotrophs. Through the "extended" CTA we have been able to solve some ambiguities related to trophic position in the presence of multiple non-living nodes connected to living compartments. Instead of considering non-living matter as part of the system, we propose to place it as import flows (see Figure 3.4). Setting the nutrient pool trophic level to 0 avoids the strange trophic position 2 that happens to characterize benthic macrophytes, phytoplankton and benthic microflora, changing it to a more reasonable first level (Table 3.4). Conversely, if suspended POC and sediment POC trophic levels are equal to 1 we obtain ecologically robust outcomes for meiofauna, zooplankton and invertebrate feeders (herbivores with trophic level 2). What comes out from suspended POC and sediment POC (Figure 3.4) substantially corresponds to what Gaedke and Straile (1997) call autochthonous dead organic matter. Although our choice of considering suspended POC and sediment POC as inflows into the system matches with their "biospheric point of view" that abolish the distinction between allochthonous and autochthonous material, still a difference remains in assigning these compartments to a trophic level. We set up these flows (to herbivores) at trophic level 1, whereas the above cited authors keep them at trophic level 0. Their approach seems beneficial for the clarity and meaningfulness of various trophic measures such as trophic position as indicative of number of assimilation events, transfer efficiency between trophic levels, homogeneity of trophic level composition. However, it seems that the same achievements could be obtained by setting up dead organic material (irrespective of its endogenous or exogenous origin) to trophic level 1 or to the respective trophic position of its sources, as likely, what matters in their context is the difficulty of treating dead organic material as part of the system. We reiterate, however, that dead organic material, as external source of energy (e.g. inorganic nutrients), should be allocated to trophic level 0 if inconsistencies involving the trophic positions of plants have been noticed, while it should be set up to 1 when it just apportions its outflows to heterotrophic compartments. Lindeman (1942) allocated dead organic material at trophic level 1, and this choice was criticized as an arbitrary convention (Cousins, 1987). However, while we show that in doing so one avoids many inconsistencies in the trophic aggregation scheme and that the entire CTA acquires internal coherence, we also point out that a distinction should be made between detritus (i.e. suspended POC and sediment POC in our example) and dead organic material. In many cases detritus includes living organisms and in this case particular attention should be devoted in separating living from non-living things before calculating the trophic position of this entity (Gaedke and Straile, 1997). The same authors draw their conclusion discussing the trophic position of autochthonous dead organic material from a logical-empirical point of view. They make evident how the various approaches used in the literature to calculate trophic position in ecosystems generate inconsistencies when dead organic material is included in the computation and show how these inconsistencies can be resolved. Although they provide a logical -step by step- reconstruction of the potential consequences of the different assumptions concerning the trophic position of the dead organic matter, they do not generalize their point of view in a rigorous procedure for calculation. Because we developed our arguments within the canonical trophic aggregation framework, which is basically manipulation of ecosystem flow matrices, our outcomes are fairly general. Despite the fact that conclusions drawn in this paper differ from Gaedke and Straile's, rather than seeing the two approaches in contrast with one another we perceive the potential for their integration. In recent times, the shift of emphasis from local stability dynamics towards a more comprehensive analysis of ecosystem persistence and nonlinear dynamics, has allowed to reconsider the importance of cycles in networks. As a consequence, the fact that we used simple acyclic ecosystem models to derive our "extended" CTA may cast doubts on the realism of our approach. We stress, however, that our aim was to resolve the inconsistencies arising in CTA. Therefore, we started from the very same set of assumptions (Ulanowicz, 1995, 2004), and we treated flows from non-living compartments as imports to living. The choice of using an acyclic model is coherent within this framework. Since cycles play a major role in ecosystems, and recent studies show a greater evidence of cycles involving living nodes (Dunne et al., 2002a, 2004; Williams and Martinez, 2000), next amendment to CTA should deal with them, replacing decyclization routine that over simplifies the networks. We have not tested the consequences of our approach on the trophic attributes of real ecosystems such as efficiency of transfer, trophic pyramids, interpretation of analysis of the overall energy flow. Applications to real ecosystems are required to perform these investigations, and they will certainly be carried out in the near future. Also a more functional CTA should explore the effect of cycling but this is a question that we inherited from the original formulation of the canonical trophic aggregation. Whether the concept of trophic position has resolved many ambiguities associated with that of trophic level, this latter idea remains as a reference point because the calculation of the trophic position is made in relation to a classification scheme that sets plants as the first level.

3.6 Appendix TPs

3.6.1 TPs and canonical trophic aggregation

In ecosystem network analysis (ENA) the effective trophic position of each compartment is calculated using canonical trophic aggregation (CTA), starting from the so-called trophic transformation matrix $[A]$ (Ulanowicz and Kemp, 1979; Ulanowicz, 1995). The method is based on the trophic transformation matrix $[A]$ and requires one to assign trophic positions to compartments according to their diet composition. Consider the hypothetical simple 5-species system depicted in Figure 3.1.The corresponding $[T]$ matrix, and (Z), (E) and (R) vectors, will be:

$$\begin{bmatrix} \textit{Species A} \\ \textit{Species B} \\ \textit{Species C} \\ \textit{Species D} \\ \textit{Species E} \end{bmatrix} \mathbf{T} = \begin{bmatrix} 0 & 300 & 0 & 0 & 200 \\ 0 & 0 & 0 & 0 & 150 \\ 0 & 0 & 0 & 40 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} 1000 \\ 10 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 0 \\ 0 \\ 20 \\ 10 \\ 100 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 500 \\ 160 \\ 40 \\ 30 \\ 150 \end{bmatrix}$$

and the matrix of diet coefficients (also called partial feeding coefficients) $[G]$ becomes:

$$\mathbf{G} = \begin{bmatrix} 0 & 0.968 & 0 & 0 & 0.571 \\ 0 & 0 & 0 & 0 & 0.429 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

The matrix of partial feeding coefficients $[G]$ and a normalized import vector (N) are the basic tools for the CTA. The normalized import vector elements (n_j), can be computed as:

$$n_j = \frac{z_j}{T_{\cdot j}} \tag{3.9}$$

where

$$T_{\cdot j} = z_j + \sum_{i=1}^{k} t_{ij} \tag{3.10}$$

and $T.j$ is an element of the inflow vector $(T.j)$, which, for each node, represents the sum of the correspondent column fluxes in $[T]$ plus the inputs from outside the system (z_j); k is the number of compartments. Assuming the theoretical trophic level for imports to be 0, we can distinguish between three main cases: (a) when n_j is 0 the compartment j does not receive input from the outside; (b) if n_j is equal to 1 the node j depends only on imports; (c) with $0 < n_j < 1$ the node j receives inputs both from outside the system and from internal exchanges. For our 5-species system we obtain:

$$\mathbf{N} = \begin{bmatrix} 1.000 \\ 0.032 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The next step is to compute the trophic transformation matrix $[A]$ using a row by row procedure. Every row is computed as

$$a_i = N^T \times G^{i-1} \tag{3.11}$$

where ^{T} superscript stands for transpose, and powers of $[G]$ are defined as repeated product of $[G]$: $[G]^0$ stands for identity matrix $[I]$, $[G]^1 = [G]$, $[G]^2 = [G] \times [G]$, and so on. The matrix $[A]$ in our example is:

$$\mathbf{A} = \begin{bmatrix} 1 & 0.032 & 0 & 0 & 0 \\ 0 & 0.968 & 0 & 0 & 0.585 \\ 0 & 0 & 0.585 & 0 & 0.415 \\ 0 & 0 & 0.415 & 0.585 & 0 \\ 0 & 0 & 0 & 0.415 & 0 \end{bmatrix}$$

The trophic behaviour of each jth species can be inferred by the elements of the j column of $[A]$, where the trophic position of the jth species (TP_j) is:

$$TP_j = \sum_{i=1}^{k-1} a_{ij} \cdot i \tag{3.12}$$

Accordingly, trophic positions in the model of Figure 3.1 are

- *Species A*: $TP_1 = 1 \cdot 1 = 1$
- *Species B*: $TP_2 = 0.032 \cdot 1 + 0.968 \cdot 2 = 1.968$
- *Species C*: $TP_3 = 0.585 \cdot 3 + 0.415 \cdot 4 = 3.415$
- *Species D*: $TP_4 = 0.585 \cdot 4 + 0.415 \cdot 5 = 4.415$

- *Species E*: $TP_5 = 0.585 \cdot 2 + 0.415 \cdot 3 = 2.415$

Transposing the row vector (TP) we get:

$$\mathbf{TP} = \begin{bmatrix} 1 \\ 1.968 \\ 3.415 \\ 4.415 \\ 2.415 \end{bmatrix}$$

In the presence of cycles, CTA becomes difficult because the powers of $[G]$ form an infinite sequence, and it is not clear how and where to truncate the number of rows in $[A]$. Ulanowicz (1995) proposes to resolve the system into two constitutive networks, one containing cyclic paths and another with only once-through flows, using only the latter for CTA. If cycles constitute a large amount of total system activity, information is lost during network decyclization. This inaccuracy may result in distorted values of trophic positions. Generally detritus and nutrient pools appear in ecological flow networks as non-living nodes. The distinction between living and non-living components allows to identify $[T]_{living}$ and $[G]_{living}$ sub-matrices that account only for exchanges between living compartments. CTA considers that in living sub-networks cycles are scarce, as the bulk of cycling involves non-living components (Pimm, 1982; May, 1983; Ulanowicz, 1995). The distortion imposed by the decyclization algorithm would be, in this case, somehow marginal. However, recent studies (Dunne et al., 2002a, 2004; Williams and Martinez, 2000) highlight the presence of simple cycles between living compartments. In this perspective, do not treat cycles excluding non-living nodes, and applying decyclization algorithms (Ulanowicz, 1983), seems a weak point in CTA procedure that should be implemented in future studies. In the present work we focussed on a general framework for CTA, and therefore we treat non-living nodes as imports to living subsystem. Ulanowicz (1995) assimilates inflows coming from these nodes to primary production assigning virtually all inputs to trophic level 0 and, consequently, detritus and nutrient pools to level 1 (Kay et al., 1989). Other studies treat inflows from non-living nodes as imports, with virtual trophic position 0 (Gaedke and Straile, 1997), or include non-living compartments in the network unfolding procedure, to obtain the corresponding trophic macrochains. In the latter case, trophic positions of non-living (and living) nodes may count both trophic and non-trophic flows (Higashi et al., 1989, 1991, 1992), or exclusively transfers among compartments with energy assimilation (Burns et al., 1991; Whipple and Patten, 1993; Whipple, 1998).

3.6.2 Limitations of the CTA

Migratory imports

As already explained in CTA all the imports are assigned to trophic level 0, as they were assimilated to energy or nutrients received by primary producers, but this is not always the case. When preys migrate from another area to the system, the inflows should be considered with effective trophic level greater than 0. Therefore, to settle all the inflows from outside the system to trophic level 0, will affect the results.

Consider the simple 2-species system in Figure 3.5,

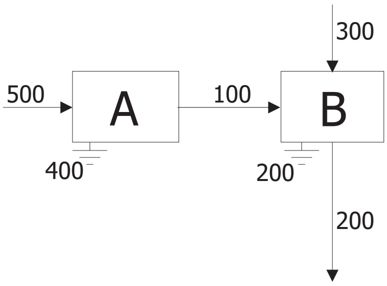


Figure 3.5: A simple 2 node network.

where the import to species B (henceforth called B) is made up of individuals that feed at a trophic level equal to, say, 3. We might simply derive its trophic position (TP_B):

$$TP_B = \frac{100}{400} \cdot 2 + \frac{300}{400} \cdot 4 = 0.25 \cdot 2 + 0.75 \cdot 4 = 3.5 \tag{3.13}$$

Using CTA one obtains:

$$\begin{bmatrix} A \\ B \end{bmatrix} \mathbf{T} = \begin{bmatrix} 0 & 100 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} 500 \\ 300 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 0 \\ 200 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 400 \\ 200 \end{bmatrix}$$

$$\mathbf{N} = \begin{bmatrix} 1 \\ 0.75 \end{bmatrix} \mathbf{G} = \begin{bmatrix} 0 & 0.25 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0.75 \\ 0 & 0.25 \end{bmatrix} \mathbf{TP} = \begin{bmatrix} 1 \\ 1.25 \end{bmatrix}$$

Calculation yields trophic positions 1.25 for B, underestimating its trophic position. This problem arises because both imports are assigned to trophic level 0.

Scale dependency

Inconsistencies emerge also when computing the effective trophic position of compartments in a portion of a known network. In this situation all flows to the nodes comprised in the selected subsystem, coming from other compartments not included, become imports. Even if these latter compartments possess a defined and known trophic position, flows coming out from them are assimilated to external input and allocated to trophic level 0. Consider the example given in Figure 3.1, trophic positions of the compartments are

$$\begin{bmatrix} \textit{Species A} \\ \textit{Species B} \\ \textit{Species C} \\ \textit{Species D} \\ \textit{Species E} \end{bmatrix} \mathbf{TP} = \begin{bmatrix} 1 \\ 1.968 \\ 3.415 \\ 4.415 \\ 2.415 \end{bmatrix}$$

Focussing on the sub-network made up of C and D (Figure 3.6), the 100 carbon units formerly flowing from E to C become now external input to C. We then apply CTA to this subsystem and the outcomes are:

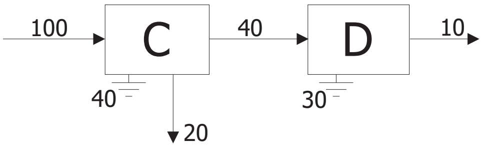


Figure 3.6: Subsystem of the Figure 3.1 network.

$$\begin{bmatrix} C \\ D \end{bmatrix} \mathbf{T} = \begin{bmatrix} 0 & 40 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} 100 \\ 0 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 20 \\ 10 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 40 \\ 30 \end{bmatrix}$$

$$\mathbf{N} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{G} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{TP} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

C and D are assigned, respectively, to trophic level 1 and 2, whereas, when considering the whole system these numbers were larger; results are summarized in Table 3.5.

Node	TP_{whole}	TP_{sub}
C	3.415	1
D	4.415	2

Table 3.5: Trophic positions calculated, using "classical" CTA, in the network of Figure 3.1 (TP_{whole}), and in the subsystem of Figure 3.6 (TP_{sub}).

Non-living compartment trophic levels

Trophic level assigned to non-living nodes may affect the trophic position of living compartments. Ulanowicz (1995) places non-living components to trophic level 1; this choice does not influence the trophic position of living components when the inorganic nutrients are not included in the network. This is often the case when currency is carbon. In doing so, plants, that receive inputs only from outside the system, remain at trophic level 1; non-living nodes, usually detritus, provide matter to detritivores and bacteria that, accordingly, occupy trophic level 2. In a nitrogen (or phosphorus) based network, non-living compartments, such as inorganic nutrients, are included, providing inputs to primary producers. In this case setting non-living trophic level to 1 increases plant trophic level to 2, shifting the whole trophic positions along grazing chain. Consider the whole simple linear nitrogen based system in Figure 3.7

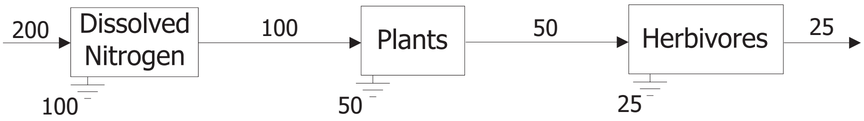


Figure 3.7: Hypothetical linear chain starting with a non-living node (dissolved nitrogen) preceding a primary producer (plants).

$[T]$ matrix, (Z), (E) and (R) vectors, are:

$$\begin{bmatrix} Plants \\ Herbivores \\ Dissolved\ Nitrogen \end{bmatrix} \mathbf{T} = \begin{bmatrix} 0 & 50 & 0 \\ 0 & 0 & 0 \\ 100 & 0 & 0 \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} 0 \\ 0 \\ 200 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 0 \\ 25 \\ 0 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 50 \\ 25 \\ 100 \end{bmatrix}$$

Discriminating between living and non-living compartments we set the trophic level of non-living ones (Dissolved Nitrogen in our example) to 1 (Ulanowicz, 1995). We define vector $(N)_{living}$, matrix $[G]_{living}$ and the associated matrix $[A]_{living}$, which gives rise to (TP):

$$\mathbf{N_{living}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \mathbf{G_{living}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{A_{living}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{TP} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This illogical situation derives from imports that are exclusively located on non-living nodes. But, by definition, CTA sets the trophic position of non-living compartments to 1. As a consequence, the correct trophic transformation matrix becomes:

$$\mathbf{A_{living}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since non-living node (Dissolved Nitrogen) shows a trophic position equal to 1, the one of the following compartment should be 2. It results that the suitable formula to derive trophic position of each jth species is:

$$TP_j = \sum_{i=1}^{k-1} a_{ij} \cdot (i+1) \tag{3.14}$$

where TP_j is the trophic position of jth species, k is the number of living compartments and $[A]$ is the trophic transformation matrix. In equation (3.14) the second term of the product (i) is increased by 1 unit (i+1) because the linear chain starts with a non-living node ($TP = 1$). Applying equation (3.14) to the network 3.7 we calculated the vector of trophic positions:

$$\mathbf{TP} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Obtaining an ecologically unlikely result: Plants are apportioned to trophic level 2 and Herbivores are allocated at trophic level 3.

3.6.3 Solving incongruities by “extended” CTA

Migratory imports and sensitivity analysis

The “extended” CTA allows to consider the effective trophic position of various imports, taking into account the difference between input of energy or inorganic matter (trophic level 0) and prey migration (with an effective trophic position far from 0). Nevertheless, it is not likely to know the trophic level of the immigrant prey; to work out this problem we presume a range of trophic position in which this species is likely to feed and we carry out 10,000 random samplings into this interval. For each sampled value, considered as a possible trophic position for the import, we apply the “extended” CTA, to calculate trophic positions of the different compartments of the system depicted in Figure 3.1. As an example, consider an import flow of $10\ gC\ m^{-2}\ yr^{-1}$ to compartment B, assuming this inflow having a trophic position between 2 and 3. Table 3.6 shows that the effect of this import impacts the system in two ways: it has a direct consequence on the receiving compartment (species B), and indirect effects on the other compartments, that depend directly (species E) and indirectly (species C and D) on B.

	A	**B**	**C**	**D**	**E**
Min.	1	2.032	3.442	4.442	2.442
1st Qu.	1	2.041	3.446	4.446	2.446
Median	1	2.048	3.449	4.449	2.449
Mean	1	2.049	3.449	4.449	2.449
3rd Qu.	1	2.057	3.453	4.453	2.453
Max.	1	2.065	3.456	4.456	2.456
μ	1	2.049	3.449	4.449	2.449
sd	0	0.009	0.004	0.004	0.004
CV	0	0.451	0.115	0.089	0.162

Table 3.6: Trophic position sensitivity analysis in network 3.1, related to migratory imports to B, where μ stands for average values, sd is the standard deviation and CV is the coefficient of variation.

Compartment A, which receives an import with trophic level equal to 0, is not affected ($\mu = 1$, $sd = 0$ and $CV = 0$). Standard deviation is bigger for B (0.009) than for the other compartments (0.004). This difference depends on the relative magnitude of the incoming flow in relation with the compartment throughput (10/310). Then this effect is passed along the linear chain $B \rightarrow E \rightarrow C \rightarrow D$, but its relative contribution is further modified by the presence of an extra inflow to E coming from A. Dividing each flow by the total throughput (350) we define the effect of A and B on the trophic position of E. The weight of $B \rightarrow E$ is equal to 0.429 ($w_{BE} = 150/350$), while $A \rightarrow E$ is 0.571 ($w_{AE} = 200/350$). E standard deviation then becomes (sd_E):

$$sd_E = sd_B \cdot w_{BE} + sd_A \cdot w_{AE} \tag{3.15}$$

$$sd_E = 0.009 \cdot 0.429 + 0 \cdot 0.571 = 0.004 \tag{3.16}$$

where sd_E, sd_B and sd_A are, respectively, trophic level standard deviations of E, B and A. Species C and D totally depend on E for their requisite medium: the former directly, and the latter indirectly; their standard deviations are

$$sd_C = sd_E \cdot w_{EC} = 0.004 \cdot 1 = 0.004 \tag{3.17}$$

$$sd_D = sd_C \cdot w_{CD} = 0.004 \cdot 1 = 0.004 \tag{3.18}$$

Moving along the chain $B \rightarrow E \rightarrow C \rightarrow D$ the effect of prey migration on trophic positions decreases and this is highlighted by the reduction of the coefficients of variation (CV). To complete our analysis we vary the magnitude of the flow. For each inflow we repeat 10,000 random samplings as explained before. Here we discuss two extreme situations: (a) all the excess import to B (with respect to the starting condition of 10 $gC\ m^{-2}\ yr^{-1}$) is dissipated as respiration by B (option called *dissipative simulation*); (b) all the excess import is passed along the pathway $B \rightarrow E \rightarrow C \rightarrow D$ until it goes out as export from D (called *maximum efficient option*). Outcomes are depicted in Tables 3.7 and 3.8.

	A	**B**	**C**	**D**	**E**
10 gC m^{-2} yr^{-1}					
mean	1	2.049	3.449	4.449	2.449
sd	0	0.009	0.004	0.004	0.004
CV	0	0.451	0.115	0.089	0.162
20 gC m^{-2} yr^{-1}					
mean	1	2.094	3.469	4.469	2.469
sd	0	0.018	0.008	0.008	0.008
CV	0	0.863	0.223	0.173	0.314
40 gC m^{-2} yr^{-1}					
mean	1	2.176	3.504	4.504	2.504
sd	0	0.034	0.015	0.015	0.015
CV	0	1.572	0.418	0.325	0.585
100 gC m^{-2} yr^{-1}					
mean	1	2.376	3.590	4.590	2.590
sd	0	0.072	0.031	0.031	0.031
CV	0	3.043	0.863	0.675	1.196
200 gC m^{-2} yr^{-1}					
mean	1	2.600	3.686	4.686	2.686
sd	0	0.115	0.049	0.049	0.049
CV	0	4.442	1.343	1.056	1.843
280 gC m^{-2} yr^{-1}					
mean	1	2.724	3.739	4.739	2.739
sd	0	0.139	0.060	0.060	0.060
CV	0	5.113	1.597	1.260	2.180
350 gC m^{-2} yr^{-1}					
mean	1	2.810	3.776	4.776	2.776
sd	0	0.156	0.067	0.067	0.067
CV	0	5.538	1.766	1.396	2.403

Table 3.7: Trophic position sensitivity analysis, in network 3.1, related to migratory imports to B. All inputs dissipated as respirations in compartment B.

Considering species B, all coefficients do coincide in the two cases, irrespective of the magnitude of the import. Compartment B, as entry point

	A	B	C	D	E
10 gC m^{-2} yr^{-1}					
mean	1	2.048	3.449	4.449	2.449
sd	0	0.009	0.004	0.004	0.004
CV	0	0.458	0.117	0.090	0.164
20 gC m^{-2} yr^{-1}					
mean	1	2.097	3.486	4.486	2.486
sd	0	0.018	0.008	0.008	0.008
CV	0	0.864	0.231	0.179	0.323
40 gC m^{-2} yr^{-1}					
mean	1	2.176	3.557	4.557	2.557
sd	0	0.034	0.016	0.016	0.016
CV	0	1.569	0.455	0.355	0.633
100 gC m^{-2} yr^{-1}					
mean	1	2.375	3.750	4.750	2.750
sd	0	0.072	0.039	0.039	0.039
CV	0	3.037	1.049	0.828	1.430
200 gC m^{-2} yr^{-1}					
mean	1	2.599	4.007	5.007	3.007
sd	0	0.116	0.073	0.073	0.073
CV	0	4.455	1.819	1.456	2.424
280 gC m^{-2} yr^{-1}					
mean	1	2.724	4.168	5.168	3.168
sd	0	0.138	0.093	0.093	0.093
CV	0	5.064	2.242	1.808	2.950
350 gC m^{-2} yr^{-1}					
mean	1	2.809	4.285	5.285	3.285
sd	0	0.156	0.111	0.111	0.111
CV	0	5.562	2.590	2.100	3.378

Table 3.8: Trophic position sensitivity analysis, in network 3.1, related to migratory imports to B. All inputs dissipated along the chain $B \rightarrow E \rightarrow C \rightarrow D$.

of the import, is not affected by the form in which this energy leaves the compartment. For the other species, coefficients are always higher in the *maximum efficient option*. To understand this outcome it is useful to compare inflows to E, C and D with the ones resulting from the first simulation (with an import on B equal to $10gC\ m^{-2}\ yr^{-1}$). In the *dissipative simulation* the amount that reaches species E does not change because all the excess import to B is dissipated. It increases in the *maximum efficient option*. In both *dissipative simulation* and *maximum efficient option*, for every different import to B, no variations in standard deviations are observed for species C, D and E (but obviously values differ between the two simulations). On the contrary, the coefficients of variations change because mean values also vary.

Scale independence

In Figure 3.1, single compartment trophic positions are:

$$\begin{bmatrix} Species\ A \\ Species\ B \\ Species\ C \\ Species\ D \\ Species\ E \end{bmatrix} \mathbf{TP} = \begin{bmatrix} 1 \\ 1.968 \\ 3.415 \\ 4.415 \\ 2.415 \end{bmatrix}$$

If we select a subsystem made up of compartments C and D (Figure 3.6) "extended" CTA imposes that we maintain for the inflow to C (which now appears like imports from outside) the same trophic position it had in the whole network, that is 2.415 (it comes from species E). The calculation yields:

$$\begin{bmatrix} C \\ D \end{bmatrix} \mathbf{T} = \begin{bmatrix} 0 & 40 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{Z_{2.415}} = \begin{bmatrix} 100 \\ 0 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 20 \\ 10 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 40 \\ 30 \end{bmatrix}$$

$$\mathbf{N_{2.415}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{G} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{A_{2.415}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{TP_{2.415}} = \begin{bmatrix} 3.415 \\ 4.415 \end{bmatrix}$$

The vector $(TP)_{2.415}$ confirms that "extended" CTA computes trophic positions as scale independent features: C and D occupy the same trophic positions in the 5-species network and in 2-species sub-network. Table 3.9 shows trophic positions for C and D in the two cases.

Node	TP_{whole}	TP_{sub}
C	3.415	3.415
D	4.415	4.415

Table 3.9: Trophic positions calculated, using "extended" CTA, in the network of Figure 3.1 (TP_{whole}), and in the subsystem of Figure 3.6 (TP_{sub}).

Non-living compartment trophic levels

The connection between living and non-living compartments may heavily affect the trophic positions of living nodes, depending on the identity of non living compartments included in the system. "Extended" CTA overcomes this difficulty by considering non-living compartments as input source for the system, setting their trophic level to 1, in case they do not apportion outflows to primary producers, or to 0 when representing inorganic nutrients.

Considering the simple nitrogen based system of Figure 3.7 we can write the following $[T]_{living}$ matrix, $(K)_0$, (E) and (R) vectors:

$$\begin{bmatrix} Plants \\ Herbivores \end{bmatrix} \mathbf{T_{living}} = \begin{bmatrix} 0 & 50 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{K_0} = \begin{bmatrix} 100 \\ 0 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 0 \\ 25 \end{bmatrix} \mathbf{R} = \begin{bmatrix} 50 \\ 25 \end{bmatrix}$$

The normalized $(W)_0$ vector and $[G]_{living}$ matrix are

$$\mathbf{W_0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{G_{living}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

We finally calculate $[B]$ matrix and the vector $(TP)_0$:

$$\mathbf{B_0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{TP_0} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

from which plants and herbivores assume a trophic position coherent with ecological conventions.

4

Decomposing system complexity

The following is a bridge chapter introducing the topic of species contribution to flow topology and consequent characteristics. Only theoretical and methodological premises are given, without any results and conclusions.

4.1 The average mutual information

The question whether complexity affects ecosystem stability has long been central in ecology. MacArthur (1955) applied Shannon's information measure to the flows in ecosystem networks, with S species, as

$$H = -k \sum_{i=1}^{S} \sum_{j=1}^{S} \frac{t_{ij}}{T_{..}} \, log_2 \left(\frac{t_{ij}}{T_{..}} \right) \tag{4.1}$$

where H is the diversity of flows in the network, k is a scalar constant, t_{ij} denotes flows from node i to node j and $T_{..}$ is the total system throughput (TST)

$$TST = \sum_{i=1}^{S} \sum_{j=1}^{S} t_{ij} = T_{..} \tag{4.2}$$

The increasing consensus around this index stimulated its application to the more accessible stocks of biomass, shifting the discussion on diversity and stability to effects of biomass diversity. Unfortunately, May (1972) demonstrated that a higher biodiversity in linear dynamical systems was more likely to result in instability than in the reverse. Thereafter, ecologists quickly abandoned their interest towards information theory, maintaining the same prejudice even after Rutledge et al. (1976) applied a Bayesian emendation of Shannon's measure to MacArthur's index of flow diversity. Rutledge et al. (1976) employed the notion of conditional probability to decompose MacArthur's index into two complementary terms. In particular, the joint probability that an arbitrary quantum both leaves i and enters j can be estimated by the quotient $t_{ij}/T_{..}$. The conditional probability that a quantum proceeds to compartment j, when it already left i, is defined by equation (4.3); similarly, the equation (4.4) summarizes the conditional probability associated to a flow exiting the node i respect to the total input to compartment j.

$$\frac{t_{ij}}{\sum_{v=0}^{S} t_{iv}} = \frac{t_{ij}}{T_{i\cdot}} \tag{4.3}$$

$$\frac{t_{ij}}{\sum_{r=0}^{S} t_{rj}} = \frac{t_{ij}}{T_{\cdot j}} \tag{4.4}$$

The measure of total flow diversity can be amended as follows

$$H = AMI + H_c \tag{4.5}$$

where the average mutual information (AMI) quantifies the amount of diversity that is encumbered by structural constraints

$$AMI = k \sum_{i=1}^{S} \sum_{j=1}^{S} \frac{t_{ij}}{T_{..}} \, log_2 \left(\frac{t_{ij} \, T_{..}}{T_{i\cdot} \, T_{\cdot j}} \right) \tag{4.6}$$

and H_c represents the amount of "choice" (residual diversity/freedom) pertaining to both the inputs and outputs of an average node in the network.

$$H_c = -k \sum_{i=1}^{S} \sum_{j=1}^{S} \frac{t_{ij}}{T_{..}} \, log_2 \left(\frac{t_{ij}^2}{T_{i\cdot} \, T_{\cdot j}} \right) \tag{4.7}$$

Therefore, the overall complexity of the flow structure, as measured by the MacArthur's index, can be resolved into two components: (a) AMI that estimates how orderly and coherently the flows are connected; (b) H_c that gauges the disorder and freedom that is preserved.

Rutledge et al. (1976) proposed H_c as a more appropriate measure of ecosystem maturity as conceived by Odum (1969), than the MacArthur's ambiguous index, but further studies (Atlan, 1974; Ulanowicz, 1980) suggested that AMI is more indicative of the developmental status of an ecological network.

