Coping with Dating Errors in Causality Estimation

D.A. Smirnov[1,2], N. Marwan[3], S.F.M. Breitenbach[4], F. Lechleitner[5,6], and J. Kurths[3,2]

[1] *Saratov Branch of V.A. Kotel'nikov Institute of RadioEngineering and Electronics of the Russian Academy of Sciences – 38 Zelyonaya St., Saratov 410019, Russia*
[2] *Institute of Applied Physics of the Russian Academy of Sciences, 46 Ulyanova St., Nizhny Novgorod 603950, Russia*
[3] *Potsdam Institute for Climate Impact Research, Telegraphenberg A31, Potsdam 14473, Germany*
[4] *Institute for Geology, Mineralogy & Geophysics, Ruhr-Universität Bochum, Universitätsstr. 150, 44801, Bochum, Germany*
[5] *Department of Earth Sciences, ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland*
[6] *Department of Earth Sciences, Durham University, Durham, DH1 3LE, UK*

Abstract –We consider the problem of estimating causal influences between observed processes from time series possibly corrupted by errors in the time variable (dating errors) which are typical in palaeoclimatology, planetary science and astrophysics. "Causality ratio" based on the Wiener – Granger causality is proposed and studied for a paradigmatic class of model systems to reveal conditions under which it correctly indicates directionality of unidirectional coupling. It is argued that in case of *a priori* known directionality, the causality ratio allows a characterization of dating errors and observational noise. Finally, we apply the developed approach to palaeoclimatic data and quantify the influence of solar activity on tropical Atlantic climate dynamics over the last two millennia. A stronger solar influence in the first millennium A.D. is inferred. The results also suggest a dating error of about 20 years in the solar proxy time series over the same period.

Introduction. – Revealing cause-and-effect relationships between observed processes at various time scales is an important step in understanding many physical, biological, physiological and geophysical systems [1–8]. Frequently, this issue must be addressed with rather limited knowledge about the systems under study, amounts of observational data, and dating accuracy. A general approach to detect and quantify causal couplings, i.e., to find out "who drives whom", is the Wiener – Granger (WG) causality [9,10]. In its simplest version, the idea is to check whether a present value of one process (X) can be predicted more accurately using the past of a second process (Y) in comparison with predictions based solely on the past of X. In fact, this concept generalizes a conditional (partial) cross-correlation [11] and has been followed by a number of elaborations such as information-theoretic measures [3,12–15] and various nonlinear approximations [16]. Despite some limitations and obstacles [17–20], the WG causality appears quite useful in practice, allowing meaningful dynamical interpretations [21,22] and becoming increasingly widely used in different fields, such as biomedicine [1,5,8] and geophysics [6].

Causal coupling estimation is also of great value in cli-mate science, where temporal changes of climatically sensitive proxies [23] are the main source of information about past climate dynamics over long time intervals. The stalagmite YOK-I from the Yok Balum Cave in Southern Belize is especially well dated [24] and provides a high-resolution reconstruction of low-latitudinal Atlantic moisture variations [25]. Making use of solar irradiance reconstructions (e.g. [26]), one can ask "How do variations in solar activity affect regional Atlantic climate?". Answering this question helps further delineating the time-variant processes that drive climate variations. However, this question leads directly to the main difficulty with such data: dating accuracy of the reconstructions used. Uncertainties inherent to sampling and dating methods limit our knowledge of the time instant of each proxy observation, so that temporal ordering of the observations from the two time series may be distorted uniformly or irregularly in the course of time. This makes questionable any application of the WG causality approach, which essentially requires a clear distinction between the future and the past.

In this Letter, we propose a solution with an appropriate specification of the problem setting and adaptation of the WG causality characteristics. We consider a situation

where it is known in advance that the coupling between two processes underlying the observed time series is unidirectional, and the problem reduces to identifying the coupling directionality. Observational noise and dating errors may strongly affect the results of any coupling analysis. In particular, the usual cross-correlation function (CCF) is obviously insufficient since even a uniform dating error moves the location of the CCF maximum along the time axis, so that "lead – lag" information is lost. We note, however, that the WG causality approach provides two coupling characteristics corresponding to the two directions $X \to Y$ and $Y \to X$, which is a richer characterization than a single CCF value. To make the WG causality work in case of dating errors, we suggest its modification involving the definition of the *causality ratio* $r_{Y \to X}$ which is the ratio of maximized time-lagged truncated WG causalities in the directions $Y \to X$ and $X \to Y$. We argue that if a coupling indeed exists in the direction $Y \to X$, then under certain conditions $r_{Y \to X} > 1$, i.e., the causality ratio is an indicator of the coupling directionality.

We study the conditions under which this causality ratio allows us to extract information on directionality of unidirectional coupling or, knowing the directionality, to characterize dating errors and observational noise in the analyzed time series. As for the latter task, the mentioned palaeoclimate problem is a relevant example where coupling is unidirectional from solar activity variations to regional climate (reflected in proxy reconstructions), while dating errors and observational noise in the proxy signals remain largely unknown. Here, we (i) determine the causality ratio for a class of model systems exactly, (ii) analyze statistical properties of its estimator in numerical simulations, and (iii) apply the approach to palaeoclimate data using the two records mentioned above to assess their dating accuracy and quantify the time-variant influence of solar activity on the tropical Atlantic climate. Further details of the method and additional results are given in [27].

