Oikos

O17073
Scotti, M., Bondavalli, C., Bodini, A. and Allesina, S. 2009. Using trophic hierarchy to understand food web structure. – Oikos 118: 1695–1702.

Appendix 1.

Trophic position (TP) and average mutual information (AMI)

Trophic position
Trophic position is determined by the type and magnitude of exchange flows. These flows are given in the matrix form that is presented in Fig. S1.

This matrix can be easily partitioned in sub-matrices with flows involving only living components (species or groups of species) and non-living components (receiving non-trophic flows). To apply the canonical trophic aggregation (C) one extracts the sub-matrix of inter-compartmental exchanges between living nodes (Fig. S1), assigning non-living matter and detritus to the first trophic level (Ulanowicz 1995, Scotti et al. 2006). For the original matrix-based network unfolding method (Higashi et al. 1989) and the path-based network unfolding analysis (Whipple 1998) the whole set of flows in the matrix is considered (decyclization algorithms are not required) and this prevent detritus from being assigned an arbitrary trophic level (Fig. S1).

In C, TP calculation is performed as in Eq. S1. In case of a binary food web the trophic activity would be evenly distributed between the prey (Eq. S2).

$$TP_j = 1 + \sum_{i=1}^{S} TP_i \times \frac{t_{ij}}{T_{.j}} \quad \text{(S1)}$$

$$TP_j = 1 + \sum_{i=1}^{S} TP_i \times \frac{l_{ij}}{n_{.j}} \quad \text{(S2)}$$

S is the total number of nodes, TP_i and TP_j are, respectively, the trophic positions of nodes i and j, while the ratio between t_{ij} and $T_{.j}$ is the fraction by which species i (t_{ij}) enters the diet of species j ($T_{.j}$).

The calculation for binary data makes use of $n_{.j}$, the total number of links entering the species j, and l_{ij} that is 1 if species j consumes species i and 0 if not. The diet composition of each species can be inferred by evenly assigning the whole intake among its prey, in case of unweighted food webs, or proportionally distributing the total input according to the strength of trophic links. So, if a species is partially primary carnivore for half of its energy intake, and herbivore for the remaining half, calculation yields a TP of 2.5 = 0.5 × 2 + 0.5 × 3.

Average mutual information (AMI)
The question whether complexity affects ecosystem stability has long been central in ecology. MacArthur (1955) applied Shannon's information index to the flows in ecosystem networks as

$$H = -k \sum_{i=1}^{S} \sum_{j=1}^{S} \frac{t_{ij}}{T_{..}} \log_2 \left(\frac{t_{ij}}{T_{..}} \right) \quad \text{(S3)}$$

where H and S are the diversity of flows and the number of species in the network, respectively; k is a scalar constant, and t_{ij} is the flow from node i to node j, with $T_{..}$ that indicates the total amount of energy throughout the network (total system throughput, TST)

$$TST = \sum_{i=1}^{S} \sum_{j=1}^{S} t_{ij} = T_{..} \quad \text{(S4)}$$

The increasing consensus around this index stimulated its application to the more accessible stocks of biomass, shifting the discussion from flow diversity to biomass diversity and its effects on stability. Unfortunately, May (1972) demonstrated that a higher biodiversity in linear dynamical systems was more likely to result in instability and ecologists quickly abandoned applications of information theory to food webs, maintaining the same prejudice also when Rutledge et al. (1976) applied a Bayesian emendation of Shannon's measure to MacArthur's index of flow diversity. These authors used the notion of conditional probability and split MacArthur's index in two complementary terms. The joint probability that an arbitrary elementary unit of currency both leaves i and enters j can be estimated by the quotient ($t_{ij}/T_{..}$), whereas the conditional probability that the unit goes to compartment j, given it already left i ($T_{i.}$), or that exhibited by a flow exiting the node i in respect to the total input to compartment j ($T_{.j}$) are defined, respectively, as

$$\frac{t_{ij}}{\sum_{z=1}^{S} t_{iz}} = \frac{t_{ij}}{T_{i.}} \quad \text{(S5)}$$

$$\frac{t_{ij}}{\sum_{r=1}^{S} t_{rj}} = \frac{t_{ij}}{T_{.j}} \quad \text{(S6)}$$

