What do networks have to do with climate?

Anastasios A. Tsonis, Kyle. L. Swanson and Paul. J. Roebber

Department of Mathematical Sciences, Atmospheric Sciences Group,

University of Wisconsin-Milwaukee, Milwaukee, WI 53201-0413

Corresponding author:

Anastasios Tsonis

Tel: 414-229-5373

Fax: 414-229-4907

Email: aatsonis@uwm.edu

Submitted to BAMS

CAPSULE: Networks offer a new way to study the collective behavior of interactive systems. Advances into understanding the coupling architecture of complex networks have already resulted in new insights in many areas of science.

ABSTRACT

The study of networks has recently exploded into a major research tool in many areas of sciences. The discovery of 'small-world' and scale-free networks has led to many new insights about the collective behavior of a large number of interacting agents and complex systems. Here we introduce the basic ideas behind networks as well as some initial applications of networks to the climate system. Our results suggest that the climate system exhibits aspects of 'small-world' networks as well as of scale-free networks with super nodes corresponding to major teleconnection patterns. This result suggests that the organization of teleconnections may play a role in the stability of the climate system. In addition, preliminary work suggests that temporal changes in the network's architecture may be used to identify signatures of global change. These and other applications suggest that networks provide a new tool for investigating and reconstructing climate dynamics from both models and observations.

Last summer one of us (AAT) visited the Greek island of Corfu. During summer the population of the island is 100,000 people. Before he went there, he knew only two Corfians. The first night of his stay he dined in one of the restaurants close to his hotel. Greeks being the friendly people they are, always open a discussion about who you are, what you do, and so on. So, as the waiter takes his order they start talking about such things. What do you think is the probability that the waiter knows one of the two Corfians? This problem has an analytical solution but the fact is that the waiter did know one of the two Corfians. Otherwise stated, it only took one connection between two persons that did not know each other to arrive to a common link. It is a small world after all! Through the work of the American psychologist Stanley Milgram and other subsequent investigations we know that any two of the six billion people on Earth are linked by a trail of only six people (Milgram 1967). This is referred to as the *six degrees of separation*.

Insights in such strange but otherwise common connections have been provided by the study of networks. A network is a system of interacting agents. In the literature an agent is called a node. The nodes in a network can be anything. For example, in the network of actors, the nodes are actors that are connected to other actors if they have appeared together in a movie. In a network of species the nodes are species that are connected to other species they interact with. In the network of scientists, the nodes are scientists that are connected to other scientists if they have collaborated. In the grand network of humans each node is an individual, which is connected to people he or she knows.

There are four basic types of networks.

a) Regular (ordered) networks. These networks are networks with a fixed number of nodes, each node having the same number of links connecting it in a specific way to a number of neighboring nodes (Fig. 1, left panel). If each node is linked to all other nodes in the network, then the network is a fully connected network.

.

b) Classical random networks. In these networks (Erdos and Renyi 1960) the nodes are connected at random (Fig. 1, right panel). In this case the degree distribution is a Poisson distribution (the degree distribution, p_k, gives the probability that a node in the network is connected to k other nodes). The problem with these networks is that they are not very stable. Removal of a number of nodes at random, may fracture the network to non-communicating parts.

c) Small-world networks. Regular networks are locally clustered, which means that, unless they are fully wired, it takes many steps to go from a node to another node away from its immediate neighborhood. 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. Thus, random networks exhibit efficient information transfer and regular networks do not (unless they are fully connected). 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. Work in this direction led to a new type of network, which was proposed a few years ago by the American mathematicians Duncan Watts and Steven Strogatz (1998) and is called *'small-world'* networks. A 'small-world' network is a superposition of regular and classical random graphs. Such 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. Those long-range connections do not have to be designed. A few long-range connections added at random will do the trick (Fig. 1, middle panel). Both random and ‘small-world’ networks are rather homogeneous networks in which each node has approximately the same number of links $<k>$. Both have nearly Poisson degree distributions that peak at $<k>$ and decay exponentially for large k.

d) Networks with a given degree distribution. The ‘small-world’ architecture can explain phenomena such as the six-degrees of separation (most people are friends with their immediate neighbors but we all have one or two friends a long way away), but it really is not a model found often in the real world. In the real world the architecture of a network is neither random nor ‘small-world’ but it comes in a variety of distributions such as truncated power-law distributions (Newmann 2001), Gaussian distributions (Amaral et al. 2000), power-law distributions (Faloutsos et al. 1999), and distributions consisting of two power-laws separated by a cutoff value of k (Dorogovtsev and Mendes 2001; Ferrer and Sole 2001). The last two types emerge in certain families of networks that grow in time (Dorogovtsev and Mendes 2001; Barabasi and Albert 1999).

The most interesting and common of such networks are the so-called *scale-free* networks, in which the degree distribution is the power law $p_k \sim k^{-\gamma}$. Consider a map showing an airline’s routes (Fig. 2). This map has a few hubs connecting with many other points (super nodes) and many points connected to only a few other points, a property associated with power law distributions. Such a map is highly clustered, yet it allows motion from a point to another far away point with just a few connections. As such, this network has the *property* of ‘small-world’ networks, but this property is not achieved by local clustering and a few random connections. It is achieved by having a few elements with large number of links and many elements having very few links. Thus, even though they share the same property, the architecture of scale-free networks is different than that of ‘small-world’ networks. Such inhomogeneous networks have been found to pervade biological, social, ecological, and economic systems, the internet, and other systems (Albert et al.1999; Jeong, et al. 2000; Liljeros et al. 2001; Jeong et al. 2001; Pastor-Satorras and Vespignani 2001; Bouchaud and Mezard 2000; Farkas et al. 2003; Barabasi and Bonabeau 2003; Albert and Barabasi 2002). These networks are referred to as scale-free because they show a power-law distribution of the number of links per node. Lately, it was also shown that, in addition to the power-law degree distribution, many real scale-free networks consist of self-repeating patterns on all length scales. This result is achieved by the application of a renormalization procedure that coarse-grains the system into boxes containing nodes within a given 'size' (Song et al. 2005). In other words scale-free networks also exhibit fractal geometry. These properties are very important because they imply some kind of self-organization within the network. Scale-free networks are not only efficient in transferring information, but due to the high degree of local clustering they are also very stable (Barabasi and Bonabeau 2003). Because there are only a few super nodes, chances are that accidental removal of some nodes will not include the super nodes. In this case the network would not become disconnected. This is not the case with random and to a lesser degree with ‘small-world’ networks, where accidental removal of the same percentage of nodes makes them more prone to failure (Barabasi and Bonabeau 2003; Albert et al. 2000). A scale-free network is vulnerable only when a super node is ‘attacked’. Note that scale-free networks have properties of ‘small-world’ networks, but ‘small-world’ networks a la Watts and Strogatz are not scale-free. An example of such a network is given in Fig. 3, which shows the network of interactions between the proteins in the yeast *Saccharomyces cerevisiae*, otherwise known as baker's yeast (Jeong et al. 2001). By looking at the connectivity of each protein to other proteins, the authors were able to determine that more than 90% of the proteins in the network have less than five links and only one in five of these were essential to the survival of the yeast. In other words removing these proteins did not affect the function of this organism. In contrast, they found that less than 0.7% of the proteins were hubs having many more than fifteen connections. For these hubs they found that removal of any hub resulted in the death of the organism. Such findings, which can only be delineated by constructing the network, can be extremely useful as they may lead to ways to protect the organism from microbes by specifically protecting the hubs. In other areas, the presence of scale-free networks has led to strategies to slow the spread of diseases (Lijeros et al. 2001) and strategies to secure the internet (Barabasi and Bonabeau 2003).

