https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022


Geoscientific Model Development Discussions
Open Access

# Empirical Assessment of Normalized Information Flow for Quantifying Causal Contributions

Chin-Hsien Cheng[1,2], Simon A. T. Redfern[1]

[1]Asian School of the Environment, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
[2]Joint International Research Laboratory of Climate and Environment Change, Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China

*Correspondence to*: Simon A. T. Redfern (simon.redfern@ntu.edu.sg)

**Abstract.** To understand the plethora of important processes that are characterized by their complexity, such as global climate change, it is important to quantify causal contributions between time series variables. Here, we examine the hypothesis that the normalized causal sensitivity (*nCS*) can be measured by the (modified) normalized information flow, *nIF* (or *mdnIF*). The instantaneous causal sensitivity is defined by absolute causal contributions to the effect variable over the change in cause variable. The *nCS* needs to be comparable among i) causes, ii) at different times and iii) from various locations. Therefore, if our hypothesis holds, the *nIF* must also fulfil these three requirements. We verify, empirically, that the causal contributions between variables can be reasonably estimated by the product of a constant "maximal causal sensitivity" and a modified *nIF*. Between opposite causal directions, causal sensitivity can be further normalized by the larger "maximal causal sensitivity". Our method is useful when there are: i) strong but hard-to-quantify noise contributions to the effect variable, ii) significant causal time-lags with a need to estimate the lag, iii) many causes from various locations to an overall mean effect with a need to differentiate their causal contributions, or iv) causal contributions at higher order.

**Keywords:** causality, information flow, causal contributions, climate change, Earth system models

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development Discussions

## 1 Introduction

Causality is one of the foundations of scientific understanding and progress. It has continued to expand its application in various disciplines in recent years, including in biomedical science (Russo and Williamson, 2011; Rasmussen et al., 2016; Lin and Ikram, 2020; Friston et al., 2020), neuroscience (Seth et al., 2015; Chen et al., 2016; Stokes and Purdon, 2017; Hill et al., 2017; Barnett et al., 2018), artificial intelligence (Pearl, 2019; Luo et al., 2020), and economics (Granger, 1969; Varian, 2016; Athey and Imbens, 2017; Andor and Fels, 2018). Within Earth sciences, causation is important, for example, for detecting causal signals and testing models against observed data (Sugihara et al., 2012; Stips et al., 2016; Runge et al., 2019a; Winkler et al., 2021), evaluating, constraining, and improving climate models (Cox et al., 2018; Bai et al., 2018; Hall et al., 2019; Verbitsky et al., 2019; Runge et al., 2019a; Vázquez-Patiño et al., 2020; Nowack et al., 2020; Docquier et al.), and estimating attribution of extreme or local events to climate or other global changes (Ornes, 2018; Pfrommer et al., 2019; Swain et al., 2020). The application of various causal methods to Earth sciences has been reviewed by Runge et al. (2019b), where the challenges of such methods are discussed, especially those arising from the nonlinear and spatiotemporal variation of complex processes. Runge et al. (2019b) also suggested a way forward for Earth sciences, by combining observational causal inference and physical modelling. While process-based models attempt to quantify the complex interactions between, for example, anthropogenic activities and multiple natural processes, they could potentially overlook or misinterpret some important processes. On the other hand, statistical models extrapolate historical trends into the future through statistical tools, but may still lack insight into the physical underlying processes. Intuitively, methods that are capable of quantifying physical causal contributions between observational time series would plug the gap between process-based and statistical models, providing a key to unlocking and understanding causality in Earth systems science processes.

The progress of causal research has been fuelled by the continued development and improvement of analytical tools for assessing causal influences, from the Nobel-prize winning Granger causality developed in 1960s (Granger, 1969) to the Shannon entropy-based information transfer (flow) (Schreiber, 2000) in the 21st century. Among various methods, information flow (*IF*) (Liang, 2014, 2016, 2018, 2021b, a), and its normalized form (*nIF*) (Liang, 2015, 2016, 2021a; Liang and Yang, 2021) derived by Liang, is a relatively new, yet established, measure of causality between two dynamical events realized in time series. Currently, *IF* and *nIF* have principally been applied in Earth sciences, with examples of its application including confirmation of the contribution of anthropogenic greenhouse gases (GHGs) to global warming in the post-industrial period (Stips et al., 2016), forecasting tropical cyclone genesis (Bai et al., 2018), and the central-Pacific type of El Niño (Liang et al., 2021). The method has proven capable in reconstructing causal graphs with single and bi-directional causality as well as with confounding processes (Liang, 2021a). In contrast, the most commonly applied causal analysis, Granger causality, faces challenges when dealing with contemporaneous effects and feedback cycles, which are unfortunately ubiquitous in Earth systems (Runge et al., 2019b). Nevertheless, although *IF* and *nIF* appear good quantitative measures of causality strength, they are often applied in concert with other statistical models. For example, such methods have been employed to improve regression-based correlation or/and neural network models containing multiple potential factors, by highlighting only those factors with significant causal influence (Bai et al., 2018; Liang et al., 2021). The effectiveness of information flow methodologies for directly quantifying causal contributions or for building semi-process-based causal models, by employing the magnitudes of *IF* or *nIF* (especially for coupled feedback processes) has not, however, been examined. Furthermore, the rather complicated theoretical derivation and underlying understanding of *IF* and *nIF* still limits the application of these methodologies by research communities unfamiliar with them. Here, we adopt an empirical approach to test the hypothesis that the normalized causal sensitivity between time-dependent variables can be described by normalized information flow, and explore the conditions under which such an approach is effective.

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development Discussions
Open Access

## 2. Concepts and Methods

### 2.1 Testing Framework

Before we examine whether any causal method, such as *IF* or *nIF*, can be used to directly quantify causal contribution, normalized causal sensitivity (*nCS*) must be defined in order to allow comparison of causal strengths from different causes at various times and locations. Ideally, such normalization should allow comparison among various systems and causal directions. Nevertheless, for simplicity's sake, we will first explore the normalization of causal sensitivity from one variable (*X*) to the other (*Y*) across various times and locations in a system.

The causal contribution of variable *X* to the rate of change of variable *Y* can be expressed as $\partial Y(X)/\partial t$, while the total changes in *X* and *Y* are expressed as total derivatives $dX/dt$ and $dY/dt$, respectively. Here, the causal contribution is expressed as a partial derivative, i.e. $\partial Y(X)/\partial t$, since it describes the rate of change of *Y* as a function of variation in *X* under the conditions that other (non-*X*) variables do not contribute to $dY/dt$, equivalent to them being held constant. All non-*X* contributions to $dY/dt$ can be considered as the "noise contribution".

The ratio of $|\partial Y(X)/\partial t|$ to $|dX/dt|$ reflects the instantaneous causal sensitivity of *Y* to changing *X*. The causal sensitivity can then be normalized to the maximal causal sensitivity of a system over various *Xs* during interested periods and from interested locations, thus *nCS* ranges between 0-1:

$$nCS_{(X\rightarrow Y)} = \left|\frac{\frac{\partial Y(X)}{\partial t}}{\frac{dX}{dt}}\right| \div max\left|\frac{\frac{\partial Y(X)}{\partial t}}{\frac{dX}{dt}}\right| = \frac{causal\ sensitivity\ of\ Y\ to\ changing\ X}{maximal\ causal\ sensitivity\ of\ Y\ to\ changing\ X} \quad (1)$$

Equation 1 defines *nCS*, the normalized sensitivity of a specific cause from various times and locations (e.g. $nCS(X_1 \rightarrow Y)$ and $nCS(X_2 \rightarrow Y)$) on the effect variable. Note that the maximal causal sensitivity does not necessarily occur when $|\partial Y(X)/\partial t|$ approaches $|dY/dt|$. If there are persistently strong contributions from noise variables, $|\partial Y(X)/\partial t|$ may always be smaller than $|dY/dt|$. Conversely, if the maximal causal sensitivity occurs when the noise contribution acts in opposition to $\partial Y(X)/\partial t$ and $dY/dt$ (hence $\partial Y(noise)/\partial t$ has opposite sign to $dY/dt$), then $|\partial Y(X)/\partial t| > |dY/dt|$. Hence, this normalization is from the perspective of the cause variable *X* rather than the effect variable *Y*, and direct comparison of different systems (e.g. between $nCS(X \rightarrow Y)$ and $nCS(A \rightarrow B)$) or of opposite causal directions (e.g. between $nCS(X \rightarrow Y)$ and $nCS(Y \rightarrow X)$) is not allowed.

A causal analysis capable of estimating comparable causal strengths should reflect *nCS*. In other words, *nCS* can be used as the testing framework to assess how good a causal method is in estimating a comparable causal strength.

