

Quantitative methods for ecological network analysis

Robert E. Ulanowicz*

Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, MD 20688-0038, USA

Received 2 September 2004; accepted 5 September 2004

Abstract

The analysis of networks of ecological trophic transfers is a useful complement to simulation modeling in the quest for understanding whole-ecosystem dynamics. Trophic networks can be studied in quantitative and systematic fashion at several levels. Indirect relationships between any two individual taxa in an ecosystem, which often differ in either nature or magnitude from their direct influences, can be assayed using techniques from linear algebra. The same mathematics can also be employed to ascertain where along the trophic continuum any individual taxon is operating, or to map the web of connections into a virtual linear chain that summarizes trophodynamic performance by the system. Backtracking algorithms with pruning have been written which identify pathways for the recycle of materials and energy within the system. The pattern of such cycling often reveals modes of control or types of functions exhibited by various groups of taxa. The performance of the system as a whole at processing material and energy can be quantified using information theory. In particular, the complexity of process interactions can be parsed into separate terms that distinguish organized, efficient performance from the capacity for further development and recovery from disturbance. Finally, the sensitivities of the information-theoretic system indices appear to identify the dynamical bottlenecks in ecosystem functioning.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Ascendency; Cycling in ecosystems; Ecosystem status; Flow networks; Information theory; Input–output analysis; Nutrient limitation; Trophic levels

1. Why network analysis?

Ecologists have precious few tools at their disposal to represent phenomena that transpire at the level of the whole-ecosystem. Yet, they increasingly are being exhorted to approach environmental problems at the level of the whole ecosystem (NSF, 1999). To date, the most common tool for quantifying systems-level events is simulation modeling. Before one can simulate an ecosystem, it is necessary first to identify the relevant taxa that comprise it. Thereupon, the investigator must parse out the significant interactions among those taxa. It is only after these preliminaries that the actual modeling commences, as the modeler then quantifies each such interaction in algorithmic fashion. The aggregated formulae are then executed under an appropriate shell on some suitable platform. Finally, there follow the arduous and non-

trivial tasks of calibrating the model according to some set of system observations and validating it against another independent set of data.

Unfortunately, grave problems, both practical and conceptual, beset such whole-ecosystem simulation modeling. On a practical level, the results of the endeavor appear to leave much to be desired (Sheffer and Beets, 1994). Which is not to ignore some successful models consisting of one or a few processes, but as the number of interacting processes increases, problems multiply disproportionately (Platt et al., 1981). Some argue that difficulties in calibrating and validating simulation models arise from the propagation of errors across the nexus of interacting processes. Whence, more accurate and precise values for the model parameters should ameliorate the problem. Others, however, are convinced that interacting nonlinear processes inevitably lose their power to predict (Lorenz, 1963). Furthermore, prediction ability wanes as the interacting processes increase in number and/or nonlinearity (Ulanowicz, 1979). Hence, in order to avoid

* Tel.: +1 410 326 7266; fax: +1 410 326 7378.

E-mail address: ulan@cbl.umces.edu.

pathologies that inevitably crop up in the behaviors of coupled process models, biologically unrealistic assumptions are made out of mathematical convenience (to maintain stability), thereby further impairing prediction ability.

It seems frustrating that nature thwarts attempts to model ecosystems accurately, and thereby control them, but modelers seldom draw the conclusion that the conceptual framework that supports the simulation process is itself flawed. After all, multi-process models work quite well for physical and chemical systems, and the modeling process is built upon the selfsame assumptions that guide the larger body of science—namely, closure, atomism, reversibility, determinism and universality (Depew and Weber, 1994; Ulanowicz, 1999). Why should matters go suddenly awry with ecosystems? It happens that, for a while now, proponents of “deep-ecology” have maintained that ecology is truly different from other disciplines (Naess, 1988); that it forces the investigator to alter profoundly his/her conceptions about how nature operates. Only recently have some of the features been articulated that distinguish ecological phenomena from more fundamental physical and chemical events (Ulanowicz, 1999). These matters are beyond the scope of this exposition, suffice it here to mention: (a) an inversion in the magnitudes of characteristic lifetimes (and with it the direction of causality) as one passes from the individual organism to the ecosystem, and (b) the singular nature of and crucial role played by aleatoric events over the histories of ecosystems.

Leaving these intriguing academic issues aside, the very practical question arises, “If simulation modeling is less than satisfactory, what alternative approach might culminate in a deeper and more satisfactory understanding of whole ecosystems?” Almost a quarter century ago SCOR Working Group 59 (Platt et al., 1981) suggested that systems ecologists simply forego the latter stages of the modeling process altogether. They suggested that one concentrate instead on what can be inferred from the identification and parsing tasks alone; that one pay more attention to processes (flows) than to objects (stocks). This advice eventually was acted upon by a successor working group (#73) that assembled an inchoate set of analytical tools which collectively became known as “ecosystem network analysis” (ENA) (Wulff et al., 1989).

An assumption underlying ENA has been that the configuration of processes represents the “anatomy” of the ecosystem; and that, as in medical practice, such anatomy will reveal much about the history, current status and workings of the ecosystem.¹ The remainder of this work will discuss how to create ecological flow networks and analyze them in systematic fashion.

¹ It should be noted in passing that an independent school of investigations into the structure of “foodwebs” has evolved in parallel with ENA (e.g., Yodzis, 1989; Cohen et al., 1990; Polis and Winemiller, 1995). Foodweb research treats unweighted (binary) networks of trophic interactions, whereas ENA addresses the [sometimes large] differences in the magnitudes of the connections.

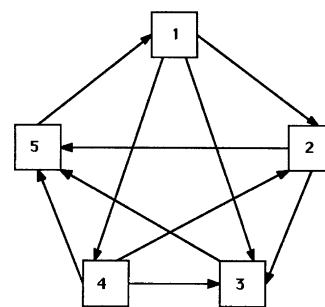


Fig. 1. A directed flow graph.

2. Required information

A working definition of an ecological network is that it is a representation of the answers to two questions: (1) who eats whom?, and (2) at what rate? As with simulation modeling, preliminary to the first question it becomes necessary to identify the significant taxa or *nodes* comprising the ecosystem. For each of the taxa one then needs to know which of the other nodes are present in its diet. Once this qualitative information is known for all taxa, the result can be represented in one of two ways: (a) one can present the system graphically as a directed graph, or digraph. On the digraph the nodes are usually represented as boxes, and each transaction is represented as an arrow that originates out of the prey taxon and terminates (with an arrowhead) at the predator node (Fig. 1); (b) the connection topology among n nodes can likewise be represented as an $n \times n$ square *adjacency matrix*, where a one in the entry for row i and column j means that an amount of material flows from predator i to prey j .² By contrast, a zero entry signifies that no palpable transfer occurs between i and j (hence the notion of a *binary network*). The advantage of using the matrix approach and associated linear algebra is that one may then deal with systems of arbitrary dimension.

Now, ecosystems are necessarily open, meaning that they exchange material and energy with their surroundings. These exogenous transfers require that one emend both the graphical and analytical representations. Without loss of generality, one can assume that all inputs to a particular taxon be lumped as a single entry. (If such is not the case, multiple types of inputs can be treated as will presently be discussed for exogenous outputs.) Actual physical inputs take the form of primary production, immigration or inbound advection of material or energy. The lumped input is represented graphically by an arrow that originates out of no visible taxon and terminates (with an arrowhead) into the actual receiving node (Fig. 2).

For reasons that will become obvious presently, it is useful to distinguish between two different types of exogenous outputs: the first is the export of material or energy still useful

² Several other investigators (e.g., Patten, 1985) reverse the order of the subscripts. That is, flow is in the direction from column j to row i .

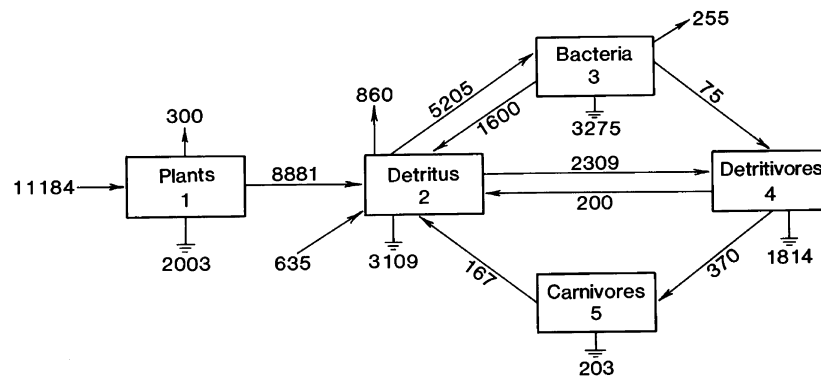


Fig. 2. The trophic exchanges of energy ($\text{kcal m}^{-2} \text{y}^{-1}$) in the Cone Spring ecosystem (Tilly, 1968). Arrows not originating from a box represent exogenous inputs. Arrows not terminating in a box portray exogenous outputs. Ground symbols represent dissipations.

to other ecosystems of comparable scale. Examples of such exports are emigration, harvesting by humans, and advection out of the system. Similar to the external inputs, these flows can be represented as arrows that originate from the given taxon but terminate in empty space (Fig. 2). In addition to useful exports, some energy may be dissipated into heat (respiration) or some material may be degraded into its lowest-energy form (e.g., denitrification to N_2). The occurrence of such dissipations is required by the second law of thermodynamics for energy. With pictorial representations (Odum, 1971), the common convention is to depict dissipation leaving a compartment by the “ground symbol” of electrical circuit symbology (Fig. 2).

For an ecosystem of n taxa, the maximum number of flows possible under the typology just described is $n^2 + 3n$. Whenever n exceeds about 20, the number of actual transfers usually falls well below this limit. With systems having many compartments, the number of nonzero entries into the adjacency matrix usually is less than 15% of n^2 . Hence, one is dealing mostly with sparse matrices.

The second query, “at what rate?”, requires the investigator to quantify each palpable exchange. Quantification usually is in terms of a single conservative medium. That medium may be energy, or it could be any one of the chemical elements, such as C, N, P, S, etc. ENA usually allows only one medium per network (although some investigators allow several (e.g., Hannon et al., 1991)). In systems where the dynamics of several elements appear to be significant, the practice is to represent each medium by its own network, and the analysis of multiple parallel networks is covered in Section 7. Till then, all systems will be quantified using a single medium.

Having chosen a suitable medium, the investigator can proceed with the quantification of the flows. For this purpose, the rate of the internal transfer from donor (prey) taxon i to recipient (predator) taxon j will be represented as T_{ij} . The rate of exogenous input to taxon i will be denoted by X_i ; the rate of loss of useful medium from taxon i to the outside world will be signified by E_i ; and the dissipation from i will be R_i . If bookkeeping is kept over a sufficiently long interval, medium is likely to be balanced around each compartment. That is,

for each taxon, i ,

$$X_i + \sum_{j=1}^n T_{ji} = \sum_{k=1}^n T_{ik} + E_i + R_i \quad (1)$$

or

$$X_i + T_{i.} = T_{i.} + E_i + R_i,$$

where a dot has been used to indicate summation over the full range of an index. (Henceforth, the sigma convention will be used only when necessary for clarity.)

At any particular time, however, the inputs to each taxon may not be balanced by its outputs. Some of the methods to be discussed, such as the analysis of cycling (Section 5) or the information-theoretic indices (Section 6) do not require that components be balanced. The assessment of indirect effects (Section 3) and the description of trophic levels (Section 4) can, however, yield spurious results whenever any taxon is not in steady state.

Another stratagem is to define separate vectors to represent instantaneous storages (G) and loss (L) of biomass, respectively. Then the artificial balance becomes

$$X_i + G_i + T_{i.} = T_{i.} + L_i + E_i + R_i.$$

Ideally, one determines of the magnitude of each rate of transfer from measurements made in the field. For some flows that characterize the activity of the system, such as the primary productions, this is usually feasible and is highly recommended. Because of the large number of transfers occurring in most systems, however, it is usually unfeasible to reckon the magnitudes of all flows directly. It becomes necessary, therefore, to estimate the magnitudes of some rates by indirect means. One helpful device for estimating unknown flows is to assume the balance of inputs and outputs around each taxon.