However, Ulanowicz and Wolff (1991) adopted H_c as a tool to estimate effective connectance per node in ecosystems. In particular, dividing H_c into two terms reveals more about its mathematical meaning

$$\begin{aligned}
H_c &= -\sum_{i=1}^{S}\sum_{j=1}^{S}\frac{t_{ij}}{T_{\cdot\cdot}}\, log_2\left(\frac{t_{ij}}{T_{i\cdot}}\right) - \sum_{i=1}^{S}\sum_{j=1}^{S}\frac{t_{ij}}{T_{\cdot\cdot}}\, log_2\left(\frac{t_{ij}}{T_{\cdot j}}\right) \\
&= \sum_{i=1}^{S}\frac{T_{i\cdot}}{T_{\cdot\cdot}}\left[-\sum_{j=1}^{S}\frac{t_{ij}}{T_{i\cdot}}\, log_2\left(\frac{t_{ij}}{T_{i\cdot}}\right)\right] + \sum_{j=1}^{S}\frac{T_{j\cdot}}{T_{\cdot\cdot}}\left[-\sum_{i=1}^{S}\frac{t_{ij}}{T_{\cdot j}}\, log_2\left(\frac{t_{ij}}{T_{\cdot j}}\right)\right] \\
&= \sum_{i=1}^{S}\frac{T_{i\cdot}}{T_{\cdot\cdot}}H_{i\cdot} + \sum_{i=j}^{S}\frac{T_{\cdot j}}{T_{\cdot\cdot}}H_{\cdot j} \qquad (4.8)
\end{aligned}$$

with output diversity at node i ($H_{i\cdot}$) and input diversity at node j ($H_{\cdot j}$) calculated as

$$H_{i\cdot} = -\sum_{j=1}^{S}\frac{t_{ij}}{T_{i\cdot}}\, log_2\left(\frac{t_{ij}}{T_{i\cdot}}\right) \qquad (4.9)$$

and

$$H_{\cdot j} = -\sum_{i=1}^{S}\frac{t_{ij}}{T_{\cdot j}}\, log_2\left(\frac{t_{ij}}{T_{\cdot j}}\right) \qquad (4.10)$$

Average diversity of the biomass going to consumers, weighted by total outputs ($T_{i\cdot}$), and average diversity of inflows, weighted by total inputs ($T_{\cdot j}$), are the constitutive terms of equation (4.8), with the average diversity over both input and output that can be written as $H_c/2$. Since the diversity of pathways through a decision tree is an exponential function of the number of branch points that generate the tree, the mean number of flows from a typical node in the network should be

$$m = 2^{H_c/2} \qquad (4.11)$$

Similarly to what proposed by Ulanowicz and Wolff (1991), Bersier et al. (2002) applied the diversity of input and output biomass flows to compute a sort of effective connectance index called link density (LD).
First, they introduced the equivalent numbers of consumers for taxon i ($n_{i\cdot}$) and preys for taxon j ($n_{\cdot j}$), computed as the reciprocals of $H_{i\cdot}$ and $H_{\cdot j}$, respectively

$$n_{i\cdot} = \begin{cases} 2^{H_{i\cdot}} & \\ 0 & \text{if } T_{i\cdot} = 0 \end{cases} \qquad (4.12)$$

$$n_{\cdot j} = \begin{cases} 2^{H_{\cdot j}} & \\ 0 & \text{if } T_{\cdot j} = 0 \end{cases} \tag{4.13}$$

These equivalent numbers of consumers and preys represent the number of events that, occurring in equal proportion, would produce the same values of outflow and inflow diversity measured in a given ecosystem.
The link density is then computed averaging the equivalent numbers of consumers and preys over all the species and weighting their values by relative outflows and inflows

$$LD = \frac{1}{2}\left(\sum_{i=1}^{S} \frac{T_{i\cdot}}{T_{\cdot\cdot}} 2^{H_{i\cdot}} + \sum_{i=j}^{S} \frac{T_{\cdot j}}{T_{\cdot\cdot}} 2^{H_{\cdot j}}\right) \tag{4.14}$$

Then, the difference between the effective connectance (m) proposed by Ulanowicz and Wolff (1991) and the link density (LD) formulated by Bersier et al. (2002) resides solely in the weighting which applies, in the first case, to outflow and inflow diversities, and to taxa's equivalent numbers of consumers and preys in the latter. In particular, the effect of weighting is larger when m is computed, being applied to diversities used as exponents in the geometric mean of the input and output effective connectances.
Therefore, both the applications developed by Ulanowicz and Wolff (1991) and Bersier et al. (2002) derived from output (4.9) and input (4.10) diversities, aiming to identify average connectance per node. In particular, they refer to the equation (4.7), employing information on residual diversity (H_c) for total equivalent links (both entering and exiting each node).

Here we discuss an alternative approach, focussing on average mutual information (4.6) which accounts for constraints in the flow structure. We broke up the whole index into the relative contribution of flows entering in each node, weighting their effect by the corresponding throughput ($T_{\cdot j}$)

$$AMI_{\cdot j} = \frac{1}{T_{\cdot j}}\left(\sum_{i=1}^{S} t_{ij}\ log_2\left(\frac{t_{ij}\ T_{\cdot\cdot}}{T_{i\cdot}\ T_{\cdot j}}\right)\right) \tag{4.15}$$

The information, supplied as *bit*, is correlated to the level of input flow articulation, for each node j.
In general, in the presence of a generalist trophic behaviour, reflecting more indeterminacy in flow structure, we expect lower $AMI_{\cdot j}$ than in case of specialist diets. Since the catalyst for the formulation of $AMI_{\cdot j}$ is the Shannon measure of entropy (Shannon, 1948), $AMI_{\cdot j}$ reaches its minimum when all the input flows to node j occur in equal intensity, while the maximum is a function of the energy/matter distribution in each event. Moreover, $AMI_{\cdot j}$ and its contribution to the whole AMI depend on the fraction of throughput processed by each node j ($T_{\cdot j}$) respect to TST ($T_{\cdot\cdot}$)

$$\begin{aligned} AMI &= \sum_{j=1}^{S} \frac{T_{\cdot j}}{T_{\cdot\cdot}} \, AMI_{\cdot j} \\ &= \sum_{j=1}^{S} \frac{T_{\cdot j}}{T_{\cdot\cdot}} \, \frac{1}{T_{\cdot j}} \left[\sum_{i=1}^{S} t_{ij} \, log_2 \left(\frac{t_{ij} \, T_{\cdot\cdot}}{T_{i\cdot} \, T_{\cdot j}} \right) \right] \\ &= \sum_{i=1}^{S} \sum_{j=1}^{S} \frac{t_{ij}}{T_{\cdot\cdot}} \, log_2 \left(\frac{t_{ij} \, T_{\cdot\cdot}}{T_{i\cdot} \, T_{\cdot j}} \right) \qquad (4.16) \end{aligned}$$

4.2 Food web models

Within the simple topological descriptions of ecological communities as food webs illustrating "who eats whom", information on diversity, species composition, trophic structure, energy and material flows, and species interaction are summarized (Williams and Martinez, 2000). For this reason, discovering patterns in food web topology has long been a prominent topic in ecology (Cohen, 1978) but to unveil mechanisms lying behind community structure we are asked to identify general rules interpreting such architectures. Progresses towards this knowledge would help advances on applied issues as understanding the potentially catastrophic consequences of species loss on cascades of further extinctions (Pimm, 1980; Allesina and Bodini, 2004) and predicting the direct and indirect impacts of invasive species (Woodward and Hildrew, 2001).

The existence of these empirical patterns stimulated analysis to identify mechanisms that underlie food webs. It has been shown unambiguously that real food webs are different from randomly connected networks (Solow and Beet, 1998) and several models, using only two input parameters as species richness (S) and connectance ($C = l/S^2$, where l is the total number of directed trophic links), predict in detail many regularities of food web topology (Cohen and Newman, 1985; Williams and Martinez, 2000; Cattin et al., 2004; Stouffer et al., 2006).

In particular, two recent "static" models, the cascade model (Cohen et al., 1990) and niche model (Williams and Martinez, 2000) estimates key structural properties of the most complex food webs in nature.

4.2.1 Cascade model

Cascade model assigns species a random value extracted uniformly from the interval $[0, 1]$. Each taxa is ranked from x_0 to x_{S-1} ($S =$ total number of species), with x_0 identifying a primary producer and the remaining species x_i feeding on strictly previous taxa ($x_0, \ldots, x_{i-1}$) with probability

$$p = \frac{2SC}{S-1} \tag{4.17}$$

Diets of consumers are not restricted to contiguous range. As a consequence, moving towards nodes with higher rank the number of preys consumed increase (tendency to behave as top predators), while lower hierarchy taxa tend to be primary producers (without preys). Webs simulated by this model show, at least, one primary producer (x_0) and one top predator (x_{S-1}). The diet of each node is composed of taxa with strictly lower hierarchy position so, given two species of rank i and j, we can assume that

$$TP_i < TP_j \qquad \text{with } \; i < j \tag{4.18}$$

with TP_i and TP_j as trophic positions (TPs) of nodes i and j (see Figure 4.1). Moreover, a taxa x_i has a probability $p = 2SC/(S-1)$ of consuming only species with rank values less than its own and an expected number of preys $E_{prey}(x_i) = p(i-1)$. In conclusion, cascade model simulates food webs characterized by nodes with an increasing number of topological connections as the trophic position raises (e.g. top predators with higher rank are more likely generalist than basal species).

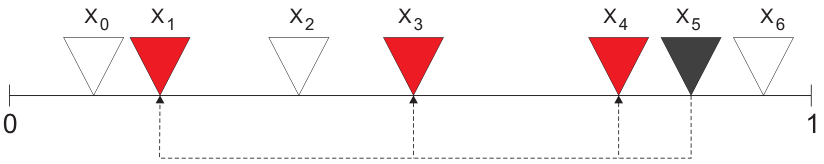


Figure 4.1: Cascade model. Species are ranked from x_0 (primary producers) to x_6 (top predator). Each taxa can exclusively consume preys with values less than its own. The diet of species x_5 (black triangle) is composed of nodes x_1, x_3 and x_4 (red triangles). Food webs generated by this model are not strictly interval (e.g. the diet of x_5 is not restricted to a contiguous range, since x_2 is excluded).

4.2.2 Niche model

The niche model assigns to each species i a niche value n_i uniformly drawn from the interval $[0, 1]$ and a parameter w_i extracted from a beta distribution with

$$p(w) = \beta(1-w)^{(\beta-1)} \qquad \text{with } \; 0 \leq w \leq 1 \tag{4.19}$$

Species i consumes all taxa falling in a range

$$r_i = n_i w_i \tag{4.20}$$

with centre c_i randomly chosen from a uniform distribution $[r_i/2, n_i]$. This restricts diets to being contiguous, allowing cannibalism and looping (since up to half a consumer's range could include species with higher niche values). Although Williams and Martinez (2000) chosen the functional equation (4.19), Camacho et al. (2002b) demonstrated how the predictions of the model are robust to changes in the specific form of $p(w)$. Linkage density ($z = l/S$) and directed connectance ($C = l/S^2$), derived from parameters $\beta = (1/2C) - 1$ and S, give rise to the analytical expression for the distribution number k of preys

$$p_{prey}(k) = \frac{1}{2z} E_1 \left(\frac{k}{2z} \right) \qquad \text{for } S \gg 1 \quad \text{and } C \ll 1 \tag{4.21}$$

where $E_1(x)$ is the exponential-integral function (Gradstheyn and Ryzhik, 2000).

For large S, we can calculate that the preys of a species i are $k_i = Sr_i$, so that the expected number of preys is given directly by the distribution of r_i and, as a consequence of equation (4.20), it depends on the niche value n_i. Specifically, with higher n_i, there is a proportional increase for r_i and the related number of preys consumed $E_{prey}(n_i) \propto (n_i)$.

As showed by Figure 4.2, species can be ordered according to their niche value, $n_0 < n_1 < \ldots < n_{S-1}$, that is, the niche values form a totally ordered set, and the range r_i constrains all consumers to eat within a contiguous interval a fixed sequence of species (generating the so called "interval" webs). Contiguity of consumers diets, absence of weighting (constructed food webs are binary) and the condition

$$c_i < n_i \tag{4.22}$$

permit to outline a good relation between niche values and tropic positions.

In conclusion, assemblages of species simulated by the niche model exhibit more generalist diets with higher trophic positions.

4.3 Connectivity and energy flow constraints

In the next chapter, 26 ecological networks are analyzed to estimate trophic position and contribution to the whole average mutual information of each species.

Average constraints displayed by inflows to ecosystem nodes ($AMI_{\cdot j}$), distinguishing relative importance of each link by means of weighting, have been considered. We emphasize the complementary character of this index in respect to previously introduced measures of effective connectance

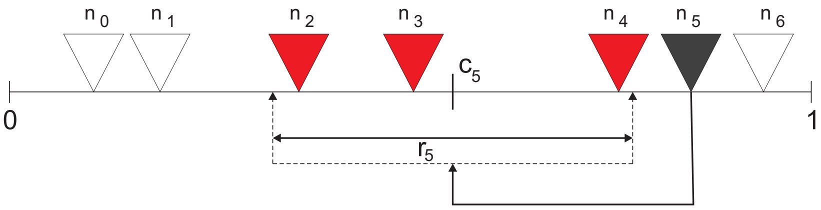


Figure 4.2: Niche model. Here species are ranked using their niche value (n_i), with interval diets (species consuming preys set in a contiguous interval) defined by a range (r_i) and its corresponding centre (c_i). Species with niche value n_5 preys upon taxa n_2, n_3 and n_4 falling into the range r_5 with centre c_5.

and link density (Ulanowicz and Wolff, 1991; Zorach and Ulanowicz, 2003; Bersier et al., 2002). Both m and LD identify, in different ways, the average number of connections per node. These approaches yield a refined and amended version of the degree centrality (Wassermann and Faust, 1994): (a) using Shannon's information formula (Shannon, 1948); (b) weighting input flows to each taxa by its specific throughput. While previous studies derived geometric means of input and output connectances from the residual diversity of flow structure (the conditional entropy H_c), here we gauged the amount by which constraints encumber the potential complexity of total inflows to each node ($AMI._j$).

Moving from primary producers to higher trophic positions we identified increasing values of $AMI._j$. This trend highlights the importance of weighting trophic links in food webs, unveiling an otherwise hidden pattern. Moreover, $AMI._j$ is not a simple weighted version of in-degree centrality, but also considers the relative importance of constraints associated to throughflow of node j respect to TST processed by the whole system.
Our analysis seems to describe higher specialization for top predators and a shift to more generalist behaviours proceeding towards species with lower trophic positions (see Chapters 5 and 6).

Despite many structural features identified using unweighted trophic links (Solé and Montoya, 2001; Dunne et al., 2002b, 2004), other studies put in evidence the importance of weighting in food webs (Paine, 1992; McCann et al., 1998; Jordán et al., 2006; Vasas and Jordán, 2006). However, webs generated by cascade (Cohen and Newman, 1985) and niche models (Williams and Martinez, 2000) powerfully predict main topological properties (e.g. the fractions of top, intermediate and basal species; the mean and standard deviation of food chain lengths and the fraction of species that are cannibals), using only qualitative data and two empirical parameters: species number (S) and connectance (C).
In our study on trophic networks (Chapters 5 and 6), we found an increasing tendency to specialistic behaviour moving from herbivores to top predators.

Despite its relevance, this linear correlation accounting for constraints to matter (or energy) flow could remain masked by the exclusive use of raw binary data. In this sense, while cascade and niche models are useful to describe topological structure of real ecosystems, they underestimate the effective specialistic activity of species at higher trophic positions, without considering the higher degree of freedom displayed by flows at the beginning of trophic chain.

We do not think that our results contradict premise on which simulated food webs are built on by cascade and niche models, but we stress the importance of quantifying link strengths to better understand consequences of ecosystem topology on functioning. Moreover, having measure that are sufficiently sensitive to provide key patterns of natural ecosystem is a fundamental target and, often, *...current measures of food web topology lack sensitivity because they take the form of averages, sums, standard deviations or proportions, calculated across all species or across subparts of a web..."* (Fox, 2006). Then, $AMI_{.j}$ and its logarithmic formulation appears as an ecologically sound index combining both structural and energetic constraints.

As suggested by Dunne et al. (2004), we perceive the need of strengthen structural studies measuring interaction intensities. In fact, when we are analyzing a weighted network we are including "functional" information in addition to the topological one and this functional information can obscure significantly the topological characteristics of the network. As a consequence, we could put further efforts to extend the original formulation of cascade and niche models on qualitative food webs, including details on link strength.

4.4 Number of roles and development capacity

[1]Number of nodes, flow organization and transfer intensities are strongly affected by the level of network detail. Therefore, indices as TST, AMI, A, and C depend on arbitrary choices taken when an ecosystem network is built. For instance, if we are interested in decomposition activities, we will emphasize resolution of microbial and non-living nodes while to evaluate the susceptibility to environmental conditions (water and nutrient availability, solar radiation, temperature, wind intensity, salinity levels, etc.), each compartment will group species showing the same behaviour.

To define development capacity variations as a function of the network framework, the connectivity (χ), calculated as flows per node, and the number of roles (ρ) are introduced

$$\chi = \frac{f}{S} \tag{4.23}$$

---

[1] Published section: Scotti, M., in press. Ecological Indicators: Development Capacity and Overhead. For Jørgensen, S. E., Encyclopedia of Ecology - Elsevier.

$$\rho = \frac{S}{\chi} = \frac{S^2}{f} = \frac{f}{\chi^2} \tag{4.24}$$

with S = number of nodes and f = total number of flows.
It appears evident how connectivity and number of flows are directly affected by network topology and, in what follows, as an alternative way to define whole-system indices are proposed.

Although the concept of role has been fundamental to develop ecological niche idea and food web research, it has never been formally defined. Trophic position (computed as the sum of the fractions of trophic activity that each species performs at different trophic levels), trophic niche (the "ecological function" carried out by species in a given ecosystem), ecological guild (two species belong to the same ecological guild when they exploit the same class of environmental resources and in a similar way) and trophospecies (recently interpreted as a set of species with similar diet or predators) describe, in a slightly different way, the trophic role in an ecosystem, but none is completely satisfactory in wholly capturing relation respect to food and enemies. An effective definition can be obtained adopting concepts from social network analysis, where the role is seen as a specialized function joining structurally equivalent nodes. In the framework of ecological flow networks this means that species belonging to the same role take input from one source and show outflows to a single destination. Then, in a recursive definition, two species (or group of species) are regular equivalent, exhibiting the same role, when eaten from and feeding on equivalent species. Nevertheless, when role is calculated as a function of connectivity and number of flows (4.23) remains a different concept respect to regular equivalence. Rather than considering the two approaches in contrast with one another it is interesting to stress the potential for their integration. The relationship between regular equivalence and number of roles is clearest when the network contains no cycles. In a linear chain with S nodes we identify S roles, while in a fully connected topology of S compartments there is only one role (see Figure 4.3).

The example of Figure 4.3 represents two hypothetical unweighted networks, not really suitable to describe the real world. In what follows, a method to extend the role calculation when dealing with real ecosystems (where flows have unequal size, with sometimes extraordinary differences) is suggested. In weighted networks, the effective connectivity is estimated accounting for the weights of flows and the portion of TST processed by each node. Applying a weighted geometric mean, connectivity (χ) and total system flows (f) become

$$\chi = \prod_{i=0}^{S+2}\prod_{j=0}^{S+2}\left(\frac{t_{ij}^{*2}}{\sum_{r=0}^{S+2} t_{rj}^{*}\sum_{v=0}^{S+2} t_{iv}^{*}}\right)^{-(1/2)\cdot(t_{ij}^{*}/TST)} \tag{4.25}$$

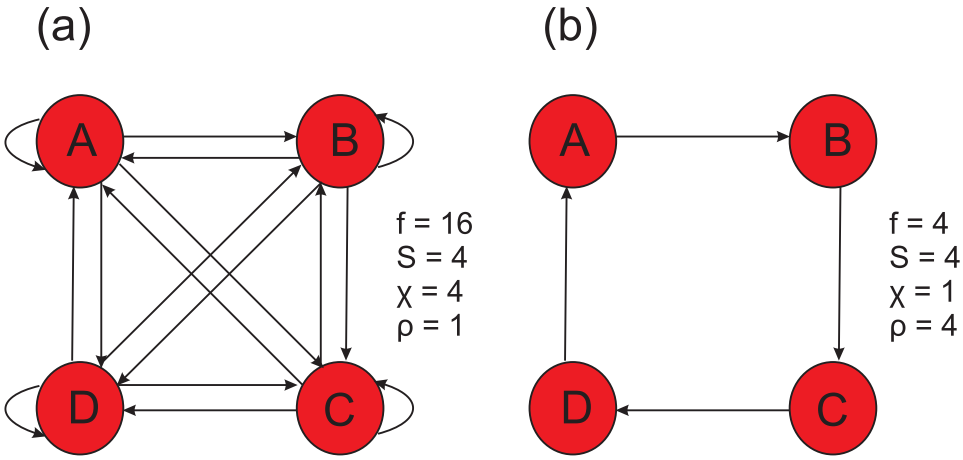


Figure 4.3: (a) The fully connected topology with 4 compartments ($S = 4$) and 16 flows ($f = 16$): in this case we calculated an average number of 4 flows per node ($\chi = 4$), with only one role ($\rho = 1$). (b) The more articulated network with $S = 4$ nodes, $f = 4$ flows, connectivity $\chi = 1$ flow per node, number of roles $= S = 4$ ($\rho = S = 4$).

$$f = \prod_{i=0}^{S+2} \prod_{j=0}^{S+2} \left(\frac{t_{ij}^{*2}}{TST} \right)^{-(t_{ij}^{*}/TST)} \tag{4.26}$$

and the derived number of roles (ρ) is

$$\rho = \frac{f}{\chi^2} = \prod_{i=0}^{S+2} \prod_{j=0}^{S+2} \left(\frac{t_{ij}^{*2}\, TST}{\sum_{r=0}^{S+2} t_{rj}^{*} \sum_{v=0}^{S+2} t_{iv}^{*}} \right)^{(t_{ij}^{*}/TST)} \tag{4.27}$$

The logarithm of ρ can also be a convenient value to measure AMI and ascendency of the system

$$log(\rho) = \frac{K}{TST} \sum_{i=0}^{S+2} \sum_{j=0}^{S+2} t_{ij}^{*}\, log \left[\frac{t_{ij}^{*} TST}{\sum_{r=0}^{S+2} t_{rj}^{*} \sum_{v=0}^{S+2} t_{iv}^{*}} \right] = AMI \tag{4.28}$$

$$A = TST \cdot log(\rho) \tag{4.29}$$

The concept of role establishes a relationship between number of nodes, connectivity and information theory indices (AMI, A and C).

Given the relation (4.28), when AMI of a system is known, the corresponding number of roles is deduced from the AMI power of e (Nepero number). As a consequence, with assigned ascendency and TST, the number of roles results rising e to the power of A/TST. Moreover, dividing the value of C by TST, it is also possible to infer the logarithm of roles when the maximum development is reached and, accordingly, the associated number of roles (ρ_C)

$$\rho = e^{AMI} \tag{4.30}$$

$$\rho = e^{\frac{A}{TST}} \tag{4.31}$$

$$\rho_C = e^{\frac{C}{TST}} \tag{4.32}$$

In Table 4.1, TST, AMI, ascendency, development capacity and number of roles (corresponding to A and C) in real ecosystems are summarized.

	TST	**AMI**	**A**	**C**	**Roles$_\mathbf{A}$**	**Roles$_\mathbf{C}$**
Charca de Masp. $mgC\ m^{-2}\ d^{-1}$	7496600	2.250	16871000	39886000	9.492	204.495
Chesapeake Bay Meso. $mgC\ m^{-2}\ sum^{-1}$	4116200	2.088	8593800	19655000	8.067	118.514
Crystal River C. (c) $mgC\ m^{-2}\ d^{-1}$	22420	1.264	28340	70712	3.540	23.4289
Crystal River C. (dT) $mgC\ m^{-2}\ d^{-1}$	18050	1.243	22434	56315	3.466	22.645
Everglades Gram. (ws) $gC\ m^{-2}\ yr^{-1}$	19949	1.937	38643	79572	6.939	53.989
Florida Bay (ws) $gC\ m^{-2}\ yr^{-1}$	3459	2.025	7004	18540	7.573	212.578
L. Chesapeake Bay $mgC\ m^{-2}\ sum^{-1}$	1451200	2.044	2966500	7713700	7.723	203.445
M. Chesapeake Bay $mgC\ m^{-2}\ sum^{-1}$	1879000	2.061	3872600	9328300	7.854	143.237
U. Chesapeake Bay $mgC\ m^{-2}\ sum^{-1}$	854330	2.133	1822300	4583700	8.440	213.846
St. Marks River $mgC\ m^{-2}\ d^{-1}$	2064	1.805	3726	11264	6.079	234.293
Lake Michigan $mgC\ m^{-2}\ d^{-1}$	36985	1.775	65649	140690	5.900	44.879
Mondego Estuary $gAFDW\ m^{-2}\ yr^{-1}$	10852	1.525	16547	39126	4.594	36.797
Final Narragansett $mgC\ m^{-2}\ yr^{-1}$	4611300	1.628	7506700	20464000	5.093	84.588
Ythan Estuary $gC\ m^{-2}\ yr^{-1}$	5440	1.592	8663	23397	4.915	73.744

Table 4.1: Total system throughput (TST), average mutual information (AMI), ascendency (A) and development capacity (C) are summarized for 14 ecosystems. Number of roles related to ascendency and development capacity are computed (Roles$_A$ and Roles$_C$, respectively) adopting formula (4.29) as a reference.

For the 14 ecosystems of Table 4.1, when the number of roles is computed starting from the ascendancy, through formula (4.31), it ranges between 3.466 (Crystal River Creek - delta temp.) and 9.492 (Charca de Maspalomas), with an average value of 6.405. These data confirm how the concept of role shows several relationship to that of trophic position (TP), since both these indices are limited by the maximum number of passages that energy can experience from the ultimate source of energy (imports - outside environment). Values of 10% $\div$ 20% are rather accepted as average living compartment efficiency in processing food, setting to 6 $\div$ 8 the maximum number of energy steps (corresponding to TPs) in real networks. While TP is a property depending exclusively on compartment feeding activity, both prey items and predators affect the role that is assigned to each node. In ecosystems, even though procedures computing TPs and roles slightly differ, they are deeply settled by energy and efficiency constraints. Therefore, it appears evident how the number of roles computed through the development capacity (4.32) is overestimated (min = 22.645; max = 234.293; mean = 119.320), being associated to an ecologically unfeasible topology that does not consider constraints to energy transfer efficiency in natural systems.

5

Trophic positions and flow structure

5.1 The higher the better?

[1]Food webs are complex networks of trophic interactions among diverse species in ecosystems. Ecologists have expended much effort to explore and understand this complexity. Empirical patterns and trends (e.g. predator/prey ratio, fraction of species that occupy definite trophic positions, number of connections per species, food chain length, to cite a few) emerged from the analysis of food web data (Cohen et al., 1990; Cohen and Newman, 1991) and this phenomenology was explained through general models that reproduce those regularities (Cohen et al., 1990; Williams and Martinez, 2004; Cattin et al., 2004).

Such regularities depend on food web structure, that is how many links are there and how they are distributed between the nodes. Accordingly, unveiling patterns of link arrangements may shed light on the ultimate cause of these regularities, thus contributing to grasp essential organizing principles of food webs and the relationship between their structure and functions. Because number and distribution of links depend on species trophic habits, investigating whether link organization can be associated to the trophic structure of the food webs can help in this direction.

---

[1] "Poster session" Chapter: Scotti, M., Bondavalli, C., Bodini, A, 2007. Linking trophic positions and flow structure constraints in ecological networks: efficiency or topology effect? ECEM'07, Trieste, Italy.

The relationship between species trophic activity and food web topology, although implicitly acknowledged, has never been explored in detail. Null models for food web structure (Cohen et al., 1990; Williams and Martinez, 2004; Cattin et al., 2004) postulate the existence of a niche axis that determines a hierarchy among the species composing a food web. In this hierarchy, for mere probability reasons, the higher the trophic level of a species, the higher the number of prey it should have. In other words, species that feed higher should contribute with more connections to the system topology. This deductive outcome has inspired this research whose main objective is to investigate whether web topology forms patterns throughout ecosystem that can be explained by the species trophic position.

Models of food web structure are strictly qualitative in nature, as they were developed considering only the presence/absence of links between species. According to this, as a first step of investigation, we performed our analysis using qualitative food webs to assess whether the relationship between species trophic position and link distribution could be accommodated within the logic that characterize food web models. On the other hand, link quantification has been called for as an essential ingredient for a more rigorous approach to food web analysis (Borer et al., 2002; Berlow et al., 2004) and in this framework we performed a quantitative investigation of the same relationship. The aim here was to see whether the criteria that comes out from food web null models hold when quantification of links is included in the analysis.

For this dual approach we used food webs extracted from 26 weighted ecological flow networks, with links that are quantitatively expressed as energetic contributions from resources to consumers. To conduct our analysis we calculated the index of trophic position (TP; Ulanowicz, 1995; Scotti et al., 2006) for each species, as defined by network analysis. Also, we computed an unweighted version for this index similar to the binary link-based structural index used in other contexts (Williams and Martinez, 2004; Dunne, 2006). The contribution to web topology is assessed by the average mutual information (AMI, henceforth; Ulanowicz, 2004), another index taken from the suite of network analysis indices. Its value depends on the multiplicity of links targeted to each component weighted by their relative magnitude. This index measures the contribution that each compartment, through its trophic connections, gives to the organization of system connections. An unweighted version of AMI have been calculated as well, assuming all the links of the same magnitude.

5.2 TPs and AMIs of single compartments

5.2.1 Weighted food webs

Weighted ecological networks of 26 ecosystems, at disposal in two websites, were analyzed. One set (8 graphs) comes from the project site [2]ATLSS and the other 18 models are included in the Prof. [3]Ulanowicz's web page. These two groups are identified, respectively, as ATLSS and NETWRK. List and some description of the 26 networks are provided in Table 5.1.