### 2.2 Key Hypothesis

We wish to explore the hypothesis that the normalized causal sensitivity between time-dependent variables can be described by normalized information flow:

$$nCS_{(X\rightarrow Y)} \approx \frac{flow\ of\ uncertainty\ from\ X\ to\ Y}{overall\ flow\ of\ uncertainty\ to\ Y\ from\ X,\ \ nonX,\ \ and\ Y\ itself} = \frac{|IF_{(X\rightarrow Y)}|}{|IF_{(X,nonX,Y\rightarrow Y)}|} = |nIF_{(X\rightarrow Y)}| \quad (2)$$

$$IF_{a,(X\rightarrow Y)} = |IF_{(X\rightarrow Y)}| \times (\pm 1, based\ on\ R_{XY}) \quad (3)$$

$$nIF_{a,(X\rightarrow Y)} = |nIF_{(X\rightarrow Y)}| \times (\pm 1, based\ on\ R_{XY}) \quad (4)$$

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022



In information theory, the flows of the amount of information (*IF*) or degree of uncertainty, equivalent to flows of Shannon entropy, represent the causality, with unit in nat per time (Liang, 2014, 2015, 2016, 2018). Intuitively, we assume a stronger causal sensitivity corresponds to a proportionally stronger *IF*. Therefore, the normalized causal sensitivity may be estimated from the normalized uncertainty or normalized information flow. For normalization of information flow between only two time series, we can only categorize the information flow received by *Y* into three sources: from *X*, not from *X* (i.e. from non-*X*), and from *Y* itself (equation 2). We term this as *nIF* and hypothesize it represents *nCS*, denoted by the approximate equal sign. As discussed in 2.1, *nCS* requires comparable causal sensitivity at different times and from various locations. In other words, for the hypothesis to be valid, the *nIF* must also be comparable over times, locations, and among different causes (at least comparable between *X*, non-*X*, and *Y* itself in equation 2). In fact, the normalization of *nCS* over maximal causal sensitivity that only occurs at a specific time and location, is hypothesized to be interchangeable with the normalization of *nIF* over causes at different times and locations. We will show later that the interchangeability of normalization over cause, time, and location is indeed the greatest strength of this method. It enables the identification of particular causes, and their locations and timing, associated with effects. Such an approach is common in Earth sciences; for example, methane-climate feedback sensitivity can be expressed as the dependence of the increase in naturally contributed atmospheric methane concentration with increasing global temperature.

In addition, while the absolute magnitudes of *IF* and *nIF* represent the causality, the physical interpretation of their signs remains unclear. The sign, in this case, may reflect whether the cause-variable influences the effect-variable through an increasing (positive *IF*) or decreasing (negative *IF*) uncertainty (Liang, 2018). To remove the absolute operator and estimate the direction between $\partial Y(X)/\partial t$ and $dX/dt$, which is needed for *nIF* to represent *nCS*, the sign of correlation coefficient in the regression (i.e. $R_{XY}$) is assigned to the *IF* and *nIF* which we then denote as $IF_a$ and $nIF_a$ (equations 3-4, subscript "a" stands for "adjusted"). This becomes useful to indicate either positive or negative feedback.

For a given linear model of the effect of *X* on *Y*, the maximum likelihood estimator of $IF_{(X\rightarrow Y)}$ is given by (Liang, 2014):

$$IF_{(X\rightarrow Y)} = \frac{C_{YY}C_{YX}C_{X,dY} - C_{YX}^2 C_{Y,dY}}{C_{YY}^2 C_{XX} - C_{YY}C_{YX}^2} \tag{5}$$

where $C_{YX}$ is the covariance between variables *Y* and *X*, and $C_{X,dY}$ is the covariance between *X* and $\dot{Y}$, given by the series approximation of $dY/dt$ using Euler forward differencing ($\dot{Y}_n = (Y_{n+1} - Y_n)/\Delta t$). The same system of notation applies to $C_{XX}$, $C_{YY}$, and $C_{Y,dY}$. The normalized information flow (Liang, 2015) was proposed by dividing the $|IF|$ by a normalizing factor, *Z*.

$$|nIF_{(X\rightarrow Y)}| = |IF_{(X\rightarrow Y)}| / Z_{(X\rightarrow Y)} \tag{6}$$

$$Z_{(X\rightarrow Y)} = |IF_{(X\rightarrow Y)}| + \left|\frac{dH_Y^*}{dt}\right| + \left|\frac{dH_Y^{noise}}{dt}\right| \tag{7}$$

where $\left|\frac{dH_Y^*}{dt}\right| + \left|\frac{dH_Y^{noise}}{dt}\right|$ is the estimated increase in marginal entropy (extent of uncertainty) $H_Y$, which includes the contribution of $H_Y$ due to *Y* itself (first term) and the contribution from noise (second term), as given by equations 8- 10.

$$\left|\frac{dH_Y^*}{dt}\right| = \frac{C_{XX}C_{Y,dY} - C_{YX}C_{X,dY}}{C_{YY}C_{XX} - C_{YX}^2} = p \tag{8}$$

$$\left|\frac{dH_Y^{noise}}{dt}\right| = \frac{\Delta t}{2C_{YY}}\left(C_{dY,dY} + p^2 C_{YY} + q^2 C_{XX} - 2pC_{dY,Y} - 2qC_{dY,X} + 2pqC_{YX}\right) \tag{9}$$

$$q = \frac{-C_{YX}C_{Y,dY} - C_{YY}C_{X,dY}}{C_{YY}C_{XX} - C_{YX}^2} \tag{10}$$

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development
Discussions
Open Access

Comparing equations 6 and 7 to equation 2, $\left|\frac{dH_Y^{noise}}{dt}\right|$ appears to equate to $IF_{(\text{non-}X\rightarrow Y)}$, $\left|\frac{dH_Y^{*}}{dt}\right|$ equates to $IF_{(Y\rightarrow Y)}$, and $IF_{(X,\text{non-}X,Y\rightarrow Y)}$ is assumed to be the sum of the three terms in equation 7, i.e. $Z_{(X\rightarrow Y)}$. However, these three terms, especially the $IF_{(Y\rightarrow Y)}$ term should be dependent on the earlier trend of $Y$ before the specific time of interest, so that it should be partly influenced by the $IF_{(X\rightarrow Y)}$ and $IF_{(\text{non-}X\rightarrow Y)}$ terms. The overall information flow $IF_{(X,\text{non-}X,Y\rightarrow Y)}$ is hence not necessarily equal to the sum of the three terms. We will therefore empirically examine three other normalizing factors that may represent $IF_{(X,\text{non-}X,Y\rightarrow Y)}$:

$$md_1Z_{(X\rightarrow Y)} = |IF_{X\rightarrow Y}| + \left|\frac{dH_Y^{noise}}{dt}\right| \tag{11}$$

$$md_2Z_{(X\rightarrow Y)} = \left|\frac{dH_Y^{*}}{dt}\right| \tag{12}$$

$$md_3Z_{(X\rightarrow Y)} = md_1Z_{(X\rightarrow Y)} + \left|md_2Z_{(X\rightarrow Y)} - md_1Z_{(X\rightarrow Y)}\right| \tag{13}$$

Where the "*md*" denotes "modified" and the subscripts 1-3 refer to different modifications. By substituting Z in equation 6 by $md_1Z$, $md_2Z$ or $md_3Z$, we obtain $|md_1nIF|$, $|md_2nIF|$, or $|md_3nIF|$. Basically, equation 11 assumes that $IF_{(X,\text{non-}X,Y\rightarrow Y)}$ is only from $X$ and non-$X$ cause variables, and the sum of the two should have already included the $IF_{(Y\rightarrow Y)}$. Oppositely, equation 12 assumes that $IF_{(Y\rightarrow Y)}$ should have already included the information flows from $X$ and non-$X$ cause variables. Equation 13 assumes the absolute difference between $md_1Z$ and $md_2Z$ to be additional information flow on top of those from $X$ and non-$X$ cause variables.

### 2.3 Empirical Tests

To facilitate empirical investigation, we re-formulate equations 1-4 in the manner shown in equation 14. We define the "maximal causal sensitivity of $Y$ to changing $X$" as a constant α and the $nIF_{a,(X\rightarrow Y)}$ as the "multiplier". Equation 14 focuses on causal contribution (i.e. $\partial Y(X)/\partial t$) instead of causal sensitivity for practical reasons, since larger peaks of $\partial Y(X)/\partial t$ have a greater bearing on the sensitivity over long time scales (i.e. $\Delta Y(X)/\Delta X$). In other words, relatively large percentage errors in instantaneous sensitivity during periods with small $\partial Y(X)/\partial t$ and $dX/dt$ do not significantly affect the long-term $\Delta Y(X)/\Delta X$.

$$\frac{\partial Y(X)}{\partial t} = \alpha \times multiplier \times \frac{dX}{dt} \tag{14}$$

In addition to testing our hypothesis presented in equation 2, that $nIF$ is a measure of normalized causal sensitivity, we explore the effect of using $IF_a$ rather than $nIF_a$ as the "multiplier" in equation 14. For $IF_a$ to be a valid multiplier in equation 14, it requires that the maximal causal sensitivity of $Y$ to changing $X$ and the overall flow of uncertainty to $Y$ in equations 1-2 are time-independent constants. We have also compared the applicability of linear and second order regressions in the determination of the "multiplier" in equation 14. For linear regression, the "multiplier" is $mR^2$ where $m$ is given by $Y = mX + c$ and $R$ is the correlation coefficient. For second order regression, the "multiplier" is $M_2R^2$ with $M_2 = 2aX+b$, the differential of $Y = aX^2 + bX + c$ (with subscript 2 in $M_2$ denoting second order regression). The "multiplier" in such an approach also takes care of α, which is then 1. An approach to estimating α for $nIF$, as well as its modified forms and $IF$, is to visually match the estimated and the designed $\partial Y(X)/\partial t$, in which case α effectively serves as a calibration factor.