The estimation of a balanced flow network is an art form that cannot readily be formalized in all cases. There are some useful guides for the novice, however (Ulanowicz et al., 1997). It is ironical, for example, that a methodology that so emphasizes processes over objects, should use as its starting

point some measure of the biomass or density of each taxon. Because it is so much more difficult to measure metabolic rates in the field than it is to estimate the sizes of each population, the common procedure has become the assessment of the stock density of a taxon (in terms of the chosen medium) so that that stock can be multiplied by an appropriate tabulated metabolic quotient to obtain the desired rate. For example, if one should measure a microzooplankton density of 150 mg C m^{-2} and discover from published reports that the average consumption rate by these organisms is 160% per diem, then the total consumption (demand) by these organisms could be roughly estimated as $87,600 \text{ mg C m}^{-2} \text{ y}^{-1}$. Similarly, one would search for physiological quotients to quantify the losses due to respiration, excretion and natural mortality. By difference, one is able to arrive at the net productivity available for consumption by predators. Balancing the network thereafter becomes a task of matching up the estimated demands to the corresponding availabilities over all indicated predator–prey interactions.

There exist several widely-used optimization routines that have been written to create such a balanced network automatically (Polovina, 1985; Christensen and Pauly, 1992; Vezina and Platt, 1988; Kavanagh et al., 2004). The problem with optimization algorithms, however, is that they tend to yield solutions that lie on one of the vertices of the constraint polyhedron. That is the mathematician's way of saying that the solution almost always zeroes out one or more of the flows known to pass into or out of each taxon. Some balancing algorithm that allows all the known exchanges to remain in the eventual apportionment is to be preferred. Because such an algorithm would involve fewer inferences than are implicit in an optimization routine, the result is referred to as a least-inference distribution.

Recently, two potential least-inference schemes have been implemented (Ulanowicz and Scharler, in preparation). The first of these, MATBLD, begins by assigning a flow magnitude to each link in joint proportion to the availability of the given prey and the demand by the corresponding predator. If the matrix of flows were fully connected, this joint apportionment would directly yield a balanced network. Ecosystem networks are usually sparsely connected; however, and it becomes necessary to inflate these joint proportions by a constant factor to yield a magnitude for the total system activity that is commensurate with the aggregated demands of all components. Even after such adjustment, the network is still not balanced, and one must invoke an equilibrating routine, such as the donor-controlled (inherently stable) balancing scheme of Allesina and Bondavalli (2003).

A second alternative, called MATLOD, “loads” the network matrix by incrementing the set of all designated flows by some uniform very small amount. After each flow has been incremented, the routine checks to see whether the availability of the prey has been exhausted or the total demand of the predator has been satisfied. If either of these checks is true, that particular flow is removed from the list of those that will receive additional increments. Matrix loading termi-

nates when the list of flows to be incremented becomes empty. This process, which resembles “mold-filling” will match the aggregate demands with the availabilities insofar as possible. Final balancing is then achieved via the Allesina and Bondavalli (2003) equilibration routine.

Comparisons of networks balanced using MATBLD with those starting with the same data and using MATLOD reveal that results differ by statistically insignificant amounts. Perhaps more interestingly, either method yields results that do not differ from published networks to any statistically significant degree (Ulanowicz and Scharler, in preparation). If true, once one obtains field estimates of population densities and a predator–prey topology, one should be able to find in the literature appropriate physiological quotients needed to invoke either MATBLD or MATLOD to create a complete, balanced, quantified network of exchanges.

3. Indirect effects

By following the methods just outlined, an estimate of the complete trophic flow structure can be achieved for most ecosystems. Many, however, might question what could possibly be accomplished with such a network? This is a legitimate question, considering how most graphical representations of networks of even moderate dimension (say 30–40 nodes) resemble tangled “spaghetti diagrams”, and their corresponding matrices fill screens and printouts with a welter of numbers. Indeed, what sense can be made of these constructs?

Fortunately, through the systematic application of linear algebra, a full picture of all the indirect trophic effects and the overall trophic structure can be realized. To see how this is done, it is helpful to turn to matrix notation. One begins by defining a matrix of dietary proportions, $[G]$.³ The elements of $[G]$ are obtained from the elements of the flow matrix, $[T]$, and the input vector, (X) , by normalizing the inter-compartmental exchanges using the total input to the receiving compartment, j ,

$$g_{ij} = \frac{T_{ij}}{(T_{\cdot j} + X_j)} \quad (2)$$

That is, the element g_{ij} represents that fraction which i comprises of the total intake by j . Reading down column j of $[G]$, one encounters the respective percentages that each dietary item i constitutes of the full intake by j .

The matrix $[G]$ and its algebraic powers have very didactic meanings as regards trophic functioning. For example, in the very simple network in Fig. 3, the six non-zero values of T_{ij}

³ The convention here will be to denote square or rectangular matrices as capital letters contained within square brackets, e.g., $[G]$, $[A]$, etc. Column vectors will be denoted by capital letters within parentheses, such as (X) or (E) . The superscript “T” after any matrix or vector signifies its transpose. That is, $[G]^T$ is a matrix formed by transposing the rows and columns of matrix $[G]$. Similarly, $(E)^T$ is the row-vector transpose of column vector (E) .

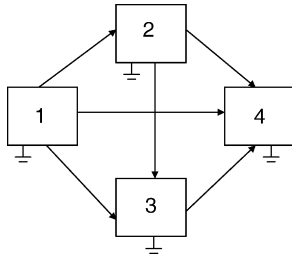


Fig. 3. A simple hypothetical network of four components.

give rise to a corresponding six elements in the matrix $[G]$.

$$[G] = \begin{bmatrix} 0 & g_{12} & g_{13} & g_{14} \\ 0 & 0 & g_{23} & g_{24} \\ 0 & 0 & 0 & g_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Multiplying the matrix $[G]$ by itself gives the following result:

$$[G]^2 = \begin{bmatrix} 0 & 0 & g_{12}g_{23} & (g_{12}g_{24} + g_{13}g_{34}) \\ 0 & 0 & 0 & g_{23}g_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The reader's attention is drawn to the fact that each of the non-zero elements of $[G]^2$ corresponds to the collection of pathways of length 2 that connect i with j . For example, the 1–3 element of $[G]^2$ reveals how much gets to 3 from 1 over the two step pathway $1 \rightarrow 2 \rightarrow 3$. The 1–4 element is comprised of two terms, signifying there are two pathways of length 2 connecting 1 with 4. The first term reports the fraction of what reaches 4 from 1 over the pathway $1 \rightarrow 2 \rightarrow 4$, while the second term gives the corresponding fractional input to 4 that traveled via the route $1 \rightarrow 3 \rightarrow 4$.

Multiplying $[G]^2$ by $[G]$ once more yields the matrix $[G]^3$,

$$[G]^3 = \begin{bmatrix} 0 & 0 & 0 & g_{12}g_{23}g_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

As can be seen, there is a single non-zero element of $[G]^3$, corresponding to the only three step pathway in the graph, $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$. The sequence of powers of $[G]$ truncates with $[G]^4 = [0]$, as there are no pathways >3 in the network.

The reader is encouraged to experiment with other simple graphs to ascertain that the elements of the m th power of $[G]$ contain contributions from each and every pathway in the graph of exactly length m .

Because the elements of $[G]$ have been normalized such that each $g_{ij} \leq 1$, it is probable that the elements in the higher powers of $[G]$ will grow progressively smaller. This, in turn,

prompts the question of whether the infinite series,

$$[I] + [G] + [G]^2 + [G]^3 + [G]^4 + \dots$$

converges to a finite limit? (The matrix $[I] \{=[G]^0\}$ is the identity matrix. It consists of ones along its diagonal and zeroes elsewhere.) Simon and Hawkins (1949) were able to show that the normalizations (2) guarantee that the series of matrices converges to the finite limit,

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

This limit, $[S] = [I - G]^{-1}$, is called the Leontief structure matrix. The i - j th component of $[S]$ provides the fraction of the total input to j that came from i over all pathways of all lengths per unit of final demand. Final demands play a key role in economic theory, and the discovery of the $[S]$ matrix enabled economists to estimate the productions in various economic sectors that were necessary in order to satisfy any vector of final demands. The formulation of $[S]$ by Leontief (1951) won for him the 1973 Nobel Prize in Economics for what eventually became known as "input-output" (I/O) theory. About the same time that Leontief received his prize, Hannon (1973) was introducing Leontief's methods into ecology.

Although final demands are central to economic planning, their homologs in ecology take the form of respiratory losses. Of course, respiration is important in ecology, but it is the intermediate transfers among system components that interests ecologists even more. That is, ecologists are more eager to know how much of what eventually arrives at j was the result of a particular activity i ? Szyrmer and Ulanowicz (1987) showed how the intermediate transfers could be reckoned via a convenient transformation of the $[S]$ matrix. Szyrmer called his result the "total dependency matrix", $[D]$, where the elements of the $[D]$ matrix could be calculated from $[S]$ and the original flows as

$$d_{ij} = (s_{ij} - \delta_{ij}) \left(\frac{\sum_{k=1}^{n+2} T_{ik}}{s_{ii} \sum_{m=0}^n T_{mj}} \right),$$

where δ_{ij} are the elements of the identity matrix, T_{0j} are the exogenous inputs to j , $T_{i,(n+1)}$ represent useable exports from i to other systems of comparable scale, and $T_{i,(n+2)}$ are the dissipative losses from i . Another way of looking at element d_{ij} is that it answers the question, "What fraction of the total diet of j passes through i along its way to j ?" Viewed in this way, the j th column of $[D]$ can be interpreted as the "indirect diet" of j , that is, the respective amounts by which j depends upon the activity of each element in the ecosystem. Because medium can visit multiple compartments on its way to j , the elements of column j sum to ≥ 1 , and the sum is indicative of the effective trophic position of j .⁴

⁴ Another useful index of system status is the average path length, APL, defined by Finn (1976) as the average number of transfers an arbitrary quantum of medium makes while traversing the system. It can be calculated simply by dividing the sum of all system inputs by the total system throughput, i.e., $APL = (T_{\bullet\bullet}/T_{0\bullet})$.

The matrix of indirect diets is sometimes useful for differentiating trophic roles. For example, in the Chesapeake mesohaline ecosystem, Striped Bass (*Morone saxatilis*) and Bluefish (*Pomotatus saltatrix*) are both piscivorous predators and one would expect them to compete heavily with each other. Baird and Ulanowicz (1989) calculated the indirect diets, $[D]$, of all major compartments of the Chesapeake system, revealing (among others) the following indirect dependencies:

Striped Bass on Zooplankton (%)	65.8
Bluefish on Zooplankton (%)	28.7
Striped Bass on Polychaetes (%)	1.8
Bluefish on Polychaetes (%)	48.0

The results indicated how the Striped Bass were ultimately highly dependent on pelagic production, whereas Bluefish owed their activity more to benthic secondary production.

Of course, normalizing $[T]$ down its columns is but one way to proceed. One could have as readily normalized the rows of $[T]$ to obtain a matrix of host coefficients, $[F]$,

$$f_{ij} = \frac{T_{ij}}{(T_i + E_i + R_i)}.$$

That is, f_{ij} is the fraction of the total activity of i that flows directly to element j . The corresponding output structure matrix, $[\Sigma] = (I - F^T)^{-1}$, was formulated by the economist Augustinovic (1970) to estimate the activity of each sector j that is generated by a unit of primary input to i . As with final demands, primary inputs are of substantial interest to ecologists, but at times they might prefer to know the intermediate contributions of all the compartments. That is, how much of what leaves i eventually flows as input to j ? Accordingly, Szyrmer defined the total contribution matrix, $[C]$, with the elements⁵

$$c_{ij} = (\sigma_{ji} - f_{ji}) \left(\frac{T_i}{\sigma_{ii} T_{i,j}} \right).$$

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. In particular, one might want to use as an indicator of overall ecosystem efficiency how much primary production makes its way through the ecosystem and finally reaches the upper trophic components. For example, Ulanowicz (1984) compared networks of trophic exchanges in two tidal marsh ecosystems near the Crystal River nuclear power plant on the west coast of Florida. One network was adjacent to the thermal outfall from the plant ($\Delta T \approx 6^\circ\text{C}$), and the second, virtually identical marsh (control), was distant from any thermal influence. Calculation of the total contribu-

tion coefficients of the primary producers to the Gulf Killifish (*Fundulus grandis*) and two Needlefish (*Strongylura marina* and *S. notata*) yielded the following results:

	Control	ΔT
Gulf Killifish	0.147E-02	0.670E-03 (−54%)
Needlefish	0.338E-03	0.140E-03 (−59%)

One concludes that the impact of the thermal effluent was to decrement the overall efficiency of the ecosystem at producing top carnivores by some 55–60%.

