Homage to Ramon Margalef; or, Why there is such pleasure in studying nature 387
(J. D. Ros & N. Prat, eds.). 1991. *Oecologia aquatica*, 10: 387-391

FLOW-NETWORK ORGANIZATION IN ECOSYSTEMS AND THE MATHEMATICAL THEORY OF INFORMATION

JORGE WAGENSBERG[1], RICARD V. SOLÉ[2] & AMBROSIO GARCÍA[1]

[1] Departament de Física Fonamental. Facultat de Fisica. Universitat de Barcelona. Av. Diagonal, 647. 08028 Barcelona. Spain
[2] Departament de Física i Enginyeria Nuclear. Universitat Politècnica de Catalunya. Pau Gargallo, 5. 08026 Barcelona. Spain

Received: February 1990

ABSTRACT

Ecological flow networks are described by means of the Theory of Graphs and characterised by two quantities of the Mathematical Theory of Information, namely the Internal Information Transfer (a measure of the Specialization) and the Joint Entropy (a measure of Connectivity). Both indexes fall into a narrow interval of values. This tendency of real stable ecosystems is interpreted in the light of two variational adaptative principles.

KEY WORDS: Flow-network, information, ecosystem modelling.

INTRODUCTION

The use of Shannon's entropy,

$$S(p_1, ..., p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad (1)$$

as a measure of ecological diversity was first introduced in the 50s by MacArthur and Margalef and extensively used later in the study of real ecosystems. Applications of this concept are today numerous and varied. Shannon's entropy provides, in physical sciences (WEHRL, 1978), a good approach to the equilibrium states in isolated systems. The final statistical structure in these situations can be obtained by means of the maximization of (1) under some physical constraints. For example, if p_i is the probablility of finding a given particle of a dilute gas (confined in a box) in the interval of energy (E_j, $E_j+\Delta E_j$), we can use the constraints of normalization of $\{p_i\}$ -set,

$$\sum_{i=1}^{n} p_i = 1 \qquad (2)$$

and the energy-conservation condition,

$$\sum_{i=1}^{n} p_i E_i = \overline{E} \qquad (3)$$

(here $\overline{E}$ is the mean amount of energy) to find the equilibrium probability distribution, $\{p_{ieq}\}$. We then have,

$$\delta \left[S + \alpha \left(1 - \sum_{i=1}^{n} p_i\right) + \beta\left(\overline{E} - \sum_{i=1}^{n} p_i E_i\right)\right]=0 \qquad (4)$$

and performing the variation of (4), we

obtain the best known Gibbs canonical distribution

$$P_i = \frac{1}{z} e^{-Ei/kT} \quad (5)$$

where Z is the sum $\sum_{i=1}^{n} e^{-Ei/kT}$, called the partition function.

Entropy-like functions (1) can be defined in several systems where $\{p_i\}$ is known. Then, with appropriate constraints, it is possible to derive the equilibrium distributions in, for example, ecological systems when an extensive quantity is defined. The maximum entropy (MAXENT) principle, due to Jaynes, allows one to make these unbiased estimates of which otherwise only some averages (corresponding to macroscopic measures) are known. The MAXENT principle provides very elegant access to basic relations and concepts of thermodynamics and other fields (JAYNES, 1985) and has been applied in recent years in the study of biomass distribution and ecological adaptation. The use of correlation i.e. $F_{ij} = \langle q_i q_j \rangle$ as constraints allows us to extend this approach to systems far from equilibrium (HAKEN, 1985, 1987) as for example in laser physics, in a quite straightforward form.

The MAXENT algorithm has been used (LURIE & WAGENSBERG, 1983) in the study of the statistical structure of biomass diversity in fish populations. Defining probabilities as $p_i = P\,[m_i, m_i + \Delta m]$, and using (2) and constant value of mean biomass,

$$\sum_{i=1}^{n} p_i m_i = \overline{M} \quad (6)$$

as constraints, it can be shown, by using this variational procedure, that

$$p_i = \frac{1}{z} e^{-mi/\overline{m}} \quad (7)$$

where now Z is the equivalent partition function defined as

$$Z = \sum_{i=1}^{n} e^{-mi/\overline{m}}$$

and this predicted probability distribution was proved to conform closely to fishery data. Other applications of MAXENT theory also showed the possibility of recovering other real macroscopic results from this approach (WAGENSBERG & VALLS, 1988). We will now consider the problem of structure and organization of flow networks in ecosystems, using some concepts of information theory and physics.