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 (some times in the literature growing networks are classified as a new type of network). Whatever the type of the network, its underlying topology provides clues about the collective dynamics of the network. The basic 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 and is illustrated in Fig. 4: Assume that a node i is connected to k_i other nodes. Now consider the k_i closest nodes of i. This defines the neighborhood of i. Then count the number of links, Δ_i, between any two nodes of the neighborhood (excluding node i). The clustering coefficient of node i is then given by $C_i=2\Delta_i/k_i(k_i-1)$. Since there can be at most $k_i(k_i-1)/2$ links between k_i nodes (which will happen if they formed a fully connected subnetwork), the clustering coefficient is normalized on the interval [0,1]. The average C_i over all nodes provides C. As such C provides a measure of local "cliqueness". The diameter of the network is defined by the number of connections in the shortest path between two nodes in the network averaged over all pairs of nodes. For a random network having the same average number of connections per node, $<k>$, it can be shown analytically (Bollabas 1985; Watts and Strogatz 1998; Albert and Barabasi 2001) that $L_{random}=lnN/ln<k>$ and $C_{random}=<k>/N$. The 'small-world' property requires that $C>> C_{\rm random}$ and $L\geq L_{\rm random}$. There are other measures and ways to investigate networks. Examples include minimum spanning trees (Mantegna 1998), asset trees and asset graphs (Onnela et al. 2004), tree length and occupation levels (Onnela et al. 2003), intensity and coherence of networks (Onnela et al. 2005), but here we will stick with the basic principles.

CLIMATE NETWORKS. How can these ideas be extended to a system like the climate? One way is to assume that interacting dynamical systems can also form a network. Consider for example the results shown in Fig. 5. In this figure we start with a number of limit-cycle (periodic) oscillators with distributed natural frequencies. The state of each oscillator is represented as a dot in the complex plane. The amplitude and phase of each oscillation correspond to the radius and angle in polar coordinates. The color of each oscillator indicates its natural frequency (ranging from violet to red or from high to low frequency). If the oscillators are not coupled, then each oscillator will settle onto its limit cycle and will rotate at its natural frequency. When they are coupled, however, the oscillators appear to self-organize and rotate as a synchronized group with locked amplitudes and phases. When the oscillators are more complex than a limit cycle (for example chaotic) the situation can be more complicated, but studies have shown that synchronization is possible in this case as well (Strogatz 2001). This synchronization translates into links between the individual oscillators that define the structure of the network of these dynamical systems. Thus, one way to apply networks to climate system is to assume that climate is represented by a grid of oscillators each one of them representing a dynamical system varying in some complex way. What we are then interested in, is the collective behavior of these interacting dynamical systems and the structure of the resulting network. Next, we will present an example, which will introduce us to the applications and promise of networks in atmospheric sciences. Some of these ideas have been presented in two recent publications by Tsonis and Roebber (2004) and Tsonis (2004).

We start by considering the global National Center for Environmental Prediction/ National Center for Atmospheric Research (NCEP/NCAR) reanalysis 500 hPa data set (Kistler et al. 2001). A 500 hPa value indicates the height of the 500 hPa 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^o latitude x 5^o longitude. For each grid point monthly values from 1950 to 2004 are available. 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. These 2664 points will be assumed to be the nodes of the network. From the monthly 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 660 anomaly values. In order to define the "connections" between the nodes, 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] is estimated. 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". Even though most of the annual cycle is removed by producing anomaly values, some of it is still present as the amplitude of the anomalies is greater in the winter than in the summer. For this reason, in order to avoid spurious high values of *r*, only the values for December, January and February in each year were considered. It follows that the estimation of the correlation coefficient between any two time series is based on a sample size of 165. Note that since the values are monthly anomalies there is very little autocorrelation in the time series. A pair is considered as connected if their correlation $|r| \geq 0.5$. This criterion is based on parametric and non-parametric significance tests. According to the t-test with N=165, a value of r=0.5 is statistically significant above 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. The use of the correlation coefficient to define links in networks is not new. Correlation coefficients have been used to successfully derive the topology of gene expression networks (Farkas et al. 2003; de la Fuente et al. 2002; Featherstone and Broadie 2002; Agrawal, 2002), and to study financial markets (Mantegna 1999). The choice of r=0.5 while it guarantees statistical significance is somewhat arbitrary. The effect of different correlation threshold is discussed in Tsonis and Roebber (2003). In any case, one may in fact consider all pairs as connected and study the so called weighted properties of the network where each link is assigned a weight proportional to its corresponding correlation coefficient (Onnela et al. 2003, 2004). For the scope of this paper, however, we will keep things simple and consider that a pair is connected if the correlation coefficient is above a threshold.

Once we have decided what constitutes a link, we are ready to look at the architecture of this network and how does it relate to dynamics. Figure 6 provides a first insight to this question. It shows the area weighted number of total links (connections) at each geographic location. More accurately it shows the fraction of the total global area that a point is connected to. This is a more appropriate way to show the architecture of the network because the network is a continuous network defined on a sphere, rather than a discrete network defined on a two-dimensional grid. Thus, if a node *i* is connected to *N* other nodes at λ_N latitudes then its area weighted connectivity, $\tilde{C}_i$, is defined as

$$\tilde{C}_i = \sum_{j=1}^{N} \cos\lambda_j \Delta A \Big/ \sum_{over\ all\ \lambda\ and\ \phi} \cos\lambda\Delta A \qquad (1)$$

where ΔA is the grid area at the equator and φ is the longitude. Once we have the information displayed in Fig. 6 we can estimate *C* and *L*. According to the definition of connectivity (equation 1), in order to find the clustering coefficient of node *i*, C_i, we consider a circular area on the sphere centered on *i* which is equal to $\tilde{C}_i$. Then C_i is the fraction of this circular area that is connected (for a fully connected area, i.e. all pairs of nodes are connected, C_i=1 for all *i*). The average C_i over all nodes provides the clustering coefficient of the network, C. Note also that according to the definition of $\tilde{C}$, the average $\tilde{C}_i$ over all nodes gives the clustering coefficient C_{random}. Concerning the estimation of L, rather than finding the number of connections in the shortest path between two points, we estimate the distance of this path on the sphere. For this network we find that $L \cong 9{,}600$ Km and C=0.56. For a random network with the same specifications (number of nodes, and average links per node) it is estimated that $L_{\text{random}} \cong 7{,}500$ Km and C_{random}=0.19. These values indicate that indeed $L \geq L_{\text{random}}$ and $C > C_{\text{random}}$ (by a factor of three), which will make this global network close to a 'small-world' network. There is, however, more to this global network than what these values suggest.

