Stochastic assessment of temperature – CO_2 causal relationship in climate from the Phanerozoic through modern times

Preprint · May 2024

DOI: 10.13140/RG.2.2.10126.98883

CITATIONS
0

READS
61

1 author:

Demetris Koutsoyiannis
National Technical University of Athens
984 PUBLICATIONS **17,874** CITATIONS

SEE PROFILE

PREPRINT IN REVIEW

Article

Stochastic assessment of temperature – CO_2 causal relationship in climate from the Phanerozoic through modern times

Demetris Koutsoyiannis*

Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Zographou, Greece, http://www.itia.ntua.gr/dk

* **Correspondence:** dk@itia.ntua.gr

Abstract: As a result of recent research, a new stochastic methodology of assessing causality was developed. Its application to instrumental measurements of temperature (T) and atmospheric carbon dioxide concentration ($[CO_2]$) over the last seven decades provided evidence for a unidirectional, potentially causal link between T as the cause and $[CO_2]$ as the effect. Here we refine and extend this methodology, and apply it to both paleoclimatic proxy data and instrumental data of T and $[CO_2]$. Several proxy series, extending over the entire Phanerozoic or parts of it, gradually improving in accuracy and temporal resolution up to the modern period of accurate records, are compiled, paired and analyzed. The extensive analyses made converge to the single inference that change in temperature leads, and that in carbon dioxide concentration lags. This conclusion is valid for both proxy and instrumental data in all time scales and time spans. The time scales examined start from annual and decadal for the modern period (instrumental data) and the last two millennia (proxy data), and reach one million years for the most sparse time series for the entire Phanerozoic. The type of causality appears to be unidirectional, $T \rightarrow [CO_2]$, as in earlier studies. The time lags found depend on the time span and time scale and are of the same order of magnitude as the latter. These results contradict the conventional wisdom, according to which the temperature rise is caused by $[CO_2]$ increase.

Keywords: causality; causal systems; stochastics; impulse response function; geophysics; climate

Crime begins with propaganda, even if such propaganda is for a good cause

Hans Fritzsche [1]

1 Introduction

While causality is a fundamental notion in science and life, there also exist fundamental problems, on philosophical, scientific and practical grounds, in its essence and its identification [2]. These problems are manifested in controversies, which are not only theoretical but have important social, political and economic implications. One of the most controversial issues of our time is the causal relationship between atmospheric temperature (T) and carbon dioxide (CO_2) concentration ($[CO_2]$) in Earth's climate.

Earth's climate has varied in all times and on all time scales. Its variability should not be regarded a puzzle, given its huge complexity and the connections of the climatic system with numerous agents of change, internal, such as atmospheric composition, hydrological processes and biosphere evolution, or external, such as geologic activity and tectonic changes, solar activity, galactic cosmic ray flux, and orbital changes. Rather, the puzzling issue is the relative stability (small variation) of Earth's climate. Specifically, geological evidence presented by Veizer [3-5] suggests the presence of running water as far back as we have a record, up to 3.8 or even 4.2 billion years, while it was thought that Earth would be in an ice ball state up to about 1 billion years ago because of the much smaller solar irradiance. This is known as the faint young sun puzzle. In addition, the carbon isotope record suggests that the fundamentals of the planetary carbon cycle and life may have been established as early as 3.5 billion years ago, which would necessitate the presence of oceans on Earth in turbulent motion.

However, the reconstructed climate information of the long distant past, before the geologic period known as the Phanerozoic eon (541 million years), is unclear. As strange as it may seem, at a time horizon of the order of 100 million years and beyond, even the location and orbit of the Earth become uncertain as the solar system becomes chaotic and unpredictable in such long horizons [6-10]. Certain orbital aspects influence climate. The exploration of the latter was pioneered, during the Hellenistic period by Hipparchus (Ἵππαρχος ὁ Νικαεύς; c. 190–c. 120 BC), who discovered and calculated the precession of the equinoxes (μετάπτωσις ἰσημεριών) by studying measurements on several stars. In the 20th century, this precession would be found to be related to the climate, constituting one of the so-called Milanković cycles, after the Serbian civil engineer Milutin Milanković (1879 – 1958) [11-13]. These cycles, i.e., oscillations of the Earth's eccentricity, axial tilt, and precession, are important drivers of climate.

The influence of orbital cycles on climate for the late Quaternary period, namely the last half million years, has been substantial. The relevant research has been pioneered by Hays et al. [14], who used three indices of global climate and found agreement of these records with dominant periods of the Earth's solar orbit, namely 23, 42 and approximately 100 thousand years. He concluded that changes in the Earth's orbital geometry are the fundamental cause of the succession of Quaternary ice ages. More recently, Roe [15] demonstrated that it is the effect of the Milanković cycles that explains the glaciation process in the Quaternary. To this aim, he effectively used reconstructions of global ice volume during the last 750 thousand years, which relied on measurements of oxygen isotopes

incorporated into the shells of foraminifera recovered from deep-sea sediment cores.

However, if we go further back to the past, to cover the entire Phanerozoic eon, for which reliable reconstructions exist, or merely the Cenozoic era (the last 66 million years), in which the reconstructions have superior resolution and reliability, other climate drivers may become more important. In such a long period, the tectonics play important role, as the Earth's continents and oceans are moving (see Figure 1 for a couple of examples) and clearly the distribution of land and sea on the globe affects climate. In addition, the solar radiation was not constant but rose consistently, with a total 5% increase over the Phanerozoic [16]. It is also accepted that the cosmic ray flux has a large effect on the climate and this flux had variations, including a cycle with a period of about 145 million years, corresponding to the passage of the solar system through one of the two sets of spiral arms of the Milky Way [16]. Furthermore, life on Earth has evolved and this influenced climate substantially, as the biosphere processes control the chemical synthesis on the atmosphere and the water solvents in the oceans.

A particular substance in the chemical synthesis of the atmosphere that has been regarded as an essential climate driver is CO_2. It has been generally thought that it is a strong greenhouse gas and its concentration in the atmosphere determines temperature. In the last decades, the observed atmospheric [CO_2] increase was presented as the cause of the (also observed) temperature rise. In addition, the [CO_2] increase in the last century was blamed on human actions, namely the burning of fossil fuels. Huge amounts of funding were directed to the relevant research on the relationship of CO_2 with climate, and the climate change has been identified with CO_2 emission change. According to the established narrative (a product of this funding), which is simplistic and negligent of the huge complexity of the climatic system, humans emit CO_2 by burning fossil fuels, and the emissions accumulate into the atmosphere and increase the greenhouse effect causing increase in temperature. However, it is healthy to question whether these premises and the entire narrative are valid. Is [CO_2] a climate driver or an internal state variable of the climatic system? Does [CO_2], with a share of about 5% in the greenhouse effect [17], have a decisive role in it? Has the intensity of the greenhouse effect increased in the last century? Have humans become a climate driver by pumping CO_2 into the atmosphere through burning fossil fuels? In particular, is the human share in CO_2 emissions, i.e. 4% of the total, decisive in determining [CO_2]? Or are the biosphere processes, with a share of 96% in emissions, more important than human actions? Does [CO_2] change cause temperature change ($[\mathrm{CO_2}] \rightarrow T$), or is the causality direction the exact opposite ($T \rightarrow [\mathrm{CO_2}]$) or bidirectional (also known as Hen-Or-Egg; HOE; $T \leftrightarrow [\mathrm{CO_2}]$)?

These questions have been investigated in a number of recent publications [2, 18-25] and the answers given mostly contradict the established narrative. Here we study the last question using proxy data from the entire Phanerozoic and modern instrumental data, and a stochastic methodology. This question has already been studied in a similar (stochastic) framework [2,20,22,23] but here we use a richer data set and we propose some improvements of the methodology.

