Available online at www.sciencedirect.com

SCIENCE DIRECT®

Ecological Modelling 165 (2003) 221–229

ECOLOGICAL MODELLING

www.elsevier.com/locate/ecolmodel

Steady state of ecosystem flow networks: a comparison between balancing procedures

Stefano Allesina*, Cristina Bondavalli

Department of Environmental Sciences, University of Parma, Parco Area delle Scienze, 33/A, 43100 Parma, Italy

Received 9 August 2002; received in revised form 30 January 2003; accepted 19 February 2003

Abstract

When modelling real ecosystems, a number of techniques need steady state condition to proceed in their analysis. In ecosystem network models this means that energy entering the system exactly balances the output. Steady state, however, is not a straightforward outcome of network construction and, to have this condition satisfied, network analysis uses balancing procedure. This operation leads to restructuring the weighted network, changing the values of some network flows; this can affect drastically the results of the analysis. Presently, two algorithms are used for balancing ecosystem networks: input-based approach and output-based approach. In the former input flows are kept constant while outputs and transfer coefficients are manipulated; the latter requires that inputs and intercompartmental flows are modified. This paper discusses the effects of these algorithms on some products of network analysis, in particular system level indices such as total system throughput (TST) and ascendency. Also it suggests four new procedures that, while balancing the networks, can minimise changes on measured flows and distortion on results of the analysis.

Keywords: Balancing algorithms; Ecological networks; Ecosystems; Simulation; Steady state

1. Introduction

Network analysis (Ulanowicz, 1986) investigates ecosystem's properties by studying the network of relationships among different compartments (species, nutrient pools, etc.). This methodology, which is taking ground among ecologists (Abarca-Arenas and Ulanowicz, 2002; Heymans et al., 2002; Nielsen and Ulanowicz, 2000), requires that every compartment as well as the ecosystem as a whole is in steady state conditions. In particular, this means that flows of matter and energy entering and leaving a given compartment are equal, so that there is no increase or decrease in mass:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = 0, \quad \forall x_i \in S$$

where x_i is variable-compartment that belongs to the system S.

For network construction flows must be quantified, but in many cases, direct quantification is not possible. In all these cases, values are determined by combining:

- literature screening to gather information on diet, biomasses or population densities, metabolic parameters, growth functions, population dynamics of single compartments (guilds, populations);
- estimation by using parameters found in literature and data gathered through sampling or observation;

* Corresponding author. Tel.: +39-0521-905614; fax: +39-0521-905402.
E-mail address: alle@dsa.unipr.it (S. Allesina).

0304-3800/03/$ – see front matter
doi:10.1016/S0304-3800(03)00075-9

- contacting experts on specific taxonomic groups to clarify controversies emerging from the literature.

The outcome of this investigation is an extreme heterogeneous dataset that reflects the different approaches used by the various authors (Ulanowicz, 1989); it inevitably leads to networks that are not in steady state. Furthermore, errors associated with measures, as well as various approximations, contribute to unbalance the network.

To apply network analysis this problem is usually overcome by some balancing procedure that changes some flow values. Depending on how such balancing is done, there can be effects on model's properties, with repercussions on the results of network analysis. The main goal of this paper is to discuss implications of balancing methods on the performance of network analysis. In this framework specific objectives are: (1) highlight problems associated with balancing procedure that are presently in use; and (2) suggest and test new algorithms for balancing the networks, to minimise changes on measured flows and the disturbance on results of network analysis.

In particular, a comparison between the different balancing procedure is performed by considering their effects on the computation of two system level indices that measure the size and organisation of flow networks: total system throughput (TST) and ascendency (Hirata and Ulanowicz, 1984; Ulanowicz, 1986).

2. Methods

2.1. Balancing algorithms: state of the art

Balancing can be achieved by different methods. For example the ECOPATH routine (Christensen and Pauly, 1992, see also Ducklow et al., 1989) uses a method like the Singular Value Decomposition (Polovina, 1984; Savenkoff et al., 2001), which is essentially based on a mass balance approach built around linear metabolic equations for the compartments. Such equations are solved via matrix algebra. No unique solution exists to these models and a balanced network with ECOPATH is obtained basically by trial and error procedure that can make use of Monte Carlo methods. The inverse approach (Parker, 1977) is also based on a system of linear equations. Ecological constraints are imposed on their parameters to constrain the range of possible solutions in terms of flows between the compartments.

