Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Physica A 333 (2004) 497–504
PHYSICA A
www.elsevier.com/locate/physa

The architecture of the climate network

A.A. Tsonis*, P.J. Roebber

Department of Mathematical Sciences, Atmospheric Sciences Group, University of Wisconsin-Milwaukee, Milwaukee, WI 53201-0413, USA

Received 7 May 2003

Abstract

We consider climate as a network of many dynamical systems and apply ideas from graph theory to a global data set to study its collective behavior. We find that the network has properties of 'small-world' networks (Nature 393 (1999) 440). A detailed investigation of the coupling architecture of this network reveals that the overall dynamics emerge from the interaction of two interweaved subnetworks. One subnetwork operates in the tropics and the other at higher latitudes with the equatorial one acting as an agent that establishes links between the two hemispheres. Both subsystems are 'small-world' networks, but there are distinct differences between the two subsystems. The tropical one is an almost fully connected network, whereas the mid-latitude one is more like a scale-free network characterized by dominant super nodes, and multifractal properties. This unique architecture may lead to new insights not only about the dynamics of the climate system but of other spatially extended complex systems with a large number of degrees of freedom.

PACS: 89.75.−k; 89.75.Da; 89.75.Fb; 89.75.Hc; 92.60.Bh; 92.70.Gt

Keywords: Small-world networks; Climate; Subsystems; Complex systems; Multifractal networks

1. Introduction

A network is a set of points (nodes) that interact. In the past, the study of networks was restricted to either regular (ordered) networks, where each node has the same number of links connecting it in a specific way to a small number of neighboring nodes, or to random networks, where each node is haphazardly connected to a few nodes that can be anywhere in the network. Regular networks are thus highly clustered,

* Corresponding author. Tel.: +1-414-229-5373; fax: +1-414-229-4907.
E-mail address: aatsonis@uwm.edu (A.A. Tsonis).

doi:10.1016/j.physa.2003.10.045

which means that it takes many steps to go from a node to another node away from its immediate neighborhood. In addition, because of the high degree of local clustering, if some of the connections are removed the network does not become fragmented into disconnected parts. On the contrary, random networks do not exhibit local clustering. Far away nodes can be connected as easily as nearby nodes. In this case information may be transported all over the network much more efficiently than in ordered networks. This is an important property but random networks are also highly unstable. The loss or addition of several links may completely devastate the system [2,3]. Thus, random networks imply efficient information transfer and regular networks imply stability. This dichotomy of networks as either regular or random is undesirable since one could expect that in nature networks should be efficient in processing information and at the same time be stable. Networks that combine both stability and efficient information propagation are called 'small-world' networks and were discovered a few years ago by Watts and Strogatz [1]. 'Small-world' networks exhibit a high degree of local clustering but a small number of long-range connections make them as efficient in transferring information as random networks. Since the original discovery, such networks have been found to pervade biological, social, ecological, and economic systems, the internet, and other systems [1,4–9].

The nodes in a network can be anything (for example, actors that are connected to other actors if they have appeared together in a movie, species that are connected to other species they interact with, individual dynamical systems, etc.). The networks can be either fixed, where the number of nodes and links remains the same, or evolving, where the network grows as more nodes and links are added.

Whatever the type of the network, its underlying topology provides clues about the collective dynamics of the network. The structural properties of networks are delineated by the clustering coefficient C and the characteristic path length (or diameter) L of the network. The clustering coefficient is defined as follows: For each node we assume a number of neighbors k satisfying the requirement $n \gg k \gg \ln n$, where n is the number of nodes in the network. Then at most $C_{\max} = k(k-1)/2$ connections can exist between them. We then find the number of actual connections and we calculate $C'_{\mathrm{v}} = C_{\mathrm{v}}/C_{\max}$. The average C'_{v} over all nodes provides C. As such C provides a measure of local "cliqueness". The diameter of the network is defined by the average number of connections needed to connect any two nodes in the network. Graph theory predicts that for random networks $L_{\mathrm{random}} \approx \ln n/\ln k$ and $C_{\mathrm{random}} \approx k/n$ (see Ref. [1]). Thus, for a 'small-world' network we require that $C \gg C_{\mathrm{random}}$ and $L \geqslant L_{\mathrm{random}}$.