Ecosystem	S	nl	**currency**
NETWRK			
Aggregated Baltic Sea	15	3	mgC m^{-2} day^{-1}
Cedar Bog Lake	9	1	cal cm^{-2} $year^{-1}$
Charca de Maspalomas	21	3	mgC m^{-2} day^{-1}
Chesapeake Mesohaline Ecosystem	15	3	mgC m^{-2} day^{-1}
Chesapeake Mesohaline Network	36	3	mgC m^{-2} $summer^{-1}$
Crystal River Creek (control)	21	1	mgC m^{-2} day^{-1}
Crystal River Creek (delta temp.)	21	1	mgC m^{-2} day^{-1}
Lower Chesapeake Bay in Summer	34	3	mgC m^{-2} $summer^{-1}$
St.Marks River (Florida) Flow Network	51	3	mgC m^{-2} day^{-1}
Lake Michigan Control Network	36	1	gC m^{-2} $year^{-1}$
Middle Chesapeake Bay in Summer	34	3	mgC m^{-2} $summer^{-1}$
Mondego Estuary	43	1	gAFDW m^{-2} $year^{-1}$
Final Narraganasett Bay Model	32	1	mgC m^{-2} $year^{-1}$
North Sea	10	0	kcal m^{-2} $year^{-1}$
Somme Estuary	9	1	gC m^{-2} $year^{-1}$
Upper Chesapeake Bay in Summer	34	3	mgC m^{-2} $summer^{-1}$
Upper Chesapeake Bay	12	2	gC m^{-2} $year^{-1}$
Ythan Estuary	13	3	gC m^{-2} $year^{-1}$
ATLSS			
Cypress Wetlands dry season	68	3	gC m^{-2} $year^{-1}$
Cypress Wetlands wet season	68	3	gC m^{-2} $year^{-1}$
Marshes and Sloughs dry season	66	3	gC m^{-2} $year^{-1}$
Marshes and Sloughs wet season	66	3	gC m^{-2} $year^{-1}$
Florida Bay dry season	125	3	gC m^{-2} $year^{-1}$
Florida Bay wet season	125	3	gC m^{-2} $year^{-1}$
Mangroves dry season	94	3	gC m^{-2} $year^{-1}$
Mangroves wet season	94	3	gC m^{-2} $year^{-1}$

Table 5.1: List of the ecological networks considered in the analysis. Total number of nodes (S) and number of non-living compartments (nl) are given. Flow intensities are measured as energy (i.e. kcal m^{-2} $year^{-1}$ or cal cm^{-2} $year^{-1}$), carbon (i.e. gC m^{-2} $year^{-1}$ or mgC m^{-2} $summer^{-1}$) and ash-free dry weight (gAFDW m^{-2} $year^{-1}$).

5.2.2 Trophic position

In ecosystem network analysis every species trophic position is computed through a suite of matrix manipulations called canonical trophic aggregation (CTA; Ulanowicz, 2002; Scotti et al., 2006). Simply put, every species feeding activity is apportioned to a series of discrete trophic levels sensu

---

[2] Across Trophic Level System Simulation; http://www.cbl.umces.edu/∼atlss/

[3] http://www.cbl.umces.edu/∼ulan/

Lindeman and the effective trophic position comes out as the sum of these fractions. While the Chapter 3 provides essential mathematical details to follow this discussion, suffice here to say that the effective trophic position is defined as the weighted average length of all the pathways that originate from the outside environment (primary source of energy) and reach a given compartment. In other words, it is the weighted average distance of a given node from the external source of energy. Accordingly, autotrophs will have distance (and therefore TP) equal to 1, herbivores 2 and so forth. Usually TPs are fractionary: for a species that bases half of its diet on primary producers and the other half on herbivores, the TP will be $2 \times 0.5 + 3 \times 0.5 = 2.5$ (one pathway of length 2: outside $\longrightarrow$ primary producers $\longrightarrow$ herbivore; one of length 3: outside $\longrightarrow$ primary producers $\longrightarrow$ herbivore $\longrightarrow$ consumer). Organisms feeding on this latter species would have $TP = 3.5$.

Adopting this "flow-based" approach (Williams and Martinez, 2004), the trophic position of the j^{th} species is computed as

$$TP_j = 1 + \sum_{i=1}^{S} TP_i \cdot p_{ij} \tag{5.1}$$

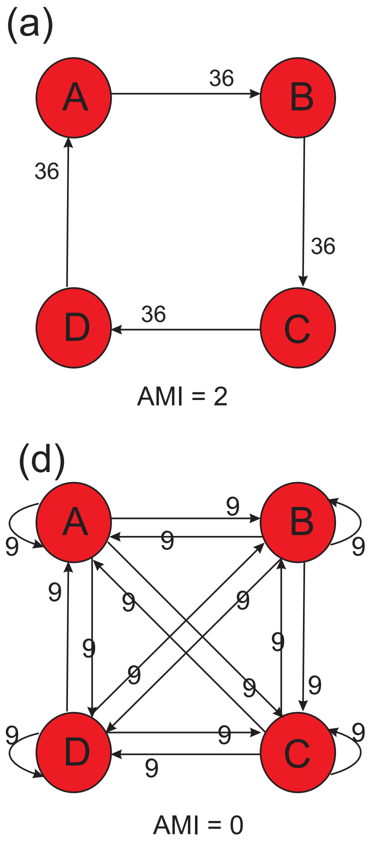
where TP_j and TP_i are the trophic positions of species i and j, respectively, S stands for the total number of compartments and p_{ij} is the fraction that species i constitutes within the diet of species j. The unweighted counterpart is the "prey-averaged" trophic position (Williams and Martinez, 2004), which assumes that a consumer feeds on all its prey species equally. Then, in a S species network, the trophic position of a j species (TP_j) connected by $l_{ij} = 1$ with its m_j preys is

$$TP_j = 1 + \sum_{i=1}^{S} l_{ij} \cdot \frac{TP_i}{m_j} \tag{5.2}$$

5.2.3 Average mutual information

The average mutual information (AMI) computes the average amount of constraints exerted upon an arbitrary quantum of currency when passing from any one compartment to the next (Rutledge et al., 1976; Ulanowicz, 1986). This depends on both the number of links and their magnitude and can be explained using Figure 5.1 as reference.

The upper left configuration (Figure 5.1a) shows the maximally constrained distribution of the same total amount of currency (total system throughput); no alternative routes exist to the movement of a quantum of matter: it moves in an exactly predictable way. On the contrary, network (d) possesses the most equivocal distribution of the 144 units of transfer among the four system components. According to the meaning of this index and the way it is computed, network (a) depicted in Figure 5.1 possesses



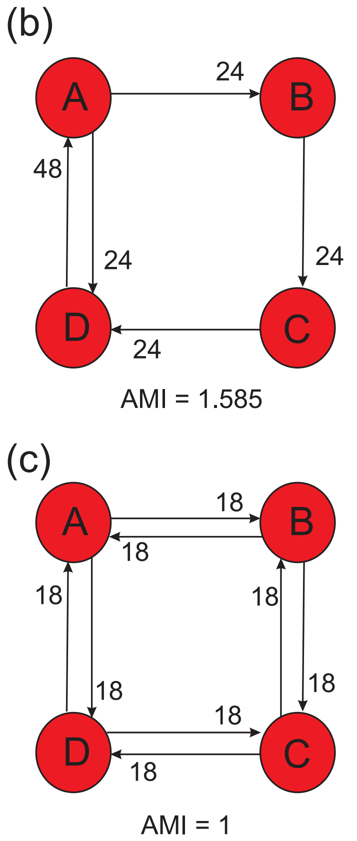


Figure 5.1: Hypothetical networks depicted in 4 different topologies: (a) maximally constrained (AMI maximum); (b) and (c) intermediate levels of constraints; (d) most equivocal distribution of links ($AMI = 0$). Every node is at steady state and the total system throughput is preserved ($TST = 144$).

the highest AMI, whereas the lowest is for network (d) (see Chapter 4 for details).

This example shows the intrinsic whole-system character of this index. However, Ulanowicz (unpublished manuscript) has shown how the individual contribution of the compartments can be estimated summing up all the terms generated by inflows to each of them, and dividing this value by the corresponding throughput (sum of all inflows to a given compartment). In particular, each compartment contribution to the system flow structure is estimated in terms of inflow topology, scaled by the ratio of total system throughput (TST) processed. This index represents a measure that is similar, although complementary, to effective connectance (m; Ulanowicz and Wolff, 1991) and link density (LD; Bersier et al., 2002) previously proposed (see Chapter 4, section 4.1). In the present chapter, the contribution of each species to the whole AMI is denoted with $AMI_{.j}$, adopting it as a proxy of the flow structure.

5.2.4 Statistical analysis of food web structure

We calculated the trophic position (TP) and the average mutual information ($AMI_{\cdot j}$) for each of the living components in the 26 weighted food webs and their corresponding unweighted versions (imposing magnitude equal to 1 to all links). Then we evaluated the correlation between the two indices using the the Pearson's product-moment correlation (ρ). Differences in the correlation coefficients obtained for the two groups of networks (NETWRK and ATLSS) were studied by applying the Wilcoxon-Mann-Whitney test (Wilcoxon, 1945; Mann and Whitney, 1947).

5.3 Comparing binary and weighted food webs

Table 5.2 reports the correlation coefficients between TP and AMI for the 26 ecosystems under investigation.

Ecosystem	$\rho_{\mathbf{q}}$	**p**	$\rho_{\mathbf{Q}}$	**p**
NETWRK				
Aggregated Baltic Sea	-0.639	0.025	0.552	0.063
Cedar Bog Lake	-0.858	0.006	0.811	0.015
Charca de Maspalomas	-0.552	0.017	0.327	0.185
Chesapeake Mesohaline Ecosystem	-0.681	0.015	0.527	0.079
Chesapeake Mesohaline Network	-0.456	0.008	0.692	$\ll 0.001$
Crystal River Creek (control)	-0.693	0.001	0.774	$\ll 0.001$
Crystal River Creek (delta temp.)	-0.636	0.003	0.751	$\ll 0.001$
Lower Chesapeake Bay in Summer	-0.220	0.233	0.442	0.013
St.Marks River (Florida) Flow Network	-0.556	$\ll 0.001$	0.358	0.012
Lake Michigan Control Network	-0.736	$\ll 0.001$	0.901	$\ll 0.001$
Middle Chesapeake Bay in Summer	-0.537	0.002	0.715	$\ll 0.001$
Mondego Estuary	-0.872	$\ll 0.001$	0.516	$\ll 0.001$
Final Narraganasett Bay Model	-0.737	$\ll 0.001$	0.798	$\ll 0.001$
North Sea	-0.577	0.081	0.976	$\ll 0.001$
Somme Estuary	-0.930	0.001	0.836	0.010
Upper Chesapeake Bay in Summer	-0.432	0.015	0.639	$\ll 0.001$
Upper Chesapeake Bay	-0.440	0.203	0.832	0.003
Ythan Estuary	-0.887	0.001	0.886	0.001
ATLSS				
Cypress Wetlands dry season	-0.794	$\ll 0.001$	0.719	$\ll 0.001$
Cypress Wetlands wet season	-0.809	$\ll 0.001$	0.457	$\ll 0.001$
Marshes and Sloughs dry season	-0.675	$\ll 0.001$	0.758	$\ll 0.001$
Marshes and Sloughs wet season	-0.674	$\ll 0.001$	0.745	$\ll 0.001$
Florida Bay dry season	-0.882	$\ll 0.001$	0.914	$\ll 0.001$
Florida Bay wet season	-0.878	$\ll 0.001$	0.896	$\ll 0.001$
Mangroves dry season	-0.821	$\ll 0.001$	0.657	$\ll 0.001$
Mangroves wet season	-0.820	$\ll 0.001$	0.655	$\ll 0.001$

Table 5.2: Correlation coefficients for plots of TPs against AMIs in food webs ($\rho_{\mathbf{q}}$) and trophic networks ($\rho_{\mathbf{Q}}$) with their p-values

Plots of the relationship between TP and AMI are organized in Figures from 5.2 to 5.5.

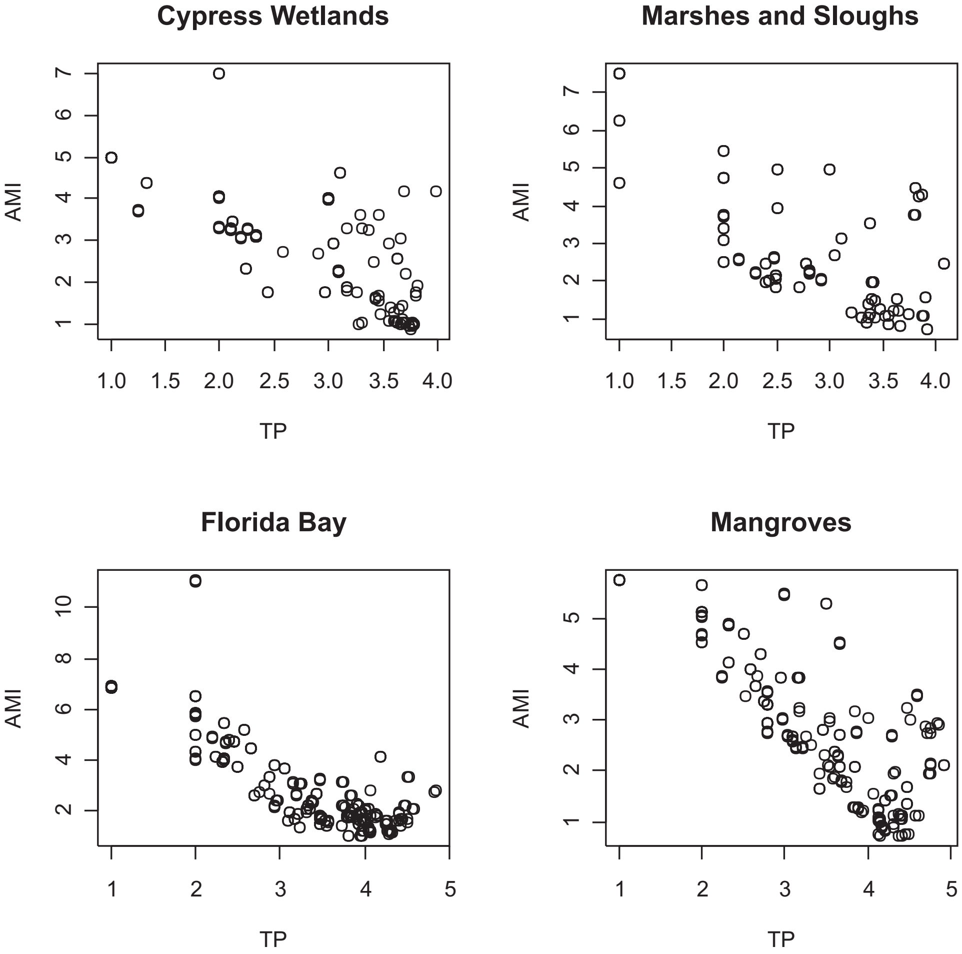


Figure 5.2: Florida Bay dataset, binary data.

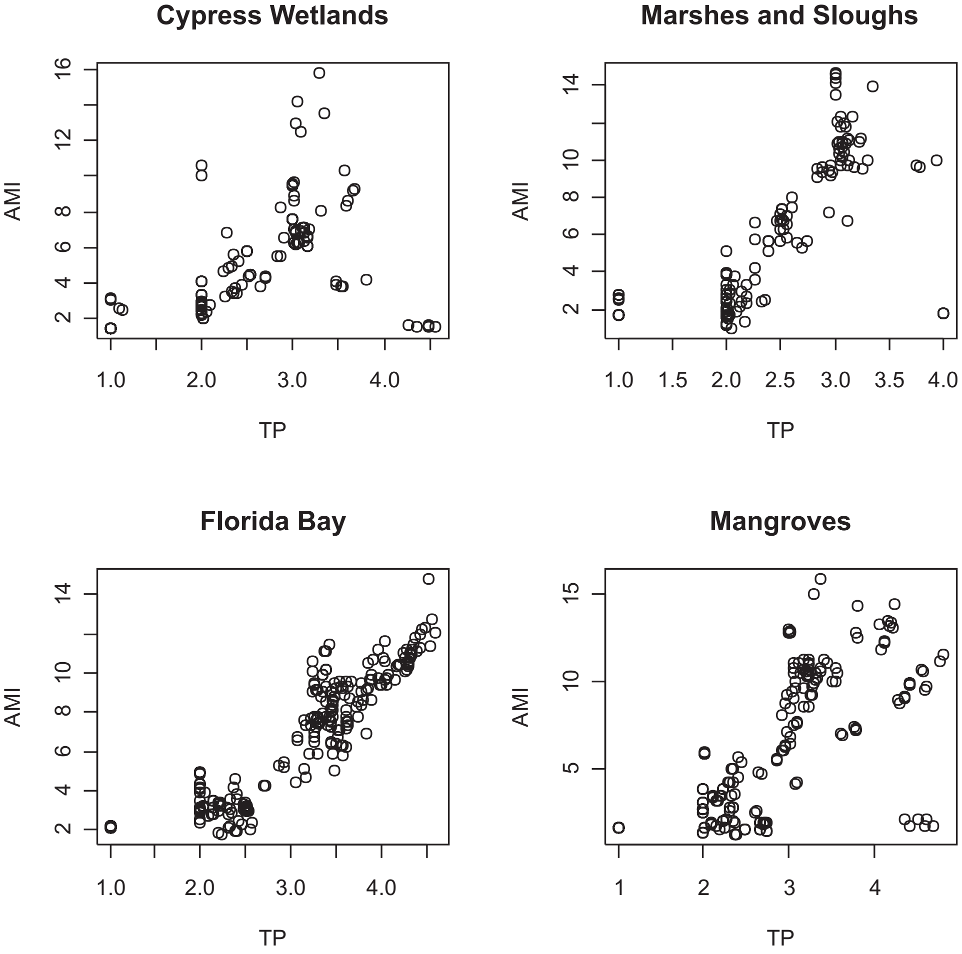


Figure 5.3: Florida Bay dataset, weighted links.

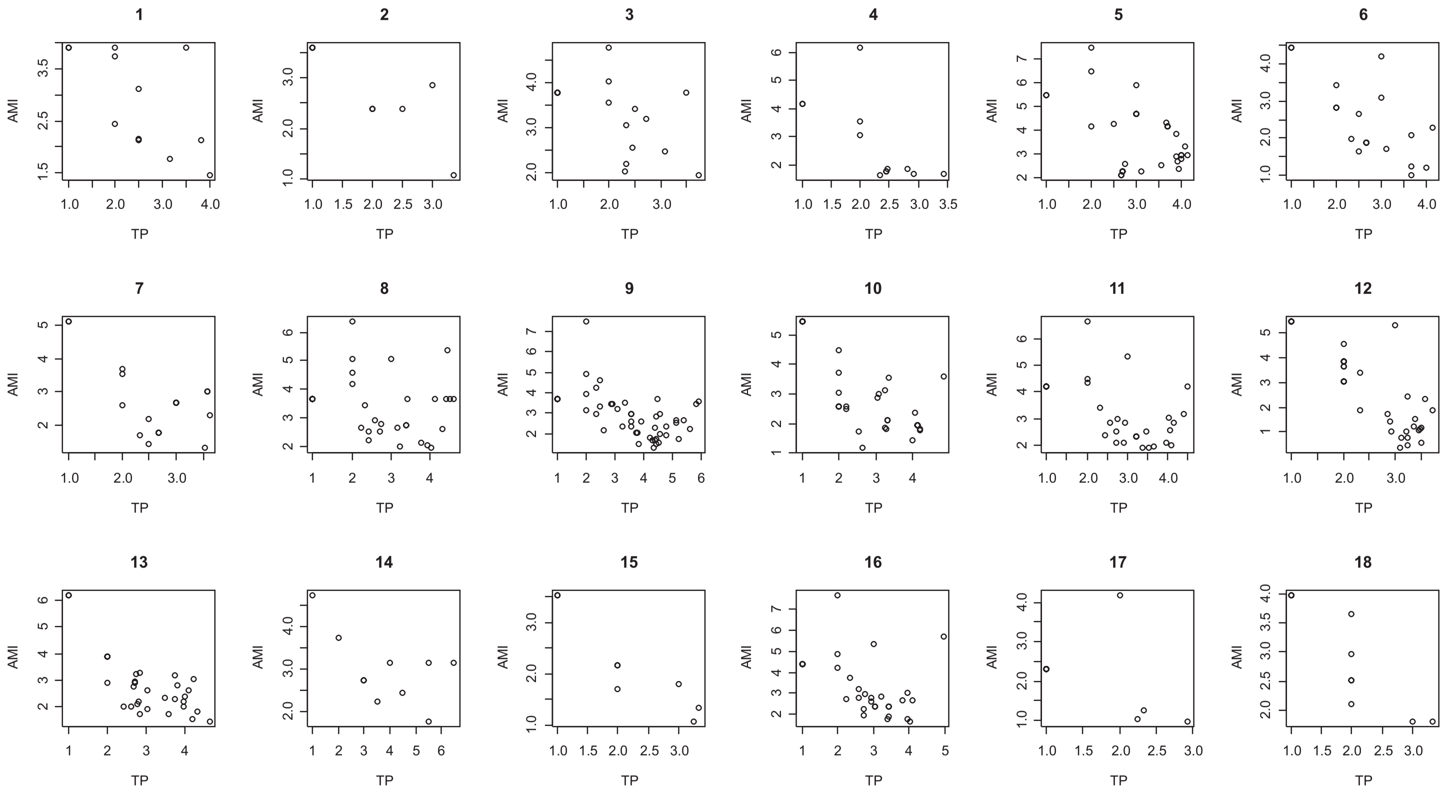


Figure 5.4: NETWRK database, binary data. Numbers above each plot stand for ecosystem names: (1) Aggregated Baltic Sea; (2) Cedar Bog Lake; (3) Charca de Maspalomas; (4) Chesapeake Mesohaline Ecosystem; (5) Chesapeake Mesohaline Network; (6) Crystal River Creek (control); (7) Crystal River Creek (delta temp.); (8) Lower Chesapeake Bay in Summer; (9) St. Marks River; (10) Lake Michigan; (11) Middle Chesapeake Bay in Summer; (12) Mondego Estuary; (13) Final Narraganasett Bay Model; (14) North Sea; (15) Somme Estuary; (16) Upper Chesapeake Bay in Summer; (17) Upper Chesapeake Bay; (18) Ythan Estuary.

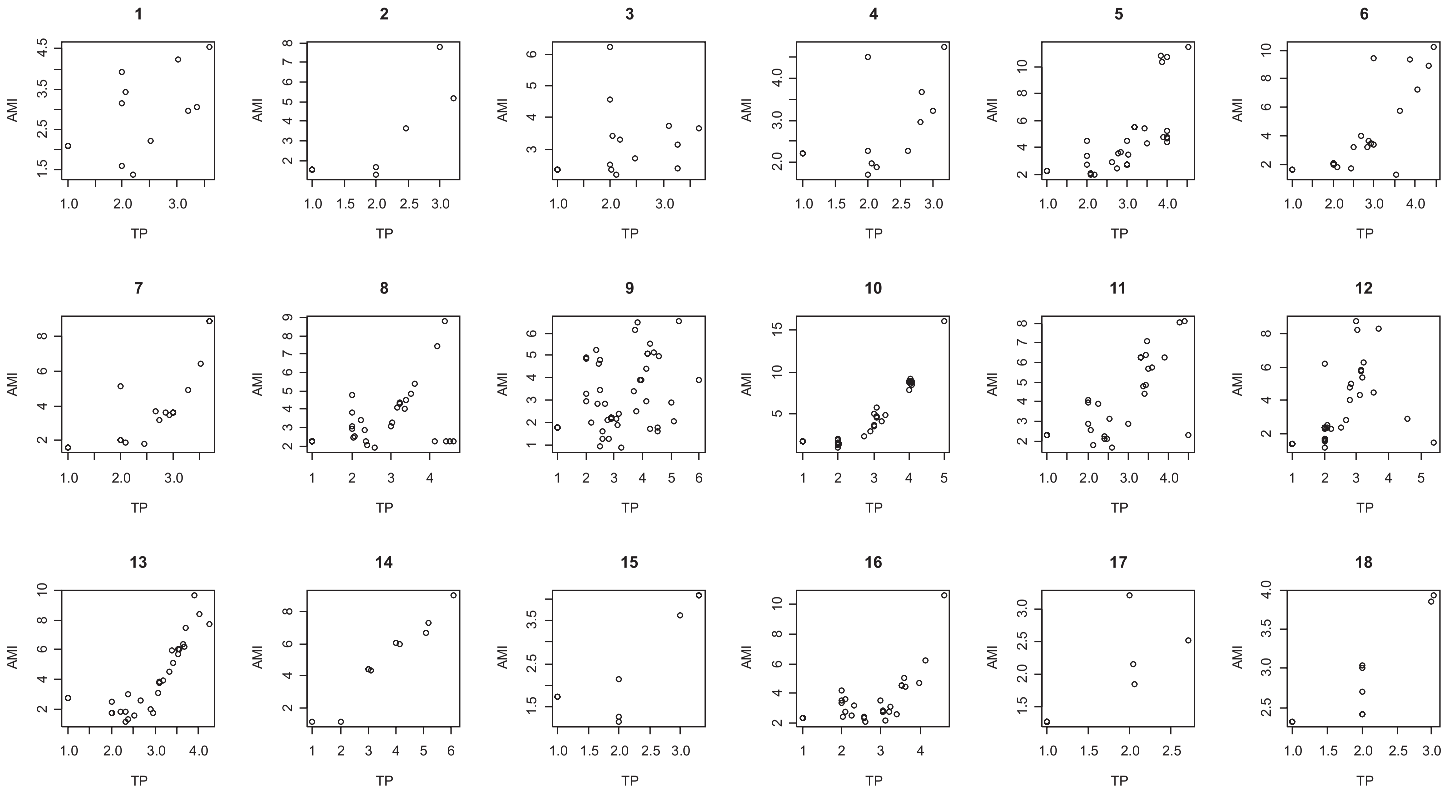


Figure 5.5: NETWRK database, weighted links. Numbers above each plot stand for ecosystem names: (1) Aggregated Baltic Sea; (2) Cedar Bog Lake; (3) Charca de Maspalomas; (4) Chesapeake Mesohaline Ecosystem; (5) Chesapeake Mesohaline Network; (6) Crystal River Creek (control); (7) Crystal River Creek (delta temp.); (8) Lower Chesapeake Bay in Summer; (9) St. Marks River; (10) Lake Michigan; (11) Middle Chesapeake Bay in Summer; (12) Mondego Estuary; (13) Final Narraganasett Bay Model; (14) North Sea; (15) Somme Estuary; (16) Upper Chesapeake Bay in Summer; (17) Upper Chesapeake Bay; (18) Ythan Estuary.

In particular, those classified as ATLSS are the most resolved networks among the 26 we have examined and their plots are reproduced separately in Figure 5.2 (qualitative) and Figure 5.3 (quantitative). For the networks of the group NETWRK, characterized by more heterogeneous levels of resolution into trophic species, we summarized the results into two figures: one for qualitative indices (Figure 5.4) and the other for quantitative ones (Figure 5.5).

When trophic position and average mutual information of each node were computed and plotted for the food web version of analyzed ecosystems (flows between species simply characterized as present or absent), a strong negative correlation was estimated. All the 26 ecosystems but Lower Chesapeake Bay in Summer ($\rho_{qualitative} = -0.220$, $p = 0.233$), North Sea ($\rho_{qualitative} = -0.577$, $p = 0.081$) and Upper Chesapeake Bay ($\rho_{qualitative} = -0.440$, $p = 0.203$) showed significant correlation ($\mu = -0.720$) and results were not affected by systems belonging to different datasets ($Z = 1.581$, $p > 0.05$).

In ecological networks for which interaction strength were measured, the positive correlation between TP and AMI is observed in 25 ecosystem networks out of 26 ($p \ll 0.05$), with the exception of Charca de Maspalomas ($\rho_Q = 0.327$, $p = 0.185$). Most (60%) of the correlation coefficients are greater than 0.7, with a mean value equal to 0.712. No significant difference was found between the two groups of networks (ATLSS and NETWRK) as for correlation coefficients ($Z = 0.146$, $p \gg 0.05$).

For each ecosystem, we plotted the Pearson's product-moment correlation against the number of nodes (S) and no scale dependence has been noticed both for unweighted and weighted data (see Figure 5.6: food webs, $R^2 = 0.093$; ecological networks, $R^2 = -0.036$).

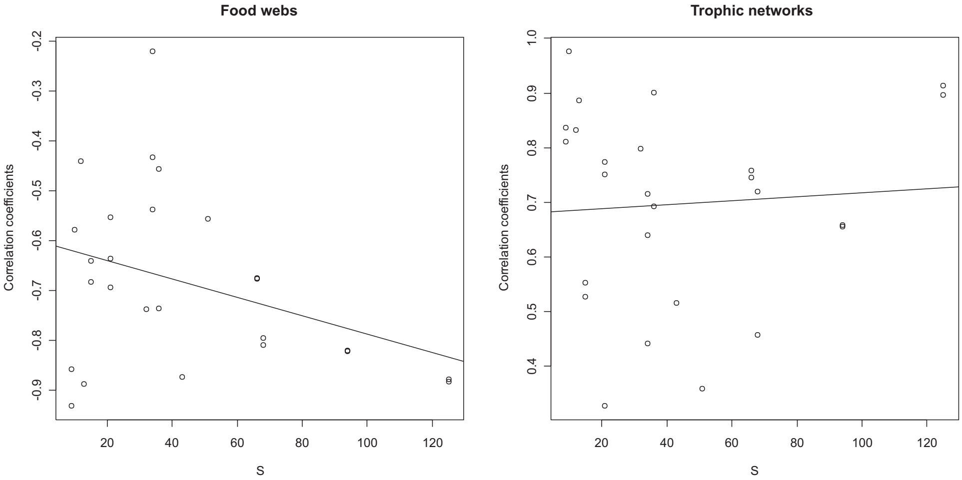


Figure 5.6: Absence of scale dependence effect (linear regression line is depicted) when the correlation coefficients of 26 whole ecosystems (ρ) are plotted against their number of nodes (S), both using unweighted (on the left) and weighted (on the right) data.

5.4 From topology to functioning

The relationship between trophic position and link distribution forms defined patterns. By insisting on food web models (Cohen et al., 1990; Williams and Martinez, 2004), the outcomes of the hierarchy they implicitly assume, that is the higher the trophic level of a species the higher the number of prey it will have, seems to be confirmed when the analysis of food webs is conducted on binary (presence/absence of interaction) data. A negative correlation, in fact, constantly emerge between TP and AMI and because the latter index is inversely proportional to the randomness of flows, that is it scores higher when connections entering a node are few, the result obtained for the 25 systems tells us that the number of links would increase towards the top of the food chain.

On the other hand, taking into account the magnitude of interactions reverses this outcome. When link strength is included, the amount of information grows with the trophic level of the species: this, in general, can be obtained either by having a lower number of preys or a very skewed distribution of interactions; however, accounting for the results on binary webs, the only possible explanation for this reversal remains the uneven distribution of link intensity, whose effect remain masked by the exclusive use of raw binary data.

Our results communicate the sobering message that in food webs top species establish more interactions than basal species but most of these interactions are weak. This supports the idea that these interactions prevail in ecosystems, a point of view rather diffuse in the ecological literature, and built up with contributions from different approaches: from energy flow studies (de Ruiter et al., 1995; Raffaelli and Hall, 1996) to researches on food web stability (McCann et al., 1998; Neutel et al., 2002) and investigations involving field and lab experiments (Ives et al., 1999; Sala and Graham, 2002).

The outcomes presented here do not discard the validity of food web models. On the contrary, from a pure topological approach this research adds another evidence about the potential of these architectures, that are extremely simple in their premises to capture the very essence link arrangements in food webs. However, if one passes from the bare connective structure to a more functional characterization of ecosystems, the binary approach holds only marginally, if at all. Studies on network robustness and secondary extinction (Dunne et al., 2002b; Allesina and Bodini, 2004) have revealed that a pure topological approach provides average insight but fails to make accurate predictions on the impact of species loss. Adding link magnitude changes significantly the scenario (Allesina et al., 2006).