One frequent criticism of I/O methods is that they pertain only to steady-state or temporally-averaged configurations. There have been several efforts to extend I/O theory to non-steady-state systems. Matis et al. (1979) treated the imbalance as another category of flow, while Matis and Patten (1981) normalized the flows by the storages of the source compartments and inflated the diagonal elements to reflect the relative storage times. This author prefers to treat temporal dynamics by applying information-theoretic methods to time series of network snapshots (see Section 7 below).

Another deficiency of conventional input–output analysis is that it deals only with positive flows of medium. Although contributions and dependencies are key features of ecosystem dynamics, ecologists are also interested in the negative effects of predation and how these might propagate throughout the system. Ideally, one wishes to follow simultaneously both the positive and negative impacts of trophic exchanges. Fortunately, the linear algebra of I/O analysis lends itself to evaluating the net trophic impact (+ or −) of any one species upon any other (Ulanowicz and Puccia, 1990). For starters, one notes as how g_{ij} quantifies the positive impact of prey i upon predator j . The corresponding f_{ij} , however, seems to underestimate the negative impact that j has upon i , because it has been normalized by the total output from i , rather than just the predatory losses that i sustains. Accordingly, f_{ij} is normalized as

$$f_{ij}^* = \frac{T_{ij}}{\sum_{m=1}^{nl} T_{im}},$$

where the index m is summed only over the subset of living members ($nl \leq n$) of the n ecosystem components. One can then define the net direct effect that i has upon j as

$$q_{ij} = g_{ij} - f_{ji}^*,$$

where $-1 \leq q_{ij} \leq 1$.⁶

In almost all instances, the series consisting of the powers of $[Q]$ converges, just as the powers of $[G]$ converge in conventional I/O analysis:

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

⁵ In general $d_{ij} \neq c_{ij}$, as they are the answers to slightly different questions. At steady-state, $d_{ii} = c_{ii}$, however.

⁶ Fath and Patten (1998) retain f_{ij} as normalized by the total flow through i , which makes the elements of their $[Q]$ matrix (on average) more positive. See also Fath and Borret (in review).

(In rare [usually degenerate] instances the matrix $[I - Q]$ can be singular, and no convergence obtains). The overall net matrix of total impacts, $[M]$, then $= [I - Q]^{-1} - [I]$. For any connected graph, the matrix $[M]$ will possess the full complement of n^2 non-zero components. This jumble of impacts would become overwhelming (e.g., Hillebrand and Shurin, *in press*) without some systematic way of presenting the results. The original algorithm for indirect trophic influences is called IMPACTS, and it allows the user to focus upon a particular compartment by ranking the impacts the given species has on other system components, along with the impacts that other species exert on the compartment in question.

Earlier, when he first addressed temporal changes in ecological networks, Patten (1982a) surmised a tendency for negative, competitive direct interactions to become positive indirect accommodations or cooperation. Fath and Patten (1998) demonstrated how the elements of $[M]$ are usually more positive than the direct effects, $[Q]$. As noted above, the way Fath and Patten normalize their $[Q]$ matrix makes the elements of their version of $[M]$ more positive than the ones defined here. These differences notwithstanding, it is easy to cite instances of negative direct interactions becoming positive indirect influences, even under the more conservative assumptions used above (Bondavalli and Ulanowicz, 1999).

4. Trophic considerations

The concept of “trophic level” arose from the very simplistic image of an ecosystem as a trophic chain or pyramid. That is, energy is fixed by autotrophs or primary producers at the first level, and some (but not all) of those resources are transmitted to the herbivores at the second level. Further transfer occurs to carnivores at the third trophic level and to top carnivores at the fourth. There is occasional mention of tertiary carnivores, but hardly ever any reference to anything beyond the fifth level. It was originally thought that losses of energy and material at each transfer limited trophic chains to less than five members, but that causal inference turned out to be problematical (Pimm and Lawton, 1977).

In any event, nature is almost never so simple as the chain or pyramid just described, and considering the full manifold combinations of circumstances soon becomes impractical (Hillebrand and Shurin, *in press*). One must either devise a systematic way of applying the trophic level concept to complicated webs or abandon the notion altogether. While some species do behave as obligate autotrophs or herbivores, many heterotrophs feed at multiple levels, and no simple one-for-one mapping of taxa to integer trophic levels appears to be feasible. Cousins (1985), for example, points out that a hawk can feed at five different trophic levels. Thus, one must either emend the notion of trophic level to apply to the continuum of trophic positions in complicated networks, or else one might seek some way of mapping the actual web into an abstract chain of integer trophic levels. It happens the tools

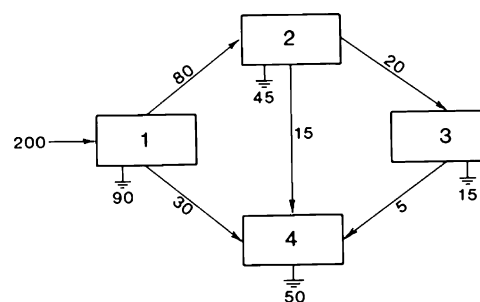


Fig. 4. A hypothetical trophic network. Units are arbitrary.

of linear algebra developed for I/O analysis make both tasks possible.

In the previous section, it was mentioned how the sums of the columns of the structure matrix, $[S]$, were related to the number of trophic transfers that medium had incurred along its way to the target compartment. Although this sum is generally not an integer, Levine (1980) suggested that it be regarded as the average or effective trophic level at which that particular taxon is feeding. A simple demonstration of the intuitive appeal of this definition appears in Fig. 4. In this hypothetical example, compartments 1, 2 and 3 are connected in chain-like fashion. Compartment 4, however, receives only 5 of its 50 units of total activity at the fourth trophic level. It receives 30% of its sustenance at the third level and 60% at the second. Its effective trophic level becomes $(0.6 \times 2) + (0.3 \times 3) + (0.1 \times 4) = 2.5$. It is left as an exercise for the reader to calculate the $[G]$ matrix and the structure matrix, $[S]$ for this network. The sums of the first three columns of the resulting $[S]$ are 1.0, 2.0 and 3.0, respectively, whereas the fourth column sums to 2.5. The calculation is quite general and applies to even the most complicated flow topologies, regardless of the number and position of cycles in the system. Christian and Luczkovich (1999) combined this notion of effective trophic level with the trophic impact analysis just discussed to help establish the direction of causality (bottom-up versus top-down) at work in aquatic vegetation ecosystems in St. Mark's Refuge, Florida.

The obverse task of apportioning the activities of species into virtual integer trophic levels requires somewhat more complicated manipulations (Ulanowicz and Kemp, 1979). One begins by noting that the fractions by which each compartment feeds directly from primary sources are

$$g_{0i} = \frac{T_{0i}}{\sum_{m=0}^n T_{mi}},$$

which elements can be arrayed as a row vector, call it $(L_1)^T = (g_0)^T$. That is, the elements of this row vector reveal the fraction of the activity of each taxon that consists of primary production (assuming for the moment that all exogenous inputs correspond to primary production. The more general case will be covered in a footnote below). To calculate the amounts by which each taxon feeds as a herbivore, one simply multiplies $[G]$ by $(L_1)^T$ on the left to generate

the row vector $(L_2)^T = (L_1)^T[G]$. Similarly, the fractions by which each compartment acts as a carnivore (the third level) become $(L_3)^T = (L_2)^T[G]$. By induction, one may calculate the fractions at the m th trophic level as $(L_m)^T = (L_1)^T[G]^{m-1}$. If there are no cycles in the network, the sequence of row vectors, $(L_1)^T, (L_2)^T, (L_3)^T, \dots$ will truncate (become all zeroes) in $n - 1$ or fewer steps (because $[G]^{m-1}$ will become a matrix of zeroes, once no pathways of length $m - 1$ exist). Therefore, when cycling is absent, one may define a Lindeman trophic transformation matrix, $[L]$, such that the i th row of $[L]$ is $(L_i)^T$. Reading down the j th column of $[L]$ reveals the fraction by which the j th taxon feeds at each corresponding trophic level. Reading across the i th row provides the composition of that virtual trophic level.

The simplistic network in Fig. 4, for example, yields the Lindeman trophic transformation matrix,

$$[L] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0.6 \\ 0 & 0 & 1 & 0.3 \\ 0 & 0 & 0 & 0.1 \end{bmatrix}.$$

Reading down the columns, one sees as how the first three compartments act entirely at their respective trophic levels, whereas compartment 4 is partitioned as described above. Reading across the rows, one encounters those fractions that the activity of component 4 contributes to the throughflows of the second, third and fourth virtual trophic levels.

At this point, the reader might interject that almost all trophic networks contain cycles. This is a valid observation; however, Pimm (1982) observed how cycles strictly among living, feeding taxa are relatively rare and usually are small in magnitude. Put in other words, almost all cycles in ecosystems contain at least one non-living compartment. If, then, one orders the ecosystem's components so that all living compartments (say nl of them) appear first, then the $nl \times nl$ submatrix of transfers among the living compartments should contain a very small number of cycles having very insignificant activities. It will be described in the next section how cycles can be removed from networks without altering the balance around any of the members of the cycle. If all of these cycles are small in magnitude (which is almost always the case), then the cycles among the living components can be removed without changing the $nl \times nl$ submatrix to any significant degree. Because there are no longer any cycles among the first nl taxa, one can now define an $nl \times nl$ Lindeman trophic transformation matrix for the living components as just described (Ulanowicz, 1995).⁷

Having devised a scheme to apportion the activities of all the original compartments among a set of integer trophic levels, it now becomes possible to apportion the individual flows along a chain of transfers linking the aggregated, integer

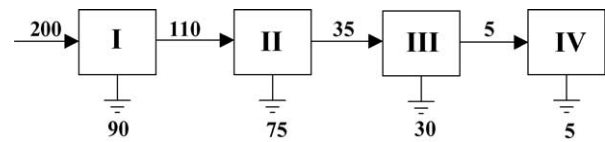


Fig. 5. The Lindeman Spine of the network in Fig. 4.

trophic levels. This follows by treating $[L]$ as a conventional transform in linear algebra. Thus, one defines the canonical inputs⁸, (Ψ) (Greek upper case “psi”) as

$$(\Psi) = [L](X),$$

In similar fashion, the useful exports, (E) , and dissipations, (R) map as

$$(Y) = [L](E),$$

and

$$(P) = [L](R).$$

(Y) and (P) (Greek upper case “*psilon*” and “*rho*”) are the “canonical” exports and respirations, respectively.

It should then follow that the matrix of flows $[T]$ would transform as

$$[\Theta] = [L][T][L]^T.$$

There is a problem, however, in that $[L]$ is not orthonormal, so that the transformed matrix of flows, $[\Theta]$, is not diagonal in form. All is not lost, however, because $[\Theta]$ possesses a peculiar form of symmetry that allows one to abstract from it a surrogate diagonal matrix, $[\mathcal{E}]$. $[\mathcal{E}]$ has as its i th diagonal element the sum of the i th row of $[\Theta]$ and zeroes elsewhere, i.e., $\xi_{ii} = \sum_k \theta_{ik}$, and $\xi_{ij} = 0$ otherwise. Then homologous to equation (1), one can balance around each integer “trophic level”, λ ,

$$\Psi_\lambda + \xi_\lambda = \xi_{(\lambda+1)} + \nu_\lambda + \rho_\lambda.$$

For example, the Lindeman matrix $[L]$ shown above for the simple network in Fig. 4 transforms that network into the equivalent straight-chain in Fig. 5.

Again, the canonical transformation works in straightforward manner whenever there are no cycles in the entire network. As mentioned earlier, there are almost always significant cycles in the network as a whole, but almost all involve at least one non-living compartment. Thus, arranging the n compartments so that the nl living species appear first allows one to remove all of the insignificant cycles comprised of only living predators. One may then treat the revised $nl \times nl$ acyclic submatrix to create an $nl \times nl$ Lindeman matrix, $[L']$, for the living species alone. To create a Lindeman transformation of the entire system (Fig. 6), one begins with $[L']$ as

⁷ It should be noted here how Higashi et al. (1991) elect *not* to remove the cycles from the networks they treat, with the result that their trophic chains do not truncate at some level $< n$, but continue on indefinitely.