#### 2.3.1 Assessing the Hypothesis

To explore the hypothesis that normalized information represents normalized causal sensitivity (section 2.2) and to identify which normalizing factor performs best, we first perform a series of tests based on designed mock-up datasets with general expression as

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development
Open Access
Discussions

$$dY/dt = \partial Y(X)/\partial t + \partial Y(n)/\partial t = f(dX/dt,\, t) + n(Y,\, t), \text{ and } dX/dt = \partial X(Y)/\partial t + \partial X(n)/\partial t = g(dY/dt,\, t) + o(X,\, t) \quad (15)$$

where $f$ and $g$ are the interdependent contributions, representing *cyclic causal interferences* or a *feedback loop* as in a format described by equation 14, such that the designed "α x multiplier" includes trigonometric terms with varying frequencies or/and a constant. The trigonometric terms mimic typical climate oscillations resulting from alternating positive and negative feedbacks, such as the famous El Niño–Southern Oscillation (ENSO) cycle (Im et al., 2015). Similarly, $n$ and $o$ are functions of other cause-variables and termed as noises. They include i) a self-dependent term which often tends to stabilize the fluctuations of effect-variables, mimicking negative Earth system feedbacks, e.g. the carbon-concentration feedback (Arora et al., 2020), as well as ii) other noise independent of $X$ and $Y$, but potentially varying with time $t$.

The first series of tests include:

1) a 1- dimensional (1D) example with a constant independent noise-contribution and a single causal direction;
2) a 1D example with fluctuating independent noise-contribution and a single causal direction;
3) a 1D example with fluctuating self-dependency noise-contribution and a single causal direction;
4) a 1D example with strong bidirectional causality but very weak self-dependency and independent contributions;
5) a 1D example with bidirectional causality and highly fluctuating independent noise contributions;
6) a 1D example with moderate contributions from all terms, together with 21-steps of time-lag (i.e. 21% of each analyzed time-window) for the interdependency term;
7) additional 1D examples with time-gaps for different terms and directions;
8) 3- dimensional (3D) examples with and without teleconnections.

Tests 1 to 5 examine whether the respective "multiplier" is able to reflect the causal contributions under various types of noise contributions, and which normalizing factor Z (equations 7, 11-13) performs best. Tests 6-8 examine whether such normalization applies across causes over time and space, so that the method has the potential to estimate where and when a change in cause-variable contributes to a change in effect-variable. This is the core requirement for our hypothesis to be valid and applicable.

Furthermore, for each 1D test, in order to examine if the "multiplier" measures $nCS$ with a constant calibration factor α representing the maximal causal sensitivity, we have chosen a 1:2:3 ratio for $\partial X1(Y1)/\partial t : \partial X2(Y1)/\partial t : \partial X3(Y1)/\partial t$ and $\partial Y1(X1)/\partial t : \partial Y2(X2)/\partial t : \partial Y3(X1)/\partial t$ for interdependent contributions between $X1$, $X2$, $X3$ and $Y1$, as well as between $Y1$, $Y2$, $Y3$, and $X1$. Note that the 1:2:3 ratio neither applies to $dX1/dt : dX2/dt : dX3/dt$ nor to $dY1/dt : dY2/dt : dY3/dt$, since a common noise function $\partial X(n)/\partial t$ is applied to all $dX1/dt$, $dX2/dt$, $dX3/dt$ and another common $\partial Y(n)/\partial t$ is applied to all $dY1/dt$, $dY2/dt$, $dY3/dt$. Comparing equation 14 to equations 1-2, the value of α for $IF_a$ and $nIF_a$ should then follow the same 1:2:3 ratio. Therefore, we have set the 1:2:3 ratio for α and examined if the estimated $\partial X(Y1)/\partial t$ and $\partial Y(X1)/\partial t$ given by equation 14 also reflect that ratio.

In the 3D context we make an analogy to climate systems, which typically involve data expressed in terms of longitude (lon), latitude (lat), and time ($t$) coordinates across the globe's surface (Fig. 1). In our empirical assessment, we have produced cause-

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022



maps of 3D-to-global-mean-1D variables (distribution and variability of contribution from causes, as illustrated in Fig. 1d), and effect-maps of global-mean-1D-to-3D variables (distribution and variability of contribution as effects illustrated in Fig. 1e). We further consider the presence or absence of interdependent teleconnection, by assigning the interdependent function based on values from the opposite side of the hemisphere (e.g. interdependency between *dX/dt* at 60°N and *dY/dt* at 60°S, see Fig. 1c) or from the same grid (Fig. 1b), respectively. If normalized information represents normalized causal sensitivity, the results should reflect the teleconnection from the opposite hemisphere for the cause-maps.

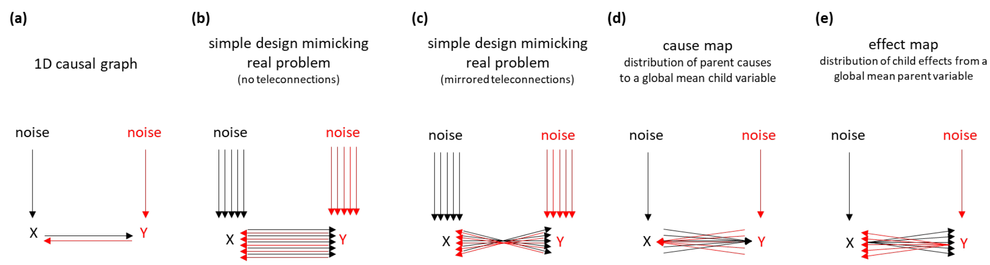


**Figure 1.** Illustrative causal graphs of designed 1D and multi-D causally interdependent variables X and Y, with/without teleconnection, and the basis of estimates for cause map and effect map.

### 2.3.2 Higher Order Dependency

The above testing framework via equation 14 and designed mock-up data via equation 15 corresponds to "rate-dependent" causal sensitivity described by equation 1, which sets a proportional relationship between changing cause and its changing contribution to the effect variables. Such "rate-dependent" causal sensitivity may best describe hysteresis of cause variable on effect variable. However, problems may arise when the causal dependency is "state-dependent", or a combination of both "rate-dependency" and "state-dependency". For example, the rate of natural carbon sink ($dC_{CO2}/dt$) is temperature-dependent, but the long-term dependency may be mainly due to temperature (*T*) (Arora et al., 2020), i.e. state-dependency, while its interannual variability and hysteresis may be associated with the initial condition and the rate of changing temperature (*dT/dt*) that links to drought, flood, and/or rewetting (Obermeier et al., 2017; Barnard et al., 2020), i.e. rate-dependency. In addition, the maximal likelihood of information flow in equation 5 applies to a linear model between *X* and *Y* *(Liang, 2014)* (hence between $\partial X(Y)/\partial t$ and $dY/dt$). Therefore, to cater for such higher order dependencies, equation 14 needs to be split into equations 16-18. Two maximal causal sensitivities are needed: the $\alpha_{hys}$ in equation 17 represents the maximal instantaneous sensitivity due to hysteresis, as in equation 14; and the $\alpha_{long}$ in equation 18 represents the maximal instantaneous sensitivity due to the long-term impact. With respect to the two maximal causal sensitivities, two different multipliers are also needed. By breaking down the second order causal dependency into two first-order equations, the $IF_a$ and $nIF_a$ are hence estimated based on time series *X* and *Y* (for the hysteresis in equation 17), and time series *dX/dt* and *Y* (for estimating the $\partial^2 X(Y)/\partial t^2$ followed by integration into the long-term $\partial X_{long}(Y)/\partial t$ in equation 18). With this testing framework, the designed equation for mock-up data assessment is modified accordingly (equation 19).

$$\frac{\partial X(Y)}{\partial t} = \frac{\partial X_{hys}(Y)}{\partial t} + \frac{\partial X_{long}(Y)}{\partial t} = \alpha_{hys}\left(multiplier_{hys}\right)\left(\frac{dY}{dt}\right) + \alpha_{long}\left(multiplier_{long}\right)(Y) \qquad (16)$$

$$\frac{\partial X_{hys}(Y)}{\partial t} = \alpha_{hys}\left(multiplier_{hys}\right)\left(\frac{dY}{dt}\right) \qquad (17)$$

$$\frac{\partial^2 X_{long}(Y)}{\partial t^2} = \alpha_{long}\left(multiplier_{long}\right)\left(\frac{dY}{dt}\right) \text{ and } \frac{\partial X_{long}(Y)}{\partial t} = \alpha_{long}\int_{t0}^{t}\left(multiplier_{long}\right)\left(\frac{dY}{dt}\right)dt \qquad (18)$$

$$dX/dt = \partial X(Y)/\partial t + \partial X(n)/\partial t = f(Y, dY/dt, t) + n(X, t) \qquad (19)$$