However, it must be considered that a functional characterization of food webs depends very much on the way interaction strength is measured. Because this characterization refers to the functional role occupied by species, when link intensity is expressed as energy transfer from resources to consumer, like in this work, keystone species are those that act as bottleneck for energy delivery throughout the ecosystem (Allesina et al., 2006). On the other hand, the same species may not be as important, say, in a top-down perspective, in which other types of regulatory flows are at work in ecosystems (Paine, 1980, 1992; Raffaelli and Hall, 1996). According to this, despite its importance, including link quantification into food web analysis will contribute to make progresses towards an insightful synthesis, only if all the different aspects connected to link weighting will be taken into account and accommodated in a coherent framework (Berlow et al., 2004).

6

Species trophic hierarchy

6.1 Trophic hierarchy as a criterion for link distribution around the energy flow in ecosystems

The trophic level concept of a simple linear chain of energy passages that had great appeal as an easy and intuitive description of complex communities has recently been challenged. The structural intricacy of food webs (e.g. richness and topology of connections between species) would not easily accommodate into the framework of the "green world" theories (Polis and Strong, 1996). Spatial and temporal heterogeneity, diffuse omnivory and nutrient cycling, shunt the flow of matter and energy away from adjacent trophic components thus challenging the idea that populations aggregate into discrete homogeneous trophic levels (Vadas Jr., 1990; Winemiller, 1990; Polis, 1991, 1994), although recent investigations reevaluated the concept of trophic level and its scientific utility (Williams and Martinez, 2004).

If the Lindeman tropho-dynamic sequence of discrete trophic levels is a too rigid model to represent ecosystems, multiple interconnections between species in food webs yield the perception that link topology is not organized around the observed gradient of trophic positions (Ulanowicz, 1995; Scotti et al., 2006), and trophic hierarchy seems not a criterion explaining patterns of link distribution in food webs.

Nonetheless, links in food webs are representations of who eats whom in ecosystems, and their topology is unequivocally determined by the species feeding behaviour. Accordingly, if link distribution shows patterns that can be put in relation with the gradient of trophic positions from producers to consumers, and holds across ecosystems, the trophic hierarchy becomes a key to explain food web structure.

To search for one such relation is the objective of this paper. To pursue it, we analyzed food webs extracted from 26 weighted ecological flow networks, with links quantitatively expressed as energetic contributions from resources to consumers. In particular, in our analysis we considered species trophic position (TP; Ulanowicz, 1995; Scotti et al., 2006) as an index of trophic classification for the various compartments, and the decomposition of the average mutual information (AMI, henceforth; Ulanowicz, 2004), another index taken from network analysis, as a metric of link density. Its value depends on the multiplicity of links targeted to each node weighted by their relative magnitude. This index measures the contribution that each compartment, through its trophic connections and their relative importance, gives to the organization of energy flow in ecosystems.
In particular, we address our objective by answering the following questions: (a) is the density of links for each species related to its trophic position? (b) if so, can patterns be identified among ecosystems? (c) is there evidence that the attitude towards specialized or redundant connections is more a taxonomically based attribute than a functional one?

Contrarily to most food web studies centred on topology (Montoya and Solé, 2002; Dunne et al., 2002a,b; Allesina and Bodini, 2004)), here we use directed, weighted food webs. This choice roots in the idea that crude linkage density should be amended to include, besides the number of links, also their contribution in terms of energy flow. Accordingly, many links with same magnitude define higher linkage density than few links and uneven distribution of medium among them (Ulanowicz, 1997). In doing this we propose a shift in focus: from the bare connective structure to the function associated to links, that is energy distribution.

6.2 TPs and AMIs of single compartments

6.2.1 Weighted food webs

In the present chapter we analyzed data described in Chapter 5, section 5.2.1 (see Table 5.1 for details).

6.2.2 TPs and AMIs

Descriptions of the indices measuring steps experienced by energy to reach a given node (TP) and contribution, of each species, to the flow structure ($AMI_{.j}$), are summarized in Chapter 5, sections 5.2.2 and 5.2.3.

6.2.3 Statistical analysis of food web structure

Trophic position (TP) and average mutual information ($AMI_{\cdot j}$) for each of the living components in the 26 weighted food webs were calculated. Then, the correlation between the two indices has been evaluated using the Pearson's product-moment correlation (ρ) both for pooled data and for every single ecosystem. Because the systems used in this analysis were extracted from two datasets, it is likely that different rules of constructions were used to assemble the networks. To verify whether this had some effect on the final results (i.e. the correlation between TP and $AMI_{\cdot j}$), differences in the correlation coefficients obtained for the two groups of networks were studied by applying the Wilcoxon-Mann-Whitney test (Wilcoxon, 1945; Mann and Whitney, 1947).

Correlation between TP and $AMI_{\cdot j}$ for system subgroups including taxonomically related species was computed; they are: microfauna, macroinvertebrates, herpetofauna, fishes, avifauna and mammals. This was done in order to observe whether values of correlation coefficients produced a pattern reflecting taxonomic classification. These six subgroups represent a common conventional subdivision of compartments for all the 26 networks; inclusion of species or compartments in each group was decided according to taxonomic classification provided by the authors who compiled the networks. For these groups, Pearson's product-moment correlations were ranked, then their order has been compared applying the Kruskal-Wallis test (Kruskal and Wallis, 1952).

6.3 Trophic hierarchy and flow structure

For ease of explanation, results in detail are presented only for the Florida Bay (wet season) ecosystem, one of the 8 networks included in the ATLSS database. This ecosystem is composed of 125 compartments of which 3 stand for non-living nodes (water POC, benthic POC and DOC). Focussing on living nodes, there are 14 primary producers, 12 microfauna species, 27 macroinvertebrates, 48 fish, 4 herpetofauna species, 16 birds species (avifauna) and 2 mammals. Each compartment name is detailed in Table 6.1 and, for each of them, trophic position (TP) and average mutual information ($AMI_{\cdot j}$) are summarized. Compartments are also ranked according to the values of these two coefficients. The relationship between the two indices is plotted in Figure 6.1.

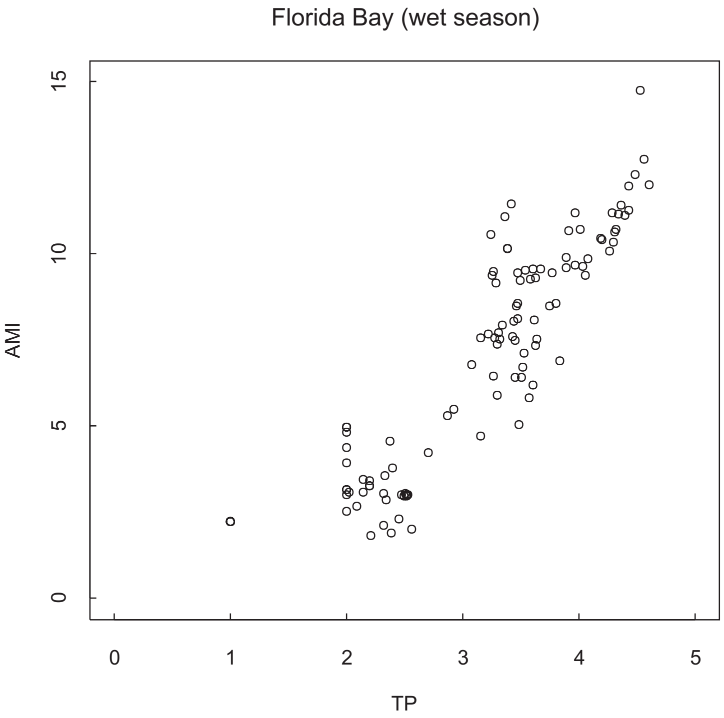


Figure 6.1: Plot of the distribution of the Florida Bay (wet season) compartments according to their average mutual information (AMI) and trophic position (TP). All the 14 primary producers share the same values ($TP = 1$ and $AMI_{\cdot j} = 2.228$). The Pearson's product-moment correlation is $\rho = 0.896$).

	nodes	TP	$AMI_{\cdot j}$	**rank** TP	**rank**$AMI_{\cdot j}$
	Primary producers				
1	2um Spherical Phytoplankton	1.000	2.228	109	105
2	*Synedococcus* sp.	1.000	2.228	109	105
3	*Oscillatoria* sp.	1.000	2.228	109	105
4	Small Diatoms (¡20um)	1.000	2.228	109	105
5	Big Diatoms (¿20um)	1.000	2.228	109	105
6	Dinoflagellates	1.000	2.228	109	105
7	Other Phytoplankton	1.000	2.228	109	105
8	Benthic microalgae	1.000	2.228	109	105
9	*Thalassia testudium*	1.000	2.228	109	105
10	*Halodule wrightii*	1.000	2.228	109	105
11	*Syringodium filiforme*	1.000	2.228	109	105
12	Roots	1.000	2.228	109	105
13	Drift Algae	1.000	2.228	109	105
14	Epiphytes	1.000	2.228	109	105
	Microfauna				
15	Free Bacteria	2.000	4.382	100	78
16	Water Flagellates	2.092	2.671	98	102
17	Water Ciliates	2.555	2.009	75	120
18	*Acartia tonsa*	2.517	3.001	79	96
19	*Oithona nana*	2.517	3.000	77	98
20	*Paracalanus crassirostris*	2.517	3.000	78	97
21	Other Copepoda	2.507	3.048	81	92
22	Meroplankton	3.484	5.055	44	72

Continued on next page

Continued from previous page

23	Other Zooplankton	2.531	3.003	76	95
24	Benthic Flagellates	2.000	2.536	100	103
25	Benthic Ciliates	2.450	2.306	84	104
26	Meiofauna	2.317	2.119	91	119
	Macroinvertebrates				
27	Sponges	2.140	3.461	96	83
28	Coral	3.293	5.910	61	68
29	Other Cnidaria	3.341	7.955	57	49
30	Echinodermata	2.375	4.588	87	77
31	Bivalves	2.199	3.418	95	84
32	Detritivorous Gastropods	2.024	3.102	99	89
33	Epiphyte Grazing Gastropods	2.000	4.975	100	73
34	Predatory Gastropods	3.415	11.480	53	6
35	Detritivorous Polychaetes	2.324	3.051	90	91
36	Predatory Polychaetes	3.154	4.705	69	76
37	Pelagic Feeding Polychaetes	2.201	3.294	94	85
38	Macrobenthos	2.476	3.031	83	94
39	Benthic Crustaceans	2.514	2.993	80	100
40	Detrivorous Amphipods	2.491	2.994	82	99
41	Herbivorous Amphipods	2.000	3.035	100	93
42	Isopods	2.000	3.177	100	87
43	Herbivorous Shrimp	2.000	4.848	100	75
44	Predaceous Shrimp	3.296	7.401	60	58
45	Pink Shrimp	2.390	1.920	86	121
46	*Thor floridanus*	2.000	4.975	100	73
47	Spiny Lobster	3.262	6.467	64	64
48	Detritivorous Crabs	2.332	3.576	89	82
49	Omnivorous Crabs	2.342	2.883	88	101
50	Predatory Crabs	3.379	10.172	55	22
51	*Callinectes* spp.	3.379	10.169	54	23
52	Stone Crab	3.828	6.911	26	61
	Fishes				
53	Sharks	4.521	14.759	3	1
54	Rays	3.457	8.494	48	45
55	Tarpon and Ladyfish	3.962	11.220	21	9
56	Bonefish	3.364	11.093	56	13
57	Sardines	3.314	7.544	58	55
58	Anchovies	3.525	7.116	40	60
59	Bay Anchovy	3.566	5.837	38	69
60	Lizardfish	3.887	9.618	24	29
61	Catfishes	3.466	8.592	46	42
62	Eels	3.905	10.669	23	16
63	Toadfish	3.669	9.562	30	31
64	Brotulas and Batfishes	3.494	9.228	43	40
65	Halfbeaks and Flyingfish	2.140	3.101	97	90
66	Needlefishes	4.012	10.734	20	14
67	Killifishes	3.151	7.579	70	53
68	*Floridychthys carpio*	3.447	7.506	49	57
69	*Lucania parva*	3.223	7.693	68	51
70	Snooks	3.887	9.895	25	25
71	Poecilids	2.000	3.933	100	80
72	Silversides	3.603	6.218	35	67
73	Seahorses and Pipefishes	3.287	9.170	62	41
74	*Sygnathus scovelli*	3.302	7.723	59	50
75	*Hippocampus zosterae*	3.257	9.488	65	33
76	Groupers	3.961	9.699	22	27
77	Jacks and Runners	4.029	9.637	19	28
78	Pompano and Permits	3.422	7.601	52	52
79	Snappers	3.619	9.331	32	38
80	Gray Snapper	3.772	9.462	28	34

Continued on next page

Continued from previous page

81	Mojarras and Jennies	3.443	8.047	51	48
82	Grunts	3.444	6.408	50	66
83	Porgies	3.277	7.574	63	54
84	Pinfish	3.073	6.780	71	62
85	Sciaenid fishes	3.619	7.336	33	59
86	Spotted Seatrout	4.054	9.398	18	36
87	Red Drum	3.535	9.547	39	32
88	Spadefish	3.577	9.264	37	39
89	Parrotfishes	2.705	4.225	74	79
90	Mackerels	4.555	12.750	2	2
91	Mullets	2.205	1.836	92	122
92	Barracudas	4.422	11.964	6	5
93	Blennies	3.256	9.383	66	37
94	Code Goby	3.510	6.711	41	63
95	Clown Goby	3.506	6.416	42	65
96	Flatfishes	3.633	7.525	31	56
97	Filefishes and Trigger fishes	3.475	9.448	45	35
98	Puffers and Burrfishes	3.600	9.584	36	30
99	Other Pelagics	4.482	12.317	4	3
100	Other Demersals	2.865	5.314	73	71
	Avifauna				
101	Loons	4.311	10.729	10	15
102	Grebes	4.289	10.356	12	21
103	Pelicans	4.387	11.135	7	12
104	Cormorants	4.335	11.181	9	11
105	Big Herons and Egrets	4.258	10.075	14	24
106	Small Herons and Egrets	4.186	10.447	16	19
107	Ibis	4.078	9.866	17	26
108	Roseate Spoonbill	4.197	10.444	15	20
109	Herbivorous Ducks	2.399	3.796	85	81
110	Omnivorous Ducks	2.918	5.491	72	70
111	Predaceous Ducks	4.302	10.630	11	17
112	Raptors	4.599	12.007	1	4
113	Gruiformes	3.741	8.515	29	44
114	Small Shorebirdss	3.610	8.098	34	47
115	Gulls and Terns	3.798	8.591	27	43
116	Kingfishers	4.280	11.192	13	10
	Herpetofauna				
117	*Crocodylus acutus*	4.357	11.413	8	7
118	*Caretta caretta*	3.466	8.136	47	46
119	*Chelonia mydas*	2.000	3.152	100	88
120	*Eretmochelys imbricata*	3.235	10.587	67	18
	Mammals				
121	*Tursiops truncatus*	4.424	11.261	5	8
122	*Trichechus manatus*	2.203	3.289	93	86

Table 6.1: Results for Florida Bay (wet season) compartments: trophic position (TP), average mutual information ($AMI_{\cdot j}$) and compartment ranking for both indices.

At the lowest trophic position, as expected, ranked the 14 primary producers ($TP = 1$). Minimum values of $AMI_{\cdot j}$ are exhibited by mullets (1.836), pink shrimp (1.920), water ciliates (2.009) and meiofauna (2.119). The highest, both in the TP and $AMI_{\cdot j}$ ranking, are performed by sharks, mackerels, other pelagic fishes and raptors. Nodes called predatory gastropods and tarpon and ladyfish (*Megalops atlanticus* and *Elops saurus*) show high $AMI_{\cdot j}$ values (they rank 6 and 9, respectively) but are not included in the first 20 highest TP values (positions 21 and 53, respectively).

Besides these latter nodes, the general tendency is that of a linear increase in $AMI_{.j}$ with TP, according to the Pearson's product-moment correlation ($\rho = 0.896$, $p \ll 0.05$).
The same trend observed for the whole network appeared when indices are plotted separately for the 6 subgroups (see Figure 6.2). Although the pattern is preserved, the correlation increased from microfauna ($\rho = 0.487$, $p = 0.109$) to avifauna ($\rho = 0.993$, $p \ll 0.001$).

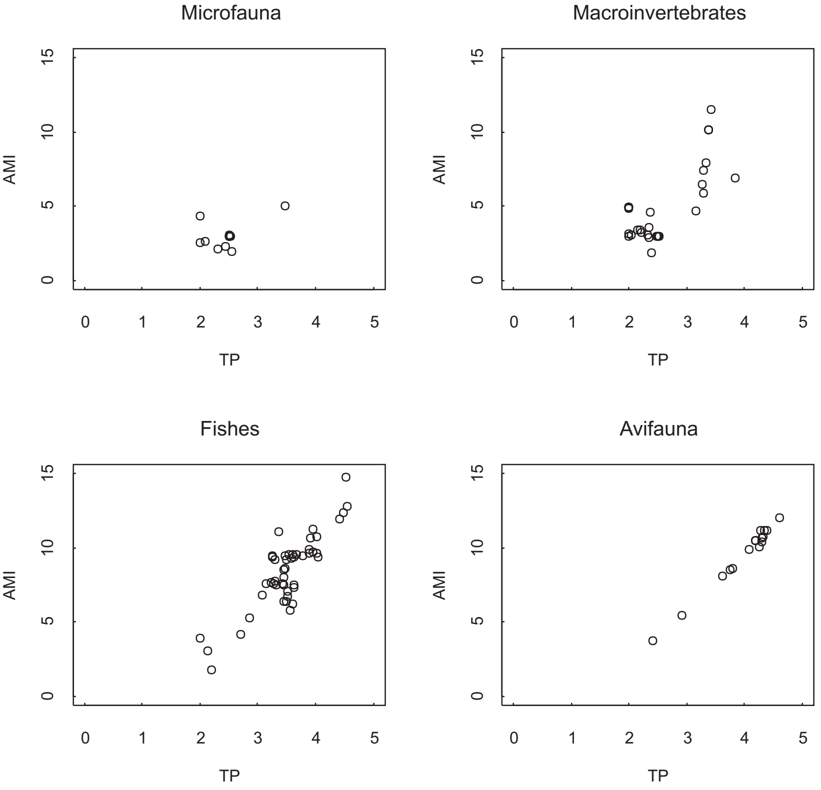


Figure 6.2: Florida Bay (wet season) data of TP and $AMI_{.j}$ are plotted into separated graphs, according to taxonomic groups. Correlation coefficient for microfauna is the lowest($\rho = 0.487$, $p = 0.109$) and increases for macroinvertebrates ($\rho = 0.772$, $p \ll 0.001$), fishes ($\rho = 0.862$, $p \ll 0.001$) and avifauna ($\rho = 0.993$, $p \ll 0.001$).

Positive correlation between TP and $AMI_{.j}$ observed for Florida Bay (wet season) is confirmed in 25 ecosystem networks out of 26 ($p \ll 0.05$), with the exception of Charca de Maspalomas ($\rho = 0.327$, $p = 0.185$; see ρ_Q in Table 5.2). Most (68%) of the correlation coefficients are greater than 0.7, with a mean value equal to 0.735. No significant difference was found between the two groups of networks (ATLSS and NETWRK) as for correlation coefficients ($z = 1.078$, $p \gg 0.05$), although the values for the South Florida networks are less dispersed (see Figure 6.3).

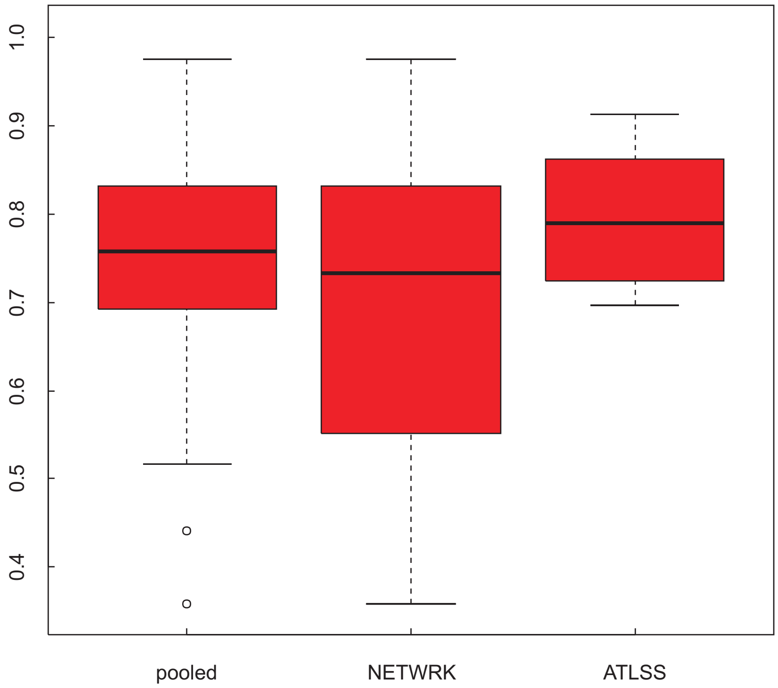


Figure 6.3: Distributions of the significant correlation coefficients as boxplots: pooled (left), NETWRK (centre) and ATLSS (right). Central values for the distributions are (a) "NETWRK" *median* = 0.751 and $\mu = 0.706$; (b) "ATLSS" *median* = 0.791 and $\mu = 0.796$. Inside the boxes, black thick lines show median values

As many as 77 subgroups (15 microfauna, 20 macroinvertebrates, 8 herpetofauna, 18 fishes, 10 avifauna and 6 mammals) originated from the whole database. The rank order correlation coefficients obtained for each group produced the barplot of Figure 6.4. The most controversial results are showed by microfauna, with 10 negative correlations coefficients out of 18. The great majority of coefficients produced by the other groups are positive with the exception of few cases: macrofauna (1), herpetofauna (1), fishes (2) and avifauna (2).

The Kruskal-Wallis test applied to the subgroups was significant ($\chi^2 = 18.969$, $p \ll 0.005$) and microfauna appeared clearly separated from all the other subsets ($median = -0.387$, $\mu = -0.096$). As for the other groups, correlation coefficients for macroinvertebrates ($median = 0.826$, $\mu = 0.741$) were different from those of herpetofauna ($median = 0.724$, $\mu = 0.656$) and avifauna ($median = 0.806$, $\mu = 0.454$), but not from fish. The correlations of the fish group ($median = 0.826$, $\mu = 0.672$) were different from those of herpetofauna and avifauna. The mammals subgroup ($median = 0.857$, $\mu = 0.840$) exhibited correlations not different from those of the other groups but microfauna. However, a tendency of mammals to behave as macroinvertebrates and fishes (differing from herpetofauna and avifauna) can be outlined (see Figure 6.5). Subgroups seem to be split into 3 main clusters (plus mammals) by these results: (a) microfauna; (b) herpetofauna and avifauna; (c) macroinvertebrates and fishes.

Figure 6.4: Barplot showing the rank order of the Pearson's product-moment correlation estimated for microfauna (M, white), macroinvertebrates (K, yellow), herpetofauna (H, grey), fishes (F, blue), avifauna (A, red) and mammals (S, green). The first 16 bars on the left correspond to negative correlation coefficients and microfauna is the prevalent subgroup.

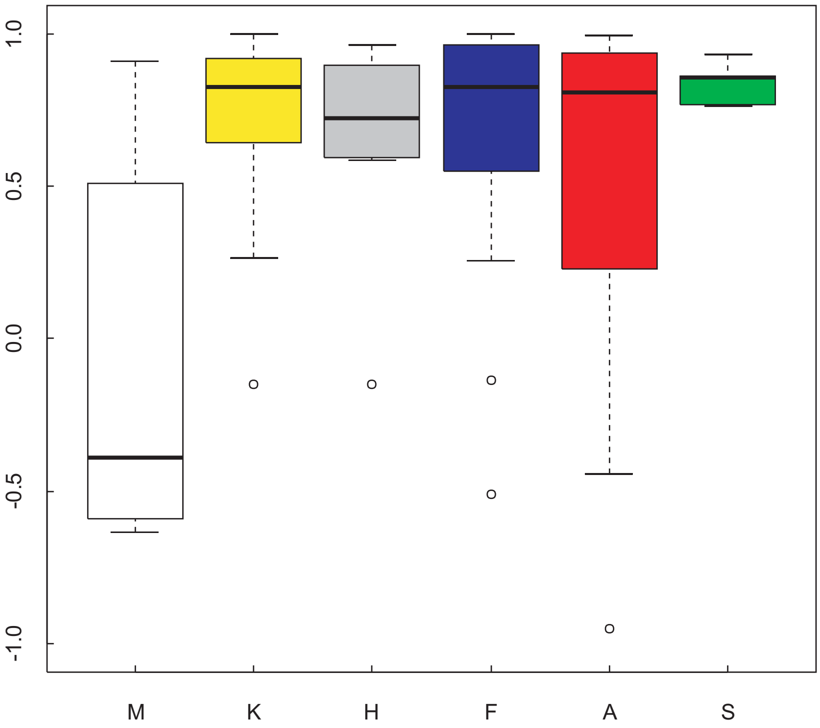


Figure 6.5: Distributions of estimated ρ coefficients for the taxonomic subgroups. Microfauna (M, white) shows more dispersed coefficients and negative median and mean values ($median = -0.387$, $\mu = -0.096$). The other distributions are for macroinvertebrates (K, yellow), fishes (F, blue), herpetofauna (H, grey), dispersed avifauna (A, red), mammals (S, green).

6.4 Energy delivery as a functional issue

In all the networks but one (Charca de Maspalomas ecosystem) $AMI_{.j}$ values increase from primary producers to top consumers. Because this index is inversely proportional to the randomness of flows, that is it scores higher when connections entering a node are few and/or total inflow is unevenly distributed among them, it follows that linkage density, as defined by the $AMI_{\cdot j}$, is attenuated for species that feed higher in the food chain. Thus, our analysis indicates that species feeding at higher position tend to trophic specialization and a more generalist alimentary behaviour characterizes species at the bottom of the trophic hierarchy.

Link density seems to form a pattern around the energy flow from resources to consumers and the trophic hierarchy becomes a key lecture for food web topology. Accordingly, although the concept of trophic level may not be sufficient to explain ecosystem organization, and food web complexity spreads the effects of productivity and consumption throughout the web (Polis and Strong, 1996), such spreading would be constrained by the trophic hierarchy expressed by the gradient of species trophic position.

The metric used for linkage density, the $AMI_{\cdot j}$, includes both the number of links per node and their magnitude. We performed the same analysis described here using a qualitative version of the index (see Chapter 5), that is we considered only the number of connections as a measure of link density (all the links were taken as equal to one).

Plotted against the species trophic position it yielded a negative correlation. This means that by looking solely at the number of links, top species seem to possess more connections than basal ones. However, if community organization has to be considered in relation to energy distribution, what really matters is what species receive as energy contribution and this justifies the need to pass from qualitative to quantitative analysis including interaction strength (Berlow et al., 2004). In this perspective, our results suggest that energy flows towards the top consumers are also unevenly distributed among the links. Evidences accumulated in the literature related to the distribution of interaction strength are mostly in favour of a strong skew towards weaker interactions (McCann et al., 1998; Sala and Graham, 2002; Emmerson and Yearsley, 2004), which our results tend to confirm. However, from what presented here, this distribution shows asymmetry within food webs that seems to be in relation with the gradient of trophic positions.

The distribution of linkage density depends upon a trade-off between specialization and generalization that is a complex matter to investigate. This topic is intriguing and deserves further attention. In particular, it would be interesting studying it in the light of evidences from previous investigations on evolutionary patterns in diet specialization (Sih et al., 1998; Rana et al., 2002). They hypothesized that trophic specialists would have stronger interactions with individual prey species than more generalized species. The results we obtained have indirect implications on the debate on omnivory. The discussion about this alimentary behaviour focussed mainly on its degree of occurrence in ecosystem food webs (Yodzis, 1984; Williams and Martinez, 2004) and it has been fuelled by the difficulty to accommodate it within the findings of local stability analysis (Pimm and Lawton, 1978;

Pimm et al., 1991; Fagan, 1997). Nonetheless, in food web literature there are studies exploring whether omnivory would be more common among taxa at higher or lower trophic positions. We can find evidence for both points in the literature, for example, studies on the Ythan estuary (Hall and Raffaelli, 1991; Raffaelli and Hall, 1996) showed that omnivory should be less common towards the base of a community web because organisms feeding higher would have more opportunity to feed at several trophic levels. On the other hand, Yodzis (1984), by studying the 40 Briand's food webs (Briand, 1983), showed that the number of loop-forming omnivore links (both observed and expected) decreased as the trophic position of the predators augmented. More recently, other studies confirmed that omnivory tends to predominate above the herbivore trophic level (Thompson et al., 2007). All the works cited above analyzed pure qualitative food webs.

For our ecosystems the $AMI_{.j}$ values indicate more randomness of flows, that is higher tendency to omnivory, at the bottom of the trophic hierarchy. Once again quantification of links and the use of $AMI_{.j}$ combining both structural and functional features may reveal patterns about omnivory that would remain hidden in the qualitative approach.

With respect to the correlation between TP and $AMI_{.j}$ certain components are outliers. These nodes show higher (or lower) contribution to the whole system organization (AMI) than expected from the correlation with trophic position. In the case of Florida Bay (wet season), $AMI_{.j}$ of herbivorous shrimp and *Thor floridanus* is, respectively, 4.848 and 4.975. These values are uncommonly high within herbivores, showing how these components are characterized by more specialized trophic strategies (the latter exclusively eats epiphytes whereas herbivorous shrimp diets is based on epiphytes - 89% and benthic microalgae - 11%). In a similar way, the selected number of preys eaten by predatory gastropods and tarpon and ladyfish more than compensate their intermediate efficiency of interactions, resulting in strong contribution to system organization. These results suggest that, although ecologists have been more interested in organisms feeding at the top of food chain than in those feeding at or near the bottom (e.g. all the bulk of theory and experiments related to trophic cascade are based on top species manipulation; Pace et al., 1999; Schmitz et al., 2000), there might be species at intermediate trophic positions that, with their low degree of connections, may represent bottlenecks for the flow of energy from producers to consumers with, possibly, destabilizing effects if removed (Dunne et al., 2004; Allesina and Bodini, 2004).

The extremely high linear correlations within subgroups are consistent with the idea that trophically similar species do not cleanly split along taxonomic lines, such that members of one taxon may occupy different trophic roles in the same ecosystem (Polis, 1984). Indeed, all subgroups but microfauna showed a large range of tropic positions with associated skewed distribution of $AMI_{.j}$ values. The statistical difference separating microfauna (with a majority of negative correlation coefficients) from other subgroups originates from: (a) the ever-coarser degree of node aggregation, and (b) the tendency to occupy a low and narrowed trophic position range. Finally, the three clusters in which subgroups (microfauna; herpetofauna and avifauna; macroinvertebrates and fishes) seem to aggregate are in accordance with well known phylogenetic and evolutionary schemes so that future and more detailed studies should be conducted to demonstrate these patterns.

