Package **NetIndices**, network indices and food web descriptors in **R**

Karline Soetaert
Netherlands Institute of Ecology
the Netherlands

Julius Kipkeygon Kones
University of Nairobi
Kenya

July 17, 2008

R package **NetIndices** is designed to estimate the most common network indices.

In this vignette we first deal with conventions adopted, after which network functions are briefly discussed. The formulations for all network indices are represented in several tables.

1 Notations and flow matrix conventions

The descriptions of symbols used in network indices computations are shown in Table 1. [1] As in Latham (2006) we adopt for these tables the convention as described in [9].

We assume that a system has n biotic and abiotic compartments. The flow value T_{ij} is defined as a destination-source flow (i.e. j ->i) .

Quantitative flows between compartments of a web are classified into four types [7]:

- exogenous inputs (imports),
- inter-compartmental exchanges,
- exports of useable medium, and
- dissipation of unusable medium .

The source compartment of imports to the internal network is labeled with number 0 (zero), the destination of usable exports (secondary production) is labeled n+1 and the destination of unusable exports (respiration/dissipation) is labeled n+1 (sensu Hirata and Ulanowicz 1984). The flow matrix, with source compartments in columns and destination compartments in rows, has dimensions $0 <= j <= n$ and $1 <= i <= n+2$. A matrix containing all flows within a web has dimensions of $1 <= i <= n$ and $1 <= j <= n$.

2 Arguments to network index functions

In all functions of **NetIndices**, the network can be inputted in two ways:

[1]As our work generally involves food webs, our notation/terminology will be skewed to this field; hence we will use the term ”web” where others might use ”network”, and ”flow” instead of ”link”

- $Flow$, a matrix defined as source i -> destination j
- T_{ij}, the transpose of $Flow$, i.e. a matrix defined as destination i <- source j

Internally the calculation uses T_{ij}

If present, the row- and -or column names of $Flow$ or T_{ij} are used to label the compartments. This is recommended.

All functions distinguish between internal components and external components, Externals are either specified by their name (more general, only applicable if the compartments have been labelled) or by a number (error-prone):

- Import, externals that are a source to the network. If specified by numbers they should refer to *columns* of T_{ij} (or rows of Flow)
- Export, externals that are a sink to the network. If specified by numbers they should refer to *rows* of T_{ij} (or columns of Flow)

3 Network indices

The R-functions for computing network indices are in Tables 2-8. They fall in several categories:

- function *GenInd*. General network indices. In this category we consider a number of general systems' properties. [14]
- function *UncInd*. Network Uncertainty indices, based on communication theory. [25]
- function *AscInd*. System's growth and development. They are the ascendancy, development capacity and overhead. e.g. [26] They can similarly be defined at four decomposed stages of a system: import (state 0), internal (between the compartments), export and dissipation [28].
- function *PathInd* Path analysis. Identifies the direct and indirect pathways in a network. (e.g. [10])
- function *EnvInd* Environ network indices. ([20])
- function *TrophInd*. Trophic level and Omnivory index ([4]. The trophic level of a consumer equals 1 + the weighted average of the trophic levels of its food. Primary producers and the compartments labeled as "detritus" are assumed to have trophic level of 1. The omnivory index measures the variation in trophic levels of the food sources of a consumer.
- function *Dependency* The dependency matrix estimates the direct + indirect dependence of a consumer on a resource.

Note: Most of the index calculations were based on the paper and the software written by Latham ([14]), who did a very commendable (if not heroic) job in gathering all the mathematical formulations of these indices.

However, there were a couple of inconsistencies in the paper of Latham:

(1) The Connectance index ([15]): The L reported in [14] should be L_{int}, because Connectance is only calculated on internal links.

(2) The value of TSTbar in figure (2) of the article was shown incorrectly (as T/n, when it should have been TST/n. It was however correctly described in the paper.

