Geophysical Research Abstracts
Vol. 13, EGU2011-11552, 2011
EGU General Assembly 2011

Information-theoretic methods to find possible causalities and relevant time lags between multiple time series

Jakob Runge (1,2), Jobst Heitzig (1), Jürgen Kurths (1,2)

(1) Potsdam Institute for Climate Impact Research, Research Domain IV: Transdisciplinary Concepts and Methods, Potsdam, Germany (jakobrunge@gmail.com), (2) Institute of Physics, Humboldt University Berlin

When looking for physical mechanisms that cause a climate phenomenon like the Indian Monsoon, typically a large number of variables have to be considered, and couplings might involve time-lags and cannot be assumed to be linear a priori. The aim of this paper is to develop a statistical method to identify possible causal relationships and relevant time-lags between multiple (e.g., geophysical) time series under the following conditions. When using only pairwise statistical associations, many of these observed correlations will be due to common drivers or reflect only indirect couplings. In order to help the search for physical mechanisms that underlie the observed associations, one thus needs methods to narrow down the space of possible explanatory variables and time-lags. We therefore analyze lagged coupling measures and seek to detect the precise coupling delay and strength and whether a link is only indirect or due to common drivers.

Common approaches based on Granger causality or Bayesian networks usually make assumptions about the underlying dynamics by invoking models. But we wish to keep the analysis as general as possible to account for a large class of causal mechanisms. Information theory offers this general framework, and Schreibers transfer entropy has been widely applied in many fields. But transfer entropy is a bivariate measure and is also not intended to yield information about the precise coupling delay. Our coupling measure is based on lagged conditional mutual information with conditions chosen such that misleading effects of auto-dependencies within each time series and the influence of other possible explanatory variables are excluded. Thus the resulting network of significant couplings contains only statistical relationships that remain after other factors of influence have been excluded. This is not to say that the resulting network yields the true causal links as there still can be unobserved variables.

To estimate coupling measures, one faces the problem of high dimensional probability distributions, which will be also be addressed in this paper. Furthermore, we invoked appropriate statistical significance tests to analyze the robustness of our results. To illustrate the method, we use some model systems. Finally, we apply the method to some climate indices. The results give insights into the coupling structure of climatic processes with precise delays and thus help to find physical mechanisms that underlie this information flow.