Returning to Fig. 6 we observe that it displays two very interesting features. In the tropics it appears that all nodes posses more or less the same number of connections, which is a characteristic of fully connected networks. In the extratropics it appears that certain nodes posses more connections than the rest, which is a characteristic of scale-free networks. In the northern hemisphere we clearly see the presence of regions where such super nodes exist in China, North America and Northeast Pacific Ocean. Similarly several super nodes are visible in the southern hemisphere. These differences between tropics and extratropics are clearly delineated in the corresponding degree distributions. Figure 7 shows, on a double logarithmic plot, the distribution of nodes according to how many links they possess (i.e. p_k against k). Given the definition of a link in our case, this figure indicates the fraction of the total area covered as a function of the connectivity, $\tilde{C}$. More specifically, Fig. 7a shows the distribution of nodes in the extratropical region of 30N-65N and 30S-65S and Fig. 7b shows the corresponding distribution of nodes in the tropics (20N-20S). The region from 20N-30N and 20S-30S is a transition between the two regimes and was left out for better delineation of the properties in the tropics and extratropics. Figure 7a appears to exhibit a scaling regime similar to those observed in scale-free networks. In fact, the slope of this graph is around -2.0 in agreement with other scale-free networks (Barabasi and Albert 1999). In Fig. 7b no such regime is identifiable. The distribution is basically a narrow peak at about $\tilde{C} = 0.5$ indicating that most points possess the same large number of connections, a characteristic of regular almost fully connected networks. Deviations from the power law (manifesting as a peak at about $\tilde{C} = 0.4$ in 7a) and uniformity (the four points below $\tilde{C} = 0.1$ in 7b) are due to the fact that the two subnetworks are interwoven; a node in one subnetwork may be connected to a node or nodes in the other subnetwork. Note that very similar results are obtained if instead of the 30-65 N and 30-65 S belts the whole extratropical area (35-90N and 35-90 S) is considered. It thus appears that the overall network is a "fusion" of a fully connected tropical network and a scale free extratropical network. As is the case with all scale-free networks, the extratropical subnetwork is also a 'small-world' network. Indeed, we find that for points in the extratropics, the clustering coefficient is much greater than that of a corresponding random network (by a factor of nine). The collective behavior of the individual dynamical systems in the complete network is not described by a single type but it self-organizes into two coupled subnetworks, one regular almost fully connected operating in the tropics and one scale-free/'small-world' operating in the higher latitudes. The extratropics are considered as one subsystem, even though they are physically separated. Whether or not we consider them as one or two subsystems it 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 (Tomas and Webster 1994; Love 1985; Compo et al. 1999; Meehl et al. 1996; Carrera 2001) and with the notion of sub-systems in climate proposed in the late 1980s (Tsonis and Elsner 1989; Lorenz 1991), and with recent studies on synchronized chaos in the climate system (Duane et al. 1999).

An interesting observation in Fig. 6 is that super nodes may be associated with major teleconnection patterns. For example, the super nodes in North America and Northeast Pacific Ocean coincide with the well-known Pacific North America (PNA) pattern (Wallace and Gutzler 1981). In the southern hemisphere we also see super nodes over the southern tip of South America, Antarctica and South Indian Ocean that are consistent with some of the features of the Pacific South America (PSA) pattern (Mo and Higgins 1998). Interestingly, no such super nodes are evident where the other major pattern, the North Atlantic Oscillation (NAO) (Thompson and Wallace 1998; Pozo-Vazquez et al. 2001; Huang et al. 1998) is found. This does not indicate that NAO is not a significant feature of the climate system. Since NAO is not strongly connected to the tropics, the high connectivity of the tropics with other regions is masking NAO out. In fact, if we consider a network with only nodes north of the 30 N latitude, we find (Fig. 8) that a dipole consistent with NAO is not only present but it is also a prominent feature of the network. It should be noted here that in their pioneering paper Wallace and Gutzler (1981) defined teleconnectivity at each grid point as the strongest negative correlation between a grid point and all other points. This brings out teleconnection patterns associated with waves such as the trough-ridge-trough PNA pattern. However, because of the requirement of strongest negative correlation (which occurs between a negative anomaly center and a positive anomaly center), this approach can only delineate long-range connections. As such, information about clustering and connectivity at other spatial scales is lost. In the network approach all the links at a point are considered and as such much more information (clustering coefficients, diameter, scaling properties etc) can be obtained. The similarities between Wallace and Gutzler's results and the network results arise from the fact that grid points with many long-range links will most likely stand out.

The physical interpretation of the results is that the climate system (as represented by the 500 hPa field) 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. These fluctuations will tend to destabilize the source region. For example, dynamical connections between the ocean and the atmosphere during an El Nino may make the climate over tropical Pacific less stable. However, the 'small-world' as well as the scale-free property of the extratropical network and the fully connected tropical network allow the system to respond quickly and coherently to fluctuations introduced into the system. This "information" transfer diffuses local fluctuations thereby reducing the possibility of prolonged local extremes and providing greater stability for the global climate system. An important consequence of this property is that local events may have global implications. The fact that the climate system may be inherently stable may not come as a surprise to some, but it is interesting that we find that a stable climate may require teleconnection patterns.

Unlike networks where a node is solidly defined (think of social networks where a node is a person), here a node is a point on a grid, which is defined rather arbitrarily and/or can be represented at various resolutions. Strictly speaking our network has infinite nodes. Due, however, to spatial correlations the "effective" number of nodes is much less. In fact, we find that reproducing Fig. 6 but with the full grid of 2.5^ox2.5^o resolution results in virtually the same network architecture. Other studies have demonstrated this as well. For example, in the *www* network an increase of the number of nodes by a factor of 2,500 results in increase of *L* by a factor of only 1.6 and the architecture (vis-à-vis degree distribution) remains identical (Albert and Barabasi 2002). Other meteorological fields may also exhibit small-world or scale-free properties. As an additional example we used upper tropospheric streamfunction. Figure 9 is like Fig. 8 but for the streamfunction. Again here we observe super nodes. This network is also a scale-free and a 'small-world' network.

DISCUSSION. From this initial application of networks to climate it appears that atmospheric fields can be thought as a network of interacting points whose collective behavior may exhibit properties of 'small-world' networks. This ensures efficient transfer of information. In addition, the scale-free architectures guarantee stability. Furthermore, super nodes in the network identify teleconnection patterns. As was demonstrated in Tsonis (2004), these teleconnections are not static phenomena but their spatiotemporal variability is affected by large (global) changes. The 55-year period used to produce Fig. 5 can be divided into two distinct periods each of length of 27 years (1951-1977 and

1978-2004). During the first period the global temperature shows no significant overall trend. During the second period, however, a very strong positive trend is present. Since there is a distinct change in the global property of the system, how does this affect the dynamics of the global network? To answer this question *C* and *L* for the two periods were estimated. It was found that *C* is about 5% smaller and *L* is about 4% smaller in the second period. This result will indicate that during the warming of the planet the network has acquired more long-range connections and less small range-connections. This is clearly shown in Fig. 10, which shows the distribution of connections according to their distance (as calculated on the sphere). The solid line represents the distribution in the first period and the dashed line the distribution in the second period. This figure shows that the frequency of long-range connections (>6,000 km) has increased whereas the frequency of shorter-range connections (<6,000 Km) has decreased. The differences between the two distributions may not appear impressive but with hundreds of thousands of connections involved, these differences are, according to the Kolmogorov-Smirnov test, statistically significant at the 99% confidence level (see also Tsonis 2004). Even though this is only one example, this result suggests that monitoring the properties of such networks may provide an additional tool to identify or verify climate changes.