2. Data analysis and results

For our analysis we considered the global National Center for Environmental Prediction/Nationall Center for Atmospheric Research (NCEP/NCAR) reanalysis 500 hPa data set [10]. A 500 hPa value indicates the height of the 500 mb pressure level and provides a good representation of the general circulation (wind flow) of the atmosphere. The data used here are arranged on a grid with a resolution of 5° latitude ×5° longitude. This results in 72 points in the east-west direction and 37 points in the north-south direction for a total of $n = 2664$ points. The values are monthly values from 1948 to 1999. From those values we produced anomaly values (actual value minus the climatological average for each month). Thus, for each grid point we have a time series of 624 anomaly values. Each grid point is a node in the network and it is considered to be a dynamical system whose state varies in time in a complex way. In order to define the "connections" between the nodes, we first estimated the correlation coefficient at lag zero (r) between the time series of all possible pairs of nodes [$n(n-1)/2 = 3,547,116$ pairs]. Note that even though r is calculated at zero lag, a "connection" should not be thought of as 'instantaneous'. The fact that the values are monthly introduces a time scale of at least a month to each "connection". In order to avoid spurious high values of r introduced by the presence of a strong annual cycle we considered only the values for December, January and February in each year (i.e., the estimation of the correlation coefficient is based on a sample size of 156). A pair is considered as connected if their correlation $r \geqslant 0.5$. This criterion is based on parametric and non-parametric significance tests. We found that according to the t-test with $N = 156$, a value of $r = 0.5$ is statistically significant at the 99% level. In addition, randomization experiments where the values of the time series of one node are scrambled and then are correlated to the unscrambled values of the time series of the other node indicate that a value of $r = 0.5$ will not arise by chance. More discussion on the correlation coefficient as an indicator of a connection is presented later.

Having decided which nodes are connected, we then estimated C and L. In our calculations we assumed that $k = 48$, which for these calculations effectively reduces the number of nodes in the network from 37×72 to 31×66. This choice guarantees that a random graph with the same specifications (number of nodes, and average links per node) will be connected [11]. For our network we found that $L = 2.7$ and $C = 0.688$. For a random network we estimated that $L_{\text{random}} = 1.97$ and $C_{\text{random}} = 0.023$. These values indicate that indeed $L \geqslant L_{\text{random}}$ and $C \gg C_{\text{random}}$ (a factor of 30). Thus, our network clearly satisfies the 'small-world' network criteria.

An additional analysis was to check the effect of the polar data on the results. In actuality, the polar latitudes have a much smaller length. When the data are projected onto a grid with equal number of grid points at each latitude, a bias is introduced. However, we find that this is not affecting the results significantly. Exclusion of polar regions does not significantly alter the average values of C and L.

The physical interpretation of this result is that the climate system exhibits properties of stable networks and of networks where information is transferred efficiently. In the case of the climate system, 'information' should be regarded as 'fluctuations' from any source (for example, the tropics, El Niño, etc.). The 'small-world' architecture of this climate network allows the system to respond quickly and coherently to fluctuations introduced into the system. Since $L \approx 3$ and each connection is associated with a time scale of one month, the general circulation synchronizes (adjusts) so that fluctuations diffuse over the whole globe in just a few months. This reduces the possibility of prolonged local extremes and provides greater stability for the climate system. The fact that the climate system may be inherently stable may not come as a surprise to some, but it is interesting to find that this stability is the result of long-range connections. From all possible pairs we found that for $r = 0.5$ about 400,000 pairs are connected.

