

Geosci. Model Dev. Discuss., referee comment RC2  
<https://doi.org/10.5194/gmd-2021-196-RC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on gmd-2021-196

Anonymous Referee #2

---

Referee comment on "Quantifying Causal Contributions in Earth Systems by Normalized Information Flow" by Chin-Hsien Cheng and Simon A. T. Redfern, Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-196-RC2>, 2021

---

The paper uses the methodology developed of Liang over the last years for the information flow in coupled systems. It advocates the use of normalized information flow, or nIF, to (i) characterize causal relationships of the coupled variables and (ii) to assess the quality of model simulations from a causality perspective.

The paper is timely in the sense that causal methods are increasingly popular, with very different approaches nicely summarized e.g. by the Runge et al. (2019) paper cited in this manuscript where, as far as the reviewer remembers, the information flow of Liang is missing. So this could be a welcome contribution to causal approaches in the geosciences.

Unfortunately, the paper is excessively difficult to read. It starts with the equation (2) - which is eq. (5) of Liang (2021) in Entropy, which should be referenced to, but this has been mentioned already in the comments - which to the uninitiated comes out of nowhere, and is not motivated but only described. Why should this particular algebraic combination of covariances be called "information flow"? In the same spirit, how should a reader easily grasp the meaning of the normalization factor Z from eq. (4)? How do you separate the changes in marginal entropy into the self-dependent and the noise term for observed (as opposed to generated) time series; in other words, how do you calculate Z and nIF in this situation? There is no hint given in ch. 2.1.

Next, extensive use is made of an artificial example (called mock-up data), but should the reader be interested in this, she has to refer to the Supplementary Material. Here, she finds a system of two coupled first-order equations for the variables X1 and Y1 - written in a manner more complicated than necessary, since the overall term  $dY1/dt$  is not taken out but repeated three times, and with seemingly arbitrary numerical constants (1.1, 1.5, 300, 1.8, 0.000005 and so on) - which is not motivated by any means. The reviewer also wonders in which sense the noise terms on the right side of the Table deserve that name - being a sum of a constant, a deterministic trigonometric function of time, and the function itself? How could that be noisy, where is a stochastic process involved?

In the main text, there is talk about X and Y (not X1 and Y1) and in Figure 2, we suddenly have  $d(\text{partial}) X_1$ ,  $dX_2$  and  $dX_3$  - what are these? In the heading of the Figure, it says that  $dX_2=dX_1$  and  $dX_3=dX_1$ , so it is only one quantity after all? The graphs show three curves, so they are again different? The reader should look for red boxes according to the legend - there are no red boxes in Fig. 2. And what do we see on the y axis, actually? The variables themselves? Their partial derivatives? The information flow? What does the axis legend "by  $mR^2$ " even mean?

At the beginning of ch. 3.1, there is talk of "where causality becomes more important" - as opposed to what? And how do you know that, given observations and measurements (only)?

The reviewer didn't get the concept of the "1:2:3 ratio" either nor how this could convert (apparently at high noise level) to a "-1:-2:-3 ratio" - isn't that the same since pairwise signs would cancel out?

For the comparison of observations and model runs, here for CH4 growth rates, the reviewer has a hard time to discern the upper panels of Fig. 6 and 7. They look exactly identical, apart from the fact that the time axis for Fig. 7 is shorter (up to 2012). CESM2 can't be that "perfect"? Also, the observations/simulations give a rather blurred image along the latitudes and the time axis, whereas the estimates have a finer resolution. How is that possible, and how do the authors come to the conclusion that nIF is doing best, and that CESM2 fails to reproduce the spatial pattern?

The paper raises a lot of questions. It requires a substantial revision (major) before having the chance to come close to be readable. The potential of the method is present, but it has to be motivated much more explicitly and the examples have to be explained, shown in the main paper (a time series graph of the very target variables X and Y would be handy) and then a clearer demonstration and indication why nIF and its variants are superior to a regression approach, or in other words, why correlation and causality are different concepts, to the extent that you can have two causally connected variables with a Pearson correlaton coefficient of zero.

A promising application of the framework seems to be to determine the effective lag between cause and effect by lagging one of the two such that the nIF value (or one of its variants) is maximized, and compare this to a conventional cross correlation analysis. This is merely a suggestion since these lags are to be expected, in particular in the context of teleconnected variables.