As a consequence, the measure of total flow diversity is amended as follows

$$H = AMI + H_c \quad \text{(S7)}$$

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

$$AMI = k \sum_{i=1}^{S} \sum_{j=1}^{S} \frac{t_{ij}}{T_{..}} \log_2 \left(\frac{t_{ij} T_{..}}{T_{i.} T_{.j}} \right) \quad \text{(S8)}$$

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

$$H_C = -k\sum_{i=1}^{s}\sum_{j=1}^{s}\frac{t_{ij}}{T_{..}}\log_2\left(\frac{t_{ij}^2}{T_{i\cdot}T_{\cdot j}}\right) \tag{S9}$$

Therefore, the overall complexity of the flow structure, as measured by the MacArthur's index, can be divided in two parts: a) AMI that estimates how orderly and coherently flows are connected; b) H_c that gauges the disorder and freedom that is preserved. Rutledge et al. (1976) proposed H_c as an appropriate measure of ecosystem maturity (Odum 1969), but further studies (Atlan 1974, Ulanowicz 1980) suggested AMI as more reliable index to describe the developmental status of an ecological network. However, Ulanowicz and Wolff (1991) adopted H_c as a tool to estimate effective connectance per node in ecosystems. In particular, dividing H_c in two terms reveals more about its meaning:

$$\begin{aligned} H_c &= -\sum_{i=1}^{s}\sum_{j=1}^{s}\frac{t_{ij}}{T_{..}}\log_2\left(\frac{t_{ij}}{T_{i\cdot}}\right) - \sum_{i=1}^{s}\sum_{j=1}^{s}\frac{t_{ij}}{T_{..}}\log_2\left(\frac{t_{ij}}{T_{\cdot j}}\right) = \\ &= \sum_{i=1}^{s}\frac{T_{i\cdot}}{T_{..}}\left[-\sum_{j=1}^{s}\frac{t_{ij}}{T_{i\cdot}}\log_2\left(\frac{t_{ij}}{T_{i\cdot}}\right)\right] + \sum_{j=1}^{s}\frac{T_{\cdot j}}{T_{..}}\left[-\sum_{i=1}^{s}\frac{t_{ij}}{T_{\cdot j}}\log_2\left(\frac{t_{ij}}{T_{\cdot j}}\right)\right] = \\ &= \sum_{i=1}^{s}\frac{T_{i\cdot}}{T_{..}}H_{i\cdot} + \sum_{j=1}^{s}\frac{T_{\cdot j}}{T_{..}}H_{\cdot j} \end{aligned} \tag{S10}$$

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

$$H_{i\cdot} = -\sum_{j=1}^{s}\frac{t_{ij}}{T_{i\cdot}}\log_2\left(\frac{t_{ij}}{T_{i\cdot}}\right) \tag{S11}$$

and

$$H_{\cdot j} = -\sum_{i=1}^{s}\frac{t_{ij}}{T_{\cdot j}}\log_2\left(\frac{t_{ij}}{T_{\cdot j}}\right) \tag{S12}$$

Average diversity of the biomass going to consumers, weighted by total outputs ($T_{i.}$), and average diversity of inflows, weighted by total inputs ($T_{.j}$), constitute Eq. S10, with the average diversity over both input and output that can be written as $H_c/2$. Because the diversity of pathways through a decision tree is an exponential function of the number of branch points that generate the tree, the mean number of flows from a typical node in the network should be

$$m = 2^{H_c/2} \tag{S13}$$

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

$$n_{i\cdot} = \begin{cases} 2^{H_{i\cdot}} & \\ 0 & \text{if } T_{i\cdot} = 0 \end{cases} \tag{S14}$$

$$n_{\cdot j} = \begin{cases} 2^{H_{\cdot j}} & \\ 0 & \text{if } T_{\cdot j} = 0 \end{cases} \tag{S15}$$