Table 1
This squared $M \times M$ (M = no. of compartments + 3) matrix embeds all the fluxes between system compartments and the outside environment

	1	2	…	N	N+1	N+2	N+3
1	Transfers Between Compartments				Exports	Respirations	0
2							0
…							…
N							0
N+1	0	0	…	0	0	0	0
N+2	0	0	…	0	0	0	0
N+3	Imports				0	0	0

The $N \times N$ part is the internal transfer matrix, while in the $N+1$th, $N+2$th, $N+3$th rows and columns the exchanges with the outside are stored. All the coefficients represent fluxes of matter or energy from row-compartments to the column ones.

This study consider balancing procedures used in network analysis (Ulanowicz and Kay, 1991; Ulanowicz, 1989), which usually do not require metabolic equations to be solved and where the only entities considered are flow matrices:

- *input*, or donor-based approach, in which inputs are kept constant while outputs and flow transfer coefficients are manipulated;
- *output*, or predator-based approach, in which manipulation is performed on inputs and intercompartmental flows.

These methods use the same basic procedure, which starts from the so called extended transfer matrix (T^*), whose general form is given in Table 1.

For ease of explanation a formal description of the general *input-based* algorithm is given below, based on a two-compartment network (Fig. 1).

Step 1. In the extended transfer matrix T^* of the network given in Fig. 1, the sums of the ith row (outflows from the ith compartment) and the ith column (inflows to the ith compartment) are not equal, and so the network is not at steady state (unbalanced).

	1	2	$N+1$	$N+2$	$N+3$	Sum
1	0	73.161	25.504	24.029	0	122.694
2	25.640	0	25.958	23.899	0	75.497
$N+1$	0	0	0	0	0	
$N+2$	0	0	0	0	0	
$N+3$	104.383	0	0	0	0	
Sum	130.023	73.161				

Step 2. Divide all the coefficients t_{ij}, $i \in [1, \ldots, N]$, $j \in [1, \ldots, N+3]$ by the ith-row sum to obtain matrix $F^*[f^*_{ij}]$:

$$f^*_{ij} = \frac{t^*_{ij}}{\sum_k^{N+3} t^*_{ik}}$$

	1	2	$N+1$	$N+2$	$N+3$
1	0	0.596	0.208	0.196	0
2	0.340	0	0.344	0.317	0

Step 3. Transpose the $N \times N$ part of matrix F^* and subtract the identity matrix to get matrix R:

$$R \leftarrow F^{*\mathrm{T}} - I$$

	1	2
1	−1	0.340
2	0.596	−1

Step 4. Then invert matrix R:

$$R^{-1} \leftarrow R$$

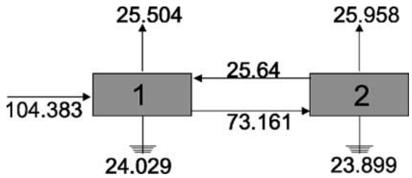

Fig. 1. A two compartments network not in steady state. The ground symbol identifies energy that is dissipated by metabolic processes.

	1	2
1	−1.254	−0.426
2	−0.748	−1.254

Step 5. Multiply every r_{ij} coefficient by the corresponding jth input in the T^* matrix, and change its sign:

$$r_{ij} \leftarrow -r_{ij}(t^*_{N+1,j} + t^*_{N+2,j} + t^*_{N+3,j})$$

	1	2
1	130.889	0
2	78.048	0

Step 6. Sum ith's row to build vector $U[u_i]$:

$$u_i = \sum_{k=1}^{N} r_{ik}$$

Step 7. Multiply each f^*_{ij}, $i \in [1, \ldots, N]$, $j \in [1, \ldots, N+3]$ (Step 2) by the corresponding u_i to obtain a balanced form ($T^{*\mathrm{Bal}}_{\mathrm{In}}$) of the T^* matrix (Table 2).

The output-based approach can be implemented from the input-based one by performing two further transpositions: one on the T^* matrix before Step 2; the second on the resulting balanced matrix. The form of the final $T^{*\mathrm{Bal}}_{\mathrm{Out}}$ related to the network of Fig. 1 is shown in Table 3.