The designed $\partial X_{long}(Y)/\partial t$ tends to grow together with the $Y$, hence it could behave as a growing noise influencing the estimates of the "hysteresis" $IF_a$ and $nIF_a$. It is hence important to preliminarily minimize the influence of independent noise and $\partial X_{long}(Y)/\partial t$ on the $\partial X_{hys}(Y)/\partial t$ estimation. The potentially improved $IF_a$ and $nIF_a$ are obtained between $Y$ and an adjusted $X$, i.e. $X_{adj}$, and the $X_{adj}$ is obtained via equations 20-22: where $X_{adj}$ is obtained by adding an adjusted $dX_{adj}/dt$ time series to the initial $X_0$ (equation 20); $dX_{adj}/dt$ is obtained by removing a reference $dX_{ref}/dt$ from the $dX/dt$ (equation 21); and the $dX_{ref}/dt$ serves as a preliminary approximation of a rather constant or constantly growing $\partial X(noise)/\partial t$ and/or $\partial X_{long}(Y)/\partial t$, for example by assuming the value of $dX/dt$ at 25-75% split of the time-window as the $dX_{ref}/dt$ (equation 22), so that most (75%) of the $dX_{adj}/dt$ falls behind the $dX_{ref}/dt$ to reflect the causal effect on $\partial X_{hys}(Y)/\partial t$ (see Data Processing in Supplementary Information).

$$X_{adj} = X_0 + dX_{adj}/dt \tag{20}$$

$$dX_{adj}/dt = dX/dt - dX_{ref}/dt \tag{21}$$

$$dX_{ref}/dt = dX/dt \text{ at 25-75\% split of the time-window for calculating } IF \text{ and } nIF \tag{22}$$

## 3. Results and Discussion

### 3.1 Validating the Hypothesis and Method Advantages

Among all the "multipliers" tested, we find that causal contributions estimated based on $md_3nIF_a$ (i.e. replacing the $Z$ in equations 6 by $md_3Z$ in equation 13) best represent the designed causal contributions. Hence, for the 1D tests, we only present the designed and the $|md_3nIF|$-estimated causal contributions as a key comparison here (comparison with estimates based on other "multipliers" can be found in Supplementary Information).

Figure 2 shows the designed and $md_3nIF_a$-estimated causal contributions for tests 1-5 (section 2.3.1). When the influence of independent and self-dependent noise is insignificant (Fig 2 m-p), the estimates reflect the designed trends well. When there is strong influence from the independent-noise (Fig. 2 a-h, q-t), the major issue is that the correlation sign, when incorrect, misinterprets the feedback direction and causal contribution (Fig. 2 c, g, s, t). Nevertheless, we would like to highlight that the 1:2:3 ratio of the absolute contribution is approximately retained even when the correlation sign is wrong, suggesting the validity of our proposed hypothesis. Furthermore, a secondary issue is that even without a strong independent noise contribution, a strong influence on the effect variable via self-dependency terms may also affect the peak-to-peak ratio (Fig. 2 k).

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022



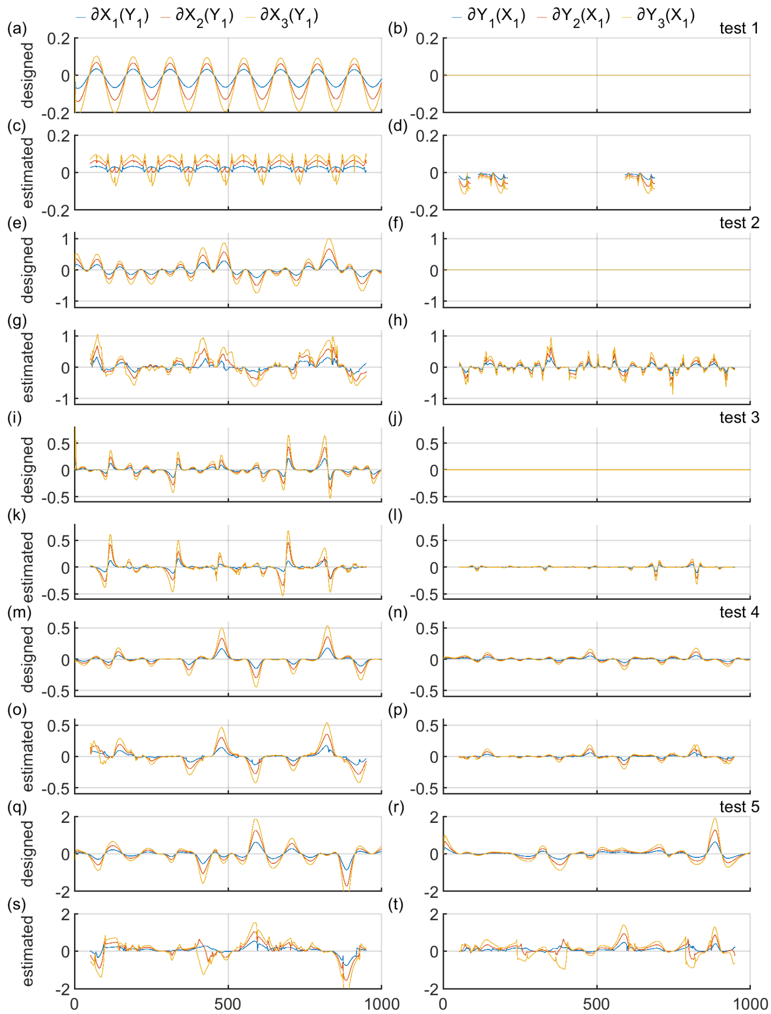


**Figure 2.** The basic 1D tests 1-5: two rows for each test. The odd and even rows are the designed and *md₃nIFₐ*-estimated causal contributions, respectively. Highlight of each test: strong but constant independent noise (a-d), strong and fluctuating independent noise (e-h), strong self-dependency noise (i-l), coupled-feedback with insignificant noise (m-p), and coupled-feedback with highly fluctuating independent noise (q-t). See Figs. S1-S5 for comparisons of estimates based on various "multipliers".

The estimates given by other "multipliers" for the same designed causal contributions in Fig. 2 are shown in Supplementary Information (Figs. S1-S5). For estimates given by regressions, the 1:2:3 ratio is strongly affected by the independent noise. For example, in a scenario with a misinterpreted correlation sign, a designed -1:-2:-3 ratio can be incorrectly reflected as ~3:2:1 in the estimated causal contributions (Fig. S1). Even if the correlation sign is correct, this designed ratio of causal sensitivities may still be lost and be reflected as ~1:1:1 with complete failure of the estimated peak-to-peak ratio under strong influence by independent noise contribution (Fig. S2). This shows that *$md_3nIF_a$* works better than regressions in the presence of hard-to-estimate noise contributions. The self-dependency contributions also affect the estimates given by regressions more than the estimates given by *$md_3nIF_a$* (Fig. S3). Estimates based on *$IF_a$* may better reflect the single directional causality (Fig. S1), however, their results for the 1:2:3 ratio as well as the peak-to-peak ratio are badly affected by the self-dependency terms (Fig.

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022


S3) and coupled feedback (Fig. S4-S5). This is because the 1:2:3 ratio is occasionally reflected in the absolute value of $|IF|$ (sub-Figs o, p of Figs. S3-S5) rather than being solely reflected by the calibration constant α.

Among the various (modified) $nIF_a$, $md_3nIF_a$ performs the best. It tends to minimize the error due to incorrect correlation sign, giving smaller estimated causal contributions when this occurs (Figs S1, S5). Firstly, the estimates given by $md_2nIF_a$ show unreasonably sharp fluctuations (e.g. sub-Fig. y of Fig. S1). Such fluctuations are better reflected by the absolute values of $|md_2nIF|$, with sudden change occurring between ~0 and ~1. This can be best explained by the arguments of Liang and Yang (2021) which highlight the impossibility of distinguishing cause and effect between two identical oscillating functions with a time-gap (e.g. sin(x) and sin(x-π)). Secondly, when the normalized causal sensitivity approaches its maximum, the $|nIF|$ proposed by Liang (2015) (equations 6-7) tends to approach 0.5 rather than 1 (sub-Figs s, t of Figs. S1-S5), while our proposed $|mdnIF|$ is close to 1. This highlights the dependence of the $IF_{(Y\rightarrow Y)}$ (or $\frac{dH_Y^*}{dt}$) term on the $IF_{(X\rightarrow Y)}$ and $IF_{(\mathrm{non}X\rightarrow Y)}$ (or $\frac{dH_Y^{noise}}{dt}$) terms. In other words, when the normalized causal sensitivity is at its largest, the information flow from the effect variable to itself ($md_2Z$) may actually mean the information flow from cause-variable and noise ($md_1Z$). Thirdly, our observation that $md_3nIF_a$ gives better estimates of causal contributions than $md_1nIF_a$ may imply that the difference between the $IF$ from cause-variable and noise ($md_1Z$) and that from the effect variable ($md_2Z$) could be the actual "additional" $IF$ that the effect-variable perceives (see equation 13 for $md_3Z$). Compared to typical material or energy balance equations with no output, this "additional" $IF$ is similar to a "generation" term (equation 23). Using this analogy, we assume that $md_2Z \sim md_1Z$ and $|md_2Z - md_1Z|$ is negligible when the causal sensitivity is strong, but when the causal sensitivity weakens, $|md_2Z - md_1Z|$ increases as does the perceived information flow by the effect-variable ($md_3Z$). This dilutes the normalized information flow, minimizing the error due to incorrect assignment of correlation sign.

$$Input\ (md_1Z_{X\rightarrow Y}) + Generation\ (|md_2Z_{X\rightarrow Y} - md_1Z_{X\rightarrow Y}|) = Accumulation\ (or\ perceived\ md_3Z_{X\rightarrow Y}) \qquad (23)$$