FLOW-NETWORKS IN ECOLOGY

An ecosystem can be shown as a set of

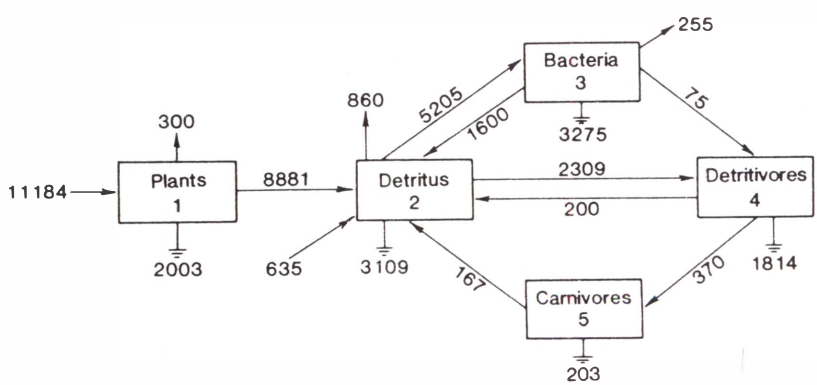

FIGURE 1. Flow energy in the Cone Spring ecosystem. Reprinted from ULANOWICZ (1986). Flows are given in kcal m^{-2} y^{-1}

components $\{C_i\}$ $i = 1,2, ..., n$ between which an extensive quantity (energy, carbon, etc.) is exchanged. This picture can be represented by an oriented graph as is shown, for example, in figure 1. Here the arrows give the direction and magnitude of the flows. In our formal description, we will consider the boundaries (i.e. the origin and sink of resources) as another compartment and the system of figure 1 will now be represented by a graph as in figure 2. Physically, this consideration enables us to talk of a closed system. Two sets of probabilities of emission or reception of energy at nodes $i = 1, 2, ..., n$ and a set of conditional probabilities, $\{p_{ij}\}$ that gives the amount of energy transfer between nodes i and j ($i,j = 1, 2, ..., n$) are necessary in our description.

Some studies of matter and energy flow in an ecosystem (ULANOWICZ, 1986) provide some regularities for this structure. But is it possible to define macroscopic quantities useful to describe a network-flow in an ecosystem? It has been found (WAGENSBERG *et al.*, 1989) that two central magnitudes of Mathematical Theory of Information such as the joint entropy H(X,Y):

$$H(X,Y)= -\sum_j p_j \log_2 p_j - \sum_i \sum_j p_j p_{ij} \log_2 p_{ij} \quad (8)$$

and the information transfer, I (X,Y),

$$I(X,Y)= -\sum_j p_j \log_2 p_j - \sum_i \sum_j p_j p_{ij} \log_2 p_{ij} \quad (9)$$

provide us a general and interesting approach to this problem.

Calculations of H(X,Y) and I(X,Y) from some real ecosystems showed important regularities (Table I). Real data provides us with the set of probabilities p_{ij} of the interaction matrix, say Σ. We can now generate all the possible states compatible with Σ, as is shown in figure 3. The steady state appears on this "space of states" indicated by a cross. Two essential facts should be emphasized after inspection of Table I: (1) H_{stat} always falls in a quite a narrow interval around the 3 bits and (2) I_{stat} tends to attain the maximum value that is compatible with the given H_{stat} except in the case of Wingra Lake. The case of this lake is quite important in our interpretation of these data. We believe that I (X,Y) values are some kind of measure of the global fitness of the ecosystem, as the result of a mechanism of selection. Wingra Lake is an eutrophic lake, and in this case a

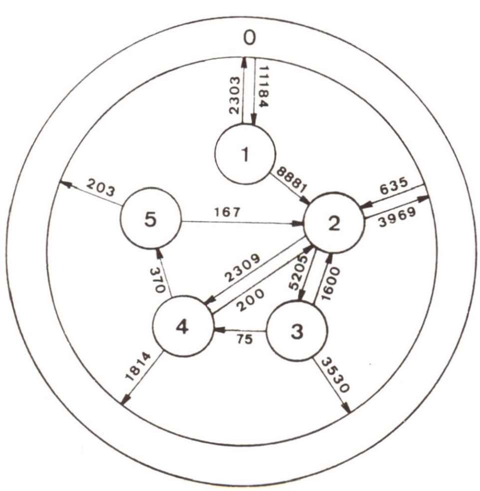

FIGURE 2. Graph representation of Cone Spring ecosystem including surroundings. 1, plants. 2, detritus. 3, bacteria. 4, detritivores. 5, carnivores.