Table 2
$T^{*\text{Bal}}_{\text{In}}$ is the balanced matrix derived from the original unbalanced matrix T^* via the input-based algorithm

	1	2	N+1	N+2	N+3	Sum
1	0	78.048	27.208	25.634	0	130.889
2	26.506	0	26.835	24.706	0	78.048
N+1	0	0	0	0	0	
N+2	0	0	0	0	0	
N+3	104.383	0	0	0	0	
Sum	130.889	78.048				

All the coefficients, except input ones ($N+1$th, $N+2$th, $N+3$th rows), have been changed, making ith row sum equal to ith column sum.

2.2. Derived approaches

Four new algorithms have been derived from the two illustrated before:

(1) AVG: it calculates average coefficients using the corresponding values obtained with input-based and output-based approach.

$$T^{*\text{Bal}}_{\text{AVG}}[i, j] = \frac{1}{2}(T^{*\text{Bal}}_{\text{In}}[i, j] + T^{*\text{Bal}}_{\text{Out}}[i, j])$$

(2) IO: starting from the unbalanced matrix T^*, the input-based approach is applied to $(1/2)T^*$, the result is summed with $(1/2)T^*$. The resulting matrix is completely balanced using the output-based algorithm.

$$T^{*\text{Bal}}_{\text{IO}}[i, j] = T^{*\text{Bal}}_{\text{Out}}\left(\frac{1}{2}T^{*\text{Bal}}_{\text{In}}[i, j] + \frac{1}{2}T^*[i, j]\right)$$

Table 3
$T^{*\text{Bal}}_{\text{Out}}$ is the balanced matrix derived from the original unbalanced matrix T^* via the output-based algorithm

	1	2	N+1	N+2	N+3	Sum
1	0	74.271	25.504	24.029	0	123.804
2	24.414	0	25.958	23.899	0	74.271
N+1	0	0	0	0	0	
N+2	0	0	0	0	0	
N+3	99.390	0	0	0	0	
Sum	123.804	74.271				

All the coefficients, except output ones ($N+1$th, $N+2$th, $N+3$th columns), have been changed, making ith row sum equal to ith column sum.

(3) OI: is the same of IO but first output-based approach is implemented, and then input-based.

$$T^{*\text{Bal}}_{\text{OI}}[i, j] = T^{*\text{Bal}}_{\text{In}}\left(\frac{1}{2}T^{*\text{Bal}}_{\text{Out}}[i, j] + \frac{1}{2}T^*[i, j]\right)$$

(4) AVG2: the same of AVG but with $T^{*\text{Bal}}_{\text{IO}}$ and $T^{*\text{Bal}}_{\text{IO}}$ matrices.

$$T^{*\text{Bal}}_{\text{AVG2}}[i, j] = \frac{1}{2}(T^{*\text{Bal}}_{\text{IO}}[i, j] + T^{*\text{Bal}}_{\text{OI}}[i, j])$$

2.3. Testing the performances

To compare the six different algorithms, unbalanced, realistic matrices have been used. In particular, three systems previously investigated, namely the Cone Spring ecosystem (five compartments) (Tilly, 1968), the Chesapeake Bay (36 compartments) (Wulff and Ulanowicz, 1989) and the Gramminoid Marshes Dry (66 compartments) (Heymans et al., 2002) have been considered. Flow matrices of these ecosystems were randomly unbalanced by changing every coefficient of the extended matrix by ±10%, multiplying them for a random number in uniform distribution [0.9, 1.1]. This operation has been repeated on each matrix 1000 times. For each system the 1000 unbalanced matrices T^* have been treated with the different algorithms.

3. Results

A computer program has been developed to handle the huge amount of computations necessary to carry out the investigation. This software, developed by the authors in MS Visual Basic, stores in a database, for each system: the initial matrix, all its 1000 "randomly unbalanced versions", and the 6 balanced matrices for each simulation, that give rise to 18,000 records (3 systems, 6 algorithms, 1000 repetitions).