Wiener – Granger causality. – Let $(X(t), Y(t))$ be a bivariate random process with realizations $(x(t), y(t))$. Denote $x_n = x(t_n)$, $y_n = y(t_n)$, where $t_n = nh$, $n \in \mathbf{Z}$, and h is the sampling interval. Consider the self-predictor $x_n^{ind} = E[X(t_n)|x_{n-1}, x_{n-2}, \ldots]$ where the expectation $E[\cdot|\cdot]$ is conditioned on the infinite past $\{x_{n-1}, x_{n-2}, \ldots\}$. Its mean-squared error is $\sigma_{X,ind}^2 = E[(X(t_n) - x_n^{ind})^2]$ where the expectation is taken over all x_n and all $\{x_{n-1}, x_{n-2}, \ldots\}$. This error is the least over all self-predictors for X. The joint predictor $x_n^{joint} = E[X(t_n)|x_{n-1}, y_{n-1}, x_{n-2}, y_{n-2}, \ldots]$ gives the least error $\sigma_{X,joint}^2$ over all joint predictors. The prediction improvement (PI) $G_{Y \to X} = (\sigma_{X,ind}^2 - \sigma_{X,joint}^2)/\sigma_{X,ind}^2$ is a measure of WG causality in the direction $Y \to X$. Everything is analogous for the direction $X \to Y$.

The WG idea was first realized for stationary Gaussian processes [10]. Then, when estimating $G_{Y \to X}$ from a finite time series $\{x_n, y_n\}_{n=1}^N$, one truncates the (conditioning) infinite pasts at finite numbers of terms l_X and l_{XY} and

fits univariate and bivariate linear autoregressive models of the orders l_X and (l_X, l_{XY}) to the data via the ordinary least-squares technique. In other words, one uses the predictors $x_{n,l_X}^{ind} = E[X_n|x_{n-1}, x_{n-2}, \ldots, x_{n-l_X}]$ and $x_{n,l_X,l_{XY}}^{joint} = E[X_n|x_{n-1}, \ldots, x_{n-l_X}, y_{n-1}, \ldots, y_{n-l_{XY}}]$ and gets the truncated WG causality $G_{Y \to X}^{tr}$. The latter is often a good approximation of $G_{Y \to X}$ even at small l_X and l_{XY}. The model orders can be selected via the Schwarz criterion [28] and statistical significance can be checked via Fisher's F-test [29].

Causality ratio. – Consider a more general setting with the original processes X_0 and Y_0, whose observed versions X and Y are distorted along two lines. First, due to an amplitude noise: $X(t) = X_0(t) + \Xi(t)$ and $Y(t) = Y_0(t) + \Psi(t)$ where $\Xi(t)$ and $\Psi(t)$ are independent observational noises with variances σ_Ξ^2 and σ_Ψ^2, whose discrete time realizations ξ_n and ψ_n are white noises. Second, due to time uncertainty: genuine (*a priori* unknown) observation instants t_n^X and t_n^Y deviate from the supposed regular equidistant series $t_n = nh$: $x_n = x(t_n^X) + \xi_n$ and $y_n = y(t_n^Y) + \psi_n$ with $t_n^X + \delta_n^X = nh$ and $t_n^Y + \delta_n^Y = nh$, where δ_n^X and δ_n^Y stay for the time axis (i.e. dating) errors. The latter may be rapidly fluctuating or slowly varying and may be defined either as random processes or deterministic functions of time. To account for the dating errors and retain sensitivity to coupling, we use the time-lagged WG causality: namely, $G_{Y \to X}^{tr}(\Delta)$ is defined as prediction improvement of x_n when using the segment $\{y_{n-\Delta/h}, \ldots, y_{n-(l_{XY}-1)-\Delta/h}\}$. Then, we suggest to determine its maximum over an interval of positive and negative time lags of some width $2\Delta_m$: $G_{Y \to X}^{tr,max} = \max_{-\Delta_m \le \Delta \le \Delta_m} G_{Y \to X}^{tr}(\Delta)$. Analogously we define $G_{X \to Y}^{tr,max}$. Finally, the causality ratio in the direction $Y \to X$ reads

$$r_{Y \to X} = \frac{G_{Y \to X}^{tr,max}}{G_{X \to Y}^{tr,max}}. \tag{1}$$

Obviously, $r_{X \to Y} = 1/r_{Y \to X}$. The value of Δ_m should be chosen so as to exceed a maximal possible dating error to avoid missing the maximal PIs. If, moreover, the coupling is time-delayed, locations of the PIs maxima are shifted along the Δ-axis by the value of this delay. Hence, if one expects a time delay, then the value of Δ_m should be selected so as to exceed the sum of the absolute values of the coupling delay and the dating error.

We conjecture that for unidirectional coupling $Y \to X$ and similar individual characteristics of the processes X and Y, the ratio $r_{Y \to X}$ is considerably greater than unity. However, dating errors and observational noise along with estimates fluctuations due to shortness of time series may somewhat decrease $r_{Y \to X}$, which is studied below.

Model system. – Since the value of $r_{Y \to X}$ may depend on many features of the processes under study (such as characteristic times and sampling interval) and parameters of the estimation technique (such as l_X), we need to

choose a reasonably simple system and a narrow range of the parameters for which the causality ratio can be studied in detail. As such a testing system, we use coupled "relaxators" (first-order decay processes):

$$
\begin{aligned}
dX_0/dt &= -\alpha X_0(t) + kY_0(t) + \zeta_X(t), \\
dY_0/dt &= -\alpha Y_0(t) + \zeta_Y(t),
\end{aligned}
\tag{2}
$$

where α determines the characteristic relaxation time $\tau = 1/\alpha$, k is the coupling coefficient, and ζ_X and ζ_Y are independent zero-mean white noises with autocorrelation functions $E[\zeta_X(t_1)\zeta_X(t_2)] = E[\zeta_Y(t_1)\zeta_Y(t_2)] = \delta(t_1 - t_2)$ where δ is Dirac's delta. Eqs. (2) represent a simple, but basic class of systems which still exhibit irregular temporal behavior and are often encountered in different fields (e.g. [30]). The squared zero-lag CCF reads here $C^2_{X_0Y_0,0} = (\beta/4)/(1 + \beta/2)$ where $\beta = k^2/\alpha^2$ is a non-dimensional coupling strength. $C^2_{X_0Y_0,0}$ ranges from 0 (for $k = 0$) to 0.5 (for $k \to \infty$) and can be used to parameterize the coupling strength as well. The sampling rate can be conveniently characterised by the ratio h/τ.

