SUPPLEMENTARY MATERIALS

Quantifying Causal Contributions in Earth Systems by Normalized Information Flow

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)

Table S1. Designed functions assessed with results presented in the Figures

Fig. #	Interdependent function	Noise function (inc. self-dependency)
One-Dimensional: nt = 2 to 3000; (or split into 2-1000, 1001-2000, 2001-3000) initial conditions (i.e. nt = 1) are zeros; $\vert_{nt}$ or $\vert_{nt-1}$ refers to nt^{th} or $(nt-1)^{th}$ term on time-axis		
2 & S1 - low noise	$\left.\frac{\partial X1_{Y1}}{\partial t}\right\vert_{nt} = \frac{\cos\left[\frac{\left(2nt+\sum_2^{nt}\frac{1}{nt}\right)}{120\pi}\right]}{\mathbf{1.1}} \times \left.\frac{dY1}{dt}\right\vert_{nt-1} + \frac{\sin\left[\frac{\left(2nt-\sum_2^{nt}\frac{1}{nt}\right)}{120\pi}\right]}{\mathbf{1.5}} \times \left.\frac{dY1}{dt}\right\vert_{nt-1} + \frac{1}{\mathbf{300}}\left.\frac{dY1}{dt}\right\vert_{nt-1}$	$\left.\frac{\partial X1_{noise}}{\partial t}\right\vert_{nt} = \mathbf{0.00001} + \left\vert\frac{\cos\left(\frac{4nt}{200\pi}\right)}{0.9}\right\vert - \frac{X1\vert_{nt-1}}{5000}$
	$\left.\frac{\partial Y1_{X1}}{\partial t}\right\vert_{nt} = \frac{\cos\left[\frac{\left(nt+\sum_2^{nt}\frac{1}{nt}\right)}{310\pi}\right]}{\mathbf{1.8}} \times \left.\frac{dX1}{dt}\right\vert_{nt-1} - \frac{\sin\left[\frac{\left(2nt-\sum_2^{nt}\frac{1}{nt}\right)}{240\pi}\right]}{\mathbf{2.5}} \times \left.\frac{dX1}{dt}\right\vert_{nt-1} + \frac{1}{\mathbf{500}}\left.\frac{dX1}{dt}\right\vert_{nt-1}$	$\left.\frac{\partial Y1_{noise}}{\partial t}\right\vert_{nt} = \mathbf{0.000005} + \left\vert\frac{\sin\left(\frac{2nt}{300\pi}\right)}{0.8}\right\vert - \frac{Y1\vert_{nt-1}}{5000}$
S4 - medium noise	$\left.\frac{\partial X1_{Y1}}{\partial t}\right\vert_{nt} = \frac{\cos\left[\frac{\left(2nt+\sum_2^{nt}\frac{1}{nt}\right)}{120\pi}\right]}{\mathbf{1.8}} \times \left.\frac{dY1}{dt}\right\vert_{nt-1} + \frac{\sin\left[\frac{\left(2nt-\sum_2^{nt}\frac{1}{nt}\right)}{120\pi}\right]}{\mathbf{2.5}} \times \left.\frac{dY1}{dt}\right\vert_{nt-1} + \frac{1}{\mathbf{500}}\left.\frac{dY1}{dt}\right\vert_{nt-1}$	$\left.\frac{\partial X1_{noise}}{\partial t}\right\vert_{nt} = \mathbf{0.01} + \left\vert\frac{\cos\left(\frac{4nt}{\mathbf{500}\pi}\right)}{\mathbf{0.2}}\right\vert - \frac{X1\vert_{nt-1}}{5000}$
	$\left.\frac{\partial Y1_{X1}}{\partial t}\right\vert_{nt} = \frac{\cos\left[\frac{\left(nt+\sum_2^{nt}\frac{1}{nt}\right)}{310\pi}\right]}{\mathbf{2.9}} \times \left.\frac{dX1}{dt}\right\vert_{nt-1} - \frac{\sin\left[\frac{\left(2nt-\sum_2^{nt}\frac{1}{nt}\right)}{240\pi}\right]}{\mathbf{3.9}} \times \left.\frac{dX1}{dt}\right\vert_{nt-1} + \frac{1}{\mathbf{700}}\left.\frac{dX1}{dt}\right\vert_{nt-1}$	$\left.\frac{\partial Y1_{noise}}{\partial t}\right\vert_{nt} = \mathbf{0.005} + \left\vert\frac{\sin\left(\frac{2nt}{\mathbf{600}\pi}\right)}{\mathbf{0.17}}\right\vert - \frac{Y1\vert_{nt-1}}{5000}$
4 & S3 - strong noise	$\left.\frac{\partial X1_{Y1}}{\partial t}\right\vert_{nt} = \frac{\cos\left[\frac{\left(2nt+\sum_2^{nt}\frac{1}{nt}\right)}{120\pi}\right]}{\mathbf{1.8}} \times \left.\frac{dY1}{dt}\right\vert_{nt-1} + \frac{\sin\left[\frac{\left(2nt-\sum_2^{nt}\frac{1}{nt}\right)}{120\pi}\right]}{\mathbf{3.2}} \times \left.\frac{dY1}{dt}\right\vert_{nt-1} + \frac{1}{\mathbf{500}}\left.\frac{dY1}{dt}\right\vert_{nt-1}$	$\left.\frac{\partial X1_{noise}}{\partial t}\right\vert_{nt} = \mathbf{0.001} + \left\vert\frac{\cos\left(\frac{4nt}{\mathbf{550}\pi}\right)}{\mathbf{0.15}}\right\vert - \frac{X1\vert_{nt-1}}{\mathbf{50000}}$

	$\left.\frac{\partial Y1_{X1}}{\partial t}\right\|_{nt} = \frac{\cos\left[\frac{\left(nt+\sum_{2}^{nt}\frac{1}{nt}\right)}{310\pi}\right]}{\mathbf{2.9}} \times \left.\frac{dX1}{dt}\right\|_{nt-1} - \frac{\sin\left[\frac{\left(2nt-\sum_{2}^{nt}\frac{1}{nt}\right)}{240\pi}\right]}{\mathbf{4.2}} \times \left.\frac{dX1}{dt}\right\|_{nt-1} + \frac{1}{\mathbf{700}}\left.\frac{dX1}{dt}\right\|_{nt-1}$	$\left.\frac{\partial Y1_{noise}}{\partial t}\right\|_{nt} = \mathbf{0.0005} + \left\|\frac{\sin\left(\frac{2nt}{\mathbf{650}\pi}\right)}{\mathbf{0.13}}\right\| - \frac{Y1\|_{nt-1}}{\mathbf{50000}}$
3 & S2 - low noise	Same as for Fig. 2 & S1, except for nt = 1001-2000, replace $\|_{nt-1}$ by $\|_{nt-20}$ for nt = 2001-3000, replace $\|_{nt-1}$ by $\|_{nt-40}$	Same as for Fig. 2 & S1
S5- medium noise	Same as for Fig. S4, except for nt = 1001-2000, replace $\|_{nt-1}$ by $\|_{nt-20}$ for nt = 2001-3000, replace $\|_{nt-1}$ by $\|_{nt-40}$	Same as for Fig. S4
S6- strong noise	Same as for Fig. 4 & S3, except for nt = 1001-2000, replace $\|_{nt-1}$ by $\|_{nt-20}$ for nt = 2001-3000, replace $\|_{nt-1}$ by $\|_{nt-40}$	Same as for Fig. 4 & S3
Three-Dimensional (nlon, nlat, t) with nlon = 1 to 360, nlat = 1 to 180 in loop calculations initial conditions (t = 1) are zeros; for N-S teleconnection, the cause-variable are from **181-nlat**		
7 & S10 N-S tele-connection moderate-frequency	$\left.\frac{\partial X_Y}{\partial t}\right\|_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+2t}{\mathbf{270}\pi}\right]}{\mathbf{9}} \times \left.\frac{dY}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{\sin\left[\frac{(\mathbf{181-nlat})+3t}{\mathbf{280}\pi}\right]}{\mathbf{10}} \times \left.\frac{dY}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{1}{\mathbf{3}}\left.\frac{dY}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}}$ $\left.\frac{\partial Y_X}{\partial t}\right\|_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+3t}{\mathbf{290}\pi}\right]}{\mathbf{9}} \times \left.\frac{dX}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{\sin\left[\frac{(\mathbf{181-nlat})+2t}{\mathbf{260}\pi}\right]}{\mathbf{4}} \times \left.\frac{dX}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{1}{\mathbf{1.4}}\left.\frac{dX}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}}$	$\left.\frac{\partial X_{noise}}{\partial t}\right\|_{t,nlon,nlat} = \mathbf{0.03} + \left\|\frac{\cos\left(\frac{nlon+\mathbf{5}t}{\mathbf{320}\pi}\right)}{\mathbf{1.2}}\right\| + \left\|\frac{\sin\left(\frac{nlat+\mathbf{8}t}{\mathbf{330}\pi}\right)}{\mathbf{2}}\right\|$ $\left.\frac{\partial Y_{noise}}{\partial t}\right\|_{t,nlon,nlat} = \mathbf{0.02} + \left\|\frac{\cos\left(\frac{nlon+\mathbf{4}t}{\mathbf{350}\pi}\right)}{\mathbf{2}}\right\| + \left\|\frac{\sin\left(\frac{nlat+\mathbf{12}t}{\mathbf{340}\pi}\right)}{\mathbf{3}}\right\|$
S7 & S11 N-S tele-connection High-frequency	$\left.\frac{\partial X_Y}{\partial t}\right\|_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+2t}{\mathbf{120}\pi}\right]}{\mathbf{7}} \times \left.\frac{dY}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{\sin\left[\frac{(\mathbf{181-nlat})+3t}{\mathbf{140}\pi}\right]}{\mathbf{8}} \times \left.\frac{dY}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{1}{\mathbf{2.5}}\left.\frac{dY}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}}$ $\left.\frac{\partial Y_X}{\partial t}\right\|_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+3t}{\mathbf{135}\pi}\right]}{\mathbf{2.5}} \times \left.\frac{dX}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{\sin\left[\frac{(\mathbf{181-nlat})+2t}{\mathbf{125}\pi}\right]}{\mathbf{3}} \times \left.\frac{dX}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}} + \frac{1}{\mathbf{1.3}}\left.\frac{dX}{dt}\right\|_{t-1,nlon,\mathbf{181-nlat}}$	$\left.\frac{\partial X_{noise}}{\partial t}\right\|_{t,nlon,nlat} = \mathbf{0.02} + \left\|\frac{\cos\left(\frac{nlon+\mathbf{8}t}{\mathbf{105}\pi}\right)}{\mathbf{1.2}}\right\| + \left\|\frac{\sin\left(\frac{nlat+\mathbf{9}t}{\mathbf{120}\pi}\right)}{\mathbf{1.7}}\right\|$ $\left.\frac{\partial Y_{noise}}{\partial t}\right\|_{t,nlon,nlat} = \mathbf{0.01} + \left\|\frac{\cos\left(\frac{nlon+\mathbf{6}t}{\mathbf{115}\pi}\right)}{\mathbf{2}}\right\| + \left\|\frac{\sin\left(\frac{nlat+\mathbf{12}t}{\mathbf{125}\pi}\right)}{\mathbf{3}}\right\|$