Furthermore, mapping atmospheric fields into networks appears to bring out properties of the general circulation. Thus, it may provide an alternative approach to study atmospheric phenomena and dynamics. Moreover, just because in the case of 500 hPa teleconnections the network approach brings out what has been found by linear approaches (such as EOF analysis), it does not mean that it will always produce what linear approaches produce.

For example, scale-free phenomena are associated with nonlinear dynamics. As such linear approaches, such as EOF analysis *cannot* bring out this property. The fact that our network approach recovers the scale-free characteristic is a strong indication that it will *not* always produce the same result as a linear approach and that in fact it may produce novel insights. The following presents another example where linear approaches would not yield certain properties. The current standard for understanding non-local interactions in the atmosphere is linear inverse modeling (e.g. Winkler et al. 2001). In plain terms, linear inverse modeling is similar to a least squares fit linking the values of certain dynamical quantities at one time (e.g., a subset of empirical orthogonal functions) with their values at some future time. In practice, this involves making the assumption that the dynamics are sufficiently approximated by a linear, stable, stochastic dynamical system, and the propagator for that system is then calculated from data at some fixed time lag. It is not clear that such linear approaches are necessarily optimal, however, and specifically, whether they distort the network structure of the atmosphere. For an example of such distortion, consider a one dimensional discrete dynamical system similar to that introduced by Lorenz and Emanuel (1998) which is chaotic and mimics zonal wave propagation in the atmosphere. The relevant time-scale in this system is the error doubling time, which we take as one model day. This model has 40 nodes, and is modified to possess a long range spatial correlation by allowing nodes 20 and 40 to force each other. Given this set up, we calculated the correlation for each node with every other node at a lag of 5 model days and defined links for $|r| \geq 0.5$. Our rationale in doing this was to see how effectively information is transferred into the future. The solid line in Fig. 11 shows the number of links of each node. We observe that most of the nodes possess just about 4-5 links but a few nodes stand above this level with more links. Interestingly, as expected from a nonlinear system, it is not necessarily nodes 20 and 40 that possess most of the links. However, a linear inverse model (broken lines) constructed to provide a 5-day forecast does not possess a similar connectivity pattern. Apparently the linear inverse approach does not preserve the properties of the actual nonlinear model.

True enough the correlation coefficient is a linear measure. One may question why we should use a linear measure to study nonlinear dynamics. This is a legitimate question. A nonlinear measure that could be used instead is the mutual information, but the problem is that its accurate estimation requires much more data than we have available. As such there is very little choice but to use the correlation coefficient as an indicator of a link. However, the correlation coefficient is used only as a construction tool. In a similar way it has been used successfully in the past to delineate nonlinear dynamics. For example, in reconstructing attractors from time series, the state space is reconstructed using coordinates that are shifts of the original time series. The optimum shift is often assumed to be the lag at which the autocorrelation function first becomes zero. This approach reproduces the properties of the attractor quite reasonably (Tsonis 1992, Tsonis and Elsner 1989).

OUTLOOK. Complex network describe many natural and social dynamical systems and their study has revealed interesting mechanisms underlying their function. The novelty of networks is that they bring out topological/geometrical aspects which are related to the physics of the dynamical system in question thus providing a new and innovative way to treat and investigate nonlinear systems and data. While several advances have been made, this area is still young and the future is wide open. This introductory paper presented some fundamental aspects of networks and some preliminary results of the application of networks to climatic data, which indicate that networks delineate some key features of the climate system. This suggests that networks have the potential to become a new and useful tool in climate research.

REFERENCES

Albert, R., H. Jeong, and A-L Barabasi, 1999: Diameter of the World Wide Web. *Nature* **401**, 130-131.

Albert, R., H. Jeong, and A.-L.Barabasi, 2000: Error and attack tolerance of complex networks. *Nature* **406**, 378-382.

Albert, R., and A.-L. Barabasi, A-L. 2002: Statistical mechanics of complex networks. *Rev. Mod. Phys***.74***,* 47-101.

Amaral, L.A.N., A. Scala, M. Barthelemy, and H.E. Stanley, 2000: Classes of behavior of small world networks. *Proc. Natl. Acad. Sci. USA* **97**, 11149-11152.

Agrawal, H., 2002: Extreme self-organization in networks constructed from gene expression data. *Phys. Rev. Lett.* **89**, 268702.

Barabasi, A-L., and R. Albert, 1999: Emergence of scaling in random networks. *Science* **286**, 509-512.

Barabasi, A.-L., and E. Bonabeau, 2003: Scale-free networks. *Scientific American* **288**, 60-69.

Bollabas, B., 2001: *Random Graphs*, 2nd Ed., Cambridge, 498 pp.

Bouchaud, J-P., and M. Mezard, 2000: Wealth condensation in a simple model of economy. *Physica A* **282**, 536-540.

Carrera, M.L., 2001: Significant events of inter-hemispheric atmospheric mass exchange. Ph.D. thesis, McGill University, Montreal, Canada.

Compo, G.P., G.N. Kiladis, and P.J. Webster, 1999: The horizontal and vertical structure of east Asian winter monsoon pressure surges. *Quart J. Roy Meteor. Soc*. **125**, 29-54.

Dorogovtsev, S.N., and J.F.F. Mendes, 2001: Language as an evolving word web. *Proc. Roy. Soc. London B* ***268***, 2603-2602.

Duane, G.S., P.J. Webster, and J.B. Weiss, 1999: Co-occurrence of northern and southern hemisphere blocks as partially synchronized chaos. *J. Atmos Sci.* **56**, 4183-4205.

Erdos, P., and A. Renyi, 1960: On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17-61.

Faloutsos, M., P. Faloutsos, and C. Faloutsos, 1999: On power-law relationships of the internet topology. *Comp. Comm. Rev.* **29**, 251-260.

Farkas, I. J., H. Jeong, T. Vicsek, A.-L. Barabási, and Z.N. Oltvai, 2003: The topology of the transcription regulatory network in the yeast Saccharomyces cerevisiae. *Physica A* **318**, 601-612.

Featherstone, D. E., and K. Broadie, 2002: Wrestling with pleiotropy: Genomic and topological analysis of the yeast gene expression network. *Bioessays* **24**, 267-274.

Ferrer, R., and R.V. Sole, 2001: The small-world of human language. *Proc. Roy. Soc. London B* **268**, 2261-2266.

de la Fuente, A., P. Brazhnik, and P. Mendes, 2002:Linking the genes: Inferring quantitative gene networks from microarray data. *Trends in Genetics* **18**, 395-398.

Huang, J.P., K. Higuchi, and A. Shabbar, 1998: The relationship between the North Atlantic Oscillation and El Nino Southern Oscillation. *Geophys. Res. Lett.*. **25**, 2707-2710.

Jeong, H., S. Mason, A.-L. Barabasi, and Z.N. Oltvai, 2001: Lethability and centrality in protein networks. *Nature* **411**, 41-42.

Jeong, H., B. Tombor, R. Albert, A.N. Oltvai, and A.-L. Barabasi, 2000: The large scale organization of metabolic networks. *Nature* **407**, 651-654,

Kistler, R., and Coauthors, 2001: The NCEP/NCAR 50-year reanalysis: monthly means, CD-ROM and documentation. *Bull. Amer. Meteor. Soc.* **82**, 247-267.