For system (2) it appears possible to confine ourselves with the orders $l_X = l_{XY} = l_Y = l_{YX} = 1$. It can be argued that $G^{tr}_{Y \to X}(\Delta)$ obtained at $l_X = l_{XY} = 1$ is close to $G^{tr}_{Y \to X}(\Delta)$ obtained at $l_X = \infty$ and $l_{XY} = 1$, if the sampling interval h is not too small (e.g. $\geq 0.2\tau$) [21]. In numerical simulations here, we also find that the results for $G^{tr}_{Y \to X}(\Delta)$ with $l_X = 1$ are close to those obtained with l_X selected via the Schwarz criterion (difference of the order of 1%). Similar arguments hold for l_{XY}. Then, the quantity $G^{tr}_{Y \to X}(\Delta)$ can be expressed via the autocorrelation function (ACF) $C_{XX}(h)$ and the CCF $C_{XY}(\Delta)$ and $C_{XY}(\Delta - h)$ [20, 27]. Having found ACFs and CCF analytically, we compute the time-lagged truncated WG causalities versus Δ and select their maxima to calculate the causality ratio. Such a precise analysis is performed for various coupling coefficient values, sampling intervals, observational noise and dating error levels, while statistical properties of the causality ratio estimator are investigated in numerical simulations. We check if indeed $r_{Y \to X} > 1$ and assess how small $r_{Y \to X}$ can be at all. A closer attention is paid to cases with $0.1 \leq C^2_{XY,max} \leq 0.2$ and WG causalities $0.01 \leq G^{tr,max}_{Y \to X} \leq 0.03$ which are reminiscent of those often observed in climate data analysis in cases of statistically significant coupling detection (e.g. [31] and the palaeoclimate example below).

Exact study of possible causality ratio values. – Before considering the central point of dating errors identification, it is necessary to study the case of undistorted observations $X = X_0$ and $Y = Y_0$. For the most practically interesting situations of not too sparse sampling (e.g. $h \leq 0.2\tau$), $r_{Y \to X}$ is well above unity, confidently indicating the correct coupling direction. Namely, $r_{Y \to X} = 1.6$ for $h = 0.2\tau$ and a moderately strong coupling of $C^2_{X_0Y_0,0} = 0.1$. For rather sparse samplings of $h \geq \tau$, the ratio $r_{Y \to X}$ gets close to unity and, hence, cannot reliably reveal coupling directionality. This is similar for any coupling strength: in particular, at $h/\tau = 0.2$ the causality ratio remains almost constant ($r_{Y \to X} \approx 1.6$) in the wide range of $0 < C^2_{XY,0} < 0.3$. For stronger couplings, $r_{Y \to X}$ becomes even greater, up to ≈ 3 at $C^2_{XY,0} = 0.5$. Thus, if the sampling is not too sparse, $r_{Y \to X}$ correctly detects coupling directionality. More details are given in [27].

Though there can be different types of dating errors, their basic effect can be studied on a simple example where dating errors equal a constant temporal shift half the time (e.g. for an older half of a palaeoclimate record where accurate dating is more difficult) and zero otherwise. Regardless which signal is erroneously dated, only the relative dating errors matter in causality estimation. For definiteness, we introduce the dating errors only into the driving signal: $\delta^Y_n = const = \delta^Y$ half time (for $n = 1, \ldots, N/2$) and $\delta^Y_n = 0$ otherwise (for $n = N/2 + 1, \ldots, N$). The "average CCF" of such a nonstationary process (X, Y) can be defined as the expectation of the sample CCF computed over the entire time span and equals an arithmetic mean of the CCFs for the two stationary halves. The usual WG causalities defined for the entire time span are expressed via such an average CCF in the same way as before. Figs. 1,a,b show that the shape of the plots for the time-lagged WG causalities and locations of their maxima change strongly when the dating error becomes comparable with the relaxation time τ. Then, the "correct" $G^{tr,max}_{Y \to X}$ decreases almost two times as compared to zero dating error, while the opposite $G^{tr,max}_{X \to Y}$ decreases only 1.5 times. At that, the causality ratio becomes close to unity and may even fall down to 0.9 for the dating error greater than τ. If a smaller or a larger portion of a time series suffers from a uniform dating error, then the effect of the latter on the causality ratio and the respective distortions of the plots $G^{tr}_{Y \to X}(\Delta)$ are weaker [27], in particular, they vanish if the entire time series is characterized with a uniform dating error since the causality ratio involves maximization over temporal shifts.

Principally, dating errors may be distributed in a complicated manner determined both by random walk-like stochastic contribution, analytical limitations and global contribution induced by incorrect tie points as, e.g., erroneous attribution of volcanic eruption dates due to incorrect identification of individual eruptions [32]. Still, we have obtained results very similar to Fig. 1 for dating errors linearly increasing with age, even with a superimposed random-walk component whose values become of the order of τ for ages of the order of 100τ as motivated by palaeoclimate applications. Thus, the described effect of the dating errors is robust, being observed just for reasonably large dating errors without any other, specific conditions.