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

$$LD = \frac{1}{2}\left(\sum_{i=1}^{s}\frac{T_{i\cdot}}{T_{..}}2^{H_{i\cdot}} + \sum_{j=1}^{s}\frac{T_{\cdot j}}{T_{..}}2^{H_{\cdot j}}\right) \tag{S16}$$

Then, the difference between the effective connectance (m) proposed by Ulanowicz and Wolff (1991) and the link density (LD) formulated by Bersier et al. (2002) resides solely in the weighting which applies, in the first case, to outflow and inflow diversities, and to taxa's equivalent numbers of consumers and prey in the latter. In particular, the effect of weighting is larger when m is computed, being applied to diversities as exponents in the geometric mean of input and output effective connectance.

Both the applications developed by Ulanowicz and Wolff (1991) and Bersier et al. (2002) are obtained from output (Eq. S11) and input (Eq. S12) diversities, aiming to identify average connectance per node. In particular, they refer to the Eq. S9, making use of information on residual diversity (H_c) for total equivalent links (both entering and exiting each node).

In the present manuscript we discuss an alternative approach, focusing on average mutual information (Eq. S8) which accounts for constraints in the flow structure. We split the whole index into relative contribution of flows entering or exiting each node, weighting their effect with the corresponding throughput ($T_{.j}$ or $T_{i.}$)

$$AMI_{\cdot j} = \frac{1}{T_{\cdot j}}\left[\sum_{i=1}^{s}t_{ij}\log_2\left(\frac{t_{ij}T_{..}}{T_{i\cdot}T_{\cdot j}}\right)\right] \tag{S17}$$

$$AMI_{i\cdot} = \frac{1}{T_{i\cdot}}\left[\sum_{j=1}^{s}t_{ij}\log_2\left(\frac{t_{ij}T_{..}}{T_{i\cdot}T_{\cdot j}}\right)\right] \tag{S18}$$

The information is correlated both to the level of input flow articulation, for each node j, and to outflow diversity of i prey when the AMI on incoming links is estimated (Eq. S17), while adopting its counterpart on outgoing links, evenness of flows exiting each node i and entering its predators j is measured (Eq. S18).

A generalist trophic behavior, meaning more indeterminacy of flow structure, is described by a lower AMI on inflows ($AMI_{.j}$) than in case of specialized diets, while the tendency to avoid sharing natural enemies (apparent competition) with similar species, represented by peculiar pathways linking a prey to its predators, is associated to higher AMI on outflows ($AMI_{i.}$).

Since the catalyst for the formulation of AMI on inflows and outflows is the Shannon measure of entropy (Shannon 1948), these indices reach their minimum when all the input flows to node j, or output links from compartment i, occur in equal intensity, while the maximum is a function of the energy/matter distribution in each event. Moreover, their contribution to the whole AMI depends on the fraction of throughput processed by each node j ($T_{.j}$), or i ($T_{i.}$), as regards to TST ($T_{..}$)

$$\begin{aligned}\mathrm{AMI} &= \sum_{j=1}^{s} \frac{T_{\cdot j}}{T_{\cdot\cdot}} \mathrm{AMI}_{\cdot j} = \\ &= \sum_{j=1}^{s} \frac{T_{\cdot j}}{T_{\cdot\cdot}} \frac{1}{T_{\cdot j}} \left[\sum_{i=1}^{s} t_{ij} \log_2 \left(\frac{t_{ij} T_{\cdot\cdot}}{T_{i\cdot} T_{\cdot j}} \right) \right] = \\ &= \sum_{i=1}^{s} \sum_{j=1}^{s} \frac{t_{ij}}{T_{\cdot\cdot}} \log_2 \left(\frac{t_{ij} T_{\cdot\cdot}}{T_{i\cdot} T_{\cdot j}} \right) \end{aligned} \tag{S19}$$