No scale dependence seems to characterize the relationship between TP and $AMI_{\cdot j}$ which, instead, might be affected by criteria of taxonomic aggregation used to define network compartments. For example, in the Florida Bay (wet season) the same level of resolution was applied to main ecosystem subgroups (microfauna, macroinvertebrates, herpetofauna, fishes and avifauna) and the correlation between the two indices was quite high ($\rho = 0.896$). On the contrary, a much lower correlation ($\rho = 0.358$) was found for the St. Marks river ecosystem. Interestingly enough, different resolution was used in parsing this network: a finer taxonomic resolution was applied to create macroinvertebrate compartments, whereas other groups of species (i.e. benthos-eating birds, fish-eating birds and fish & crustacean eating bird) were assigned to a unique compartment.

More generally, the observed relation is not affected by species richness (see Figure 5.6 in Chapter 5 for trophic networks), as noted for linkage density and food chain length (Schoener, 1989; Cohen and Newman, 1991; Hall and Raffaelli, 1991; Pimm et al., 1991; Martinez, 1991). However, it seems sensitive to a common aggregation scheme within any single network. The question whether the sensibility to trophic and functional rather than taxonomic aggregations should be addressed in future studies, although we hypothesize the correlation should be mainly biased by the latter option because of the skewed distribution of $AMI_{\cdot j}$ with increasing trophic positions. Food web analysis has shown how link topology presents patterns that hold across various ecological systems (Camacho et al., 2002a; Dunne et al., 2002a; Montoya and Solé, 2002; Williams and Martinez, 2004). Because structure and functions in ecosystems are inextricably interlinked, there have been several attempts to understand such patterns in the context of ecosystem stability (MacArthur, 1955; May, 1973; Yodzis, 1981; McCann, 2000). By considering the energy delivery as a functional issue, we highlighted that food web links are distributed in a characteristic pattern along the flow of energy from producers to consumers. This suggests that the trophic hierarchy, delineated by species trophic position, may become a key that contributes reading ecosystem organization, representing not only a heuristic concept. Furthermore, once topological attributes are considered in relation to the different ecosystem functions there will be the opportunity to shed further light on the principles of ecological organization.

7

Trophic positions and keystone species

7.1 Are topological keystone species necessarily top predators?

[1]The focus of conservation biology recently started to shift from identifying rarity of species to characterizing their relative importance (keystone species) and effects on ecosystem functioning.
In this context, studying the complexity of multi-species trophic interactions through the network approach looks like a helpful perspective to identify species roles and importance (e.g. centrality in the network, cf Margalef, 1991).
Since Lindeman (1942) introduced the concept of trophic level, food webs have long interested ecologists. Trophic structure has been commonly studied as food chain length (Wulff and Ulanowicz, 1989; Bondavalli et al., 2006) but other applications used effective trophic position (or trophic height) to infer controlling factors at the whole system level (Christian and Luczkovich, 1999).
Beside trophic analysis, many recent investigations calculated local indices specific to a single species as a function of its topological position (Jordán et al., 2006, 2007). However, emerging patterns are often related to binary

---

[1] “Oral presentation” Chapter: Scotti, M., Jordán, F., 2007. On the relation between centralities and trophic positions in ecological networks. “Marchetti Prize” at the Joint AIOL-SItE meeting, Ancona, Italy.

data (e.g. link distribution; Montoya and Solé, 2002; Dunne et al., 2002b; Abarca-Arenas et al., 2007) and no relations between trophic structure and centralities have been studied.

In the present work we studied 19 ecosystem networks, calculating trophic positions and node centralities both in unweighted and weighted versions. We aimed to clarify the following questions: (a) Are there general patterns linking species feeding behaviour and their relative importance estimated with centrality measures? (b) How are distributed species densities of frequency as regards to centrality and trophic height values? (e.g. do prevail nodes with lowest centrality or are they normally distributed with highest frequencies in corresponding to intermediate values?) (c) Which are the effects of weighting on trophic heights and centrality indices?

Finally, we briefly sketched on consequences that centrality distribution along trophic chain may have on ecosystem functioning.

7.2 TPs and centralities

Data

We analyzed 19 ecological networks based on predator-prey interactions. Data are freely available at the CBL website[2]. In the studied webs, flow intensities are measured as energy (i.e. $kcal\ m^{-2}\ year^{-1}$ or $cal\ cm^{-2}\ year^{-1}$) or matter ($gC\ m^{-2}\ year^{-1}$ or $mgC\ m^{-2}\ summer^{-1}$). The number of compartments range from 21 (e.g. Charca de Maspalomas network) to 125 (e.g. wet version of the Florida Bay network). Table 7.1 provides detailed information on trophic networks.

Methods I - Trophic analysis

Ecological trophic networks describe ecosystems as boxes (species or trophospecies) connected by weighted arrows (the amount of nutrients or energy transferred by feeding relationships).

We applied ecosystem network analysis (ENA; Ulanowicz, 1986), a collection of quantitative methods mapping the intricacy of energy flows. Within this framework we calculated trophic positions (TP) by canonical trophic aggregation (CTA; Ulanowicz and Kemp, 1979; Scotti et al., 2007), a suite of matrix manipulations apportioning every species to a series of discrete trophic levels sensu Lindeman (1942). The effective trophic position is defined as the weighted average length of all the loopless pathways that originate from outside the system and reach a given living compartment (i.e. the average weighted distance between producers and a given node, plus one). Following this approach, primary producers (autotrophs) receiving energy

---

[2] www.cbl.umces.edu, collected mostly by R. E. Ulanowicz and colleagues

Trophic network	S	nl	l	C	**currency**
Charca de Maspalomas	21	18	55	0.125	$mgC\ m^{-2}\ day^{-1}$
Chesapeake Mesohaline Network	36	33	122	0.094	$mgC\ m^{-2}\ summer^{-1}$
Crystal River Creek (control)	21	20	82	0.186	$mgC\ m^{-2}\ day^{-1}$
Crystal River Creek (delta temp.)	21	20	61	0.138	$mgC\ m^{-2}\ day^{-1}$
Lower Chesapeake Bay in Summer	34	31	115	0.099	$mgC\ m^{-2}\ summer^{-1}$
St. Marks River Flow Network	51	48	270	0.104	$mgC\ m^{-2}\ day^{-1}$
Lake Michigan Control Network	36	35	172	0.133	$gC\ m^{-2}\ year^{-1}$
Middle Chesapeake Bay in Summer	34	31	149	0.129	$mgC\ m^{-2}\ summer^{-1}$
Mondego Estuary	43	42	348	0.188	$gAFDW\ m^{-2}\ year^{-1}$
Final Narraganasett Bay Model	32	31	158	0.154	$mgC\ m^{-2}\ year^{-1}$
Upper Chesapeake Bay in Summer	34	31	158	0.137	$mgC\ m^{-2}\ summer^{-1}$
Cypress, Dry Season	68	65	554	0.120	$gC\ m^{-2}\ year^{-1}$
Cypress, Wet Season	68	65	545	0.118	$gC\ m^{-2}\ year^{-1}$
Everglades Graminoids, Dry Season	66	63	793	0.182	$gC\ m^{-2}\ year^{-1}$
Everglades Graminoids, Wet Season	66	63	793	0.182	$gC\ m^{-2}\ year^{-1}$
Florida Bay, Dry Season	125	122	1969	0.126	$gC\ m^{-2}\ year^{-1}$
Florida Bay, Wet Season	125	122	1938	0.124	$gC\ m^{-2}\ year^{-1}$
Mangrove Estuary, Dry Season	94	91	1339	0.152	$gC\ m^{-2}\ year^{-1}$
Mangrove Estuary, Wet Season	94	91	1340	0.152	$gC\ m^{-2}\ year^{-1}$

Table 7.1: For each trophic network we present the number of nodes (S) the number of nodes representing living compartments (nl); the number of links (l); directed connectivity ($C = l/S^2$) and currency used to quantify flow intensities.

from outside the system are set to TP equal to 1, herbivores to 2 and so forth. If we consider an omnivorous species with a fractionary diet based for 10% on a primary producer and for 90% on a herbivore, its TP gives rise from the composition of two pathways: (a) external source $\longrightarrow$ primary producer $\longrightarrow$ omnivore (0.1×2 steps $= 0.2$); (b) external source $\longrightarrow$ primary producer $\longrightarrow$ herbivore $\longrightarrow$ omnivore (0.9×3 steps $= 2.7$). The final TP of the omnivore equals 2.9 ($0.2 + 2.7 = 2.9$).

Alternative version of the TP is the unweighted counterpart ($unwTP$), estimated using only topological information, without considering weights on trophic links. In this case, total inputs to a node are equally distributed between entering flows (in the previous example, $TP = 2.5$ for the omnivore). Theoretically, also the shortest ($unwTPmin$) and longest ($unwTPmax$) energy pathways can be computed but here we have not used these indices. Details on the four measures of trophic position are described in Figure 7.1 and Table 7.2.

Methods II - Centrality indices

Although trophic position is a property of a species in a directed trophic network (representing flow of energy from producers to consumers), in order to better understand the ecological role a species plays in a community we may be interested also in the undirected network of interactions. Doing so, we can evaluate the top-down and horizontal interaction structure of species

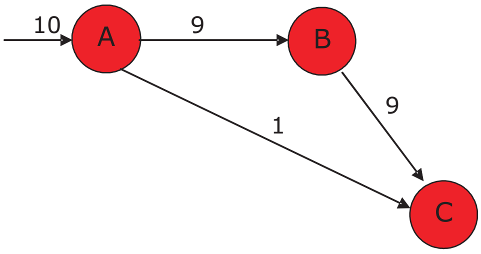


Figure 7.1: Illustration of four different trophic positions of a node in a hypothetical three-compartment network. Compartment A, receiving exclusively energy from outside, is a primary producer ($TP = 1$), while B, feeding only on A, is a herbivore ($TP = 2$). Node C, feeding both on A (10%) and B (90%), have a TP equal to $2.9 = 0.1 \times 2$ (as a herbivore in 10%) $+ 0.9 \times 3$ (as a primary carnivore in 90%). In the unweighted form of TP ($unwTP$), C receives 50% of energy from A and 50% from B, and its $unwTP$ is 2.5 ($0.5 \times 2 + 0.5 \times 3$). Shortest pathway to C is 2 steps (outside ⟶ A ⟶ C) while longest is 3 steps (outside ⟶ A ⟶ B ⟶ C): $unwTPmin = 2$, $unwTPmax = 3$.

Trophic position	path	calculus	value
TP	(9/10) outside ⟶ A ⟶ B ⟶ C (is 90% a primary consumer)	$0.9 \times 3 = 2.7$	2.9
	(1/10) outside ⟶ A ⟶ C (is 10% a herbivore)	$0.1 \times 2 = 0.2$	
$unwTP$	(1/2) outside ⟶ A ⟶ B ⟶ C (is 50% a primary consumer)	$0.5 \times 3 = 1.5$	2.5
	(1/2) outside ⟶ A ⟶ C (is 50% a herbivore)	$0.5 \times 2 = 1$	
$unwTPmin$	outside ⟶ A ⟶ C	2 steps	2
$unwTPmax$	outside ⟶ A ⟶ B ⟶ C	3 steps	3

Table 7.2: Here we summarize, for the node C depicted in Figure 7.1, pathways (path) and calculations (calculus) giving rise to different trophic positions (value), depending on the form of trophic height considered (Trophic position).

(e.g. trophic cascade and apparent competition, respectively). Shortly, energy flows are represented by a digraph, while interspecific interactions must be represented by an undirected graph (alternatively, by a digraph where there are two arrows of different direction between each pair of nodes). There is a wealth of topological indices for characterizing the centrality of nodes in networks. Recently, these have been applied in ecological literature, in order to outline a quantitative context for identifying keystone species (Jordán et al., 1999, 2006, 2007; Estrada, 2007). Different centrality indices characterize different aspects of node centrality, depending on the nature of the network (Vasas and Jordán, 2006). Key nodes can be identified in both directed and undirected, both weighted and unweighted, both signed and unsigned networks, as well as either considering or not indirect effects. Here we used 13 measures of centrality as follows: degree centrality (D; Wassermann and Faust, 1994), weighted degree centrality (wD; Wassermann and Faust, 1994), betweenness centrality (BC; Wassermann and Faust, 1994), undirected betweenness centrality ($undBC$; Wassermann and Faust, 1994), topological importance index for $n = 1$, 2, 3 and 8 steps (TI^n; Jordán et al., 2003) and its weighted version for the same n values (WI^n; Jordán et al., 2003). Weight of links is considered by wD and WI^n, direction of links is considered by BC and indirect effects are considered by all except for D and wD. For technical details, see also Jordán and Scheuring (2004). Centrality values were normalized setting the maximal value to 1.

Methods III - Comparison of indices

First we plotted each centrality index against the two definitions of trophic height adopted (TP and $unwTP$), both for pooled data and single systems. In this first step we investigated if centrality patterns are associated to particular feeding activities (e.g. are top predators more specialist than other species? Are primary producers more central than omnivores?).

Then, for every network, we ranked nodes and extracted 3 species showing highest values (doing so, we collected the TP values of 57 nodes for each index). With TPs of selected compartments we constructed histograms of density, estimating their functions of distribution along the trophic chain. Fitting goodness were measured by Shapiro-Francia test (W index; Shapiro and Francia, 1972), in case of normal distribution, or by χ^2 otherwise (with H_0: the data follow a specified distribution; H_A: the data do not follow the specified distribution). In this way, we checked if more central nodes, for different definitions of centrality: (a) are associated to basal, intermediate or top species; (b) correspond to a well defined range of TP; (c) display a trend caused by trophic and energetic constraints.

Finally, using pooled data, we depicted histograms of density for all the centralities and trophic heights. In particular, we used histogram of TPs, in comparison to density patterns showed by more central nodes, to test if there are significant differences between the distribution of whole database TPs and that displayed by the 57 nodes with higher centralities. With pooled data we also showed densities of nodes respect to each centrality index, studying their relative distribution in ecosystems (e.g. are more central species rare? Are nodes normally distributed between the whole range of normalized centrality values exhibited?).

7.3 Identifying keystone species

When centrality pooled data were studied respect to trophic heights, no trends were identified. However, dealing with keystone species means to

focus on nodes with higher centralities and we decided to extract, for each ecosystem, subgroups of the 3 most central nodes, identifying their TPs. In Table 7.3 and Figure 7.2 we illustrated this procedure applied to Crystal River Creek "control".
Despite scattered points plotted with pooled data, densities of TPs corresponding to highly central nodes give rise to well defined trends (Figure 7.3): (a) considering wD and D, TP densities decrease with a negative power function but highly represented species are at $1 < TP \leq 2$ in case of wD, while central nodes are at $2 < TP \leq 3$ with the unweighted version; (b) normal distribution well fits TP densities for BC computed with directed and undirected data; (c) TP densities associated to higher TIs are normally distributed and similarly happens for topological importance up to more steps (TI^2 - TI^8); (d) WI does not show any trend of TP density when indirect effects are computed up to shorter pathways (WI^1 and WI^2), approaching the D distribution as indirect effects become longer.

Species	D	$undBC$	TI^1	WI^8	TP
1 - Microphytes	0.417	0.267	0.396	0.945	1.000
3 - Zooplankton	0.583	0.367	0.603	1.000	2.000
4 - Benthic invertebrates	1.000	1.000	1.000	0.677	2.000
5 - Blacktip shark	0.083	0.000	0.065	0.048	3.000
6 - Stingray	0.417	0.075	0.269	0.019	3.862
7 - Striped anchovy	0.167	0.006	0.074	0.002	2.667
8 - Bay anchovy	0.500	0.117	0.280	0.127	3.537
9 - Needlefish	0.583	0.304	0.399	0.051	4.329
10 - Sheepshead killifish	0.417	0.087	0.235	0.014	2.109
11 - Goldspotted killifish	0.500	0.130	0.282	0.035	2.446
12 - Gulf killifish	0.750	0.444	0.543	0.191	3.623
13 - Longnosed killifish	0.250	0.004	0.110	0.160	3.000
14 - Silverside	0.583	0.212	0.345	0.528	2.937
15 - Moharra	0.500	0.117	0.280	0.145	2.859
16 - Silver jenny	0.167	0.006	0.074	0.003	2.846
17 - Sheepshead	0.167	0.012	0.093	0.007	2.500
18 - Pinfish	0.667	0.186	0.432	0.077	4.052
19 - Mullet	0.417	0.512	0.541	0.204	2.000
20 - Gulf flounder	0.333	0.013	0.197	0.004	4.455

Table 7.3: Normalized centralities (maximal value = 1) for 19 living nodes in Crystal River Creek "control" network and their TP values. We excluded one node (2 - Macrophytes) because it showed no connections with other living compartments. For each centrality index, the 3 nodes of highest values are highlighted: D in green, $undBC$ in orange, TI^1 in light blue and WI^8 in yellow.

We have also analyzed farness centrality in undirected networks but no clear density distribution of TPs were detected. In particular, negative power functions used with TP densities for more central nodes, in case of wD ($density = TP^{-2.505} + 0.25$) and D ($density = TP^{-1.25} - 1.830$), are significant: (a) wD, $\chi^2 = 4$ with $p = 0.26$; (b) D, $\chi^2 = 6.7$ with $p = 0.35$. Normal distributions observed for TP histograms of density corresponding to nodes with highest betwenness and topological importance centralities,

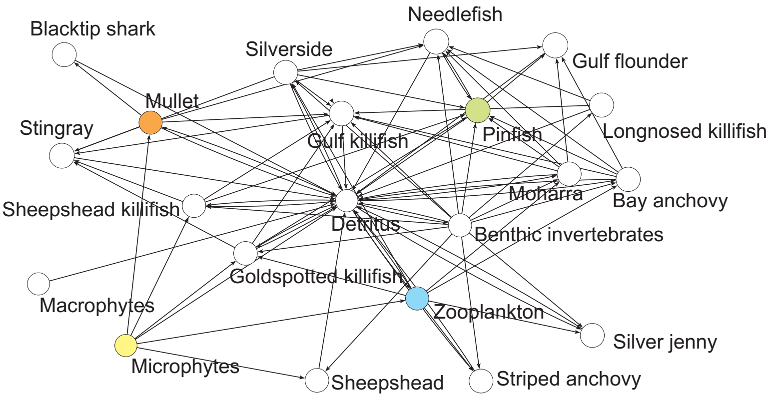


Figure 7.2: Crystal River Creek "control", 19 nodes. Nodes with a single color in Table 7.3 are colored; the figure helps understanding the indices: D in green (many neighbours) = 18 - Pinfish; $undBC$ in orange (many exclusive shortest pathways crossing) = 19 - Mullet; TI^1 in light blue (many low-degree neighbours) = 3 - Zooplankton; WI^8 in yellow (evidently big flows down there) = 1 - Macrophytes.

were always significant: BC ($W = 0.963$, $p = 0.08$), $undBC$ ($W = 0.960$, $p = 0.06$), TI^1 ($W = 0.960$, $p = 0.06$), TI^2 ($W = 0.961$, $p = 0.06$) and TI^8 ($W = 0.960$, $p = 0.06$). It is interesting to notice how central nodes do not behave as others, showing different relative densities along the trophic chain, respect to the whole number of species (Figure 7.4).

We repeated the same studies, plotting centralities against the unweighted version of TP ($unwTP$), focussing on the importance of weighting link in trophic analysis. In Figure 7.5 histograms of $unwTP$ densities are depicted for the same indices illustrated in Figure 7.3. In general, considering topology structure without information on link strength, we observed that: (a) negative power function characterizing degree centralities is preserved with wD, becoming more confusing in D; nonetheless, more representative trophic height interval remain $1 < unwTP \leq 2$ for the first and $2 < unwTP \leq 3$ for the latter; (b) normal distribution of TP densities is maintained by $undBC$ and lost in case of BC, with most central nodes confined to the interval $2 < unwTP \leq 3.5$; (c) with topological importance up to n steps ($n = 1, ..., 8$), $unwTP$ frequency distribution is normal, as with TP, suggesting scarce effects of weighting connections; (d) weighted topological importance tends to be equally distributed when computed on few steps (WI^1 - WI^3), while unwTPs = 2.5 are more central in case of longer steps (WI^8); also in this case, tendencies displayed by TPs are smoothed.

Finally, we used normalized pooled data to highlight which are frequency distributions of nodes, using different definitions of centrality: (a) in case of D, intermediate centralities tend to prevail; (b) wD, BC and $undBC$ showed a huge number of nodes with extremely low centralities and the few remaining compartments with higher values, suggesting that key nodes

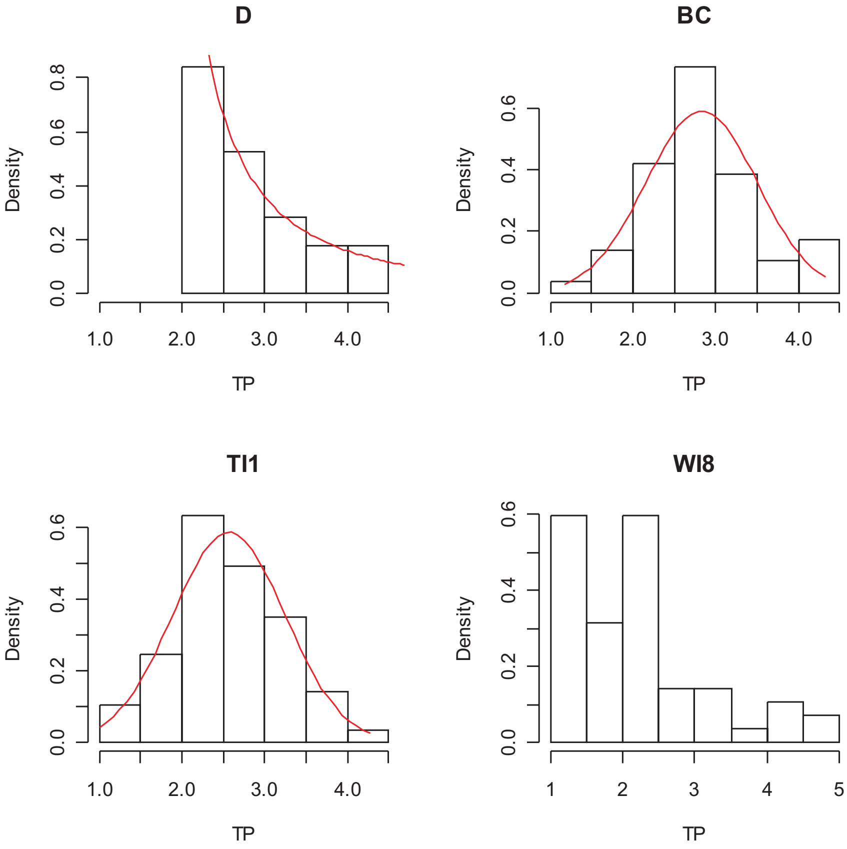


Figure 7.3: Distribution of densities for the trophic positions (TP) associated to the 57 nodes with the 3 highest centralities in the 19 ecosystems studied. Here we illustrated trends for unweighted degree centrality (D), betweenness centrality computed with directed links (BC), topological importance index up to 1 step (TI^1) and weighted topological importance up to 8 steps (WI^8). Negative power function for D and Gaussian curve describing BC and TI^1 were added in red to histograms.

do not represent the total set of nodes; (c) topological importance indices display a tendency of prevalence for nodes with low-intermediate values; when increasing steps of indirect effects (moving from TI^1 to TI^8), there is a shift, with intermediate and slightly higher centralities more represented; (d) in the case of weighted version of TI, the number of nodes decreases with a negative power law when centrality values are increasing. In Figure 7.6, frequency distributions for D, BC, TI^1 and WI^8 are shown.

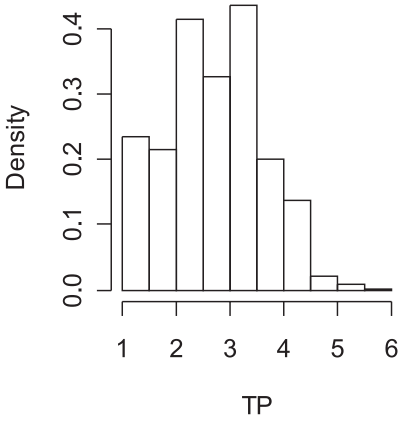


Figure 7.4: Relative distribution of all of the 1001 nodes in the 19 networks, in classes of trophic position.

7.4 Species feeding behaviour and centralities

In general, we observed no regularities when centrality indices are studied as a function of trophic position (both TP and $unwTP$), for the whole number of food web species. However, theoretical ecology and conservation biology aim to identify the most important species in ecosystems (Paine, 1969; Mills et al., 1993) and Jordán et al. (1999) and Jordán (2001) demonstrated how network perspective and centrality measures can be fruitfully adopted in this sense. We analyzed TP trends associated to more central nodes, unveiling interesting features. As suggested by Scotti et al. (2007), weighting links is particularly important in food webs for direct centrality indices, and this is confirmed by the present work: (a) negative power distribution describing TP densities of more central nodes is preserved, switching from the weighted to unweighted version of trophic height, in presence of wD but not with D; (b) the most representative TP classes of density are different using the nodes with highest degree centralities ($1 < TP \leq 2$ in case of wD and $2 < TP \leq 3$ with D), while they are the same with indirect indices; (c) trends for histograms associated to TI are the same with TPs and unwTPs. Still, at $TP = 1$ there is a prominent number of central nodes for wD but none for D. This reflects a small number of huge flows resulted from aggregation. Moreover, higher species are not very central in the D case (for wD it is not surprising because of the small flows up there; see Hairston Jr. and Hairston Sr., 1993), suggesting more constrained pathways characterizing top predators (they are more specialized

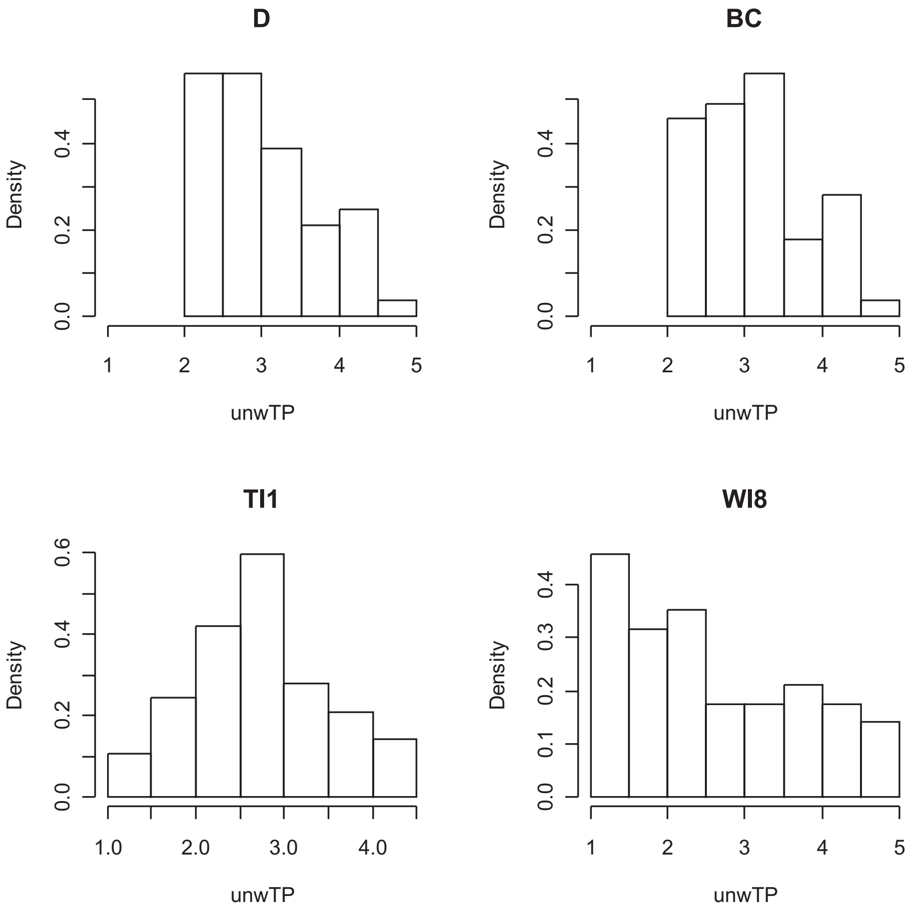


Figure 7.5: The same histograms as in Figure 7.3 but for unweighted trophic position (*unwTP*) densities. Trends are similar to what displayed by TP densities for TI and WI, while the negative power function for D and the normal distribution for BC disappeared.

than highly connected intermediate nodes). The majority of nodes with highest unweighted degree centralities have TPs included between 2 and 3 (both TP and $unwTP$) and this is consistent with trend of patterns exhibited by wasp-waist ecosystems (Jordán et al., 2005). Beside evidences on weighting, histograms estimated for BC and $undBC$ display how direction is not so important when we use weighted data for trophic heights. We also emphasized how topological importance tends to D with increasing number of steps up to which is estimated, while weighted importance approaches the wD distribution as indirect effects become longer (see Figure 7.6). Our study contradicts the common idea identifying charismatic megafauna and top predators as the main objectives of conservation ecology, setting many basal and intermediate nodes as "keystone species" in natural systems. Finally, we propose that protection activities should focus their attention both on central and rare species, considering that TP distribution of key nodes do not follow trends exhibited by the whole food web species.

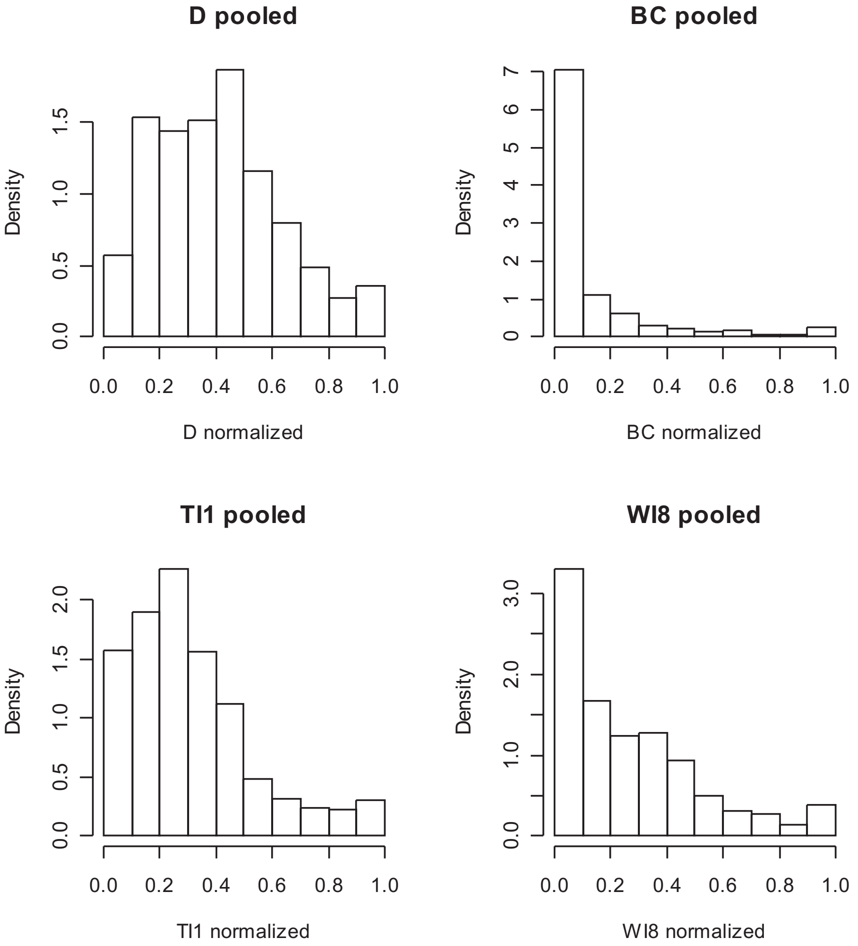


Figure 7.6: Relative distribution of densities when D, BC, TI^1 and WI^8 are calculated for the 1001 living nodes included into the 19 ecosystems studied. While WI^8 and BC are centralities with few extremely central nodes and many less important, D displays a prevalence of compartments with intermediate values and TI tends to a normal distribution with increasing number of steps up to which is measured.

