EGU24-22158, updated on 08 Oct 2024
https://doi.org/10.5194/egusphere-egu24-22158
EGU General Assembly 2024

Spatiotemporal Causal Effect Estimation

Rebecca Herman[1] and Jakob Runge[1,2]
[1]Deutsches Zentrum für Luft- und Raumfahrt
[2]Technische Universität Berlin

Causal discovery and effect estimation for time series provide scientists with a way to extract causal information from observational studies when possible. But the high dimensionality of raw climate data causes computational problems for most analysis methods, and causal inference is no exception. To address this problem, climate scientists usually pre-process climate data using dimension reduction techniques (including seasonal and regional averaging and principle component analysis) that may result in the loss of valuable information before the true analysis even begins. For example, climate scientists often represent El Niño Southern Oscillation variability (ENSO) using the uni-variate Nino3.4 index, which cannot distinguish between central Pacific and eastern Pacific El Niño events, which are believed to impact global climate varaibility in different ways. This study introduces a method for avoiding premature data dimension reduction in causal effect estimation, implemented in tigramite. The method allows the researcher to define multi-variate climate indices, reducing the dimensionality of the causal inference problem via the causal assumptions instead of losing information from the data itself. To investigate the performance of this approach on climate data, we examine the effect of ENSO on the North Atlantic Oscillation (NAO) in simulated data from the Coupled Model Intercomparison Project, phase 6. We choose this as our case study because different types of El Nino are believed to have very different effects on NAO, to the extent that the impact may be completely undetectable in observations when no distinction between the types of ENSO is made. By comparing estimated effects using uni- and multi-variate climate indices, we demonstrate that this method retains valuable information that would be lost in uni-variate analysis, and make recommendations for best practices when using multi-variate climate indices in causal effect estimation.