$$\begin{aligned}\mathrm{AMI} &= \sum_{i=1}^{s} \frac{T_{i\cdot}}{T_{\cdot\cdot}} \mathrm{AMI}_{i\cdot} = \\ &= \sum_{i=1}^{s} \frac{T_{i\cdot}}{T_{\cdot\cdot}} \frac{1}{T_{i\cdot}} \left[\sum_{j=1}^{s} t_{ij} \log_2 \left(\frac{t_{ij} T_{\cdot\cdot}}{T_{i\cdot} T_{\cdot j}} \right) \right] = \\ &= \sum_{i=1}^{s} \sum_{j=1}^{s} \frac{t_{ij}}{T_{\cdot\cdot}} \log_2 \left(\frac{t_{ij} T_{\cdot\cdot}}{T_{i\cdot} T_{\cdot j}} \right) \end{aligned} \tag{S20}$$

References

Almunia, J. et al. 1999. Benthic-pelagic switching in a coastal subtropical lagoon. – Estuarine Coastal Shelf Sci. 49: 363–384.

Atlan, H. 1974. On a formal definition of organization. – J. Theor. Biol. 45: 295–304.

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

Baird, D. and Ulanowicz, R. E. 1989. The seasonal dynamics of the Chesapeake Bay ecosystem. – Ecol. Monogr. 59: 329–364.

Baird, D. et al. 1998. Assessment of spatial and temporal variability in ecosystem attributes of the St Marks National Wildlife Refuge, Apalachee Bay, Florida. – Estuarine Coastal Shelf Sci. 47: 329–349.

Bersier, L.-F. et al. 2002. Quantitative descriptors of food-web matrices. – Ecology 83: 2394–2407.

Hagy, J. D. 2002. Eutrophication, hypoxia and trophic transfer efficiency in Chesapeake Bay. PhD thesis. – Univ. of Maryland at College Park.

Higashi, M. et al. 1989. Food network unfolding – an extension of trophic dynamics for application to natural ecosystems. – J. Theor. Biol. 140: 243–261.

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

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

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

Monaco, M. E. and Ulanowicz, R. E. 1997. Comparative ecosystem trophic structure of three US mid-Atlantic estuaries. – Mar. Ecol. Prog. Ser. 161: 239–254.

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

Patrício, J. et al. 2004. Ascendency as an ecological indicator: a case study of estuarine pulse eutrophication. – Estuarine Coastal Shelf Sci. 60: 23–35.

Rutledge, R. W. et al. 1976. Ecological stability: an information theory viewpoint. – J. Theor. Biol. 57: 355–371.

Scotti, M. et al. 2006. Effective trophic positions in ecological acyclic networks. – Ecol. Modell. 198: 495–505.

Shannon, C. E. 1948. A mathematical theory of communications. – Bell System Tech. J. 27: 379–423.

Steele, J. H. 1974. The structure of marine ecosystems. – Harvard Univ. Press.

Ulanowicz, R. E. 1980. An hypothesis on the development of natural communities. – J. Theor. Biol. 85: 223–245.

Ulanowicz, R. E. 1986. Growth and development: ecosystems phenomenology. – Springer.

Ulanowicz, R. E. 1995. Ecosystem trophic foundations: Lindeman exonerata. – In: Patten, B. C. and Jorgensen, S. E. (eds), Complex ecology: the part-whole relation in ecosystems. Prentice-Hall, pp. 549–560.

Ulanowicz, R. E. and Wolff, W. F. 1991. Ecosystem flow networks: loaded dice? – Math. Biosci. 103: 45–68.

Ulanowicz, R. E. et al. 1997. Network analysis of trophic dynamics in south Florida ecosystems, FY 96: The Cypress Wetland ecosystem. – Tech. Rep. [UMCES] CBL 97-075, Chesapeake Biol. Lab., Solomons.

Ulanowicz, R. E. et al. 1998. Network analysis of trophic dynamics in south Florida ecosystems, FY 97: The Florida Bay ecosystem. – Tech. Rep. [UMCES] CBL 98-123, Chesapeake Biol. Lab., Solomons.

Ulanowicz, R. E. et al. 1999. Network analysis of trophic dynamics in south Florida ecosystem, FY 98: the mangrove ecosystem. – Tech. Rep. [UMCES] CBL 99-0073, Chesapeake Biol. Lab., Solomons.