8

Centralities in ecological networks

8.1 Network perspective and keystone species

[1] Conservation biology is being shifted from protecting species to protecting interspecific interactions and communities. In order to better understand the nature of interaction networks, we need comparative analysis of different interaction types. Ecological complexity comprises the diversity of both species and interspecific interactions. Different types of interactions, such as predator-prey or plant-pollinator interactions are of different character in their ecology and evolution (Thompson, 1991). Since the majority of ecological networks studied so far are food webs (or trophic networks), we should re-examine many classical questions for other network types as well. These basic problems include the importance of weighting, the relevance of indirect interactions and the scale dependence of network properties.

The systematic analysis of ecological networks involves three steps: (a) data collection, (b) network construction and (c) network analysis *sensu stricto*. A number of problems are relevant only to one of these steps, while others bridge over the whole process. The mostly practical question whether and how to consider weights on links (Ulanowicz, 1986; Baird and Ulanowicz, 1993; Paine, 1980) concerns step 1. The problems of aggregation, network resolution and scale dependence (Martinez, 1991; Allesina and Bodini,

---

[1] Published Chapter: Scotti, M., Podani, J., Jordán, F., 2007. Weighting, scale dependence and indirect effects in ecological networks: A comparative study. Ecological Complexity, 4(3):148-159.

2005; Allesina et al., 2005b) concern step 2. Finally, a possibly more technical question whether to neglect or explicitly study indirect interactions spreading over these networks (Menge, 1995; Wootton, 1994) concerns step 3. Each problem has a long history and has been discussed by a number of authors, and is typically investigated separately (but see Vasas and Jordán, 2006). In this paper, we demonstrate that these questions are intimately related.

We studied the order of importance of nodes (corresponding mostly to species) in 53 ecological interaction networks. First, topological importance is quantified by degree centrality which measures the number of neighbours (i.e. interacting species; see Jordán et al., 1999; Dunne, 2006; Dunne et al., 2002b). Then, we compared these orders to others obtained by considering (a) weights on links, (b) indirect effects and (c) both (see Figure 8.1, for our scenario). Our questions are: (a) how important is to consider weighting from the viewpoint of the centrality rank, (b) how important is to consider indirect effects (and up to what length), (c) what is the relationship between these approaches, (d) whether network size influences answers to the above questions, and (e) what is the difference between various types of networks? Our broader interest is to investigate how the results are related to the evolutionary stability of these interaction types (Thompson, 1982, 1991).

8.2 Ecological networks

8.2.1 Data

We analyzed highly standardized data on ecological networks. The ecological networks collected, freely available from the NCEAS data base (*www.nceas.ucsb.edu/interactionweb*), are fairly homogeneous methodologically. We studied all of the weighted webs therein (except for "kat", for technical reasons). Since all plant-herbivore networks and food webs are of binary nature in that source, so that none of them could be used in this study, we decided to add 20 weighted food webs from another standard source (*www.cbl.umces.edu/∼ulan/*). Finally, we had a total of 53 weighted networks representing 5 host-parasite, 3 plant-ant, 20 plant-pollinator, 5 plant-seed disperser and 20 food webs (Table 8.1 provides the identification number and name of the webs, the number of nodes and the currency for weighting, while Appendix I - 53 ecological networks - summarizes the original reference for each).

In the 5 host-parasite systems, weighted relations imply infection intensity in terms of the average number of parasites per host individual (average parasite load). Analogously, the currency used in plant-ant webs is the number of visits of ant species to a given plant. In case of plant-pollinator networks, the intensity of contacts may be measured as the number of individuals caught, the frequency of visits or the number of visits of each pol-

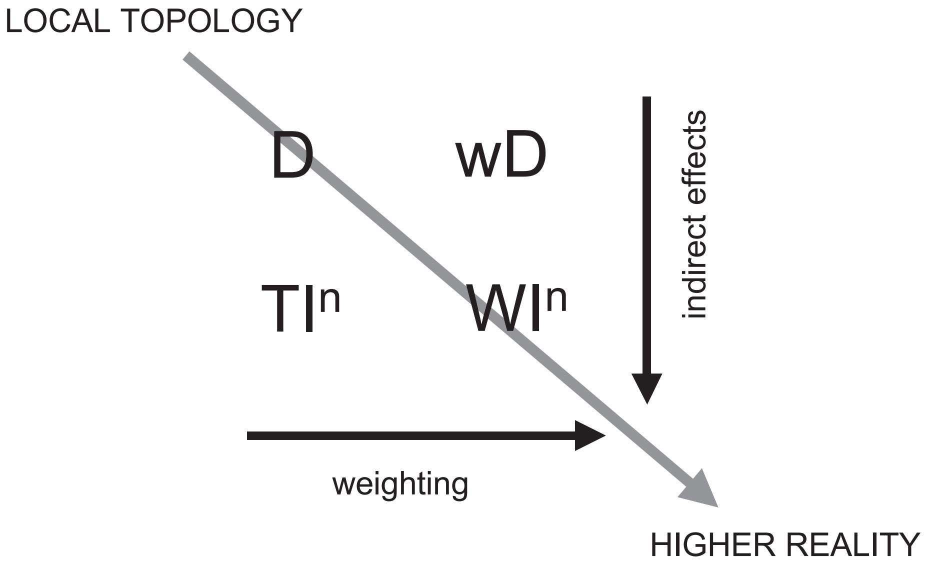


Figure 8.1: Illustration of the logic of the present approach. In the simplest topological case a node is characterized only by the number of its neighbours, i.e. its degree (D). Pure topology was made more realistic by considering either weights on links or indirect effects (up to n steps). The former case can be quantified by the weighted degree of nodes (wD), while the latter is measured by the topological importance index (TI^n). Their combination is a weighted indirect positional measure (WI^n). We can also compare the TI^n and WI^n values for different values of n.

linator to a plant species. One of the networks ("vaz": Evergreen Montane Forest in Argentina) is a plant-pollinator system composed of 103 nodes but split into 8 sub-networks. These 8 subsystems are studied separately and together with the remaining 12 webs constitute the plant-pollinator dataset. The number of visits and fruits removed are the two methods used to weight links in plant-seed disperser systems. The currency commonly adopted in food webs is carbon (expressed as $mgC\ m^{-2}\ day^{-1}$ or $gC\ m^{-2}\ year^{-1}$) or energy flow ($cal\ cm^{-2}\ year^{-1}$ and $kcal\ m^{-2}\ year^{-1}$). For obtaining unweighted webs we simply replaced every non-zero weight by one.

8.2.2 Methods I - Network analysis

Since the ecological networks analyzed are of different size (defined as the number of nodes in the network, representing either species or other functionally relevant units), we made them comparable by calculating normalized values for every measure (i.e. the rank of a node was divided by the number of nodes in the network). We were interested in the centrality rank of nodes in each network for each index, for both the weighted and the

Web#		S	**Currency**
	Host - Parasite webs (5)		
1	Aishihik Lake	36	intensity of infection
2	Cold Lake	50	intensity of infection
3	McGregor River	65	intensity of infection
4	Parnsip River	70	intensity of infection
5	Small Wood Reservoir	31	intensity of infection
	Plant - Ant webs (3)		
6	Rainforest (Peru)	10	# of visits to each plant
7	Tropical Forest (Costa Rica)	24	# of visits to each plant
8	Amazon Rainforest	41	# of visits to each plant
	Plant - Pollinator webs (20)		
9	Boreal Forest (Canada)	114	individuals caught
10	Alpine Subartic Community (Sweden)	141	# of visits to each plant
11	Beech Forest (Japan)	133	individuals caught
12	Montane Forest (Australia)	147	# of visits to each plant
13	Medow (Bristol, U.K.)	104	frequency of visits
14	Arctic Community (Canada)	29	individuals caught
15	Deciduous Forest (U.S.A.)	57	# of visits to each plant
16	Coastal Forest (Mauritius Island)	27	# of visits to each plant
17	Rocky Cliff and Open Herb Community	22	# of visits to each plant
18	Upland Grassland (South Africa)	65	individuals caught
19	Maple-Oak Woodland (U.S.A.)	39	# of visits to each plant
20	Peat Bog (Canada)	47	individuals caught
21-28	Evergreen Montane Forest - 8 subwebs	104 total	# of visits to each plant
	Plant - Seed Disperser webs (5)		
29	Forest (Papa New Guinea)	40	# of visit to each plant
30	Semideciduous Tropical Forest	24	fruits removed
31	Neotropical Forest (Trinidad)	79	# of visit to each plant
32	Neotropical Forest (Trinidad)	48	# of visit to each plant
33	Temperate Woodland (Britain)	25	# of visit to each plant
	Food webs (20)		
34	Aggregated Baltic Sea	15	$mgC\ m^{-2}\ day^{-1}$
35	Cedar Bog Lake	9	$cal\ cm^{-2}\ year^{-1}$
36	Charca de Maspalomas	21	$mgC\ m^{-2}\ day^{-1}$
37	Chesapeake Mesohaline Ecosystem	15	$mgC\ m^{-2}\ day^{-1}$
38	Chesapeake Mesohaline Network	36	$mgC\ m^{-2}\ summer^{-1}$
39	Crystal River Creek (control)	21	$mgC\ m^{-2}\ day^{-1}$
40	Crystal River Creek (delta temp.)	21	$mgC\ m^{-2}\ day^{-1}$
41	Everglades Graminoids (wet season)	66	$gC\ m^{-2}\ year^{-1}$
42	Florida Bay Ecosystem	125	$gC\ m^{-2}\ year^{-1}$
43	Lower Chesapeake Bay in Summer	29	$mgC\ m^{-2}\ summer^{-1}$
44	St. Marks River (Florida) Flow Network	51	$mgC\ m^{-2}\ day^{-1}$
45	Lake Michigan Control Network	34	$gC\ m^{-2}\ year^{-1}$
46	Middle Chesapeake Bay in Summer	32	$mgC\ m^{-2}\ summer^{-1}$
47	Mondego Estuary	43	$gAFDW\ m^{-2}\ year^{-1}$
48	Final Narraganasett Bay Model	32	$mgC\ m^{-2}\ year^{-1}$
49	North Sea	10	$kcal\ m^{-2}\ year^{-1}$
50	Somme Estuary	9	$gC\ m^{-2}\ year^{-1}$
51	Upper Chesapeake Bay in Summer	33	$mgC\ m^{-2}\ summer^{-1}$
52	Upper Chesapeake Bay	12	$gC\ m^{-2}\ year^{-1}$
53	Ythan Estuary	13	$gC\ m^{-2}\ year^{-1}$

Table 8.1: List of the studied webs. Here we summarize the names of the different types of webs, classified into 5 subgroups, depending on the relations they account for (host - parasite; plant - ant; plant - pollinator; plant - seed disperser and food webs). Then, for each system, we present the number of nodes (S) and the currency used.

unweighted case (Jordán et al., 2006, 2007; Estrada, 2007).

The simplest approach takes node degree (D) as the measure of centrality. The rank order of D values for networks provides a very basic quantification of the structural (topological) importance of nodes in the network. Its weighted form (wD) also considers weights on links to neighbours. Weighted degree provides information as to which nodes are characterized by the largest flows (inflows plus outflows) or by the strongest direct effects from and to other nodes (see Figure 8.2, for illustration; Figure 8.3 and 8.4 as actual examples; while Table 8.2 and Appendix II - Aishihik Lake species - provide background information on the network exemplified).

Alternative indices characterizing positional importance of species in unweighted and weighted networks, considering maximum n-step long indirect effects, are TI^n and WI^n, respectively (Jordán et al., 2003). This approach is derived from the analysis of two-step long, horizontal, apparent competition interactions in weighted host-parasitoid networks (Müller et al., 1999). In an unweighted network, we define $a_{n,ij}$ as the effect of species j on species i when i can be reached from j in n steps. The simplest mode of calculating $a_{n,ij}$ is if $n = 1$ (i.e. the effect of j on i in 1 step): $a_{1,ij} = 1/D_i$, where D_i is the degree of species i. When the effect of step n is considered, the effect received by species i from all species in the same network is equal to 1 (i.e. each species is affected by the same unit effect). Furthermore, we define the n-step effect originated from species i by the following formula

$$\sigma_{n,i} = \sum_{j=1}^{S} a_{n,ji} \tag{8.1}$$

where S is the number of species in the network. Effects originated from different species are typically different. Here, we define the topological importance of species i when effects "up to" n steps are considered as follows

$$TI_i^n = \frac{\sum_{m=1}^{n} \sigma_{m,i}}{n} = \frac{\sum_{m=1}^{n} \sum_{j=1}^{S} a_{m,ji}}{n} \tag{8.2}$$

which is simply the sum of effects originated from species i up to n steps $(1+2+3+\ldots+n)$ averaged over the maximum number of steps considered (i.e. n).

In our study, we extend the TI index to directed networks. In this case, we define the direct effect of species j upon i as

$$a_{1,ij} = \frac{\alpha_{ij}}{\sum_{j=1}^{D} \alpha_{ij}} \tag{8.3}$$

where $\alpha_{ij} = 1/D_{i,out}$ if species j is a predator and $\alpha_{ij} = 1/D_{i,in}$ if node j is a prey of i. In the equations $D_{i,in}$ is the number of i's preys, $D_{i,out}$ is the number of i's predators and D_i is their sum, while the other calculations remain the same. The method assumes that the effects each species receives from its predators and preys are equal.

For a weighted network, all the effects are defined in the same way as above with the exception of calculating α_{ij}, which is computed as:

$$\alpha_{ij} = \frac{\epsilon_{ij}}{\mu_i} \tag{8.4}$$

where μ_i is the sum of strength values of links pointing to species i, if species j is a prey of species i or the sum of the strength values of links originated from species i, if species j is a predator of species i. ϵ_{ij} is the strength of the link connecting species i and j.

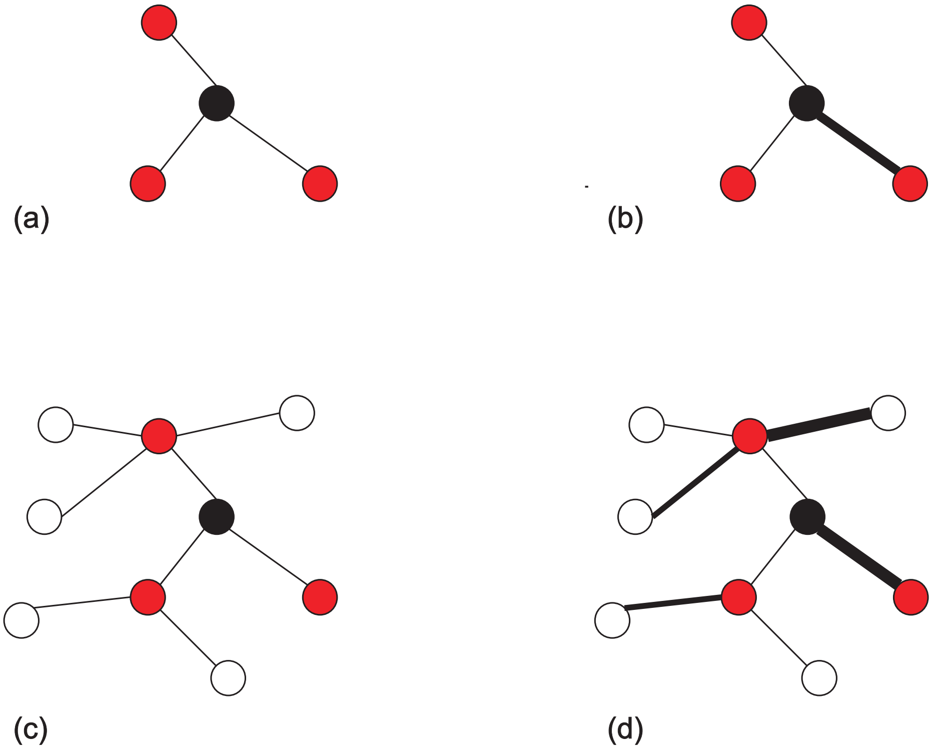


Figure 8.2: Illustration of the calculation of D, wD, TI^n and WI^n (for simplicity, links are weighted by 1, 2 and 3, as shown by the different thickness of edges). In (a), the black node has three red neighbours ($D = 1+1+1 = 3$). In (b), it has links of different strength to its red neighbours ($wD = 1+1+3 = 5$). In (c), we also know the white neighbours of its red neighbours (but forget about the weights; $TI^1 = 1 + 0.25 + 0.33 = 1.58$). In (d), we consider both weighting and indirect effects ($WI^1 = 3/3 + 1/7 + 1/4 = 1.4$).

8.2.3 Methods II - Statistical analysis

We calculated the normalized values of D, wD, TI^1, TI^2, TI^3, TI^4, WI^1, WI^2, WI^3 and WI^4 for all of the nodes in the 53 webs. Then, we ranked the nodes according to each normalized index in each web. We compared weighted and unweighted ranks, direct and indirect ranks, as well as indirect ranks of different lengths in case of every web. Also, we compared the five types of networks.

Weighted and unweighted node rank orders for each network, based on each index, were compared by the Goodman-Kruskal lambda (Goodman and Kruskal, 1954; Podani, 2000). Based on each index, network types were compared by the Kruskal-Wallis test (Sokal and Rohlf, 1981, pg. 429) to see whether the types differ significantly in their average ranks. The performance of the five indices was evaluated by standardized Principal Components Analysis (PCA), performed by SYN-TAX 2000 (Podani, 2000, 2001). In this analysis, the starting data matrix contained 53 columns, representing networks, and 5 rows, each representing the rank correlations between the networks based on the weighted and unweighted versions of a given index.

The scale dependence of lambda was evaluated graphically, by plotting lambda against the number of nodes, such that linear regression lines were added to enhance visual interpretation.

8.3 Central nodes in ecological networks

8.3.1 Lake Aishihik: a case study

The 7 hosts and 29 parasites in the Lake Aishihik community are of different positional importance in the interaction network. The simplest approach to quantifying this is based on degree (Figure 8.3a). This reveals that the structural key species include three species of highest degree ($D = 16$ for species #5, #6 and #7), while we may still mention species #2 ($D = 15$), #4 ($D = 12$) and #3 ($D = 10$; the D values are presented in the first column of Table 8.2 and species names are given in Appendix II - Aishihik Lake species). Some species have only a single neighbour (like species #9). If we consider weights on links (reflecting intensity of infection, Table 8.1), then a different rank order is provided (second column in Table 8.2, see also Figure 8.3b showing weights on links). In the weighted network, some species are of higher importance than before (e.g. species #12 and #16), while others are of lower importance (e.g. species #5 and #7). The topological dominance of hosts (species from #1 to #7) has been changed in the weighted network.

If indirect effects up to three steps are considered rather than weights, a new importance rank is provided (third column in Table 8.2, see also Figure 8.4a). Species #4 is of higher importance than species #2, compared to the direct unweighted case, but otherwise the two rank orders are very similar in case of this web. If indirect effects are considered and the web is weighted, the most realistic results are obtained (last column of Table 8.2, see also Figure 8.4b). This rank order combines the effects of the two

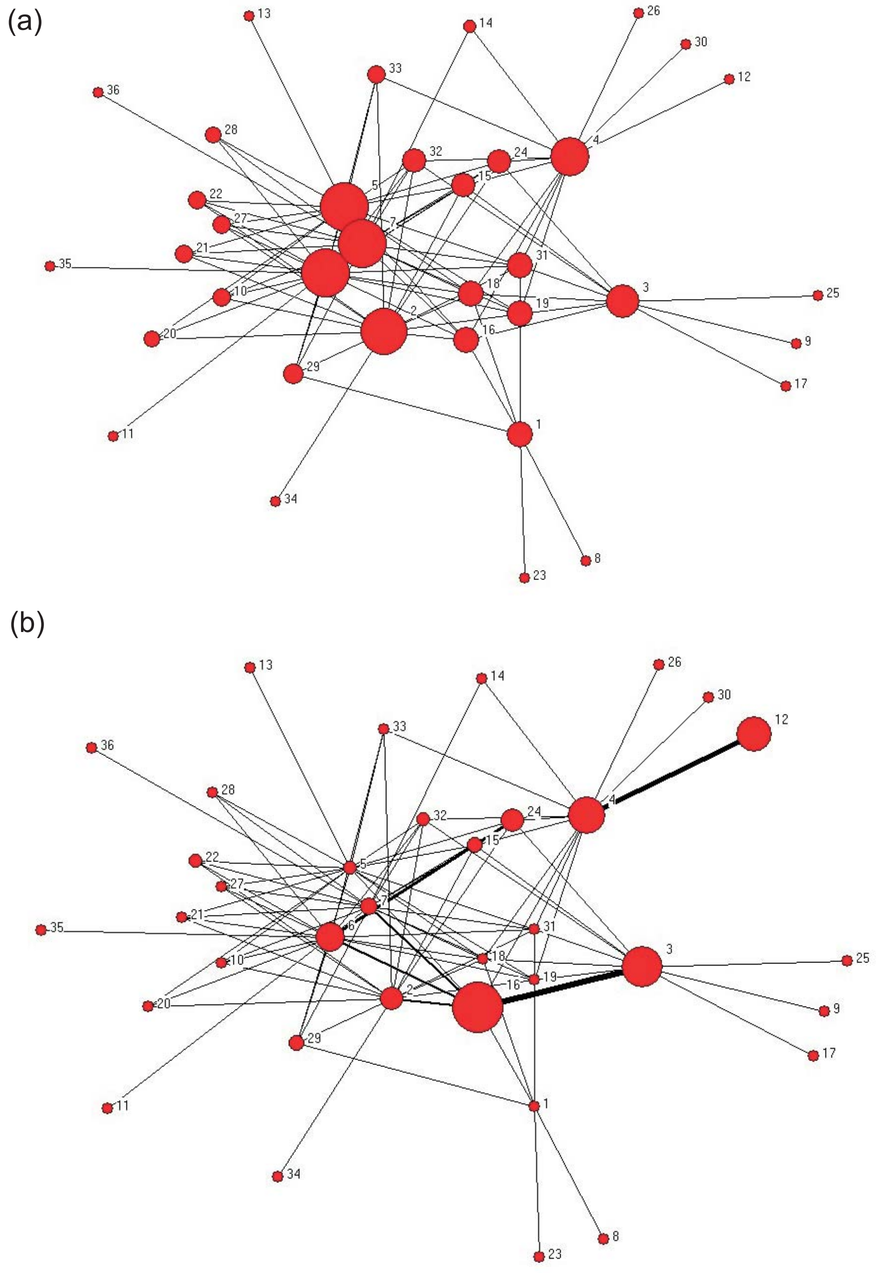


Figure 8.3: The Lake Aishihik host-parasite interaction network, containing 36 species (7 hosts and 29 parasites). Figure (a) shows the unweighted (binary) web and the size of nodes is proportional to the value of node degree D. Figure (b) shows the same web weighted by intensity: the size of nodes is proportional to the value of weighted node degree wD. All webs are drawn by UCINET (Borgatti et al., 2002).

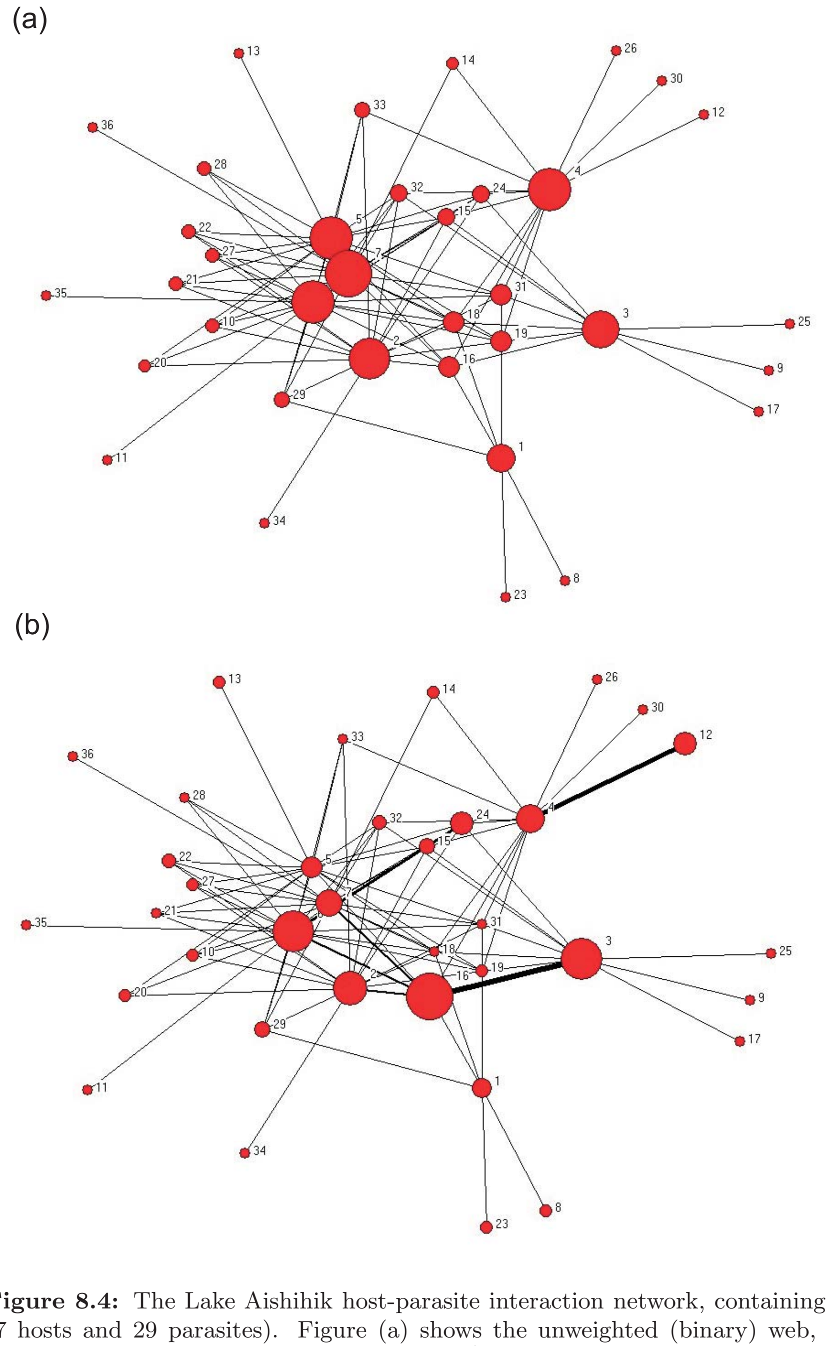


Figure 8.4: The Lake Aishihik host-parasite interaction network, containing 36 species (7 hosts and 29 parasites). Figure (a) shows the unweighted (binary) web, the size of nodes is proportional to the value of the TI^3 topological importance index. Figure (b) shows the web weighted by intensity: the size of nodes is proportional to the value of the WI^3 topological importance index. All webs are drawn by UCINET (Borgatti et al., 2002).

node	D	**node**	wD	**node**	TI^3	**node**	WI^3
5	16	16	1464.7	7	11.20	16	16.58
6	16	3	1079.4	5	10.10	6	14.50
7	16	4	924.7	6	10.10	3	14.23
2	15	12	879.6	4	9.90	2	10.85
4	12	6	681.9	2	9.42	4	8.60
3	10	2	420.9	3	8.41	7	7.68
1	7	24	417.1	1	5.88	24	5.75
16	7	7	211.0	16	3.43	12	5.46
19	7	15	152.9	18	3.43	5	5.22
31	7	29	138.1	19	3.43	1	3.92
18	7	5	80.8	31	3.43	15	2.71
24	6	32	77.6	15	2.80	29	2.42
15	6	22	55.0	24	2.80	32	1.36
32	6	20	25.4	32	2.80	22	1.34
29	5	10	25.2	29	2.28	13	1.18
10	4	26	24.7	33	1.75	8	1.00
22	4	27	22.3	10	1.64	10	0.72
27	4	13	20.5	21	1.64	23	0.60
21	4	1	19.8	22	1.64	14	0.47
33	4	14	16.6	27	1.64	19	0.45
28	3	9	12.1	28	1.24	20	0.44
20	3	21	11.2	20	1.19	27	0.43
14	2	19	11.0	14	1.02	31	0.40
26	1	31	10.0	8	0.64	18	0.34
11	1	11	8.9	23	0.64	21	0.33
25	1	18	8.0	9	0.59	11	0.24
17	1	28	6.7	17	0.59	36	0.18
34	1	36	6.3	25	0.59	28	0.16
35	1	8	5.8	12	0.56	33	0.15
36	1	33	5.8	26	0.56	26	0.15
12	1	30	4.1	30	0.56	9	0.08
13	1	23	3.0	11	0.45	30	0.02
30	1	17	2.0	36	0.45	34	0.02
23	1	35	1.9	34	0.40	17	0.01
8	1	34	1.0	13	0.40	35	0.01
9	1	25	1.0	35	0.40	25	0.01

Table 8.2: Positional importance ranks of nodes in the Lake Aishihik network. We characterized positional importance of nodes by node degree (D), weighted node degree (wD), the topological importance index for indirect effects up to three steps (TI^3) and its weighted form (WI^3).

different approaches. Species #16 is of the highest importance because of its weights, while species #6 is ranked high because of the number of its neighbours. We performed the same calculations for all of the 53 webs, and studied also the pooled results for the five types of networks, beyond comparing individual webs.

8.3.2 General results for 53 networks

Weighting always affects the ranking very seriously: the lambda measure of rank correlation is typically near zero or strongly negative for node orderings based on the weighted and unweighted forms of the same index. Positive correlations were very rare (10 out of 265, all insignificant). The effect of weighting is the largest for food webs and the smallest for plant-pollinator networks.

The rank order of 53 lambda values for five pairs of coefficients is visualized by the barplots of Figure 8.5. Based on D and wD, food webs are significantly different from the other network types ($\chi^2 = 36$, $p \ll 0.05$), as also depicted by the first row of Figure 8.5.

