



## Comment on hess-2021-445

Anonymous Referee #3

---

Referee comment on "Detecting hydrological connectivity using causal inference from time-series: synthetic and real karstic study cases" by Damien Delforge et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-445-RC4>, 2021

---

The authors compared four causal inference methods in a hydrological issue, which is probably the first method-comparison study on revealing causality in hydrology. The following four methods are used: cross-correlation, convergent cross mapping, partial correlation-based PCMCI, and conditional mutual information-based PCMCI. However, there are two main issues that concern me (as described below). Therefore, I recommend a major revision.

**Comparison between CCM and CMI-based PCMCI.** The current comparison based on noisy data is unfair to CCM, because CCM is more suitable for deterministic dynamics and does not work well in a stochastic system. Also, the authors used different levels of noises in the synthetic study. Still, only the averaged results are reported in Figure 3, and the noise impact on the performances of the four methods remains unknown. Therefore, I suggest, at least in the synthetic case study, performing the comparison based on a noise-free/deterministic system and a thorough evaluation of the noise impact.

**Limited data points for computing CMI.** In the real case study, the inferred causality from CMI-based PCMCI is much less trustable, given only 465 datapoints of 7 variables (what is the maximum allowed number of conditioned variables set in PCMCI by the way?). In fact, it is somehow expected that the CMI-based PCMCI does not work well using this limited dataset (even for a three-dimensional CMI estimation, several hundred data points might not be sufficient). Although the authors acknowledged this limitation, I strongly recommend a corresponding synthetic study to evaluate the impact of dataset size and the number of variables in CMI-based PCMCI, which is very critical to guide the current and future causality analysis in earth science inferred by the PCMCI algorithm.

### Other minor revisions/comments:

- Lines 67 and 68: Please spell out ParCorr and CMI.
- Line 144: "entropy transfer" -> "transfer entropy"
- Lines 147 and 148: "The nearest-neighbor estimator is recommended for time-series below 1000 samples" ... under what dimensionality?
- Line 238:  $Q_b$  ->  $Q_B$

- Line 356: "constraint causal inference" -> "constrain causal inference"