Ulanowicz, R. E. et al. 2000. Network analysis of trophic dynamics in south Florida ecosystems FY 99: the graminoid ecosystem. – Tech. Rep. [UMCES] CBL 00-0176, Chesapeake Biol. Lab., Solomons.

Whipple, S. J. 1998. Path-based network unfolding: a solution for the problem of mixed trophic and non-trophic processes in trophic dynamic analysis. – J. Theor. Biol. 190: 263–276.

Wulff, F. and Ulanowicz, R. E. 1989. A comparative anatomy of the Baltic Sea and Chesapeake Bay ecosystems. – In: Wulff, F. et al. (eds), Network analysis in marine ecology: methods and applications. Vol. 32 of Coastal and Estuarine Studies. Springer.

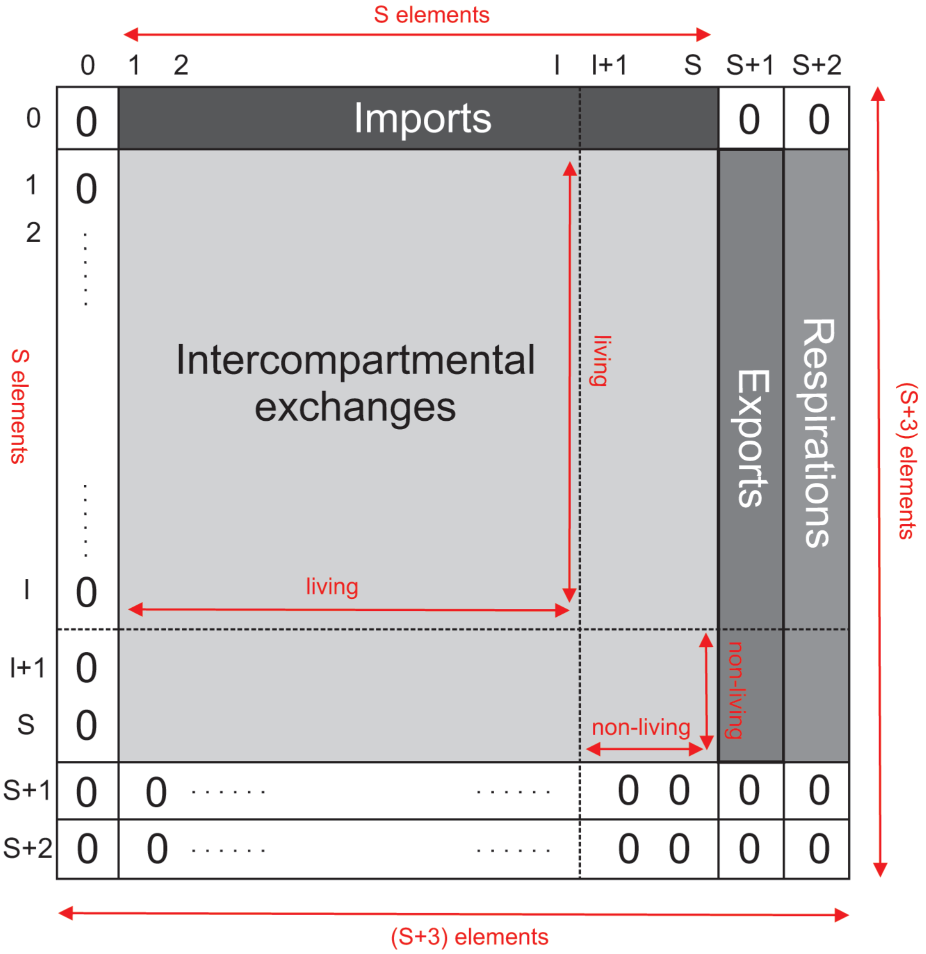

Figure S1. A squared (S+3) × (S+3) matrix is another form to represent ecological flow networks. It is comprised of exchange flows between system compartments (l living and nl non-living) and the outside environment. Flows from and to the outside environment are called imports, exports and respirations (non-usable energy). Food webs, in this work, originated from ecological flow networks and link intensities were determined by flow values of the corresponding network matrices.

