Revisiting causality using stochastics: 1. Theory

Supplementary Information

Demetris Koutsoyiannis[1], Christian Onof[2], Antonis Christofidis[1] and Zbigniew W. Kundzewicz[3]

[1] Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens (dk@itia.ntua.gr)

[2] Department of Civil and Environmental Engineering, Faculty of Engineering, Imperial College London

[3] Meteorology Lab, Department of Construction and Geoengineering, Faculty of Environmental Engineering and Mechanical Engineering, Poznan University of Life Sciences, Poznań, Poland

Abstract This report contains Supplementary Information, namely, mathematical derivations, justifications and illustrations, for the paper series *Revisiting causality using stochastics* (Koutsoyiannis et al., 2022a,b) and in particular its first part, *Theory* (Koutsoyiannis et al., 2022a). It comprises three sections, namely *Relationship of continuous- and discrete-time impulse response functions* (SI1.1), *Justification of the linear causal system* (SI1.2), *Derivation of the relationships of autocorrelations and cross-correlations* (SI1.3) and *Temporal moments of the* impulse response function (SI1.4).

SI1.1 Relationship of continuous- and discrete-time impulse response functions

The basic equation for the causal system in continuous (natural) time is:

$$\underline{y}(t) = \int_{-\infty}^{\infty} g(h)\underline{x}(t-h)\mathrm{d}h + \underline{v}(t) \quad \text{(SI1.1)}$$

Changing variables, i.e., setting $u = t - h$, we can also write equation (SI1.1) as

$$\underline{y}(t) = \int_{-\infty}^{\infty} g(t-u)\underline{x}(u)\mathrm{d}u + \underline{v}(t) \quad \text{(SI1.2)}$$

Now we assume that we do not know the processes $\underline{x}(t), \underline{v}(t), \underline{y}(t)$ in continuous time but only at discrete times $t_\tau = \tau D$, where t is an integer. We can distinguish the following three cases.

Case 1. The instantaneous processes are known at time instants t_τ.

In this case we may assume that $\underline{x}(t)$ is formed as a series of instantaneous impulses at $t_i = iD$, each of magnitude $\underline{x}_i D\delta(t - iD)$, where $\underline{x}_i$ is the intensity of the process at time t_i and $\delta(\,)$ is the Dirac delta function. Hence from equation (SI1.2) we get

$$\underline{y}_\tau = \sum_{i=-\infty}^{\infty} \int_{-\infty}^{\infty} g(\tau D - u)\, \underline{x}_i D\delta(u - iD)\mathrm{d}u + \underline{v}_\tau \tag{SI1.3}$$

This yields:

$$\underline{y}_\tau = \sum_{i=-\infty}^{\infty} Dg(\tau D - iD)\, \underline{x}_i + \underline{v}_\tau \tag{SI1.4}$$

and finally, by changing variables again ($i = \tau - j$),

$$\underline{y}_\tau = \sum_{j=-\infty}^{\infty} g_j\, \underline{x}_{\tau-j} + \underline{v}_\tau, \qquad g_j = Dg(jD) \tag{SI1.5}$$

Case 2. The process $\underline{x}(t)$ is known as time averaged, while the other processes are known at times instants t_τ.

In this case we assume that the input $\underline{x}(t)$ in the period $((i-1)D, iD)$ is constant, equal to the time average

$$\underline{x}_i \coloneqq \frac{1}{D} \int_{(i-1)D}^{iD} \underline{x}(t)\mathrm{d}t \tag{SI1.6}$$

Thus, from equation (SI1.2) we get

$$\underline{y}_\tau = \sum_{i=-\infty}^{\infty} \int_{(i-1)D}^{iD} g(\tau D - u)\, \underline{x}_i \mathrm{d}u + \underline{v}_\tau = \sum_{i=-\infty}^{\infty} \underline{x}_i \int_{(i-1)D}^{iD} g(\tau D - u)\, \mathrm{d}u + \underline{v}_\tau \tag{SI1.7}$$

