EGU22-6958
https://doi.org/10.5194/egusphere-egu22-6958
EGU General Assembly 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Causal Discovery in Ensembles of Climate Time Series

Andreas Gerhardus[1] and Jakob Runge[2,3]
[1]German Aerospace Center (DLR), Institute of Data Science, Jena, Germany (andreas.gerhardus@dlr.de)
[2]German Aerospace Center (DLR), Institute of Data Science, Jena, Germany (jakob.runge@dlr.de)
[3]Technische Universität Berlin, 10623 Berlin, Germany

Understanding the cause and effect relationships that govern natural phenomena is central to the scientific inquiry. While being the gold standard for inferring causal relationships, there are many scenarios in which controlled experiments are not possible. This is for example the case for most aspects of Earth's complex climate system. Causal relationships then have to be learned from statistical dependencies in observational data, a task that is commonly referred to as (observational) causal discovery.

When applied to time series data for learning causal relationships in dynamical systems, methods for causal discovery face additional statistical challenges. This is so because, as licensed by an assumption of stationarity, samples are taken in a sliding window fashion and hence autocorrelated rather than iid. Moreover, strong autocorrelations also often occlude other relevant causal links. The recent PCMCI algorithm (Runge et al., 2019) and its variants PCMCI+ (Runge, 2020) and LPCMCI (Gerhardus and Runge, 2020) address and to some extent alleviate theses issues.

In this contribution we present the Ensemble-PCMCI method, an adaption of PCMCI (and its variants PCMCI+ and LPCMCI) to cases in which the data comprises several time series, i.e., measurements of several instances of the same underlying dynamical system. Samples can then be taken from these different time series instead of a in a sliding window fashion, thus avoiding the issue of autocorrelation and also allowing to relax the stationarity assumption. In particular, this opens the possibility to analyze temporal changes in the underlying causal mechanisms. A potential domain of application are ensemble forecasts.

Related references:
Jakob Runge et al. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances 5 eaau4996.

Jakob Runge (2020). Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. In Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI). Proceedings of Machine Learning Research 124 1388–1397. PMLR.

Andreas Gerhardus and Jakob Runge (2020). High-recall causal discovery for autocorrelated time series with latent confounders. In Advances in Neural Information Processing Systems 33

12615–12625. Curran Associates, Inc.