S8 & S12 no tele-connection moderate-frequency	$\frac{\partial X_Y}{\partial t}\Big\rvert_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+\mathbf{2}t}{\mathbf{240}\pi}\right]}{\mathbf{8}} \times \frac{dY}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{\sin\left[\frac{\boldsymbol{nlat}+\mathbf{2}t}{\mathbf{250}\pi}\right]}{\mathbf{9}} \times \frac{dY}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{1}{\mathbf{4}}\frac{dY}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}}$ $\frac{\partial Y_X}{\partial t}\Big\rvert_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+\mathbf{4}t}{\mathbf{230}\pi}\right]}{\mathbf{3.5}} \times \frac{dX}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{\sin\left[\frac{\boldsymbol{nlat}+\mathbf{2}t}{\mathbf{260}\pi}\right]}{\mathbf{2.5}} \times \frac{dX}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{1}{\mathbf{1.5}}\frac{dX}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}}$	$\frac{\partial X_{noise}}{\partial t}\Big\rvert_{t,nlon,nlat} = \mathbf{0.0005} + \left\lvert\frac{\cos\left(\frac{nlon+\mathbf{4}t}{\mathbf{220}\pi}\right)}{\mathbf{1.2}}\right\rvert + \left\lvert\frac{\sin\left(\frac{nlat+\mathbf{4}t}{\mathbf{210}\pi}\right)}{\mathbf{1.1}}\right\rvert$ $\frac{\partial Y_{noise}}{\partial t}\Big\rvert_{t,nlon,nlat} = \mathbf{0.001} + \left\lvert\frac{\cos\left(\frac{nlon+\mathbf{10}t}{\mathbf{230}\pi}\right)}{\mathbf{1.7}}\right\rvert + \left\lvert\frac{\sin\left(\frac{nlat+\mathbf{2}t}{\mathbf{200}\pi}\right)}{\mathbf{2.5}}\right\rvert$
S9 & S13 no tele-connection high-frequency	$\frac{\partial X_Y}{\partial t}\Big\rvert_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+\mathbf{4}t}{\mathbf{120}\pi}\right]}{\mathbf{8}} \times \frac{dY}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{\sin\left[\frac{\boldsymbol{nlat}+\mathbf{3}t}{\mathbf{130}\pi}\right]}{\mathbf{9}} \times \frac{dY}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{1}{\mathbf{4}}\frac{dY}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}}$ $\frac{\partial Y_X}{\partial t}\Big\rvert_{t,nlon,nlat} = \frac{\cos\left[\frac{nlon+\mathbf{4}t}{\mathbf{110}\pi}\right]}{\mathbf{3}} \times \frac{dX}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{\sin\left[\frac{\boldsymbol{nlat}+\mathbf{2}t}{\mathbf{110}\pi}\right]}{\mathbf{4}} \times \frac{dX}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}} + \frac{1}{\mathbf{1.2}}\frac{dX}{dt}\Big\rvert_{t-1,nlon,\boldsymbol{nlat}}$	$\frac{\partial X_{noise}}{\partial t}\Big\rvert_{t,nlon,nlat} = \mathbf{0.003} + \left\lvert\frac{\cos\left(\frac{nlon+\mathbf{15}t}{\mathbf{100}\pi}\right)}{\mathbf{1.2}}\right\rvert + \left\lvert\frac{\sin\left(\frac{nlat+\mathbf{8}t}{\mathbf{90}\pi}\right)}{\mathbf{1.1}}\right\rvert$ $\frac{\partial Y_{noise}}{\partial t}\Big\rvert_{t,nlon,nlat} = \mathbf{0.006} + \left\lvert\frac{\cos\left(\frac{nlon+\mathbf{10}t}{\mathbf{80}\pi}\right)}{\mathbf{2}}\right\rvert + \left\lvert\frac{\sin\left(\frac{nlat+\mathbf{12}t}{\mathbf{110}\pi}\right)}{\mathbf{3}}\right\rvert$