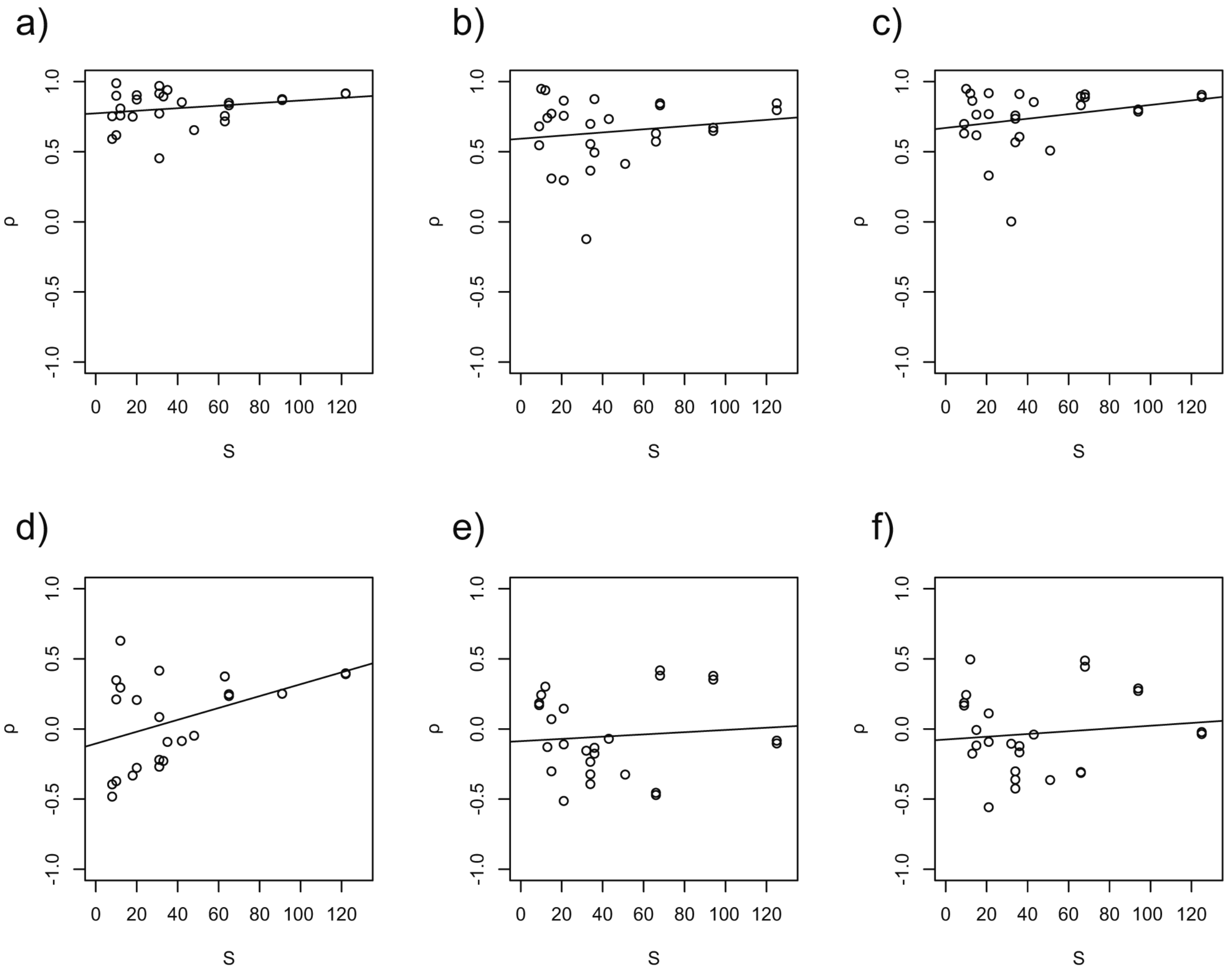

Figure S2. Plots of Spearman's rho values against food web size in presence of weighted data. Plots (a), (b) and (c) refer to the Input scenario (AMI computed on incoming links). Plots (d), (e) and (f) refer to the Output scenario (AMI on outflows). In plots (a) and (d) correlation coefficients are based on AMI vs C; correlation coefficients that make plots (b) and (e) are based on AMI vs H; finally, Spearman's rho values in plots (c) and (f) were computed as AMI vs W. Linear regression lines are shown (S = number of compartments)