Liljeros, F., C. Edling, L.N. Amaral, H.E. Stanley, and Y. Aberg, 2001: The web of human sexual contacts. *Nature* **411**, 907-908.

Lorenz, E.N., 1991: Dimension of weather and climate attractors. *Nature* **353**, 241-244.

Lorenz, E.N., and K.A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.* **55**, 399–414.

Love, G., 1985: Cross-equatorial influence of winter hemisphere subtropical cold surges. *Mon. Wea. Rev.* 113, 1487-1498.

Mantegna, R.N., 1999: Hierarchical structure in financial markets. *Eur. Phys. J.* B **11**, 193-197.

Meehl, G.A., and Coauthors1996: Modulation of equatorial subseasonal convective episodes by tropical-extratropical interaction in the Indian and Pacific ocean regions. *J. Geophys. Res.* **101**, 15, 033-15, 049.

Milgram, S., 1967: The small-world problem. *Psychology Today* **1**, 60-67.

Mo, K.C., and R.W. Higgins, 1998: The Pacific-South America modes and tropical convection during the southern hemisphere winter. *Mon. Wea. Rev*. **126**, 1581-1596.

Newmann, M.E.J., 2001: The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci USA* **98**, 404-409.

Onnela, J.-P., A. Chakraborti, K. Kaski, J. Kertesz, and A. Kanto, 2003: Dynamics of market correlations: Taxonomy and portfolio analysis. *Phys. Rev. E* **68**, 056110.

Onnela, J.-P., K. Kaski, and J. Kertesz, 2004: Clustering and information in correlation based financial networks. *Eur. Phys. J.* B **38**, 353-362.

Onnela, J.-P., J. Saramaki, J. Kertesz, and K. Kaski, 2005: Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E***71**, 065103.

Pastor-Satorras, R., and A. Vespignani, 2001: Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200-3203.

Pozo-Vazquez, D., M.J. Esteban-Parra, F.S. Rodrigo, and Y. Castro-Diez, 2001: The association between ENSO and winter atmospheric circulation and temperature in the North Atlantic region. *J Climate* **14,** 3408 - 3420.

Song, C., S. Havlin, and H.A. Makse, 2005: Self-similarity of complex networks. *Nature* **433,** 392–395.

Strogatz, S.H., 2001: Exploring complex networks. *Nature* **410**, 268-276.

Thompson, D.W.J., and J.M. Wallace, 1998: The Artic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett* **25(9)**, 1297-1300.

Tomas, R.A., and P.J. Webster, 1994: Horizontal and vertical structure of cross-equatorial wave propagation. *J. Atmos. Sci.* **51**, 1417-1430.

Tsonis, A.A., and J.B. Elsner, 1989: Chaos, strange attractors, and weather. *Bull. Amer. Meteor. Soc.* **70**, 16-23.

Tsonis, A.A., 1992: *Chaos: From theory to applications*, Plenum Press, 274 pp.

Tsonis, A.A., and P.J. Roebber, 2003: The architecture of the climate network. *Physica A* **333**, 497-504.

Tsonis, A.A., 2004: Does global warming inject randomness into the climate system? *EOS*, **85(38**), 361-364

Wallace, J.M., and D.S. Gutzler, 1981: Teleconnections in the geopotential height field during the northern hemisphere winter. *Mon. Wea. Rev*. **109**, 784-812.

Watts, D.J., and S.H. Strogatz, 1999: Collective dynamics of 'small-world' networks. *Nature* **393**, 440-442.

Winkler, C. R., M. Newman, and P.D. Sardeshmukh, 2001: A Linear Model of Wintertime Low-Frequency Variability. Part I: Formulation and Forecast Skill. *J. Climate* **14**, 4474–4494.

Figure captions

Fig. 1: Illustration of a regular, a small-world and a random network (after Watts and Strogatz, 1998, reproduced with permission from Science News).

Fig. 2: Route map for Continental Airlines (courtesy of Continental Airlines)

Fig. 3: The network of interactions between the proteins in the yeast Saccharomyces cerevisiae, otherwise known as baker's yeast (courtesy of A.-L. Barabasi).

Fig. 4: Illustration of how to estimate the clustering coefficient. In (a) a node *i* is connected to k_i=8 other nodes (solid lines). In (b) we consider the k_i=8 closest nodes of *i*. This defines the neighborhood of *i*. In this neighborhood there exist Δ_i=4 connections between nodes (excluding node *i*) (broken lines). The clustering coefficient of node *i* is then given by $C_i=2\Delta_i/k_i(k_i-1)=0.143$. The average C_i over all nodes in the network provides the clustering coefficient of the network, *C*.

Fig. 5: Synchronization of several coupled limit-cycle oscillators. Each oscillators starts from a random initial condition but soon they all self-organize and rotate as a synchronized group (after Strogatz, 2001, reproduced with permission form the author).

Figure 6: Total number of links (connections) at each geographic location. The uniformity observed in the tropics indicates that each node possesses the same number of connections. This is not the case in the extratropics where certain nodes possess more links than the rest. See text for details on how this figure was produced.

Fig. 7: Degree distribution of (a) extratropical (30N-65N, 30S-65S) grid points, and (b) tropical (20N-20S) grid points.

Fig. 8: Same as Fig. 5 but when we start with a network with nodes north of the 30 N latitude only.

Fig. 9: Same as Fig. 7 but for the streamfunction.

Fig. 10: The relative frequency distribution of the connections according to their distance for the period 1951-1977 (solid line) and for the period 1978-2004 (dashed line). This result indicates that during periods of warming, the network acquires more long-range connections and less small range-connections.

Fig. 11: An example demonstrating that linear approaches may not yield the true result of a nonlinear process (see text for details).