To highlight to what extent balancing algorithms may change flow values, average and maximum change, in absolute value, were considered for each matrix. The maximum changes were obtained, by considering the greatest magnitude (in absolute value) of the variation between every coefficient obtained after balancing and its unbalanced counterpart, whereas the average changes were calculated for each matrix by simply averaging such variations:

$$\alpha_{ij} = \left| \frac{T_{ij}^{*\mathrm{Bal}}}{T_{ij}^{*\mathrm{UnBal}}} - 1 \right|, \quad \forall T_{ij}^{*\mathrm{Bal}} \neq 0$$

$$\text{Max. change} = \max[\alpha_{11}, \ldots, \alpha_{NN}]$$

$$\text{Avg. change} = \frac{\sum_{i,j} \alpha_{ij}}{\tau}$$

where τ is the number of $T_{ij}^{*\mathrm{Bal}} \neq 0$.

Fig. 2 gives the results of this calculation in pictorial terms. For every balancing algorithm the 1000 maximum changes (one for every matrix) and 1000 average changes are depicted as small circles.

To compare the effects of balancing procedures on network analysis performances, two system level indices have been considered:

(a) TST (Ulanowicz, 1986): It measures the total amount of flows that pass through the system in a

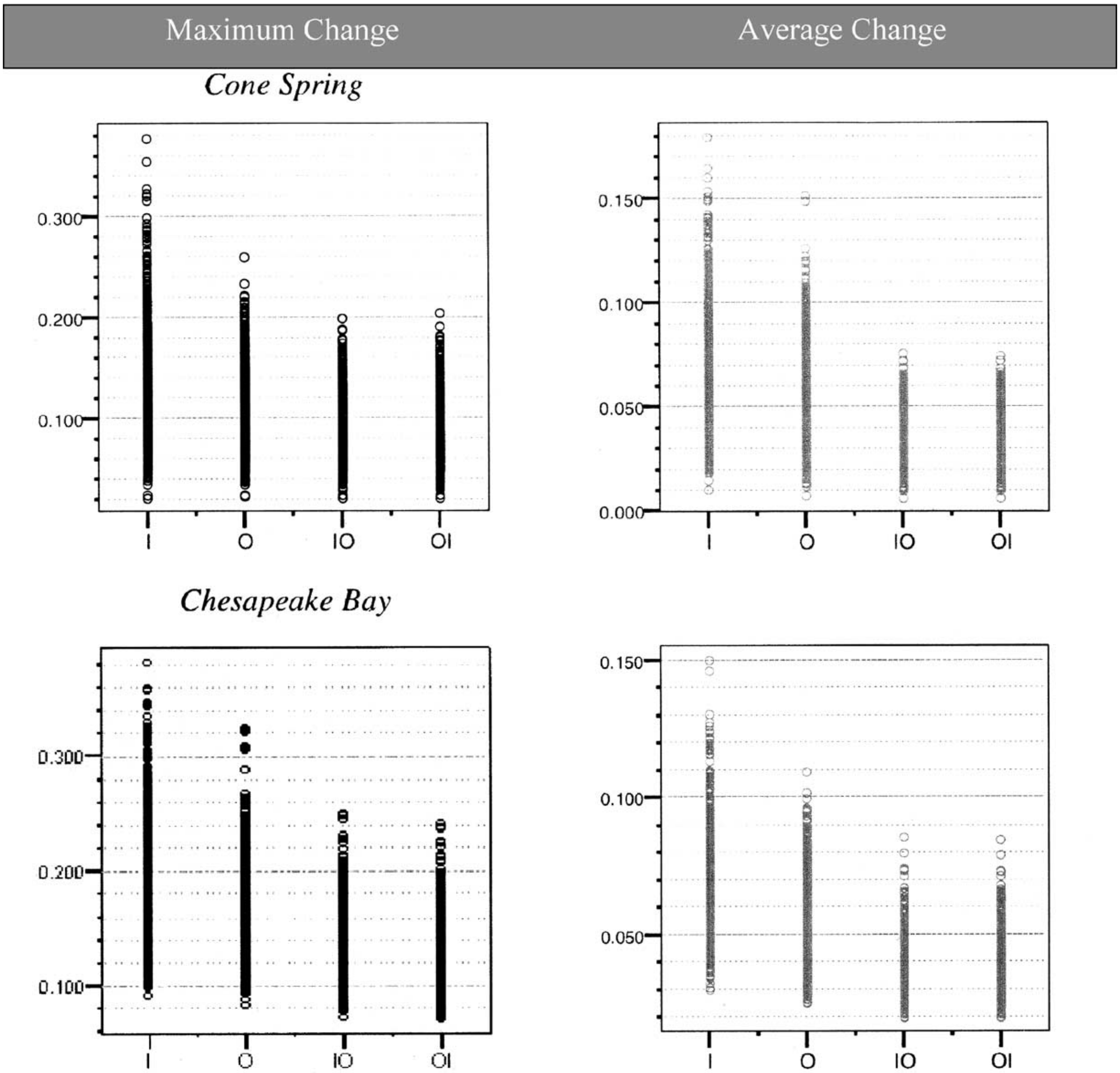