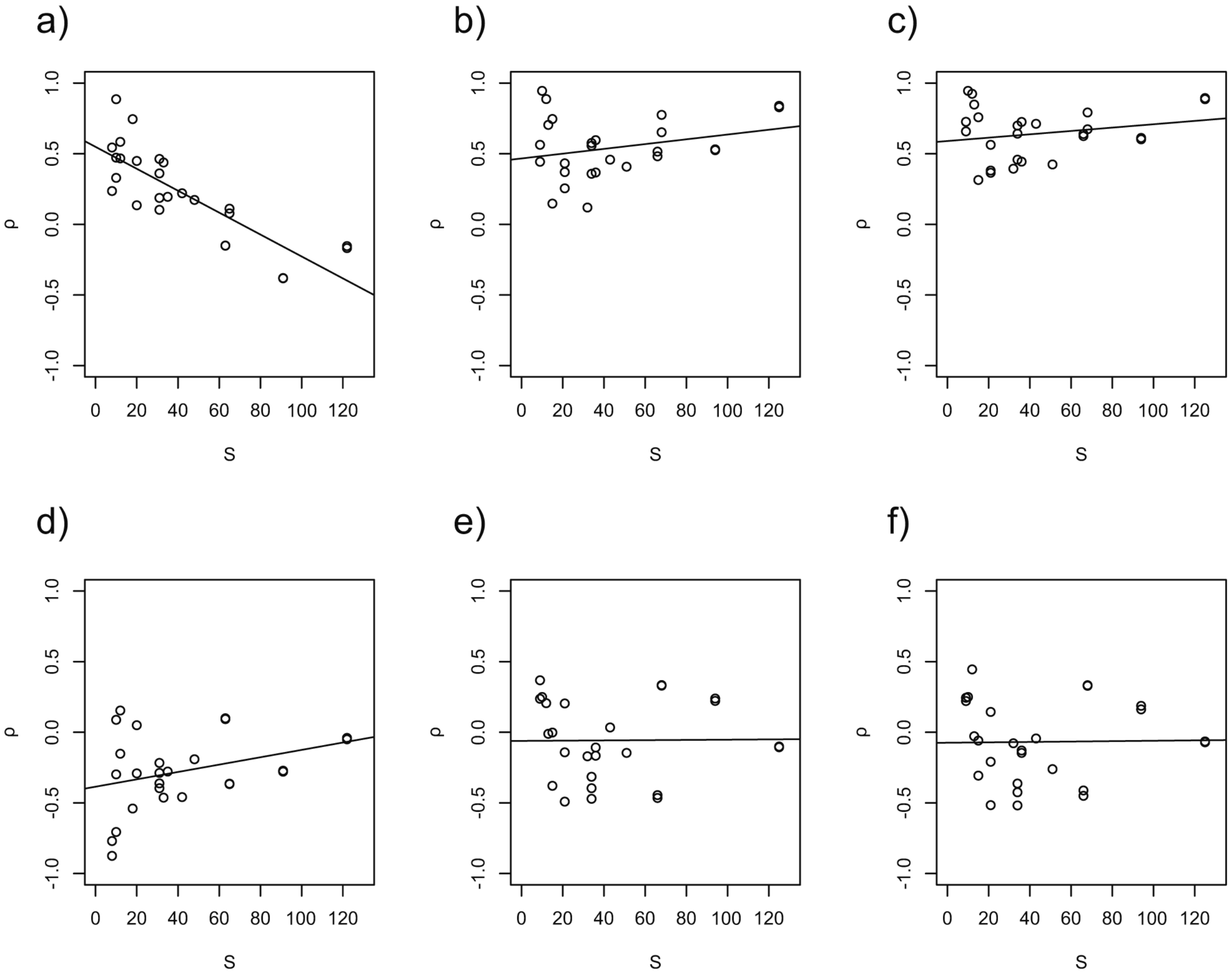

Figure S3. Effect of network size (S = number of compartments) on the correlations between AMI and TP in the 26 binary food webs. The keys for plots are the same as in Fig. S2. All the correlations are scale-independent but the first case, when strictly living exchanges are considered in the comparison between AMI on incoming links and TP ($R^2 = 0.665$, $p << 0.001$). Linear regression lines are depicted.

Table S1. List of the ecological networks considered in the analysis. Total number of nodes (S) and number of non-living compartments (nl) are given. Flow intensities are measured as energy (i.e, kcal m^{-2} $year^{-1}$ or cal cm^{-2} $year^{-1}$), carbon (i.e. g C m^{-2} $year^{-1}$ or mg C m^{-2} $summer^{-1}$) and ash free dry weight (g AFDW m^{-2} $year^{-1}$). The last column summarizes the references to networks analyzed.

Ecosystems	S	nl	Flow units	References
NETWRK				
Aggregated Baltic Sea	15	3	mg C m^{-2} day^{-1}	Wulff and Ulanowicz 1989
Cedar Bog Lake	9	1	cal cm^{-2} $year^{-1}$	Lindeman 1942
Charca de Maspalomas	21	3	mg C m^{-2} day^{-1}	Almunia et al. 1999
Chesapeake Mesohaline Ecosystem	15	3	mg C m^{-2} day^{-1}	Wulff and Ulanowicz 1989
Chesapeake Mesohaline Network	36	3	mg C m^{-2} $summer^{-1}$	Baird and Ulanowicz 1989
Crystal River Creek (control)	21	1	mg C m^{-2} day^{-1}	M. Homer W. M. and Kemp unpubl., Ulanowicz 1986
Crystal River Creek (delta temp.)	21	1	mg C m^{-2} day^{-1}	M. Homer W. M. and Kemp unpubl., Ulanowicz 1986