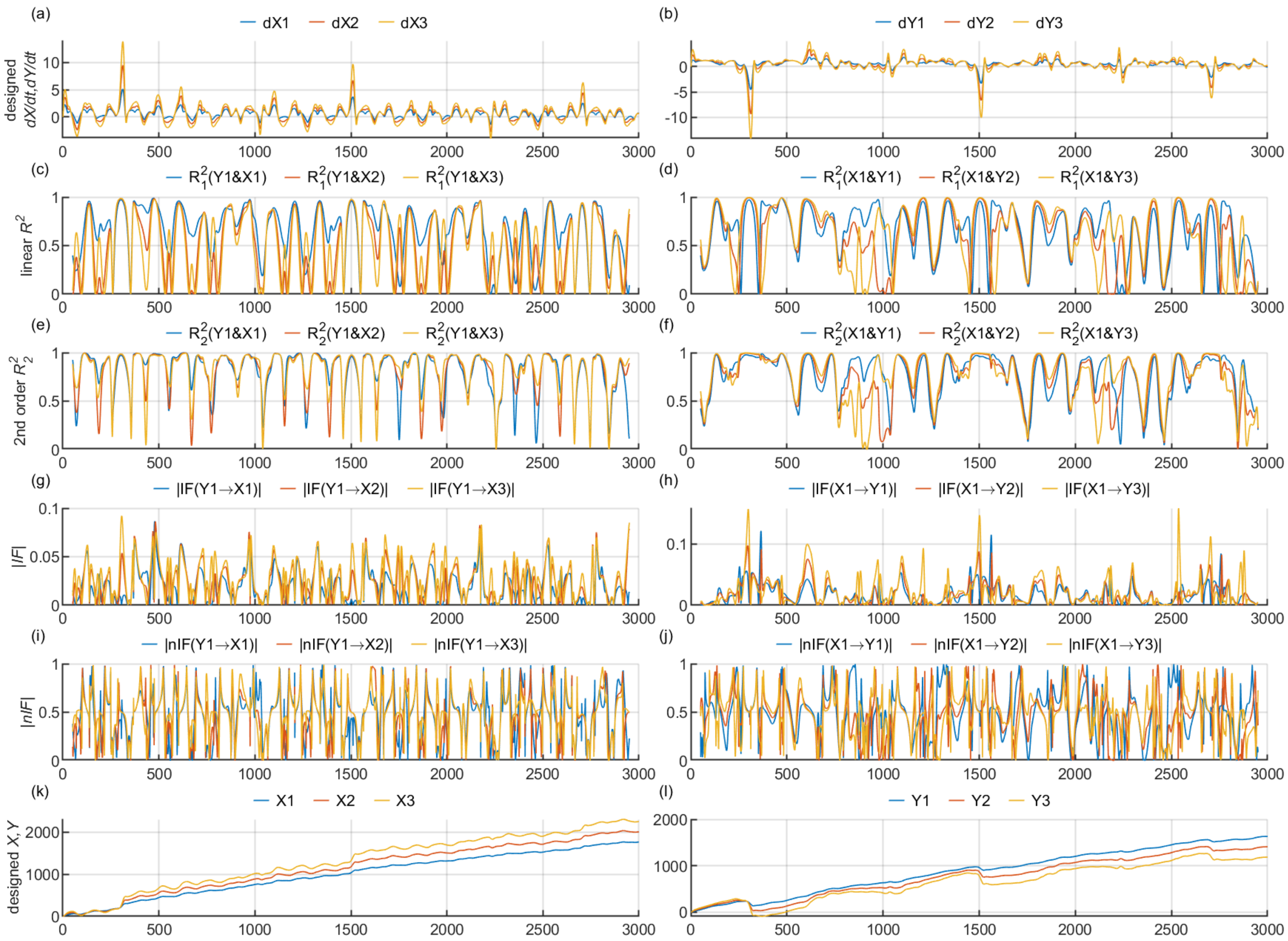


Figure S1. The overall trend of designed interdependent variables (a, b, k, l) and the quantifiers in various methods (c-j), as supplementary information for Fig. 2. Note that the $|nIF|$ is roughly independent from the linear 1:2:3 interdependency ratio (i, k). On the other hand, the $|IF|$, however, occasionally become dependent with this linear ratio.

0

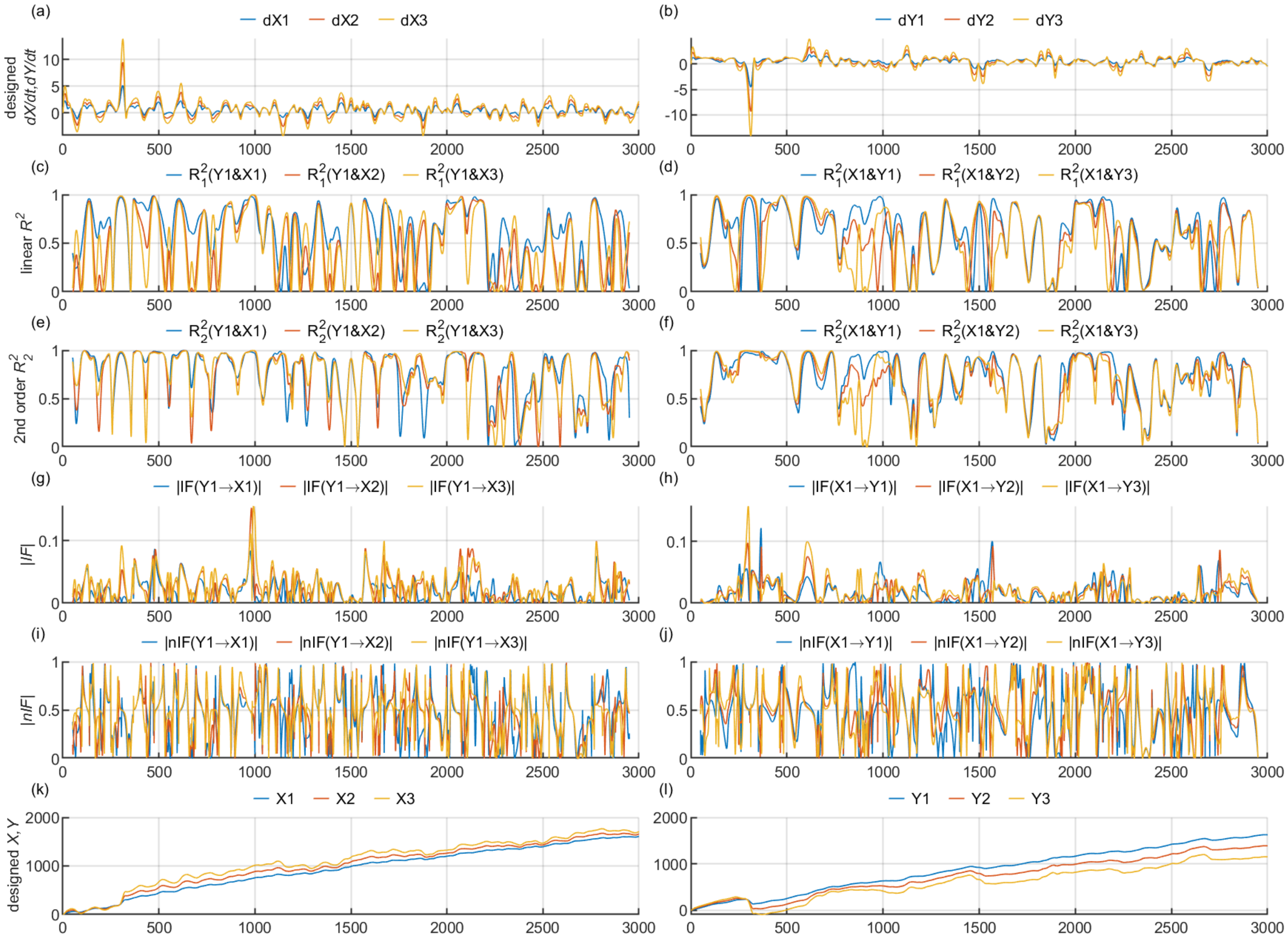


Figure S2. The overall trend of designed interdependent variables (a, b, k, l) and the quantifiers in various methods (c-j), as supplementary information for Fig. 3.

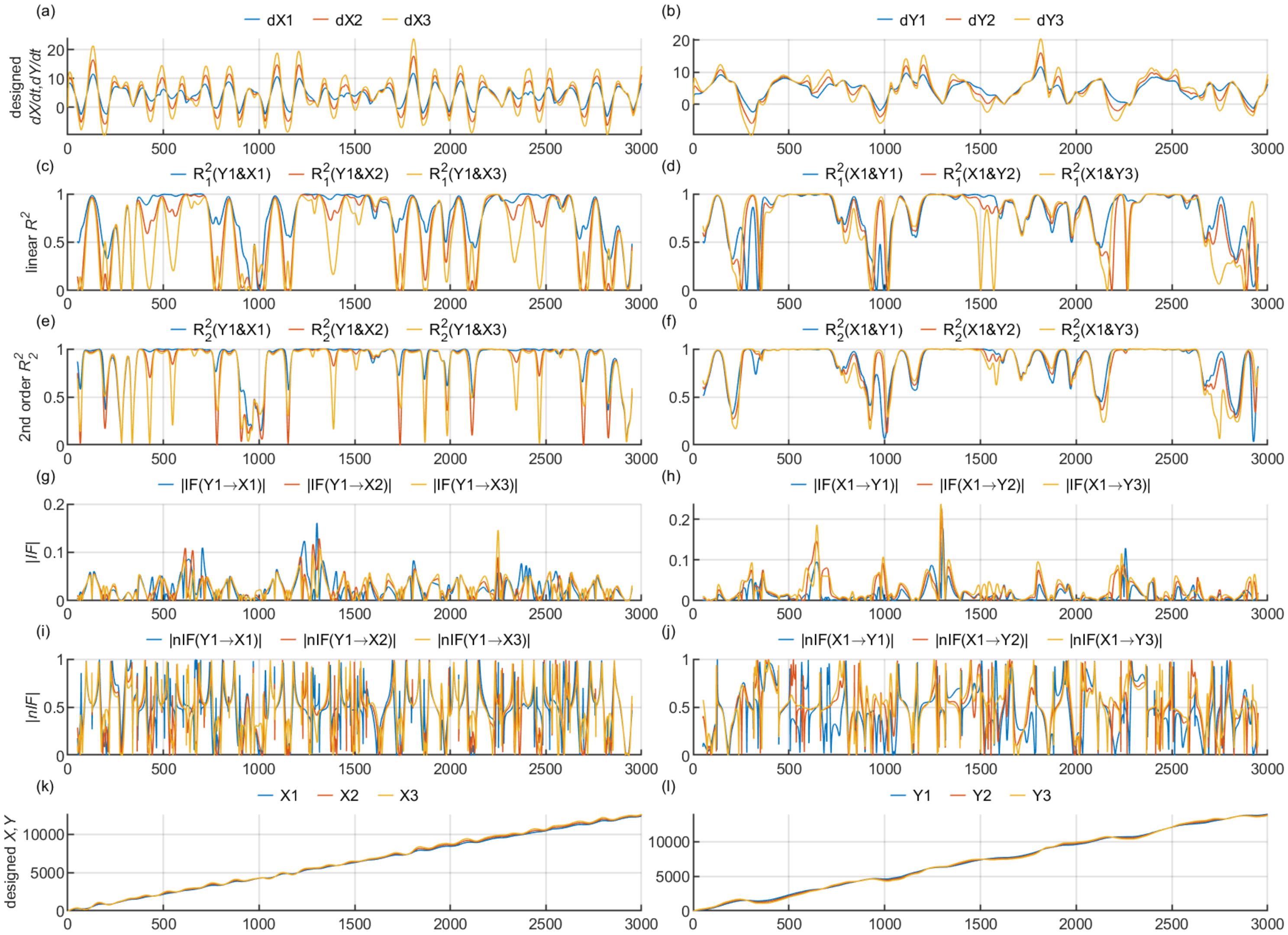


Figure S3. The overall trend of designed interdependent variables (a, b, k, l) and the quantifiers in various methods (c-j), as supplementary information for Fig. 4. Note that under the large positive noise influence, the $|IF|$ becomes roughly independent from the 1:2:3 ratio.