Fig. 2. Maximum and average change observed for extended transfer matrix coefficients caused by the different balancing algorithms. Only the first four algorithms are depicted.

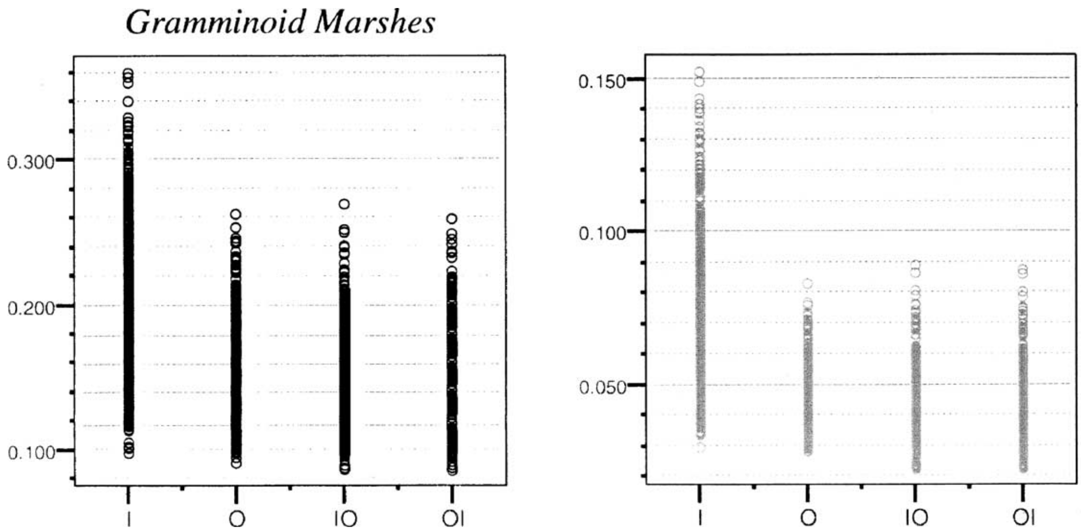

Fig. 2. (*Continued*).

certain period. It is the sum of all coefficients of the extended transfer matrix T^*:

$$\mathrm{TST} = \sum_{i=1}^{N+3}\sum_{j=1}^{N+3} T_{ij}^{*\mathrm{Bal}}$$

(b) Ascendency (Ulanowicz, 1995): This index is derived from Shannon's information theory and is a measure of ecosystem's organization. It is obtained by subtracting the actual ecosystem's information entropy from the maximum possible entropy for the system:

$$A = \sum_{i=1}^{N+3}\sum_{j=1}^{N+3} T_{ij}^{*\mathrm{Bal}} \times \log_2 \frac{T_{ij}^{*\mathrm{Bal}}\,\mathrm{TST}}{\left(\sum_{k=1}^{N+3} T_{ik}^{*\mathrm{Bal}}\right)\left(\sum_{k=1}^{N+3} T_{kj}^{*\mathrm{Bal}}\right)}$$

These two indices have been selected because size and organisation are the two most fundamental attributes that describe an ecosystem. In particular, ascendency combines in a unique measure both size and organisation, but since its value is deeply affected by the magnitude of TST it has been decided to use both indices for comparing the effects of the six balancing procedures.

