

Comparing deterministic and stochastic methods to infer causal Earth system interactions

Niclas Schilling^{1,2}, Ingo Fetzer^{1,3}, Kira Rehfeld^{2,4,5}, and Hannah Zoller¹



1. Introduction

- Understanding interactions among Earth system processes is essential for mitigating cascading Anthropocene risks and avoiding critical tipping points
- To detect such interactions from observational time series, **causal inference methods** are increasingly applied to complex, nonlinear systems
- We compare two tools from different paradigms: **Convergent Cross Mapping**¹ (CCM), a state-space approach for deterministic dynamics, and **Transfer Entropy**² (TE), an information-theoretic approach for coupled stochastic processes
- Our goal:** Determine **how reliably the methods detect interactions** under typical challenges such as noise, strong coupling and latent drivers



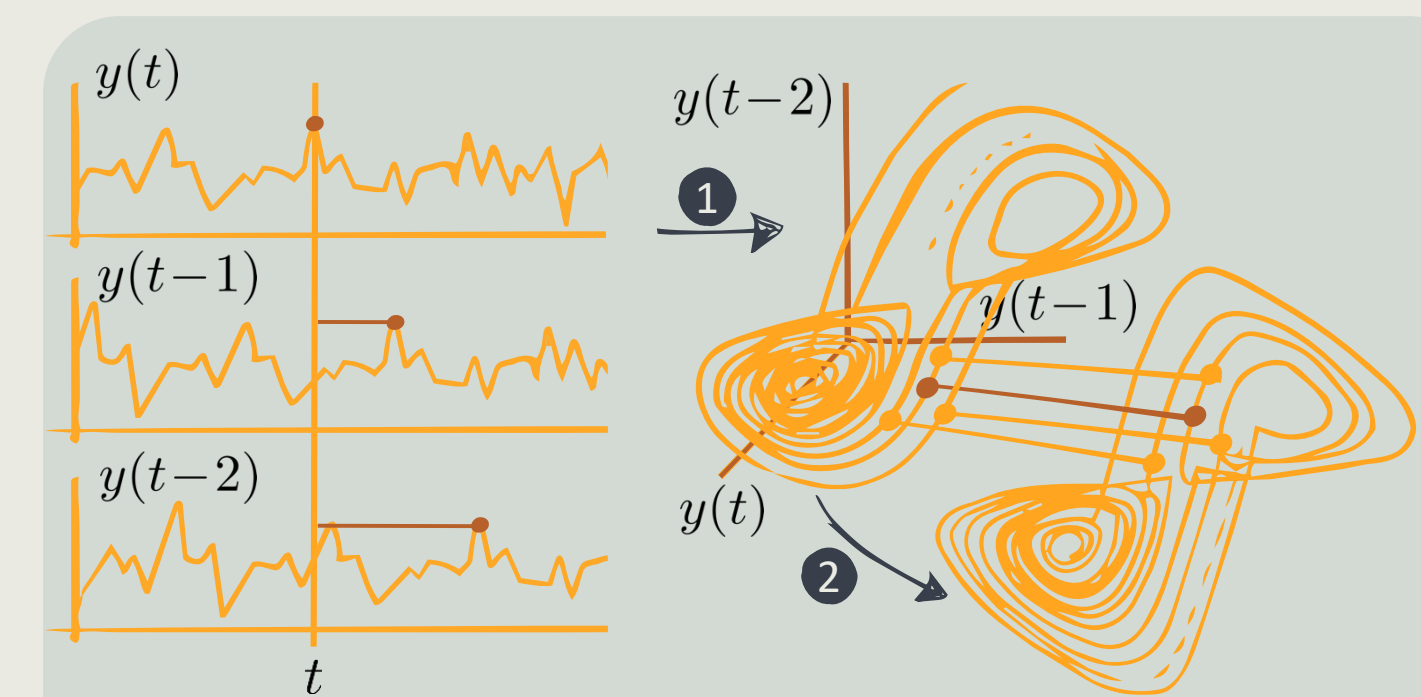
2.1 Methods – In a Nutshell

Given: Two observed time series $x(t) = x_1, \dots, x_n$ and $y(t) = y_1, \dots, y_n$ of variables X and Y
Question: Can we detect a directed, nonlinear coupling from X to Y ?