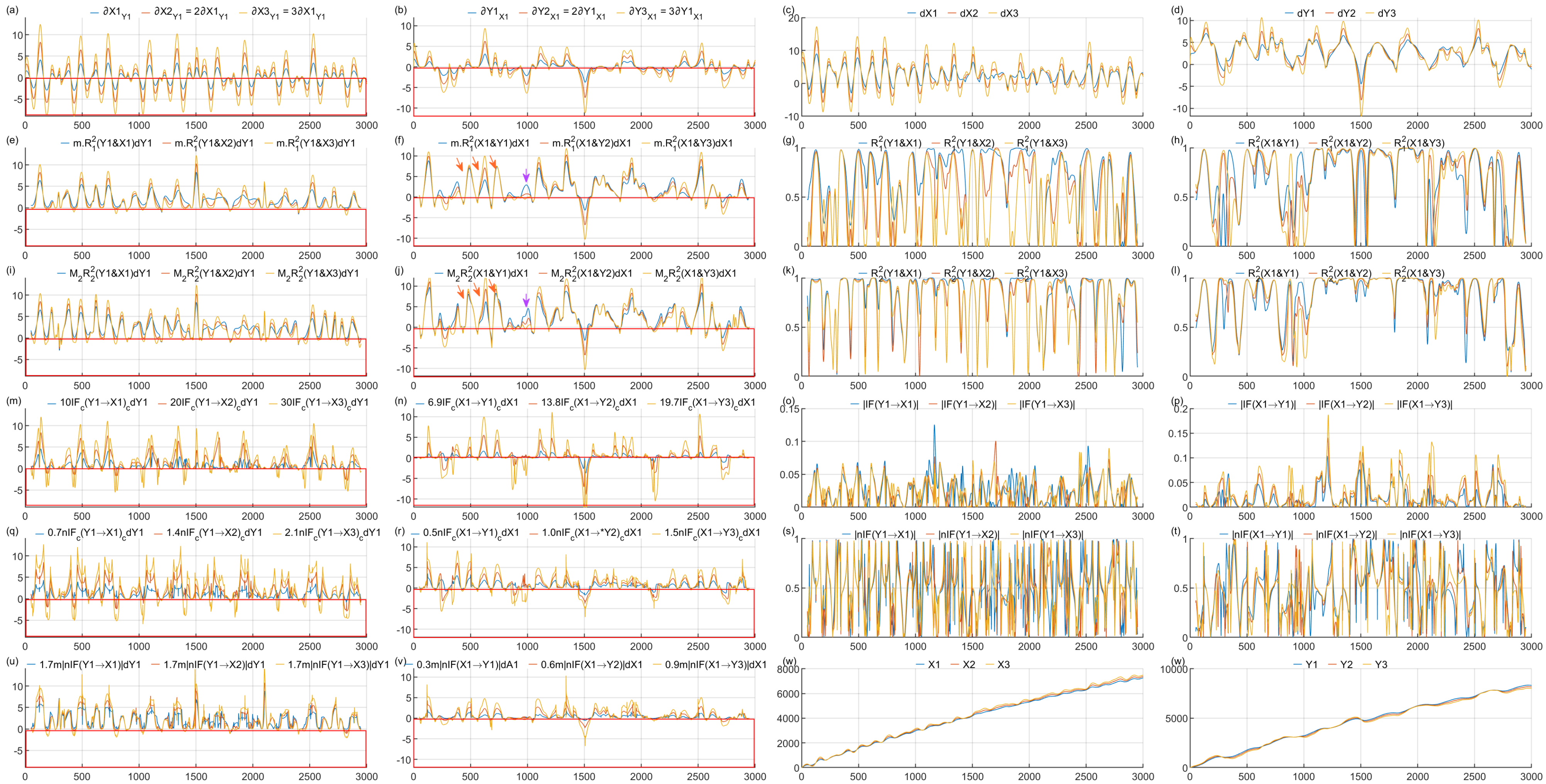


Figure S4. Assessment of methods for estimating interdependent contribution between designed 1D variables *X* and *Y*, highlighting the conditions of medium noise-contributions. The figure layout is as the combination of Fig. 2 and Fig. S1: the first two columns compare the designed and estimated causal contributions, and the last two columns show the designed *dX/dt*, *dY/dt*, *X*, *Y*, and multipliers of various methods. Note the additional occurrence of opposite (negative) contributions to the common (positive) trend captured by nIF_c (highlighted in red boxes as compared to Fig.2). The orange arrows in f and j show inconsistent capability representing the 1:2:3 ratio especially when the signals are falsely amplified. The purple arrows indicate signals with a wrong feedback sign due to opposite causal signals to the common trend.