On the other hand, setting $w = \tau D - u$, we find

$$\int_{(i-1)D}^{iD} g(\tau D - u)\, \mathrm{d}u = \int_{(\tau-i)D}^{(\tau-i+1)D} g(w)\, \mathrm{d}w = g_{\mathrm{I}}\big((\tau - i + 1)D\big) - g_{\mathrm{I}}\big((\tau - i)D\big) \tag{SI1.8}$$

where

$$g_{\mathrm{I}}(a) \coloneqq \int_{-\infty}^{a} g(h)\mathrm{d}h \tag{SI1.9}$$

Hence,

$$\begin{aligned} \underline{y}_\tau &= \sum_{i=-\infty}^{\infty} \underline{x}_i \int_{(i-1)D}^{iD} g(\tau D - u)\, \mathrm{d}u + \underline{v}_\tau \\ &= \sum_{i=-\infty}^{\infty} \underline{x}_i \Big(g_{\mathrm{I}}\big((\tau - i + 1)D\big) - g_{\mathrm{I}}\big((\tau - i)D\big) \Big) + \underline{v}_\tau \end{aligned} \tag{SI1.10}$$

and finally, by changing variables ($i = \tau - j$),

$$\underline{y}_\tau = \sum_{j=-\infty}^{\infty} g_j \, \underline{x}_{\tau-j} + \underline{v}_\tau, \qquad g_j = g_j = g_\mathrm{I}\big((j+1)D\big) - g_\mathrm{I}(jD) \tag{SI1.11}$$

Case 3. All processes are known as time averaged in discrete time

In this case, which constitutes the most reasonable choice and is most frequently met in practice, we have

$$\underline{y}_\tau \coloneqq \frac{1}{D} \int_{(\tau-1)D}^{\tau D} \underline{y}(t) \mathrm{d}t \tag{SI1.12}$$

where, similarly to (SI1.7),

$$\begin{aligned} \underline{y}(t) &= \sum_{i=-\infty}^{\infty} \underline{x}_i \int_{(i-1)D}^{iD} g(y-u)\,\mathrm{d}u + \underline{v}_t = \sum_{i=-\infty}^{\infty} \underline{x}_i \int_{(i-1)D}^{iD} g(y-u)\,\mathrm{d}u + \underline{v}_t = \\ &= \sum_{i=-\infty}^{\infty} \underline{x}_i \big(g_\mathrm{I}(t-(i-1)D) - g_\mathrm{I}(t-iD)\big) + \underline{v}_\tau \end{aligned} \tag{SI1.13}$$

Hence,

$$\begin{aligned} \underline{y}_\tau &\coloneqq \frac{1}{D} \int_{(\tau-1)D}^{\tau D} \underline{y}(t) \mathrm{d}t + \frac{1}{D} \int_{(\tau-1)D}^{\tau D} \underline{v}(t) \mathrm{d}t \\ &= \frac{1}{D} \int_{(\tau-1)D}^{\tau D} \sum_{i=-\infty}^{\infty} \underline{x}_i \big(g_\mathrm{I}(t-(i-1)D) - g_\mathrm{I}(t-\tau)\big)\,\mathrm{d}t + \frac{1}{D} \int_{(\tau-1)D}^{\tau D} \underline{v}(t) \mathrm{d}t \end{aligned} \tag{SI1.14}$$

where the order of the integral and the sum can be interchanged. On the other hand,

$$\begin{aligned} \int_{(\tau-1)D}^{\tau D} \underline{x}_i g_\mathrm{I}(t-(i-1)D) \mathrm{d}t &= \underline{x}_i \int_{(\tau-i)D}^{(\tau-i+1)D} g_\mathrm{I}(w) \mathrm{d}w \\ &= \underline{x}_i \Big(G\big((\tau-i+1)D\big) - G\big((\tau-i)D\big) \Big) \end{aligned} \tag{SI1.15}$$

where

$$G(b) \coloneqq \int_{-\infty}^{b} g_\mathrm{I}(a) \mathrm{d}a = \int_{-\infty}^{b} \int_{-\infty}^{a} g(h) \mathrm{d}h\,\mathrm{d}a \;\Leftrightarrow\; g_\mathrm{I}(a) = G'(a) \tag{SI1.16}$$

After algebraic manipulations on (SI1.14), also considering (SI1.15), we find

$$\underline{y}_\tau = \sum_{j=-\infty}^{\infty} g_j \underline{x}_{\tau-j} + \underline{v}_\tau, \qquad g_j = \frac{1}{D}\Big(G\big((j-1)D\big) - 2G(jD) + G\big((j+1)D\big)\Big) \tag{SI1.17}$$

where now $\underline{y}_\tau$ and $\underline{v}_\tau$ are time averaged, rather than instantaneous process as were in the previous cases.