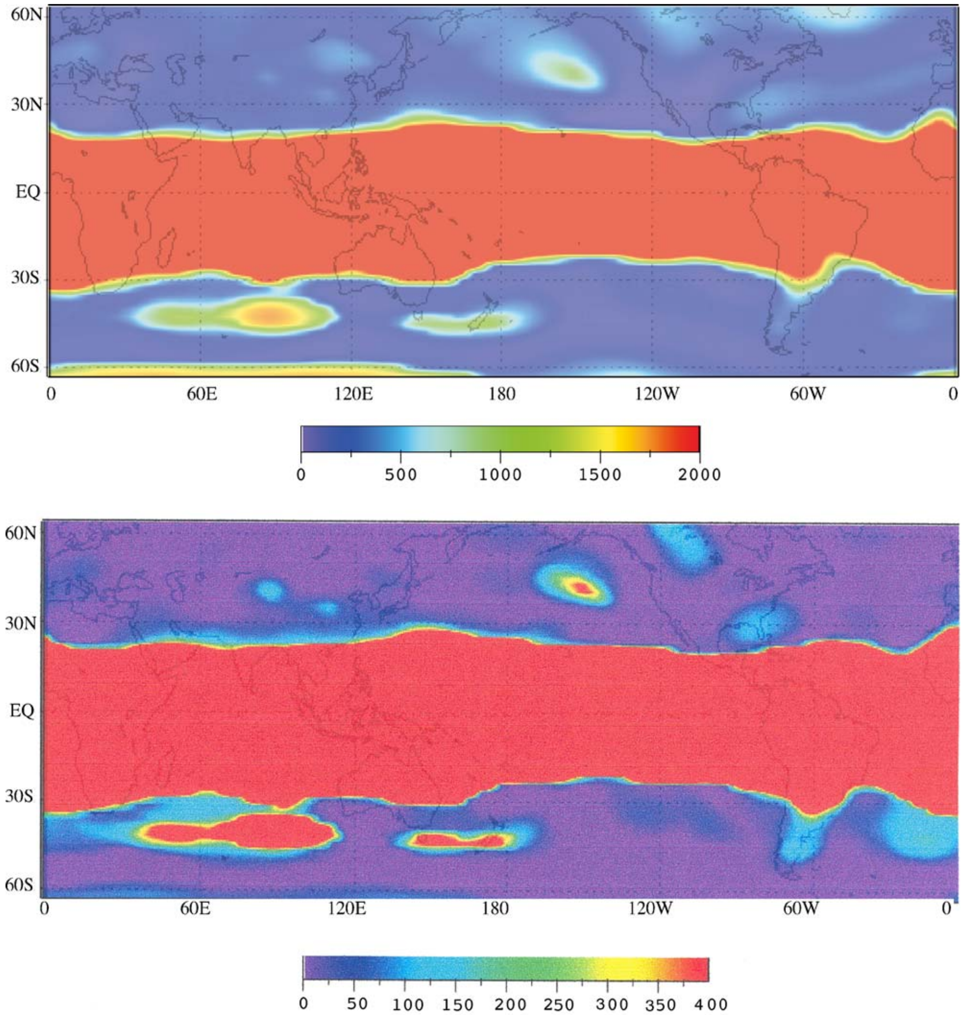

Fig. 1. Top: Total number of links (connections) at each geographic location. Clearly, equatorial nodes exhibit a larger number of connections than midlatitude nodes. Bottom: Same as top but the number of links (connections) with distance of at least 5000 km at each geographic location is shown. The uniformity observed in the tropics indicates that each node possesses the same number of connections regardless the length of the connection. This is not the case in the midlatitudes where certain nodes posses more links than the rest (see text for details).

A detailed study of all these connections and their significance in atmospheric sciences is beyond the scope of this paper and will be reported in the near future in an atmospheric sciences journal. The main concern here is the unique architecture pertaining to this network.

Fig. 1 (top) shows the number of total links (connections) at each geographic location in the range 65N to 65S. Fig. 1 (bottom) shows the number of connections

