




# Comments on identifying causal relationships in nonlinear dynamical systems via empirical mode decomposition

 Chun-Wei Chang <sup>1,2</sup>, Stephan B. Munch<sup>3</sup> & Chih-hao Hsieh <sup>1,2,4,5</sup> 


 ARISING FROM Albert C. Yang et al. *Nature Communications* <https://doi.org/10.1038/s41467-018-05845-7> (2018)

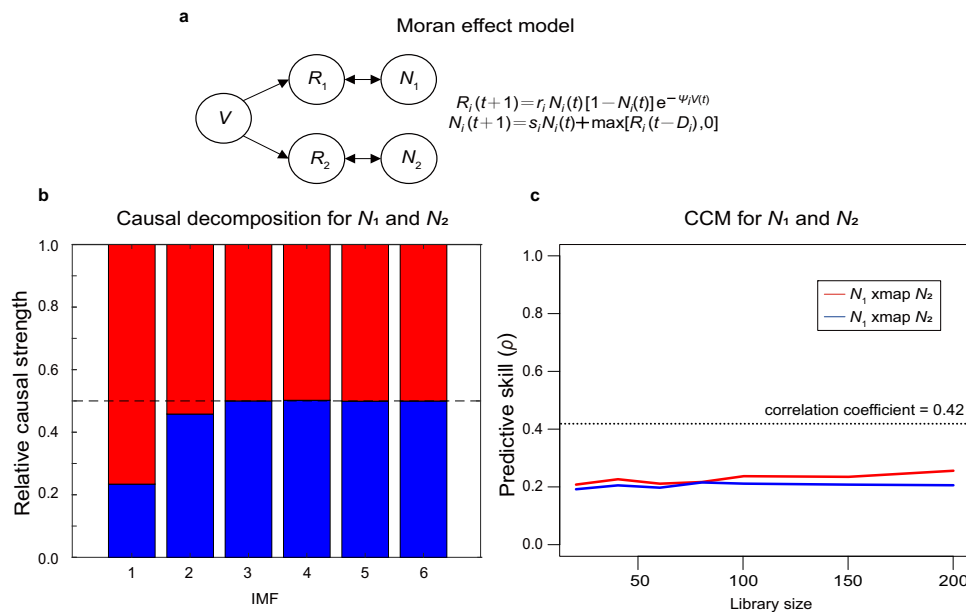
The empirical mode decomposition (EMD) method proposed in Yang et al.<sup>1</sup> fails to correctly identify causal relationships for a system of two independent variables driven by a shared external forcing (aka Moran effect). Using a simple, two-species Moran effect model (Fig. 1a), it is obvious that the EMD method erroneously concludes that  $N_1$  and  $N_2$  have a causal relationship (IMF 1 in Fig. 1b), although in fact they do not. This is because effects of external forcing were recorded in both  $N_1$  and  $N_2$  time series, and at least one IMF decomposed from  $N_1$  and  $N_2$  time series is associated with that external forcing. Therefore, removal of such IMF from  $N_1$  makes the remaining less coherent with  $N_2$ , and vice versa. As such, EMD methods based on diminished coherence due to IMF removals fail to falsify spurious causations caused by sharing external forcing. In contrast, convergent cross mapping (CCM) correctly identifies the lack-of causal relationship between  $N_1$  and  $N_2$  (Fig. 1c). The efficacy of CCM to distinguish Moran effects depends on the strength of the shared external forcing. In extreme cases, when external forcing is too strong,  $N_1$  and  $N_2$  are synchronized (the pathological case, as noted in a previous study<sup>2</sup>). Methods to cope with such situation have also been developed<sup>3,4</sup>.

Second, Yang et al.<sup>1</sup> argued that CCM provides incorrect causal relationships. However, Yang et al.<sup>1</sup> used CCM in a manner that the original authors<sup>2</sup> did not intend, producing incorrect conclusions. Specifically, following McCracken et al.<sup>5</sup>, Yang et al.<sup>1</sup> used correlation difference ( $X$  cross-map  $Y - Y$  cross-map  $X$ ) as the definition of CCM causation (Fig. 3 in ref. <sup>1</sup>) without examining the convergence of the cross-mapping skill; this is an incorrect definition. The correct definition of causation under CCM is improvement of cross-mapping skill with increasing time series length (i.e., convergence). In addition, as in real systems, CCM causation can be bidirectional<sup>2</sup>. Yang et al.<sup>1</sup> used an incorrect definition (i.e.,