Lower Chesapeake Bay in Summer	34	3	mg C m^{-2} $summer^{-1}$	Hagy 2002
St. Marks River (Florida) Flow Network	51	3	mg C m^{-2} day^{-1}	Baird et al. 1998
Lake Michigan Control Network	36	1	g C m^{-2} $year^{-1}$	A. E. Krause and D. M. Mason unpubl.
Middle Chesapeake Bay in Summer	34	3	mg C m^{-2} $summer^{-1}$	Hagy 2002
Mondego Estuary	43	1	g AFDW m^{-2} $year^{-1}$	Patrício et al. 2004
Final Narraganasett Bay Model	32	1	mg C m^{-2} $year^{-1}$	Monaco and Ulanowicz 1997
North Sea	10	0	kcal m^{-2} $year^{-1}$	Steele 1974
Somme Estuary	9	1	g C m^{-2} $year^{-1}$	H. Rybarczyk unpubl.
Upper Chesapeake Bay in Summer	34	3	mg C m^{-2} $summer^{-1}$	Hagy 2002
Upper Chesapeake Bay	12	2	g C m^{-2} $year^{-1}$	A. Osgood unpubl.
Ythan Estuary	13	3	g C m^{-2} $year^{-1}$	Baird and Milne 1981
ATLSS				
Cypress Wetlands (dry season)	68	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 1997
Cypress Wetlands (wet season)	68	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 1997
Marshes and Sloughs (dry season)	66	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 2000
Marshes and Sloughs (wet season)	66	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 2000
Florida Bay (dry season)	125	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 1998
Florida Bay (wet season)	125	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 1998
Mangroves (dry season)	94	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 1999
Mangroves (wet season)	94	3	g C m^{-2} $year^{-1}$	Ulanowicz et al. 1999