Physics Letters A 375 (2011) 801–804

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Analyzing complex networks evolution through Information Theory quantifiers

Laura C. Carpi [a,b], Osvaldo A. Rosso [b,c], Patricia M. Saco [a,d], Martín Gómez Ravetti [e,*]

[a] Civil, Surveying and Environmental Engineering, The University of Newcastle, University Drive, Callaghan NSW 2308, Australia
[b] Departamento de Física, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Belo Horizonte (31270-901), MG, Brazil
[c] Chaos & Biology Group, Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón II, Ciudad Universitaria, 1428 Ciudad de Buenos Aires, Argentina
[d] Departamento de Hidráulica, Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario, Avenida Pellegrini 250, Rosario, Argentina
[e] Departamento de Engenharia de Produção, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Belo Horizonte (31270-901), MG, Brazil

ARTICLE INFO

Article history:
Received 26 September 2010
Received in revised form 14 December 2010
Accepted 15 December 2010
Available online 21 December 2010
Communicated by C.R. Doering

Keywords:
Complex networks
Jensen–Shannon divergence
Statistical complexity

ABSTRACT

A methodology to analyze dynamical changes in complex networks based on Information Theory quantifiers is proposed. The square root of the Jensen–Shannon divergence, a measure of dissimilarity between two probability distributions, and the MPR Statistical Complexity are used to quantify states in the network evolution process. Three cases are analyzed, the Watts–Strogatz model, a gene network during the progression of Alzheimer's disease and a climate network for the Tropical Pacific region to study the El Niño/Southern Oscillation (ENSO) dynamic. We find that the proposed quantifiers are able not only to capture changes in the dynamics of the processes but also to quantify and compare states in their evolution.

1. Introduction

There are a huge number of dynamical systems that can be modeled through networks, and that is why the importance of having a proper characterization methodology able to capture either their structural and dynamical properties. Until now, Information Theory quantifiers have been poorly explored as tools to study complex networks behavior, only a small number of works have been reported using the concept of entropy [1,2] and complexity [3]. In this work, we propose the use of complexity measures and the square root of the Jensen–Shannon divergence to characterize the evolution of networks by means of their degree distribution. Our work is motivated by two real applications, the analysis of the evolution of the El Niño/Southern Oscillation (ENSO) phenomenon and the analysis of the gene-network progression during the development of Alzheimer's disease. Both cases can be modeled through networks. In the first case, coordinates in a gridded dataset are considered nodes, and edges are defined by correlations between pairs of data points. In a similar way, a gene network will consider each gene as a node and links will be created depending on the gene-expression correlation values. Those types of networks are well known to possess complex network attributes [4–6] and can be modeled with a fixed number of nodes during their evolution process (grid points and genes considered). To first test our methodology we use the Watts–Strogatz (WS) model [7].

The degree of a node in a network is the number of edges incident to it. As not all nodes have the same number of edges, the spread in the number of edges a node has, is characterized by a distribution function $P(k)$, which gives the probability that a randomly selected node has exactly k edges [8]. Then, $P(k)$ is defined as n_k/n, being "n" the total number of nodes in the network.

The WS model starts with a regular network, and during subsequent steps, each edge can be rewired to a randomly chosen vertex with a given probability p. By using this model, several intermediate states from the initial regular network (all nodes with k incident edges) to a random network are obtained. At each step of the process, the probability p is increased and the network walks towards a random graph. The probability of not rewiring a specific edge of the regular lattice is then given by $(1-p)+p.(k+1)/n$, that is, the sum of the probability of not allowing that edge to change plus the probability that the chosen target node is already linked to the edge (this change is therefore prohibited). This result can be generalized for the complete network: the probability of a regular network not changing its structure in one step of the process is $P=((1-p)+p.(k+1)/n)^{nk/2}$. The binomial condition of the process should lead to a Poisson distribution, but the fact that the model considers an on going evolution of the network, that is, changes in the network remain for the following steps, alters the shape of the final average degree probability distribution.

An alternative model, which we call herein modified WS (mWS) is also considered. The difference between the WS and mWS mod-

* Corresponding author. Tel.: +55 31 3409 3481; fax: +55 31 3409 4883.
E-mail addresses: Laura.Carpi@studentmail.newcastle.edu.au (L.C. Carpi), rosso@fisica.ufmg.br (O.A. Rosso), Patricia.Saco@newcastle.edu.au (P.M. Saco), martin.ravetti@dep.ufmg.br (M.G. Ravetti).