When dating errors are present, it is natural to expect also an observational noise. Let us first show how the latter affects the causality ratio for zero dating errors. It appears that the noise Ψ in the driving signal can significantly decrease $r_{Y \to X}$. Thus, at moderate $h/\tau = 0.2$,
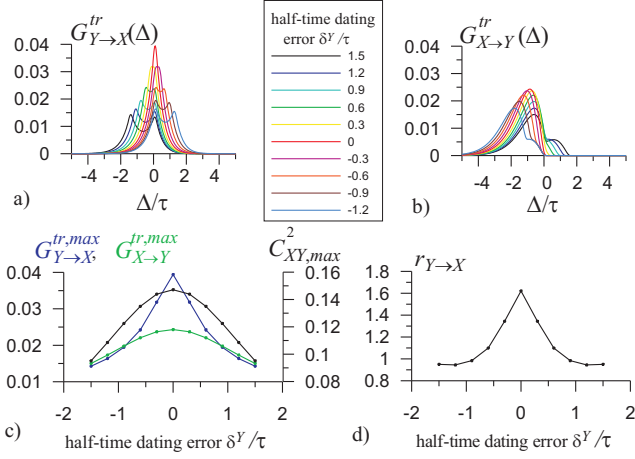
Fig. 1: Causality measures depending on dating error δ^Y for the system (2) at $h/\tau = 0.2$ and $\sigma_\Xi^2 = \sigma_\Psi^2 = 0$, k/α is such that $C_{X_0 Y_0, 0}^2 = 0.1$, $l_X = l_{XY} = l_Y = l_{YX} = 1$: (a,b) truncated WG causalities versus time lag for different dating errors; (c) maximal truncated WG causalities (blue and green) and maximum CCF value (black) and (d) causality ratio versus δ^Y.

Fig. 2: Causality measures depending on observational noise level σ_Ψ^2 for the system (2) at $h/\tau = 0.2$, $\sigma_\Xi^2 = 0$, $\delta^Y = 0$, $C_{X_0 Y_0, 0}^2 = 0.1$, and $l_X = l_{XY} = l_Y = l_{YX} = 1$: (a,b) truncated WG causalities versus time lag for different noise levels; (c) maximal truncated WG causalities (blue and green) and maximal CCF value (black) and (d) causality ratio versus $\sigma_\Psi^2/\sigma_{Y_0}^2$.

$C_{X_0 Y_0,0}^2 = 0.1$ and $\sigma_\Xi^2 = 0$, the "correct" $G_{Y \to X}^{tr,max}$ decreases with σ_Ψ^2 faster than $G_{X \to Y}^{tr,max}$ so that $r_{Y \to X}$ approaches unity at $\sigma_\Psi^2/\sigma_{Y_0}^2 > 0.5$ (Fig. 2). However, the noise Ξ in the driven signal increases $r_{Y \to X}$ apart from unity, which becomes quite visible as soon as $\sigma_\Xi^2/\sigma_{X_0}^2$ exceeds just 0.1. To summarize, large values of $\sigma_\Psi^2/\sigma_{Y_0}^2$ (50% and greater) along with small $\sigma_\Xi^2/\sigma_{X_0}^2$ (less than 10%) at moderate coupling strengths make the ratio $r_{Y \to X}$ close to unity. Hence, such a specific combination of noise levels can complicate inference of coupling direction from $r_{Y \to X}$.

To distinguish between impacts of observational noise and dating error from data, we can use either (i) assumptions about possible levels of both factors or (ii) shapes of the plots $G^{tr}(\Delta)$. For example, (i) if the noise is hardly greater than 20% in terms of variance, then $r_{Y \to X} < 1.1$ may be induced only by a dating error greater than $\tau/2$ (Figs. 2,c,d); (ii) if shapes of the plots $G_{Y \to X}^{tr}(\Delta)$ and $G_{X \to Y}^{tr}(\Delta)$ strongly differ from each other (cf. Figs. 1,a,b and 2,a,b), this is a sign of dating errors rather than observational noise. Being based on exact values of the causality ratio, such considerations are valid only for long enough time series, where statistical fluctuations can be neglected.

Note of causality ratio estimation. – Much smaller causality ratio estimates (e.g., 0.5) could appear in practice either due to a violation of Eq. (2) or too short time series. To give an analytic guess for possible statistical fluctuations of time series-based estimates, we note that the estimator $(N/l_{XY})\hat{G}_{Y \to X}^{tr}(\Delta)$ for sufficiently large N roughly follows χ^2 distribution with l_{XY} degrees of freedom, so the amplitude of its deviations from the mean for $l_{XY} = 1$ equals $3/N$ (the latter is the distance from 0.95-quantile to the mean) [29]. After maximization over a reasonable interval of the width $2\Delta_m = 4\tau$, the difference
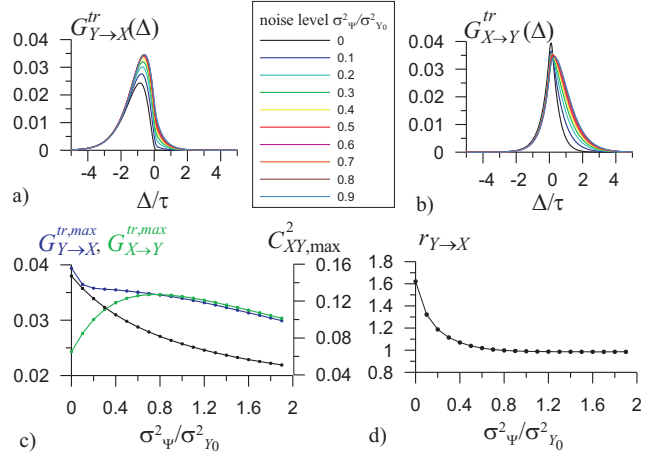
$\delta\hat{G} = \hat{G}_{Y \to X}^{tr,max}(\Delta) - \hat{G}_{X \to Y}^{tr,max}(\Delta)$ for $l_{XY} = l_{YX} = 1$ fluctuates with an amplitude of $(3/N)\sqrt{4 \cdot 2} \approx 9/N$. Denote the expectation of this difference δG. Then, $\delta\hat{G}$ and hence the estimator $\hat{r}_{Y \to X} = \hat{G}_{Y \to X}^{tr,max}/\hat{G}_{X \to Y}^{tr,max}$ are slightly affected by statistical fluctuations if the time series length is $N \gg 9/\delta G$. Hence, for a typical $\delta\hat{G} \approx 0.01$ (as in the following example) one should require $N \gg 900$. If a time series is shorter, the role of statistical fluctuations may well appear strong. For a detailed numerical study of such small sample effects, let us focus on situations close to the properties of the palaeoclimate data analyzed below.