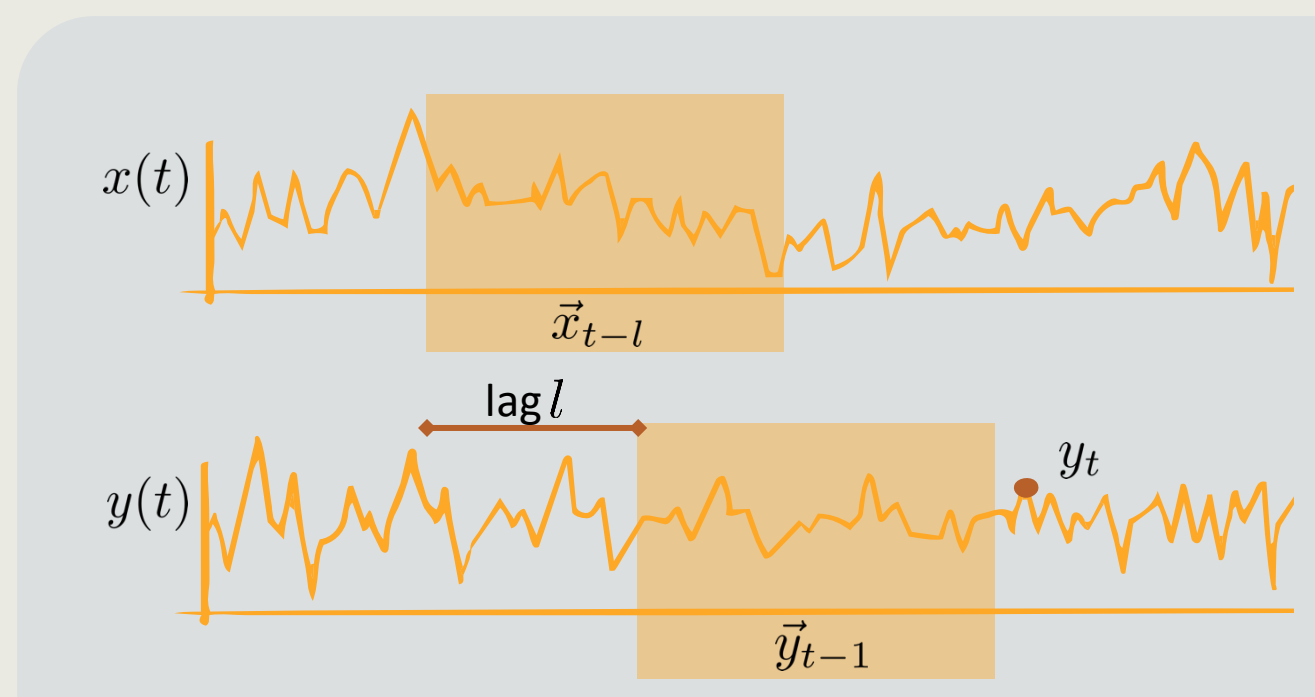
Convergent Cross Mapping (CCM)	Transfer Entropy (TE)
If $X \rightarrow Y$, then the reconstructed state-space of Y contains information about X	Main Idea If $X \rightarrow Y$, then X 's past provides information about Y 's future (beyond Y 's past)
(Partially) deterministic, stationary, non-separable systems	Suitable Systems (Partially) stochastic, stationary, separable systems
Link detected if CCM score peaks at lag > 0 (i.e. past $x(t-l)$ influences $y(t)$) ³ and is significant against reshuffled $x(t)$	Detection Criteria Link detected if TE is significant against scores from reshuffled $x(t)$ (surrogate test)
Partial Cross Mapping ⁴ (PCM) tests whether X 's signature in Y 's state-space persists after accounting for indirect paths via Z	Conditional Extensions Conditional TE ⁵ (cTE) tests whether X 's past retains predictive power over Y 's future after conditioning also on Z 's past