(3) The Synergism index both in the text and the equations were wrong. See Table (7) for how it is correctly estimated

Table 1: Nomenclature for equations

Term	Description
n	Number of internal compartments in the network, excluding 0 (zero),n+1 and n+2
$j = 0$	External source
$i = n+1$	Usable export from the network
$i = n+2$	Unusable export from the network (respiration, dissipation)
T_{ij}	Flow from compartment j to i, where j represents the columns of the flow matrix and i the rows
T^*_{ij}	Flow matrix, excluding flows to and from external
$T_{i.}$	Total inflows to compartment i
$T_{.j}$	Total outflows from compartment j
T_i	Total inflows to compartment i, excluding inflow from external sources
T_j	Total outflows from compartment j, excluding outflow to external sinks
$(\dot{x_i})_-$	A negative state derivative, considered as a gain to the system pool of mobile energy
$(\dot{x_i})_+$	A positive state derivative, considered as a loss from the system pool of mobile energy
z_{i0}	Flow into compartment i from outside the network
$y_{n+,j}$	Flow out of the network for compartment j to compartments n+1 and n+2
c_{ij}	The number of species with which both i and j interact divided by the number of species with which either i or j interact
I, δ_{ij}	Identity matrix and its elements

Table 2: General Network indices

Index name	Code	Formula	Source(s)
Total system throughflow	TST	$\sum_{i=1}^{n}\sum_{j=1}^{n}\left[T_{ij}+z_{i0}-(\dot{x_i})_{-}\right]$	
		$=\sum_{i=1}^{n}\sum_{j=1}^{n}\left[T_{ij}+y_{n+,j}+(\dot{x_j})_{+}\right]$	[14]
Total system throughput	$T..$	$\sum_{i=1}^{n+2}\sum_{j=0}^{n}T_{ij}$	[9]
Number of links	L_{tot}	$\sum_{i=1}^{n+2}\sum_{j=0}^{n}(T_{ij}>0)$	
Number of internal links	L_{int}	$\sum_{i=1}^{n}\sum_{j=1}^{n}(T_{ij}>0)$	
Link density	LD	$\frac{L_t ot}{n}$	[14]
Connectance	C	$\frac{L_{int}}{n\cdot(n-1)}$	[14, 15]
Average link weight	$\overline{T}_{ij}$	$\frac{T_{..}}{L_{tot}}$	[14]
Average compartment throughflow	$\overline{TST}$	$\frac{TST}{n}$	[14]
Compartimentalization	$\overline{C}$	$\frac{1}{n\cdot(n-1)}\cdot\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}c_{ij}$	[21]

Table 3: Network uncertainty indices

Index name	Code	Formula	Source(s)
Average mutual information	AMI	$\sum_{i=1}^{n+2}\sum_{j=0}^{n}\frac{T_{ij}}{T_{..}}log_2\frac{T_{ij}T_{..}}{T_{i.}T_{.j}}$	[27, 23, 3, 13, 25]
Statistical uncertainty	H_R	$-\sum_{j=0}^{n}\frac{T_{.j}}{T_{..}}log_2\frac{T_{.j}}{T_{..}}$	[14, 28]
Conditional uncertainty	D_R	$H_R - AMI$	[14, 28]
Realized uncertainty	RU_R	$\frac{AMI}{H_R}$	[14, 28]
Network uncertainty	H_{max}	$\sum_{i=1}^{n} log_2(n+2)$	[14, 28]
Network efficiency	H_{sys}	$-\sum_{i=1}^{n+2}\sum_{j=1}^{n}\frac{T_{ij}}{T_{..}}log_2\frac{Tij}{T_{.j}}$	[14, 28]
Constraint information	H_c	$H_{max} - H_{sys}$	[14, 28]
Constraint efficiency	CE	$\frac{H_c}{H_{max}}$	[14, 28]

Table 4: System growth and development indices

Index name	Code	Formula	Source(s)
Ascendency	A	$\sum_{i=1}^{n+2}\sum_{j=0}^{n}T_{ij}log_2\frac{T_{ij}T_{..}}{T_{i.}T_{.j}}$	[22, 28]
Development capacity	DC	$-\sum_{i=1}^{n+2}\sum_{j=0}^{n}T_{ij}log_2\frac{T_{ij}}{T_{..}}$	[22, 28]
Overhead	ϕ	$DC - A$	[22, 28]
Extent of development	AC	$\frac{A}{DC}$	[22, 28]