Causality estimates from palaeoclimate data. –
A key problem in Climate Sciences is to understand and evaluate relative contributions of different factors to observed global and regional climate variations over time scales on the order of decades and longer. The best sources of such information from the pre-instrumental era are palaeoclimate proxies from different natural archives. One well-dated high-resolution reconstruction has been extracted from the stalagmite YOK-I from Yok Balum Cave (Southern Belize) [25]. The $\delta^{18}O$ record represents local to regional hydroclimate variations in that Atlantic region over the last two millennia with a mean temporal resolution of half a year and is characterized by very low dating errors (up to 17 yrs for ages about 2000 yrs). This time series (x signal) is examined here in parallel with the reconstruction of the total solar irradiance (TSI) based on ^{10}Be measurements on ice cores [26] to extract information on a possible influence of solar activity (y signal) on the Belize climate over the last two millennia.

The time series are presented in Figs. 3,a,b. The TSI data (Fig. 3,b) have originally been processed to remove the 11-yr solar cycle [26] and sampled in steps of $h = 5$

a) x [$\delta^{18}O$, permil VPDB], time, yr AD
b) y [TSI, W/m^2], time, yr AD
c) $\hat{C}_{XX}$, $\hat{C}_{YY}$, $\Delta/25$ yr
d) $\hat{C}_{XY}$, $\Delta/25$ yr
e) $\hat{G}^{tr}_{Y\to X}(\Delta)$, $\hat{G}^{tr}_{X\to Y}(\Delta)$, $\Delta/25$ yr
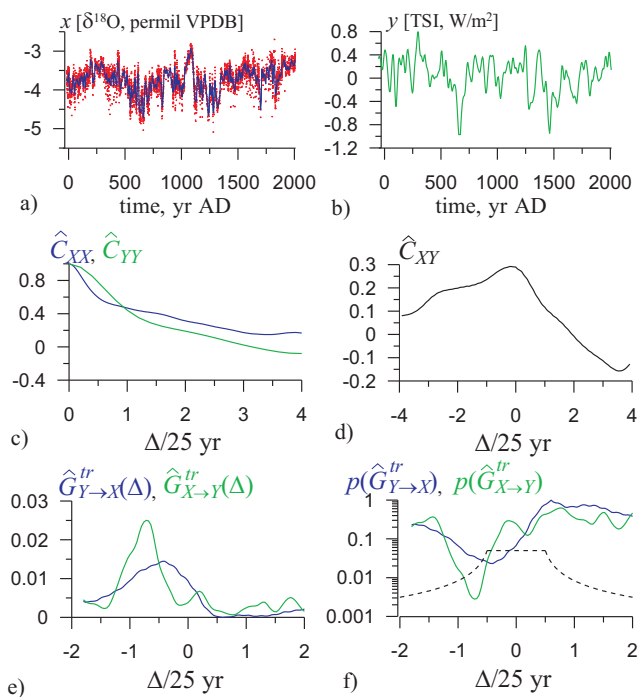f) $p(\hat{G}^{tr}_{Y\to X})$, $p(\hat{G}^{tr}_{X\to Y})$, $\Delta/25$ yr

Fig. 3: Estimation from palaeoclimate data over the period [15 yr BC - 2010 yr AD]: (a) time series of $\delta^{18}O$ from a speleothem representing local climate (moisture) in the Atlantic region, red points denote the original data, blue line – smoothed signal; (b) time series of solar activity (total solar irradiance); (c) sample ACF for the signals x (blue) and y (green); (d) sample CCF; (e) truncated WG causalities in the directions TSI $\to$ Belize climate (blue) and Belize climate $\to$ TSI (green) for $l_X = 3$, $l_{XY} = 1$, $l_Y = 4$, $l_{YX} = 1$; (f) the respective pointwise p-levels for the positivity of $\hat{G}^{tr}_{Y\to X}$ (blue) and $\hat{G}^{tr}_{X\to Y}$ (green), black dashed lines show the pointwise p-levels corresponding to the total p-level of 0.05 and obtained via the Bonferroni correction [33] with a pre-defined order of tests.

yrs. The original, nonequidistantly sampled YOK-I $\delta^{18}O$ values are shown as red dots in Fig. 3,a, the blue line shows the Gaussian kernel-based filtered [34] record (efficient width of 5 yrs) sampled equidistantly in smaller steps of 1 yr. The sample ACFs of both signals (Fig. 3,c) and their CCF (Fig. 3,d, $\hat{C}^2_{XY,max} = 0.09$) agree reasonably well with the hypothesis of the relaxators (2) with $\tau \approx 25$ yrs; some deviations may be attributed to statistical fluctuations. The resulting time series length is $N = 400$: the signal duration is 80τ, the sampling interval is 0.2τ.