For every ecosystem, all randomly unbalanced matrices were treated using the six algorithms; then TST and ascendency were computed for each balanced matrix. This computation yielded 1000 values for TST and ascendency for a given procedure. To compare the outcomes of the six algorithms, all the values for TST and ascendency were divided by their original counterparts calculated from the initial randomly unbalanced matrices.

Distributions of values for TST and ascendency are graphically summarised in Fig. 3, according to the box-plot method (Tuckey, 1977). All the box-plots have been obtained with SPSS for Windows (Darren and Mallery, 2000). Graphs A, B and C refer to Cone Spring, Chesapeake Bay and Gramminoid Marshes, respectively.

4. Discussion

The analysis of coefficients variation points out that balancing can introduce noticeable distortion in flow values, and that the six algorithms may have different outcomes in this respect. Considering maximum change (Fig. 2, left column), the input-based approach can produce variations up to 30–40% of the original flow values in all three networks. The magnitude of maximum change decreases using the output-based

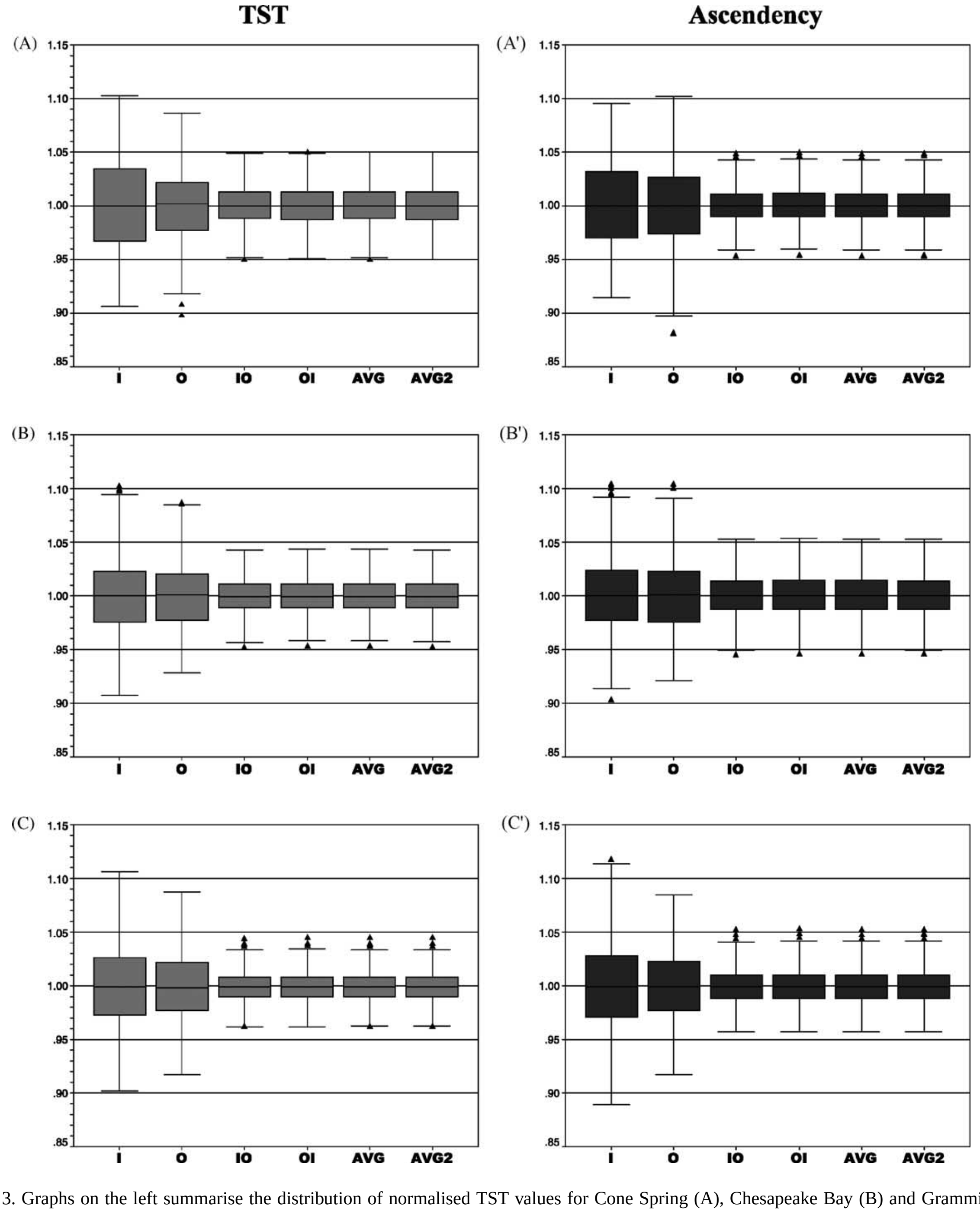

Fig. 3. Graphs on the left summarise the distribution of normalised TST values for Cone Spring (A), Chesapeake Bay (B) and Gramminoid Marshes (C). The distributions of normalised Ascendencies are given in graphs A′, B′ and C′. In each graph there are six distributions, produced by the different algorithms used to balance the initial randomly unbalanced matrices. Solid boxes contain 50% of the values of the distribution; the median of the distribution is marked by a line within the box. The box length is the interquartile range and the two vertical lines (whiskers) outside the box extend to include the smallest and the largest observations within 1.5 times the interquartile range.

approach, while better outcomes are obtained when the four methods proposed here are applied. However, when the number of compartment raises, the output-based approach performs as well as the new algorithms (Fig. 2, Gramminoid Marshes).

On average (Fig. 2, right column) coefficients change less when balancing procedure is performed by the new algorithms, but for large networks the output-based approach may give a better performance. The input-based approach, still, shows the worst result: on average, coefficients may change up to 20%, in comparison with average variations less than 10% of original flow values that are obtained by the other methods.

To summarise, distortion in flow values following balancing procedure seems to be the highest whenever the input-based approach is used. In fact this method cause the greatest variation in single coefficients and has the highest average change. The output-based approach performs less well than IO and OI algorithms, but could be preferred when system size increases.