2.2 Methods – Deep Dive

Steps to detect an interaction $X(t-l) \rightarrow Y(t)$ at lag l



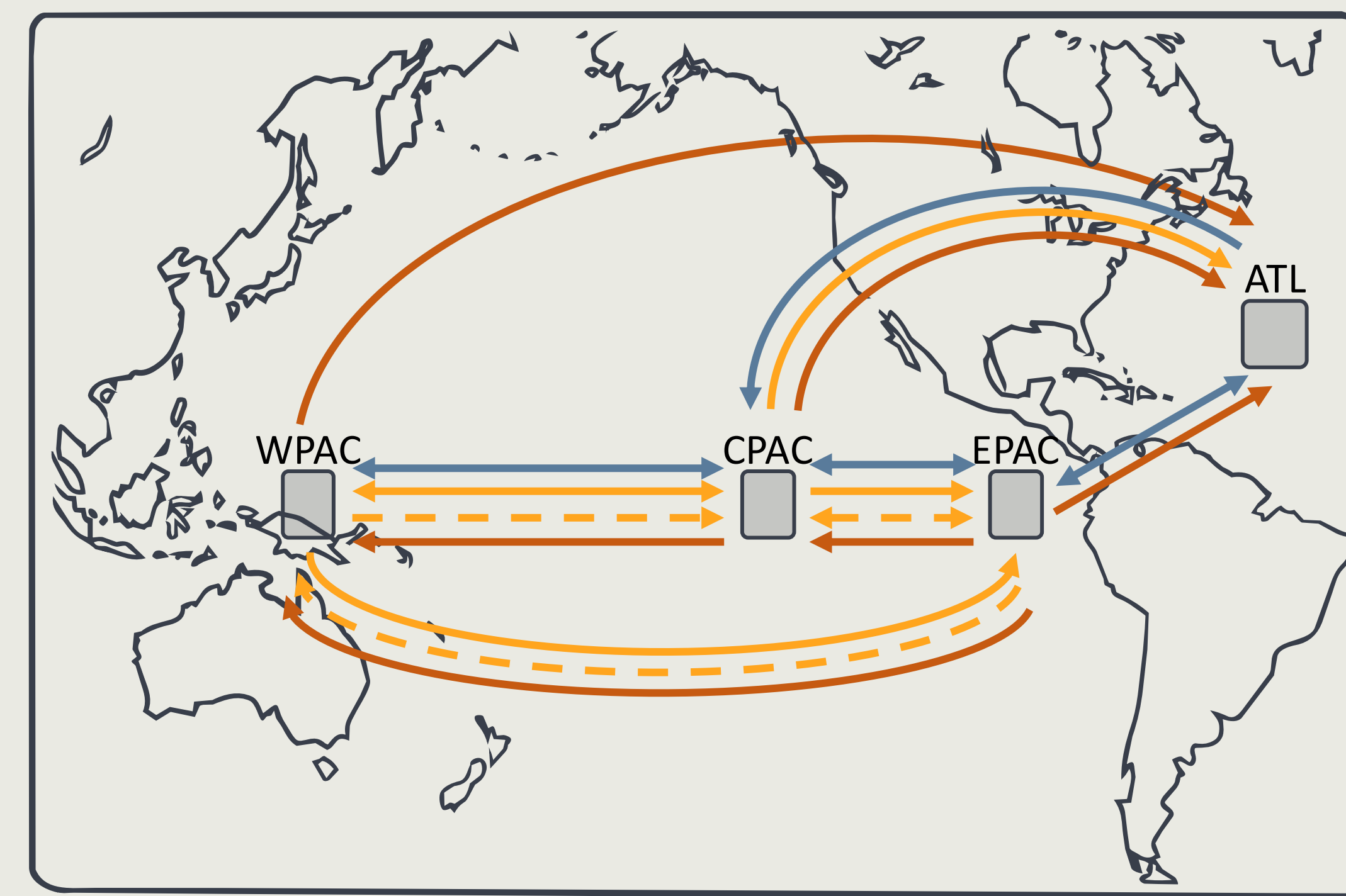
- Reconstruct manifold M_Y
 $\vec{y}_i = [y_i, y_{i-1}, \dots, y_{i-(E-1)}]$
 $M_Y = \{\vec{y}_i\}$ with $|M_Y| = L$
- Estimate x_{t-l} from neighbours of \vec{y}_t
 $\hat{x}_{t-l} = \sum_i w_i x_{i-l}$
 $w_i \propto \exp(-\text{dist}(\vec{y}_t, \vec{y}_i))$
- Quantify predictive score
 $\rho_{X \rightarrow Y} = \text{CORR}(x_{t-l}, \hat{x}_{t-l})$
- Lag-scan**³: Determine ρ at different positive and negative (causal and non-causal) lags l