To focus on the most statistically reliable results, we use the model orders selected via the Schwarz criterion for these data ($l_X = 3$ and $l_Y = 4$), even though everything is similar for the unit orders. The WG causality estimates differ from zero at least at the level of 0.05: $\hat{G}^{tr,max}_{Y\to X} = 0.014$ and $\hat{G}^{tr,max}_{X\to Y} = 0.025$ (Figs. 3,e,f). Since $\hat{G}^{tr}_{Y\to X}(\Delta)$ for the direction TSI $\to$ Belize climate is maximal at negative time lag Δ instead of an expected nonnegative lag, a possible dating error can be assumed. It is surprising that the causality ratio from TSI to Belize

climate is $\hat{r}_{Y\to X} = 0.56$, though we would expect much greater $r_{Y\to X} > 1.5$ without observational noise and dating errors and $r_{Y\to X} > 0.9$ with those distortions (Figs. 1, 2). Below, we study causality estimators for the same time series length and other parameters and check if statistical fluctuations suffice to explain such a low $\hat{r}_{Y\to X}$.

Causality estimates from short time series. – Taking $N = 400$ and $h/\tau = 0.2$, we generated an ensemble of 1000 time series by integrating Eqs. (2) with the Euler – Maruyama technique at time step of $\tau/300$ and imposing (or not) observational noise and dating errors. From each time series, we estimated WG causalities and causality ratio (for $l_X = 3$, $l_Y = 4$, $l_{XY} = l_{YX} = 1$). Then we calculated their mean values and probabilities to exceed threshold values equal to the respective palaeoclimate estimates [27]. The result is that for this data amount the effect of statistical fluctuations on the causality estimates is considerably stronger than that of dating errors (the second place) and observational noise (the third place).

Without observational noise and dating errors, we specify $k/\alpha = 0.45$ which gives CCF close to the palaeoclimate estimate. For smaller k/α (e.g. ≤ 0.3) the WG causality estimates are insignificant according to the F-test, while for greater k/α (e.g. ≥ 0.6) the CCF and WG causalities estimates considerably exceed the respective palaeoclimate values. The estimation shows that typically $\hat{r}_{y\to x} > 1$. A less typical case of $r_{y\to x} < 1$ (even down to 0.7) is observed in fewer than 10% of time series in an ensemble. Both WG causality estimates are significant at least at $p = 0.05$ in more than 90% of the time series. Appearance of the plots $\hat{G}^{tr}_{Y\to X}(\Delta)$ and $\hat{G}^{tr}_{X\to Y}(\Delta)$ is similar to Figs. 3,e,f, except for the locations of the maxima [27]: statistically significant $\hat{G}^{tr}_{Y\to X}(\Delta)$ has a maximum near zero, not at a negative lag. However, the half-time dating error $\delta^Y = 0.8\tau = 20$ yrs moves the maximum of $\hat{G}^{tr}_{Y\to X}(\Delta)$ to negative lags of $\Delta \approx -\delta^Y$ which is observed in about 50% of the ensemble. Thus, the system (2) with dating errors is closer to our palaeoclimate example.

The values of $\hat{r}_{Y\to X}$ depend on various factors [27]. For zero observational noise and zero dating errors the mean of $\hat{r}_{Y\to X}$ is 1.2 which is already low enough as compared to the theoretical $r_{Y\to X} = 1.6$, i.e., statistical fluctuations of the estimate already play the role of noise. The ratio $r_{Y\to X}$ decreases very slightly under increasing noise in the driving signal σ^2_Ψ even up to a very large 100% level (at zero noise in the driven signal). The probability to observe values of $\hat{r}_{Y\to X} \leq 0.56$ rises with σ^2_Ψ from 0.03 only up to to 0.05. The estimates of $r_{Y\to X}$ appear more sensitive to the dating error and their mean falls down to 1.1 already for moderate $\delta^Y = -0.8\tau$ and the probability of observing $\hat{r}_{Y\to X} \leq 0.56$ rises from 0.03 to 0.06 at $\delta^Y = -0.8\tau$ and even to 0.08 at $\delta^Y = -2\tau$ suggesting that the dating error is more probable to be of importance here than the observational noise. Overall, for a time series of the considered moderate length, statistical fluctuations are more influential than observational noise and dating errors: the former

decrease the causality ratio from 1.6 to 1.2, as compared to the change of the order of 0.1 induced by the dating error and 0.05 by observational noise. Thus, the time series length seems to be the main factor limiting the accuracy of the estimation for the palaeoclimate data at hand. Yet, as justified above, the relative importance of each factor depends on the time series length. In practice, it can be checked *ad hoc* for a time series at hand as is done here.

To develop a standard test for statistical significance, we note that under the null hypothesis of uncoupled processes the estimator $\hat{r}_{Y \to X}$ resembles the ratio of two χ^2-distributed quantities with l_{XY} and l_{YX} degrees of freedom. Maximization of $G^{tr}(\Delta)$ over an interval of width $2\Delta_m = 4\tau$ consisting of four independent segments corresponds to maximization of χ^2-distributed quantity over four independent trials. Numerical simulations show that for $l_{XY} = 1$ such a maximization results in the distribution which can be approximated by the χ^2 law with two degrees of freedom. Then, $\hat{r}_{Y \to X}$ is distributed according to Fisher's F-law with $(2, 2)$ degrees of freedom. However, quality of the approximation reduces for short time series, where Monte-Carlo based estimation seems more reliable.