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022



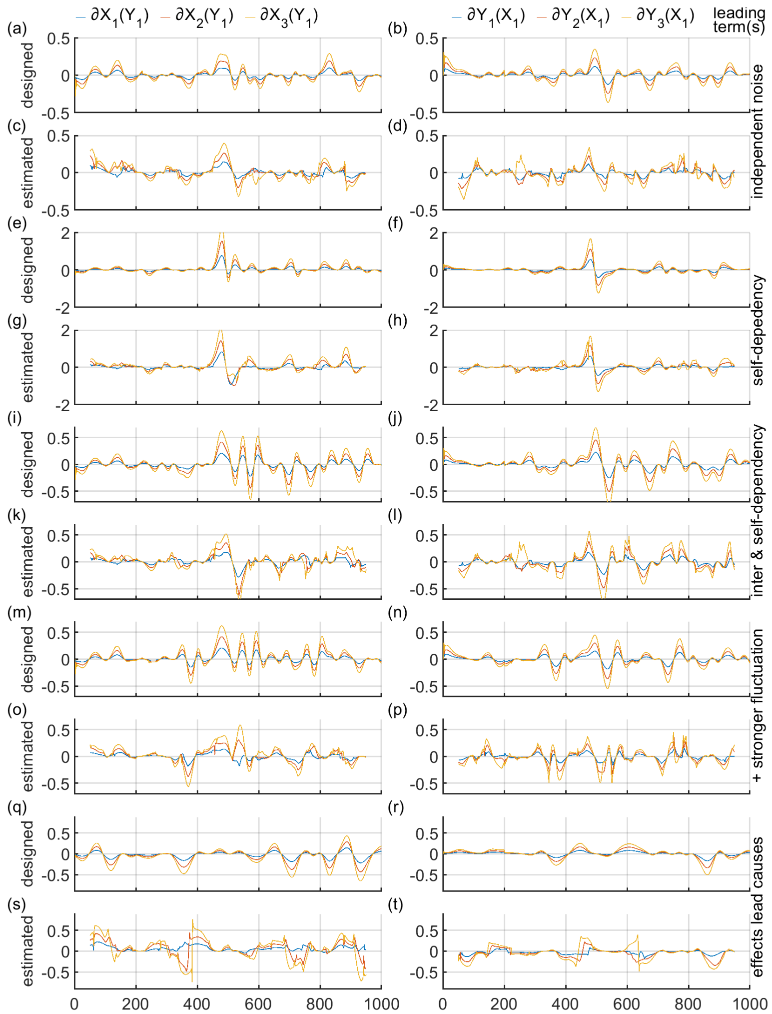


**Figure 3.** The 1D tests examining the effect of various time-gaps introduced from t > 200. (a-d): 21-unit time-lead for the interdependent terms only; (e-h): 21-unit time-lead for the self-dependency terms only; (i-l): 21-unit time-lead for both interdependent and self-dependency terms; (m-p): slightly increased (compared to i-l) fluctuation for the independent noise-contribution; and (q-t): 31-unit reverse time-lag (effect lead cause). See Figs. S6-S11 for the comparison between estimates based on various "multipliers" as well as a controlled test (S7) where all time-gaps are removed.

We have examined whether causal sensitivity with different time-lag can be normalized, and whether our method has the potential to estimate a time-lag (or even reverse time-lag). Figure 3 shows the designed and *md₃nIFₐ*-estimated causal contributions for different time gap configurations, with further comparison among different "multipliers" given in Figs. S6-S11.

The results show that the time-lag for the interdependency terms can indeed be (approximately) captured: the estimated causal contributions tend to occur at the time when the "cause" influences the designed "effect". For example, the estimated causal contributions tend to lead the designed effect by ~21-unit shown in Fig. 3 a-d and i-p. Even the reverse 31-unit "time-lag" in Fig. 3 q-t is partly captured, with the designed effect apparently leading the estimated causal contributions. While the

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development
Discussions
Open Access
EGU

time-gap applies only to the self-dependency terms, such a gap is not reflected by the estimated interdependent causal contributions (Fig. 3 e-h). However, the time-gap is not always correctly captured. In particular, the presence of a causal time-lag could lead to misinterpretation of the correlation sign, which may also tend to split or merge the effect of causal contributions, resulting in incorrect estimates of a time-lag, especially for high-frequency noise-contribution fluctuation (Fig. 3 m-p). We clearly need to be cautious when interpreting the time-gap between the estimated causal contributions and the designed (or observed) effects. For example, when the designed effects appear to precede the estimated causal contributions (supposedly around the time of cause), this should not simply be interpreted as “effect leading cause”.

When using other “multipliers”, the estimates given by regressions are particularly badly affected by time-gaps between the interdependency terms (Figs. S6. S9-S11). For $IF_a$ and other (modified) $nIF_a$, while the presence of time-gaps also affects the estimates, and the general issues discussed in Fig. 2 (and Figs. S1-S5) remain.

We have also studied designed and estimated contributions between two 3D variables with teleconnection operating from the opposite hemispheres (e.g. $X$ at 60°N is interdependent with $Y$ at 60°S), projected onto two dimensions using the zonal means (Fig. 4). The first row in Fig. 4 gives the designed distributions of $dX/dt$ and $dY/dt$, and the second row shows the designed values of interdependent $\partial X(Y)/\partial t$ and $\partial Y(X)/\partial t$ (effects). Thus, the second row corresponds to effects without any additional noise, while the first row represents the sum of contributions from effects and noise. Further rows give the estimated cause-maps. Two levels of designed noise-contribution have been applied. The obvious difference between Fig. 4a and 4c corresponds to stronger noise contributions compared to Fig. 4b and 4d. The noise alternates between positive and negative with a rather insignificant positive bias, hence the conditional advantage for $nIF_a$ (large noise contributions) is insignificant, at least for weak-noise case (right hand column). Furthermore, the time-lag is only one time-unit over the running window of the causal analysis time series data, with 49-time units in each window. Hence, this example focuses on the spatial causal contributions: in view of the mirrored teleconnection between north and south hemispheres, the best estimates of the cause-map should also be a mirrored image of the second row. This mirror characteristic can be best seen in the estimates given by $md_3nIF_a$.

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022


Geoscientific Model Development
Discussions
Open Access
EGU

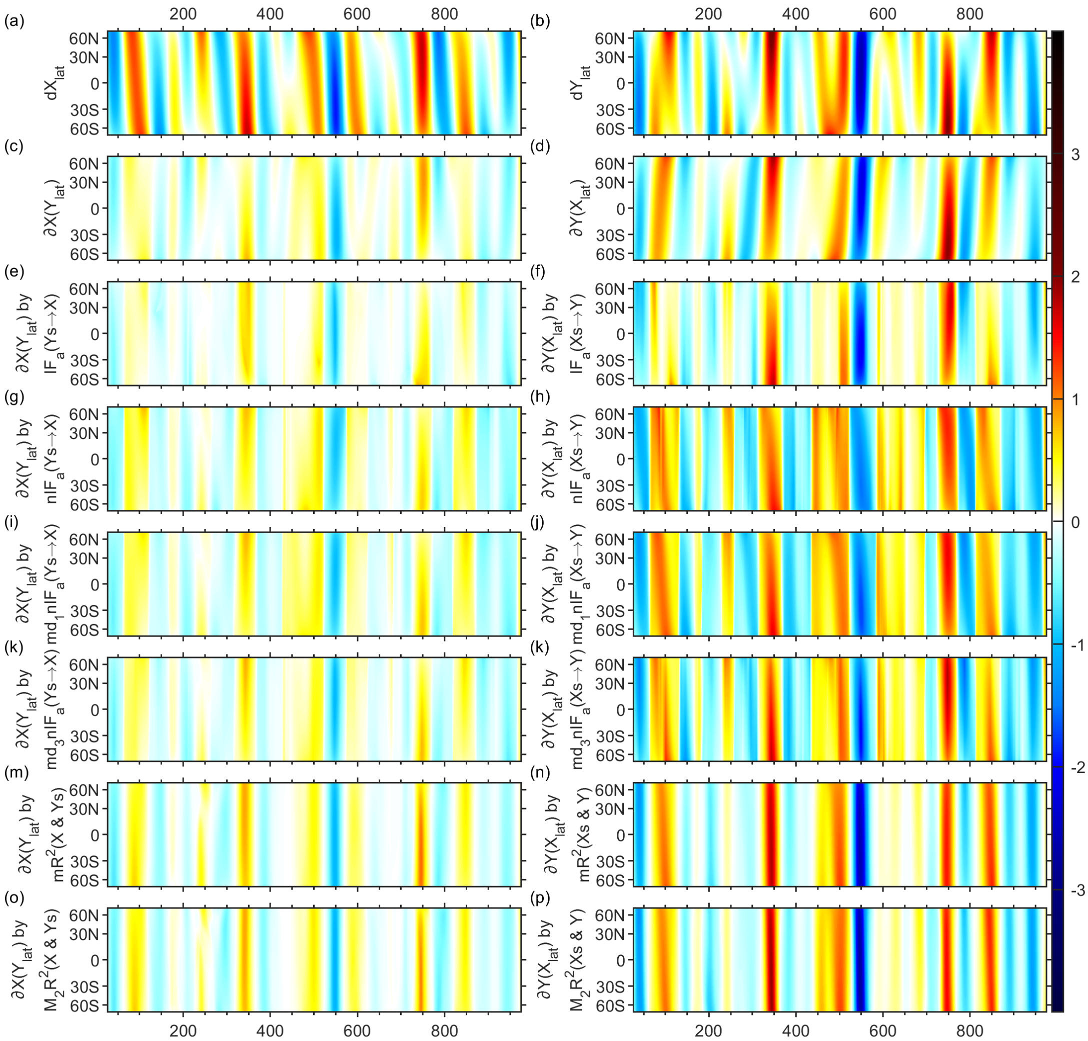


**Figure 4.** 3D-to-1D estimated cause-maps (3rd row and below) with the N-S teleconnections, which are supposed to generate N-S mirrored image of designed effect maps (2nd row). The 1st row shows the designed rates of changes, including contributions from noises. The mirroring teleconnection is best captured by $md_3nIF_a$. Also refer to Fig. S12 for cause-maps without the N-S teleconnections, as well as Fig. S13-S14 for the 1D-to-3D estimated effect-maps