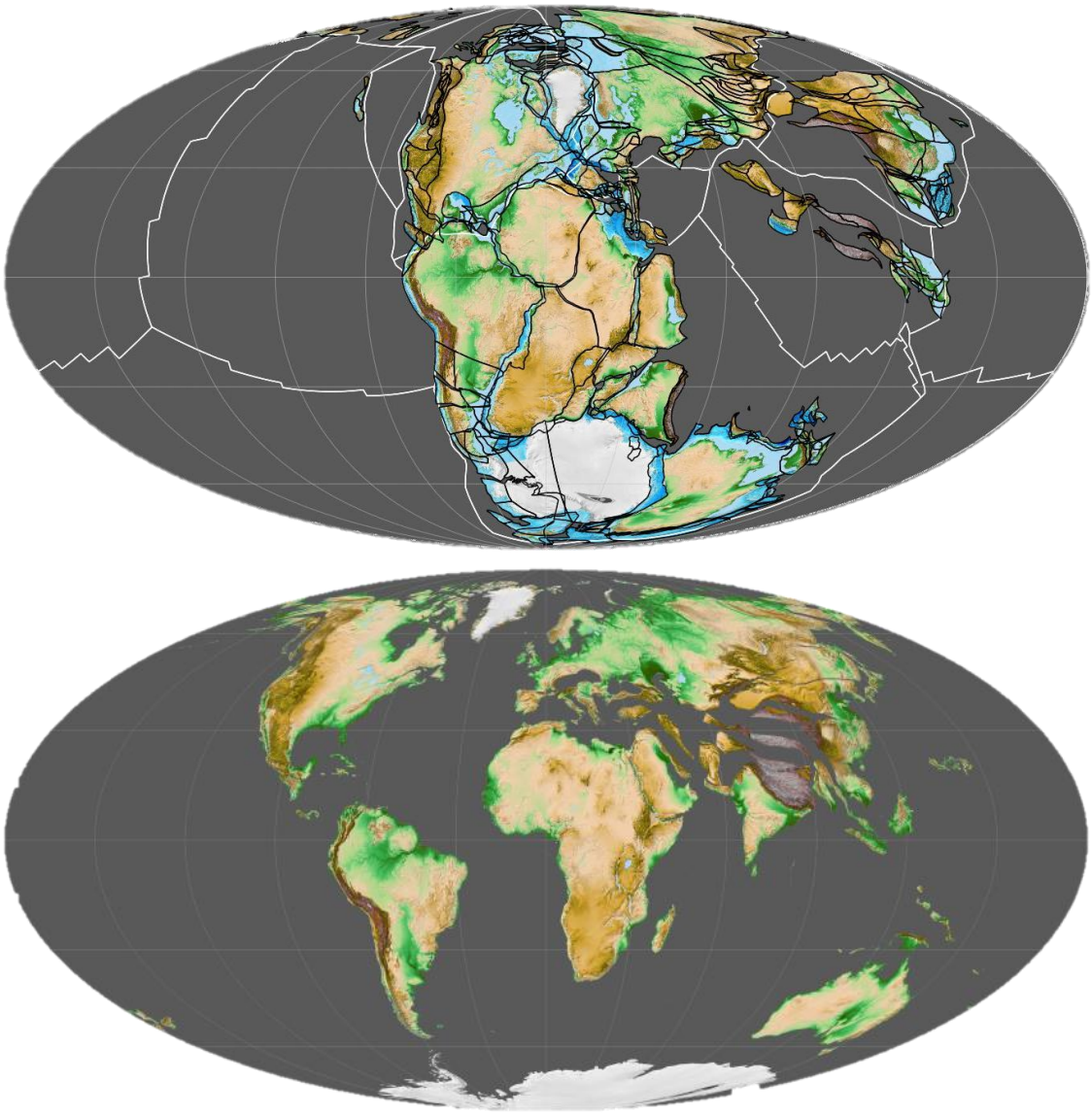

Figure 1 Reconstruction of Earth's continents and oceans in the Phanerozoic eon, (**upper**) during the Mesozoic era, about 200 million years ago (Pangaea) and (**lower**) during the Cenozoic era (Eocene epoch), about 40 million years ago (images retrieved, respectively, from https://en.wikipedia.org/wiki/Pangaea#/media/File:Pangaea_200Ma.jpg and https://upload.wikimedia.org/wikipedia/commons/archive/c/c6/20230814125217%2140_Ma_paleoglobe.png, both licensed under the Creative Commons Attribution-Share Alike 4.0 International license).

The question is crucial for the established climate narrative and thus, not surprisingly, several works tried to verify the $[CO_2] \rightarrow T$ direction using paleoclimatic reconstructions and with methodologies different from the stochastic framework, but without success. Yet, a study by Shakun et al. [26] claimed to have shown "*that temperature is correlated with and generally lags CO_2 during the last (that is, the most recent) deglaciation*". But even this claim was for a short period and only for the north hemisphere as the lead-lag relationships were found different for the two hemispheres. Given the unsuccessful attempts to support the $[CO_2] \rightarrow T$ case, other studies sought a synchrony claim, on

the basis that the estimated positive lags of $[CO_2]$ from T are often within the 95% uncertainty range [27], that the asynchrony is not significant [28], or that a "*short lead of CO_2 over temperature cannot be excluded*" [29].

In a more neutral approach, Berner and Kothavala [30] studied the entire phanerozoic (the last 530 million years) asserting that "*over the long term there is indeed a correlation between CO_2 and paleotemperature*", and emphasized the "*importance of considering ALL factors affecting CO_2 when modelling the long term carbon cycle and not concentrating* [on] *only one cause*". On the other hand, Veizer et al. [31] presented evidence for decoupling of atmospheric CO_2 and global climate during the phanerozoic, questioning the role of the CO_2 as the main driving force of past global (long-term) climate changes, at least during two of the four main cool climate modes of the phanerozoic. Later, referring to the Cenozoic, he asserted [5] that "*the temperature signal appears to precede carbon signal, arguing for climate being the cause of variability in the carbon cycle*". Referring to the latest one million years, he asserted that *"while ice cores demonstrate conclusively that on multimillennial time scales climate and greenhouse gases correlate, they again show that the shifts in climate precede the trends in gas concentrations*." For the last 750 thousand years, Roe [15] found that "*variations in atmospheric CO_2 appear to lag the rate of change of global ice volume. This implies only a secondary role for CO_2* […] *in driving changes in global ice volume.*" For the most recent ~400 thousand years, several studies, based on paleoclimatic reconstructions and mostly the Vostok ice cores (see Section 2.3), have also identified the precedence of temperature change over that in $[CO_2]$, with estimates of the time lag varying from 50 to 1000 years, depending on the time period and the particular study [20,22,32,33]. Finally, for the recent period covered by instrumental data with monthly time step, the most recent studies by Koutsoyiannis et al. [22,23] found a unidirectional, potentially causal link between temperature as the cause and $[CO_2]$ as the effect with time lags of the order of 0.5 to 1 year for time scales from 1 to 16 years, thus corroborating earlier studies with similar findings using different methods [34].

Here, as described in Section 2, we use temperature and $[CO_2]$ data series for all these time frames, namely proxy data for the Phanerozoic, the Cenozoic, the late Quaternary and the Common Era, and instrumental data for the last seven decades. In Section 3 we explain the inappropriateness of deterministic methodologies for detecting causality and we adapt the stochastic methodology proposed in [2,22] to be used with the above data series. In Section 4 we present in detail the results of the methodology application to each of the data series, which we summarize and further discuss in Section 5. As will be seen in Section 6, these results allow drawing sound conclusions.