correlation difference) and incorrectly concluded that CCM misidentified the lynx versus hare and *Didinium* versus *Paramecium* interactions as top-down control systems and the Lotka Volterra predator–prey model and wolf versus moose interactions as either no or confusing causation. However, in each of these examples, CCM exhibits clear convergence with increasing library size in both directions. By using the convergence definition of causation in CCM, we concluded that these prey–predator systems exhibited bidirectional causation. That is, CCM correctly identified the reciprocal nature of predator–prey interactions in all of these systems (Fig. 5 in ref. <sup>1</sup>). As such, we suggest any description of coupling in predator–prey systems as “directional” claimed by Yang et al.<sup>1</sup> may be misleading (Fig. 3 in ref. <sup>1</sup>) because predators causally influence prey by consuming them and prey causally influence predators by providing them the energy needed for population growth. Moreover, the relative strength of each direction can be quantified based on the rate of convergence<sup>2,6</sup>, with proper consideration of potential lagged effects<sup>3</sup>. We also disagree with the claim that CCM incorrectly identified causal coupling in white noise. Again, this also stems from using an incorrect definition of CCM. In contrast, we find no evidence of convergence when applying CCM to paired white noise signals (Fig. 2a) and a false positive rate consistent with  $p = 0.05$  as the level of significance (Fig. 2b, c).

Several additional misunderstandings about CCM in Yang et al.<sup>1</sup> warrant clarification. (1) CCM does not rely on predictability as the criterion. Rather, CCM relies on information recovering<sup>2,6</sup> that identifies whether the present state of an effect variable contains information about the present state of causal variables<sup>7</sup> (i.e., nowcast) and thus enables CCM to identify simultaneous influences. (2) Oscillatory dynamics may confound the efficacy of CCM; however, methods to

<sup>1</sup>National Center for Theoretical Sciences, Taipei 10617, Taiwan. <sup>2</sup>Research Center for Environmental Changes, Academia Sinica, Taipei 11529, Taiwan. <sup>3</sup>Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA. <sup>4</sup>Institute of Oceanography, National Taiwan University, Taipei 10617, Taiwan. <sup>5</sup>Institute of Ecology and Evolutionary Biology, Department of Life Sciences, National Taiwan University, Taipei 10617, Taiwan. email: [chsieh@ntu.edu.tw](mailto:chsieh@ntu.edu.tw)



**Fig. 1 Causal decomposition fails to falsify spurious causations presented in Moran effect model.** **a** Moran effect model is a 5-variate difference equation model in which variables  $N_1$  and  $N_2$  have no causal relationship, but have a significant correlation in their time series due to shared external forcing,  $V$ . We ran this model for 10,000 time steps with the parameter set [ $r_1 = 3.4$ ,  $r_2 = 2.9$ ,  $\psi_1 = 0.5$ ,  $\psi_2 = 0.6$ ,  $s_1 = 0.4$ ,  $s_2 = 0.35$ ,  $D_1 = 3$ ,  $D_2 = 3$ ,  $R_1(0) = R_2(0) = 1$ ,  $N_1(0) = N_2(0) = 0.5$ ], but retaining only the last 200 steps for analysis. Because of the strong correlation between  $N_1$  and  $N_2$ . **b** The causal decomposition method, incorrectly concluded causation according to IMF 1 and 2, even though  $N_1$  and  $N_2$  do not interact. Here, causal decomposition is performed under 1000 ensemble EMD with noise level  $r = 0.085$  selected based on the criteria of maximizing the separability but maintaining orthogonality of the IMFs, following the Matlab codes provided in Yang et al.<sup>1</sup>. In contrast, **c** CCM had no convergence (i.e., no improvement in CCM skill with increasing library size) in cross-mapping between  $N_1$  and  $N_2$ , and thus correctly concluded no causation between  $N_1$  and  $N_2$ .

remove cycles or construct a null model that accounts for cycles (e.g., seasonality) have been developed<sup>4,8–10</sup>. (3) Yang et al.<sup>1</sup> stated that “CCM is developed under the constraints of perfect deterministic system”; this is incorrect. In reality, real-world systems usually contain a deterministic skeleton convolved with stochastic processes<sup>11</sup>. In fact, modeling and empirical examples have demonstrated that CCM correctly identifies causation, even when stochastic processes are convolved with deterministic signals<sup>2,6,12</sup>.