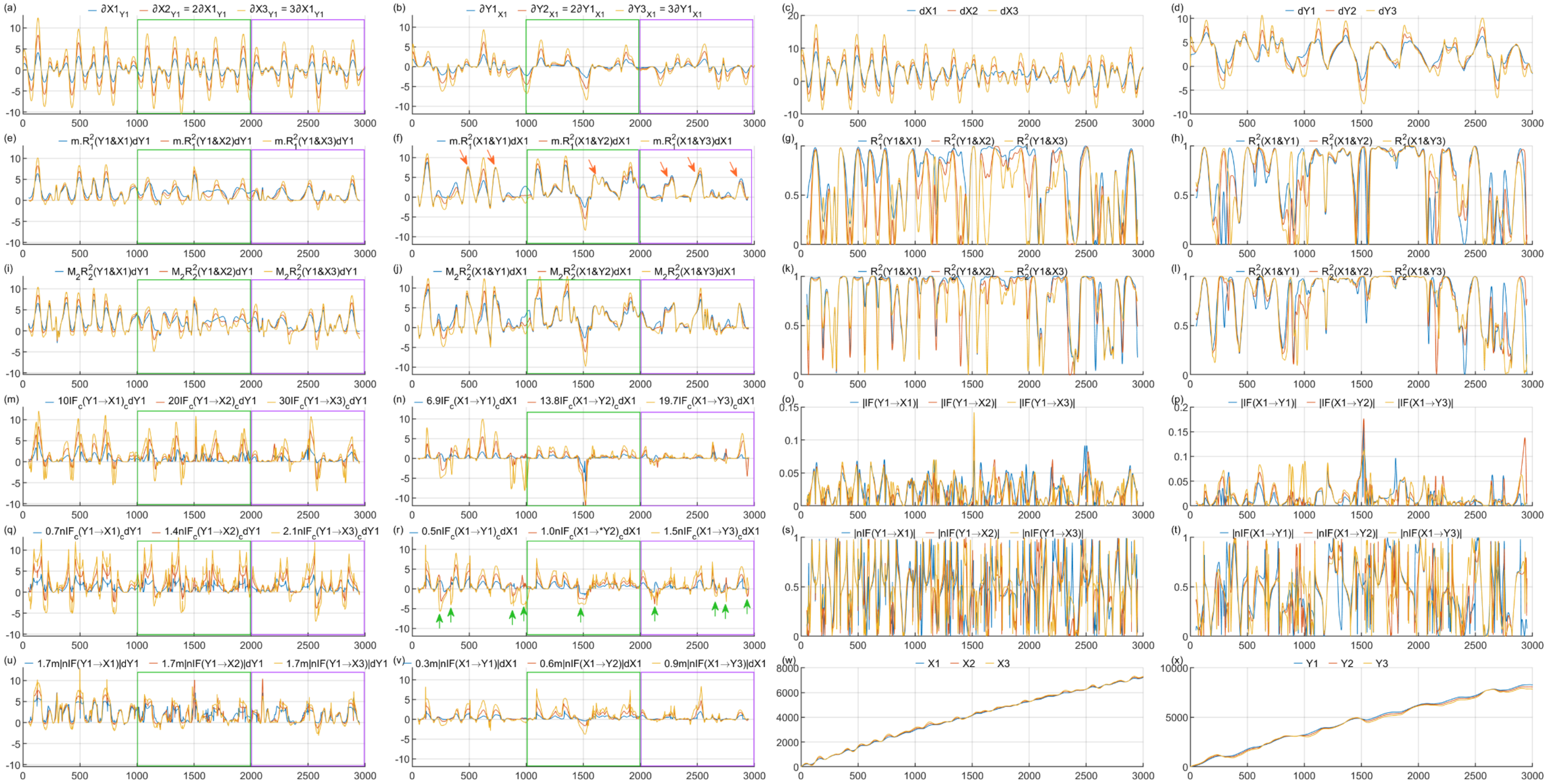
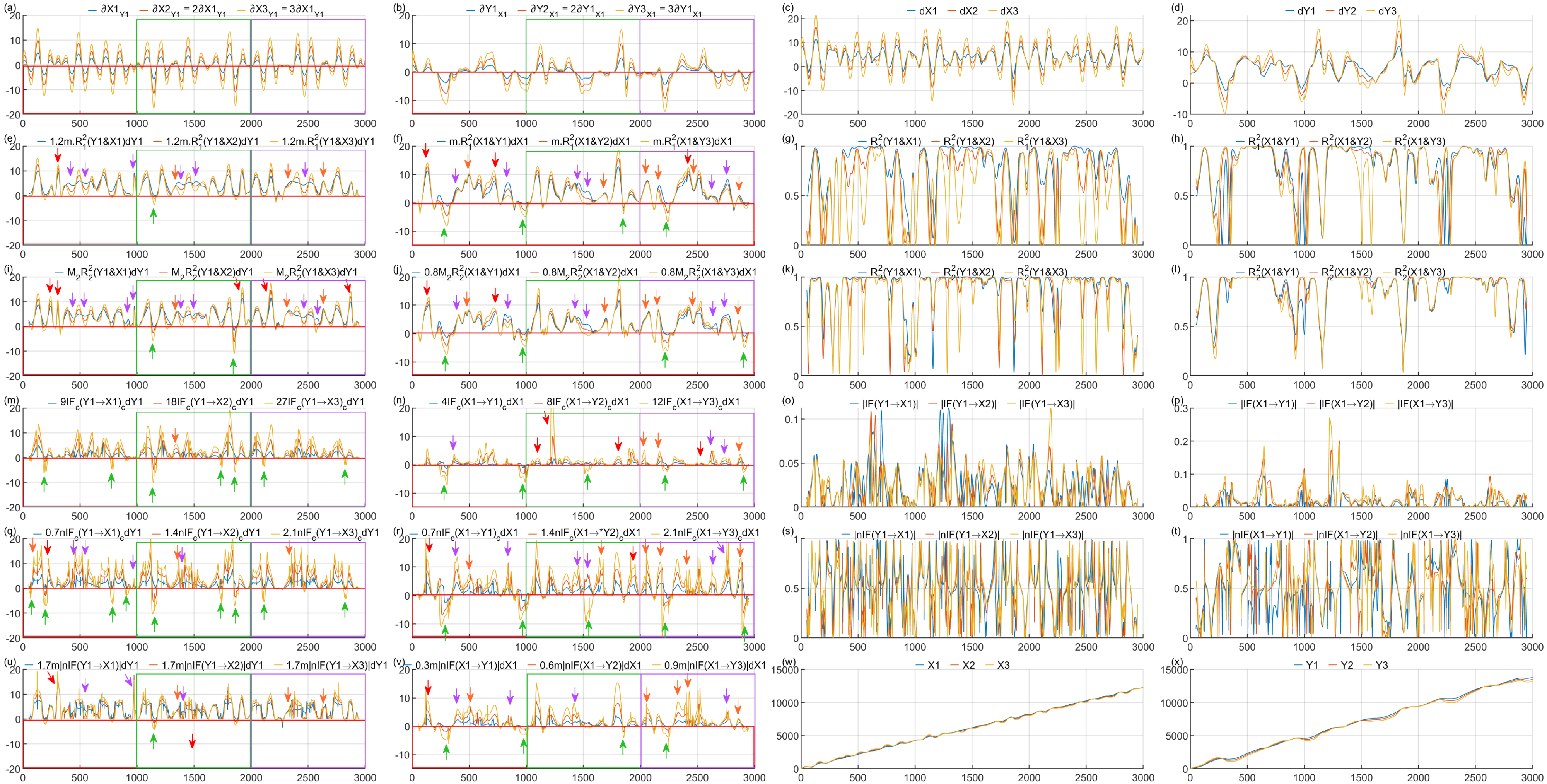


Figure S5. Assessment of methods for estimating interdependent contribution between designed 1D variables *X* and *Y*, under multiple conditions including opposite contribution to the common trend, time-lag between the cause and consequence, and medium noise-contributions. Same figure layout as in Fig. S4 and the coloured boxes with time-lag causality as in Fig. 3. The multiple conditions follow about the same trends as in Fig. 2,3, and S4: nIF_c provides better estimates with opposite contribution and time-lag causality. The increased noise-contribution results in more negative contribution with positive trend (green arrows in r, as compared to Fig. 3b/j). It also limits the capability of regression in differentiating the 1:2:3 ratio and possibly lead to more false signals or even wrong feedback sign (e.g. those indicated by orange arrows in f). Estimates by nIF_c also suffer lower accuracy.



45 **Figure S6.** Assessment of methods for estimating interdependent contribution between designed 1D variables *X* and *Y*, under multiple conditions including opposite contribution to the common trend, time-lag between the cause and consequence, and large noise-contributions. Same figure layout and arrows' meanings as in Fig. S4. In addition to the conditions described in Fig. 4, there is no significant additional influence from the time-lag causality under this strong influence of noise-signals, but it may negatively affect the relatively better estimates by IF_c and nIF_c more than the already problematic estimates by regressions, as compared to Fig. 4.

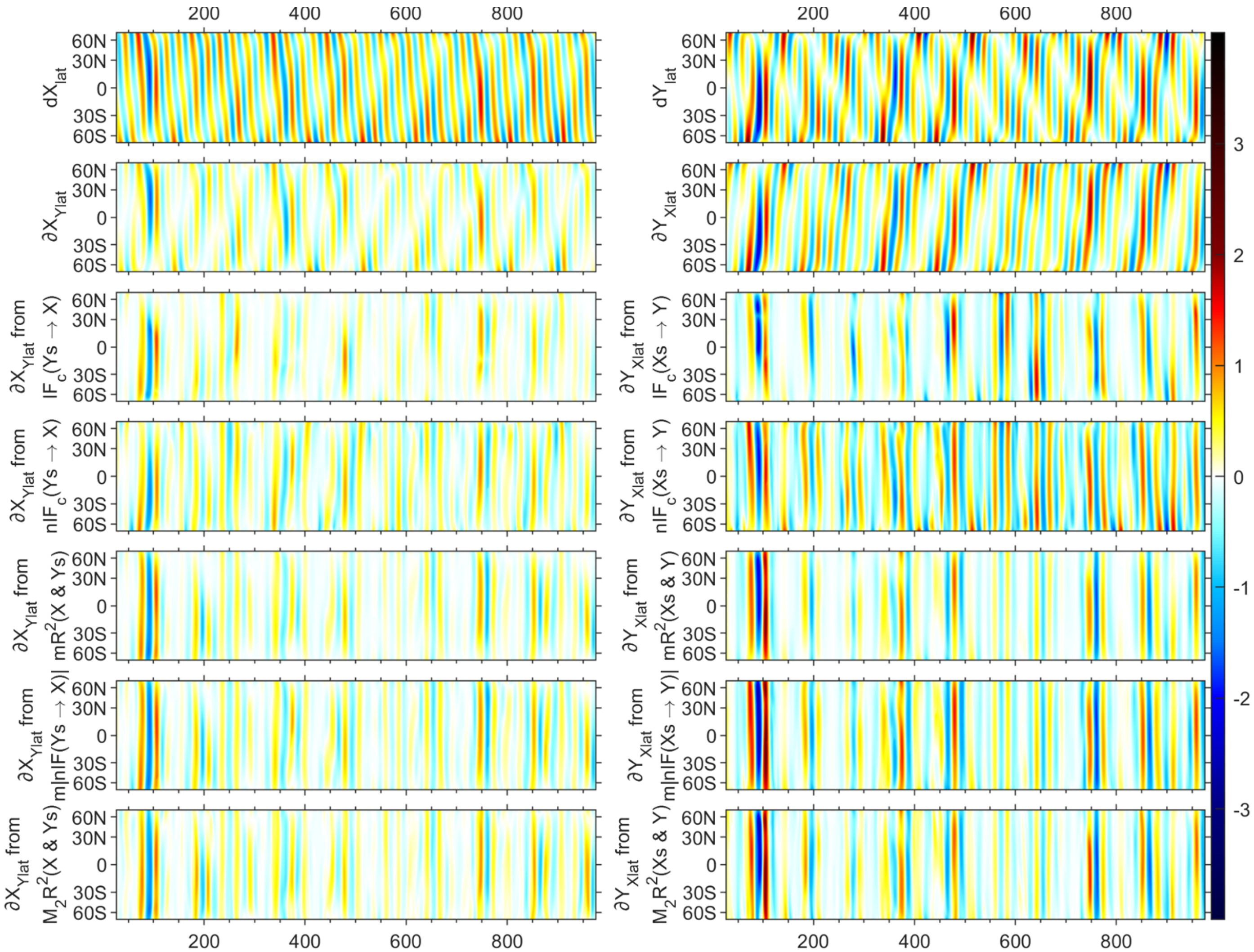


Figure S7. Assessment of methods for estimating interdependent contribution between designed variables *X* and *Y* (cause-maps), at high variability frequency (~4 *dX/dt* sign-changes per 49 time-units) with only the north-south meridional teleconnection. The distribution of estimated causes (row 3-7) is supposed to mirror the designed effects (row 2) in the meridional direction. The estimates by nIF_c appear to best capture the opposite-direction causal signals, followed by IF_c, and the regressions are the least capable of capturing this teleconnection. However, with such high variability frequency, the errors on the temporal dimension for estimates by IF_c and nIF_c are enlarged, and accuracy of estimates by regressions and $m|nIF|$ appear more consistent on the temporal dimension.