2 Data

2.1 Phanerozoic

As evident by its name, Phanerozoic (from the Greek *φανερός* and *ζωή*, visible life) started with the emergence of complex life forms, followed by the emergence of animals on land, and later was

characterized by the rise of flowering plants (angiosperms), and diversification of species, as indicated in Figure 2. Several aspects of life fossils (stomata in leaves, soils) enable broad estimates of temperature and atmospheric carbon dioxide levels. Several reconstructions of T and [CO_2] over the Phanerozoic eon have been recently published. For such very long period, temperature reconstructions are of three types [16]: (a) based on lithological indicators of climate, such as coals, evaporites, bauxites, and tillites, which provide information about Earth's past climatic zones (Köppen belts) and the pole-to-equator temperature gradient; (b) geochemical reconstructions that use oxygen isotope measurements of paleotemperature; and (c) combinations of the two. Here we retrieved four series constructed in the last 5 years, two of category (c), namely those by Scotese et al. [35] and Shaviv et al. [16] (who also reproduced the series by Scotese et al. [35])], as well two of category (b), those by a Grossman & Joachimski [36], and Song et al. [37]. These four series are depicted in Figure 2, where a general agreement is seen for the last 350 million years, but, earlier than that, the series of isotopic origin show much higher paleotemperatures than the combined ones, almost reaching 50 °C at about 500 million years before present (BP). Such high temperatures may not be implausible, as additional evidence shows huge climatic shifts, even for periods more recent than that. For example, for the Cretaceous Period (144–66 million years ago) it was estimated that the average annual low-latitude sea surface temperatures probably reached ~35 °C, that the global sea level was ~170 m higher than at present and that Antarctica was covered in rainforest [38]. Greenland was nearly ice-free for extended periods even in more recent periods, i.e. during the Pleistocene [39,40].

Atmospheric CO_2 concentrations are most often reconstructed from the GEOCARB model, pioneered by Berner and Kothavala [30] and later expanded in the GEOCARB-SULF model to describe the carbon, sulfur, and oxygen cycles, which integrates various proxy measurements of CO_2 [41,42]. A compilation of about 1500 discrete estimates of atmospheric CO_2 from 112 published studies covering the last 420 million years, was presented by Foster et al. [43]. Here, as shown in Figure 2, we retrieved three different time series published in the last 16 years, namely by Royer [44] (also contained in Davis [45]), Foster et al. [43] (also reproduced by Song et al. [37]) and Berner [46]. Most of the time series are available at a time step of 1 million years. Thus, for the application in causality assessment, we use this time step for all time series, after converting to that fixed time step those series which are given at different time step using linear interpolation.

The reconstructions of both temperature and [CO_2], depicted in Figure 2, show large variations, while the disagreements among them suggest high uncertainty. Apparently, the uncertainty is not limited to the magnitude of each quantity, but extends also to the dating of each data point. This creates even higher uncertainty in assessing the causality direction, which is the subject of this paper.

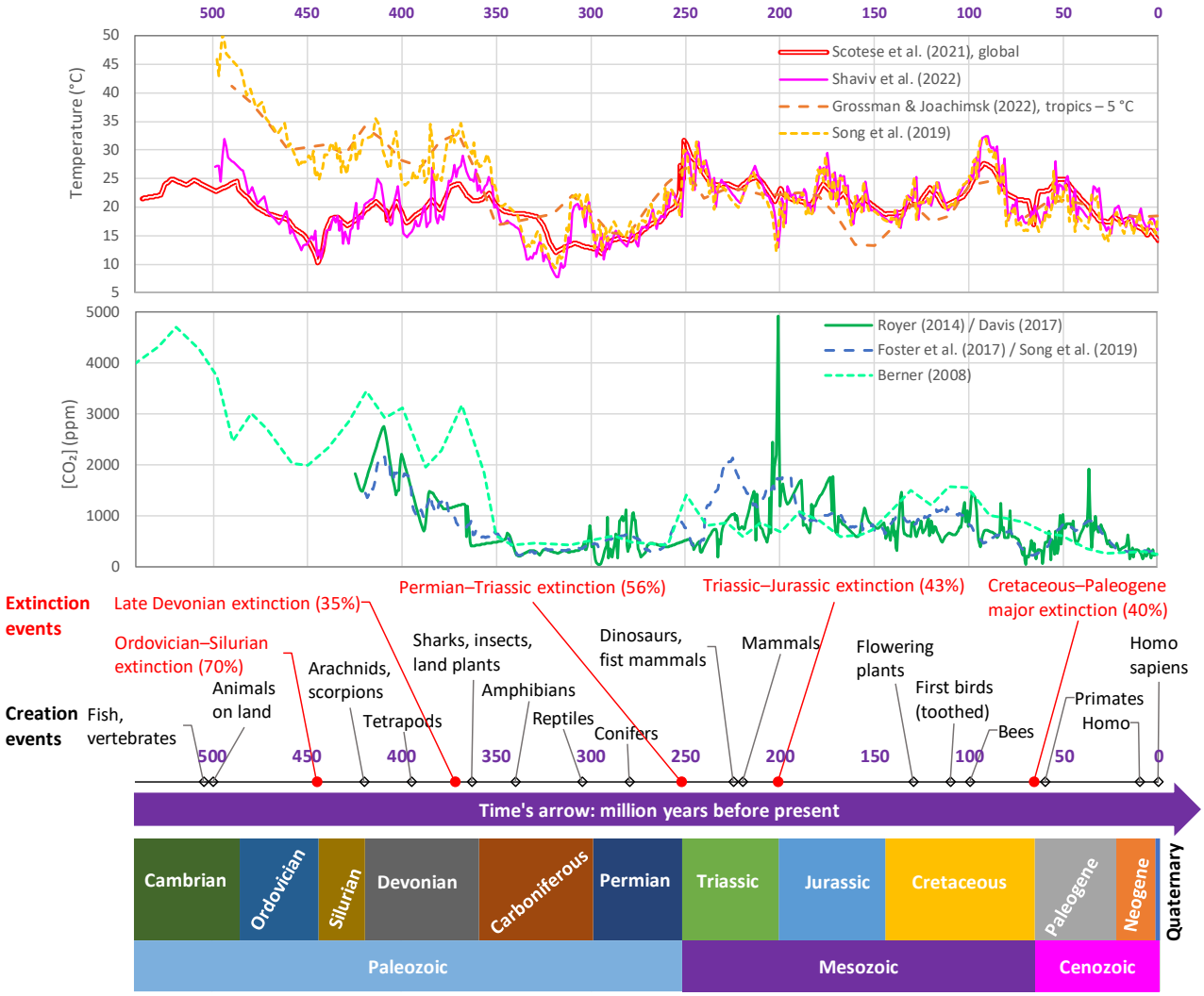

Figure 2 Recent reconstructions of atmospheric temperature and carbon dioxide concentrations over the Phanerozoic eon (541 million years). Sources of temperature series: Scotese et al. [35] (after digitization of their figure 17, Global Average Temperature); Shaviv et al. [16] (table in Supporting Information and their figure 2); Grossman & Joachimski [36] (their figure 4 and supplementary material, table S3, after offsetting by $\Delta T = -5$ °C to make it consistent with the other time series, as it originally refers to low latitude); and Song et al. [37] (their figure 3 and supplementary material, table S1). Sources of [CO_2] series: Royer [44] / Davis [45] (figure 5 in Davis [45], after digitization); Foster et al. [43] / Song et al. [37] (figure 1 and supplementary data 2 in [43]); Berner [46] (their figure 1B, after digitization). For completeness, the divisions of the Phanerozoic into three geological eras and 12 geological periods are shown, along with information on creation events (first appearance of the indicated taxa) and extinction events, with percentages of mean genus loss, taken from references [47,48].

The large variations reflect the cosmogonic changes that occurred during the Phanerozoic, both in geology and in the evolution in the biosphere. The former are exemplified in Figure 1, which shows two likely instances of Earth's continents and oceans, indicative of the huge tectonic changes that

occurred in the Phanerozoic. Figure 2 also includes the divisions of the Phanerozoic into geological eras and periods. The latter is epigrammatically noted in Figure 2 in terms of creation and extinction events. New fauna and flora (creation and expansion events) are associated to increase in diversity, which seems to align with an approximately exponential decline of the previously dominant fauna and flora, suggesting possible displacement of each evolutionary species by its successor [49].