Additional tests with simulations of a non-equidistant sampling from (2) and a subsequent Gaussian kernel-based filtering (all identical to the palaeoclimate case) show that it slightly increases the likelihood of the causality estimates obtained from the palaeoclimate data. Still, even in case of best correspondence, the system (2) exhibits characteristics similar to those in the palaeoclimate data only in 10% of all realizations. One reason for this limited agreement between the data and the stationary random process (2) can be temporal changes of some characteristics of the processes underlying the proxy records.

Nonstationarity of the palaeoclimate processes. – We have accounted for a possible nonstationarity by moving window analysis of the palaeoclimate data. The main results are presented in Fig. 4 for two non-overlapping time windows corresponding to the two subsequent millennia. Figs. 4,a,b (the first millennium A.D.) reveal a usual value of the causality ratio $r_{Y \to X} = 1.05 > 1$. Figs. 4,c,d do not reveal any significant couplings for the second millennium A.D. These results suggest a time-varying solar effect on the Belize climate. Similar analysis with moving windows of different lengths suggests that the transition between the two regimes has most probably occurred over the period 1000 to 1300 A.D. A strong influence in the first millennium A.D. would be in line with a northward position of the Intertropical Convergence Zone (ITCZ, see also [24]) and hence increased rainfall in Belize. A reduced solar influence in the second millennium A.D. could result from a southward displaced ITCZ during the Little Ice Age, and thus reduced tropical rain in Belize.

Our estimates for the first millennium A.D. (Figs. 4,a,b) show that the TSI variations lag the Belize climate proxy by about 20 yrs which seems unacceptable given that TSI should always lead the climatic signal (climatic response
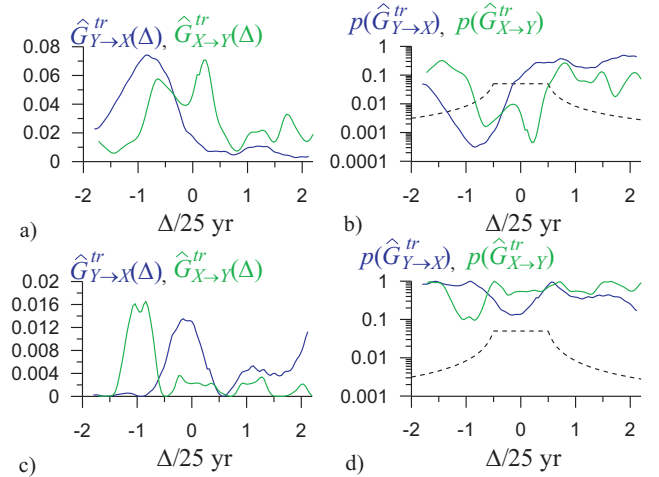

Fig. 4: Estimation from two non-overlapping 1000-yr intervals of the palaeoclimate data: (a,b) [15 yr BC – 985 yr AD]; (c,d) [985 yr AD – 1985 yr AD]. Panels (a,c) show truncated WG causalities in the directions TSI → Belize climate (blue) and Belize climate → TSI (green) for $l_X = 3$, $l_{XY} = 1$, $l_Y = 4$, $l_{YX} = 1$. Panels (b,d) show pointwise p-levels for positivity of $\hat{G}^{tr}_{Y \to X}$ (blue) and $\hat{G}^{tr}_{X \to Y}$ (green), black dashed lines show pointwise p-levels corresponding to the total p-level of 0.05.

to the Sun). Such a lag may well be determined by dating error of at least 20 yrs: Either the age of the solar signal is underestimated or the age of the cave signal is overestimated. Importantly, the question about which signal (or both) has a larger dating error is not possible to answer on the basis of bivariate data. We therefore include the best-dated ice-core based volcanic activity data [32] in our analysis (instead of the TSI data) to check whether its influence on the Belize climate (which is expected and well-accepted) is also characterized by a non-physical negative temporal shift [27]. We have found highly statistically significant volcanic forcing on speleothem $\delta^{18}O$ variations, the maximum of $G^{tr}_{Y \to X}(\Delta)$ being shifted to positive $\Delta = 2$ or 3 yrs, i.e. $\Delta \approx h/2$, that agrees with the notion of volcanic forcing delayed by no more than 1 yr. Such a small time delay is totally acceptable. Hence, the test with the volcanic record shows that there is excellent correspondence between eruptions recorded in ice cores and YOK-I which strongly supports the claim of highly accurate dating of the speleothem. Therefore, we conclude that it is the TSI record which is less accurately dated in the first millennium A.D., with a possible age underestimation of about 20 yrs.

Conclusions. – Dating errors are an almost inevitable characteristic of palaeoclimate time series which makes causality estimation even more difficult. We have proposed the causality ratio $r_{Y \to X}$ based on WG causality (1) as a relevant tool to cope with this problem. We have shown that the value of $r_{Y \to X} > 1$ correctly indicates the direction of unidirectional coupling $Y \to X$ for identical stochastic relaxators in the absence of observa-

tional noise and dating errors, if the sampling is not too sparse. Only very large observational noise in the driving signal (more than 50% in terms of variance) along with the noise-free driven signal makes $r_{Y \to X}$ close to unity and unsuitable for coupling directionality identification. The causality ratio is more sensitive to the dating error: if half a time series is dated with an error about the relaxation time τ or greater, $r_{Y \to X}$ gets close to unity again. Hence, in case of *a priori* known coupling direction, the value of $r_{Y \to X}$ allows to assess likely values of dating errors and observational noise level. However, statistical fluctuations of the estimates from sufficiently short time series may exceed the influence of dating errors and observational noise.