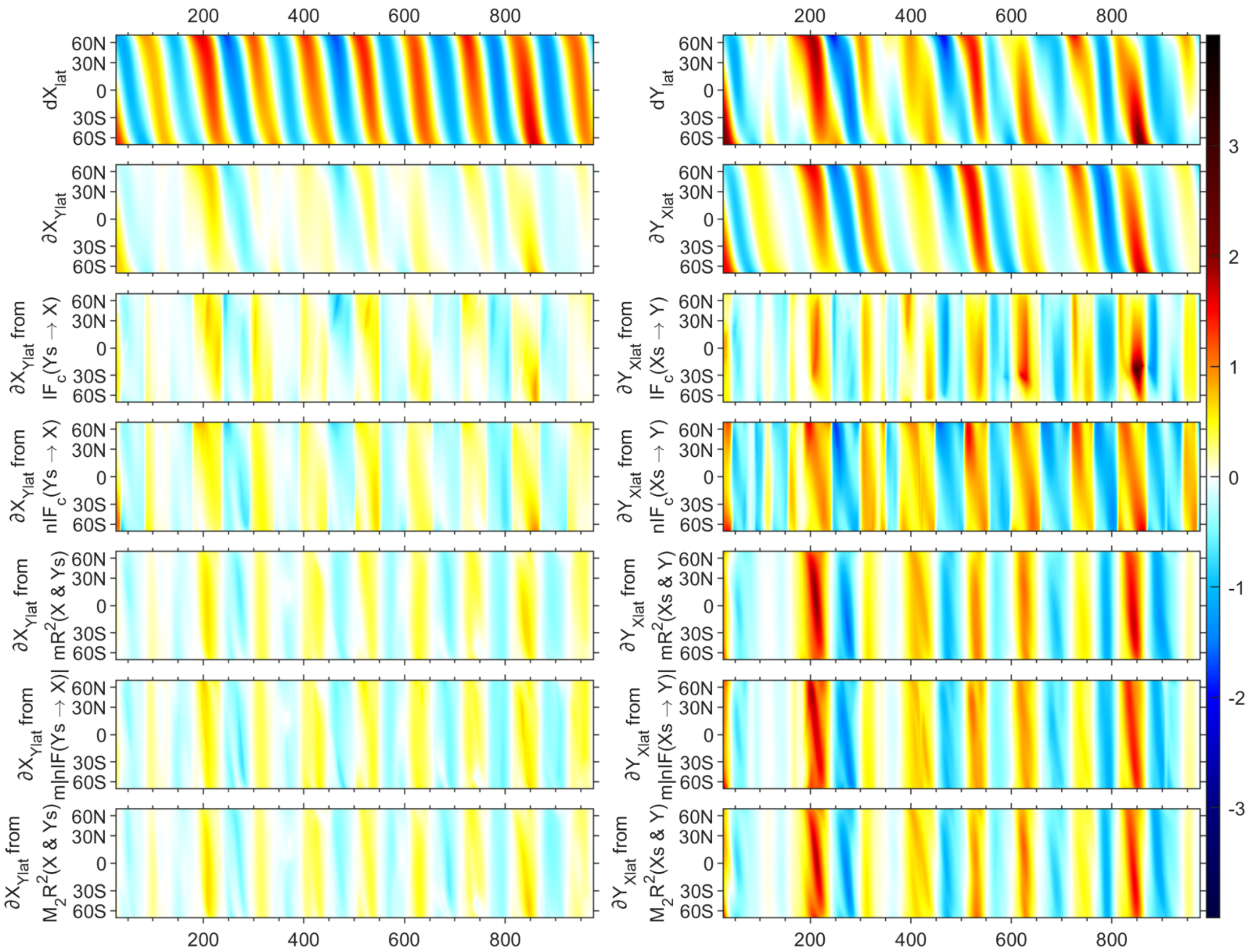


Figure S8. Assessment of methods for estimating interdependent contribution between designed variables *X* and *Y* (cause-maps), at low variability frequency (~1 *dX/dt* sign-change per 49 time-units) with no teleconnection. All methods reasonably capture the interdependent contributions, but estimates by mR^2, $m|nIF|$, and M_2R^2 may be better for temporal variation when signals are strong (dY_{Xlat}) while estimates by IF_c, and IF_c with low interdependency (dX_{Ylat}) appear to better capture the spatial distribution, and possibly temporal variation for weak dX_{Ylat} signals.

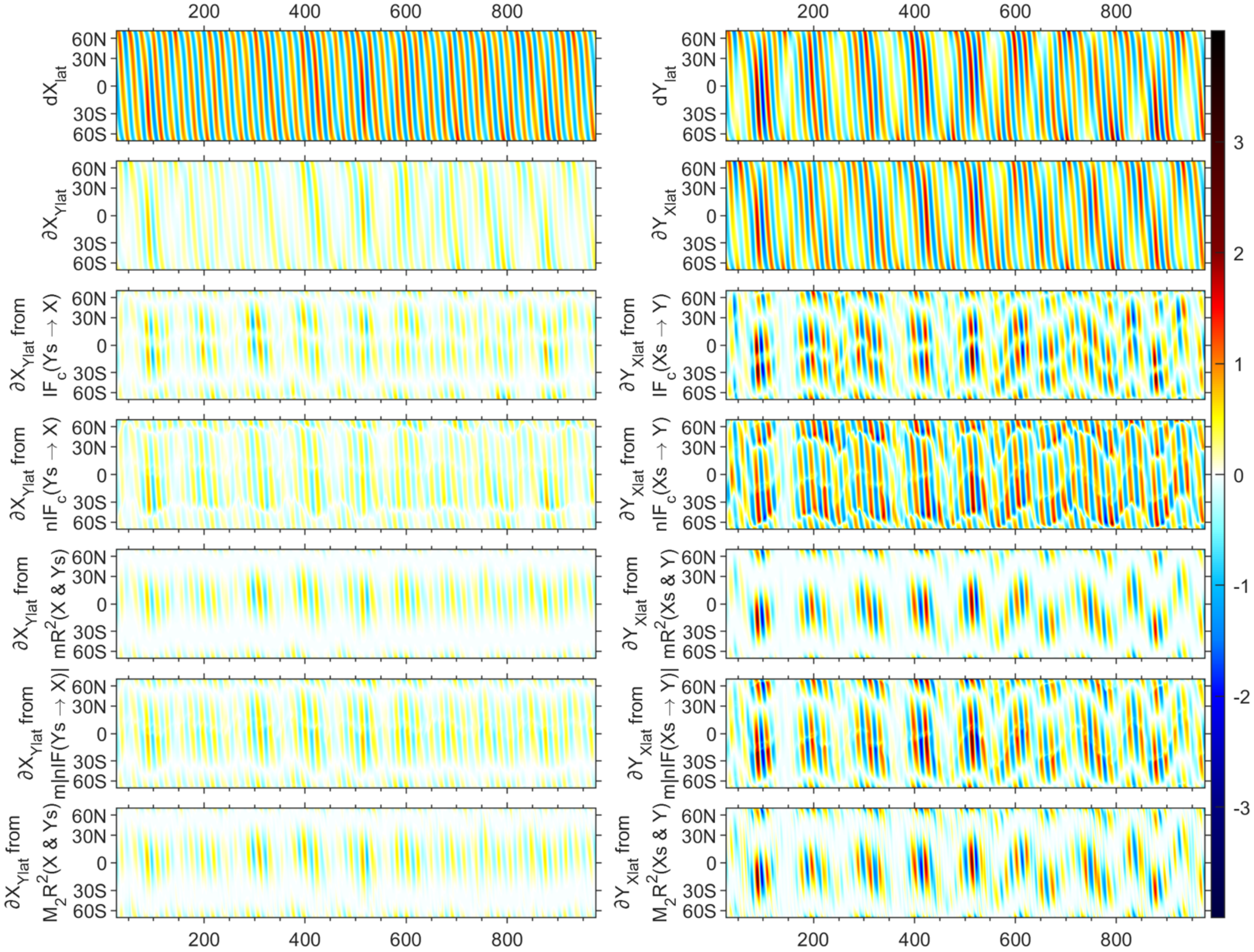


Figure S9. Assessment of methods for estimating interdependent contribution between designed variables *X* and *Y* (cause-maps), at high variability frequency (~4-5 *dX/dt* sign-changes per 49 time-units) with no teleconnection. Under this high variability frequency, the ability to capture the interdependent contributions is limited. Errors are especially large at high latitudes with trends (feedback signs) different from low-mid latitudes where represent the major trends at a specific time. Nevertheless, the estimates by nIF_c appear to better capture the distribution, especially on the spatial dimensions.