⁸ The reader will recall that only those components of (X) that represent primary production are employed in the definition of $[L]$. As a result, all the primary production figures will map into the first component of (Ψ) , while any inputs that occur at higher trophic levels will be apportioned to the remaining components of (Ψ) .

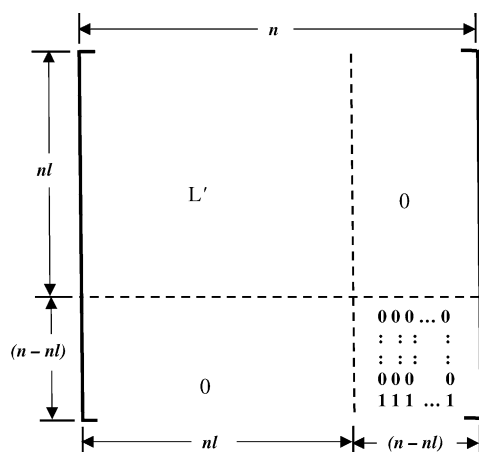


Fig. 6. The partitioned expansion of the $[L']$ matrix.

the principal submatrix within an expanded $n \times n$ matrix, $[L]$, occupying rows 1 through nl and columns 1 through nl . The remainder of the matrix is filled out with zeroes, save for the last $[n - nl]$ elements of the last row, which are all set equal to one.

After transformation by $[L]$, the $nl \times nl$ submatrix creates what is called a “Lindeman Spine” of chain-like flows. The remainder of the matrix serves to aggregate all the non-living compartments into a single detrital pool, in such a way that all of the flows connecting the detrital pool with the Lindeman Spine balance. As an example of the results, the Lindeman Spine and associated detrital pool corresponding to the network in Fig. 2 is shown in Fig. 7. A peculiarity in Fig. 7 is that the Lindeman Spine is broken between levels I and II. As a result, levels II, III, and IV are due purely to detritivory.

The advantage of transforming into “canonical trophic chains,” is that the results often allow one to observe regularities across disparate ecosystems or, conversely, changes in the same ecosystem at different times. Regarding comparisons among ecosystems, one often encounters canonical trophic chains that exceed 5 levels, ostensibly contradicting the dictum by Pimm and Lawton (1977) that trophic chains do not exceed five steps. Closer consideration, however, reveals that the amounts of medium that the system transfers beyond the 5th level are miniscule, so that any effective measure of trophic length will hardly be affected by the higher levels.

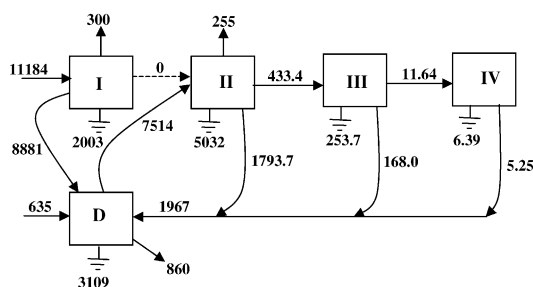


Fig. 7. The canonical representation of the Cone Spring network (Fig. 2).

Hence, when Whipple and Patten (1993) write about activities at very high trophic levels, one has to ask if anything other than a few molecules remain to carry out those activities after so many transfers. As for changes in the canonical chain over time, it appears that environmental perturbations usually act to shorten the length of the chain and diminish disproportionately the flows at higher levels.

Bernard Patten and associates have also developed an environ theory of ecosystems (Patten, 1978, 1982b; Patten and Auble, 1981; Fath and Patten, 1999; Fath and Borrett, in review). An environ constitutes the entire antecedent or subsequent history of the material currently residing in a compartment. An input environ is a quantitative description of where the current biomass has resided since it entered the system and the pathways along which it flowed to arrive at the node in question. Similarly, an output environ describes which other nodes will receive the material in question and the routes it will take to get there. Any given network can be decomposed into a suite of input and output environs. Environs are useful for scoping out the possible causes and effects of any changes in a particular population, in light of the observation that indirect effects are often more significant than direct interactions.

A useful survey of I/O and trophic methods of network analysis can be found in Higashi and Burns (1991).

5. Cycling in ecosystems

Given the scarcity of some chemical elements in the biosphere, and the refractory nature of atomic units, it is inevitable that the same material be used repeatedly by many biotic species. Such necessity can be said to drive the trophic exchanges themselves. Within a given locality, chances are that the same medium will be used more than once by the same biotic species. This is called “recycling” or more often simply “cycling” of medium. If the chemical element in question is scarce and is necessary to many biotic agents, the probability of its reuse (cycling) by the same compartments increases.

The appearance of the words “scarcity” and “necessity” in the previous paragraph implies that cycling in ecosystems is indicative of controls being exercised upon biotic species and processes. If a chemical element is both scarce and necessary, there is an increased likelihood that the loops it makes will take on an autocatalytic nature that rewards each participant for its role in keeping the scarce medium in circulation. It follows that one of the key features of interest in an ecosystem should be the pattern and magnitudes of the trophic pathways by which medium is recycled. The last statement distinguishes two different aspects of the task of understanding cycling in ecosystems. The first is qualitative—one needs to know the topology of pathways over which medium is recycled. The second is quantitative—one seeks to gain some idea of how much material is cycled over each route.

Over the course of ecosystem research, it was the quantitative issue that was first addressed. Odum (1969) had identified

the amount of cycling as one of his 24 criteria for “mature” or developed ecosystems. Patten et al. (1976) and Finn (1976), therefore, employed the I/O analysis that had just been introduced into ecology by Hannon (1973) to quantify that amount. Finn’s reasoning was straightforward: In the Leontief structure matrix, $[S]$, each diagonal element relates to the probability that a quantum of medium visits the designated compartment more than once. Finn suggested that to quantify the degree of recycling in the given network, each diagonal element should be multiplied by the total activity (throughput) of that particular taxon, and that all such products should be summed over all taxa. In time, this sum became known as the “Finn cycling index” (FCI).

Szyrmer and Ulanowicz (1987), while reworking Leontief’s scheme, noted as how the diagonal elements of both the total dependency matrix, $[D]$, and the total contribution matrix, $[C]$, are identical; and that they, rather than the diagonals of $[S]$, represent more accurately the probability that a given quantum leaves a particular taxon and returns to it. They recommended that the diagonal elements of either $[C]$ or $[D]$ replace the diagonal elements of $[S]$ in Finn’s calculations. Han (1997) independently recapitulated Szyrmer’s calculations. Finally, Allesina and Ulanowicz (2004) pointed out how all previous calculations ignore some of the compound cycling that is inherent in the off-diagonal terms. They presented a tedious method for calculating the entire amount involved in cycling. Fortunately, their time-consuming algorithm could be spared, because comparisons showed that the FCI under-represented the actual amount of cycling by an amount that did not vary much from 14% in all the cases examined. These later modifications of Finn’s index essentially were fine-tunings of the original index, which continues to be a useful estimate, when appropriately modified.

The larger problem with the Finn index is that it doesn’t always track the developmental status of an ecosystem. Ulanowicz (1984), for example, showed how perturbing an ecosystem could increase the FCI (see also Wulff and Ulanowicz, 1989). He suggested that the increased amount of cycling was a homeostatic response of the ecosystem to stress. Impacts tended to free medium from storage in the biomass of higher trophic level taxa, and the homeostatic response of the system was to prevent insofar as possible the loss of the resource by cycling it among the lower trophic taxa.

A key element in Ulanowicz’s exegesis of why the Finn index increases with stress was his detailed knowledge of how the pattern of the pathways of recycle had changed. (Longer, slower cycles involving the higher taxa had been replaced by shorter, faster cycles among lower components.) That is, he relied upon the qualitative, topological aspect of cycling (Ulanowicz, 1983). In order to elucidate the topology of cycling, one begins with the task of how to find all the simple cycles in a network. Simple cycles are those in which no element is repeated more than once. Any compound cycle can be viewed as a superposition of simple cycles. It is surprisingly easy to write a short algorithm to find all the simple cycles

Table 1

The order of the nodes in a “depth-first” search

Level	Compartment				
n	[1]	[2]	[3]	...	$[n]$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
2	[1]	[2]	[3]	...	$[n]$
1	[1]	[2]	[3]	...	$[n]$
0	[1]	[2]	[3]	...	$[n]$

in a network. One approach is to perform a depth-first search with backtracking (Mateti and Deo, 1976).

In backtracking algorithms, one orders the nodes in some convenient way (described below) and imagines the same order of n nodes to be repeated at n levels as in Table 1 (Ulanowicz, 1986). One begins at a given node in the zeroeth level (designated as the pivot element) and searches from left to right among nodes in the next level until an existing flow connection is found. One then jumps to the chosen node in the next level and proceeds searching (left to right) the level above it in an attempt to move higher. As one progresses to higher levels, the last node visited in each previous level is stored in a vector array which describes a current pathway. Before advancing to a higher level one checks to make sure that the new node has not already appeared in the current pathway (to confine the search to simple cycles). One ascends to as high a level as possible until interrupted by one of two circumstances: (1) if a link exists to the pivot element in the next level, then a simple directed cycle has been identified. Its description is read from the current pathway, and the search continues; (2) if one is searching from node k in level m and all possibilities in level $m + 1$ have been exhausted (i.e., one can move no further to the right), then one backtracks (whence the name) to the node in the current pathway at the $(m - 1)$ th level and begins searching the m th level starting with the $(k + 1)$ th node. When further backtracking becomes impossible, all cycles passing through the pivot element have been identified. The pivot element may be eliminated from further consideration, thereby decreasing the dimension of the subsequent search.

For example, in order to apply the backtracking algorithm to the Cone Spring network (Fig. 2), it helps to consider the compartments in the order 2, 3, 4, 5, 1. The following mnemonic array below may help to keep track of the order of operations:

Level	Compartment				
4	2	3	4	5	1
3	2	3	4	5	1
2	2	3	4	5	1
1	2	3	4	5	1
0	2	3	4	5	1

One begins with pivot element 2 in level 0 and searches level 1 from left to right, following the instructions in the text

above. The order of the significant operation in the search and the current pathways they generate are as follows:

Operation	Current pathway
Begin at pivot element	2
Advance to level 1	2–3
Report cycle 1	2–3–2
Advance to level 2	2–3–4
Report cycle 2	2–3–4–2
Advance to level 3	2–3–4–5
Report cycle 3	2–3–4–5–2
Backtrack to level 2	2–3–4
Backtrack to level 1	2–3
Backtrack to level 0	2
Advance to level 1	2–4
Report cycle 4	2–4–2
Advance to level 2	2–4–5
Report cycle 5	2–4–5–2
Backtrack to level 1	2–4
Backtrack to level 0	2
Further backtracking impossible	–

It accidentally happens that all cycles contain node 2, so that searches starting from the remaining four pivot elements uncover no further cycles. If the first pivot element in this example had been component 1, time would have been wasted searching for cycles containing the plants (there are none). It is obvious that the order of the pivot elements can influence the time it takes to complete the full search.

The program just described works fine on systems of small dimension (say <10 or so), but when it was applied to more realistic networks of 20 or more, the algorithm would not terminate, because it became caught up in a combinatoric, exponential search. It drastically increases the run time if the algorithm gets caught up searching from a pivot element, if that node participates in no cycles whatsoever. One seeks, therefore, some way of identifying and eliminating from the search those nodes that do not engage in cycling. It would be of further help if one could at the same time identify those nodes that are more likely than others to participate in cycling. Fortunately, one can address both of these tasks by doing a preliminary depth-first search to count the numbers of cycle arcs incident to each node (Knuth, 1973). (A cycle arc with respect to a given node is any connection from a descendent node to one of that descendant's ancestors.) This initial screening will take at most n^3 steps, and usually much fewer.⁹ Before commencing the backtracking to find simple

cycles, one rearranges the order of the nodes to place first those having the greatest number of cycle arcs. Furthermore, those nodes without any associated cycle arcs can be eliminated from consideration, thereby lowering the dimension of the search. When this backtracking method is applied to the reordered (and abbreviated) nodes in the $nl \times nl$ sub matrix (Fig. 6), it quickly locates the few simple cycles.