2.2 Cenozoic

As seen in Figure 2, Cenozoic is the latest geological era, which began about 66 million years ago. It has been a dynamic period marked by the dominance of mammals and birds, changes in vegetation and shifts in ecosystems. Some of these changes are marked in Figure 3. On the other hand, the continents assumed their modern configuration (see Figure 1). These conditions have allowed reconstruction of climate records with superior resolution and reliability.

Continuous composite astronomically dated climate reconstructions have been presented by Westerhold et al. [50] based on oxygen and carbon isotope variations in deep-sea benthic foraminifera (single-celled amoeba-like organisms, producing calcite shells, widely used in environmental reconstruction). The reconstructions are available online; here we retrieved the oxygen-18 ($\delta^{18}O$) time series, which is a temperature proxy.

For [CO_2], Rae et al. [51] compiled paleo reconstructions from marine archives based on carbon isotopes in alkenones (long-chain unsaturated ethyl and methyl ketones, commonly used as biomarkers in sediments) and boron isotopes in foraminifera. Both sources of information, as shown in figure 6c in [51], are retrieved and used here.

Both $\delta^{18}O$ and [CO_2] reconstructions are depicted in Figure 3. For better understanding, a temperature scale has also been drawn, by transforming $\delta^{18}O$ to temperature based on the Epstein et al. [52] formula. It is noted that Rae et al. [51], who also used the Westerhold et al. [50] reconstructions, converted $\delta^{18}O$ to temperature (their figure 6a) using a different algorithm and found somewhat lower temperatures than those inferred from the temperature scale of Figure 3. However, here we use the $\delta^{18}O$ time series per se (multiplying it by –1) without conversion.

We note that in the [CO_2] time series, the time distance between consecutive samples is highly variable, with an average of 73 thousand years, while this is much lower for the $\delta^{18}O$ time series. Based on this, for the application to assessment of causality we chose a fixed time step of 100 thousand years and we performed linear interpolation to covert the two time series from irregular to regular time step.

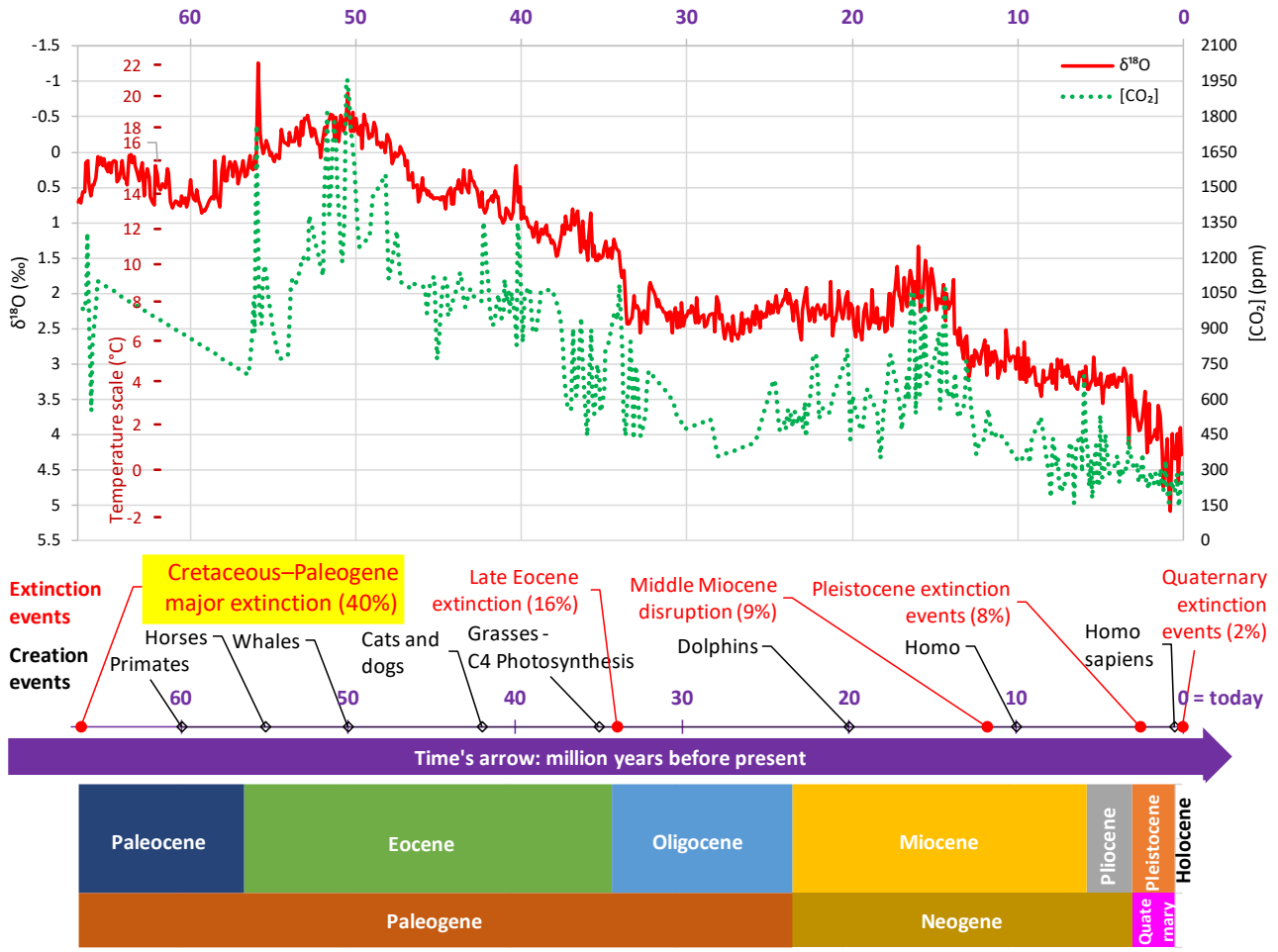

Figure 3 Recent reconstructions of the isotopic ratio $\delta^{18}O$, which is a temperature proxy with the indicated temperature scale, and $[CO_2]$ over the Cenozoic era (66 million years). Data for temperature are from Westerhold et al. [50] (available online in their Supplementary Material https://www.science.org/doi/suppl/10.1126/science.aba6853/suppl_file/aba6853_tables_s8_s34.xlsx; Table S34; column "ISOBENd18oLOESSsmooth"), and for $[CO_2]$ from Rae et al. [51] (available online in their Supplemental Material, https://www.annualreviews.org/deliver/fulltext/earth/49/1/ea49_rae_suppl_data1.xlsx; columns "xco2_84pc" and "co2" for alkenones and boron isotopes, respectively, after merger). For completeness, the divisions of the Cenozoic era into three geological periods and seven geological epochs are shown, along with information on creation events (first appearance of certain families or species) and extinction events, with percentages of mean genus loss, taken from references [47,48].

2.3 Late Quaternary

As seen in Figure 3, Quaternary is the most recent of the three geologic periods of the Cenozoic era, which spans from 2.58 million years ago to the present. Due to the relatively lower temperatures during it, also seen in Figure 3, ice has been permanent at least in Antarctica for the latest part of it. This enabled better climatic reconstructions from ice cores in Antarctica. Here we use the Vostok ice core

data, which include both temperature and [CO_2] [53,54] and whose dating studies have been most extensive.

The two time series go back to about 420 000 years before present (more precisely, before 1950) and, as seen in Figure 4, are strongly correlated to each other. They are originally given at an irregular time step, which needs to be regularized for causality assessment. The time steps are quite different for the two series. Namely, the average time step for the T and [CO_2] time series are ~128 and ~1150 years, respectively. It must be noted that the age of the gas (air bubbles) at a certain ice layer differs from the ice age. Specifically, the air extracted from the ice is younger than the surrounding ice. Extensive studies to date the air with respect to ice by Barnola et al. [55] concluded that "*the age difference between air and ice is about 6000 years during the coldest periods instead of about 4000 years, as previously assumed*". Indeed, in the data provided online, the difference between the ice and air age is about 6000 years for the coldest periods, decreasing to about 2000 years for the hottest periods and averaging at a difference of ~4070 years. Naturally, in our calculations for the CO_2 time series we used the air age, as given in the publicly available data series.