In contrast, estimates obtained by regressions do not provide any clear evidence of such mirroring, although the estimates by second-order regression may manage to reflect this slightly better than those provided by first-order regression. Note that for a fair comparison, in the 3D tests we allow a visual adjustment of a calibration factor α for estimates by regressions, while in 1D tests α for regressions is set to unity for simplicity. The failure of regressions in reproducing the expected mirroring characteristics, even with a weak noise (right-hand column), could be simply due to the absence of a common calibration factor α (except 1) since regressions do not measure the causal sensitivity. Even in the absence of N-S teleconnection, the advantage of employing $md_3nIF_a$ for estimates of the spatial distribution in 3D-to-1D cause-maps is still apparent, although to a lower extent (Fig. S12). This is because there are, altogether, 360 x 180 time series over all grids which contribute to a single global mean time series in the 3D-to-1D cause-maps, emphasizing the importance of locations of cause signals in this case. In contrast,

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022



for 1D-to-3D effect-maps (where the location-data of the causes are already merged into a global mean value), regression tends to give better estimates, regardless of the absence or presence of teleconnection (Fig. S13-S14).

### 3.2 Second Order Causal Sensitivity and Noise Minimization

Figure 5, from left to right columns, shows the estimates of overall $\partial X(Y)/\partial t$, $\partial X_{hys}(Y)/\partial t$, $\partial X_{long}(Y)/\partial t$, and $\partial Y(X)/\partial t$, given by $md_3nIF_a$ in three 1D tests using the methods discussed in 2.3.2, including the removal of $dX_{ref}/dt$ (as the preliminary estimate of the noise and $\partial X_{long}(Y)/\partial t$ contributions) into $X_{adj}$ for the estimates of $\partial X_{hys}(Y)/\partial t$. Firstly, Fig. 5a-h extends the simple example of single directional causality shown in Fig. 2a-d and Fig. S1. Secondly, Fig. 5i-p includes a long-term influence term of $\partial X_{long}(Y)/\partial t$; and finally Fig. 5q-x further includes a bidirectional feedback influence of $\partial Y(X)/\partial t$, with its estimate based on $X$ (no adjustment) and $Y$ time series. The complete comparison among various methods with and without the removal of $\partial X_{long}(Y)/\partial t$ is given in Figs. S15-S21.

As compared to Fig. 2a and c, Fig. 5b and f show improved estimates of $\partial X_{hys}(Y)/\partial t$ after removing the $dX_{ref}/dt$ and running the causal analysis based on $X_{adj}$ and $Y$ time series. Problems associated with incorrect correlation sign are minimized. Note that the 1:2:3 ratio of α is preserved, which confirms again that the (maximal) causal sensitivity is independent of the noise. With the incorporation of $\partial X_{long}(Y)/\partial t$ (Fig. 5i-p) or even the interdependent $\partial Y(X)/\partial t$ (Fig. 5q-x), the method appears reasonably capable in separating the $\partial X_{hys}(Y)/\partial t$ and $\partial X_{long}(Y)/\partial t$ estimates. This suggests the utility of splitting higher order of causal sensitivities into multiple first order causal sensitivities. Note that although the estimates based on $X_{adj}$ and $Y$ appear to improve the $\partial X_{hys}(Y)/\partial t$ (also in Fig. S16, S18, S20 vs S15, S17, and S19), the removal of running $dX_{ref}/dt$ for each window unavoidably influences the causal signals too. For example, if we estimate $\partial Y(X)/\partial t$ based on $X$ and $Y_{adj}$, with a relatively small $\partial Y(noise)/\partial t$ than the $\partial Y(X)/\partial t$ contributions, the estimates will differ more from the designed trends, as compared to estimated $\partial Y(X)/\partial t$ based on $X$ and $Y$ (Fig. S19-S20). Hence, this simple method improvement is only suitable when there is a rather large constant or constantly growing noise and/or when there are significant long-term contributions to be determined (equation 18). We have also tested another way (other than that of equation 18) for estimating $\partial X_{long}(Y)/\partial t$, by estimating the $\partial Y_{long}(X)/\partial t$ based on the $X$ and time-integral of $Y$ (so that the causal sensitivity will be $\partial X(Y)/\partial t/Y$), instead of estimating the $\partial^2 X(Y)/\partial t^2$ based on $dX/dt$ and $Y$ (so that the causal sensitivity will be $[\partial^2 X(Y)/\partial t^2]/[dY/dt]$) with subsequent integration into the long-term $\partial X_{long}(Y)/\partial t$. However, this alternate method results in a highly fluctuating $\partial X_{long}(Y)/\partial t$ with correlation sign influenced by the $\partial X_{hys}(Y)/\partial t$ (Fig. S21). This suggests that a second-order form of causal sensitivity, i.e. $[\partial^2 X(Y)/\partial t^2]/[dY/dt]$ is capable of distinguishing the influence of $\partial X_{long}(Y)/\partial t$ from $\partial X_{hys}(Y)/\partial t$, but the (relatively more) integral form of causal sensitivity, i.e. $\partial X(Y)/\partial t/Y$, is unable to separate out the causal contributions from different orders of dependency.

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development
Discussions

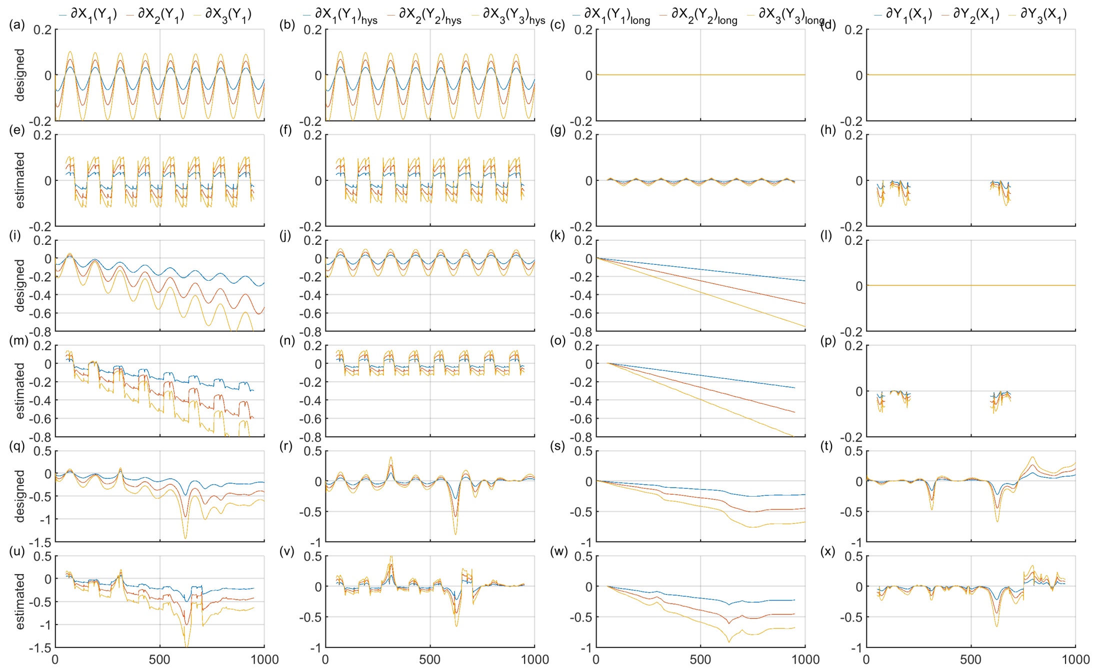


**Figure 5.** The 1D tests assessing the practicability of preliminary noise removal and breakdown of higher order causal contributions into multiple first order causal contributions: (a-h) with the same first order causal functions as in Fig.2a-d, (i-p) incorporating a growing $\partial X_{long}(Y)/\partial t$ term, (q-x) further incorporating a $\partial Y(X)/\partial t$ term. Odd row: designed contributions. Even rows: estimated contributions based on $md_3nIF_a$. Refer to section 2.3.2 for the $\partial X(Y)/\partial t$, $\partial X_{hys}(Y)/\partial t$ and $\partial X_{long}(Y)/\partial t$ in each column.