It is noteworthy that interpretation of findings based on IMF subtraction needs to be done with caution. Yang et al.<sup>1</sup> included a strong statement regarding interpreting their findings, stating that removals of causal-related IMFs enable us to exclusively recover intrinsic dynamics of the target series from the residual IMFs. Although this statement is correct when the system dynamic is a result of superposition of signals, it is not always correct for general dynamical systems in which causal influences cannot be easily separated from intrinsic dynamics. Using a simple example of prey–predator model,  $\frac{dx}{dt} = ax - \beta xy$ ;  $\frac{dy}{dt} = \delta xy - \gamma y$ , we have a clear expectation that the prey will grow exponentially in the absence of the predator, or at the very least, prey cannot oscillate. However, these expectations are not realized after subtracting causal IMFs; rather, the remaining components continue to cycle—quite at odds with intuition based on the statement provided by Yang et al.<sup>1</sup>. The salient point is that interpretation of the residual IMF is not as unambiguous as the original text implies. Certainly, IMF subtraction is not equal to mathematical subtraction. However, real-world biologists applying Yang et al.’s approach<sup>1</sup> to predator–prey systems may be confused when interpreting the results according to Yang et al.’s statement about separability in causal inference. In fact, after subtracting the effect of the predator on the prey (e.g., Fig. 1c in ref. <sup>1</sup>), prey continue to oscillate. Based on countless

chemostat experiments, prey grown in isolation reach a steady state set by the rate of nutrient input and media outflow. Thus, sustained oscillations (e.g., remaining series after accounting for predator effect) suggest the existence of other factors. Assured by their IMF analysis that these oscillations are neither driven by the predator nor by its intrinsic dynamics, biologists might conclude that there must be another variable causing the oscillation (perhaps, time-varying fluctuations in resource availability or temperature) and fruitlessly search without success. Of course, they never find one, because the oscillation is artificially introduced by performing an additive decomposition on a non-separable system. Thus, we caution potential over-interpretation of the meanings of various IMFs in EMD.

To summarize, the EMD method of Yang et al.<sup>1</sup> clearly works for systems in which superposition is obtained, but does not provide unambiguous results for non-separable, nonlinear dynamical systems.

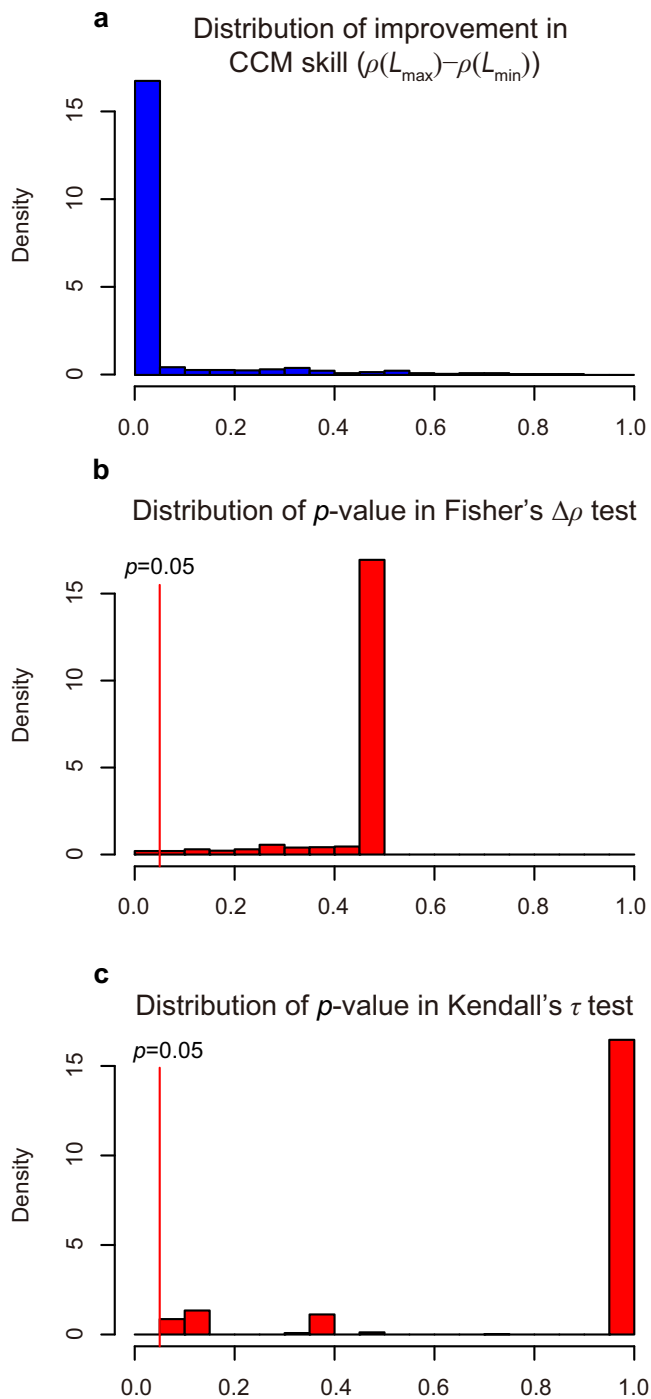
**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Empirical datasets analyzed in this study are open-access and available by following the same instruction addressed in Yang et al.<sup>1</sup>. R source codes used to generate synthetic datasets from the Moran-effect model and white noises are available on GitHub, <https://github.com/biozoo/CommentEMD>.