Applying the above results to analyze palaeoclimate data, we confirmed a strong influence of solar activity on the Belize climate over the first millennium A.D. and suggested that this influence strongly decreased in the second millennium. An unexpectedly low causality ratio appears to be determined by the shortness of the time series and, probably, the dating error in the solar proxy over the first millennium A.D. of about 20 yrs, the age of the solar data being underestimated. It seems to be an interesting and fruitful conclusion from an analysis of such a short piece of data on the basis of the adapted causality analysis.

The theoretical part of our research is based on the analysis of a simple, but basic test system (2). Further studies of the influence of dating errors and other factors on WG causalities for more general systems are relevant, including non-identical processes, higher dimensionality of state spaces, and various kinds of nonlinearity. More "inertial" couplings can be analyzed with $l_{XY} > 1$ and even with l_{XY} different temporal shifts rather than with a single Δ. All these features will possibly reveal more complicated relationships between the causality ratio and coupling directionality which can then be taken into account, extending the range of applicability of the approach to all fields where dating errors are encountered. Yet, the research presented here is valuable as the first step which already reveals that the adapted WG causality analysis is a promising tool to deal with data corrupted by dating errors and extract information about underlying causal couplings.

$* * *$

REFERENCES

[1] E. Pereda, R. Quian Quiroga, and J. Bhattacharya, Progr. Neurobiol. **77** (2005) 1.

[2] M. Winterhalder, B. Schelter, and J. Timmer (eds.), *Handbook of Time Series Analysis* (Berlin: Wiley-VCH, 2006).

[3] K. Hlavackova-Schindler, M. Palus, M. Vejmelka, and J. Bhattacharya, Phys. Rep. **441** (2007) 1.

[4] B.P. Bezruchko and D.A. Smirnov, *Extracting knowledge from time series: An introduction to nonlinear empirical modeling* (Berlin, Heidelberg: Springer-Verlag, 2010).

[5] M. Wibral, B. Rahm, M. Rieder, et al, Prog. Biophys. Mol. Biol. **105**, 80 (2011).

[6] A. Attanasio, A. Pasini, and U. Triacca, Atmospheric and Climate Sciences **3** (2013) 514.

[7] C.L. Webber and N. Marwan (eds.), *Recurrence Quantification Analysis Theory and Best Practices* (Berlin, Heidelberg: Springer-Verlag, 2015).

[8] A. Mueller, J.F. Kraemer, T. Penzel, H. Bonnemeier, J. Kurths, and N. Wessel, Physiol. Meas. **37** (2016) R46R72.

[9] N. Wiener, in E.F. Beckenbach (ed.) *Modern Mathematics for the Engineer* (New York: McGraw-Hill, 1956).

[10] C.W.J. Granger, Econometrica **37** (1969) 424.

[11] J. Runge, J. Kurths, and V. Petoukhov, J. Climate **27**(2014) 720.

[12] T. Schreiber, Phys. Rev. Lett. **85** (2000) 461.

[13] N. Ay and D. Polani, Adv. Complex Syst. **11** (2008) 17.

[14] J.T. Lizier and M. Prokopenko, Eur. Phys. J. B **73** (2010) 605.

[15] X. San Liang, Phys. Rev. E **90** (2014) 052150.

[16] A. Montalto, S. Stramaglia, L. Faes, G. Tessitore, R. Prevete, and D. Marinazzo, Neural Networks **71** (2015) 159.

[17] H. Nalatore, M. Ding, and G. Rangarajan, Phys. Rev. E **75** (2007) 031123.

[18] D.W. Hahs and S.D. Pethel, Phys. Rev. Lett. **107** (2011) 128701.

[19] D.A. Smirnov and B.P. Bezruchko, Europhys. Lett. **100** (2012) 10005.

[20] D.A. Smirnov, Phys. Rev. E **87** (2013) 042917.

[21] D.A. Smirnov, Phys. Rev. E **90** (2014) 062921.

[22] D.A. Smirnov and I.I. Mokhov, Phys. Rev. E **92** (2015) 042138.

[23] P.D. Jones, K.R. Briffa, T.J. Osborn, et al, The Holocene **19** (2009) 3.

[24] H.E. Ridley, Y. Asmerom, J.U.L. Baldini, et al, Nat. Geosci. **8** (2015) 195.

[25] D.J. Kennett, S.F.M. Breitenbach, V.V. Aquino et al, Science **338** (2012) 788.

[26] F. Steinhilber, J. Beer, and C. Froehlich, Geophys. Res. Lett. **36** (2009) L19704.

[27] Supplementary material on the web-site of Europhysics Letters https://epletters.net/.

[28] G. Schwarz, Ann. Stat. **6** (1978) 461.

[29] G.A.F. Seber, *Linear Regression Analysis* (New York: Wiley, 1977).

[30] K. Hasselmann, Tellus **28** (1976) 473.

[31] I.I. Mokhov, D.A. Smirnov, P.I. Nakonechny, S.S. Kozlenko, Ye.P. Seleznev, and J. Kurths, Geophys. Res. Lett. **38** (2011) L00F04.

[32] M. Sigl, M. Winstrup, J.R. McConnell, et al. Nature **523** (2015) 543.

[33] E.L. Lehmann, *Testing Statistical Hypotheses* (New York: Springer, 1986).

[34] K. Rehfeld, N. Marwan, S.F.M. Breitenbach, and J. Kurths, Clim. Dyn. **41** (2013) 3.