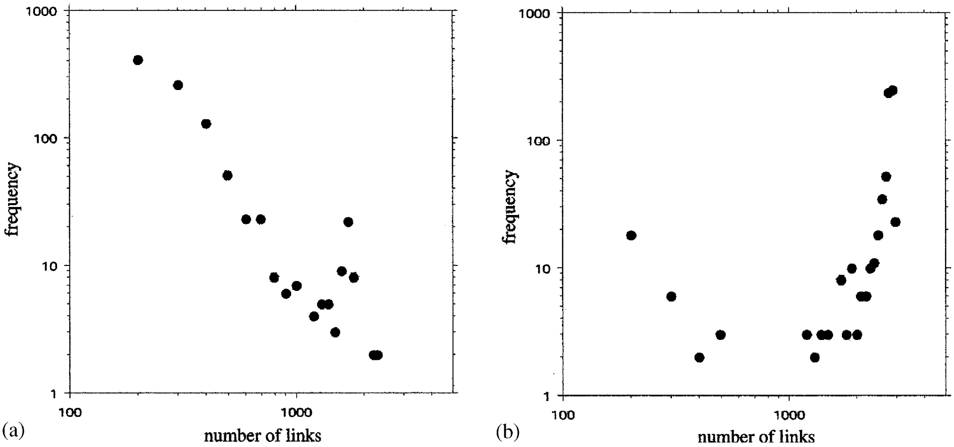

Fig. 2. Degree distribution for the network consisting (a) of mid-latitudes (30N–65N, 30S–65S) grid points; (b) of tropical (20N–20S) grid points.

with distance of at least 5000 km at each geographic location. This figure displays two very interesting features. First, in the tropics it appears that all nodes posses the same number of connections (Fig. 1a). Second, this uniformity in the tropics is present even for long-range links (Fig. 1b). This indicates that each node in the tropics possesses the same number of connections regardless the length of the connection. It thus appears that in the tropics all points tend to be connected to all other points, which will make the tropics more fully-wired than the midlatitudes. In fact, for $r=0.5$, the values of C and L for the network of points in the belt 20N–20S are 0.9 and 2.07, respectively. For a fully wired network one should expect $C=1$ and $L=1$. This tendency is not observed in the midlatitudes where it appears that certain nodes posses more connections than the rest. In the northern hemisphere, for example, we clearly see the presence of three regions where such super nodes exist in North America and Northeast Pacific Ocean. This corresponds to the well-known Pacific-North America (PNA) pattern of the atmospheric circulation [12]. In the southern hemisphere we also see centers corresponding to known southern hemisphere teleconnection patterns [13]. Considering the midlatitudes only (30N–65N and 30S–65S), we find that $L \approx 3.1$ and $C \approx 0.5$.

These differences are clearly delineated in the corresponding to the tropics and to midlatitudes degree distributions. Let p_d denote the fraction of nodes that have d links. The variable d is called the degree and p_d is called the degree distribution. A variety of forms for this distribution have been reported in the past [14]. It includes truncated power-law distributions [15], gaussian distributions [16], power-law distributions (corresponding to scale-free networks) [17], and distributions consisting of two power-laws separated by a cutoff value of d [18,19]. The last two types emerge in certain families of networks that grow in time [18,20]. Fig. 2 shows, on a double logarithmic plot, the distribution of nodes according to how many links they possess (i.e., p_d against d). More specifically, Fig. 2a shows the distribution for all nodes in the mid-latitudes

(30N–65N and 30S–65S) and Fig. 2b shows the corresponding distribution for the tropics (20N–20S). Fig. 2a appears to exhibit a scaling regime similar to those observed in scale-free networks. Scale-free networks are dominated by a few nodes that posses many more links. In fact the slope of this graph is -2.2 in agreement with other scale free networks [14]. In Fig. 2b no such regime is identifiable. The distribution is basically a narrow peak at about 3000 links indicating that most points possess the same large number of connections, a characteristic of almost fully connected networks. It thus appears that the architecture of the coupling between the individual dynamical systems in the network is self-organized into two coupled subnetworks, one operating in the tropics and the other in the higher latitudes. The tropical one is an almost fully connected network, which has a characteristic path length closer to that of a random network, and at the same time has a higher clustering coefficient. The mid-latitude one is more like a scale-free network characterized by dominant super nodes. Whether we consider them as one or two subsystems does not modify the physical interpretation which is that the equatorial network acts as an agent that connects the two hemispheres, thus allowing information to flow between them. This interpretation is consistent with the various suggested mechanisms for inter-hemispheric teleconnections [21–25], with the notion of sub-systems in climate proposed in the late 1980s [26], and with recent studies on synchronized chaos in the climate system [27].