- Define history vectors
 $\vec{y}_{t-1} = [y_{t-1}, y_{t-2}, \dots, y_{t-k}]$
 $\vec{x}_{t-l} = [x_{t-l}, x_{t-(l+1)}, \dots, x_{t-(l+k-1)}]$
- Calculate transfer entropy T
 $T_{X \rightarrow Y} = H(y_t | \vec{y}_{t-1}) - H(y_t | \vec{y}_{t-1}, \vec{x}_{t-l})$
with $H(y|x) = - \sum_{x,y} p(x,y) \log p(y|x)$
- Lag-scan**⁶: determine T at different positive (causal) lags l

3. Case Study: The Walker Circulation

- The Walker circulation is a **large-scale overturning cell in the tropical Pacific**, with warm air rising in the west & cooler air descending in the east
- Its variability is tightly coupled to El Niño and La Niña events, making it **crucial for understanding climate** and extreme weather worldwide
- Its well-understood dynamics make it a **suitable benchmark** for causal inference methods⁷
- We **analyse four variables**: West Pacific surface pressure (WPAC) and sea surface temperatures in the Central Pacific (CPAC), East Pacific (EPAC), and Atlantic (ATL) — to also probe the Pacific–Atlantic teleconnection



Expected links (La Niña/Neutral) Detected by CCM + PCM
Expected links (El Niño) Detected by TE + cTE

From the analysis, several hypotheses can be generated

- In bidirectional interactions, CCM detects only the dominant direction
- TE detects bidirectionality, but may misidentify directions in one-way links
- PCM does not reliably rule out indirect links
- Benchmarking CCM against cross-correlation gives a conservative criterion
- Both methods have difficulties in accurately estimating interaction lags

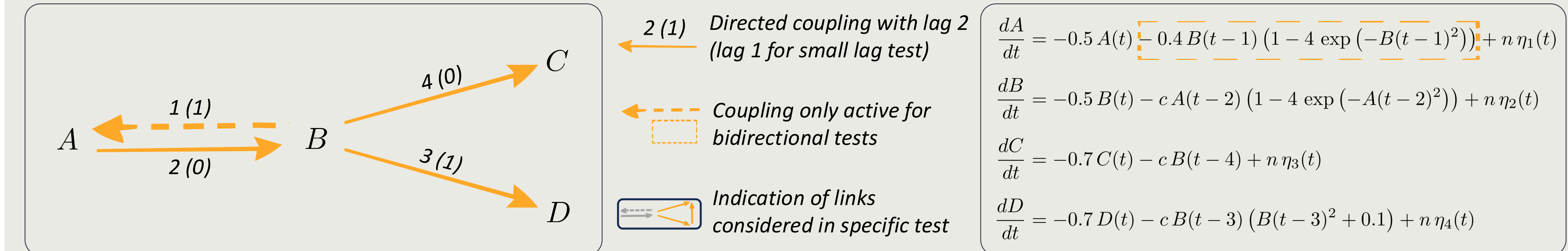
Hypotheses I-IV will be addressed in the results