Removing an *isolated* cycle from its supporting network is a straightforward task. One simply traces around the cycle in question and locates the link with the smallest magnitude ($D \rightarrow A$ in Fig. 8a). One then subtracts that least magnitude from the magnitudes of each link in the cycle. This effectively partitions the original cycle (e.g., $A \rightarrow B \rightarrow C \rightarrow D \rightarrow A$ in Fig. 8a) into a pathway fragment of unidirectional flows (Fig. 8b) and a cycle of the same length in which each link has the magnitude of the smallest arc in the original cycle (Fig. 8c).

If one now applies the methods for cycle identification that have been described up to this point to the entire $n \times n$ network of trophic flows (i.e., including those involving the non-living members), the user encounters a significant complication. In most whole-system networks the number of simple cycles far outnumbers the number of flows comprising the network. This implies that several cycles must share the same smallest link, and one can therefore group cycles sharing the same smallest link. Each such grouping is called a *nexus*. It is also likely that all the nodes and cycles belonging to given nexus are controlled by their common smallest link. Removing the smallest link by subtracting it, as in Fig. 8, from any particular cycle of the nexus would be quite arbitrary. Rather, some rational method for apportioning the magnitude of the smallest link over *all* the cycles of a nexus seems preferable.

The possibilities for apportioning the least flow are infinite. One could, for example, divide the magnitude of the smallest link by the number of cycles in the nexus and uniformly subtract that amount from each of the member cycles. There is no heuristic reason for pursuing such equipartitioning, however. A more reasonable division might be to apportion the least flow according to the fraction of the aggregate activity of the nexus that is comprised by the respective activity of each simple cycle. Perhaps an even more rational apportionment (and the one implemented) is to use the $[F]$ matrix to calculate the probability that a quantum of medium anywhere in a simple cycle will complete that particular pathway back to its starting point. (That probability is simply the product of all the f_{ij} comprising the cycle pathway). One then attributes the least link to each cycle in proportion to its contribution to the sum of all such probabilities associated with that nexus (William Silvert, personal communication).

Having thus removed all cycles from the starting network, one is left with an acyclic residual of once-through flows (a tree in the language of graph theory). It should be noted that none of the inputs, exports or dissipations in the system have been altered by the cycle extraction process. (Inputs are always paired with identical outputs in the subtraction process.)

⁹ The simplest algorithm in the worst case requires on the order of $[nodes] + [edges]$ calculations (Tarjan, 1972). Because ecological networks are sparse, that number can be considered practically linear.

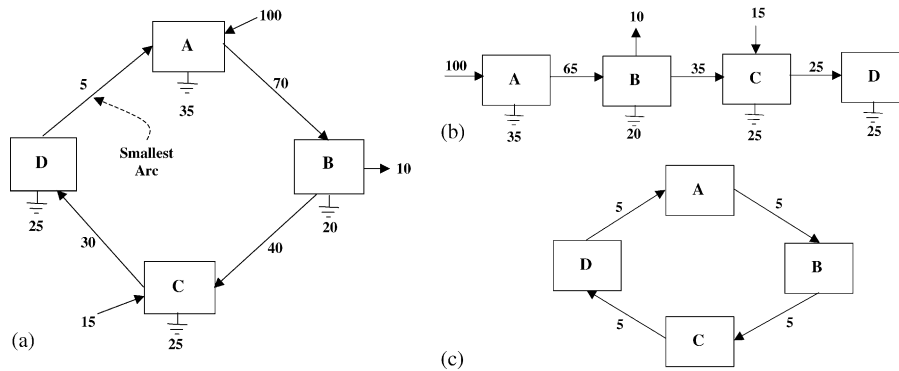


Fig. 8. (a) Hypothetical network consisting of one simple cycle. (b) Single-pass flows through the network. (c) Cycling within the network. Note that (a) = (b) + (c).

The aggregated network of cycled medium can be calculated either by accumulating the cycles that have been removed, or by taking the difference between the starting network and the residual once all have cycles have been stripped. Fig. 9a shows the residual, acyclic network from the Cone Spring example (Fig. 2) and Fig. 9b, the complementary web of purely cyclical flows.

It is true that this decomposition of a graph into its cyclic and acyclic counterparts is purely an abstract exercise. It seems, however, not to be wanting in practical value. Baird and Ulanowicz (1989), for example, extracted the cycles from a 36-compartment network of trophic exchanges in the mesohaline Chesapeake Bay ecosystem to discover that the cycled

flow separated out into two non-overlapping graphs. One grouping contained only pelagic species (those that move passively with water currents), while the other was comprised of benthic (bottom-dwelling) compartments and nektonic (actively moving through the water) species. Thus were two separate domains of control in estuaries identified, and such bifurcation seems to characterize many other estuaries (Allesina et al., in review). What was also quite revealing was that none of the filter-feeding compartments, either on the bottom or among the fishes, participated in any recycling. Rather their apparent “function” is to transfer resources from the pelagic realm into the benthic-nektonic domain. (N.b., it is maintained by some ecologists that it makes no sense to speak of the “function” of a species in the context of its particular community.) Similarly of interest, two compartments of pelagic microbes, normally associated with the “microbial loop” in open oceanic waters, participated in no recycling in the Chesapeake system. They serve rather to shunt carbon (which is fixed by excessive plant growth in Chesapeake Bay) out of the ecosystem.

6. Whole-system status

The foregoing methods have focused upon particular system components, or subsystem aggregates, but the questions that are asked in ecosystem analysis and management quite often concern how the system is performing as a whole? In fact, granting agencies increasingly are requiring that projects consider ecological problems in the context of the whole ecosystem (NSF, 1999). To facilitate such whole-system evaluations, several indices of an ad hoc nature have been developed (e.g., the popular and useful index of biotic integrity [IBI] of Karr et al., 1986). It would be preferable, however, if ecosystem assessment could proceed on more solid theoretical foundations. The question naturally arises, therefore, whether some attribute of ecosystem networks might provide an appropriate and useful indicator of system performance.

Of historical note in this regard was the widespread search during the 1960s for some connection between system diversity and stability (Woodwell and Smith, 1969). The whole

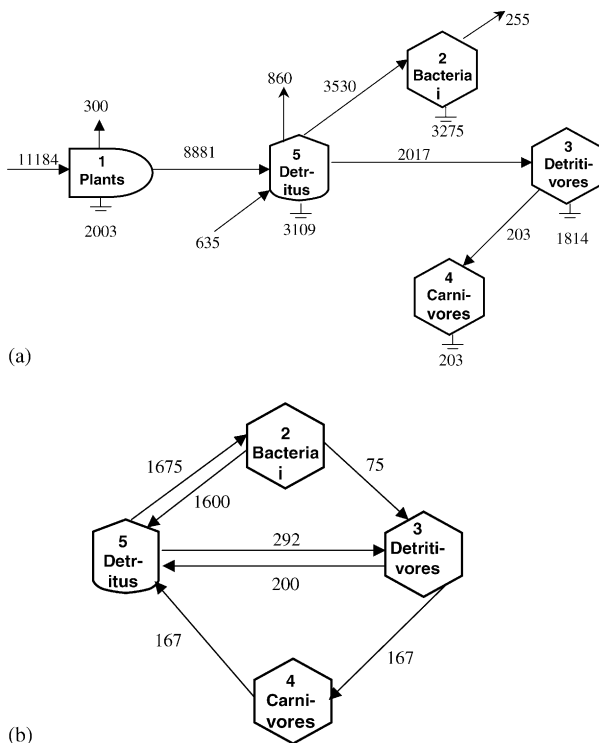


Fig. 9. (a) The residual throughflow in the Cone Spring ecosystem (Fig. 2b). The aggregate cycled flows in the same system.

issue was initiated by MacArthur (1955), who applied Shannon's information measure to the flows in an ecosystem network as,

$$H = -k \sum_{i,j} \left(\frac{T_{ij}}{T_{..}} \right) \log \left(\frac{T_{ij}}{T_{..}} \right),$$

where H is the diversity of flows in the network, k is a scalar constant, and $T_{..}$ signifies the sum of T_{ij} over all combinations i and j .

Shortly thereafter, the consensus arose that the same index could be applied instead to the more accessible stocks of biomass as,

$$D = -k \sum_i \left(\frac{B_i}{B_{..}} \right) \log \left(\frac{B_i}{B_{..}} \right),$$

where D is the information-theoretic biodiversity, B_i is the biomass of compartment i , and $B_{..}$ is the sum of all the B_i .

The subsequent discussions on diversity and stability centered around this biomass diversity. Unfortunately, the enterprise met a rather ignominious end when May (1972) demonstrated that a higher biodiversity in linear dynamical systems was more likely to result in instability than in the reverse. Thereafter, ecologists by the droves quickly abandoned the notion that information theory might be useful in ecosystems analysis — a prejudice which remains prevalent among ecologists today. As a result, little interest was shown when Rutledge et al. (1976) returned to MacArthur's original intuition to focus upon the diversity of flows, and applied a Bayesian emendation of Shannon's information measure to MacArthur's index of flow diversity so as to articulate the latter attribute more precisely.

Effectively, what Rutledge et al. did was to employ the more recent notion of conditional probability to decompose MacArthur's index into two complementary terms. If $(T_{ij}/T_{..})$ is the unconditional probability that a flow from i to j occurs, then $(T_{ij}/T_{j.})$ is the conditional probability that the quantum of flow proceeds to compartment j , given that it had issued from component i . That is, H can be decomposed as

$$H = \text{AMI} + H_c \quad (3)$$

where

$$\text{AMI} = k \sum_{i,j} \left(\frac{T_{ij}}{T_{..}} \right) \log \left(\frac{T_{ij} T_{..}}{T_{i.} T_{j.}} \right),$$

and

$$H_c = -k \sum_{i,j} \left(\frac{T_{ij}}{T_{..}} \right) \log \left(\frac{T_{ij}^2}{T_{i.} T_{j.}} \right).$$

AMI is called the *average mutual information* inherent in the flow structure, and H_c is the residual (conditional) diversity/freedom (inappropriately called the *conditional entropy* in information theory). In other words, the overall complexity of the flow structure, as measured by MacArthur's index, can be resolved into a component that gauges how orderly

and coherently the flows are connected and a residual that measures the disorder and freedom that remains. Rutledge et al. focused upon H_c as a more appropriate measure of ecosystem maturity (in the sense of Odum (1969)) than MacArthur's ambiguous index.

Ulanowicz (1980) suggested that Rutledge's focus upon H_c was perhaps misplaced. In particular, he was impressed by Atlan's (1974) demonstration that the AMI is a useful index of the organization inherent in a system. Ulanowicz proposed that AMI, instead of H_c , is more indicative of the developmental status of an ecosystem, because AMI measures the average amount of constraint exerted upon an arbitrary quantum of currency in passing from any one compartment to the next (Ulanowicz, 1997; Latham and Scully, 2002). AMI, however, has no physical dimensions. That is, given a particular value for the AMI, an investigator cannot tell whether the index pertains to a microbial community in a Petri dish or to an ecological assemblage on the Serengeti Plain. Tribus and McIrvine (1971) had commented on the scalar ambiguity of information indices and suggested that the constant, k , (usually set equal to one) be used to impart physical dimensions to the information index. Accordingly, Ulanowicz set $k = T_{..}$ (the total system throughput) and called the result the system network ascendancy,¹⁰ A , where

$$A = \sum_{i,j} T_{ij} \log \left(\frac{T_{ij} T_{..}}{T_{i.} T_{j.}} \right).$$

Ascendancy thus combines the total activity, or power of the system ($T_{..}$), with the organization by which the component processes are linked (AMI) (Latham and Scully, 2002). It gauges how well the system is performing at processing the given medium. Initially, it had been thought that an ecosystem would develop so as to maximize its ascendancy (Ulanowicz, 1980), but the variational nature of such a statement eventually was viewed as being overly mechanical and deterministic (Mueller and Leupelt, 1998). One now speaks simply of a propensity for ecosystems to increase in ascendancy (Ulanowicz, 1997).