Table 5: Effective measures indices

Index name	Code	Formula	Source(s)
Effective connectivity	CZ	$\prod_{i,j=1}^{n}(\frac{T_{ij}^2}{T_{i.}T_{.j}})^{-0.5 \cdot T_{ij}/T_{..}}$	[29]
Effective flows	FZ	$\prod_{i,j=1}^{n}(\frac{T_{ij}}{T_{..}})^{-T_{ij}/T_{..}}$	[29]
Effective nodes	NZ	$\prod_{i,j=1}^{n}(\frac{T_{..}^2}{T_{i.}T_{.j}})^{0.5 \cdot T_{ij}/T_{..}}$	[29]
Effective roles	RZ	$\prod_{i,j=1}^{n}(\frac{T_{ij}T_{..}}{T_{i.}T_{.j}})^{T_{ij}/T_{..}}$	[29]

Table 6: Pathway analysis

Index name	Code	Formula	Source(s)
Total System cycled throughflow	TST_c	$\sum_{j=1}^{n}(1-\frac{1}{q_{jj}}) \cdot T_j$	[10, 11, 12, 17, 2]
		$Q=[I-G^*]^{-1}$	
		$G^*=\left[T_{ij}^*/max(T_j,T_i)\right]$	
Total System non-cycled throughflow	TST_S	$TST-TST_c$	[10, 11, 12, 17, 2]
Finn's cycling index	FCI	$\frac{TST_c}{TST}$	[10, 11, 12, 17, 2]
Average pathlength	$\overline{PL}$	$\frac{TST}{\sum z_{i0}-\sum(\dot{x}_i)_-}$	
		$=\frac{TST}{\sum y_{n+,j}+\sum(\dot{x}_i)_+}$	[22, 28]

Table 7: Environ analysis

Index name	Code	Formula	Source(s)
Transitive closure matrix	G	$\left[T^*_{ij}/T_j\right]$	[18, 20]
Integral nondimensional matrix	N	$I+G+G^2+\ldots=[I-G]^{-1}$	[18, 17]
Non-dimensional direct flow-based utility matrix	D	$(d_{ij})=\frac{T^*_{ij}-T_{ji}*}{T_i}$	[17, 8]
Utility nondimensional matrix	U	$I+D+D^2+\ldots=[I-D]^{-1}$	[17, 8]
Coefficient of variation of N	CV(N)	$\sqrt{\frac{\sum_{i,j=1}^{n}\left(\overline{N}-N_{ij}\right)^2}{(n^2-1)\cdot\overline{N}^2}}$	[16, 6]
Coefficient of variation of G	CV(G)	$\sqrt{\frac{\sum_{i,j=1}^{n}\left(\overline{G}-G_{ij}\right)^2}{(n^2-1)\cdot\overline{G}^2}}$	[16, 6]
Homogenization	H_p	$\frac{CV(G)}{CV(N)}$	[16, 6]
Integral Utility Matrix	γ	$T_i\cdot U$	[1, 19, 5]
Synergism Index	$\frac{b}{c}$	$\frac{\sum+utility\ \ in\ \ \gamma}{\sum-utility\ \ in\ \ \gamma}$	[1, 19, 5]
Dominance indirect effects	$\frac{i}{d}$	$\frac{\sum\limits_{i,j=1}^{n}(N_{ij}-I_{ij}-G_{ij})}{\sum\limits_{i,j=1}^{n}G_{ij}}$	[1, 19, 5]

Table 8: Trophic analysis

Index name	Code	Formula	Source(s)
Diet matrix	P	$\left[\frac{T^*_{ij}}{T_i}\right]$	
Diet dependency matrix	D	$I+P+P^2+\ldots=[I-P]^{-1}$	
Trophic level of compartment i	TL_i	$1+\sum\limits_{j=1}^{n}(\frac{T^*_{ij}}{T_i}\cdot TL_j)$	[4, 24]
Omnivory index for compartment i	OI_i	$\sum\limits_{j=1}^{n}(TL_j-(TL_i-1))^2\cdot\frac{T^*_{ij}}{T_i}$	[4]

References

[1] Patten BC. Energy, emergy and environs. *Ecological Modelling*, 62:29–69, 1992.