The regularized time step was initially chosen at 1000 years, i.e., close to the average time step of the original [CO_2] time series. After a detailed investigation, it was seen that in the period from 100 to 300 thousand years BP, the data are denser in time, with an average time step of about 800 years which goes down to 500 years for some subperiods. Therefore, we also constructed a second couple of time series with time step of 500 years spanning the period 100 to 300 thousand years BP.

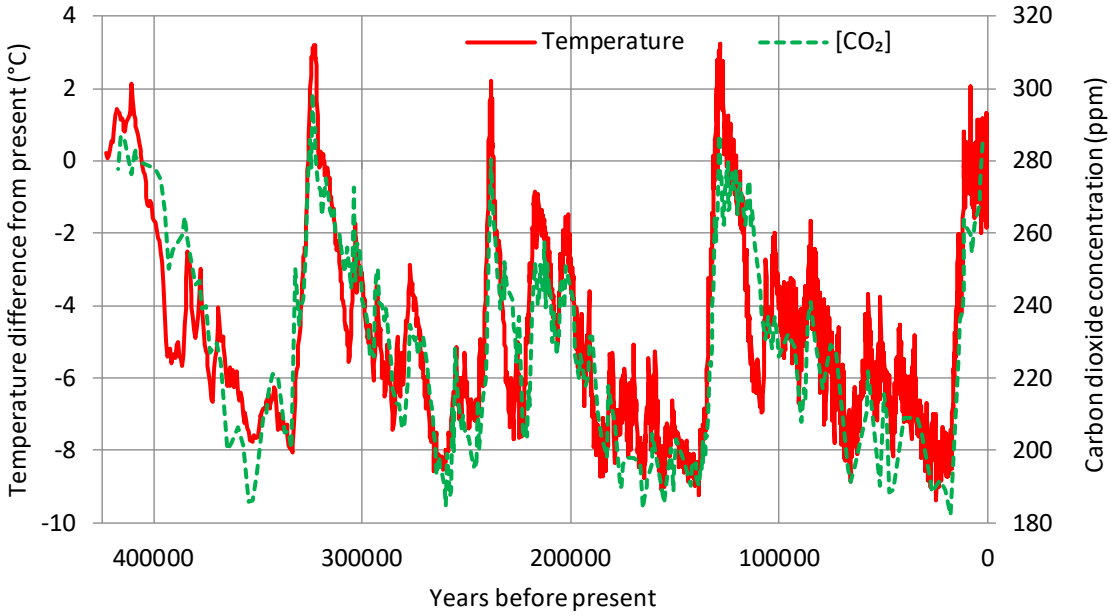

Figure 4 Time series of temperature and CO_2 concentration from the Vostok ice core, retrieved, respectively, from http://cdiac.ess-dive.lbl.gov/ftp/trends/temp/vostok/vostok.1999.temp.dat and from http://cdiac.ess-dive.lbl.gov/ftp/trends/co2/vostok.icecore.co2.

2.4 Common Era

The climate variations in the last two millennia (a period known as "Common Era") have been intensively studied and, at the same time, have been one of the most controversial issues of the last decades. Here we include this period in our analyses for its wide interest, albeit expecting high uncertainty in the results due to the controversy and the reasons behind it.

In their recent review paper, Christiansen and Ljungqvist [56], reproduced 18 reconstructions covering the last two millennia in part or completely (their figure 1). Of those, here we used the three reconstructed series that are (almost) complete for the entire Common Era. Specifically, these are those by Moberg et al. [57], who combined low-resolution proxies with tree-ring data, Loehle and McCulloch [58], who used non-tree-ring data that had been previously calibrated and converted to temperature by other authors, and Christiansen and Ljungqvist [59], who used different types of proxies and of different resolutions, from annual to decadal. As seen in Figure 5 (upper), the three series differ substantially from each other, in magnitude, resolution and smoothing.

Reconstructions of paleo-atmospheric CO_2 levels based on ice core data from Antarctica (Indermühle et al. [60] and Etheridge et al. [61]; Francey et al. [62]; Böhm et al. [63]), show a rather stable [CO_2] except after 1800 AD, when it started to increase. However, a [CO_2] reconstruction based on stomatal frequency analysis (namely, of buried Tsuga heterophylla needles) by Kouwenberg [64] reveals significant centennial-scale atmospheric CO_2 fluctuations. Some of the data from [64] are also reproduced in Kouwenberg et al. [65], namely the data post 800 AD, even though in the original figure in [64] they start at 200 AD; the reasons are not explained but perhaps can be guessed by viewing Figure 5 (lower), which reproduces them. An additional stomata-based reconstruction for the period 1000-1500 AD presented by van Hoof et al. [66] also shows fluctuation (albeit smaller than Kouwenberg's) and disagreement with the ice core data. The agreements and disagreements of all these data series are shown in Figure 5. Additional ice core data but from Greenland [67,68], not contained in Figure 5, show much higher variation with respect to the Antarctica data. For completeness, we note that stomata-based reconstructions, presented for other periods of the Holocene prior to the last two millennia [69-74] (not shown in Figure 5) also present high variation.