Interestingly, in cases 1, 2 and 3 the discrete-time impulse response function (IRF) ordinate (the coefficient g_j) is, respectively, the zeroth, first and second discrete-time derivative of the zeroth, first and second continuous-time anti-derivative (integral) of the continuous-time IRF (the function $g(h)$), multiplied by the time step D. Hence, because as $D \to 0$ the continuous and discrete derivatives become identical, in all three cases, for small D the approximation holds that

$$g_j \approx Dg(jD) \qquad \text{(SI1.18)}$$

On the other hand, this approximation is not acceptable if j is also small. In particular, for $j = 0$, if we assume a classic or potentially causal system, in which in continuous time $g(0) = 0$, then approximation (SI1.18) gives also $g_j = 0$, while equation (SI1.17) gives the better approximation

$$g_0 = \frac{1}{D}\big(G(-D) - 2G(0) + G(D)\big) \qquad \text{(SI1.19)}$$

In a (classic or potentially) causal system, only the last term is nonzero and hence

$$g_0 = \frac{G(D)}{D} \neq 0 \qquad \text{(SI1.20)}$$

SI1.2 Justification of the linear causal system

We provide two different lines of thought that justify the linear form of the causal system.

Justification 1

In our first version of justification, we start from discrete time and then generalize to continuous time. A causal relationship with the process $\underline{x}_\tau$ being the cause and $\underline{y}_\tau$ the effect should imply

$$y_\tau = f(x_\tau, x_{\tau-1}, \dots, x_{\tau-\eta}, \dots) \qquad \text{(SI1.21)}$$

while y_τ should not be functionally dependent on $x_{\tau+1}, x_{\tau+2}, \dots$. We stress that, since we are using discrete time and as explained through equation (SI1.20), there no reason to exclude x_τ from the functional form f in equation (SI1.21). At any point $(x_\tau, x_{\tau-1}, \dots, x_{\tau-\eta}, \dots) = (\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots)$ we can make a linear approximation of $f(x_\tau, x_{\tau-1}, \dots, x_{\tau-\eta}, \dots)$. Omitting the terms of order 2 or higher, we write:

$$y_\tau = f(\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots) + \sum_{\eta=0}^{\infty} \frac{\partial f}{\partial x_{\tau-\eta}} (\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots)\,(x_{\tau-\eta} - \xi_{\tau-\eta}) \qquad \text{(SI1.22)}$$

or

$$y_\tau \approx f(\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots) - \sum_{\eta=0}^{\infty} \frac{\partial f}{\partial x_{\tau-\eta}} (\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots)\, \xi_{\tau-\eta} + \sum_{\eta=0}^{\infty} \frac{\partial f}{\partial x_{\tau-\eta}} (\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots)\, x_{\tau-\eta} \tag{SI1.23}$$

We chose $(\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots)$ so that

$$f(\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots) - \sum_{\eta=0}^{\infty} \frac{\partial f}{\partial x_{\tau-\eta}} (\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots)\, \xi_{\tau-\eta} = 0 \tag{SI1.24}$$

If it happens that zero input gives zero output, i.e., $f(0,0, \dots, 0, \dots) = 0$, then the point sought is $(\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots) = (0,0, \dots, 0, \dots)$, as at this point, equation (SI1.24) is satisfied. Nonetheless, equation (SI1.24) is more general than this. For this chosen point, we have:

$$y_\tau \approx \sum_{\eta=0}^{\infty} \frac{\partial f}{\partial x_{\tau-\eta}} (\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots)\, x_{\tau-\eta} \tag{SI1.25}$$