One may likewise scale H_c by $T_{..}$ to yield what is called the system "overhead", Φ (Ulanowicz and Norden, 1990), as

$$\Phi = - \sum_{i,j} T_{ij} \log \left(\frac{T_{ij}^2}{T_{i.} T_{j.}} \right),$$

and H itself can also be scaled to produce what is termed the system's "development capacity", C ,

$$C = - \sum_{i,j} T_{ij} \log \left(\frac{T_{ij}}{T_{..}} \right). \quad (4)$$

Accordingly, relationship (3), when scaled becomes

$$C = A + \Phi. \quad (5)$$

¹⁰ The alternative spelling "ascendancy" is used here to distinguish the quantitative measure from the conventional meaning of "ascendancy".

The decomposition (5) implies that increasing ascendancy usually arises at the expense of the complementary overhead, Φ . One then naturally asks, what are the limits to how much increasing ascendancy can displace system overhead? Systems with a high preponderance of ascendancy over overhead appear to be rigidly linked and almost mechanical. Holling (1986) calls such systems “brittle” and vulnerable to collapse. This makes it important to remember that the freedom to adapt to novel perturbations is contained in the system overhead, Φ . It then becomes obvious that a “healthy” ecosystem, or one with biotic “integrity”, requires adequate amounts of *both* attributes, ascendancy and overhead (Ulanowicz, 2000a).

It happens that the balancing act between two mutually exclusive properties such as A and Φ , cannot properly be described in algorithmic fashion. The ensuing agonistic tradeoff is more reminiscent of a dialectic, i.e., a competition between two mutually exclusive properties at one level, that is resolved via their mutual necessity at the next higher level. No one has yet succeeded at programming a dialectic, so it is unlikely that the corresponding ecosystem dynamic will yield readily to mechanical simulation.

Difficulties with simulation aside, knowing the relative values of A and Φ can nonetheless indicate to the investigator the status of an ecosystem. For example, one informative property of flow networks is their degree of connectivity, sometimes expressed as the link-density, or how many links on average flow into or out of a typical node. Ulanowicz and Wolff (1991) demonstrated that taking the base of the logarithms to the power ($H_c/2$), yields a convenient measure of the effective link-density. In particular, their measure converges to the appropriate intuitive value as the network in question approaches being uniformly connected and equiponderant (i.e., a binary or topological network). Sensing significant utility behind this convergence, Bersier et al. (2002) used the calculation as a bridge between quantitative network analysis and the more popular “food web analysis”, which treats only topological networks. For each and every index that had been defined by investigators in food web analysis, Bersier was able to construct an information-theoretic homolog that converges more rapidly on the property in question than does the original, ad hoc measure.

Zorach and Ulanowicz (2003) were able to show that the connection between H_c and link-density is even deeper. They demonstrated that raising the base of the logarithm to the power ($H_c/2$) yields precisely the weighted geometric mean of the link-density that one would calculate using conventional algebra. Arguing from dimensional considerations, they then inferred how raising the logarithmic base to the power AMI should correspondingly estimate the number of trophic “roles” in the network. (That is, it corresponds roughly to the effective number of trophic levels in the network, or [looking backwards] to the “trophic depth” of the network.)

Zorach and Ulanowicz also remarked how, on a graph of the number of roles versus the link-density, all of their collection of 44 estimated ecosystem networks plotted inside

the confines of a small rectangular domain that Ulanowicz (2002) had earlier suggested demarcated a “window of vitality”. In particular, the link densities all ranged between 1.0 and about 3.0. The lower limit is obviously set by the requirement that the network remain fully connected. (A value below 1.0 would indicate that the graph consists of two or more disconnected subgraphs.) The upper limit of ca. 3 had been noted earlier by several investigators (Wagensberg et al., 1990; Pimm, 1982; Kauffman, 1991), and Ulanowicz had formulated an information-theoretic homolog of the May–Wigner stability criterion (May, 1972) which indicated that stability should wane rapidly beyond a natural limit of $e^{(e/3)}$ (≈ 3.01) links per node. In the other dimension, the number of roles varies between 2.0 and about 4.5. The lower limit expresses the requirement that all ecosystems exhibit a minimum of two functions—production and decomposition (Fiscus, 2002). As for the upper limit, it is widely accepted that virtually all ecosystems are characterized by fewer than five trophic links (Pimm and Lawton, 1977). Igor Matutinovics (personal communication) indicates the same limit seems to hold for economic systems. A formal identification of the actual boundary and the reasons behind that limit remain, however, unknown.

Often, in order to ascertain the full status of a system, it helps to decompose A and Φ further. For example, it is sometimes useful to focus upon that part of the ascendancy that is generated by the internal workings of the system. To do this, one limits consideration to the “internal ascendancy”, A_I , which is generated solely by the internal exchanges between the n system components,

$$A_I = \sum_{i,j=1}^n T_{ij} \log \left(\frac{T_{ij} T_{..}}{T_{i.} T_{.j}} \right).$$

The resulting fraction of the development capacity that is comprised by A_I then becomes a putative index of the internal development of the given ecosystem. This ratio has been used to compare ecosystems that were similarly parsed (Wulff and Ulanowicz, 1989; Baird et al., 1991, 2004; Christian et al., 2003).

It is likewise helpful to regard how the overhead is parceled out among the different types of flows. There being four basic categories of flow in ecosystems (internal exchanges, exogenous inputs, useful outputs, and dissipations, as in Section 2 above), the overhead thereby separates into four corresponding components (Ulanowicz and Norden, 1990):

$$R = - \sum_{i=1}^n \sum_{j=1}^n T_{ij} \log \left(\frac{T_{ij}^2}{T_{i.} T_{.j}} \right),$$

$$\Phi_I = - \sum_{j=1}^n T_{0j} \log \left(\frac{T_{0j}^2}{T_{0.} T_{.j}} \right),$$

$$\Phi_E = - \sum_{i=1}^n T_{i,n+1} \log \left(\frac{T_{i,n+1}^2}{T_{i..} T_{..,n+1}} \right),$$

$$\Phi_D = - \sum_{i=1}^n T_{i,n+2} \log \left(\frac{T_{i,n+2}^2}{T_{i..} T_{..,n+2}} \right),$$

where R is commonly called the flow “redundancy”, because it is strongly tied to the effective multiplicity of parallel flows by which medium passes between any two arbitrary system components. Φ_I is the overhead due to imports, while Φ_E and Φ_D are the overheads from exports and dissipations, respectively. Often, perturbations will drive up the value of R/C (Ulanowicz, 1984), indicating that an impacted system tends to become more resistant to further perturbations (the LeChâtelier–Braun Principle as applied to ecology). Heymans (2003) has used the internal overhead as a measure of the potential for adaptation and resilience.

The reader may have noticed that the indices defined thus far depend entirely upon the flow structure and take no account of the magnitudes of any stocks. Although it very well may be that the dynamics of the system depend far more on the configurations of flows within it than has been heretofore recognized, there still likely remain situations when the size of certain stocks will affect system dynamics. To consider the possibility of such controls, it becomes necessary to introduce the stock sizes, B_i , into the information-theoretic calculus. One begins by recognizing that the apriori probability that a quantum of medium in transit had originated from compartment i is $(B_i/B_{..})$. The corresponding probability that a particle of flowing medium will enter j is $(B_j/B_{..})$. Hence, the apriori joint probability that any quantum of flowing medium is en route from i to j , becomes $(B_i B_j / B_{..}^2)$. The observed, or aposteriori frequency of such flow from i to j is simply $(T_{ij}/T_{..})$. According to Kullback (1959), the average difference between an the apriori and aposteriori distributions takes the form

$$I_B = \sum_{i,j} \left(\frac{T_{ij}}{T_{..}} \right) \log \left(\frac{T_{ij} B_{..}^2}{T_{..} B_i B_j} \right),$$

where I_B is called the Kullback–Leibler information.

Scaling I_B by $T_{..}$, yields a biomass-inclusive form of the ascendancy, A_B (Ulanowicz and Abarca-Arenas, 1997) as,

$$A_B = \sum_{i,j} T_{ij} \log \left(\frac{T_{ij} B_{..}^2}{T_{..} B_i B_j} \right).$$

If A_B is considered a surrogate for the status of a system, then changes in it due to small increments in the biomass of a particular component should reveal the sensitivity of the whole system to the stock of that particular taxon. Taking the partial derivative of A_B with respect to B_p ,

$$\frac{\partial A_B}{\partial B_p} = 2 \left(\frac{T_{..}}{B_{..}} - \frac{1}{2} \frac{T_{..p} + T_{p..}}{B_p} \right)$$

reveals that such sensitivity is twice the amount by which the average turnover rate of medium in compartment p is exceeded by the overall turnover rate $(T_{..}/B_{..})$. Hence, the system is putatively more sensitive to those components having slower turnover rates of the medium in question. Furthermore, one could ask which prey resource, r , is most important as a source of medium to predator p . That source should be made apparent by taking the partial derivative of A_B with respect to each prey, r , of p (i.e., to each T_{rp}). Because A_B is a first-order homogeneous function in the T_{ij} (Courant, 1936), one may immediately invoke Euler’s relation to write the following simple result:

$$\frac{\partial A_B}{\partial T_{rp}} = \log \left(\frac{T_{rp} B_{..}^2}{T_{..} B_r B_p} \right).$$

Examination of this formula leads to the intuitively satisfying conclusion that the controlling resource is the one which is being depleted of medium most quickly relative to its available stock (i.e., the one with the highest ratio $[T_{rp}/B_r]$).

7. Higher dimensional considerations

The sensitivity analyses just outlined would be far more useful, if one could consider the simultaneous flow of several media at once. Furthermore, the entire enterprise of network analysis as described thus far pertains only to a single snapshot of the system. Over time, actual dynamics could be followed via a “motion picture” or a time-series of network snapshots. One might also want to consider a collection of networks each of which pertains to a small segment of the physical space occupied by the ecosystem. Although generalizations of the analyses detailed in Sections 3–5 to heterogeneities in time, space and media become problematical, similar extensions of the information-theoretic indices present no such conceptual obstacles and hinge more on the availability of sufficient data.

To extend the network approach to temporal changes, one denotes T_{ijk} to be the flow of medium from taxon i to taxon j during time interval k . The temporal capacity, C_T , then follows from straightforward generalization of (4) to become

$$C_T = - \sum_{i,j,k} T_{ijk} \log \left(\frac{T_{ijk}}{T_{...}} \right).$$

There remains some ambiguity, however, as to what constitutes the ascendancy in the dynamical case. To clarify how C_T can be resolved into components, it is helpful to employ Venn diagrams (Blachman, 1961). In two dimensions the situation was unambiguous.

The circle on the left of Fig. 10 represents the diversity of the source flows (prey), while the one on the right, that of the sinks (predators). Their intersection, or overlap represents the AMI. The non-overlapping remaining areas comprise the conditional uncertainty, H_C .

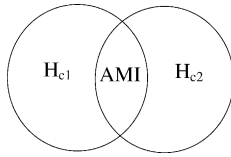


Fig. 10. Two-dimensional Venn diagram showing the AMI as the intersection of the uncertainties of sources and sinks. The conditional uncertainty, $H_c = H_{c1} + H_{c2}$.

The situation in three dimensions (Fig. 11) becomes more complex.

One sees Fig. 10 repeated in the two adjacent upper circles. The additional (lower) circle represents variation over time. The juxtaposition of the three domains results in not one, but four distinct regions of overlap — a central area where all three domains overlap (AMI_3) and three adjacent areas where only one of the three possible pairs overlap in their turn (λ_1 , λ_2 , λ_3). In information theory, the innermost circular triangle (AMI_3) represents the 3-D AMI. In terms of the fluxes T_{ijk} , it can be calculated (Abramson, 1963; Pahl-Wostl, 1995) as:

$$AMI_3 = \sum_{i,j,k} \left(\frac{T_{ijk}}{T_{...}} \right) \log \left(\frac{T_{ijk}^2 T_{...}}{T_{i..} T_{.j.} T_{.k.}} \right).$$

AMI_3 , then, represents the simultaneous coherence among sources and sinks over time. Unlike with the 2-D AMI, there is no guarantee that AMI_3 remain non-negative. AMI_3 , however, does not encompass all the coherency present in the system. In particular, the “leaflets” (λ_1 , λ_2 , λ_3), that adjoin AMI_3 represent bilateral coherencies that do not correlate with the third remaining variable. To be more inclusive of all coherencies in the system, Pahl-Wostl (1995) suggested that the appropriate measure be taken as the overlap of all bilateral AMIs (i.e., the trefoil-like region $AMI_3 + \lambda_1 + \lambda_2 + \lambda_3$). Calling this conjunction AMI_T , one calculates its magnitude in terms of the T_{ijk} as:

$$AMI_T = \sum_{i,j,k} \left(\frac{T_{ijk}}{T_{...}} \right) \log \left(\frac{T_{ijk}^2 T_{...}}{T_{ij.} T_{i.k} T_{.jk}} \right).$$

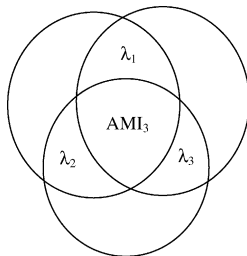


Fig. 11. The intersection of the three domains of sources sinks and time. The central area of overlap (AMI_3) represents the 3-D average mutual information, but the three areas of bilateral overlap (λ_1 , λ_2 , λ_3) also contribute to the overall system organization.