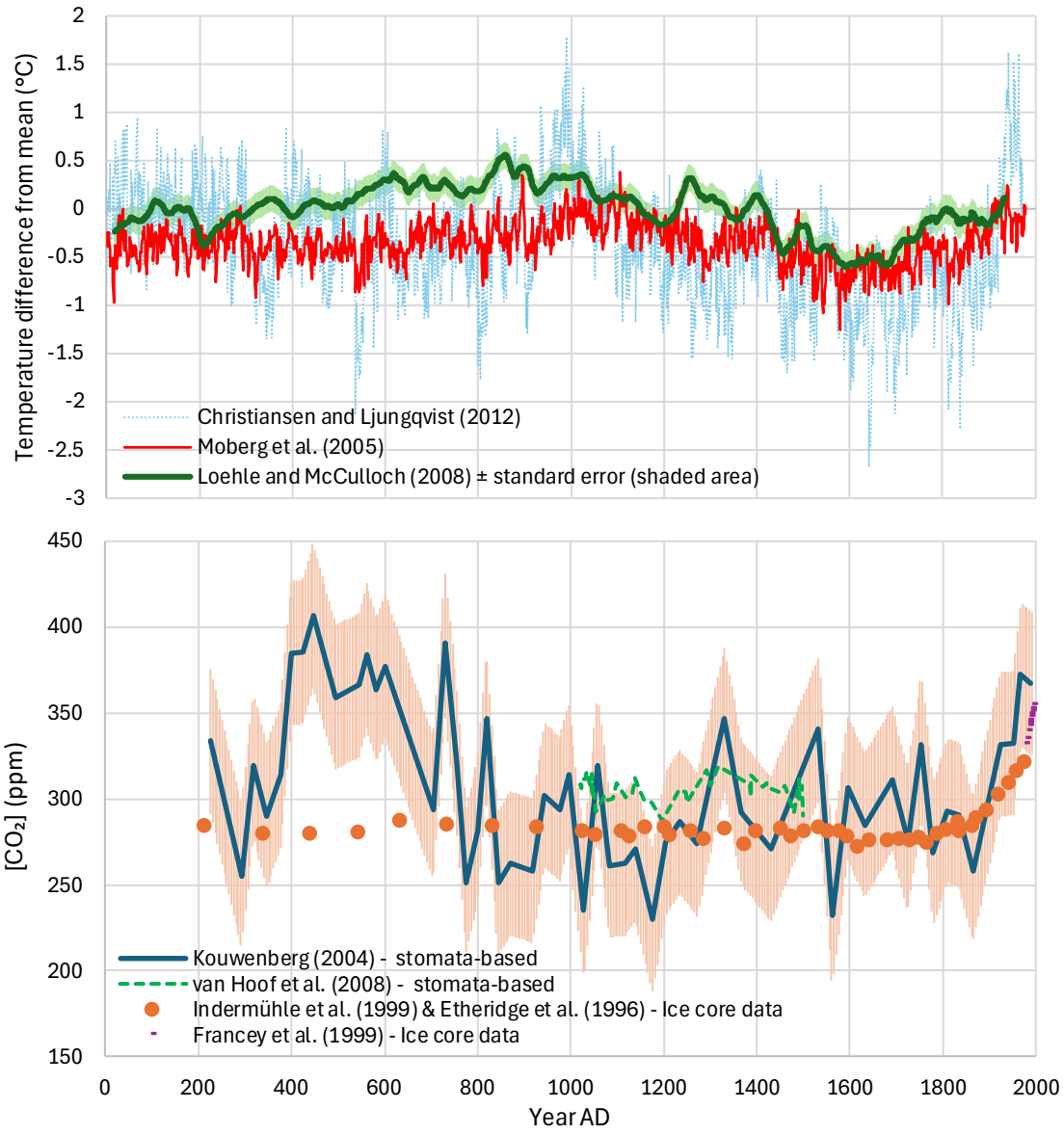

Figure 5 (**upper**) Temperature reconstructions of the last two millennia from the indicated studies. All series are available online in tabulated form and were retrieved by the following sources: Moberg et al. from [57] (Supplementary Information); Loehle and McCulloch [58] from https://www.asc.ohio-state.edu/mcculloch.2/AGW/Loehle/LoehleMcC.txt, and Christiansen and Ljungqvist [59], from [75]. (**lower**) Reconstruction of paleo-atmospheric CO_2 levels based on stomatal frequency analysis and ice cores. The stomata-based data by Kouwenberg [64] (thick blue line) are extracted by digitizing figure 6.3 of that study and are checked to be in agreement (with small differences in dating after 1500 AD) with Figure 3 of Kouwenberg et al. [65]; the standard errors (shaded orange area) are extracted by digitizing figure 5.4 of [64]. The stomata-based data by van Hoof et al. were taken from Table S1 of [66]. The ice core data of Indermühle et al. [60] and Etheridge et al. [61] were digitized from figure 5.4 of [64]. The ice core data by Francey et al. (post 1976, not contained in [64]) are taken form [62] (table 2). The ice core data were also cross-checked with figure 4 of Böhm et al. [63].

2.5 Modern Period of instrumental data

Systematic instrumental observations of hydrometeorological processes, which apparently provide the highest quality data, start from the 19th century. Based on them, several series of globally averaged temperature from instrumental measurements have been compiled. However, systematic measurements of [CO_2] are available only from the late 1950s. Therefore causality analysis based on those data cannot start before that time.

Several temperature time series, from ground stations, reanalyses and satellites have been used in predecessor related studies [20,22,23]. Here we use gridded reanalysis data, namely those of the ERA5 reanalysis [76]. This is the fifth-generation atmospheric reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF), where the name ERA refers to *ECMWF ReAnalysis*. The observation period spans from 1940 onward, with daily updates continuing forward in time. The fields are available at a horizontal resolution of 31 km on 139 levels, from the surface up to 0.01 hPa (around 80 km). It has been produced as an operational service and its fields compare well with the ECMWF operational analyses. ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables. ERA5 combines vast amounts of historical observations into global estimates using advanced modelling and data assimilation systems. The data are available for the period 1950-now at a spatial resolution of 0.5° globally.

For CO_2 concentration, here we use monthly data for two stations, compiled by the US National Oceanic and Atmospheric Administration (NOAA) and the Scripps CO_2 Program of the Scripps Institution of Oceanography, University of California [77]. These are the Mauna Loa Observatory, Hawaii (19.5° N, 155.6° W, 3397 m a.s.l.) and South Pole (90.0° S, 2810 m a.s.l.). The former we pair with the globally averaged temperature and the latter with averaged temperature both over the globe and over the southern hemisphere. The time series used in our analyses are depicted in Figure 6.

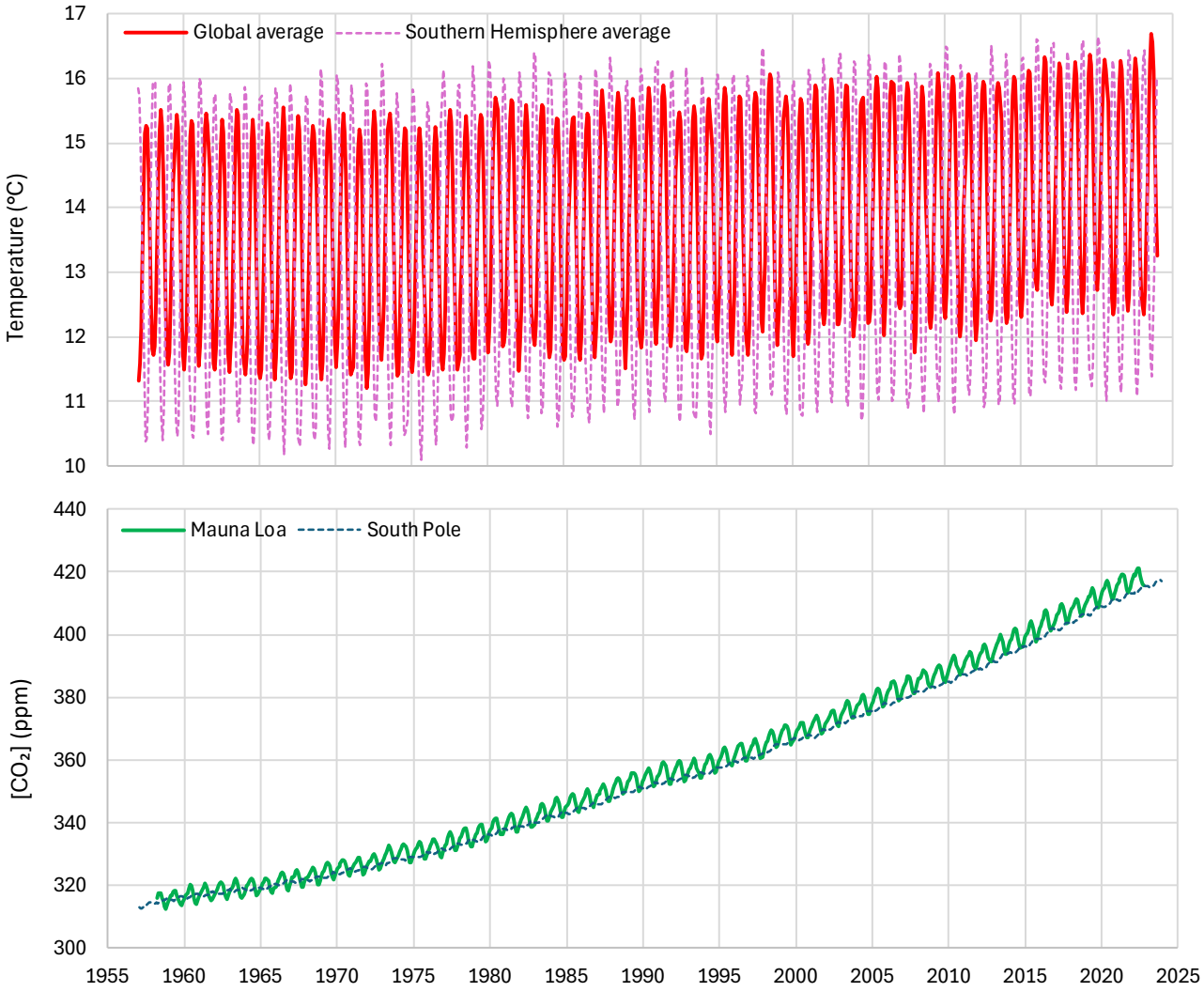

Figure 6 (**upper**) Temperature and [CO_2] time series based on instrumental observation for the indicated cases. The temperature data from the ERA5 reanalysis were retrieved and spatially aggregated through the ClimExp platform (the Climate Explorer of the Dutch Royal Netherlands Meteorological Institute—KNMI; [78]). The [CO_2] data were retrieved from the Scripps CO_2 Program, which made them available online [79].

