EGU General Assembly 2024

EGU24-14429, updated on 08 Oct 2024
https://doi.org/10.5194/egusphere-egu24-14429
EGU General Assembly 2024

Compression-complexity based estimation of Causality: Applications in Earth and Climate Sciences

Aditi Kathpalia, Pouya Manshour, and Milan Paluš
Institute of Computer Science of the Czech Academy of Sciences, Complex Systems, Czechia (kathpaliaaditi@gmail.com)

Many approaches to time series causality exist and have been inspired from fields such as statistics, information theory, physics and topology. We have proposed a method called compression-complexity causality (CCC) [1] inspired from the field of data compression in computer science. It is based on the idea that the compressibility of the 'effect' time series changes when the 'cause' time series is considered in the evolution of the future dynamics of the effect. Compressibility is estimated using compression-complexity estimator for time series called 'effort-to-compress', which employs a lossless data compression algorithm for complexity estimation. CCC makes minimal assumptions on given time series data and has been seen to work well for short length data, irregularly sampled data as well as data with low temporal resolution. We have also introduced a multidimensional version of CCC, called Permutation CCC (PCCC) [2], which uses Takens' embedding for appropriate high dimensional representation of time series. This representation is subsequently encoded using ordinal patterns before computation of CCC. PCCC formulation retains the original robustness of CCC. This is demonstrated with its application on simulated multidimensional systems. We apply this formulation to infer causality between CO_2 emissions – temperature recordings on three different time scales, El Niño–Southern Oscillation phenomena – South Asian Summer Monsoon on two different time scales, as well as North Atlantic Oscillations – European temperature recordings on two different time scales. These paleoclimate and climate datasets suffer from the issues of missing samples, low temporal resolution and short length data and so a reliable inference of these climatic interactions requires a robust causality estimator.
Finally, we explore another variation of CCC which can help to infer causality in the multivariate cases. This variation helps to infer the existence of causal influences to a particular variable (from each other variable considered) while conditioning out the additional variables present. The presence of causal influences to each variable is decided by choosing the model of least compression-complexity which can help to explain the evolution of the future of that particular variable. In case more than one model has least complexity, the smallest model is chosen. We apply this formulation to understand interactions in space-weather system, particularly the solar wind-magnetosphere-ionosphere system interactions, which manifest as geomagnetic storms and substorms. We compare the performance of CCC formulations with existing methods in case of simulations as well as real data applications.

This study is supported by the Czech Academy of Sciences, Praemium Academiae awarded to M.

Paluš.

References:
[1] Kathpalia, A., & Nagaraj, N. (2019). Data-based intervention approach for Complexity-Causality measure. PeerJ Computer Science, 5, e196.
[2] Kathpalia, A., Manshour, P., & Paluš, M. (2022). Compression complexity with ordinal patterns for robust causal inference in irregularly sampled time series. Scientific Reports, 12(1), 14170.