It is guaranteed that $AMI_T \geq 0$. Scaling AMI_T by $T_{...}$ yields the temporal ascendancy A_T ,

$$A_T = \sum_{i,j,k} T_{ijk} \log \left(\frac{T_{ijk}^2 T_{...}}{T_{ij.} T_{i.k} T_{.jk}} \right).$$

The generalization of capacity and ascendancy into 4 or more dimensions proceeds by induction. Without loss of generality, the 5-D transfer T_{ijklm} of flow from taxon i to taxon j during time interval k at spatial location l involving medium m is about as complicated as most ecological considerations become.¹¹ The overall 5-D ascendancy, A_O , then becomes

$$A_O = \sum_{i,j,k,l,m} T_{ijklm} \log \left(\frac{T_{ijklm}^4 T_{....}}{T_{.jklm} T_{i.klm} T_{ij.lm} T_{ijk.m} T_{ijkl.}} \right).$$

Of course, the number of distinct intersections among the dimensions will increase binomially with the problem dimension. The ascendancy in all cases, however, remains homogeneous in the flows. Thus, for example, the 5-D sensitivities become

$$\frac{\partial A_O}{\partial T_{pqrst}} = \log \left(\frac{T_{pqrst}^4 T_{....}}{T_{.qrst} T_{p.rst} T_{pq.st} T_{pqr.t} T_{pqrs.}} \right).$$

The biomass-inclusive overall ascendancy, A_{OB} , takes the form

$$A_{OB} = \sum_{i,j,k,l,m} T_{ijklm} \log \left(\frac{T_{ijklm} B_{iklm}^2}{B_{iklm} B_{jklm} T_{....}} \right), \quad (6)$$

where B_{iklm} is the biomass of taxon i in terms of medium m at time k in spatial segment l . The corresponding sensitivity coefficients become:

$$\frac{\partial A_{OB}}{\partial B_{prst}} = 2 \left\{ \frac{T_{....}}{B_{....}} - \frac{1}{2} \left[\frac{T_{p.rst}}{B_{prst}} + \frac{T_{prst.}}{B_{prst.}} \right] \right\}, \quad (7)$$

and

$$\frac{\partial A_{OB}}{\partial T_{pqrst}} = \log \left(\frac{T_{pqrst} B_{....}^2}{B_{prst} B_{qrst} T_{....}} \right). \quad (8)$$

For those cases where data are available on multiple media, Ulanowicz and Baird (1999) have shown analytically that the largest biomass sensitivity coefficients (7) correspond to the medium that is limiting in the sense of von Liebig (1854). Von Liebig's Law, however, gives no way of identifying the most limiting source of that nutrient. That source is given by the largest value of (8).

Ulanowicz and Baird (1999) applied (7) and (8) to parallel networks of several media to provide a new and systematic way of elucidating nutrient dynamics in ecosystems. They assembled sufficient data to estimate the trophic exchanges of

¹¹ Space is considered as only a single dimension, because any 2-D or 3-D partitioning of a finite domain of physical space always can be enumerated by a single index [in the same fashion that multi-dimensional arrays are enumerated by computer code].

carbon, nitrogen and phosphorus (C, N, and P) in the ecosystem of the mesohaline reach of Chesapeake Bay over the four seasons of a typical year (no spatial variation). They used the 4-D version of (7) to identify the limiting medium for each taxon and found that N was limiting primary production during the summer season (as has been long-known). During that same season, however, P was controlling bacterial production and secondary production by nekton (mobile fishes). Applying (8) to their networks, the investigators were able to diagram (graphical) trees that depict the hierarchy of nutrient controls during each season.

Ulanowicz (2000b) also calculated A_{OB} (6) across a simple 10×10 spatial grid of hypothetical migrating animal populations to demonstrate how information theoretic measures could be used in landscape ecology. He used cellular automata to simulate several migration patterns and then proceeded to quantify the perturbation (or augmentation) that certain constraints exercised on the organization of the patterns. Kikawada (1998) extended Ulanowicz's spatial analysis to include simultaneous trophic interactions, and discussed several problems encountered in doing so. More generally, network sensitivity analysis can be regarded as a new way to examine the network dynamics of various chemical elements being distributed across the landscape (Krivov and Ulanowicz, 2003). The method can even be applied to patterns of physical fluid flows to document the degree of organization inherent in the fluid dynamical field (Ulanowicz and Zickel, in press).

8. Concluding remarks

Some ecologists persist in believing that the creation of trophic flow networks is otiose. They perceive the resulting miasma of interconnections as far too complicated to afford the investigator any penetrating insights into the workings of an ecosystem. Visual impressions can be unduly discouraging; however, because linear algebra can be invoked to analyze network patterns in systematic fashion and to abstract from them meaningful relationships that otherwise might elude discovery. Nor should the enormous and growing power of computational machinery to search for underlying patterns (such as the cycling of materials) among ecosystem networks be discounted. Finally, the power of information theory as an alternative to conventional statistics for describing systems structure and behavior and as a bridge to thermodynamics is burgeoning and gives considerable promise of a wholly different way to apprehend and analyze living systems.

9. Software availability

The implementation of the methods described here has been a gradual and continuing process. During the 1980s and early 1990s a FORTRAN package for DOS named NETWRK4 was written to run the I/O, trophic,

and cycle analysis methods and to calculate community information measures. It is still available for downloading at <http://www.cbl.umces.edu/~ulan/ntwk/network.html>. A translation of NETWRK4 for the Windows platform has been written by the NOAA Great Lakes Environmental Research Lab and is available to the public as EcoNetwrk <http://www.glerl.noaa.gov/EcoNetwrk/>. EcoNetwrk will also include a number of Graphical User Interfaces for input and output. Stefano Allesina at the University of Parma has created a variation of NETWRK, called WAND, that is compatible with Windows ExcelTM <http://www.dsa.unipr.it/~alle/ena/>. (Allesina and Bondavalli, 2004; Fath, 2004). An optimization routine to construct balanced networks from incomplete data according to the method of Polovina (1985) is available as ECOPATH from <http://www.ecopath.org>. ECOPATH also implements the trophic impact analysis and calculates the network ascendancy and related information measures. Several experimental algorithms to implement the multi-dimensional information indices are currently in various stages of development by the author and by Stefano Allesina.

Acknowledgements

The author is most grateful to Dr. Andrzej Konopka for encouraging him to undertake this synopsis of ENA methods. He would also like to thank Stefano Allesina, Daniel Baird, Antonio Bodini, Robert Christian, Brian Fath, Johanna Heymans and Luke Latham for reviewing the initial draft of the manuscript and making valuable suggestions for its improvement. The author did not incorporate all suggestions, so any defects or omissions are entirely his responsibility. The author was supported in this endeavor by the National Science Foundation's Program on Biocomplexity (Contract No. DEB-9981328).

References

- Abramson, N., 1963. Information Theory and Coding. McGraw-Hill, New York, 201 pp.
- Allesina, S., Bondavalli, C., 2003. Steady state of ecosystem flow networks: a comparison between balancing procedures. *Ecol. Model.* 165, 221–229.
- Allesina, S., Bondavalli, C., 2004. WAND: an ecological network analysis user-friendly tool. *Environ. Model. Software* 19 (4), 337–340.
- Allesina, S., Bodini, A., Bondavalli, C., in review. Ecological subsystems via graph theory: the role of strongly connected components. *Oikos*.
- Allesina, S., Ulanowicz, R.E., 2004. Cycling in ecological networks: Finn's index revisited. *Computat. Biol. Chem.* 28, 227–233.
- Atlan, H., 1974. On a formal definition of organization. *J. theor. Biol.* 45, 295–304.
- Augustinovic, M., 1970. Methods of international and intertemporal comparison of structure. In: Carter, A.P., Brody, A. (Eds.), *Contributions to Input–Output Analysis*, vol. I. North-Holland, Amsterdam, pp. 249–269.
- Baird, D., Ulanowicz, R.E., 1989. The seasonal dynamics of the Chesapeake Bay ecosystem. *Ecol. Monogr.* 59, 329–364.