2.6 A note on varying time scales

It has been a (misleading) common practice to merge data of different time resolutions, without considering the essential differences among them, which prohibit merger. For instance, we see graphs merging instrumental data at annual or monthly time scale with paleoclimatic data at a timescale of hundreds or thousands of years. This has been practiced by most political organizations (e.g. the Word Economic Forum [80]), by news portals (e.g. The Conversation [81]), by scientific organizations (e.g. NASA [82]), and is also sometimes reproduced in scientific literature (e.g. [83]).

To illustrate how problematic the mixing of different time scales is, we need a very large time

series of measurements. Such time series are available from laboratory experiments of turbulent velocity, measured at a high frequency, so that very large samples be formed. Here we use grid data of nearly isotropic turbulence from the Corrsin Wind Tunnel at a high-Reynolds-number [84], which were made available on the Internet by the Johns Hopkins University [85]. This rich dataset consists of 40 time series with 36×10^6 data points of wind velocity along the flow direction and an equal number of time series of cross-stream velocity, all measured at a frequency of 40 kHz (sampling time interval $D = 25$ μs). Here we use part of the data, namely 1.2×10^6 data values at a single point (the first of the probes) along the flow direction, averaged at $D = 1$ ms, so that each value be the average of 40 original consecutive measurements and the length of the time series be 30 000. Like the temperature and [CO_2] time series, the turbulent velocity time series exhibits long range dependence, else known as Hurst-Kolmogorov (stochastic) dynamics [86-88].

The illustration is seen in Figure 7, where six time scales have been used, from 1 to 1024 ms, along with two methods of scale aggregation, sampling (i.e. taking the original series value at the time instance corresponding to the coarser scale) and averaging (i.e. taking the average of original series values at the coarser scale). As seen in the figure, the coarser scale series do not correspond to the one at the smallest scale. Even in the case of sampling, the variability of the original series is lost, as the highest (and less frequent) values are more likely to disappear due to coarse scale sampling. This worsens in the case of averaging, in which the shape of the coarse-scale series tends to be flat. Some small variability continues to appear at the coarse scale, but this is purely due to the Hurst-Kolmogorov dynamics.

The situation in climate proxies resembles the case of averaging, because of the diffusion that occurs, particularly when the proxies originate from ice cores [89]. In this case, the situation can be even worse than seen in Figure 7, as the diffusion time scale could be even coarser than the nominal time averaging time scale. For these reasons, in our analyses that follow we do not merge time series of different scales originating from different construction methods.

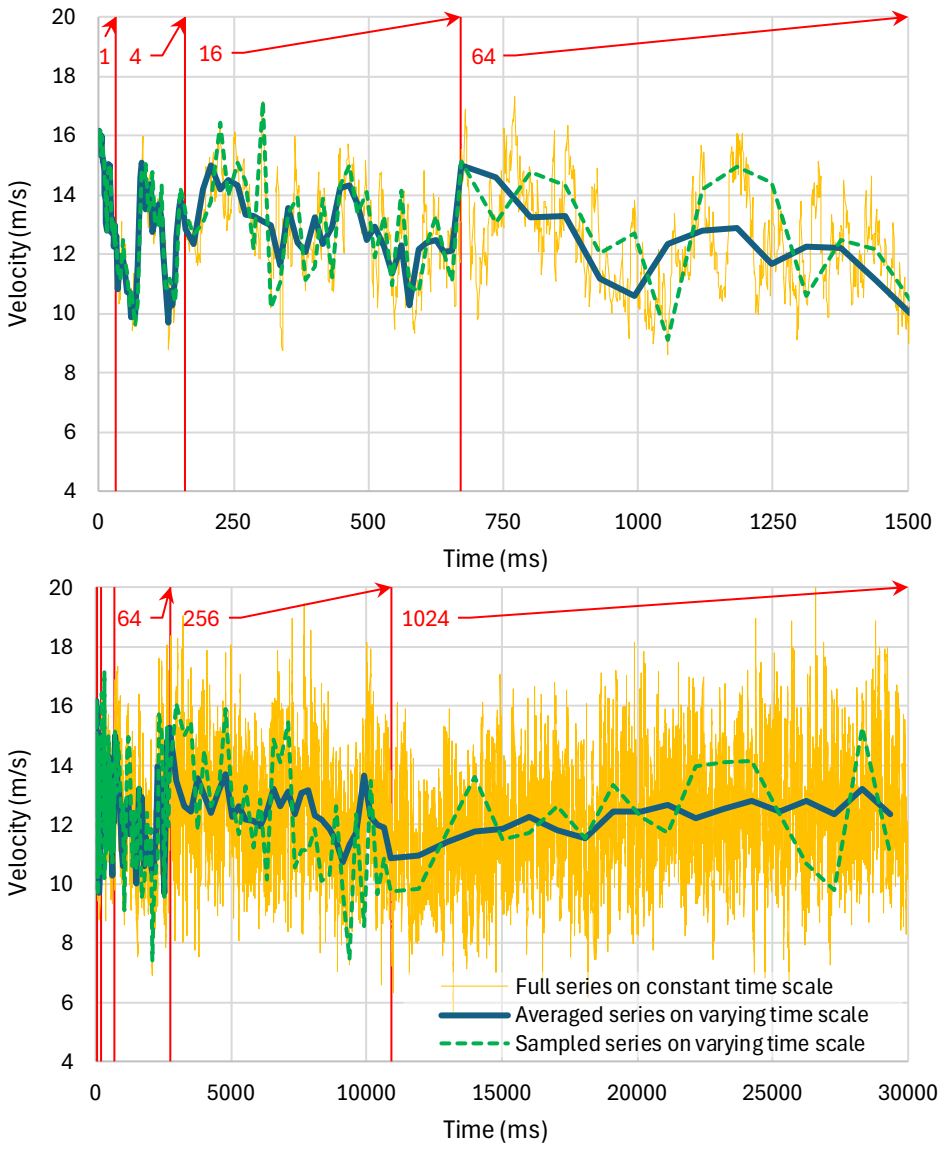

Figure 7 A long turbulent velocity time series, plotted for varying time scales, from 1 to 1024 ms, with each increased one being a quadruple of the immediate smaller time scale; (**upper**) the first 1500 terms; (**lower**) 30 000 terms.

3 Methodology

3.1 The inappropriateness of deterministic methodologies

Commonly, the methods to identify time precedence between two related processes, and thus infer the causality direction, are deterministic. As an example, Shakun et al. [26] and Humlum et al. [34] use the terms "phase" or "phasing" as if the two processes (T and [CO_2]) were harmonic oscillators, while clearly their evolution is irregular, only describable as stochastic processes with complex structure, including Hurst-Kolmogorov dynamics. Using Monte Carlo simulation (as Shakun et al. did), e.g. for testing statistical significance of lags, does not make the methodology non-deterministic, once its

fundamental characteristic in identifying these lags is based on peaks, troughs, or trends in time series.

To show the inappropriateness of the deterministic setting, we use two independent synthetic time series, x_τ and v_τ of 1000 terms (plus some necessary redundant terms) from the Hurst-Kolmogorov model [88] with Hurst coefficient $H = 0.95$. Then we calculate a third time series y_τ as follows:

$$y_\tau = \sum_{j=0}^{J} g_j x_{\tau-j} + v_\tau \tag{1}$$

where

$$J = 30, \qquad g_j = \begin{cases} \dfrac{j}{10}, & j \le 10 \\ \dfrac{30-j}{20}, & j \ge 10 \end{cases} \tag{2}$$

Clearly, the system (x_τ, y_τ) is causal with causality direction $x \to y$, as y_τ depends only on past terms of x_τ. This will be explained better in Section 3.2, while a graphical depiction of g_j will also be given in that section (Figure 9).