5. Conclusion & Outlook

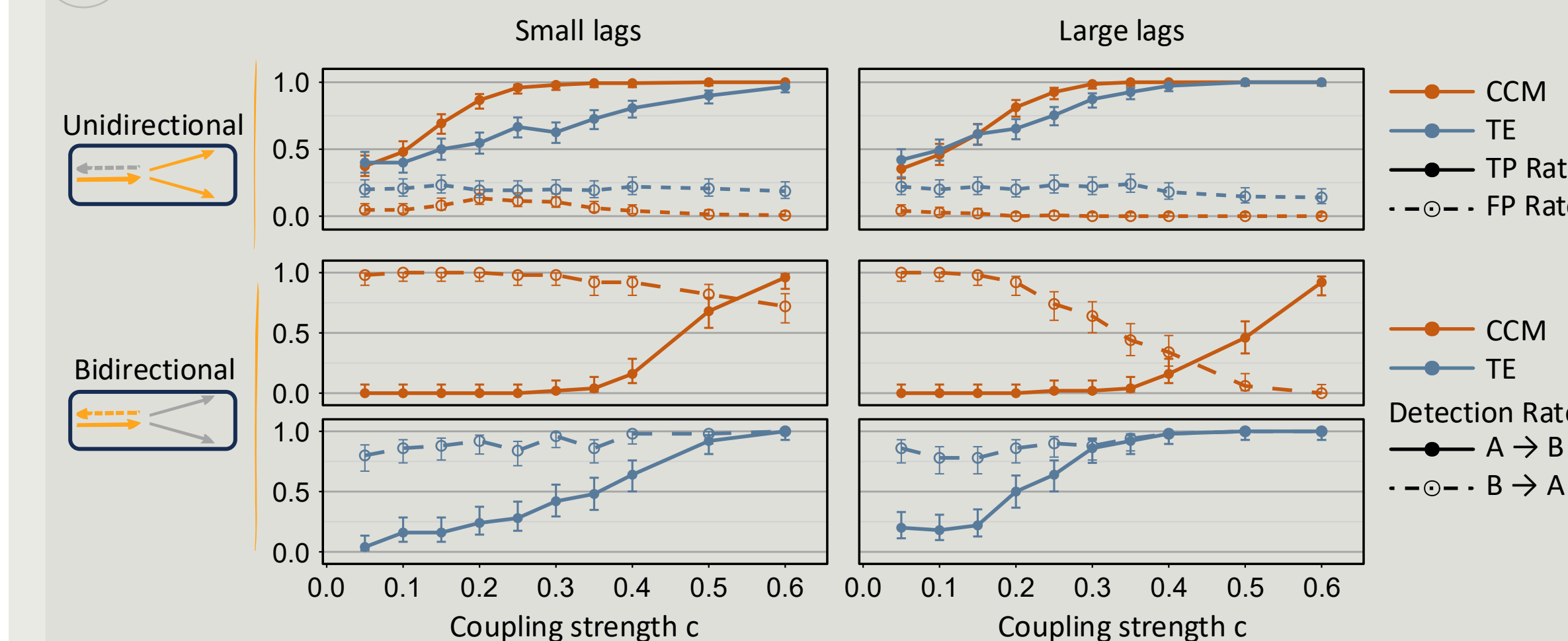
- Complementary rather than interchangeable:** CCM is more reliable for unidirectional links, while only TE consistently captures bidirectional interactions
- An open question:** A systematic evaluation of whether the methods recover not just the direction, but also correct interaction lags remains to be tested
- Quantifying interaction strengths:** The consistent scaling of TE and CCM scores with coupling strength suggests a promising pathway to observation-based estimates of Earth system couplings
- Potential application:** The Earth System Impact (ESI) score⁸ quantifies human impacts on the environment by accounting for climate–land–water interactions; coupling estimates could complement the model-based interaction strengths currently used

4. Results

Evaluating CCM and TE on a synthetic dynamical system – simulated via stochastic delay differential equations