Setting

$$g_\eta \coloneqq \frac{\partial f}{\partial x_{\tau-\eta}} (\xi_\tau, \xi_{\tau-1}, \dots, \xi_{\tau-\eta}, \dots) \tag{SI1.26}$$

we find

$$y_\tau \approx \sum_{\eta=0}^{\infty} g_\eta\, x_{\tau-\eta} \tag{SI1.27}$$

Now to move to the case of continuous (natural) time, we consider that the quantities in equation (SI1.27) are time averages of the continuous-time processes. Hence we get

$$\frac{1}{D} \int_{(\tau-1)D}^{\tau D} y(w)\mathrm{d}w \approx \sum_{\eta=0}^{\infty} g_\eta \frac{1}{D} \int_{(\tau-\eta-1)D}^{(\tau-\eta)D} x(w)\mathrm{d}w \tag{SI1.28}$$

Using approximation (SI1.18), this gives

$$\frac{1}{D} \int_{t-D}^{t} y(w)\mathrm{d}w \approx \sum_{\eta=0}^{\infty} \int_{t-\eta D-D}^{t-\eta D} g(\eta D) x(w)\mathrm{d}w = \int_{t-D}^{t} g(0) x(w)\mathrm{d}w + \int_{t-2D}^{t-D} g(D) x(w)\mathrm{d}w + \cdots \tag{SI1.29}$$

or

$$\frac{1}{D}\int_{t-D}^{t} y(w)\mathrm{d}w \approx \int_{t-D}^{t} g(t-w)x(w)\mathrm{d}w + \int_{t-2D}^{t-D} g(t-w)x(w)\mathrm{d}w + \cdots$$
$$= \int_{-\infty}^{t} g(t-w)x(w)\mathrm{d}w = -\int_{\infty}^{0} g(h)x(t-h)\mathrm{d}h \tag{SI1.30}$$

That is,

$$\frac{1}{D}\int_{t-D}^{t} y(w)\mathrm{d}w \approx \int_{0}^{\infty} g(h)x(t-h)\mathrm{d}h \tag{SI1.31}$$

Taking the limit as $D \to 0$

$$y(t) = \int_{0}^{\infty} g(h)x(t-h)\mathrm{d}h \tag{SI1.32}$$

By adding an error term $v(t)$, necessary to recover the linearization error (even if there is no actual error term in (SI1.21)), as well as the effect of processes other than $x(t)$ that may influence $y(t)$, and by substituting the processes $\underline{y}(t), \underline{x}(t), \underline{v}(t)$ for their realizations, we obtain equation (SI1.1).

Justification 2

Here we work in continuous time from the beginning and, as a starting point, we assume that a causal link is characterized by *rule-governedness* (this is the regularity discussed by Hume, and the lawfulness of Kant's analysis) and *irreversibility* (as in Kant's analysis). The rule-governedness claim entails that there is some law that links the change in magnitude of a causing factor x at time $t-h$ $(h \geq 0)$ to a change a causally impacted quantity y at time t. This would take the general form, for any t:

$$\delta y_t = f_h(\delta x_{t-h})\,\delta h \tag{SI1.33}$$

where f_h is a function not necessarily expressible in closed form. The index h is positive to reflect the claim of irreversibility according to which the effect must be completed after the cause has started acting. The inclusion of the term δh reflects the fact that a cause is never instantaneously effective. For instance, when Nadal (the famous tennis player) hits a forehand, the change of direction of the ball happens over a small time interval δh over which the energy from his arm is passed on to the ball. It is assumed that, if we were to freeze his motion after a fraction of time $\delta h/k$ $(k > 1)$ then a proportion $1/k$ of the energy would have been communicated to the ball so that the causally impacted quantity, y, would only have changed by a fraction $1/k$ of how it would have changed had the motion been allowed to last the full δh. The linear approximation will of course not work for any x, y, but (for continuous non-quantum phenomena) it is plausible that for some transformation of either x or y (here we shall want to keep y fixed), over a short duration of time δh, the effect is (to a sufficiently good approximation) reduced in proportion to the reduction of δh.