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2010.12.038

els, is that on the latter the changes at each step are not retained, that means that every step of the process starts from the regular lattice. It is interesting noticing that, the modification of the model does not alter the small-world properties of the process. However the average degree distribution converges to a true Poisson, with parameter k for the extreme case of $p = 1$.

In the experiments here presented, the average degree distribution is obtained for all the intermediate network structures of the WS and mWS models. The Shannon entropy (S), the square root of Jensen–Shannon divergence ($\mathcal{J}^{1/2}$) and the MPR Statistical Complexity ($\mathcal{C}$) are computed for all intermediate structures.

2. Information Theory quantifiers

2.1. Shannon entropy

Shannon entropy measures the degree of heterogeneity of the network [1]. Its zero value corresponds to the state of having complete knowledge of the process. In our particular case, it means that the average degree of the network is known (regular lattice). On the other hand, the maximum entropy value occurs when our knowledge of the system is minimum (uniform random network). The entropy of the degree distribution $P(k)$ can be described as $S = -\sum_k P(k) \ln P(k)$ [9].

2.2. Jensen–Shannon divergence

The Jensen–Shannon divergence is a measure of the dissimilarity between two probability distributions. It presents ultrametricity properties: (1) positive values, (2) symmetry and (3) a zero value for equal probability distributions. The only missing property to obtain a metric is the triangle inequality that can be achieved by taking its square root [10,11]. As we are interested in quantify and compare states in a network evolution, we use $\mathcal{J}^{1/2}$.

To obtain a characterization of the evolution of the WS model using the $\mathcal{J}^{1/2}$ we need to choose a reference Probability Distribution Function (PDF). Three alternatives are here analyzed: a Poisson distribution (P_o with $\lambda = k$), the uniform distribution P_e, and the PDF corresponding to the regular lattice P_r. Both P_e and P_r are extreme and invariant cases, P_e corresponds to the asymptotic random network stage and P_r to the initial stage of the process. Poisson was chosen as it is the PDF reached by the Erdös–Rényi random model [12] and it has been used by many authors as possible average degree distribution. The use of the regular lattice is interesting as it is the only extreme case that can be achieved by a single real network. On the other hand, for practical purposes the uniform PDF is appealing as its values are independent of the number of edges of the network.

We propose in this work the use of the $\mathcal{J}^{1/2}[P, P_{ref}]$ as a form of characterization of small-world networks ($P_{ref} = \{P_o, P_e, P_r\}$). The $\mathcal{J}$ is given by,

$$\mathcal{J}[P, P_{ref}] = S\big[(P + P_{ref})/2\big] - S[P]/2 - S[P_{ref}]/2$$

Figs. 1 and 2 display the $\mathcal{J}^{1/2}[P, P_{ref}]$ values for the WS and the mWS models. These figures show that the average path length (L) and the clustering coefficient (C) have similar behavior for both models and their values are in agreement with those presented in [7]. As $\mathcal{J}^{1/2}[P, P_r]$ and $\mathcal{J}^{1/2}[P, P_e]$ are extreme cases in the progression, they present unique values for each probability p independently of the model considered. The $\mathcal{J}^{1/2}[P, P_o]$ only has unique values when using the mWS model. In the WS model, what it seems to be a change of tendency shows in fact, the average degree distribution of the network getting closer to a Poisson distribution until $p = 0.5$, and diverging again from it after $p = 0.5$, confirming the results by Newman et al. [13]. This occurs because

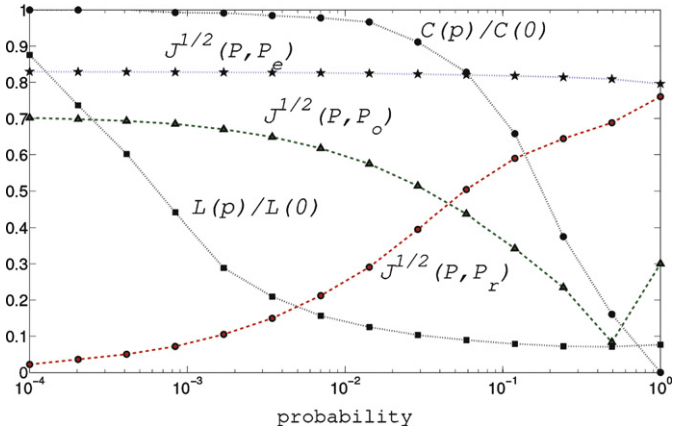

Fig. 1. Normalized characteristic path length ($L(p)/L(0)$), normalized clustering coefficient ($C(p)/C(0)$), and square root of the Jensen–Shannon divergence $\mathcal{J}^{1/2}[P, P_{ref}]$ for the WS model. The initial stage is a regular lattice of 1000 nodes, each one with degree 10. For each value of probability p (horizontal axis), 50 trials were averaged to compute the degree distribution. $\mathcal{J}^{1/2}[P, P_{ref}]$ is obtained from the average degree distribution using $P_{ref} = \{P_o, P_e, P_r\}$.

