EGU2020-9270
https://doi.org/10.5194/egusphere-egu2020-9270
EGU General Assembly 2020

Causal Discovery for Climate Time Series in the Presence of Unobserved Variables

Andreas Gerhardus and Jakob Runge
German Aerospace Center (DLR), Institute of Data Science, Jena, Germany

Scientific inquiry seeks to understand natural phenomena by understanding their underlying processes, i.e., by identifying cause and effect. In addition to mere scientific curiosity, an understanding of cause and effect relationships is necessary to predict the effect of changing dynamical regimes and for the attribution of extreme events to potential causes. It is thus an important question to ask how, in cases where controlled experiments are not feasible, causation can still be inferred from the statistical dependencies in observed time series.

A central obstacle for such an inference is the potential existence of unobserved causally relevant variables. Arguably, this is more likely to be the case than not, for example unmeasured deep oceanic variables in atmospheric processes. Unobserved variables can act as confounders (meaning they are a common cause of two or more observed variables) and thus introduce spurious, i.e., non-causal dependencies. Despite these complications, the last three decades have seen the development of so-called causal discovery algorithms (an example being FCI by Spirtes et al., 1999) that are often able to identify spurious associations and to distinguish them from genuine causation. This opens the possibility for a data-driven approach to infer cause and effect relationships among climate variables, thereby contributing to a better understanding of Earth's complex climate system.

These methods are, however, not yet well adapted to some specific challenges that climate time series often come with, e.g. strong autocorrelation, time lags and nonlinearities. To close this methodological gap, we generalize the ideas of the recent PCMCI causal discovery algorithm (Runge et al., 2019) to time series where unobserved causally relevant variables may exist (in contrast, PCMCI made the assumption of no confounding). Further, we present preliminary applications to modes of climate variability.