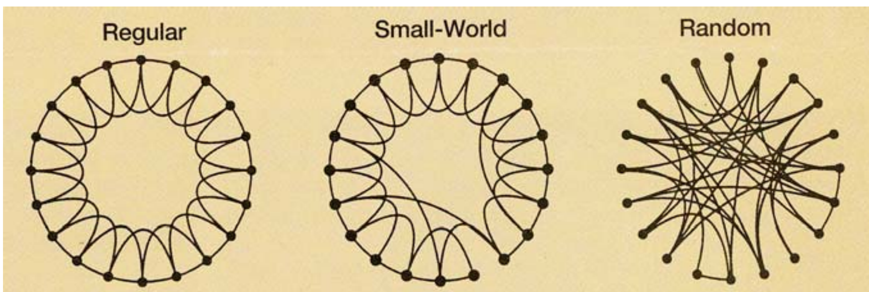


Figure 1

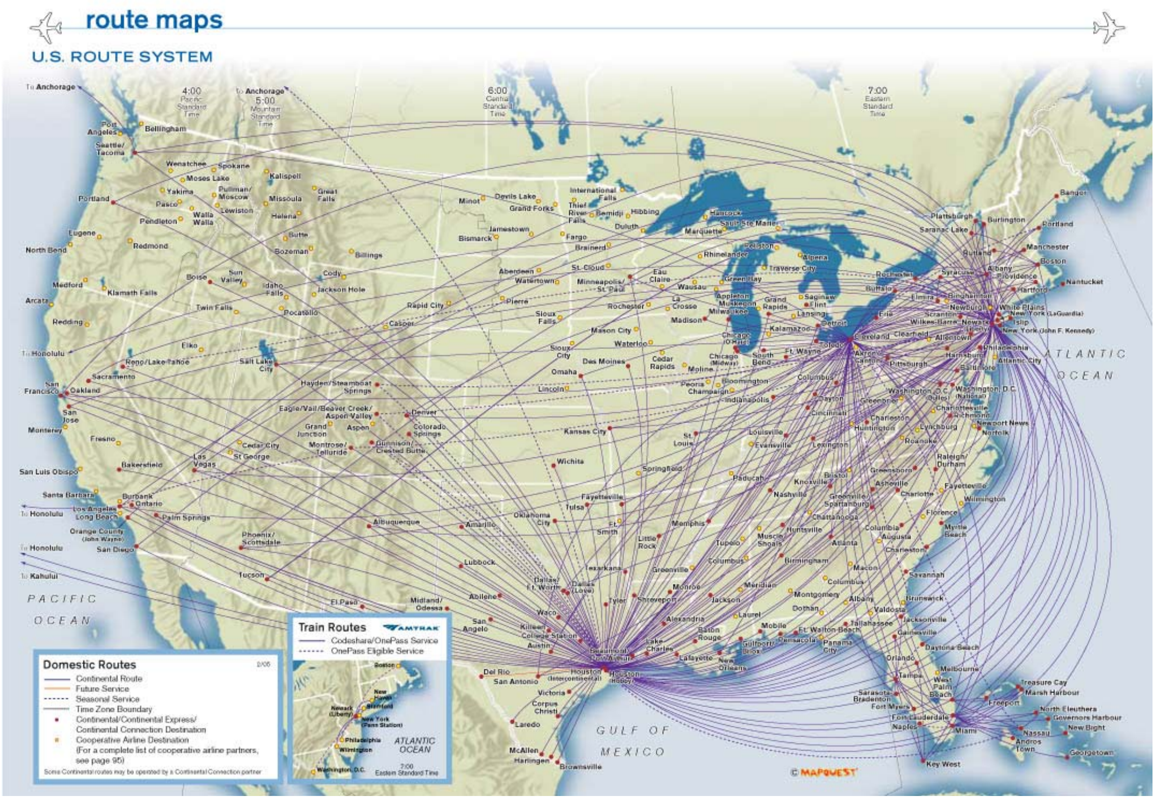

route maps
U.S. ROUTE SYSTEM
PACIFIC OCEAN
ATLANTIC OCEAN
GULF OF MEXICO
Domestic Routes
2/05
Continental Route
Future Service
Seasonal Service
Time Zone Boundary
Continental/Continental Express/
Continental Connection Destination
Cooperative Airline Destination
(For a complete list of cooperative airline partners, see page 95)
Some Continental routes may be operated by a Continental Connection partner
Train Routes
AMTRAK
Codeshare/OnePass Service
OnePass Eligible Service
7:00
Eastern Standard Time