Consequences of balancing procedures on network analysis can be assessed by considering variations in TST and ascendency indices. Because matrix coefficients, representing flow values in the network, have been randomly changed by ±10%, the distributions $TST^{original}/TST^{UnBal}$ and $ASC^{original}/ASC^{UnBal}$ are expected to vary in a range $[0.9, \ldots, 1.1]$ with average equal to 1 (the original value for the index is the one calculated by the authors in their analysis, whereas unbalanced values are those obtained after unbalancing the extended transfer matrix). What happens, instead, is something different, as one can see in Fig. 3: values for both indices (TST^{Bal}/TST^{UnBal}, ASC^{Bal}/ASC^{UnBal}) distributes over wider ranges. This highlights the fact that balancing procedures may affect system level indices in network analysis.

The form of the distributions (Fig. 3) identifies two groups within which algorithms show similar behaviour. For both ascendency and TST, input-based and output-based approach give rise to more dispersed values: 50% of the observations lie in a range that is always greater than ±0.01 away from the median. The whiskers extend well beyond the range ±0.05 from the median. In this group the output-based approach seems to perform better (narrower distribution). For network of intermediate size, such as the Chesapeake Bay ecosystem, the two algorithms show smaller difference in both the interquartile range and the whiskers. As for ascendency values, reducing the size of the network leads to narrower distribution for the input-based approach. Both the above conclusion are presented with scepticism because only one model for each size class has been investigated. To make these conclusions general, more model for each size class need to be examined, an issue that will be discussed elsewhere.

In comparison with the effects produced on flow values, the output-based approach never performs as well as the group of derived algorithms. Using these latter methods the values for both TST and ascendency lie in a range not larger than ±0.05 with respect to the median, although a little larger interval characterises ascendency values for the Chesapeake Bay. The limits of the interquartile range vary between ±0.01 and ±0.015. For both indices, the greater the size of the network the narrower the distribution of values obtained by the four derived algorithms.

In all cases the number of statistical outliers is very small (3 over 1000 values at most) and they do not add anything significant to the above conclusions.