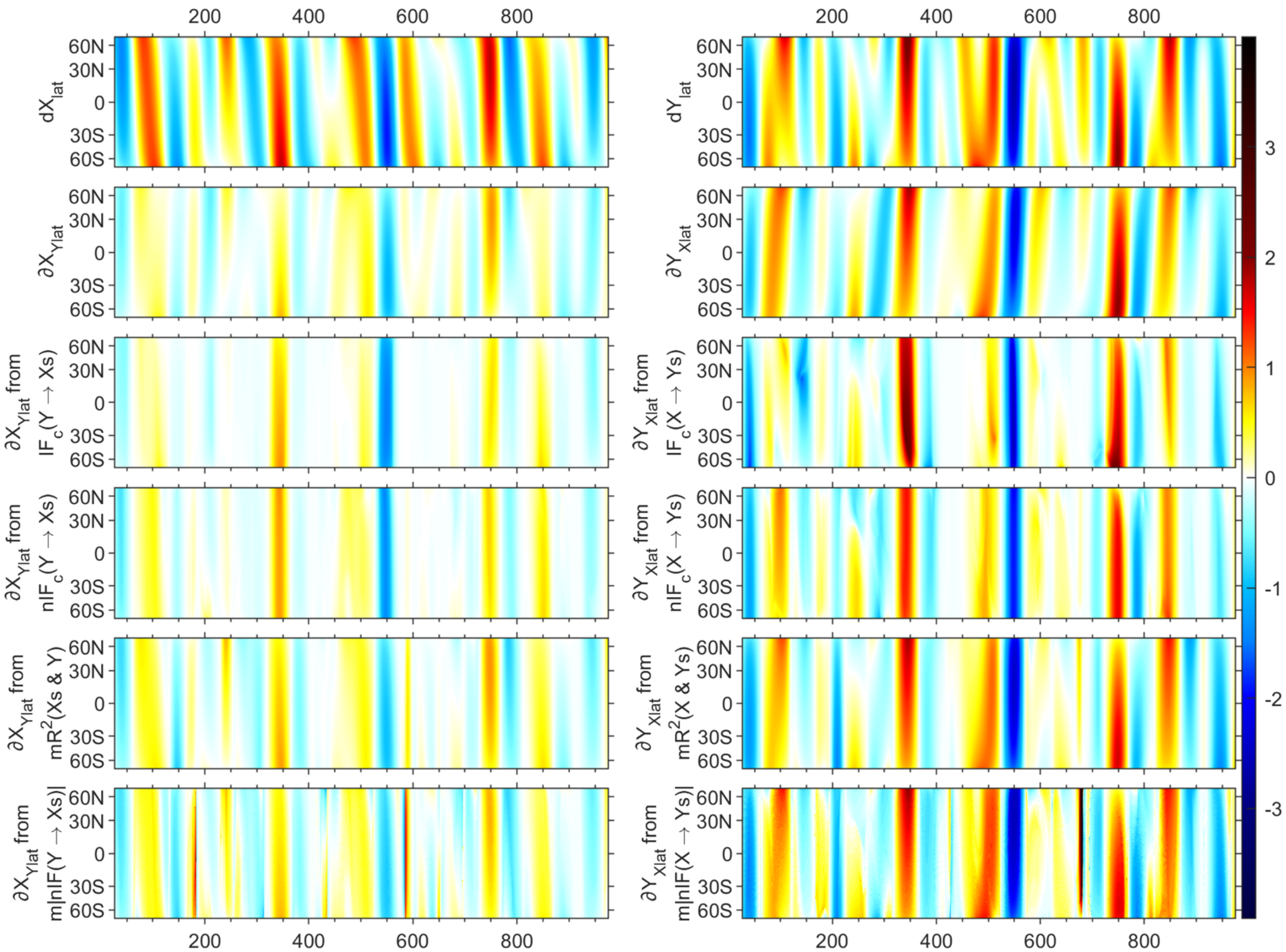


Figure S10. Assessment of methods for estimating interdependent contribution between designed variables *X* and *Y* (effect-maps), at low variability frequency (~1.1 *dX/dt* sign change per 49 time-units) with only the north-south meridional teleconnection. The distribution of estimated effects (row 3-6) is supposed to equal to the designed effects (row 2) in the meridional direction. Estimates by mR^2 appear to give the best accuracy on both temporal and spatial dimensions. There is no significant merit for other methods.

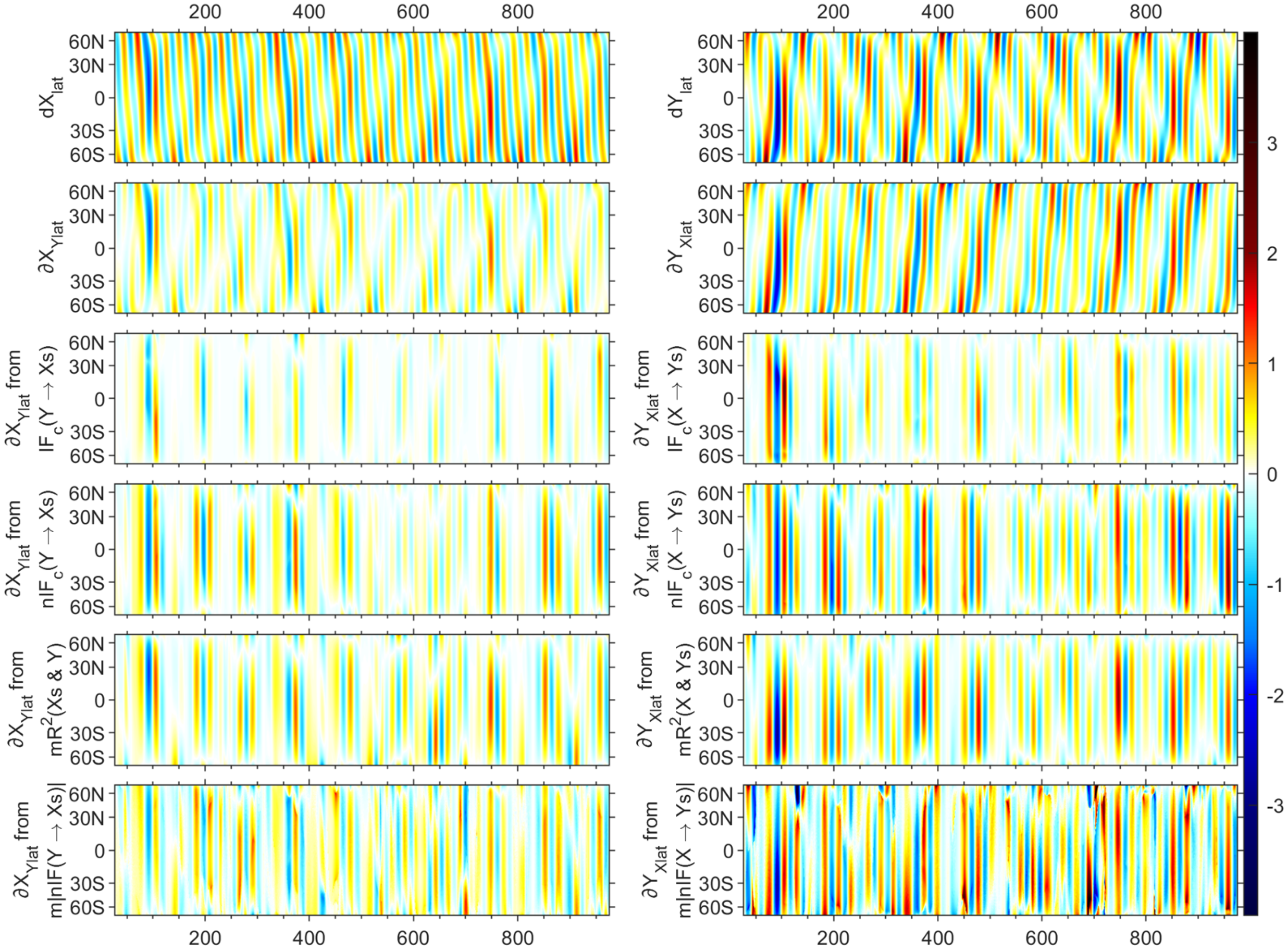


Figure S11. Assessment of methods for estimating interdependent contribution between designed variables *X* and *Y* (effect-maps), at high variability frequency (~4 *dX/dt* sign change per 49 time-units) with only the north-south meridional teleconnection. The distribution of estimated effects (row 3-6) is supposed to equal to the designed effects (row 2) in the meridional direction. Estimates by mR^2 appear to give the best accuracy on both temporal and spatial dimensions. Although the estimates by $m|nIF|$ may better capture few weak signals, it appears to introduce more errors with overestimated weak signals. For the temporal dimension, the estimates by mR^2 and $m|nIF|$ could be more reliable especially at high-frequency variability.