Figure 2

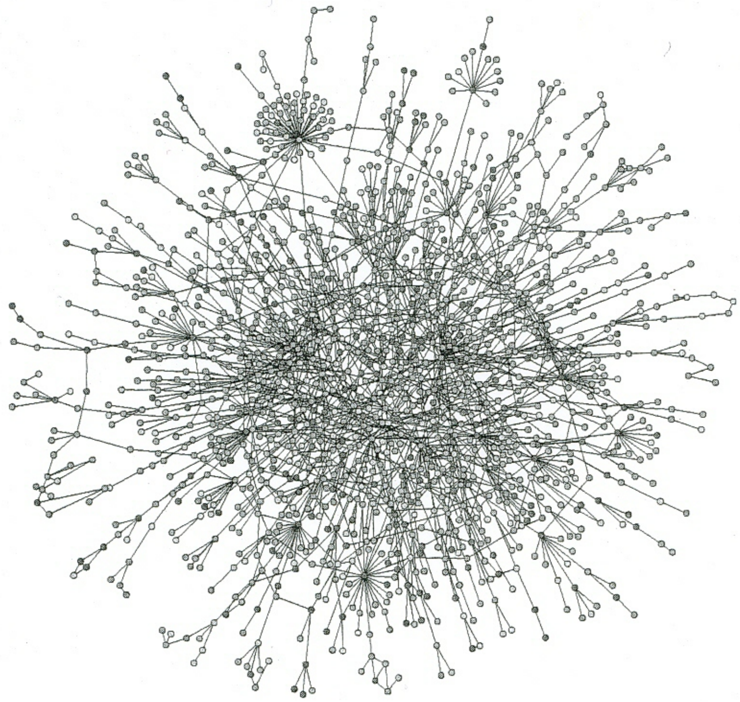


Figure 3

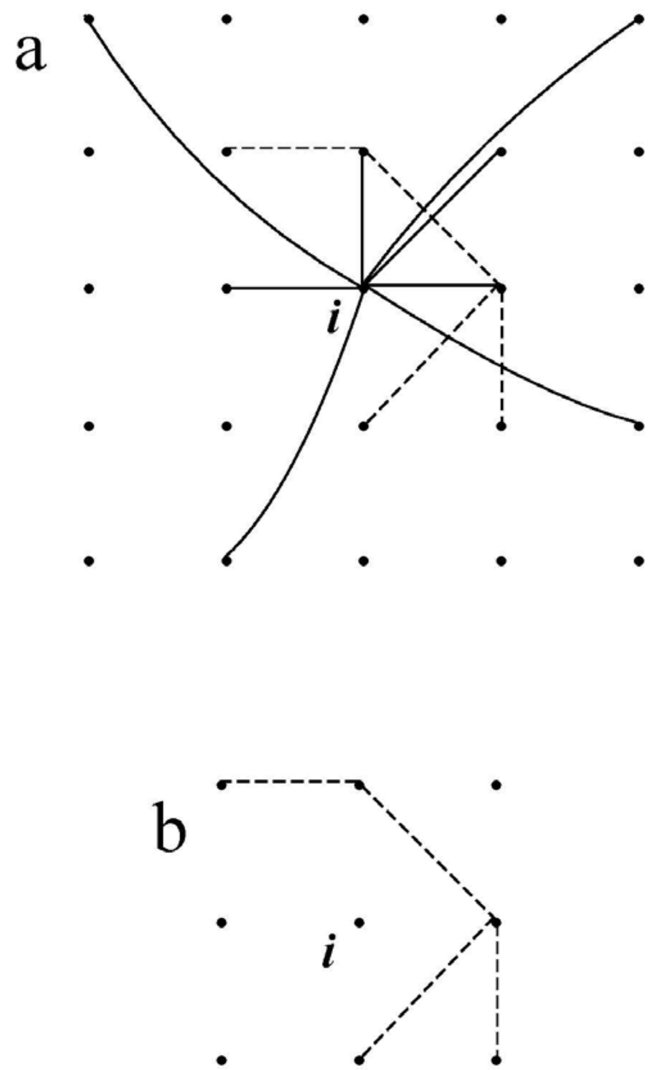


Figure 4

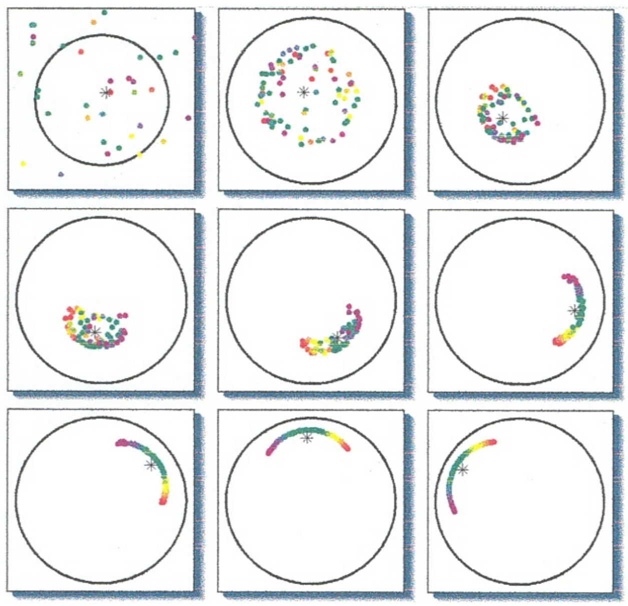
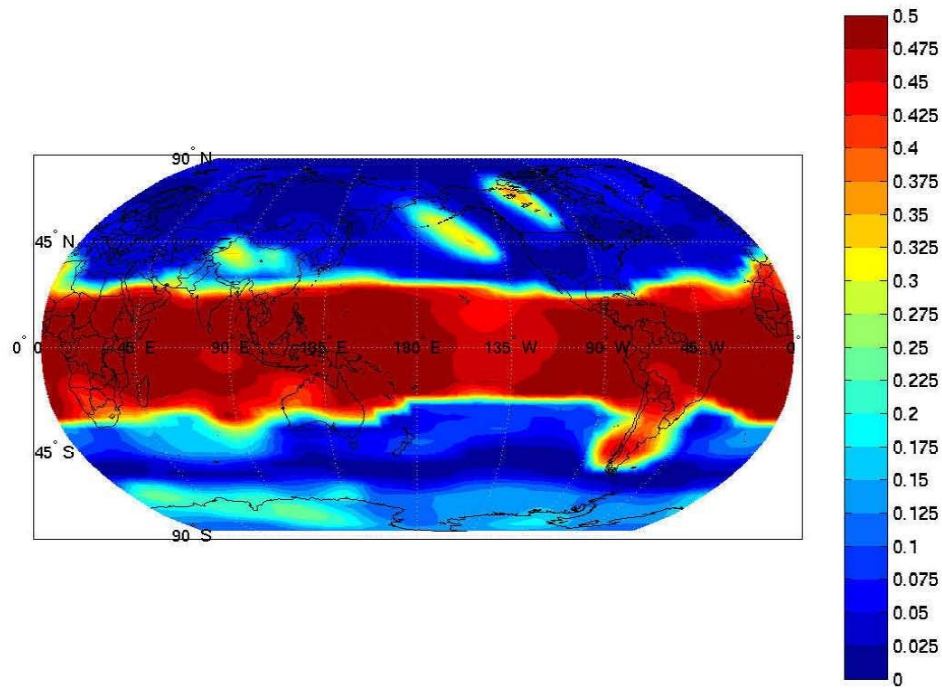
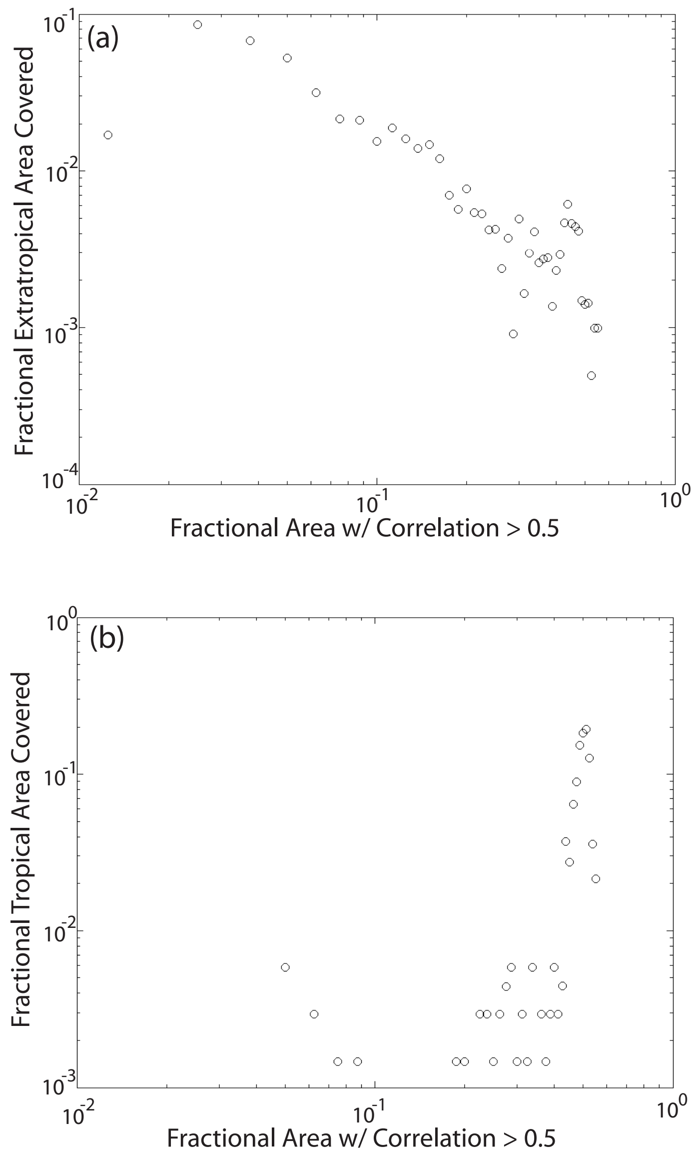
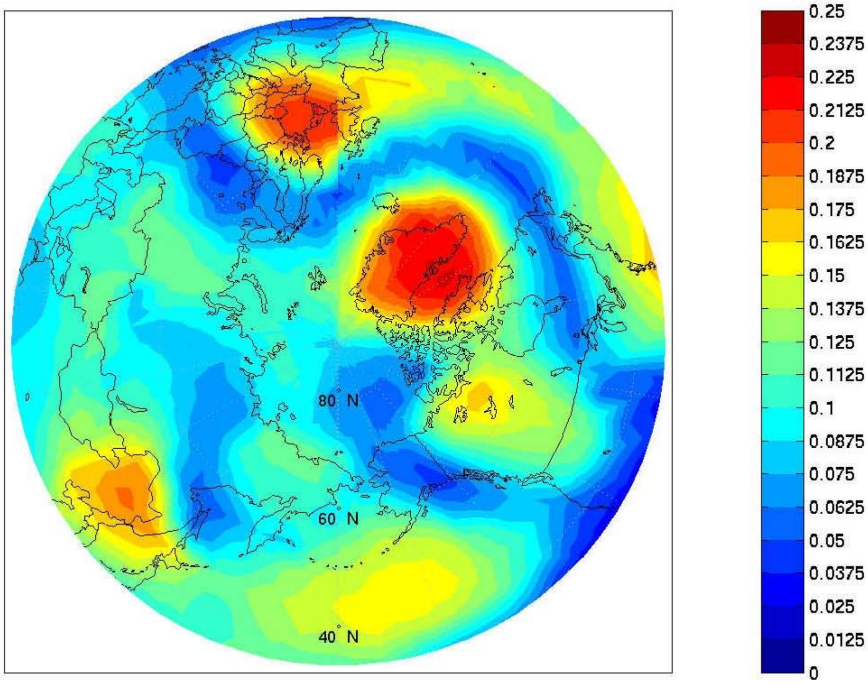