Next we performed some sensitivity experiments. Unlike networks where a connection is solidly defined (think of social networks where a connection is defined if a person knows another person), here a connection can be defined as a function of its strength (or a correlation coefficient threshold). Since the results depend on exactly how a connection is defined, we calculated C and L for various thresholds of the correlation coefficient. Fig. 3 shows C and $L^* = \sqrt{1 - \mathrm{e}^{1-L}}$ as a function of the correlation coefficient, r. The diameter of the network L was transformed to L^* for easier comparison of the results. If the correlation coefficient threshold is considered zero it means that all pairs are connected which makes $C = 1$ and $L = 1$ ($L^* = 0$). If $r = 1$, then no pairs are connected which makes $C = 0$ and $L = \infty$ ($L^* = 1$). In the same figure the two horizontal lines indicate the values of L_{random} (solid line) and $10 \times C_{\mathrm{random}}$ (broken line). These two lines refer to the criteria $L \geqslant L_{\mathrm{random}}$ and $C \gg C_{\mathrm{random}}$. From these two lines we see that the climate system is a 'small-world' network for any value of r between 0.4 and 0.9. For values of $r < 0.4$ small (and possibly non-significant) correlations overwhelm and distort the network and for values of $r > 0.9$ many significant connections are excluded making the network too disconnected. An additional insight from this sensitivity analysis is that it appears that the network exhibits yet another interesting property. We find that as the threshold of r increases the slope of the scale-free midlatitude subnetwork increases. For example if a connection is defined for $r \geqslant 0.4$ the slope of the scaling regime is -1.8. This makes sense because for a lower threshold, more longer connections are allowed. In this case the number of links in the supernodes increases thereby decreasing the slope in a graph like Fig. 2a. It thus appears that the network exhibits multiple scaling regimes. This is akin to multifractals and multifractal networks, which have lately been proposed [28]. The climate network may be one of the first such networks found in natural systems.

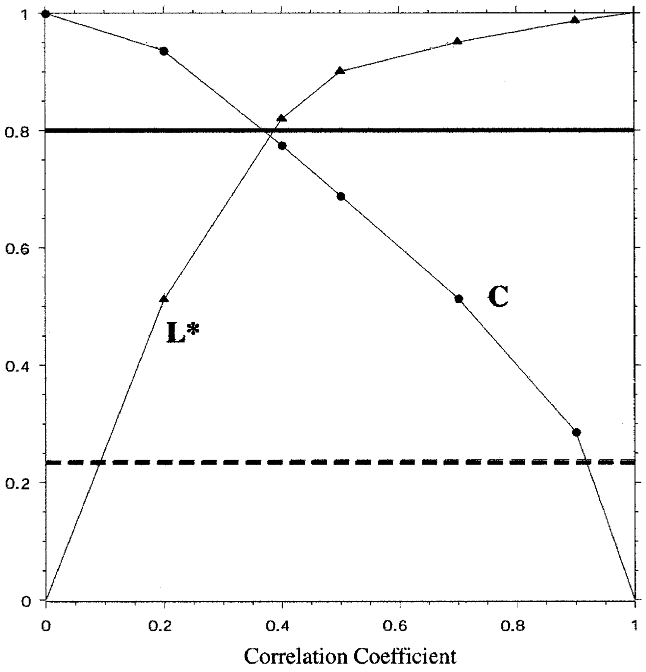

Fig. 3. The clustering coefficient, C, and transformed diameter, L^*, of the climate network as a function of the correlation coefficient. The correlation coefficient is used to determine whether a pair of nodes is connected.

3. Summary

Our analysis offers insights about the evolution and collective dynamics of a complex network consisting of many interacting nonlinear dynamical systems. While it is known that networks of identical nonlinear systems can synchronize their erratic behavior [29,30], not much is known about networks with different nonlinear systems. Here we show that even in these cases an underlying simple organization into two distinct subnetworks may dictate global dynamics. The type of 'small-world' network reported here may yet be a new kind of 'small-network' underlying the complex dynamics of other systems with many degrees of freedom and interactions at many space and time scales. We hope that our results will stimulate further work on the 'small-world' properties of such systems.