[2] Patten BC, Bosserman RW, Finn JT, and Cale WG. Propagation of cause in ecosystems. In Patten BC, editor, *Systems Analysis and Simulation in Ecology*, volume 4, pages 457–579. Academic Press, New York, 1976.

[3] Shannon CE. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[4] V Christensen and D Pauly. Ecopath ii - a software for balancing steady-state ecosystem models and calculating network characteristics. *Ecological Modelling*, 61:169–185, 1992.

[5] BD Fath. Network analysis applied to large-scale cyber-ecosystems. *Ecological Modelling*, 171:329–337, 2004.

[6] Patten BC Fath BD. Quantifying resource homogenization using network flow analysis. *Ecological Modelling*, 123:193–205, 1999.

[7] Mann KH Field JG, Wulf F. The need to analyze ecological networks. In Mann KH Wulff F, Field JG, editor, *Network analysis in marine ecology: methods and applications*, pages p3–12. Springer-Verlag, Berlin, 1989.

[8] Patten BC Higashi M. Dominance of indirect causality in ecosystems. *The American Naturalist*, 133:288–302, 1989.

[9] Ulanowicz RE Hirata H. Informational theoretical analysis of ecological networks. *The American Naturalist*, 133:288–302, 1984.

[10] Finn JT. Measures of ecosystem structure and function derived from analysis of flows. *Journal of Theoretical Biology*, 56:363–380, 1976.

[11] Finn JT. Cycling index: a general definition for cycling in compartmental models. In Brisbin IL Adriano DC, editor, *Environmental Chemistry and Cycling Processes, DOE Proceedings 45, Conf. 760429*, pages 138–165. National Technical Information Service, Springfield, VA, 1978.

[12] Finn JT. Flow analysis of models of the hubbard brook ecosystem. *Ecology*, 61:562–571, 1980.

[13] Scully EP Latham LG. Quantifying constraint to assess development in ecological networks. *Ecological Modelling*, 154:25–44, 2002.

[14] Latham LG. Network flow analysis algorithms. *Ecological Modelling*, 192:586–600, 2006.

[15] Gardner MR and Ashby WR. Connectance of large dynamic (cybernetic) systems: critical values for stabitlity. *Nature*, 228:784, 1970.

[16] Auble GT Patten BC. System theory of the ecological niche. *The American Naturalist*, 117:893–922, 1981.

[17] Higashi M Patten BC. Modified cycling index for ecological applications. *Ecological Modelling*, 25:69–83, 1984.

[18] Richardson TH Patten BC, Barber MC. Path analysis of a reservoir ecosystem model. *Canadian Water Resources Journal*, 7:252–282, 1982.

[19] Patten BC PFath BD. Network synergism: emergence of positive relations in ecological systems. *Ecological Modelling*, 107:127–143, 1998.

[20] Patten BC PFath BD. Review of the foundations of network environ analysis. *cosystems*, 2:167–179, 1999.

[21] Lawton JH Pimm SL. Are food webs divided into compartments? *Journal of Animal Ecology*, 49:879–898, 1980.

[22] Ulanowicz RE. Ascendency: a measure of ecosystem performance. In Muller F Jorgensen SE, editor, *Handbook of Ecosystem Theories and Management*, pages 303–315. Lewis Publishers, Boca Raton, 2000.

[23] Gallager RG. *Information Theory and Reliable Communication*. Wiley, New York, 1968.

[24] Lindeman RL. The trophic dynamic aspect of ecology. *Ecology*, 23:399–418, 1942.

[25] Mulholland RJ Rutledge RW, Basorre BL. Ecological stability: an information theory viewpoint. *Journal of Theoretical Biology*, 57:355–371, 1976.

[26] R.E. Ulanowicz. An hypothesis on the development of natural communities. *Journal of Theorectical Biology*, 85:223–245, 1980.

[27] RE Ulanowicz. Quantitative methods for ecological network analysis. *Computational Biology and Chemistry*, 28:321–339, 2004.

[28] Norden JS Ulanowicz RE. Symmetrical overhead in flow networks. *International Journal of System Science*, 21:429–437, 1990.

[29] Ulanowicz RE Zorach AC. Quantifying the complexity of flow networks: how many roles are there? *Complexity*, 8:68–76, 2003.