Figure 5

Figure 6


90° N
45° N
0°
45° E
90° E
135° E
180° E
135° W
90° W
45° W
45° S
90° S
0.5
0.475
0.45
0.425
0.4
0.375
0.35
0.325
0.3
0.275
0.25
0.225
0.2
0.175
0.15
0.125
0.1
0.075
0.05
0.025
0


(a)
Fractional Extratropical Area Covered
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-2}
10^{-1}
10^{0}
Fractional Area w/ Correlation > 0.5
(b)
Fractional Tropical Area Covered
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-2}
10^{-1}
10^{0}
Fractional Area w/ Correlation > 0.5



80° N
60° N
40° N
0.25
0.2375
0.225
0.2125
0.2
0.1875
0.175
0.1625
0.15
0.1375
0.125
0.1125
0.1
0.0875
0.075
0.0625
0.05
0.0375
0.025
0.0125
0

Figure 8

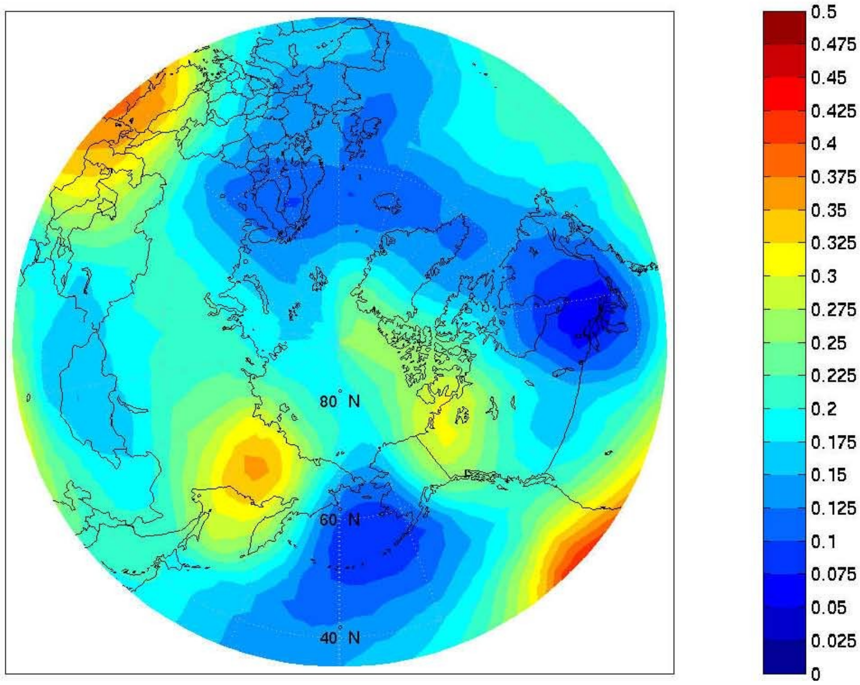
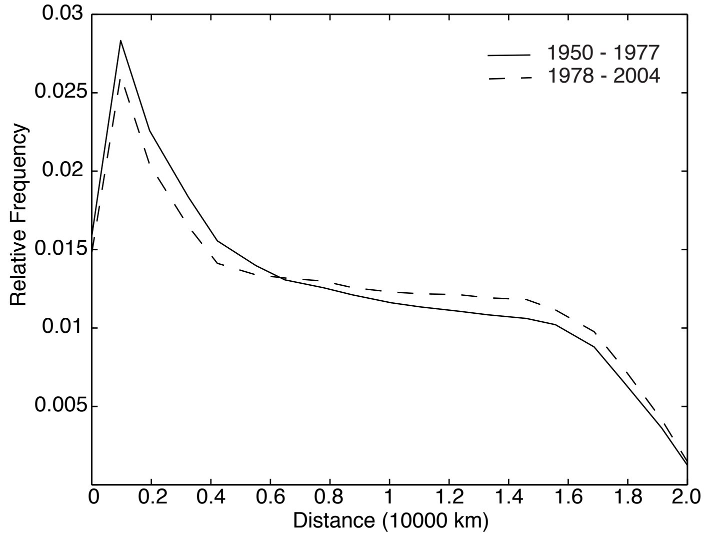
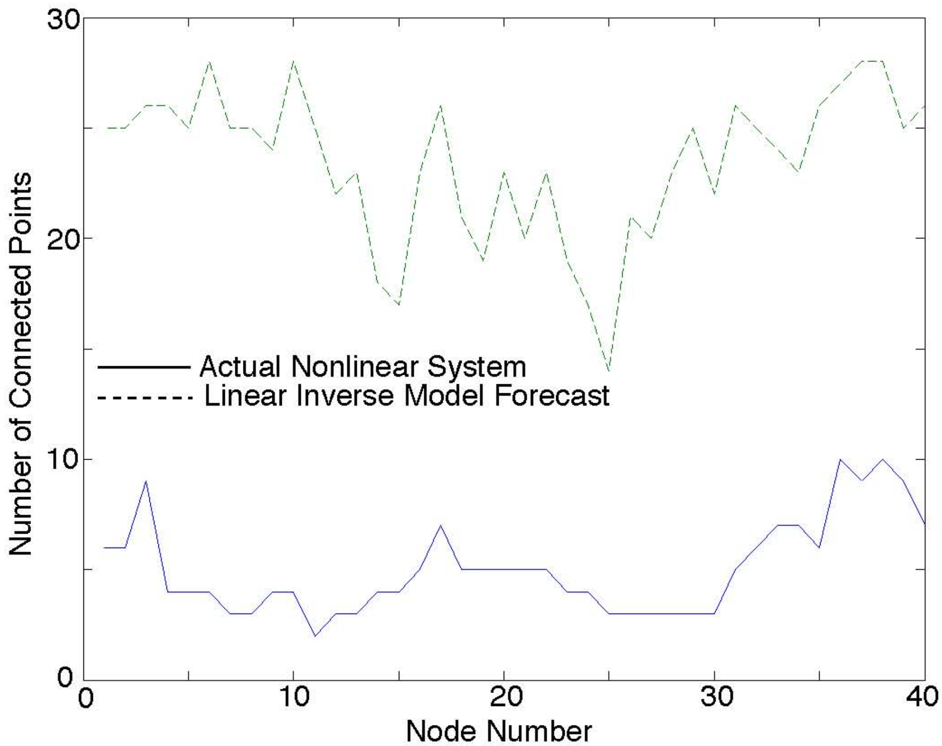

80° N
60° N
40° N
0.5
0.475
0.45
0.425
0.4
0.375
0.35
0.325
0.3
0.275
0.25
0.225
0.2
0.175
0.15
0.125
0.1
0.075
0.05
0.025
0

Figure 9



0.03
0.025
0.02
0.015
0.01
0.005
Relative Frequency
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2.0
Distance (10000 km)
1950 - 1977
1978 - 2004



Number of Connected Points
Node Number
Actual Nonlinear System
Linear Inverse Model Forecast
0
10
20
30
40

Figure 11