References

[1] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1999) 440–442.
[2] P. Yodzis, The stability of real ecosystems, Nature 289 (1981) 674–676.
[3] D. Tilman, J.A. Dowing, Biodiversity and stability in grasslands, Nature 367 (1994) 363–365.
[4] R. Albert, H. Jeong, A.-L. Barabasi, Diameter of the world wide web, Nature 401 (1999) 130–131.

[5] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabasi, The large-scale organization of metabolic networks, Nature 407 (2000) 651–654.
[6] F. Liljeros, C. Edling, L.N Amaral, H.E. Stanley, Y. Aberg, The web of human sexual contacts, Nature 411 (2001) 907–908.
[7] H. Jeong, S. Mason, A.-L. Barabasi, Z.N. Oltvai, Lethability and centrality in protein networks, Nature 411 (2001) 41–42.
[8] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (2001) 3200–3203.
[9] J.-P. Bouchaud, M. Mezard, Wealth condensation in a simple model of economy, Physica A 282 (2000) 536–540.
[10] R. Kistler, et al., The NCEP/NCAR 50-year reanalysis: monthly means, CD-ROM and documentation, Bull. Am. Meteor. Soc. 82 (2001) 247–267.
[11] B. Bollabas, Random Graphs, Academic Press, London, 1985.
[12] J.M. Wallace, D.S. Gutzler, Teleconnections in the geopotential height field during the northern hemisphere winter, Mon. Weather Rev. 109 (1981) 784–812.
[13] D.J. Karoly, Southern hemisphere circulation features associated with ENSO events, J. Climate 2 (1989) 1239–1252.
[14] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.
[15] M.E.J. Newmann, The structure of scientific collaboration networks, Proc. Natl. Acad. Sci. USA 98 (2001) 404–409.
[16] L.A.N. Amaral, A. Scala, M. Barthelemy, H.E. Stanley, Classes of behavior of small-world networks, Proc. Natl. Acad. Sci. USA 97 (2000) 11149–11152.
[17] M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the internet topology, Comp. Comm. Rev. 29 (1999) 251.
[18] S.N. Dorogovtsev, J.F.F. Mendes, Language as an evolving word web, Proc. Roy. Soc. London B 268 (2001) 2603–2606.
[19] R. Ferrer, R.V. Sole, The small-world of human language, Proc. Roy. Soc. London B 268 (2001) 2261–2266.
[20] A.-L. Barabasi, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[21] R.A. Tomas, P.J. Webster, Horizontal and vertical structure of cross-equatorial wave propagation, J. Atmos. Sci. 51 (1994) 1417–1430.
[22] G. Love, Cross-equatorial influence of winter hemisphere subtropical cold surges, Mon. Weather Rev. 113 (1985) 1487–1498.
[23] G.P. Compo, G.N. Kiladis, P.J. Webster, The horizontal and vertical structure of east Asian winter monsoon pressure surges, Q. J. Roy. Meteor. Soc. 125 (1999) 29–54.
[24] G.A. Meehl, et al., Modulation of equatorial subseasonal convective episodes by tropical-extratropical interaction in the Indian and Pacific ocean regions, J. Geophys. Res. 101 (1996) 15033–15049.
[25] M.L. Carrera, Significant events of inter-hemispheric atmospheric mass exchange, Ph.D. Thesis, McGill University, 2001.
[26] A.A. Tsonis, J.B. Elsner, Chaos, strange attractors, and weather, Bull. Am. Meteor. Soc. 70 (1989) 16–23.
[27] G.S. Duane, P.J. Webster, J.B. Weiss, Co-occurrence of northern and southern hemisphere blocks as partially synchronized chaos, J. Atmos. Sci. 56 (1999) 4183–4205.
[28] S.N. Dorogovtsev, A.N. Samunkhin, J.F.F. Mendes, Multifractal properties of growing networks, Europhys. Lett. 57 (2002) 334–338.
[29] L.M. Pecora, et al., Fundamentals of synchronization in chaotic systems: concepts and applications, Chaos 7 (1997) 520–543.
[30] G.D. VanWiggeren, R. Roy, Communication with chaotic lazers, Science 279 (1998) 1198–1200.