What can we say about f_h? If we isolate, for now, the causal link between x and y, the inactivity of the cause will imply that the effect does not come about. So

$$f_h(0) = 0 \tag{SI1.34}$$

Further, if δx_{t-h} is small, by Taylor expansion assuming that f_h is continuously differentiable, we find

$$\delta y_t = \delta x_{t-h} \frac{\mathrm{d}f_h}{\mathrm{d}x}(0)\, \delta h + O(\delta x_{t-h})\, \delta h \tag{SI1.35}$$

If we define the function g as:

$$g(h) = \frac{\mathrm{d}f_h}{\mathrm{d}x}(0) \tag{SI1.36}$$

then (SI1.35) can be written:

$$\delta y_t = \delta x_{t-h}\, g(h)\delta h + O(\delta x_{t-h})\, \delta h \tag{SI1.37}$$

This relationship holds in the neighbourhood of 0, and depending on how close to linearity f is, that neighbourhood will be more or less large. A first step in the proposal consists therefore in representing this uncertainty stochastically, so that (SI1.37) can be integrated to yield the following relation:

$$\underline{y_t} = x_{t-h}\, g(h)\delta h + \underline{w(h)}_t\, \delta h \tag{SI1.38}$$

where $\underline{w(h)}_t$ is a random variable containing the constant of integration and the uncertainty arising from the possible non-linearity of f as we move away from 0. $\underline{y_t}$ in the left-hand side of the equation must therefore now also be considered as a random variable.

While equation (SI1.38) considers the case of a single cause at time $t-h$ of the change δy_t, it can easily be generalized to include other causes. These are of two types: first, there can be further changes of x at other times prior to t which will causally impact y at time t. Consequently, we shall want to take into account all the causal impacts of x prior to t by integrating over the time h representing the delay between cause and effect. Since we are now considering processes in continuous time, the quantities x and y are now represented as functions of time. The w terms are also integrated, which yields a new random variable $\underline{v}$. This yields:

$$\underline{y}(t) = \int_0^{\infty} x(t-h)g(h)\mathrm{d}h + \underline{v}(t) \tag{SI1.39}$$

The second type of additional cause would be a cause that is not x, for example, z. This would require adding another integral term to equation (SI1.38), and many more can be added in this way:

$$\underline{y}(t) = \int_0^\infty x(t-h)g(h)\mathrm{d}h + \int_0^\infty z(t-h)r(h)\mathrm{d}h + \underline{v}(t) \tag{SI1.40}$$

In fact, the thesis of determinism would amount to claiming that such an equation can be written without the additional random variable $\underline{v}$ if all the N operative causes are accounted for:

$$y(t) = \sum_{i=1}^{N} \int_0^\infty x_i(t-h)g_i(h)\mathrm{d}h \tag{SI1.41}$$

for some functions $g_i, i = 1, \dots, N$, and where y need no longer be a random variable.

In practice, such a deterministic thesis cannot be validated. Moreover, most of the time, the interest focusses upon one particular causing process, so equation (SI1.38) is what will be most useful. Further, for any given x and y, there are two possible directions of causality, both should be considered as random processes so the existence of an equation (SI1.39) can be tested in both directions. This means that the useful equation that defines our proposal for the key necessary condition for causality is:

$$\underline{y}(t) = \int_0^\infty \underline{x}(t-h)g(h)\mathrm{d}h + \underline{v}(t) \tag{SI1.42}$$

Random variable $\underline{v}(t)$ therefore represents both the departure from linearity of the causal link x and y, and the uncertainty around the role of other causal factors than x. There is an important assumption flagged above in this proposal about the approximate linearity of f. The more we move away from that assumption, the larger the variance of the process $\underline{v}(t)$. To reflect that assumption, it is important to include the condition that this variance does not become too large when compared with the variance of the process to be explained, i.e. $\underline{y}$. Should this not be achievable, the use of a non-linear transform of $\underline{x}$ should be considered: in principle, one can only conclude to an absence of causality if (SI1.42) does not hold for $\underline{x}$ or any non-linear transform thereof.

Further, since g is a function that should have physical meaning, some condition of smoothness should be included.