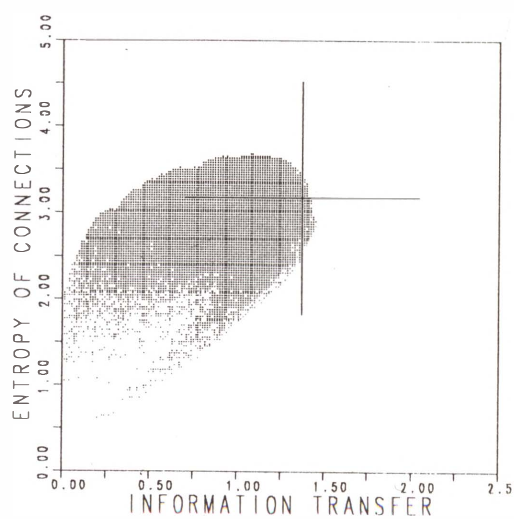

FIGURE 3. Domain of available states derived from the Cone-Spring matrix. The cross is the position of real stationary state.

TABLE I. H and I values at the stationary state (when the inflow equals the outflow at each node) for six different well known ecosystems (RICHNEY *et al.*, 1978; ULANOWICZ, 1986).

ECOSYSTEM	Nodes	Hstat	Hmax	Istat	Imax
Findlay Lake	4	3.08	3.28	1.16	1.25
Mirror Lake	5	3.17	3.57	1.25	1.38
Cone Spring	5	3.07	3.55	1.32	1.40
Wingra Lake	5	2.95	3.64	0.76	0.99
Arion Lake	5	3.01	3.63	0.97	1.13
Crystal River	7	2.93	4.08	1.25	1.35

strong hydrodynamic instability generates continuous resuspension of sediments.

These data enable us to talk about two "magic numbers" in ecological systems, in this sense of significant tendences in ecosystem organization. The first one, the joint entropy H(X,Y) (say entropy of connections) can be interpreted as the result of thermodynamical obliged dissipations at each node. If we look at the compartments as thermodynamic machines, in the light of the second law of the thermodynamics, some part of the inflow energy will always be lost in form of dissipation. The second one, the values of I(X,Y), are in agreement with our proposal of information transfer as a relevant quantity suggestive of an underlying variational principle.

MAXIMUM INFORMATION TRANSFER

We will now try to apply the MAXENT approach to our last results. Physical constraints are limiting factors on the entropy of connections H(X,Y). If we consider this fact as a constraint itself, we can find the maximum value of information transfer under the next set of conditions:

a) normalization of probabilities:

$$\{\sum_{i=1}^{n} p_{ij}\} = 1\} \quad j = 1, 2, ..., n \qquad (10.1)$$

and

$$\sum_{j=1}^{n} p_j = 1 \qquad (10.2)$$

and:

b) a fixed given value of H(X,Y), say H_0. If we perform the variation:

$$\delta\ [-\sum_i p_i \log_2 p_i + \sum_i \sum_j p_j p_{ij} \log_2 p_{ij} +$$

$$+\alpha\ (H_0 + \sum_j p_j \log_2 p_j + \sum_i \sum_j p_j p_{ij} \log_2 p_{ij}) +$$

$$+ \sum_j \beta_j (1 - \sum_i p_{ij}) + \gamma (1 - \sum_j p_j)] \qquad (11)$$

we obtain the following equations:

$$(1 + \log_2 p_i)\, p_j + p_j (1 + \log_2 p_{ij}) (1 + a) - - \beta_j = 0 \qquad (12.1)$$

and

$$- (1 + \log_2 p_{ij})\, p_{ij} - H(x/j) + \alpha (1 + \log_2 p_j) + \alpha\, H(x/j) - \gamma = 0 \qquad (12.2)$$

Equations (12) are difficult to solve, but a solution of this system with theoretical implications can be obtained by considering that: (a) In real ecosystems at stationary state, probabilities of inflows equal probabilities of outflows at each node; (b) H (X,Y) values suggest that essentially only three levels are significative in the energy transfer through the flow network, and (c) If H (X,Y) = Ho, maximization of I (X,Y) implies in fact maximization of H(X) + H(Y) + H_0, i.e. of H(X) + H(Y). It can be shown that, under the above-mentioned conditions, max {H(X) + H(Y)} holds for equal probabilities of inputs and outputs (uniform distributions) and Σ normalized by arrows and columns.