5. Conclusions

Steady state condition can be imposed to energy flow networks via balancing algorithms, but this procedure inevitably alter flow values. This effect may have important consequences. In particular, in this paper we analysed the effect of balancing algorithms on system level indices such as TST and ascendency. Algorithms presently in use, input-based and output-based procedures, by producing more disperse distributions, determine greater variability in the value of these indices, with possible distortion. As these indices measure ecosystem growth and development and they are used to assess ecosystem health and integrity, an excess distortion could lead to misleading assessment of ecosystem attributes, with potential consequences on their management. Ecosystem comparison is necessary to highlight patterns of ecosystem behaviours (i.e. relationship between maturity and cycling, controlling factors, trajectories of regional ecosystem change, effects of human disturbance) that are crucial to develop approaches for evaluating outcomes of alternative future land use, management and policy; in this framework effects of balancing procedures may make difficult to ascertain whether differences observed in system level indices are significant. In this case variation range due to balancing procedure can be seen as a point of reference for comparison: meaningful differences must exceed this range. This remark calls attention to the fact that the statistical properties of the ascendency have not been deeply investigated yet.

The comparative study presented here highlights that balancing procedures presently in use alter network attributes more severely than the other techniques introduced here as derived algorithms, no matter what the network size is. Moreover, these latter methods show more homogeneous performances, so that using different algorithms of this group may not have significant consequences on the results.

References

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

Christensen, V., Pauly, D., 1992. ECOPATH II—a software for balancing steady state ecosystem models and calculating network characteristics. Ecol. Model. 61, 169–185.

Darren, G., Mallery, P., 2000. SPSS for Windows Step by Step, 3rd ed. Boston, Allyn & Bacon.

Ducklow, H., Fasham, M.J.R., Vézina, A., 1989. Derivation and analysis of flows networks for open ocean plankton systems. In: Wulff, F., Field, J.G., Mann, K.H. (Eds.), Network Analysis in Marine Ecology: Methods and Applications. Lecture Notes on Coastal and Estuarine Studies. Springer, New York, pp. 159–205.

Heymans, J.J., Ulanowicz, R.E., Bondavalli, C., 2002. Network analysis of the South Florida Everglades Gramminoid Marshes and comparison with nearby Cypress ecosystems. Ecol. Model. 149, 5–23.

Hirata, H., Ulanowicz, R.E., 1984. Information theoretical analysis of ecological networks. Int. J. Syst. Sci. 15, 261–270.

Nielsen, S.N., Ulanowicz, R.E., 2000. On the consistency between thermodynamical and network approaches to ecosystems. Ecol. Model. 132, 23–31.

Parker, R.L., 1977. Understanding the inverse theory. Annu. Rev. Earth Planet. Sci. 5, 35–64.

Polovina, J.J., 1984. Model of a coral reef ecosystems I. The ECOPATH model and its application to French Frigate Shoals. Coral Reefs 3 (1), 1–11.

Savenkoff, C., Vézina, A.F., Bundy, A., 2001. Inverse analysis of the structure and dynamics the whole Newfoundland—Labrador Shelf ecosystem. Can. Tech. Rep. Fish. Aquat. Sci. 2354, viii, 56.

Tilly, L.J., 1968. The structure and dynamics of Cone Spring. Ecol. Monogr. 38, 169–197.

Tuckey, J.W, 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA, 688 pp.

Ulanowicz, R.E., 1986. Growth and Development: Ecosystem Phenomenology. Springer, New York, 203 pp.

Ulanowicz, R.E., 1989. A generic simulation model for treating incomplete sets of data. In: Wulff, F., Field, J.G., Mann, K.H. (Eds.), Network Analysis in Marine Ecology: Methods and Applications. Lecture Notes on Coastal and Estuarine Studies. Springer, New York, pp. 82–89.

Ulanowicz, R.E., 1995. Network growth and development. In: Patten, B.C., Jørgensen, S.E. (Eds.), Complex Ecology: The Part-Whole Relation in Ecosystems. Englewood Cliffs, pp. 643–655.

Ulanowicz, R.E., Kay, J.J., 1991. A package for the analysis of ecosystem flow networks. Environ. Software 6, 131–142.

Wulff, F., Ulanowicz, R.E., 1989. A comparative anatomy of the Baltic Sea and Chesapeake Bay ecosystem. In: Wulff, F., Field, J.G., Mann, K.H. (Eds.), Network Analysis in Marine Ecology: Methods and Applications. Lecture Notes on Coastal and Estuarine Studies. Springer, New York, pp. 82–89.