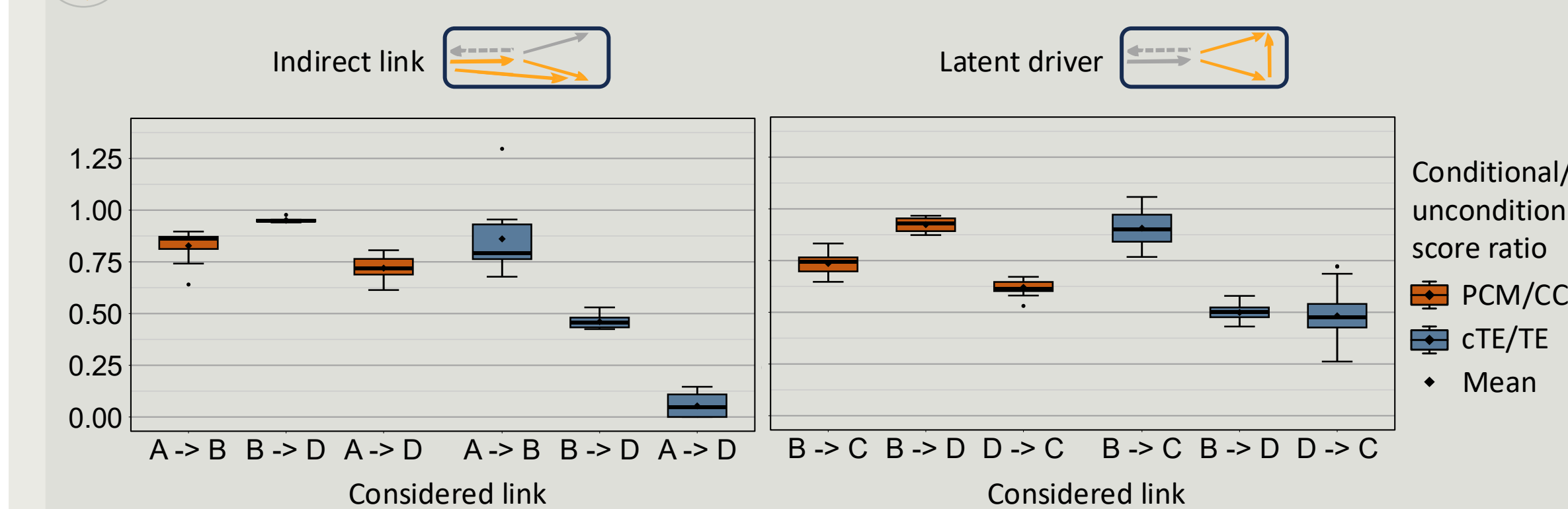


CCM misses bidirectional links, TE detects them consistently



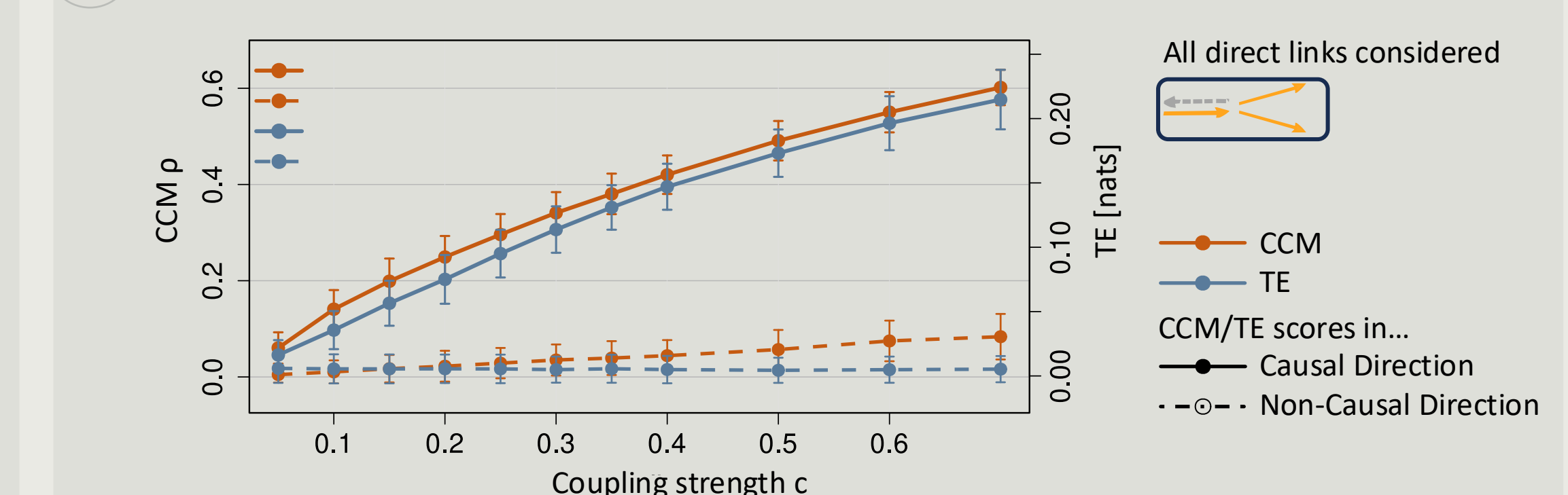
- CCM works well for unidirectional interactions, but often only finds the stronger direction for bidirectional interactions
- TE works more consistently across uni-/bidirectional links as well as short and long lags, which comes with a trade-off in detection accuracy