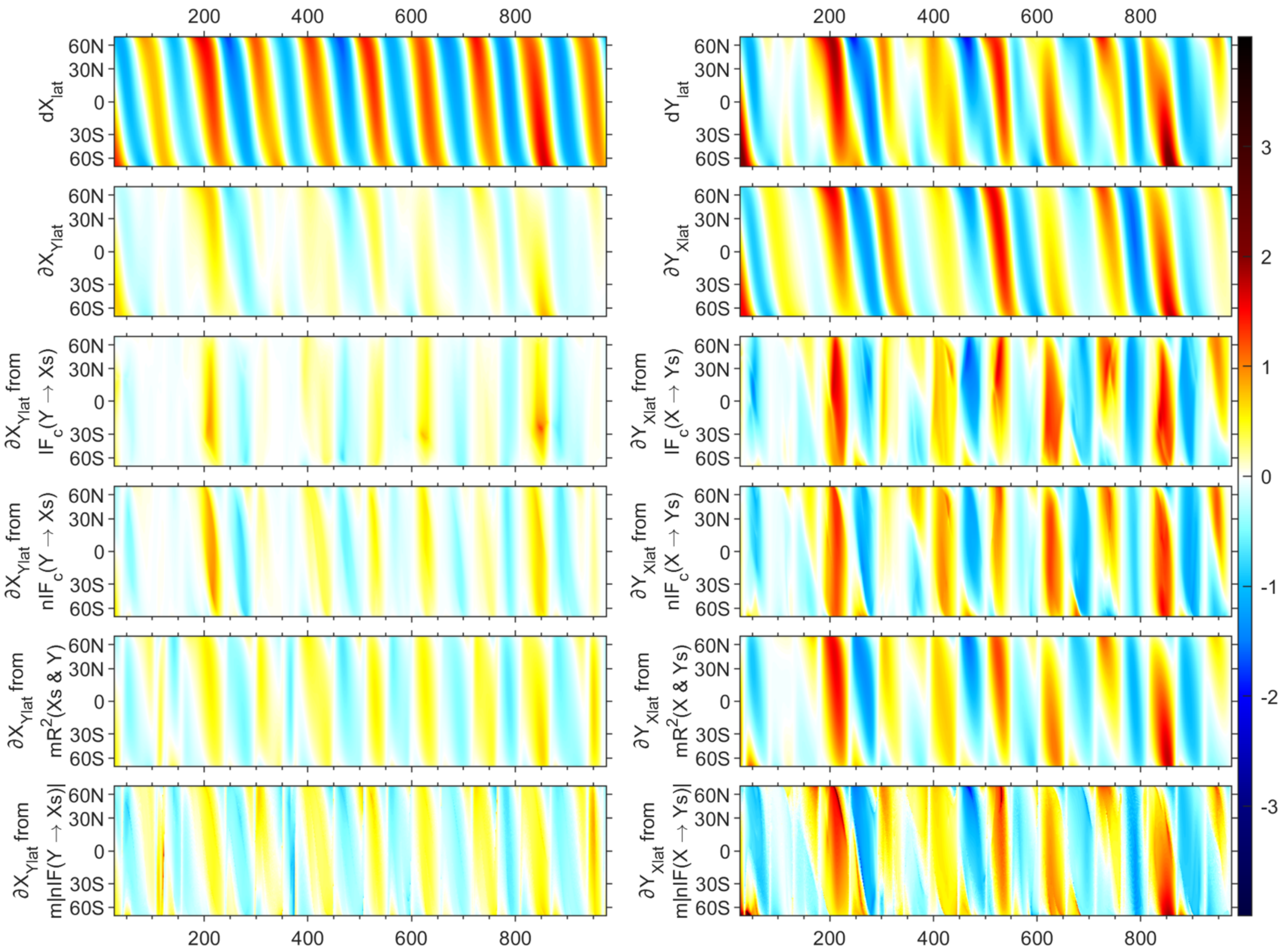


Figure S12. Assessment of methods for estimating interdependent contribution between designed variables *X* and *Y* (effect-maps), at low variability frequency (~1 *dX/dt* sign-change per 49 time-units) with no teleconnection. Estimates by mR^2 and $m|nIF|$ appear to give more consistent accuracy on both temporal and spatial dimensions when the signals are weak (dX_Y, i.e. when the noise is strong), but that advantage may be less obvious when signals are strong (dY_X).

85

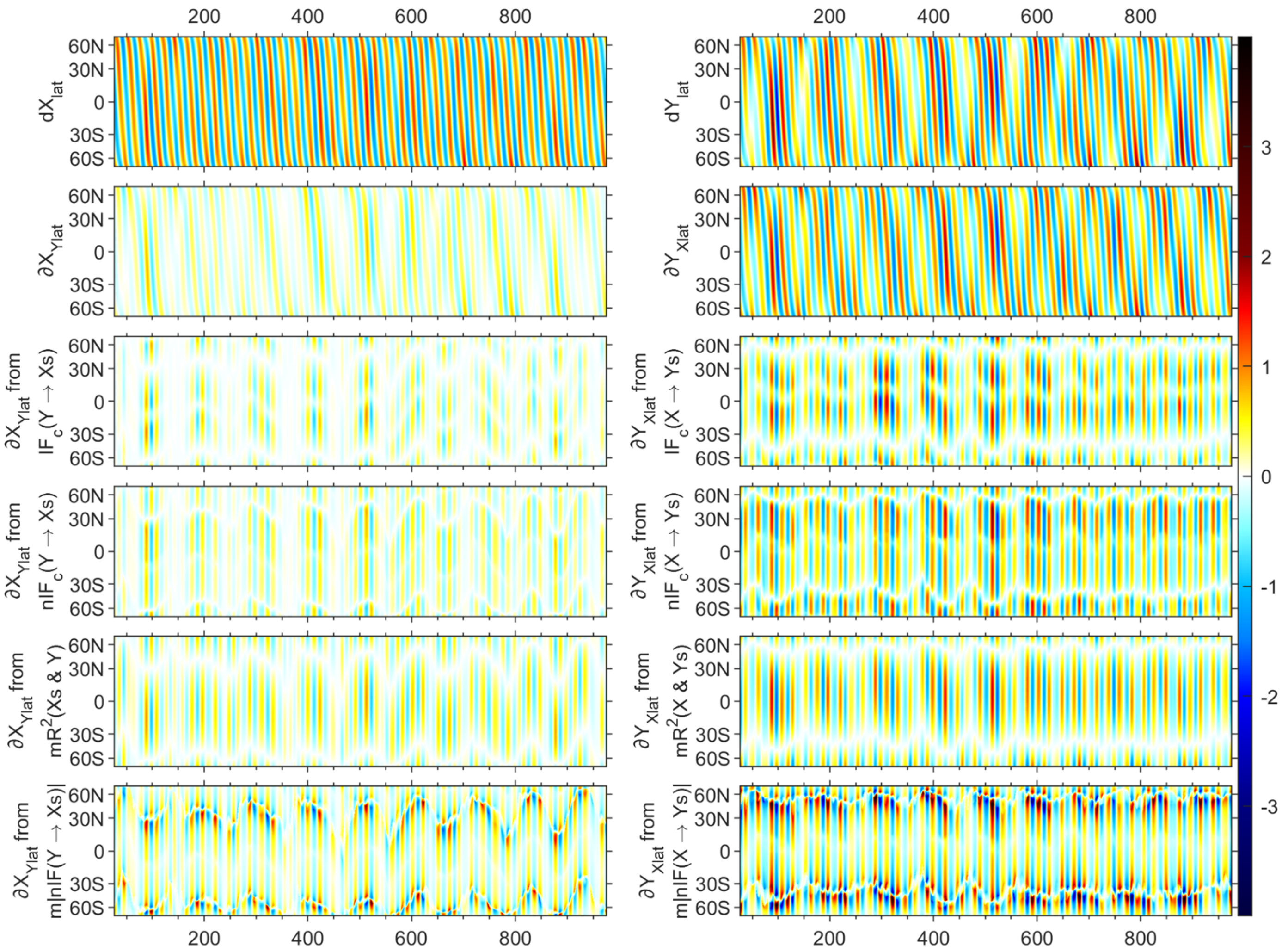


Figure S13. Assessment of methods for estimating interdependent contribution between designed variables *X* and *Y* (effect-maps), at high variability frequency (~4-5 *dX/dt* sign-change per 49 time-units) with no teleconnection. Estimates by mR^2 appear to give the best accuracy on both temporal and spatial dimensions. There is no significant merit for other methods.