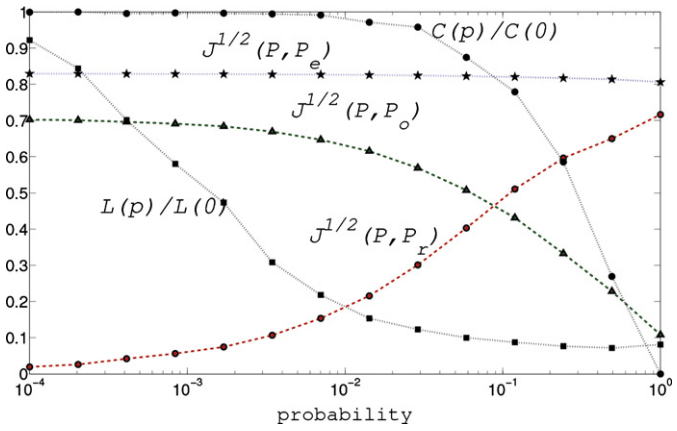

Fig. 2. Same as Fig. 1 but considering the mWS model.

the WS model retains all the changes of previous steps. The experiment here performed contemplates networks with 1000 nodes and $k = 10$ as discussed in the seminal article of Watts and Strogatz [7].

2.3. Statistical complexity

The complexity of dynamical systems has no universal definition, however statistical complexity can be understood as a measure that captures not only the system's randomness but also considers its physical components (structural correlations) [14]. In this work, we use a statistical complexity measure that is obtained by the product of the normalized Shannon entropy and the normalized Jensen–Shannon divergence, namely, MPR [15]. An important property of this measure is that it presents "zero" value for the extreme cases given by the regular lattice and the uniform distribution.

The statistical complexity is defined by [15] as $\mathcal{C}_e = \mathcal{C}[P, P_e] = H[P] \cdot Q[P, P_e]$ in which $H[P] = S[P]/S_{\max}$ is the normalized Shannon entropy ($0 \leqslant H \leqslant 1$) and $Q[P, P_e] = Q_{\max} \cdot \mathcal{J}[P, P_e]$ is the normalized disequilibrium ($0 \leqslant Q \leqslant 1$). The disequilibrium represents the distance between the PDF of the current state of the system and the PDF of reference (the uniform PDF in this case). We can also use the PDF of the regular lattice as reference, $\mathcal{C}_r = \mathcal{C}[P, P_r] = H[P] \cdot (1 - Q[P, P_r])$, in which $Q[P, P_r] = Q_{\max} \cdot \mathcal{J}[P, P_r]$. Note that the corresponding PDFs P_r and P_e are

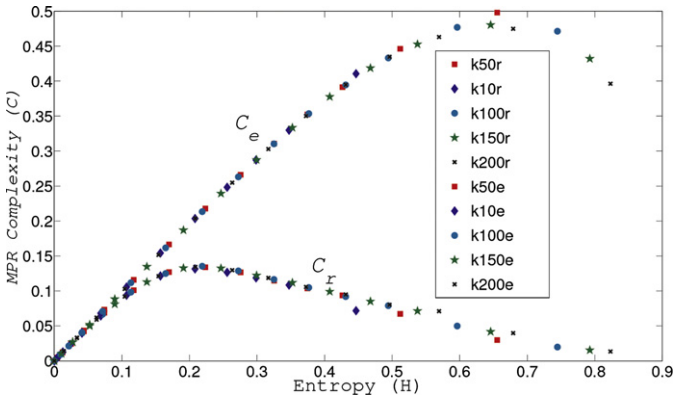

Fig. 3. MPR Statistical Complexity versus normalized Shannon entropy quantifiers are shown for intermediate states of the WS model for varying values of k (see different colors and symbols in the inset notation), considering both references, the PDF of a regular network (k_r) and a uniform distribution (k_e). MPR (C) is estimated using a PDF of a regular lattice $\mathcal{C}_r$, and a uniform distribution $\mathcal{C}_e$, as the PDF of reference. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)

extreme cases, having the same disequilibrium constant, given by $Q_{\max} = \mathcal{J}^{-1}[P_e, P_r]$.