### 3.3 Normalization of Causal Sensitivity between Opposite Causal Directions

The *nCS* and *nIF* for the 2-variate system (*X*, *Y*) can also be described by combining equations 1, 2, 6 and 8 into equation 24: the sum of causal sensitivity of *Y* on changing *X*, non-*X*, and the *Y*'s self-generated change, i.e. $md_3Z/md_3Z$, equals 1. Although equation 24 is expressed in terms of *Y* as the effect variable, we can simply swap *X* and *Y* for analysis with *X* as the effect variable. Therefore, to compare the causal sensitivity of opposite causal directions, e.g. $X \rightarrow Y$ and $Y \rightarrow X$, the key is to compare the different denominators in the respective normalization, i.e. the maximal causal sensitivity (max*CS*) for $X \rightarrow Y$ and for $Y \rightarrow X$. Further normalization of the *nCS*, and hypothetically the *nIF*, termed as *nnIF* (equation 25), is based on the larger max*CS* between the two opposite causal directions, taking into account the ratio between the two max*CS* values. This allows the causal sensitivity between two directions, different spaces and times, to be comparable.

Table 1 lists the designed and estimated max*CS* (i.e. the visually calibrated α of equation 14 based on $md_3nIF_a$) for above examples with bidirectional causal influences between *X* (or $X_1$) and *Y* (or $Y_1$), as well as the ratio of the max*CS* between opposite directions.

$$nCS_{[(X\rightarrow Y)\cup(nonX\rightarrow Y)\cup(Y\rightarrow Y)]} = nCS_{(X\rightarrow Y)} + nCS_{(nonX\rightarrow Y)} + \left|nCS_{(X\rightarrow Y)} + nCS_{(nonX\rightarrow Y)} - nCS_{(Y\rightarrow Y)}\right|$$

$$\approx nIF_{(X\rightarrow Y)} + nIF_{(nonX\rightarrow Y)} + \left|nIF_{(X\rightarrow Y)} + nIF_{(nonX\rightarrow Y)} - nIF_{(Y\rightarrow Y)}\right| = 1 \quad (24)$$

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development Discussions
Open Access EGU

If $\max CS_{(Y\to X)} > \max CS_{(X\to Y)}$, $nnIF_{(Y\to X)} = nIF_{(Y\to X)}$, but $nnIF_{(X\to Y)} = nIF_{(X\to Y)} / (\max CS_{(Y\to X)} / \max CS_{(X\to Y)})$ (25)

From the columns listing the percentage errors (i.e. (estimated – designed)/designed), the method reasonably quantifies the max*CS* in both directions as well as their ratio. This direct empirical evidence supports our hypothesis that the normalization of causal sensitivity can be represented by the normalization of information flow. However, imperfection remains:
1) The $md_3nIF_a$-estimated max*CS* tends to be slightly smaller than their designed values in most examples (this suggests that the estimated $|md_3nIF|$ may tend to slightly overestimate the actual *nCS*, and the normalizing factor $md_3Z$ may be slightly underestimated. The reason may be associated with the unstable estimate of $md_2nIF$ (and thus $md_3nIF$) and/or the analogy for defining $md_3nIF$ (equation 23).
2) The larger error in the max*CS* tends to be larger than the error in the ratio between opposite directions. The systematic error suggested in point 1) may tend to be cancelled in the ratio.
3) For the causal sensitivity in a 3D context (Figure 4, S12-S14), although earlier results in 3.1 suggests 3D-to-1D cause map could better reflect the causal contributions according to the spatial pattern of causes, Table 1 suggests the α calibrated in 1D-to-3D effect map to better reflect the designed max*CS*. In other words, for systems that described by the same set of linear causal functions with changing coefficient (due to the trigonometric term, thus changing causal strength) across space and time, the maximal causal sensitivity of such linear function could be better reflected by analysing the influence of weighted average of cause-variable on effect-variable.
4) The error in a higher order causal model (as in Figure 5q-x) tends to be larger than the error in a linear model.

**Table 1. The designed and $md_3nIF_a$-estimated maximal causal sensitivity (max*CS*) in the above examples with bidirectional feedbacks.**

| Respective figure | $\max CS_{(Y\to X)}$ | | | $\max CS_{(X\to Y)}$ | | | $\max CS_{(Y\to X)} / \max CS_{(X\to Y)}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Designed | Estimated | error | Designed | estimated | error | designed | estimated | error |
| 2m-p (S4) | 1.80 | 1.7 | 6% | 0.33 | 0.4 | 20% | 5.4 | 4.3 | 21% |
| 2q-t (S5) | 0.96 | 0.8 | 17% | 0.74 | 0.5 | 32% | 1.3 | 1.6 | 23% |
| 3a-d (S6) | 1.16 | 0.9 | 23% | 0.68 | 0.65 | 4% | 1.7 | 1.4 | 19% |
| S7 | 1.16 | 1.3 | 12% | 0.68 | 0.65 | 4% | 1.7 | 2 | 17% |
| 3e-h (S8) | 1.16 | 1.3 | 12% | 0.68 | 0.7 | 3% | 1.7 | 1.9 | 8% |
| 3i-l (S9) | 1.16 | 0.7 | 40% | 0.68 | 0.55 | 19% | 1.7 | 1.3 | 26% |
| 3m-p(S10) | 1.16 | 0.7 | 40% | 0.68 | 0.55 | 19% | 1.7 | 1.3 | 26% |
| 3q-t(S11) | 1.16 | 0.95 | 18% | 0.68 | 0.5 | 26% | 1.7 | 1.9 | 8% |
| 4 | 0.54 | 0.4 | 27% | 1.30 | 1.05 | 19% | 0.42 | 0.38 | 9% |
| S13 | 0.54 | 0.55 | 1% | 1.30 | 1.4 | 8% | 0.42 | 0.39 | 6% |
| S12 | 0.49 | 0.35 | 28% | 1.35 | 0.85 | 37% | 0.36 | 0.41 | 15% |
| S14 | 0.49 | 0.5 | 3% | 1.35 | 1.2 | 11% | 0.36 | 0.42 | 16% |
| 5q-x(hys) | 1.37 | 0.95 | 31% | 1.15 | 0.6 | 48% | 1.2 | 1.58 | 33% |
| 5q-x(long) | 0.005 | 0.013 | 160% | - | - | - | | | |

## 4. Conclusions

We have shown the applicability of (modified) normalized information flow (particularly the $md_3nIF_a$) to represent normalized causal sensitivity when estimating the causal contributions between two time series variables. The three requirements for such normalization (i.e. normalization for comparable causes, causes from different times, and different spaces) form the conditions for the method to outperform regression analysis: i) when there are strong noise contributions, especially hard-to-quantify independent noise with systematic bias; ii) when there are significant time-lags between causes and effects, especially when we would like to estimate when the causes have occurred; iii) when there are many sources of causal contributions from various spaces, especially when we would like to estimate the location of these causes. We find that (modified) normalized information flow has the potential to serve as a useful tool for understanding complex Earth system processes with multiple interacting variables occurring over various temporal and spatial scales. The estimated causal contributions could be further classified according to their correlation sign, to potentially indicate either positive or negative feedback, thus identifying potential underlying processes. This is the essence for improving Earth System Models.

Some modifications to the Liang's original normalizing factor ($Z$) for the *nIF* are proposed: i) $md_1Z$ is the sum of information flow from cause variables and noises; ii) $md_2Z$ is the self-representing uncertainty flow; and iii) $md_3Z$ is the sum of $md_1Z$ and the absolute difference between $md_1Z$ and $md_2Z$, while the original $Z$ is the sum of $md_1Z$ and $md_2Z$. Apparently, the use of $md_3Z$ helps minimize the error in estimated causal contributions when the estimated correlation sign falsely represents the direction between the two changing variables.

We have demonstrated a potential improvement of the method by preliminarily removing a large and rather constant or constantly growing noise contributions, as well as distinguishing the rate-dependent hysteresis $\partial X_{hys}(Y)/\partial t$ and the state-dependent long-term $\partial X_{long}(Y)/\partial t$ contributions from a second order causal dependency through separated estimation into two sets of linear (rate-dependent) *nIFs*, expanding the potential application of this causal method for complex systems.

Furthermore, we have also proposed the normalization of causal sensitivity between opposite causal directions based on the larger maximal causal sensitivity in two directions. This may serve as a foundation for future work on *universally normalized causal sensitivity* with multivariate systems. The respective estimations of the *IF* and *nIF* for multivariate time series were recently proposed (Liang, 2021a) but not yet tested in the form of causal sensitivities or causal contributions.

**Code and Data Availability**: Data sources are all public databases as indicated at the appropriate point in the text. There is no specific code associated with our analysis, was carried out using standard MatLab routines, with source codes obtainable at dx.doi.org/10.6084/m9.figshare.14985381.

**Author contributions**: CHC conceptualised the project, methodology, data curation and analysis. Both authors contributed to the discussion and manuscript preparation.

**Competing interests**: The authors declare that they have no conflict of interest.

**Acknowledgments:** The early phase of this study was supported by the National Key Research and Development Program of China (2017YFA0603804). This work was supported by NTU Singapore SUG to SATR.

**References**

Andor, M. A. and Fels, K. M.: Behavioral Economics and Energy Conservation – A Systematic Review of Non-price Interventions and Their Causal Effects, Ecol. Econ., 148, 178-210, https://doi.org/10.1016/j.ecolecon.2018.01.018, 2018.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C. D., Krasting, J. P., Law, R. M., Lawrence, D. M., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J. F., Wiltshire, A., Wu, T., and Ziehn, T.: Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models, Biogeosciences, 17, 4173-4222, 10.5194/bg-17-4173-2020, 2020.