### Code availability

Matlab and R source codes for conducting EMD causal decomposition and CCM analyses, respectively, are provided on GitHub, <https://github.com/biozoo/CommentEMD>.



**Fig. 2 CCM analysis for paired white noises.** White noise time series were generated from 10,000 simulations and all time series were trimmed to length = 10, following Yang et al.<sup>1</sup>. In total, we performed CCM analyses between 1000 random pairs of white noises. To evaluate convergence of CCM, we calculated three indices: **a** improvement in CCM skill from minimal ( $L = 2$ ) to maximal library length ( $L = 10$ ); **b**  $p$  value for testing the significance of the improvement in CCM skill using Fisher's  $\Delta\rho$  Z test; and **c**  $p$  value for testing the significance of monotonic increasing trend in CCM skill using Kendall's  $\tau$  test. In a majority of cases, improvements in CCM skill were very small and close to zero, indicating no convergence (**a**). As such, false positives in both Fisher's Z test (**b**) and Kendall's  $\tau$  test ( $p < 0.05$ ) (**c**) occurred, with very low probability. In summary, the probability of detecting spurious causation in paired short white noise was very low; this was opposite to conclusions of Yang et al.<sup>1</sup> based on the incorrect definition of CCM.

- Ye, H., Deyle, E. R., Gilarranz, L. J. & Sugihara, G. Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci. Rep.* **5**, 14750, <https://doi.org/10.1038/srep14750> (2015).
- Sugihara, G., Deyle, E. R. & Ye, H. Reply to Baskerville and Cobey: misconceptions about causation with synchrony and seasonal drivers. *Proc. Natl Acad. Sci. USA* **114**, E2272–E2274, <https://doi.org/10.1073/pnas.1700998114> (2017).
- McCracken, J. M. & Weigel, R. S. Convergent cross-mapping and pairwise asymmetric inference. *Phys. Rev. E* **90**, 062903, <https://doi.org/10.1103/PhysRevE.90.062903> (2014).
- BozorgMagham, A. E., Motesharrei, S., Penny, S. G. & Kalnay, E. Causality analysis: identifying the leading element in a coupled dynamical system. *PLoS ONE* **10**, e0131226, <https://doi.org/10.1371/journal.pone.0131226> (2015).
- Paluš, M., Krakovská, A., Jakubík, J. & Chvosteková, M. Causality, dynamical systems and the arrow of time. *Chaos: Interdiscip. J. Nonlinear Sci.* **28**, 075307, <https://doi.org/10.1063/1.5019944> (2018).
- van Nes, E. H. et al. Causal feedbacks in climate change. *Nat. Clim. Change* **5**, 445–448, <https://doi.org/10.1038/nclimate2568> (2015).
- Deyle, E. R., Maher, M. C., Hernandez, R. D., Basu, S. & Sugihara, G. Global environmental drivers of influenza. *Proc. Natl Acad. Sci. USA* **113**, 13081–13086, <https://doi.org/10.1073/pnas.1607747113> (2016).
- Ye, H. et al. Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proc. Natl Acad. Sci. USA* **112**, E1569–E1576, <https://doi.org/10.1073/pnas.1417063112> (2015).
- Sugihara, G. Nonlinear forecasting for the classification of natural time series. *Philos. Trans. R. Soc. A* **348**, 477–495 (1994).
- Ushio, M. et al. Fluctuating interaction network and time-varying stability of a natural fish community. *Nature* **554**, 360–363, <https://doi.org/10.1038/nature25504> (2018).

### Author contributions

C.-W.C., S.B.M., and C.-H.H. conceived the research idea; C.-W.C. designed the models and analyzed data; C.-W.C., S.B.M., and C.-H.H. wrote the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-30359-8>.

**Correspondence** and requests for materials should be addressed to Chih-hao Hsieh.

**Peer review information** *Nature Communications* thanks Dimitris Kugiumtzis, Juergen Kurths, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 17 April 2019; Accepted: 11 February 2022;  
Published online: 23 May 2022

### References

- Yang, A. C., Peng, C.-K. & Huang, N. E. Causal decomposition in the mutual causation system. *Nat. Commun.* **9**, 3378, <https://doi.org/10.1038/s41467-018-05845-7> (2018).
- Sugihara, G. et al. Detecting causality in complex ecosystems. *Science* **338**, 496–500, <https://doi.org/10.1126/science.1227079> (2012).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022