TABLE II. a) Flow matrix of the Cone Spring ecosystem. b) Interaction Matrix.

a)

Producers	Consumers	Environment	
0	0	11184	Producers
8881	2815.7	635	Consumers
2303	9516	0	Environment

b)

Producers	Consumers	Environment	
0.0000	0.0000	0.9463	Producers
0.7941	0.2283	0.0537	Consumers
0.2059	0.7717	0.0000	Environment

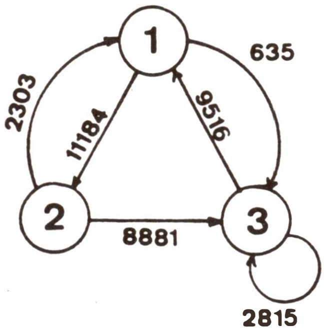

FIGURE 4. A typical three-compartment ecological graph corresponding to the Cone Spring ecosystem. 1, environment. 2, producers. 3, consumers.

If we describe the Cone Spring case using a three-level model, where only surroundings, primary producers and the others are considered (typically as in Fig. 4), approximately uniform distributions are obtained and the interaction matrix appears normalized by arrows and columns.

In Table II we can see the energy-flow matrix, the corresponding interaction matrix Σ; the probabilities of inflows (outflows) at each node in the stationary state are p(1) = 0.3165, p(2) = 0.3490 and p(3)= 0.3345.

DISCUSSION

Ecological flow-networks seem to be self-organized in a particular way, characterized by a high information Transfer value compatible with a thermodynamically limited Joint Entropy. This tendency is especially apparent when we reduce our description of an ecosystem to a graph with only three nodes: primary producers, consumers and surroundings. In the Cone Spring case, real data are close to theoretically predicted values. The preliminary analysis of some other real systems suggests that this might be a fairly general tendency in the organization of flow-networks in ecosystems. Our proposal is that some kind of global efficiency is related to I(X,Y) and has been maximized in biological evolution. Interaction between sub-levels at the third node may have been selected to provide an adequate quantity of energy cycling. A compromise between persistence and global efficiency can explain then the special features of the Σ energy transfer matrix and the tendency towards the maximization of Information Transfer.

REFERENCES

JAYNES, E.T. 1985. Where do we go from here?. In: *Maximum Entropy and Bayesian Methods in Inverse Problems* (R. Smith & W.T. Grandy, eds.): 21- 58. Reidel. Dordrecht.

HAKEN, H. 1985. Application of the maximum information entropy principle to self-organizing systems. *Z. Phys. B.*, 61: 335-338.

HAKEN, H. 1988. *Information and self-organization.* Springer. Berlin, Heidelberg.

LURIE, D. & WAGENSBERG, J. 1983. On biomass diversity in ecology. *Bull. Math. Biol.*, 45: 287-294.

LURIE, D., VALLS, J. & WAGENSBERG, J. 1983. Thermodynamic approach to biomass distribution in ecological systems. *Bull. Math. Biol.*, 45: 869-873.

MACARTHUR, M. 1955. Fluctuations of populations, and a measure of community stability. *Ecology*, 36: 533-536.

MARGALEF, R. 1958. Information theory in ecology. *Gen. Syst*, 3: 36-71.

RICHEY, J.E., WISSMAR, R.C., DEVOL, A.H., LIKENS, G.E., EATON, J.S., WETZEL, R.G., ODUM, W.E., JOHNSON, N.M., LOUCKS, O.L., PRENTKY, R.T. & RICH, P.H. 1978. Carbon flow in four lake ecosystems: a structural approach. *Science,* 202: 1183-1186.

SHANNON, C.E. & WEAVER, W. 1962. *The mathematical theory of information.* University of Illinois Press. Urbana, Ill.

ULANOWICZ, R.E. 1986. *Growth and development: ecosystems phenomenology.* Springer. New York.

WEHRL, A. 1978. General properties of entropy. *Rev. Mod. Phys.*, 50: 221-246.

WAGENSBERG, J., VALLS, J. & BERMÚDEZ, J. 1988. Biological adaptation and the mathematical theory of information. *Bull. Math. Biol.*, 50: 445-464.

WAGENSBERG, J., GARCÍA, A. & SOLÉ, R.V. 1990. Connectivity and information transfer in flow-networkks: two magic numbers in nature? *Bull. Math. Biol.*, 52: 733-740