Athey, S. and Imbens, G. W.: The State of Applied Econometrics: Causality and Policy Evaluation %J Journal of Economic Perspectives, 31, 3-32, 10.1257/jep.31.2.3, 2017.

Bai, C. Z., Zhang, R., Bao, S. L., Liang, X. S., and Guo, W. B.: Forecasting the Tropical Cyclone Genesis over the Northwest Pacific through Identifying the Causal Factors in Cyclone-Climate Interactions, J. Atmos. Ocean. Technol., 35, 247-259, 10.1175/jtech-d-17-0109.1, 2018.

Barnard, R. L., Blazewicz, S. J., and Firestone, M. K.: Rewetting of soil: Revisiting the origin of soil CO2 emissions, Soil Biology and Biochemistry, 147, 107819, https://doi.org/10.1016/j.soilbio.2020.107819, 2020.

Barnett, L., Barrett, A. B., and Seth, A. K.: Solved problems for Granger causality in neuroscience: A response to Stokes and Purdon, NeuroImage, 178, 744-748, https://doi.org/10.1016/j.neuroimage.2018.05.067, 2018.

Chen, A., Gu, Y., Liu, S., DeAngelis, G. C., and Angelaki, D. E.: Evidence for a Causal Contribution of Macaque Vestibular, But Not Intraparietal, Cortex to Heading Perception, 36, 3789-3798, 10.1523/JNEUROSCI.2485-15.2016 %J The Journal of Neuroscience, 2016.

Cox, P. M., Huntingford, C., and Williamson, M. S.: Emergent constraint on equilibrium climate sensitivity from global temperature variability, Nature, 553, 319-322, 10.1038/nature25450, 2018.

Docquier, D., Vannitsem, S., Ragone, F., Wyser, K., and Liang, X. S.: Causal links between Arctic sea ice and its potential drivers based on the rate of information transfer, Earth and Space Science Open Archive, 15, doi:10.1002/essoar.10507846.1,

Friston, K., Parr, T., Zeidman, P., Razi, A., Flandin, G., Daunizeau, J., Hulme, O., Billig, A., Litvak, V., Moran, R., Price, C., and Lambert, C.: Dynamic causal modelling of COVID-19 [version 2; peer review: 2 approved], 5, 10.12688/wellcomeopenres.15881.2, 2020.

Granger, C. W. J.: Investigating Causal Relations by Econometric Models and Cross-spectral Methods, Econometrica, 37, 424-438, 10.2307/1912791, 1969.

Hall, A., Cox, P., Huntingford, C., and Klein, S.: Progressing emergent constraints on future climate change, Nature Climate Change, 9, 269-278, 10.1038/s41558-019-0436-6, 2019.

Hill, C. A., Suzuki, S., Polania, R., Moisa, M., O'Doherty, J. P., and Ruff, C. C.: A causal account of the brain network computations underlying strategic social behavior, Nature Neuroscience, 20, 1142-1149, 10.1038/nn.4602, 2017.

Im, S.-H., An, S.-I., Kim, S. T., and Jin, F.-F.: Feedback processes responsible for El Niño-La Niña amplitude asymmetry, Geophysical Research Letters, 42, 5556-5563, https://doi.org/10.1002/2015GL064853, 2015.

Liang, X. S.: Unraveling the cause-effect relation between time series, Physical Review E, 90, 052150, 10.1103/PhysRevE.90.052150, 2014.

Liang, X. S.: Normalizing the causality between time series, Physical Review E, 92, 10.1103/PhysRevE.92.022126, 2015.

Liang, X. S.: Information flow and causality as rigorous notions ab initio, Physical Review E, 94, 052201, 10.1103/PhysRevE.94.052201, 2016.

Liang, X. S.: Causation and information flow with respect to relative entropy, Chaos, 28, 10.1063/1.5010253, 2018.

Liang, X. S.: Normalized Multivariate Time Series Causality Analysis and Causal Graph Reconstruction, Entropy, 23, 679, 2021a.

Liang, X. S.: Measuring the importance of individual units in producing the collective behavior of a complex network, 2021b.

Liang, X. S. and Yang, X.-Q.: A Note on Causation versus Correlation in an Extreme Situation, Entropy, 23, 316, 2021.

Liang, X. S., Xu, F., Rong, Y., Zhang, R., Tang, X., and Zhang, F.: El Niño Modoki can be mostly predicted more than 10 years ahead of time, Scientific Reports, 11, 17860, 10.1038/s41598-021-97111-y, 2021.

Lin, S.-H. and Ikram, M. A.: On the relationship of machine learning with causal inference, European Journal of Epidemiology, 35, 183-185, 10.1007/s10654-019-00564-9, 2020.

Luo, Y., Peng, J., and Ma, J.: When causal inference meets deep learning, Nature Machine Intelligence, 2, 426-427, 10.1038/s42256-020-0218-x, 2020.

Nowack, P., Runge, J., Eyring, V., and Haigh, J. D.: Causal networks for climate model evaluation and constrained projections, Nature Communications, 11, 1415, 10.1038/s41467-020-15195-y, 2020.

Obermeier, W. A., Lehnert, L. W., Kammann, C. I., Müller, C., Grünhage, L., Luterbacher, J., Erbs, M., Moser, G., Seibert, R., Yuan, N., and Bendix, J.: Reduced CO2 fertilization effect in temperate C3 grasslands under more extreme weather conditions, Nature Climate Change, 7, 137-141, 10.1038/nclimate3191, 2017.

Ornes, S.: Core Concept: How does climate change influence extreme weather? Impact attribution research seeks answers, Proceedings of the National Academy of Sciences, 115, 8232-8235, 10.1073/pnas.1811393115, 2018.

Pearl, J.: The seven tools of causal inference, with reflections on machine learning, 62, 54–60, 10.1145/3241036, 2019.

Pfrommer, T., Goeschl, T., Proelss, A., Carrier, M., Lenhard, J., Martin, H., Niemeier, U., and Schmidt, H.: Establishing causation in climate litigation: admissibility and reliability, Climatic Change, 152, 67-84, 10.1007/s10584-018-2362-4, 2019.

Rasmussen, S. A., Jamieson, D. J., Honein, M. A., and Petersen, L. R.: Zika Virus and Birth Defects — Reviewing the Evidence for Causality, 374, 1981-1987, 10.1056/NEJMsr1604338, 2016.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D.: Detecting and quantifying causal associations in large nonlinear time series datasets, Science Advances, 5, eaau4996, doi:10.1126/sciadv.aau4996, 2019a.

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., Zhang, K., and Zscheischler, J.: Inferring causation from time series in Earth system sciences, Nature Communications, 10, 2553, 10.1038/s41467-019-10105-3, 2019b.

Russo, F. and Williamson, J.: Epistemic Causality and Evidence-Based Medicine, History and Philosophy of the Life Sciences, 33, 563-581, 2011.

Schreiber, T.: Measuring Information Transfer, Physical Review Letters, 85, 461-464, 10.1103/PhysRevLett.85.461, 2000.

Seth, A. K., Barrett, A. B., and Barnett, L.: Granger Causality Analysis in Neuroscience and Neuroimaging, 35, 3293-3297, 10.1523/JNEUROSCI.4399-14.2015 %J The Journal of Neuroscience, 2015.

Stips, A., Macias, D., Coughlan, C., Garcia-Gorriz, E., and San Liang, X.: On the causal structure between CO2 and global temperature, Scientific Reports, 6, 21691, 10.1038/srep21691, 2016.

Stokes, P. A. and Purdon, P. L.: A study of problems encountered in Granger causality analysis from a neuroscience perspective, 114, E7063-E7072, 10.1073/pnas.1704663114 %J Proceedings of the National Academy of Sciences, 2017.

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., and Munch, S.: Detecting Causality in Complex Ecosystems, 338, 496-500, 10.1126/science.1227079 %J Science, 2012.

Swain, D. L., Singh, D., Touma, D., and Diffenbaugh, N. S.: Attributing Extreme Events to Climate Change: A New Frontier in a Warming World, One Earth, 2, 522-527, https://doi.org/10.1016/j.oneear.2020.05.011, 2020.

https://doi.org/10.5194/gmd-2022-106
Preprint. Discussion started: 25 April 2022

Geoscientific Model Development
Discussions
Open Access

Varian, H. R.: Causal inference in economics and marketing, 113, 7310-7315, 10.1073/pnas.1510479113 %J Proceedings of the National Academy of Sciences, 2016.

Vázquez-Patiño, A., Campozano, L., Mendoza, D., and Samaniego, E.: A causal flow approach for the evaluation of global climate models, 40, 4497-4517, https://doi.org/10.1002/joc.6470, 2020.

Verbitsky, M. Y., Mann, M. E., Steinman, B. A., and Volobuev, D. M.: Detecting causality signal in instrumental measurements and climate model simulations: global warming case study, Geosci. Model Dev., 12, 4053-4060, 10.5194/gmd-12-4053-2019, 2019.

Winkler, A. J., Myneni, R. B., Hannart, A., Sitch, S., Haverd, V., Lombardozzi, D., Arora, V. K., Pongratz, J., Nabel, J. E. M. S., Goll, D. S., Kato, E., Tian, H., Arneth, A., Friedlingstein, P., Jain, A. K., Zaehle, S., and Brovkin, V.: Slow-down of the greening trend in natural vegetation with further rise in atmospheric CO2, Biogeosciences Discuss., 2021, 1-36, 10.5194/bg-2021-37, 2021.