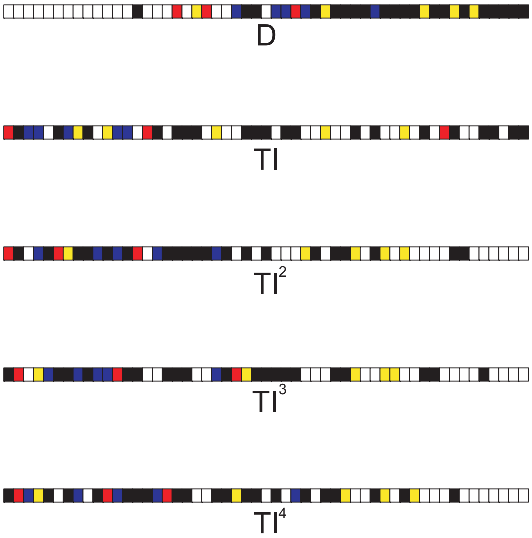


Figure 8.5: Bar plots showing the rank order of 53 networks according to the difference between unweighted and weighted centrality indices. We present results for five pairs of network indices (from the top: D and wD, TI^1 and WI^1, TI^2 and WI^2, TI^3 and WI^3 and TI^4 and WI^4). Network types are visualized by colours: white = food web; black = plant-pollinator; red = plant-ant; yellow = plant-seed disperser; blue = host-parasitoid. In webs on the left, the difference between the weighted and unweighted centrality ranks is larger, i.e. in webs on the right weighting matters less. From the top down, the length of indirect effects increases: although weighting is very important in evaluating direct interactions in food webs (white squares are on the left in the top bar), it is much less important for long indirect effects (they move to the right).

Comparing TI^1 and WI^1 provides weaker results ($\chi^2 = 10,\ p < 0.05$), such that there is no particular network type separated from all others although a notable difference between food webs and host-parasite networks is observed. Based on TI^2 and WI^2, there is significant difference between food webs and host-parasite networks and food webs and plant-ant networks ($\chi^2 = 19,\ p < 0.05$). Based on TI^3 and WI^3, food webs are significantly different from the others, although this separation is weaker than for D ($\chi^2 = 15,\ p < 0.05$). Finally, based on TI^4 and WI^4, there is significant difference between network types ($\chi^2 = 16,\ p < 0.05$).
The first two PCA dimensions represent an overwhelming majority of variation in the lambda values, i.e. 70% and 21%, respectively. The biplot (Figure 8.6) shows that the pair of local indices (D and wD) differs orthogonally from the indirect indices (TI^1 and WI^1, TI^2 and WI^2, TI^3 and WI^3 and TI^4 and WI^4). This means that considering indirect effects influences greatly the effect of weighting. Among the indirect indices, the one-step long index (TI^1 and WI^1) is separated a little bit while the others (TI^2 and WI^2, TI^3 and WI^3, TI^4 and WI^4) behave very similarly, but the overall similarity of all indirect indices is clear and there is practically nothing new after two steps (see also Table 8.3).

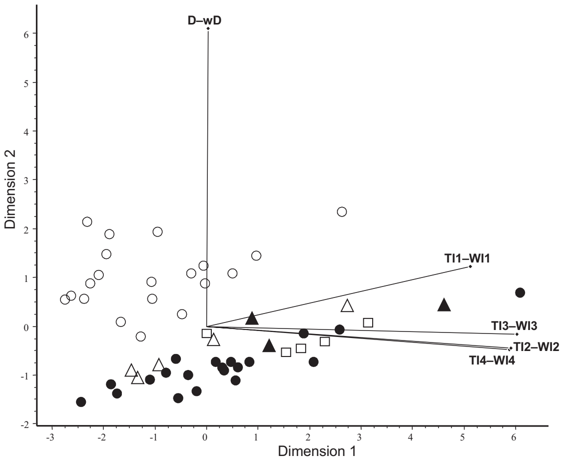


Figure 8.6: Principal components ordination biplot of the five studied indices showing that it does matter whether indirect effects are considered but there is no difference if these are longer than two steps (1: D and wD, 2: TI^1 and WI^1, 3: TI^2 and WI^2, 4: TI^3 and WI^3, 5: TI^4 and WI^4). Symbols refer to the positions of the 53 networks: (□) host-parasite webs, (▲) plant-ant webs, (•) Plant-Pollinator webs, (△) Plant-Seed disperser webs, (○) Food webs.

The arrangement of the 53 networks in the biplot is also informative, and provides a graphical summary of overall network relationships as depicted by the five index pairs. The group of food webs and plant-pollinator

	D vs. wD	TI^1 vs. WI^1	TI^2 vs. WI^2	TI^3 vs. WI^3	TI^4 vs. WI^4
D vs. wD	1.0000				
TI^1 vs. WI^1	0.1329	1.0000			
TI^2 vs. WI^2	-0.0459	0.7067	1.0000		
TI^3 vs. WI^3	-0.0150	0.7792	0.9323	1.0000	
TI^4 vs. WI^4	-0.0411	0.6669	0.9400	0.9471	1.0000

Table 8.3: Correlation matrix for the 5 sets of Goodman-Kruskal lamba-values of rank correlation between node orderings based on weighted and unweighted forms of the indices. The direct index D is very different from the others. The indirect index considering one step long effects is still different to some extent from the longer ones. Indirect indices of length two or more are essentially similar, as shown by these correlations (all being > 0.93, see also Figure 8.6).

networks are clearly recognizable on the plane, representing two endpoints in a gradient along the second axis. This axis has a strong negative correlation with network size (-0.48) showing that the smaller the network the less the difference between the weighted and unweighted ranks. The other three, smaller groups are positioned in between the two large groups. The first, more important PCA dimension corresponds with the indirect indices only and reflects increases of lambda within those. The large percentage is explained by the fact that four such indices were included. In summary, the lambda values are scale-dependent in case of the local index (D) but are absolutely scale-independent for the indirect indices, as also shown by the regression line superimposed over the point sets (Figure 8.7).

8.4 Weighting and indirect effects

Weighting links in ecological networks is costly and often problematic methodologically. Still, it can be of high interest, since a topological view on the network of interactions provides very limited information on functionality. Weights on links help us understand how the system really functions, by adding information on the extent to which particular interactions are "used". The importance of weighting and its particular methodology (frequency, normalized effect, etc.), however, depend on the actual problem. In our considerably large data base we have found that node centrality ranks are very different in weighted networks. This effect depends on network type: considering weights on links is more important in food webs, while there is no big difference between other kinds of interaction networks, especially in plant-pollinator networks. This finding may be related to the nature of prey choice (and predator avoidance) strategies: consumers seem to show strong preference for selected prey, making weights on their trophic links more variable. On the contrary, pollination seems to be more like a "yes or no" question, based on our analysis. This might be the reason why weighted pollination networks have not been in the focus of commu-

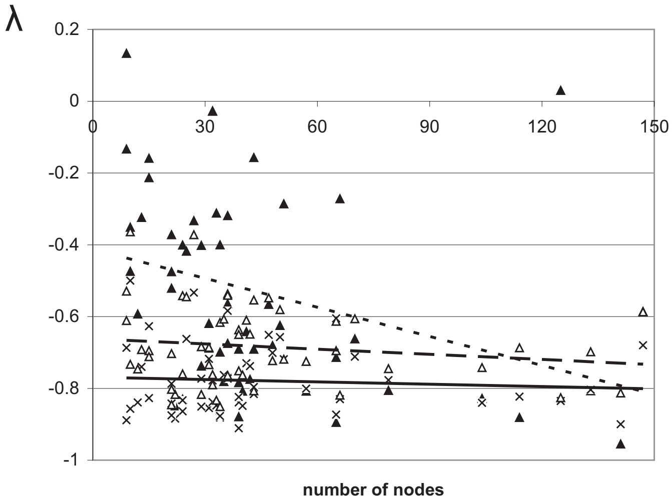


Figure 8.7: Scale dependence of lambda values (λ, characterizing the effect of weighting) for different centrality indices (linear regression lines are shown). Full triangles and dotted line correspond to direct interactions (D and wD), crosses and solid line correspond to one-step indirect effects (TI^1 and WI^1), and open triangles and the dashed line correspond to two-step indirect effects (TI^2 and WI^2). Indirect effects for three and four steps give essentially the same line as in the last case (two steps), and are therefore omitted. Note that the majority of lambda values are negative.

nity ecology for a long time. Also, these differences are more important if exclusively direct interactions are considered. Thus, if we are able to trace indirect effects in food webs, neglecting weights causes a smaller problem. It is more difficult to understand the statistical differences between food webs and host-parasitoid or plant-ant networks. Reasons can be similar to the plant-pollinator case: parasitism and mutualism are evolutionarily more stable, long-lasting interactions, their existence being more characteristic than their less variable strength (weight). Predation rates seem to be more flexible and sensitive to actual conditions, resulting in a higher variability and larger structural importance of weights. Food webs represent more transient and more flexible relationships, this is why they typically differ from other kinds of networks. Also, the effect of weighting is influenced by network size in case of direct indices but no scale-dependence is reported for indirect indices. All these imply that weighting ecological interactions is especially important in case of food webs and direct interactions. Note that the majority of recent papers on ecological networks characterize food web topology by degree and link distribution (e.g. Solé and Montoya, 2001). Our results support the surprising conclusion that weighting links is more important in analyzing complex ecological networks but, for methodological reasons, our data bases are typically of binary nature for these complex webs - while we have more data on weights in case of more aggregated webs, where it is less important.

Our results may be in agreement with some classical questions related to the evolution of ecosystems. Evolutionary ecological consequences of our results concern the long-term changes in complex networks of interspecific interactions. According to rules of thumb (based on a wealth of studies), predation and competition are relatively short-lived interactions on evolutionary time scales. Mutualism, and especially parasitism seem to hold much longer for clear reasons: to change a prey is easier than to change a host (Thompson, 1982). If these trends are correct, we may hypothesize that the strengths of the evolutionarily more stable interactions are more negligible, i.e. to consider weights in mutualistic and host-parasitoid networks is less important than to consider weights in food webs. The latter represent more transient and temporally flexible interactions, reflecting changes at faster time-scales. Thus, trophic relationships seem to be more variable and functionally weighted, while other interactions are better understandable on topological grounds. We conclude that in case of more stable interactions (host-parasite), topology is more characteristic, while in case of more transient and flexible interactions (predator-prey) the strength of links is more important.

A practical consequence of our study is the suggestion that if no weights are available (because of logistic considerations), we gain a more realistic view if we consider indirect effects. This is in concert with several suggestions by Ulanowicz and Puccia (1990) and Vasas and Jordán (2006).

8.5 Appendix I - 53 ecological networks

Web#	Web type and name	Reference
	Host - Parasite webs (5)	
1	Aishihik Lake	(Arthur et al., 1976)
2	Cold Lake	(Leong and Holmes, 1981)
3	McGregor River	(Arai and Mudry, 1983)
4	Parnsip River	(Arai and Mudry, 1983)
5	Small Wood Reservoir	(Chinniah and Threlfall, 1978)
	Plant - Ant webs (3)	
6	Rainforest (Peru)	(Davidson et al., 1989)
7	Tropical Forest (Costa Rica)	(Davidson and Fisher, 1991)
8	Amazon Rainforest	(Fonseca and Ganade, 1996)
	Plant - Pollinator webs (20)	
9	Boreal Forest (Canada)	(Barrett and Helenurm, 1987)
10	Alpine Subartic Community (Sweden)	(Elberling and Olesen, 1999)
11	Beech Forest (Japan)	(Kato et al., 1990)
12	Montane Forest (Australia)	(Inouye and Pyke, 1988)
13	Medow (Bristol, U.K.)	(Memmott, 1999)
14	Arctic Community (Canada)	(Mosquin and Martin, 1967)
15	Deciduous Forest (U.S.A.)	(Motten, 1982, 1986)
16	Coastal Forest (Mauritius Island)	(Olesen et al., 2002)
17	Rocky Cliff and Open Herb Community	(Olesen et al., 2002)
18	Upland Grassland (South Africa)	(Ollerton et al., 2003)
19	Maple-Oak Woodland (U.S.A.)	(Schemske et al., 1978)
20	Peat Bog (Canada)	(Small, 1976)
21-28	Evergreen Montane Forest - 8 subwebs	(Vázquez, 2002)
	Plant - Seed Disperser webs (5)	
29	Forest (Papa New Guinea)	(Beehler, 1983)
30	Semideciduous Tropical Forest	(Poulin et al., 1999)
31	Neotropical Forest (Trinidad)	(Snow and Snow, 1971)
32	Neotropical Forest (Trinidad)	(Snow and Snow, 1988)
33	Temperate Woodland (Britain)	(Sorensen, 1981)
	Food webs (20)	
34	Aggregated Baltic Sea	(Wulff and Ulanowicz, 1989)
35	Cedar Bog Lake	(Lindeman, 1942; Williams, 1971)
36	Charca de Maspalomas	(Almunia et al., 1999)
37	Chesapeake Mesohaline Ecosystem	(Wulff and Ulanowicz, 1989)
38	Chesapeake Mesohaline Network	(Baird and Ulanowicz, 1989)
39	Crystal River Creek (control)	(Ulanowicz, 1986)
40	Crystal River Creek (delta temp.)	(Ulanowicz, 1986)
41	Everglades Graminoids (wet season)	(Ulanowicz et al., 2000)
42	Florida Bay Ecosystem	(Ulanowicz et al., 1998)
43	Lower Chesapeake Bay in Summer	(Hagy, 2002)
44	St. Marks River (Florida) Flow Network	(Baird et al., 1998)
45	Lake Michigan Control Network	Krause and Mason (in preparation)
46	Middle Chesapeake Bay in Summer	(Hagy, 2002)
47	Mondego Estuary	(Patrício et al., 2004)
48	Final Narraganasett Bay Model	(Monaco and Ulanowicz, 1997)
49	North Sea	(Steele, 1974)
50	Somme Estuary	Rybarczyk, H., unpublished ms.
51	Upper Chesapeake Bay in Summer	(Hagy, 2002)
52	Upper Chesapeake Bay	Osgood, A., unpublished ms. CBL
53	Ythan Estuary	(Baird and Milne, 1981)

Table 8.4: References to networks analyzed in the present chapter.

8.6 Appendix II - Aishihik Lake species

	Host species
1	*Catostomus catostomus*
2	*Coregonus clupeaformis*
3	*Cottus cognatus*
4	*Esox lucius*
5	*Prosopium cylindraceum*
6	*Salvelinus namaycush*
7	*Thymallus arcticus*
	Parasite species
8	*Anonchohaptor anomalus*
9	*Dactylogyrus buddi*
10	*Discocotyle sagittata*
11	*Tetraonchus borealis*
12	*Tetraonchus monenteron*
13	*Tetraonchus variabilis*
14	*Allocreadium isosporum*
15	*Crepidostomum farionis*
16	*Diplostomum spathaceum (L)*
17	*Heterophyid meta (L)*
18	*Neascus*
19	*Tetracotyle sp. (L)*
20	*Cyathocephalus truncatus*
21	*Diphyllobothrium sp. (L)*
22	*Eubothrium salvelini*
23	*Glaridacris catostomi*
24	*Proteocephalus tumidocollus*
25	*Schistocephalus solidus (L)*
26	*Triaenophorus crassus*
27	*Triaenophorus crassus (L)*
28	*Capillaria salvelini*
29	*Cystidicola farionis*
30	*Raphidascaris acus*
31	*Raphidascaris acus (L)*
32	*Neoechinorhynchus tumidus*
33	*Piscicola milneri*
34	*Salmincola extensus*
35	*Salmincola edwardsii*
36	*Salmincola thymalli*

Table 8.5: List of species for the Aishihik Lake network (Figures 8.3 and 8.4). Species lists for the remaining webs are available at *www.nceas.ucsb.edu/interactionweb*.

9

Conclusions

9.1 From system to compartments

Although network theory principles has been introduced in ecology many years ago, advances exhibited by computer science during last decades stimulated and supported their spread and applications.

Food webs describe feeding relationships between species, depicting road maps for the flow of energy and matter in ecosystems. Information conveyed by a food web ranges from a binary map of trophic connections (with adjacency matrices associated to topological webs) to a functional description of the relative importance of each trophic connection (weighting links to quantify interaction strength). In the latter case, relative importance can be represented in terms of either the impact of consumers on resource abundance (e.g. $gC\ m^{-2}\ yr^{-1}$; $gN\ m^{-2}\ yr^{-1}$; number of fruits removed by frugivore birds) or energy flows to a consumers (e.g. $kcal\ m^{-2}\ yr^{-1}$ and $cal\ m^{-2}\ yr^{-1}$).

Few examples focussing on single species features, as emerging from roles they play in complex systems, have been stressed. In particular, such works involve topological centralities and only recent publications dealt with this approach introduced from social network analysis. By this research, starting from ENA procedure, I try to extend the knowledge on properties displayed by the node in relation to whole system properties. Brand new aspects of ecological network analysis stressed by the present work can be summarized as follows:

- Improving existing algorithms to obtain ecologically sound outcomes ("extended version" of CTA proposed in Chapter 3).

- Decomposing whole system indices into species contribution to flow structure (AMI of each compartment in Chapters 4, 5 and 6).
- Exploring how species trophic position affects food web topology (extremely high positive linear correlations between trophic position and single species AMI were found and described in Chapters 5 and 6).
- Evaluating distribution of trophic specialization and generalist trophic behaviour towards the trophic chain and defining omnivory composition (Chapter 6).
- Assessing food web model capability of predicting key structural properties of real ecological networks (see Chapters 4 and 5 for comparison of theoretical and empirical patterns of new indices introduced by my thesis).
- Studying patterns linking indices extracted from different methodologies (trophic position estimated with CTA and centralities from social analysis in Chapter 7).
- Tackling the problem of considering weights in ecological networks, since dealing with unweighted graph we are simply paying attention to the topological organization of nodes and links, while analyzing a weighted network we are including functional information that can obscure significantly the topological characteristics (food web properties were studied in Chapter 7 and a comparison between ecological networks with different interaction types is performed in Chapter 8).
- Promoting graph theory approach to ecosystems (with the simultaneous use of weighted and directed data), paying attention to differences between network application in ecology and other areas (e.g. social networks, Internet, metabolic webs and motorways). All these accomplishment are achieved in the whole thesis.

9.2 What next?

A promising approach to recognize both direct and indirect ecological interactions makes use of graph theory (Margalef, 1968, 1991). Within this framework I explored and investigated properties of each species. Although binary food webs have been deeply analyzed in recent years (Dunne et al., 2002b, 2004; Stouffer et al., 2006; Williams and Martinez, 2000), many scientists claimed the need of using weighted trophic links to unveil important dynamics, such as role of skewed interaction strengths (e.g. many weak and few strong links; Paine, 1992; McCann et al., 1998), or effects of species with few but extremely strong trophic links on secondary extinction (Dunne et al., 2002b).

In the thesis I dealt with these issues focussing on functioning rather than crude topology of trophic networks. Moreover, I demonstrated how the importance of weighting may vary depending on interaction types (Chapter 8), with for food webs more affected in comparison to the other networks. Despite link strengths vary considerably (Berlow et al., 2004), food web analysis of real networks are still limited by data availability and new efforts and expenses should be addressed to solve these inconsistencies (Cohen et al., 1993a). When qualitative descriptors are analyzed, most of these indices are extremely sensitive to different levels of sampling effort (Goldwasser and Roughgarden, 1997; Martinez et al., 1999). In particular, Bersier et al. (1999) showed how link density property tends to appear scale invariant in intrinsically scale dependent system, when low sampling effort is applied. Sampling effect is then a prominent problem inherent qualitative food webs. As previously observed, distribution of link strength is likely to be uneven (Paine, 1992; McCann et al., 1998) and the same weight assigned to each link in qualitative food webs distorts the true picture of their structure (Kenny and Loehle, 1991). These aspects emphasize the need of weighting to achieve a more sensible description of food web structure (May, 1983; Kenny and Loehle, 1991; Pimm et al., 1991; Cohen et al., 1993a; Bersier et al., 1999) although many quantitative indices of complex systems have been proposed (Ulanowicz, 1986, 1997; Hirata, 1995; Cattin et al., 2004). Stressing consequences of weighting links, I do not discard the validity of all the studies on binary data cited above; rather, I perceive the potential for their integration and propose new quantitative methods targeting species as a step ahead in the study of food webs.

Within my thesis, mathematical developments, algorithms and patterns were tested and analyzed on real data (e.g. "extended version" of CTA in Chapter 3; link and centrality distributions towards trophic hierarchy in Chapters 5 and 6; effects of weighting depending on interaction type and direct or indirect descriptors in Chapter 8) to stimulate collecting and gathering new datasets on natural systems. However, advances of computer science should continue corroborating the coupled action of empirical studies and simulations (Allesina et al., 2006; Dunne et al., 2002b).

Beside the role of weighting, I focussed on single node properties as emerging from their position in ecological networks. I decomposed indices previously analyzed as system descriptors (e.g. AMI in Chapter 4) and compared them to trophic positions (Chapters 5 and 6). Results allow to extend the discussion about certain key ecological issues such as: (a) the relation between weak and strong interactions in ecosystems (McCann et al., 1998; Sala and Graham, 2002; Emmerson and Yearsley, 2004), observing an asymmetry that seems to be in relation with the gradient of trophic positions; (b) the role of omnivory that, in accordance with other studies (Thompson et al., 2007; Yodzis, 1984), predominates above the herbivore trophic level; (c) the relation between functional and taxonomic features of species in ecological networks is consistent with the idea that members of one taxon may occupy different trophic roles in the same ecosystem (Polis, 1984).

In this framework, further studies should be addressed to solve drawbacks that are common to other food web studies as exploring effects of scale resolution (Allesina et al., 2005b), identifying the elements that comprise the ecosystem (Goldwasser and Roughgarden, 1997; Martinez et al., 1999; Abarca-Arenas and Ulanowicz, 2002) and taking decisions about the degree of resolution that, generally, is driven by the amount of available information together with the purpose of the study (Jordán, 2003).

Comparison between structure and functioning, achievable by weighting connections, has been stressed analyzing patterns of link density towards trophic hierarchy (Chapters 5 and 6) and dealing with centrality indices to detect keystone species (Chapters 7 and 8). This last issue represents an emerging approach in theoretical ecology and identifying keystone species (Paine, 1969, 1995; Mills et al., 1993; Power et al., 1996) may help ecologists answering questions of conservation biology (Jordán and Scheuring, 2002; Abarca-Arenas et al., 2007). Many recent studies discussed and adopted centrality indices for evaluating the keystone role of a species in a community (Solé and Montoya, 2001; Girvan and Newman, 2002; Jordán et al., 2005; Abarca-Arenas et al., 2007). Within my thesis I compute centrality values displayed by single nodes in real networks, exploring their patterns as regards to energy flows in food webs (Chapter 7) and interaction types (Chapter 8). Although sociometric indices resulted an important tool for the analysis of ecosystems, one can not simply translate studies on the structure of social networks into their ecological counterpart without facing some differences that make food webs peculiar (e.g. external source of energy to primary producers; directional energetic flow from producers to consumers). Nevertheless, keystone research in a network context is reasonable and the network perspective will be operative in future applications (Jordán et al., 1999; Jordán, 2000, 2001).

Centrality analysis poses important questions needing refinements. These are discussed in Chapter 8 and permeate the thinking of many ecologists in international publications: (a) depending on distance up to which direct or indirect effects are estimated, should network indices account for local, "meso" or global topological information (Jordán and Scheuring, 2002; Estrada, 2007)? (b) Can network models be constructed including different interaction types at the same time and how data should be processed (Vasas and Jordán, 2006)? (c) Which is the real ecological meaning of each centrality index (Jordán, 2001)? (d) If several centrality measures point to different species as the most central node in a network, can a generalized centrality index encompassing all the descriptors be used as a new variable characterizing the importance of a species in a food web (Estrada, 2007)?

Bibliography

Abarca-Arenas, L. G., Franco-Lopez, J., Peterson, M. S., Brown-Peterson, N. J., Valero-Pacheco, E., 2007. Sociometric analysis of the role of penaeids in the continental shelf food web off Veracruz, Mexico based on by-catch. Fisheries Research 87, 46–57.

Abarca-Arenas, L. G., Ulanowicz, R. E., 2002. The effects of taxonomic aggregation on network analysis. Ecological Modelling 149, 285–296.

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

Allesina, S., Bodini, A., 2004. Who dominates whom in the ecosystem? energy flow bottlenecks and cascading extinctions. Journal of Theoretical Biology 230, 351–358.

Allesina, S., Bodini, A., 2005. Food web networks: Scaling relation revisited. Ecological Complexity 2, 323–338.

Allesina, S., Bodini, A., Bondavalli, C., 2005a. Ecological subsystems via graph theory: the role of strongly connected components. Oikos 110, 164–176.

Allesina, S., Bodini, A., Bondavalli, C., 2006. Secondary extinctions in ecological networks: Bottlenecks unveiled. Ecological Modelling 194, 150–161.

Allesina, S., Bondavalli, C., 2003. Steady state of ecosystem flow networks: a comparison between balancing procedures. Ecological Modelling 165, 221–229.

Allesina, S., Bondavalli, C., 2004. WAND: an ecological network analysis user-friendly tool. Environmental Modelling & Software 19, 337–340.

Allesina, S., Bondavalli, C., Scharler, U. M., 2005b. The consequences of the aggregation of detritus pools in ecological networks. Ecological Modelling 189, 221–232.

Allesina, S., Ulanowicz, R. E., 2004. Cycling in ecological networks: Finn's index revisited. Computational Biology and Chemistry 28, 227–233.

Almunia, J., Basterretxea, G., Aristegui, J., Ulanowicz, R. E., 1999. Benthic-pelagic switching in a coastal subtropical lagoon. Estuarine Coastal and Shelf Science 49, 363–384.

Arai, H. P., Mudry, D. P., 1983. Protozoan and metazoan parasites of fishes from the headwaters of the Parsnip and McGregor Rivers, British Columbia: a study of possible parasite transfaunations. Canadian Journal of Fisheries and Aquatic Sciences 40, 1676–1684.

Arthur, J. R., Margolis, L., Arai, H. P., 1976. Parasites of fishes of Aishihik and Stevens Lakes, Yukon Territory, and potential consequences of their interlake transfer through a proposed water diversion for hydroelectrical purposes. Journal of the Fisheries Research Board of Canada 22, 2489–2499.

Atlan, H., 1974. On a formal definition of organization. Journal of Theoretical Biology 45, 295–304.

Augustinovic, M., 1970. Methods of international and intertemporal comparison of structure. Vol. 1. North-Holland, Amsterdam, Ch. Contributions to Input-Output Analysis, pp. 249–269.

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., Thingstad, F., 1983. The ecological role of water-column microbes in the sea. Marine Ecology Progress Series 10, 257–263.

Baird, D., Luczkovich, J. J., Christian, R. R., 1998. Assessment of spatial and temporal variability in ecosystem attributes of the St Marks National Wildlife Refuge, Apalachee Bay, Florida. Estuarine, Coastal, and Shelf Science 47, 329–349.

Baird, D., Milne, H., 1981. Energy flow in the Ythan Estuary, Aberdeenshire, Scotland. Estuarine, Coastal and Shelf Science 13, 455–472.

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

Baird, D., Ulanowicz, R. E., 1993. Comparative-study on the trophic structure, cycling and ecosystem properties of 4 tidal estuaries. Marine Ecology Progress Series 99, 221–237.

Barrett, S. C. H., Helenurm, K., 1987. The reproductive-biology of Boreal Forest herbs. 1. Breeding systems and pollination. Canadian Journal of Botany 65, 2036–2046.

Beehler, B., 1983. Frugivory and polygamy in birds of paradise. The Auk 100, 1–12.

Bennett, S. J., Sanders, R. W., Porter, K. G., 1990. Heterotrophic, autotrophic, and mixotrophic nanoflagellates: Seasonal abundances and bacterivory in a eutrophic lake. Limnology and Oceanography 35, 1821–1832.

Berlow, E. L., Neutel, A.-M., Cohen, J. E., de Ruiter, P. C., Ebenman, B., Emmerson, M., Fox, J. W., Jansen, V. A. A., Jones, J. I., Kokkoris, G. D., Logofet, D. O., McKane, A. J., Montoya, J. M., Petchey, O., 2004. Interaction strengths in food webs: issues and opportunities. Journal of Animal Ecology 73, 585–598.

Bersier, L.-F., Banasek-Richter, C., Cattin, M. F., 2002. Quantitative descriptors of food-web matrices. Ecology 83, 2394–2407.

Bersier, L.-F., Dixon, P., Sugihara, G., 1999. Scale-invariant or scale-dependent behavior of the link density property in food webs: A matter of sampling effort? American Naturalist 153, 676–682.

Bondavalli, C., Bodini, A., Rossetti, G., Allesina, S., 2006. Detecting stress at a whole ecosystem level. The case of a mountain lake: Lake Santo (Italy). Ecosystems 9, 1–56.

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

Borer, E. T., Anderson, K., Blanchette, C. A., Broitman, B., Cooper, S. D., Halpern, B. S., Seabloom, E. W., Shurin, J. B., 2002. Topological approaches to food web analyses: a few modifications may improve our insights. Oikos 99, 397401.

Borgatti, S. P., Everett, M. G., Freeman, L. C., 2002. Ucinet for Windows: Software for Social Network Analysis. Harvard: Analytic Technologies.

Bratbak, G., 1987. Carbon flow in an experimental microbial ecosystem. Marine Ecology Progress Series 36, 267–276.

Briand, F., 1983. Environmental control of food web structure. Ecology 64, 253–263.

Burns, T. P., 1989. Lindeman's contradiction and the trophic structure of ecosystems. Ecology 70, 1355–1362.

Burns, T. P., Higashi, M., Wainright, S. C., Patten, B. C., 1991. Trophic unfolding of a continental-shelf energy-flow network. Ecological Modelling 55, 1–26.

Camacho, J., Guimerà, R., Amaral, L. A. N., 2002a. Robust patterns in food web structure. Physical Review Letters 88, art. no.–228102.

Camacho, J., Guimerà, R., Nunes Amaral, L. A. N., 2002b. Analytical solution of a model for complex food webs. Physical Review E 65, article number 030901(R).

Camerano, L., 1880. Dell'equilibrio dei viventi mercé la reciproca distruzione. Atti della Reale Accademia delle Scienze 15, 393–414.

Carpenter, S. R., Kitchell, J. F., 1993. The trophic cascade in lakes. Cambridge University Press, Cambridge, UK.

Carpenter, S. R., Kitchell, J. F., Hodgson, J. R., 1986. Cascading trophic interactions and lake productivity. BioScience 35, 634–639.

Cattin, M. F., Bersier, L.-F., Banasek-Richter, C., Baltensperger, R., Gabriel, J.-P., 2004. Phylogenetic constraints and adaptation explain food-web structure. Nature 427, 835–839.

Chinniah, V. C., Threlfall, W., 1978. Metazoan parasites of fish from the Smallwood Reservoir, Labrador, Canada. Journal of Fish Biology 13, 203–213.

Christensen, V., Pauly, D., 1992. Ecopath II - a software for balancing steady-state ecosystem models and calculating network characteristics. Ecological Modelling 61, 169–185.

Christensen, V., Pauly, D. (Eds.), 1993. Trophic models of aquatic ecosystems. ICLARM Conference Proceedings 26.