SI1.3 Derivation of the relationships of autocorrelations and cross-correlations

In order to determine $c_{yx}(h) \coloneqq \operatorname{cov}\left[\underline{y}(t+h), \underline{x}(t)\right]$, first we assume, without loss of generality, that all processes have zero mean and we take into account the fact that $\underline{v}(t)$

is uncorrelated with $\underline{x}(t)$. Writing equation (SI1.1) for $\underline{y}(t+h)$, multiplying it with $\underline{x}(t)$, and taking expected values we have

$$\mathrm{E}[\underline{y}(t+h)\underline{x}(t)] = \int_{-\infty}^{\infty} g(a)\mathrm{E}\big[\underline{x}(t+h-a)\underline{x}(t)\big]\mathrm{d}a \tag{SI1.43}$$

or, since the means are zero,

$$c_{yx}(h) = \int_{-\infty}^{\infty} g(a)c_{xx}(h-a)\mathrm{d}a \tag{SI1.44}$$

To determine $c_{yy}(h) \coloneqq \mathrm{cov}\left[\underline{y}(t+h), \underline{y}(t)\right]$ we multiply $\underline{y}(t+h)$ with $\underline{y}(t)$ and obtain

$$\begin{aligned}\underline{y}(t+h)\,\underline{y}(t) = &\left(\int_{-\infty}^{\infty} g(a)\underline{x}(t+h-a)\mathrm{d}a + \underline{v}(t+h)\right) \\ &\times\left(\int_{-\infty}^{\infty} g(b)\underline{x}(t-b)\mathrm{d}b + \underline{v}(t)\right)\end{aligned} \tag{SI1.45}$$

Taking expected values, we find

$$\begin{aligned}&\mathrm{E}\left[\underline{y}(t)\underline{y}(t+h)\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(a)g(b)\mathrm{E}\big[\underline{x}(t+h-a)\underline{x}(t-b)\big]\,\mathrm{d}a\,\mathrm{d}b + \mathrm{E}\big[\underline{v}(t)\underline{v}(t+h)\big] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(a)g(b)c_{xx}(h-a+b)\,\mathrm{d}a\,\mathrm{d}b + c_{vv}(h)\end{aligned} \tag{SI1.46}$$

On the other hand, because of equation (SI1.44) we have

$$\int_{-\infty}^{\infty} g(a)c_{xx}(h+b-a)\mathrm{d}a = c_{yx}(h+b) \tag{SI1.47}$$

Hence

$$c_{yy}(h) = \int_{-\infty}^{\infty} g(b)c_{yx}(h+b)\,\mathrm{d}b + c_{vv}(h) \tag{SI1.48}$$

It is relevant to note that a measure of the magnitude of the auto- and cross-covariance functions is provided by the p-norm

$$\|c\|_p \coloneqq \left(\int_{-\infty}^{\infty} |c(h)|^p\mathrm{d}h\right)^{1/p} \tag{SI1.49}$$

where p is typically taken 1, but in case of long-range dependence (LRD), $\|c\|_1$ is infinite and so we should choose $p \geq 1/(2-2H)$, where H is the Hurst parameter, in order for the norm to be finite. As the equations (SI1.44) and (SI1.48) denote convolution of functions, using Young's convolution inequality and assuming that $\|g\| \equiv \|g\|_1$ is finite, from equation (SI1.44) we obtain

$$\|c_{yx}\|_p \leq \|g\| \, \|c_{xx}\|_p \tag{SI1.50}$$

and from equation (SI1.48)

$$\|c_{yy}\|_p - \|c_{vv}\|_p \leq \|c_{yy} - c_{vv}\|_p \leq \|g\| \, \|c_{yx}\|_p \leq \|g\|^2 \|c_{xx}\|_p \tag{SI1.51}$$

where in the leftmost part we have also used the reverse triangle inequality, along with the inequality $\|c_{yy}\|_p > \|c_{vv}\|_p$.

The last two equations can easily be transformed to norms of auto- and cross-correlation functions (e.g. $r_{yx} \coloneqq c_{yx}/(\sigma_x \sigma_y)$, where σ denotes standard deviation), obtaining

$$\|r_{yx}\|_p \leq \frac{\sigma_x}{\sigma_y} \|g\| \, \|r_{xx}\|_p \tag{SI1.52}$$

and

$$\|r_{yy}\|_p - \frac{\sigma_v^2}{\sigma_y^2} \|r_{vv}\|_p \leq \left\| r_{yy} - \frac{\sigma_v^2}{\sigma_y^2} r_{vv} \right\|_p \leq \frac{\sigma_x}{\sigma_y} \|g\| \, \|r_{yx}\|_p \leq \frac{\sigma_x^2}{\sigma_y^2} \|g\|^2 \|r_{xx}\|_p \tag{SI1.53}$$