Fig. 3 shows that the complexity values for the intermediate states of the WS model (and for all k) define a single curve. The maximum complexity value seems to be independent of the number of edges (k). However, networks with larger number of edges (larger k) achieve the maximum complexity at earlier stages of the evolution process (lower p). This maximum complexity value occurs when normalized entropy is equal to normalized disequilibrium and reflects some sort of network structure similarity.

3. Application to real networks

3.1. Gene networks – Alzheimer's disease

In [16], a group of patients in different stages of Alzheimer's disease (AD) was categorized into four groups, not AD (*notAD*), incipient AD (*incAD*), moderate AD (*modAD*) and severe AD (*sevAD*). The diagnosis was performed *post mortem* by analyzing the latest Mini-mental state exam (MMSE) and counting neurofibrillary tangles (NFT) from hippocampus' tissue samples. In a recent publication [17], and using the same data set, a thorough analysis of genes correlated with cognitive decline with AD was performed. From a total of 22,000 gene expression values and by using a bioinformatics approach, a list of genes presenting high-consensus with established markers of progression of AD was identified [17]. In this application we are interested in analyzing the network of genes and its evolution during the development of the disease.

Arbitrarily fixing the number of genes to 500, we consider four networks, one for each stage of the disease. Each node represents a gene expression vector with dimensionality equal to the number of samples (patients). The network topology is defined by computing the Spearman correlation coefficient between all pair of genes. A link connecting two nodes is used when their correlation value is greater than a given threshold (0.9). For the *notAD* network 781 edges are considered, 1312 edges for *incAD*, 966 for *modAD* and 953 for *sevAD*. For this application, the uniform distribution was chosen as reference PDF to compute $\mathcal{J}^{1/2}$ because of its independence to the number of edges, allowing a common point of reference. The networks show a noticeable behavior, together with the development of the disease, they seem to evolve toward a random network. This situation occurs until a moderate AD condition is reached (Fig. 4), afterwards the severe AD condition appears to be more regular. These results are highly consistent with AD physiology and morphology evolution. The hippocampus is one of the first regions affected by the disease; in early AD the hippocampus can loss more than 38% of its volume [18]. At a late stage of the disease there is a severe hippocampal shrinkage; the patient has already lost most of its cognitive functions, and as expected a more regular network is found, the correlated genes are widespread in the hippocampal region. The complexity values also reveal an interesting behavior. A peak is presented for incipient AD networks. Revealing that, just from the beginning of the disease the gene network seems to reflect the abnormal condition.

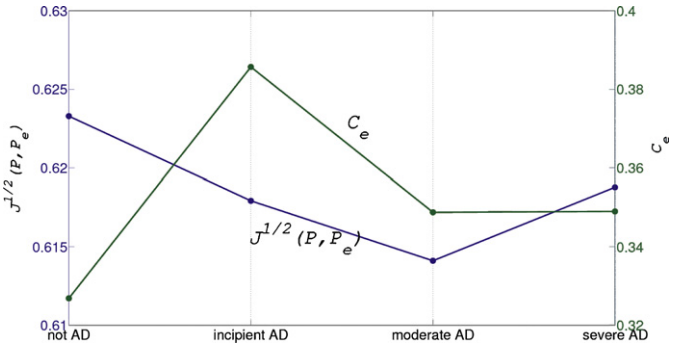

Fig. 4. Square root of the Jensen–Shannon divergence using as reference a uniform distribution $\mathcal{J}^{1/2}(P, P_e)$ and MPR Statistical Complexity for different stages of Alzheimer's disease. Results are obtained from the analysis a 500 gene network.

3.2. Climate networks – ENSO

The use of networks as a tool for the study of climate processes is recent [4,19–24]. In climate networks, nodes are identified with spatial grid points of a climate data set and edges are added between pairs of nodes depending on the degree of statistical interdependence between the corresponding pairs of anomaly time series taken from the climate data set [24].