Christensen, V., Walters, C., Pauly, D., November 2005. ECOPATH with ECOSIM: a user's Guide. Fisheries Centre, University of British Columbia, Vancouver, Canada, http://www.ecopath.org.

Christian, R. R., Luczkovich, J. J., 1999. Organizing and understanding a winter's seagrass foodweb network through effective trophic levels. Ecological Modelling 117, 99–124.

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.

Cohen, J. E., 1978. Food Webs and Niche Space. Princeton University Press.

Cohen, J. E., Beaver, R. A., Cousins, S. H., DeAngelis, D. L., Goldwasser, L., Heong, K. L., Holt, R. D., Kohn, A. J., Lawton, J. H., Martinez, N. D., Omalley, R., Page, L. M., Patten, B. C., Pimm, S. L., Polis, G. A., Rejmanek, M., Schoener, T. W., Schoenly, K., Sprules, W. G., Teal, J. M., Ulanowicz, R. E., Warren, P. H., Wilbur, H. M., Yodzis, P., 1993a. Improving food webs. Ecology 74, 252–258.

Cohen, J. E., Briand, F., 1988. Trophic links of community food webs. Proceedings of the National Academy of Sciences of the United States of America 81, 4105–4109.

Cohen, J. E., Briand, F., Newman, C. M., 1990. Community food webs: data and theory. Springer, New York.

Cohen, J. E., Newman, C. M., 1985. A stochastic theory of community food webs. I. Models and aggregated data. Proceedings of the Royal Society of London Series B 224, 421–448.

Cohen, J. E., Newman, C. M., 1991. Community area and food chain length: theoretical predictions. American Naturalist 138, 1542–1554.

Cohen, J. E., Pimm, S. L., Yodzis, P., Saldana, J., 1993b. Body sizes of animal predators and animal prey in food webs. Journal of Animal Ecology 62, 67–78.

Cousins, S. H., 1985. Ecologists build pyramids again. New Scientist 107, 50–54.

Cousins, S. H., 1987. The decline of the trophic level concept. TRENDS in Ecology and Evolution 2, 312–316.

Davidson, D. W., Fisher, B. L., 1991. Ant-Plant Interactions. Oxford University Press, Oxford, Ch. Symbiosis of ants with Cecropia as a function of light regime, pp. 289–309.

Davidson, D. W., Snelling, R. R., Longino, J. T., 1989. Competition among ants for myrmecophytes and the significance of plant trichomes. Biotropica 21, 64–73.

de Ruiter, P. C., Neutel, A., Moore, J. C., 1995. Energetics, patterns of interaction strengths, and stability in real ecosystems. Science 269, 1257–1260.

DeAngelis, D. L., 1980. Energy-flow, nutrient cycling, and ecosystem resilience. Ecology 61, 764–771.

Dunne, J. A., 2006. Ecological Networks: Linking Structure to Dynamics in Food Webs. Oxford University Press, Oxford, Ch. The network structure of food webs, pp. 27–86.

Dunne, J. A., Williams, R. J., Martinez, N. D., 2002a. Food-web structure and network theory: the role of connectance and size. Proceedings of the

National Academy of Sciences of the United States of America 99, 12917–12922.

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

Dunne, J. A., Williams, R. J., Martinez, N. D., 2004. Network structure and robustness of marine food webs. Marine Ecology Progress Series 273, 291–302.

Elberling, H., Olesen, J. M., 1999. The structure of a high latitude plant-flower visitor system: the dominance of flies. Ecography 22, 314–323.

Elton, C. S. (Ed.), 1927. Animal ecology. Sidgwick and Jackson, London, UK.

Elton, C. S. (Ed.), 1958. The Ecology of Invasions by Animals and Plants. Methuen & Co., London, UK.

Emmerson, M., Yearsley, J. M., 2004. Weak interactions, omnivory and emergent food-web properties. Proceedings of the Royal Society of London B 271, 397–405.

Erdös, P., Rényi, A., 1960. On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5, 17–61.

Estrada, E., 2007. Characterization of topological keystone species: Local, global and “meso-scale” centralities in food webs. Ecological Complexity 4, 48–57.

Fagan, W. F., 1997. Omnivory as a stabilizing feature of natural communities. American Naturalist 150, 554–567.

Fath, B. D., Patten, B. C., 1998. Network synergism: emergence of positive relations in ecological systems. Ecological Modelling 107, 127–143.

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

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.

Fonseca, C. R., Ganade, G., 1996. Asymmetries, compartments and null interactions in an amazonian ant-plant community. Journal of Animal Ecology 66, 339–347.

Fox, J. W., 2006. Current food web models cannot explain the overall topological structure of observed food webs. Oikos 115, 97–109.

Gaedke, U., Straile, D., 1997. The trophic position of dead autochthonous organic material and its treatment in trophic analysis. Environmental Modeling and Assessment 2, 13–22.

Garlaschelli, D., Caldarelli, G., Pietronero, L., 2003. Universal scaling relations in food webs. Nature 423, 165–167.

Girvan, M., Newman, M. E. J., 2002. Community structure in social and biological networks. Proceedings of the National Academy of Sciences of the United States of America 99, 7821–7826.

Goldwasser, L., Roughgarden, J., 1997. Sampling effects and estimation of food-web properties. Ecology 78, 41–54.

Goodman, L. A., Kruskal, W. H., 1954. Measures of association for cross-classifications. Journal of the American Statistical Association 49, 732–764.

Gophen, M., 1990. Biomanipulation: Retrospective and future development. Hydrobiologia 200/201, 1–11.

Gradstheyn, I. S., Ryzhik, I. M., 2000. Table of Integrals, Series and Products, 6th Edition. Academic Press, New York.

Hagy, J. D., 2002. Eutrophication, hypoxia and trophic transfer efficiency in Chesapeake Bay. Ph.D. thesis, University of Maryland at College Park (USA).

Hairston Jr., N. G., Hairston Sr., N. G., 1993. Cause-effect relationships in energy flow, trophic structure, and interspecific interactions. American Naturalist 142, 379–411.

Hall, S. J., Raffaelli, D., 1991. Food web patterns: lessons from a species-rich web. Journal of Animal Ecology 60, 823–842.

Han, B. P., 1997. On several measures concerning flow variables in ecosystems. Ecological Modelling 104, 289–302.

Hannon, B., 1973. The structure of ecosystems. Journal of Theoretical Biology 41, 535–546.

Hannon, B. M., Costanza, R., Ulanowicz, R. E., 1991. A general accounting framework for ecological-systems - a functional taxonomy for connectivist ecology. Theoretical Populatuion Biology 40, 78–104.

Havens, K., 1992. Scale and structure in natural food webs. Science 257, 1107–1109.

Herendeen, R., 1989. Energy intensity, residence time, exergy, and ascendancy in dynamic ecosystems. Ecological Modelling 48, 19–44.

Heymans, J. J., Ulanowicz, R. E., Bondavalli, C., 2002. Network analysis of the South Florida Everglades graminoid marshes and comparison with nearby cypress ecosystems. Ecological Modelling 149, 5–23.

Higashi, M., Burns, T. P., Patten, B. C., 1989. Food network unfolding: an extension of trophic dynamics for application to natural ecosystems. Journal of Theoretical Biology 140, 243–261.

Higashi, M., Burns, T. P., Patten, B. C., 1992. Trophic niches of species and trophic structure of ecosystems - complementary perspectives through food network unfolding. Journal of Theoretical Biology 154, 57–76.

Higashi, M., Patten, B. C., Burns, T. P., 1991. Network trophic dynamics: an emerging paragigm in ecosystems ecology. In: Higashi, M., Burns, T. P. (Eds.), Theoretical studies of ecosystems - the network perspective. Cambridge University Press.

Hirata, H., 1995. Complex Ecology: the part-whole relation in ecosystems. Prentice Hall, Englewood Cliffs, New Jersey, USA, Ch. Information theory and ecological networks, pp. 623–642.

Hutchinson, G. E., MacArthur, R. H., 1959. A theoretical ecological model of size distributions among species of animals. American Naturalist 93, 117–125.

Inouye, D. W., Pyke, G. H., 1988. Pollination biology in the Snowy Mountains of Australia: comparisons with montane Colorado, USA. Australian Journal of Ecology 13, 191–210.

Ives, A. R., Carpenter, S. R., Dennis, B., 1999. Community interaction webs and zooplankton responses to planktivory manipulations. Ecology 80, 1405–1421.

Jordán, F., 2000. Is the role of trophic control larger in a stressed ecosystem? Community Ecology 1, 139–146.

Jordán, F., 2001. Seasonal changes in the positional importance of components in the trophic flow network of the Chesapeake Bay. Journal of Marine Systems 27, 289–300.

Jordán, F., 2003. Comparability: the key to the applicability of food web research. Applied Ecology and Environmental Research 1, 1–18.

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

Jordán, F., Liu, W., Davis, A. J., 2006. Topological keystone species: measures of positional importance in food webs. Oikos 112, 535–546.

Jordán, F., Liu, W., Wyatt, T., 2005. Topological constraints on the dynamics of wasp-waist ecosystems. Journal of Marine Systems 57, 250–263.

Jordán, F., Liu, W.-C., van Veen, F. J. F., 2003. Quantifying the importance of species and their interactions in a host-parasitoid community. Community Ecology 4, 79–88.

Jordán, F., Scheuring, I., 2002. Searching for keystones in ecological networks. Oikos 99, 607–612.

Jordán, F., Scheuring, I., 2004. Network Ecology: topological constraints on ecosystems dynamics. Physics of Life Reviews 1, 139–172.

Jordán, F., Takács-Sánta, A., Molnár, I., 1999. A reliability theoretical quest for keystones. Oikos 86, 453–462.

Kato, M. T., Makutani, T., Inoue, T., Itino, T., 1990. Insect-flower relationship in the primary beech forest of Ashu, Kyoto: an overview of the flowering phenology and seasonal pattern of insect visits. Contr. Biol. Lab. Kyoto Univ. 27, 309–375.

Kay, J. J., Graham, L. A., Ulanowicz, R. E., 1989. A detailed guide to network analysis. Vol. Network Analysis in Marine Ecology: Methods and Applications. Springer, Berlin.

Kenny, D., Loehle, C., 1991. Are food webs randomly connected? Ecology 72, 1794–1799.

Kercher, J. R., Shugart, H. H., 1975. Trophic structure, effective trophic position, and connectivity in food webs. American Naturalist 109, 191–206.

Knuth, D. E., 1973. Fundamental Algorithms, Vol. 1. Addison-Wesley, Reading, Massachusetts.

Krause, A. E., Frank, K. A., Mason, D. M., Ulanowicz, R. E., Taylor, W. W., 2003. Compartments revealed in food-web structure. Nature 426, 282–285.

Kruskal, W. H., Wallis, W. A., 1952. Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association 48, 583–634.

Latham, L. G., Scully, E. P., 2002. Quantifying constraint to assess development in ecological networks. Ecological Modelling 154, 25–44.

Leong, T. S., Holmes, J. C., 1981. Communities of metazoan parasites in open water fishes of cold lake, alberta. Journal of Fish Biology 18, 693–713.

Leontief, W. W., 1951. The structure of the American economy, 2nd Edition. Oxford University Press, New York.

Leontief, W. W., 1963. The structure of American Economy 1919-1939. Oxford University Press, New York.

Levine, S., 1983. Several measures of trophic structure applicable to complex food webs. Journal of Theoretical Biology 83, 195–207.

Liljeros, F., Edling, C. R., Nunes Amaral, L. A. N., Stanley, H. E., Åberg, Y., 2001. The web of human sexual contacts. Nature 411, 907–908.

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

Loreau, M., 1994. Material cycling and the stability of ecosystems. American Naturalist 143, 508–513.

Luczkovic, J. J., Borgatti, D. P., Johnson, J. C., Everett, M. G., 2003. Defining and measuring trophic role similarity in food webs using regular equivalence. Journal of Theoretical Biology 220, 303–321.

MacArthur, R. H., 1955. Fluctuations of animal populations and a measure of community stability. Ecology 36, 533–536.

Mann, H. B., Whitney, D. R., 1947. On a test of whether one of two random variables is stochastically larger than the other. Annals of Mathematical Statistics 18, 50–60.

Margalef, R., 1968. Perspectives in ecological theory. University of Chicago Press.

Margalef, R., 1991. Theoretical studies of ecosystems-the network perspective. Cambridge University Press, Ch. Networks in ecology, pp. 41–57.

Martinez, N. D., 1991. Artifacts or attributes? Effects of Resolution on the Little Rock Lake food web. Ecological Monographs 61, 367–392.

Martinez, N. D., 1992. Constant connectance in community food webs. American Naturalist 139, 1208–1218.

Martinez, N. D., 1993. Effect of scale on food web structure. Science 260, 242–243.

Martinez, N. D., Hawkins, B. A., Dawah, H. A., Feifarek, B. P., 1999. Effects of sampling effort on characterization of food-web structure. Ecology 80, 1044–1055.

Mateti, P., Deo, N., 1976. On algorithms for enumerating all the circuits of a graph. SIAM Journal on Computing 5, 90–99.

May, R. M., 1972. Will a large complex system be stable? Nature 238, 413–414.

May, R. M. (Ed.), 1973. Stability and Complexity in Model Ecosystems. Princeton University Press, Princeton.

May, R. M., 1983. The structure of foodwebs. Nature 301, 566–568.

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

McCann, K., Hastings, A., Huxel, G. R., 1998. Weak trophic interactions and the balance of nature. Nature 395, 794–798.

McCann, K. S., 2000. The diversity-stability debate. Nature 405, 228–233.

Memmott, J., 1999. The structure of a plant-pollinator food web. Ecology Letters 2, 276–280.

Menge, B. A., 1995. Indirect effects in marine rocky intertidal interaction webs: paterns and importance. Ecological Monographs 65, 21–74.

Mills, L. S., Soulé, M. E., Doak, D. F., 1993. The keystone-species concept in ecology and conservation. BioScience 43, 219–224.

Monaco, M. E., Ulanowicz, R. E., 1997. Comparative ecosystem trophic structure of three U.S. mid-Atlantic estuaries. Marine Ecology Progress Series 161, 239–254.

Montoya, J. M., Pimm, S. L., Solé, R. V., 2006. Ecological networks and their fragility. Nature 440, 259–264.

Montoya, J. M., Solé, R. V., 2002. Small world patterns in food webs. Journal of Theoretical Biology 214, 405–412.

Montoya, J. M., Solé, R. V., 2003. Topological properties of food webs: from real data to community assembly models. Oikos 102, 614–622.

Mosquin, T., Martin, J. E. H., 1967. Observations on the pollination biology of plants on Melville Island, N.W.T., Canada. Canadian Field Naturalist 81, 201–205.

Motten, A. F., 1982. Pollination Ecology of the Spring Wildflower Community in the Deciduous Forests of Piedmont North Carolina. Ph.D. thesis, Duke University, Duhram, North Carolina, USA.

Motten, A. F., 1986. Pollination ecology of the spring wildflower community of a temperate deciduous forest. Ecological Monographs 56, 21–42.

Müller, C. B., Adriaanse, I. C. T., Belshaw, R., Godfray, H. C. J., 1999. The structure of an aphid-parasitoid community. Journal of Animal Ecology 68, 346–370.

Neutel, A., Heesterbeek, J. A. P., de Ruiter, P. C., 2002. Stability in real food webs: weak links in long loops. Science 296, 1120–1123.

Newman, M. E. J., 2003. The structure and function of complex networks. SIAM Rev. 45, 167–256.

NSF, 1999. Decision-making and Valuation for Environmental Policy. NSF Bulletin 99-14, National Science Foundation, Ballston, VA.

Odum, E. P., 1969. The strategy of ecosystem development. Science 164, 262–270.

Oksanen, L., 1991. Trophic level and trophic dynamics: a consensus emerging? TRENDS in Ecology and Evolution 6, 58–60.

Olesen, J. M., Eskildsen, L. I., Venkatasamy, S., 2002. Invasion of pollination networks on oceanic islands: importance of invader complexes and endemic super generalists. Diversity and Distributions 8, 181–192.

Ollerton, J., Johnson, S. D., Cranmer, L., Kellie, S., 2003. The pollination ecology of an assemblage of grassland asclepiads in South Africa. Annals of Botany 92, 807–834.

Pace, M. L., Cole, J. J., Carpenter, S. R., Kitchell, J. F., 1999. Trophic cascades revealed in diverse ecosystems. TRENDS in Ecology and Evolution 14, 483–488.

Paine, R. T., 1969. A note on trophic complexity and community stability. American Naturalist 103, 91–93.

Paine, R. T., 1980. Food webs: linkage interaction strength and community infrastructure. Journal of Animal Ecology 49, 667–685.

Paine, R. T., 1988. Food webs: road maps of interactions or grist for theoretical development. Ecology 69, 1648–1654.

Paine, R. T., 1992. Food-web analysis through field measurement of per capita interaction strength. Nature 355, 73–75.

Paine, R. T., 1995. A conversation on refining the concept of keystone species. Conservation Biology 9, 962–964.

Parker, G. G., 1977. Understanding the inverse theory. Annual review of Earth and Planetary Science 5, 35–64.

Pascual, M., Dunne, J. A., 2006. Ecological Networks: Linking Structure to Dynamics in Food Webs. Oxford University Press, Oxford, UK.

Patrício, J., Ulanowicz, R. E., Pardal, M. A., Marques, J. C., 2004. Ascendency as an ecological indicator: a case study of estuarine pulse eutrophication. Estuarine Coastal and Shelf Science 60, 23–35.

Patten, B. C., 1982. Environs - relativistic elementary-particles for ecology. American Naturalist 119, 179–219.

Patten, B. C., Bosserman, R. W., Finn, J. T., Cale, W. G., 1976. Systems Analysis and Simulation in Ecology vol. 4. Academic Press, New York, Ch. Propagation of cause in ecosystems, pp. 457–579.

Patten, B. C., Higashi, M., 1984. Modified cycling index for ecological applications. Ecological Modelling 25, 69–83.

Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., Torres, F. J., 1998. Fishing down marine food webs. Science 279, 860–863.

Pauly, D., Christensen, V., Walters, C., 2000. Ecopath, Ecosim, and Ecospace as tools for evaluating ecosystem impact of fisheries. ICES Journal of Marine Science 57, 697–706.

Pimm, S. L., 1980. Food web design and the effect of species deletion. Oikos 35, 139–149.

Pimm, S. L., 1982. Food Webs. Chapman & Hall, New York.

Pimm, S. L., 1988. Concepts of Ecosystem Ecology. Springer-Verlag, New York, Ch. Energy flow and trophic structure.

Pimm, S. L., Lawton, J. H., 1977. The number of trophic levels in ecological communities. Nature 275, 542–544.

Pimm, S. L., Lawton, J. H., 1978. On feeding on more than one trophic level. Nature 254, 542–544.

Pimm, S. L., Lawton, J. H., 1980. Are food webs divided into compartments? Journal of Animal Ecology 49, 879–898.

Pimm, S. L., Lawton, J. H., Cohen, J. E., 1991. Food web patterns and their consequences. Nature 350, 669–674.

Platt, T. C., Mann, K. H., Ulanowicz, R. E., 1981. Mathematical Models in Biological Oceanography. UNESCO Press, Paris, France.

Podani, J., 2000. Introduction to the exploration of multivariate biological data. Backhuys Publishers, Leiden, The Netherlands.

Podani, J., 2001. SYN-TAX 2000 Users Manual. Scientia, Budapest.

Polis, G. A., 1984. Age structure component of niche width and intraspecific resource partitioning: can age groups function as ecological species? American Naturalist 123, 541–564.

Polis, G. A., 1991. Complex trophic interactions in deserts: an empirical critique of food-web theory. American Naturalist 138, 123–155.

Polis, G. A., 1994. Food webs, trophic cascades and community structure. Australian Journal of Ecology 19, 121–136.

Polis, G. A., Anderson, W. B., Holt, R. D., 1997. Toward and integration of landscape and food web ecology: the dynamics of spatially subsidized food webs. Annual Review of Ecology and Systematics 28, 289–316.

Polis, G. A., Hurd, S. D., 1996. Food webs: Integration of Patterns and Dynamics. Chapman-Hall, NY, Ch. Allochthonous input across habitats, subsidized consumers, and apparent trophic cascades: examples from the ocean-land interface, pp. 275–285.

Polis, G. A., Power, M. E., Huxel, G. R., 2004. Food webs at the Landscape Level. University of Chicago Press, Chicago.

Polis, G. A., Strong, D. R., 1996. Food web complexity and community dynamics. American Naturalist 147, 813–846.

Polovina, J. J., 1984. Model of a Coral-Reef Ecosystem. 1. the Ecopath model and its application to French Frigate Shoals. Coral Reefs 3, 1–11.

Pomeroy, L. R., 1984. Current perspectives in microbial ecology. American Society for Microbiology, Washington, D.C., Ch. Significance of microorganisms in carbon and energy flow in marine ecosystems, pp. 405–411.

Post, D. M., 2002. The long and short of food-chain length. TRENDS in Ecology and Evolution 17, 269–277.

Poulin, B., Wright, S. J., L. G., Calderon, O., 1999. Interspecific synchrony and asynchrony in the fruiting phenologies of congeneric bird-dispersed plants in panama. Journal of Tropical Ecology 15, 213–227.

Power, M. E., Tilman, D., Estes, J. A., Menge, B. A., Bond, W. J., Mills, L. S., Daily, G., Castilla, J. C., Lubchenco, J., Paine, R. T., 1996. Challenges in the quest for keystones. BioScience 46, 609–620.

Raffaelli, D., Hall, S. J., 1996. Food Webs: Integration of Pattern and Dynamics. Chapman and Hall, New York, Ch. Assessing the relative importance of trophic links in food webs, pp. 185–191.

Rana, J. S., Dixon, A. F. G., Jarosik, V., 2002. Costs and benefits of prey specialization in a generalist insect predator. Journal of Animal Ecology 71, 15–22.

Rutledge, R., Basorre, B., Mulholland, R., 1976. Ecological stability: an information theory viewpoint. Journal of Theoretical Biology 57, 355–371.

Sala, E., Graham, M. H., 2002. Community-wide distribution of predator-prey interaction strength in kelp forests. Proceedings of the National Academy of Sciences of the United States of America 99, 3678–3683.

Schemske, D. W., Willson, M. F., Melampy, M. N., Miller, L. J., Verner, L., Schemske, K. M., Best, L. B., 1978. Flowering ecology of some spring woodland herbs. Ecology 59, 351–366.

Schmitz, O. J., Hambäck, P. A., Beckerman, A. P., 2000. Trophic cascades in terrestrial systems: a review of the effects of carnivore removals on plants. American Naturalist 155, 141–153.

Schoener, T. W., 1989. Food webs from the small to the large. Ecology 70, 1559–1589.

Scotti, M., Allesina, S., Bondavalli, C., Bodini, A., Abarca-Arenas, L. G., 2006. Effective trophic positions in ecological acyclic networks. Ecological Modelling 198, 495–505.

Scotti, M., Podani, J., Jordán, F., 2007. Weighting, scale dependence and indirect effects in ecological networks: A comparative study. Ecological Complexity 4, 148–159.

Shannon, C. E., 1948. A mathematical theory of communications. Bell System Technical Journal 27, 379–423.

Shapiro, J., 1990. Biomanipulation: The next phase making it stable. Hydrobiologia 200/201, 13–27.

Shapiro, S. S., Francia, R. S., 1972. An approximate analysis of variance test for normality. Journal of the American Statistical Association 67, 215–216.

Sherr, E. B., Sherr, B. F., Fessenden, L., Gosselin, M., Thibault, D., Wheeler, P. A., 2003. Heterotrophic Bacteria and Protists in the Upper Water Column of the Central Arctic Ocean. National Snow and Ice Data Center, Boulder, CO.

Sih, A., Englund, G., Wooster, D., 1998. Emergent impacts of multiple predators on prey. TRENDS in Ecology and Evolution 13, 350–355.

Simon, H. A., Hawkins, D., 1949. Some conditions of macroeconomic stability. Econometrica 17, 245–248.

Small, E., 1976. Insect pollinators of the mer bleue peat bog of ottawa. Canadian Field Naturalist 90, 22–28.

Snow, B. K., Snow, D. W., 1971. The feeding ecology of tanagers and honeycreeperes in trinidad. The Auk 88, 291–322.

Snow, B. K., Snow, D. W., 1988. Birds and Berries. Calton, England.

Sokal, R. R., Rohlf, F., 1981. Biometry, 2nd Edition. Freeman, San Francisco.

Solé, R. V., Montoya, J. M., 2001. Complexity and fragility in ecological networks. Proceedings of the Royal Society of London Series B-Biological Sciences 268, 2039–2045.

Solow, A. R., Beet, A. R., 1998. On lumping species in food webs. Ecology 79, 2013–2018.

Sorensen, A. E., 1981. Interactions between birds and fruit in a temperate woodland. Oecologia 50, 242–249.

Steele, J. H., 1974. The Structure of Marine Ecosystems. Harvard University Press, Cambridge, Massachussetts.

Stouffer, D. B., Camacho, J., Nunes Amaral, L. A. N., 2006. A robust measure of food web intervality. Proceedings of the National Academy of Sciences of the United States of America 103, 19015–19020.

Strogatz, S. H., 2001. Exploring complex networks. Nature 410, 268–275.

Szyrmer, J., Ulanowicz, R. E., 1987. Total flows in ecosystems. Ecological Modelling 35, 123–136.

Thompson, J. N., 1982. Interaction and Coevolution. Wiley and Sons, New York.

Thompson, J. N., 1991. The Coevolutionary Process. University of Chicago Press, Chicago.

Thompson, R. M., Hemberg, M., Star, 2007. Trophic levels and trophic tangles: the prevalence of omnivory in real food webs. Ecology 88, 612–617.

Tilly, L. J., 1968. The structure and dynamics of cone spring. Ecological Monographs 38, 169–197.

Tilman, D., Downing, J. A., 1994. Biodiversity and stability in grasslands. Nature 367, 363–365.

Ulanowicz, R. E., 1980. An hypothesis on the development of natural communities. Journal of Theoretical Biology 85, 223–245.

Ulanowicz, R. E., 1983. Identifying the structure of cycling in ecosystems. Mathematical Biosciences 65, 219–237.

Ulanowicz, R. E., 1986. Growth & Development: Ecosystems Phenomenology. Springer Verlag, New York.

Ulanowicz, R. E., 1995. Complex Ecology: the part-whole relation in ecosystems. Prentice-Hall, New Jersey, Ch. Ecosystem trophic foundations: Lindeman Exonerata.

Ulanowicz, R. E., 1996. Food Webs: Integration of Patterns and Dynamics. Chapman-Hall, NY, Ch. Trophic flow networks as indicators of ecosystem stress, pp. 358–368.

Ulanowicz, R. E., 1997. Ecology, the Ascendent Perspective. Columbia University Press.

Ulanowicz, R. E., 2002. NETWRK 4.2b: A Package of Computer Algorithms to Analyze Ecological Flow Networks. University of Maryland, Center for Environmental & Estuarine Studies, Chesapeake Biological Laboratory - Solomons, MD, http://www.cbl.umces.edu/∼ ulan/ntwk/network.html.

Ulanowicz, R. E., 2004. Quantitative methods for ecological network analysis. Computational Biology and Chemistry 28, 321–339.

Ulanowicz, R. E., Bondavalli, C., Egnotovich, M. S., 1998. Network Analysis of Trophic Dynamics in South Florida Ecosystems, FY 97: The Florida Bay Ecosystem. Tech. Rep. [UMCES] CBL 98-123, Chesapeake Biological Laboratory, Solomons.

Ulanowicz, R. E., Heymans, J. J., Egnotovich, M. S., 2000. Network Analysis of Trophic Dynamics in South Florida Ecosystems FY 99: The Graminoid Ecosystem. Tech. Rep. [UMCES] CBL 00-0176, Chesapeake Biological Laboratory, Solomons.

Ulanowicz, R. E., Kay, J. 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. American Naturalist 114, 871–883.

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

Ulanowicz, R. E., Wolff, W. F., 1991. Ecosystem flow networks: Loaded dice? Mathematical Biosciences 103, 45–68.

Vadas Jr., R. L., 1990. The importance of omnivory and predator regulation of prey in freshwater fish assemblages of North America. Environmental Biology of Fishes 27, 285–302.

Vasas, V., Jordán, F., 2006. Topological keystone species in ecological interaction networks: Considering link quality and non-trophic effects. Ecological Modelling 196, 365–378.

Vázquez, D. P., 2002. Interactions among introduced ungulates, plants, and pollinators: A field study in the temperate forest of the southern andes. Ph.D. thesis, University of Tennessee, Knoxville, Tennessee, USA.

Wassermann, S., Faust, K., 1994. Social Network Analysis. Cambridge University Press, Cambridge.

Watts, D., Strogatz, S. H., 1998. Collective dynamics of ‘small-world’ networks. Nature 393, 440–442.

Whipple, S. J., 1998. Path-based network unfolding: A solution for the problem of mixed trophic and non-trophic processes in trophic dynamic analysis. Journal of Theoretical Biology 190, 263–276.

Whipple, S. J., Patten, B. C., 1993. The problem of nontrophic processes in trophic ecology - toward a network unfolding solution. Journal of Theoretical Biology 163, 393–411.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics Bulletin 1, 80–83.

Williams, R. B., 1971. Systems Analysis and Simulation in Ecology. Vol I. Academic Press, New York, Ch. Computer simulation of energy flow in Cedar Bog Lake, Minnesota based on the classical studies of Lindeman, pp. 543–582.

Williams, R. J., Berlow, E. L., Dunne, J. A., Barabási, A.-L., Martinez, N. D., 2002. Two degrees of separation in complex food webs. Proceedings of the National Academy of Sciences of the United States of America 99, 12913–12916.

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

Williams, R. J., Martinez, N. D., 2004. Limits to trophic levels and omnivory in complex food webs: theory and data. American Naturalist 163, 458–468.

Winemiller, K. O., 1990. Spatial and temporal variation in tropical fish trophic networks. Ecological Monographs 60, 331–367.

Woodward, G., Hildrew, A. G., 2001. Invasion of a stream food web by a new top predator. Journal of Animal Ecology 70, 273–288.

Wootton, J. T., 1994. The nature and consequences of indirect effects in ecological communities. Annual Review of Ecology and Systematics 25, 443–466.

Wulff, F., Ulanowicz, R. E., 1989. A comparative anatomy of the Baltic Sea and Chesapeake Bay ecosystems. In: Wulff, F., Field, J., Mann, K. (Eds.), Network analysis in marine ecology - methods and applications. Vol. 32 of Coastal and Estuarine Studies. Springer-Verlag, New York.

Yodzis, P., 1981. The stability of real ecosystems. Nature 289, 674–676.

Yodzis, P., 1984. How rare is omnivory? Ecology 65, 321–323.

Yodzis, P., Winemiller, K. O., 1999. In search of operational trophospecies in a tropical aquatic food web. Oikos 87, 327–340.

Zorach, A. C., Ulanowicz, R. E., 2003. Quantifying the Complexity of Flow Networks: How many Roles are there? Complexity 8, 68–76.