However, these inequalities are too involved to provide useful information about the relative magnitude of $\|r_{xx}\|$, $\|r_{yy}\|$ and $\|r_{yx}\|$. Nonetheless, they allow us to discuss the rather common intuitive consideration that the autocorrelation of the effect $\underline{y}$ is higher than that of the cause $\underline{x}$. This is regarded to be particularly the case if $\underline{x}$ is white noise, whose autocorrelation is zero for any lag except 0. In fact, though, the autocovariance of the white noise is $c_{xx}(h) = \sigma_x^2 \delta(h)$ and its norm is not zero but $\|c_{xx}\|_1 = \sigma_x^2$, which means that $\|r_{xx}\|_1 = 1$. Hence, even assuming zero variance of $\underline{v}$, the above inequalities yield

$$\|r_{yx}\|_1 \leq \frac{\sigma_x}{\sigma_y} \|g\|, \qquad \|r_{yy}\|_1 \leq \frac{\sigma_x^2}{\sigma_y^2} \|g\|^2 \tag{SI1.54}$$

The order direction "≤" in these inequalities does not suggest an increased autocorrelation of $\underline{y}$. Rather it is more accurate to say that here we have a more diffuse shape of the autocorrelation function of $\underline{y}$ in comparison to that of $\underline{x}$.

SI1.4 Temporal moments of the impulse response function

In the main paper (Koutsoyiannis et al., 2022a) we have defined several temporal indices of the IRF and connected them to characteristics of the the processes $\underline{x}(t)$ and $\underline{y}(t)$ with equation (17). Furthermore, it is useful to define the temporal moments as:

$$H_n(T) \coloneqq \frac{1}{T}\int_0^T g(t)t^n \mathrm{d}t\,, \qquad \underline{X}_n(T) \coloneqq \frac{1}{T}\int_0^T \underline{x}(t)t^n \mathrm{d}t \tag{SI1.55}$$

and likewise for $\underline{Y}_n(T)$ and $\underline{V}_n(T)$. If T is a large time period, i.e., $T \gg J$ where J is such that $g(h) \approx 0$ for $|h| > J$, it can be shown that

$$\underline{Y}_n = \sum_{i=0}^{n} H_{n-i}\, \underline{X}_i + \underline{V}_n \tag{SI1.56}$$

where we have omitted the reference to T for notational simplification. Nb., Nash (1959) found similar relationships for the unit hydrograph.

For $n = 0$, the respective quantities are the temporal averages of the three processes, and in this case equation (SI1.56) takes the form

$$\underline{Y}_0 = H_0 \underline{X}_0 + \underline{V}_0 \tag{SI1.57}$$

which is similar to equation (17) of the main paper (Koutsoyiannis et al., 2022a). For $n = 1$ the temporal moments represent the centroids of the geometric shapes of the time series and in this case equation (SI1.56) takes the form

$$\underline{Y}_1 = H_1 \underline{X}_0 + H_0 \underline{X}_1 + \underline{V}_1 \tag{SI1.58}$$

More generally, equation (SI1.56) can be used in a recursive manner to find the temporal moments of the IRF based on those of the processes $\underline{x}(t)$ and $\underline{y}(t)$.

Acknowledgments: The constructive comments of two anonymous reviewers helped us to substantially improve our work.

Funding: This research received no external funding but was motivated by the scientific curiosity of the authors.

Conflicts of Interest: We declare no conflict of interest.

References

Koutsoyiannis, D., Onof, C. Christofidis, A., & Kundzewicz, Z. W. 2022a. Revisiting causality using stochastics: 1. Theory, *Proceedings of the Royal Society A*, in review.

Koutsoyiannis, D., Onof, C. Christofidis, A., & Kundzewicz, Z. W. 2022b. Revisiting causality using stochastics: 2. Applications, *Proceedings of the Royal Society A*, in review.

Nash, J.E. 1959 Systematic determination of unit hydrograph parameters. *J. Geoph. Res.* **64**(1), 111-115.