Spurious links shrink under conditioning – but don't vanish



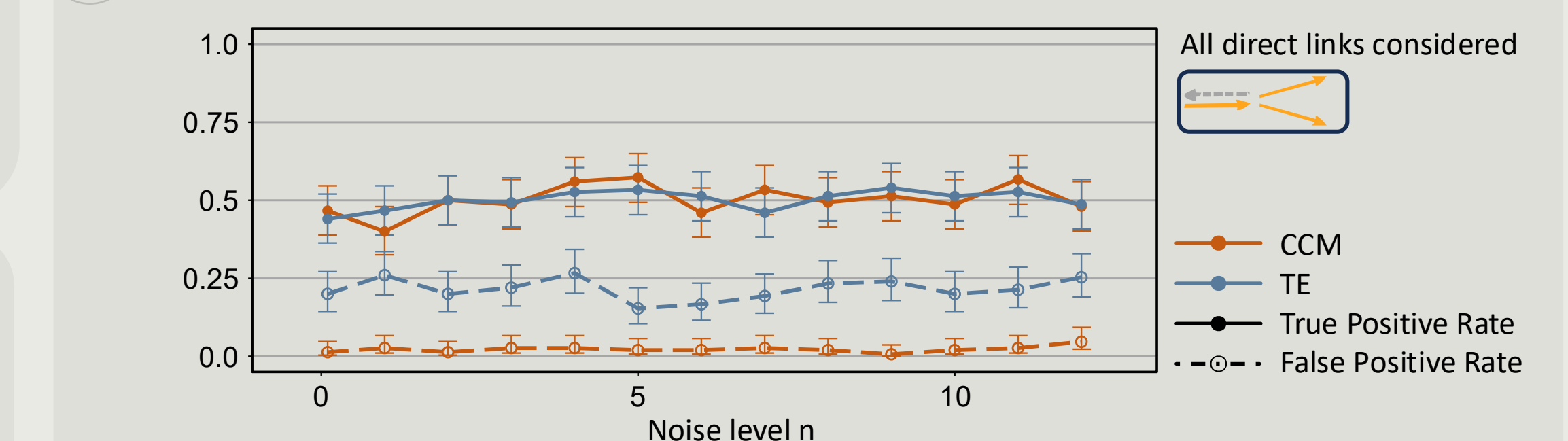
- Conditioning reduces spurious link scores more strongly than true direct links; the pattern is clearer for cTE and for indirect links
- In binary testing for (non-)detection of certain links PCM/cTE do not show any significant improvements compared to CCM/TE

Both methods scale similarly with increasing coupling strength



- Despite their different absolute scales, CCM and TE scores show similar relative increases with coupling strength

Comparable detection across the tested noise range



- Across the considered noise levels, both methods detect links at similar rates; a separate analysis indicates average CCM/TE scores rise at low noise, then plateau, without yet decreasing at high noise

Alternative CCM detection criteria and trade-offs

- Comparing CCM scores against cross-correlation, sometimes used as detection criterion, fails to separate causal and non-causal links
- A looser CCM detection criterion — accepting a link if there exists a significant CCM score at any causal lag (lag > 0) — recovers both directions in bidirectional systems, but also creates false positives under strong coupling or small lags

¹ Stockholm Resilience Centre, Stockholm University, Sweden

² Department of Physics, University of Tübingen, Germany

³ Bolín Centre for Climate Research, Stockholm University, Sweden

⁴ Department of Geosciences, University of Tübingen, Germany

⁵ TERRA Cluster of Excellence, University of Tübingen, Germany

References

[1] Sugihara G et al. (2012) Science 338: 496–500 <https://doi.org/10.1126/science.1227079>
 [2] Schreiber T (2000) Physical Review Letters 85: 461–464 <https://doi.org/10.1103/PhysRevLett.85.461>
 [3] Ye H et al. (2015) Scientific Reports 5: 14750 <https://doi.org/10.1038/srep14750>
 [4] Leng S et al. (2020) Nature Communications 11: 2632 <https://doi.org/10.1038/s41467-020-16238-0>
 [5] Lizier JT et al. (2008) Physical Review E 77: 026110 <https://doi.org/10.1103/PhysRevE.77.026110>
 [6] Wibral M et al. (2013) PLOS ONE 8: e55809 <https://doi.org/10.1371/journal.pone.0055809>
 [7] Runge J et al. (2019) Science Advances 5: eaau4996 <https://doi.org/10.1126/sciadv.aau4996>
 [8] Lade SJ et al. (2021) Environmental Research Letters 16: 115005 <https://doi.org/10.1088/1748-9326/ac2db1>

Niclas Schilling
University of Tübingen
niclas.schilling@student.uni-tuebingen.de