For this analysis, we consider the Tropical Pacific ($120E°$–$70W°$, $20N°$–$20S°$) monthly averaged surface air temperature (SAT) reanalysis data set [25], represented by 1156 grid points (nodes). We chose this data to maintain consistency with previous works [4,23], and because it captures the dynamics on the interface between ocean and atmosphere due to heat exchange [24]. This dataset is therefore appropriate to investigate the evolution of the El Niño/Southern Oscillation (ENSO). The ENSO cycle takes from 3 to 4 years (average), its warm and cold phases are called El Niño and La Niña, respectively. The horizontal line represents the average value $\mathcal{J}^{1/2} = 0.3324$ over the 62 years.

The network topology is obtained by computing the Spearman correlation coefficient between all pairs of nodes. An edge connecting two nodes is created when their correlation value exceeds a given threshold. The evolution of the network topology is captured by considering annual sliding windows without overlapping. Temporal changes are then analyzed by computing the square root of the Jensen–Shannon divergence and the MPR Statistical Complexity using the degree distribution for each network.

The determination of a good threshold depends not only on the characteristic to be analyzed but also on the size of the considered network. For this particular case, the considered region is small and highly connected. A high value threshold will work as a filter, pruning not useful information and revealing small differences in the network topology. For that reason we chose a threshold of 0.9. An interesting discussion about the choice of a threshold in climate networks can be found in [24,26].

Fig. 5 shows that the $\mathcal{J}^{1/2}$ captures the ENSO cyclic behavior. Though, the degree distribution is characterized by maintaining approximately the same distance to the reference PDF (uniform

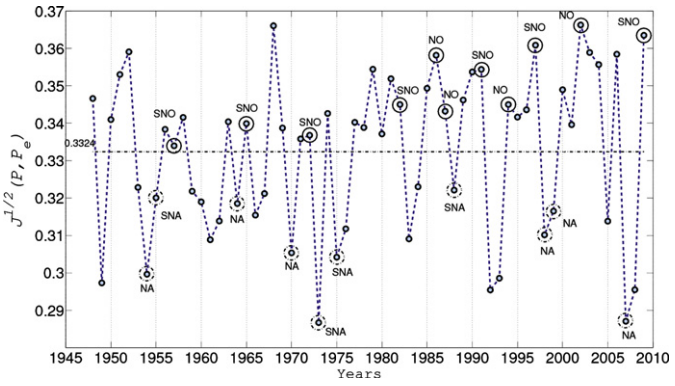

Fig. 5. Square root of the Jensen–Shannon divergence using as reference a uniform distribution $\mathcal{J}^{1/2}(P, P_e)$ for the Tropical Pacific region (1156 nodes). SNO and SNA represent strong events for El Niño and La Niña, NO and NA represent moderate events for El Niño and La Niña respectively.

distribution) throughout the studied period. All of La Niña and El Niño events display $\mathcal{J}^{1/2}$ values below and above average, respectively. The average value for the 62 years is 0.3324 (central line). Therefore, El Niño appears to be closer to a regular network, and it is therefore less efficient in transferring information because of its higher characteristic path length. The concept of efficiency is particularly interesting for climate analysis as it can be used as a measure to investigate the effects of local events at global scale. Efficiency can be estimated as the inverse of the characteristic path length [27,28]. As random networks have lower characteristic path length values then, if the $J^{1/2}[P, P_{ref}]$ is used considering a uniform PDF as the reference distribution, lower values of the $J^{1/2}[P, P_e]$ will indicate a more efficient network. The contrary happens when the chosen PDF of reference corresponds to a regular lattice (i.e., larger values for more efficient networks). This result is consistent with previous findings [21], but unlike those works here we use a quantifier that "captures" and "measures" the continuous temporal changes of the network due to ENSO.

The evolution of the Statistical Complexity measure for the case of the ENSO phenomena shows a very similar behavior to that shown in Fig. 5 for the $\mathcal{J}^{1/2}$. The reason for this is that the entropy values for the Tropical Pacific network do not change significantly with time. In this case, changes in the Jensen–Shannon divergence values ($\mathcal{J}$ values) drive the observed changes in the MPR Statistical Complexity and therefore the MPR Statistical Complexity evolution graph has a similar shape as the one presented for the $\mathcal{J}^{1/2}$.

4. Final remarks

We have proved in this work how the use of Information Theory quantifiers can characterize the progression of a system represented by a network. One of the advantages of using the $\mathcal{J}^{1/2}$ and $\mathcal{C}$ is that the existence of disconnected nodes does not interfere in their computation. Depending on the ratio between the number of nodes and neighbors (n/k), the network may become disconnected during the process, this is common on real world applications. When this happens the average path length, must be treated [29]. Another important feature of the $\mathcal{J}^{1/2}$ is its metric property. It can be used, not only as a tool to measure how far the network structure is from the chosen reference, but also to compare different states during its evolution, or to measure the distance between two different network structures.