- Baird, D., McGlade, J.M., Ulanowicz, R.E., 1991. The comparative ecology of six marine ecosystems. *Philos. Trans. R. Soc. Lond. B* 333, 15–29.
- Baird, D., Christian, R.R., Peterson, C.H., Johnson, G.A., 2004. Consequences of hypoxia on Estuarine ecosystem function: energy diversion from consumers to microbes. *Ecol. Applic.* 14 (3), 805–822.
- Bersier, L.-F., Banašek-Richter, C., Cattin, M.-F., 2002. Quantitative descriptors of food web matrices. *Ecology* 83, 2394–2407.
- Blachman, N.M., 1961. A generalization of mutual information. *Proc. IRE* 49 (8), 1331–1332.
- Bondavalli, C., Ulanowicz, R.E., 1999. Unexpected effects of predators upon their prey: the case of the American Alligator. *Ecosystems* 2, 49–63.
- Christensen, V., Pauly, D., 1992. ECOPATH II—a software for balancing steady-state models and calculating network characteristics. *Ecol. Model.* 61, 169–185.
- Christian, R.R., Luczkovich, J.J., 1999. Organizing and understanding a winter's seagrass foodweb network through effective trophic levels. *Ecol. Model.* 117, 99–124.
- Christian, R.R., Dame, J.K., Johnson, G., Peterson, C.H., Baird, D., 2003. Monitoring and modeling of the Neuse River Estuary. Phase 2. Functional assessment of environmental phenomena through network analysis. UNC-WRRI 2003-343-E, Department of Biology, East Carolina University, Greenville, North Carolina, 111 pp.
- Cohen, J.E., Briand, F., Newman, C.M. (Eds.), 1990. *Community Food Webs: Data and Theory*. Biomathematics 20. Springer-Verlag, Heidelberg, 308 pp.
- Cousins, S., 1985. Ecologists build pyramids again. *New Scientist* 107 (1463), 50–54.
- Depew, D.J., Weber, B.H., 1994. *Darwinism Evolving: Systems Dynamics and the Geneology of Natural Selection*. MIT Press, Cambridge, MA, 588 pp.
- Courant, R., 1936. P108 in *Differential and Integral Calculus*, vol. 2. Interscience Publishers, New York, 681 pp.
- Fath, B.D., 2004. Network analysis in perspective: comments on WAND: an ecological network analysis user-friendly tool. *Environ. Model. Software* 19, 341–343.
- Fath, B.D., Patten, B.C., 1998. Network synergism: emergence of positive relations in ecological systems. *Ecol. Model.* 107, 127–143.
- Fath, B.D., Patten, B.C., 1999. Review of the foundations of network environ analysis. *Ecosystems* 2, 167–179.
- Fath, B.D., Borrett, S.R., in review. A Matlab® function for network environ analysis. *Environ. Model. Software*.
- Finn, J.T., 1976. Measures of ecosystem structure and function derived from analysis of flows. *J. theor. Biol.* 56, 363–380.
- Fiscus, D.A., 2002. The ecosystemic life hypothesis. II. Four connected concepts. *ESA Bull.* 83, 94–96.
- Han, B.-P., 1997. On several measures concerning flow variables in ecosystems. *Ecol. Model.* 104, 289–302.
- Hannon, B.M., 1973. The structure of ecosystems. *J. theor. Biol.* 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. *Theor. Popul. Biol.* 40, 78–104.
- Heymans, J.J., 2003. Comparing the Newfoundland–Southern Labrador marine ecosystem models using information theory. In: Heymans, J.J. (Ed.), *Ecosystem models of Newfoundland and Southeastern Labrador (2J3KLNO): Additional Information and Analyses for “Back to the Future”*, vol. 11 (5). Fisheries Centre Research Reports, pp. 62–71.
- Higashi, M., Burns, T.D. (Eds.), 1991. *Theoretical Studies of Ecosystems: The Network Perspective*. Cambridge University Press, Cambridge, 364 pp.
- Higashi, M., Patten, B.C., Burns, T.D., 1991. Network trophic dynamics: an emerging paradigm in ecosystems ecology. In: Higashi, M., Burns, T.D. (Eds.), *Theoretical Studies of Ecosystems: The Network Perspective*. Cambridge University Press, Cambridge, pp. 117–154, 364.
- Hillebrand, H., Shurin, J.B., in press. Biodiversity and freshwater food webs. In: Belgrano, A., Scharler, U., Dunne, J., Ulanowicz, R.E. (Eds.), *Aquatic Food Webs: An Ecosystem Approach*. Oxford University Press, Oxford, Chapter 14.
- Holling, C.S., 1986. The resilience of terrestrial ecosystems: local surprise and global change. In: Clark, W.C., Munn, R.E. (Eds.), *Sustainable Development of the Biosphere*. Cambridge University Press, Cambridge, UK, pp. 292–317.
- Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., Schlosser, I.J., 1986. Assessment of Biological Integrity in Running Water: A Method and Its Rationale. Illinois Natural History Survey, Publication 5, Champaign, IL.
- Kauffman, S.A., 1991. Antichaos and adaptation. *Sci. Am.* 265, 78–84.
- Kavanagh, P., Newlands, N., Christensen, V., Pauly, D., 2004. Automated parameter optimization for Ecopath ecosystem models. *Ecol. Model.* 172, 141–149.
- Kikawada, H., 1998. Applying Network Analysis to Simulated Ecological Landscape Dynamics. Master's Thesis, University of Maryland, College Park, MD, 181 pp.
- Knuth, D.E., 1973. *Fundamental Algorithms*, vol. 1. Addison-Wesley, Reading, MA, 228 pp.
- Krivov, S., Ulanowicz, R.E., 2003. Quantitative measures of organization for multi-agent systems. *BioSystems* 69, 39–54.
- Kullback, S., 1959. *Information Theory and Statistics*. Peter Smith, Gloucester, MA, 399 pp.
- Latham, L.G., Scully, E.P., 2002. Quantifying constraint to assess development in ecological networks. *Ecol. Model.* 154, 25–44.
- Leontief, W., 1951. *The Structure of the American Economy, 1919–1939*, second ed. Oxford University Press, New York, 257 pp.
- Levine, S., 1980. Several measures of trophic structure applicable to complex food webs. *J. theor. Biol.* 83, 195–207.
- Lorenz, E.N., 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* 20, 130–141.
- MacArthur, R., 1955. Fluctuations of animal populations, and a measure of community stability. *Ecology* 36, 533–536.
- Mateti, P., Deo, N., 1976. On algorithms for enumerating all the circuits of a graph. *SIAM J. Comput.* 5, 90–99.
- Matis, J.H., Patten, B.C., White, G.C., 1979. *Compartmental Analysis of Ecosystem Models*. International Cooperative Publishing House, Fairland, MD, p. 368.
- Matis, J.H., Patten, B.C., 1981. Environ analysis of linear compartmental systems: the static, time invariant case. *Bull. Int. Stat. Inst.* 48, 527–565.
- May, R.M., 1972. Will a large complex system be stable? *Nature* 238, 413–414.
- Mueller, F., Leupelt, M. (Eds.), 1998. *Eco Targets, Goal Functions and Orientors*. Springer-Verlag, Berlin, 619 pp.
- Naess, A., 1988. Deep ecology and ultimate premises. *Ecologist* 18, 128–131.
- 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.
- Odum, H.T., 1971. *Environment Power and Society*. Wiley, New York, 331 pp.
- Pahl-Wostl, C., 1995. *The Dynamic Nature of Ecosystems: Chaos and Order Entwined*. Wiley, New York, 267 pp.
- Patten, B.C., 1978. Systems approach to the concept of environment. *Ohio J. Sci.* 78, 206–222.
- Patten, B.C., 1982a. On the quantitative dominance of indirect effects in ecosystems. In: Unpublished paper presented at the Third International Conference on State-of-the-Art in Ecological Modeling, Colorado State University, Fort Collins, Colorado, 24–28 May.
- Patten, B.C., 1982b. Environs: relativistic elementary particles for ecology. *Am. Nat.* 119, 179–219.
- Patten, B.C., 1985. Energy cycling in the ecosystem. *Ecol. Model.* 28, 1–71.

- Patten, B.C., Auble, G.T., 1981. System theory of the ecological niche. *Am. Nat.* 117, 893–922.
- Patten, B.C., Bosserman, R.W., Finn, J.T., Cale, W.G., 1976. Propagation of cause in ecosystems. In: Patten, B.C. (Ed.), *Systems Analysis and Simulation in Ecology*, vol. 4. Academic Press, New York, pp. 457–579.
- Pimm, S.L., 1982. *Food Webs*. Chapman and Hall, London, 219 pp.
- Pimm, S.L., Lawton, J.H., 1977. Number of trophic levels in ecological communities. *Nature* 268, 329–331.
- Platt, T.C., Mann, K.H., Ulanowicz, R.E. (Eds.), 1981. *Mathematical Models in Biological Oceanography*. UNESCO Press, Paris, 157 pp.
- Polis, G., Winemiller, K., 1995. *Food Webs: Integration of Patterns and Dynamics*. Chapman and Hall, New York, 472 pp.
- Polovina, J.J., 1985. An approach to estimating an ecosystem box model. *US Fish Bull.* 83 (3), 457–460.
- Rutledge, R.W., Basorre, B.L., Mulholland, R.J., 1976. Ecological stability: an information theory viewpoint. *J. theor. Biol.* 57, 355–371.
- Sheffer, M., Beets, J., 1994. Ecological models and the pitfalls of causality. *Hydrobiologia* 275–276, 115–124.
- Simon, H.A., Hawkins, D., 1949. Some conditions of macroeconomic stability. *Econometrica* 17, 245–248.
- Szyrmer, J., Ulanowicz, R.E., 1987. Total flows in ecosystems. *Ecol. Model.* 35, 123–136.
- Tarjan, R., 1972. Depth first search and linear graph algorithms. *SIAM J. Comput.* 1, 146–160.
- Tilly, L.J., 1968. The structure and dynamics of Cone Spring. *Ecol. Monogr.* 38, 169–197.
- Tribus, M., McIrvine, E.C., 1971. Energy and information. *Sci. Am.* 225, 179–188.
- Ulanowicz, R.E., 1979. Prediction chaos and ecological perspective. In: Halfon, E. (Ed.), *Theoretical Systems Ecology*. Academic Press, New York, pp. 107–117.
- Ulanowicz, R.E., 1980. An hypothesis on the development of natural communities. *J. theor. Biol.* 85, 223–245.
- Ulanowicz, R.E., 1983. Identifying the structure of cycling in ecosystems. *Math. Biosci.* 65, 219–237.
- Ulanowicz, R.E., 1984. Community measures of marine food networks and their possible applications. In: Fasham, M.J.R. (Ed.), *Flows of Energy and Materials in Marine Ecosystems*. Plenum Press, London, pp. 23–47.
- Ulanowicz, R.E., 1986. *Growth and Development: Ecosystems Phenomenology*. Springer-Verlag, New York, 203 pp.
- Ulanowicz, R.E., 1995. Ecosystem trophic foundations: Lindeman Exonerata. In: Patten, B.C., Jorgensen, S.E. (Eds.), *Complex Ecology: The Part–Whole Relation in Ecosystems*. Prentice Hall, Englewood Cliffs, NJ, pp. 549–560.
- Ulanowicz, R.E., 1997. *Ecology, the Ascendent Perspective*. Columbia University Press, New York, 201 pp.
- Ulanowicz, R.E., 1999. Life after Newton: an ecological metaphysic. *BioSystems* 50, 127–142.
- Ulanowicz, R.E., 2000a. Toward the measurement of ecological integrity. In: Pimentel, D., Westra, L., Noss, R.F. (Eds.), *Ecological Integrity: Integrating Environment, Conservation and Health*. Island Press, Washington, DC, pp. 99–113.
- Ulanowicz, R.E., 2000b. Quantifying constraints upon trophic and migratory transfers in spatially heterogeneous ecosystems. In: Sanderson, J., Harris, L.D. (Eds.), *Series in Landscape Ecology: A Top–Down Approach*. Lewis Publishers, Boca Raton, pp. 113–142.
- Ulanowicz, R.E., 2002. The balance between adaptability and adaptation. *BioSystems* 64, 13–22.
- Ulanowicz, R.E., Abarca-Arenas, L.G., 1997. An informational synthesis of ecosystem structure and function. *Ecol. Model.* 95, 1–10.
- Ulanowicz, R.E., Baird, D., 1999. Nutrient controls on ecosystem dynamics: the Chesapeake mesohaline community. *J. Mar. Syst.* 19, 159–172.
- Ulanowicz, R.E., Bondavalli, C., Egnotovich, M.S., 1997. Section 4.1 of Network Analysis of Trophic Dynamics in South Florida Ecosystems, FY 96: The Cypress Wetland Ecosystem. Ref. No. [UMCES]CBL 97-075, Chesapeake Biological Laboratory, Solomons, MD.
- Ulanowicz, R.E., Kemp, W.M., 1979. Toward canonical trophic aggregations. *Am. Nat.* 114, 871–883.
- Ulanowicz, R.E., Norden, J.S., 1990. Symmetrical overhead in flow networks. *Int. J. Syst. Sci.* 1, 429–437.
- Ulanowicz, R.E., Puccia, C.J., 1990. Mixed trophic impacts in ecosystems. *Coenoses* 5, 7–16.
- Ulanowicz, R.E., Scharler, U.M., in preparation. Least-inference methods for constructing networks of trophic flows.
- Ulanowicz, R.E., Wolff, W.F., 1991. Ecosystem flow networks: loaded dice? *Math. Biosci.* 103, 45–68.
- Ulanowicz, R.E., Zickel, M.J., in press. Using ecology to quantify fluid flows. In: Kleidon, A.M., Lorenz, R.D. (Eds.), *Thermodynamics of Life, Earth, and Beyond: Systems, Feedbacks, Organization and Evolution*. Springer, Berlin, pp. 57–66.
- Vezina, A.F., Platt, T.C., 1988. Food web dynamics in the ocean. I. Best-estimates of flow networks using inverse methods. *Mar. Ecol. Prog. Ser.* 42, 269–287.
- von Liebig, J., 1854. *Chemistry in its Application to Agriculture and Physiology*. Taylor and Walton, London, 401 pp.
- Wagensberg, J., Garcia, A., Sole, R.V., 1990. Connectivity and information transfer in flow networks: two magic numbers in ecology? *Bull. Math. Biol.* 52, 733–740.
- Whipple, S.J., Patten, B.C., 1993. The problem of nontrophic processes in trophic ecology: a network unfolding solution. *J. theor. Biol.* 163, 393–411.
- Woodwell, G.M., Smith, H.H. (Eds.), 1969. *Diversity and Stability in Ecological Systems*, vol. 22. US Brookhaven Symp. Biol., New York, 264 pp.
- Wulff, F., Ulanowicz, R.E., 1989. A comparative anatomy of the Baltic Sea and Chesapeake Bay ecosystems. In: Wulff, F., Field, J.G., Mann, K.H. (Eds.), *Flow Analysis of Marine Ecosystem*. Springer-Verlag, Berlin, pp. 232–256.
- Wulff, F., Field, J.G., Mann, K.H. (Eds.), 1989. *Flow Analysis of Marine Ecosystem*. Springer-Verlag, Berlin, 284 pp.
- Yodzis, P., 1989. Patterns in food webs. *Trends Ecol. Evol.* 4 (2), 49–50.
- Zorach, A.C., Ulanowicz, R.E., 2003. Quantifying the complexity of flow networks: how many roles are there? *Complexity* 8 (3), 68–76.