Geophysical Research Abstracts
Vol. 14, EGU2012-5073, 2012
EGU General Assembly 2012

Graph theoretical analysis of climate data

T. Zerenner and A. Hense
Meteorological Institute, University of Bonn, Germany (tanja.zerenner@yahoo.de)

Applying methods from graph and network theory to climatological data is a quite new approach and contains numerous difficulties. The atmosphere is a high dimensional and complex dynamical system which per se does not show a network-like structure. It does not consist of well-defined nodes and edges. Thus considering such a system as a network or graph inevitably involves radical simplifications and ambiguities. Nevertheless network analysis has provided useful results for different kinds of complex systems for example in biology or medical science (neural and gene interaction networks). The application of these methods on climate data provides interesting results as well. If the network construction is based on the correlation matrix of the underlying data, the resulting network structures show many well known patterns and characteristics of the atmospheric circulation (Tsonis et al. 2006, Donges et al. 2009). The interpretation of these network structures is yet questionable. Using Pearson Correlation for network construction does not allow to differ between direct and indirect dependencies. An edge does not necessarily represent a causal connection. An interpretation of these structures for instance concerning the stability of the climate system is therefore doubtful.

Gene interaction networks for example are often constructed using partial correlations (Wu et al. 2003), which makes it possible to distinguish between direct and indirect dependencies. Although a high value of partial correlation does not guarantee causality it is a step in the direction of measuring causal dependencies. This approach is known as Gaussian Graphical Models, GGMs. For high dimensional datasets such as climate data partial correlations can be obtained by calculating the precision matrix, the inverse covariance matrix. Since the maximum likelihood estimates of covariance matrices of climate datasets are singular the precision matrices can only be estimated for example by using the Graphical Lasso Algorithm, Glasso (Friedman et al. 2007). If "climate networks" are constructed in terms of GGMs and glasso, they exhibit apparently trivial structures. In this type of networks only geographically adjacent nodes are connected by edges.

These results bring up several questions which will be discussed in the talk. Are the results of the standard as well as the GGM analysis mere results of the methods? How do large scale structures arise when direct dynamical relations are restricted to geographically nearest neighbours? Is it actually possible to describe the complex dynamical interactions of the atmospheric circulation as a network structure without including additional information about physical relations between the nodes or does this method imply too heavy simplifications to describe the complex system of the atmosphere appropriately?