In the applications here presented, the $\mathcal{J}^{1/2}$ and $\mathcal{C}$ were capable to reflect the structural changes during the evolution of these processes; describing the progression from a regular to a random network in the Watts–Strogatz model, highlighting the most important structural changes in the progression of Alzheimer's disease, and discriminating structures related to El Niño and La Niña events when studying the Tropical Pacific region. We show that the use of these Information Theory quantifiers is a powerful tool to study the evolution of a wide range of natural phenomena that can be modeled through networks.

Acknowledgements

This research has been supported by a scholarship from The University of Newcastle awarded to Laura C. Carpi. O.A. Rosso acknowledges support from CONICET, Argentina and CAPES, Brazil, and M.G. Ravetti from FAPEMIG and CNPq, Brazil.

References

[1] B. Wang, H. Tang, C. Guo, Z. Xiu, Physica A 363 (2006) 591.
[2] L. Demetrius, T. Manke, Physica A 346 (2005) 682.
[3] T. Wilhelm, J. Hollunder, Physica A 385 (2007) 385.
[4] A.A. Tsonis, P.J. Roebber, Physica A 333 (2004) 497.
[5] C.J. Stam, B.F. Jones, G. Nolte, M. Breakspear, Ph. Scheltens, Cerebral Cortex. 4 (2007) 92.
[6] F. Barrenas, S. Chavali, P. Holme, R. Mobini, M. Benson, PLoS ONE 4 (2009) e8090.
[7] D.J. Watts, S.H. Strogatz, Nature 393 (1998) 440.
[8] R. Albert, A. Barabási, Rev. Modern Phys. 74 (2002) 47.
[9] L.F. da Costa, F.A. Rodrigues, G. Travieso, P.R. Villas Boas, Adv. Phys. 56 (2007) 167.
[10] F. Österreicher, I. Vajda, Ann. Inst. Statist. Math. 55 (2003) 639.
[11] D.M. Endres, J.E. Schindelin, IEEE Trans. Inform. Theory 49 (2003) 1858.
[12] P. Erdös, A. Rényi, Publ. Math. Debrecen 6 (1959) 290.
[13] M.E.J. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. E 64 (2001) 026118.
[14] J.P. Crutchfield, K. Young, Phys. Rev. Lett. 63 (1989) 105.
[15] O.A. Rosso, H.A. Larrondo, M.T. Martin, A. Plastino, M.A. Fuentes, Phys. Rev. Lett. 99 (2007) 154102.
[16] E.M. Blalock, J.W. Geddes, K.C. Chen, N.M. Porter, W.R. Markesbery, P.W. Landfield, PNAS 101 (2004) 2173.
[17] M. Gómez Ravetti, O.A. Rosso, R. Berretta, P. Moscato, PLoS ONE 5 (2010) e10153.
[18] L. Wang, J.S. Swank, I.E. Glick, M.H. Gado, M.I. Miller, J.C. Morris, J.G. Csernansky, NeuroImage 20 (2003) 667.
[19] A.A. Tsonis, K.L. Swanson, P.J. Roebber, Bull. Amer. Meteor. Soc. 87 (2006) 585.
[20] A.A. Tsonis, K.L. Swanson, G. Wang, J. Clim. 21 (2008) 2990.
[21] A.A. Tsonis, K.L. Swanson, Phys. Rev. Lett. 100 (2008) 228502.
[22] A. Gozolchiani, K. Yamasaki, O. Gazit, S. Havlin, Europhys. Lett. 83 (2008) 28005.
[23] K. Yamasaki, A. Gozolchiani, S. Havlin, Phys. Rev. Lett. 100 (2008) 228501.
[24] J.F. Donges, Y. Zou, N. Marwan, J. Kurths, Eur. Phys. J. Spec. Top. 174 (2009) 157.
[25] E. Kalnay, et al., Bull. Amer. Meteor. Soc. 77 (1996) 437.
[26] K. Steinhaeuser, N.V. Chawla, A.R. Ganguly, in: Proceedings of the 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2009.
[27] V. Latora, M. Marchiori, Phys. Rev. Lett. 87 (2001) 198701.
[28] V. Latora, M. Marchiori, Eur. Phys. J. B 32 (2003) 249.
[29] M.E.J. Newman, D.J. Watts, Phys. Lett. A 263 (1999) 341.