However, by merely inspecting the time series following a deterministic approach that looks for "phasing" in time series through trends, peaks and troughs, is misleading and inappropriate. This is illustrated in Figure 8. It is more likely to see sequences like the ones shown in the left column of Figure 8, which correctly reflect the true causality. However, sequences like the ones shown in the right column, which indicate reverse causality, are possible. This does not mean that the causality was indeed reversed: all terms of the time series were generated by the same causal system $x_\tau \to y_\tau$. It just means that irregular time series cannot be treated deterministically. At the same time, the example shows that with a proper choice of a particular time frame, such as that in the right panel of Figure 8, and with deterministic thinking helped by good imagination, one can draw conclusions opposite to reality.

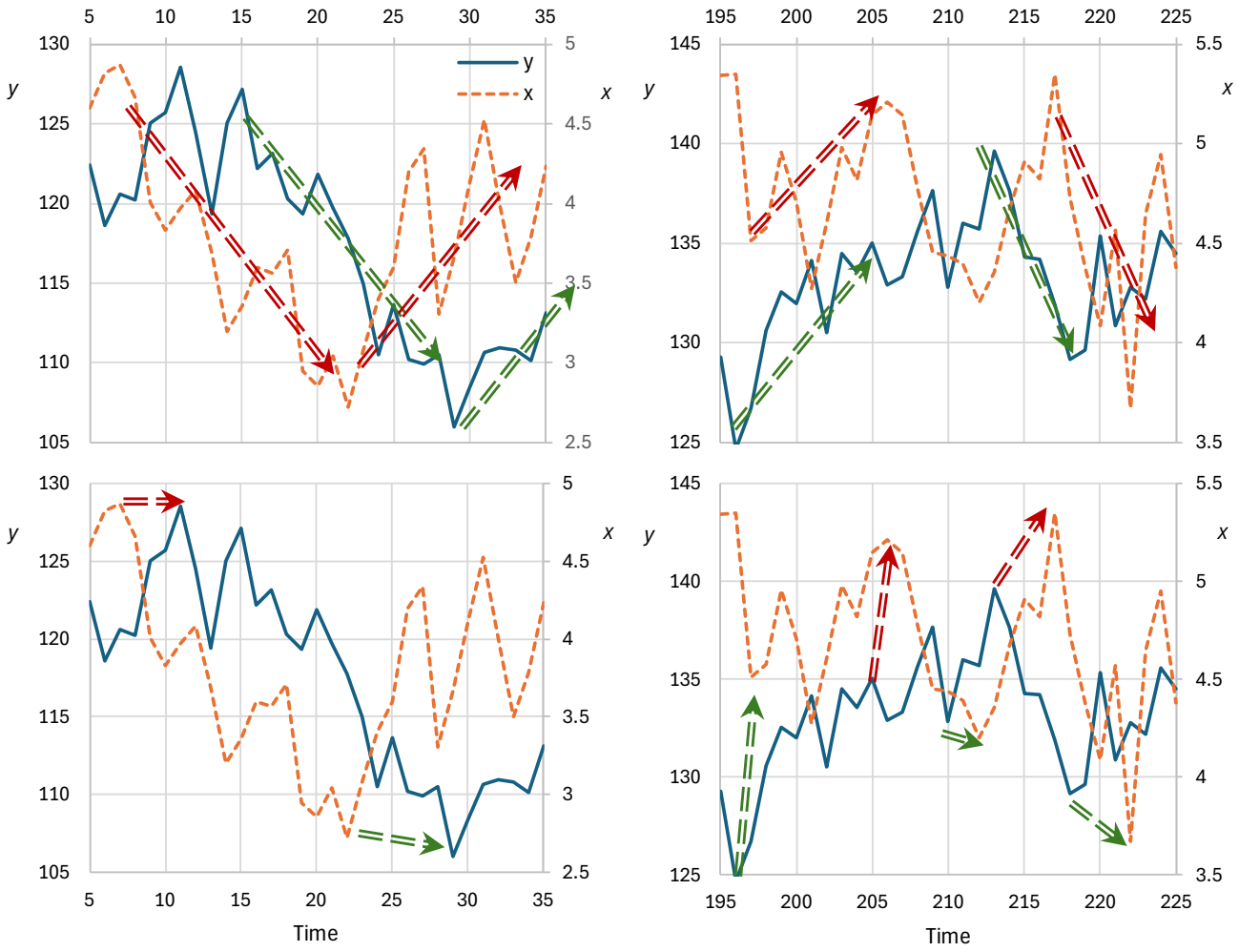

Figure 8 Two groups of consecutive values of time series from a causal system $x_\tau \rightarrow y_\tau$, each consisting of 30 items, for illustrating the inappropriateness of a deterministic approach that looks for "phasing" in time series through (**upper**) trends and (**lower**) peaks and troughs. (**left column**) A group of 30 terms preserving the correct time precedence of x_τ over y_τ; (**right column**) a group of 30 terms incorrectly suggesting the reverse time precedence.

3.2 Stochastic methodology

The stochastic methodology used here for identifying potential causal links was developed in [2,22,23] and is based on the impulse response function (IRF) between two stochastic processes $\underline{x}(t), \underline{y}(t)$, denoted as $g(h)$ where h denotes time lag, based on the convolution:

$$\underline{y}(t) = \int_{-\infty}^{\infty} g(h)\underline{x}(t-h)\mathrm{d}h + \underline{v}(t) \tag{3}$$

where $\underline{v}(t)$ is another stochastic process representing the part that is not explained by the causal link. Notice that we use the Dutch notational convention, in which stochastic variables and processes are

underlined, while common variables and functions are not.

To see that the function $g(h)$ is the impulse response function (IRF) of the system $(\underline{x}(t), \underline{y}(t))$, we set $\underline{v}(t) \equiv 0$ and $\underline{x}(t) = \delta(t)$ (the Dirac delta function, representing an impulse of infinite amplitude at $t = 0$ and attaining the value of 0 for $t \neq 0$), and we readily get $\underline{y}(t) = g(t)$. On the other hand, if we set $g(h) = b\ \delta(h - h_0)$ (with constant b and h_0), which means that the IRF is zero for every lag except for the specific lag h_0, then equation (1) becomes $\underline{y}(t) = b\underline{x}(t - h_0) + \underline{v}(t)$. This special case is equivalent to simply correlating $\underline{y}(t)$ with $\underline{x}(t - h_0)$ at any time instance t. It is easy to find (cf. linear regression) that in this case the multiplicative constant b is the correlation coefficient of $\underline{y}(t)$ and $\underline{x}(t - h_0)$ multiplied by the ratio of the standard deviations of the two processes. In general, however, we expect that the actual $g(h)$ is not a Dirac delta function but a continuous one over some domain. Thus, the IRF is a much more powerful tool than correlation, as it integrates the correlations in the entire spectrum of lags.

In applications, the continuous time representation is replaced by a discrete time one, with a time step Δ, the IRF becomes a sequence of values g_j, where j denotes the time lag, the infinite range of the time lag h becomes a finite window of time lag j, specified in the interval $[L, U]$ (the lower and upper computational lag, respectively), the integrals are replaced by sums, and the true values of statistics are replaced by estimates. Specifically, the discrete time version is

$$\underline{y}_\tau = \sum_{j=L}^{U} g_j \underline{x}_{\tau - j} + \underline{v}_\tau \tag{4}$$

and the conversion from continuous to discrete time is described in [2].

For any two processes $\underline{x}_\tau$ and $\underline{y}_\tau$, Equation (4) has infinitely many solutions in terms of the sequence g_j, and the process $\underline{v}_\tau$. An obvious and trivial one is $g_j \equiv 0, \underline{y}_\tau \equiv \underline{v}_\tau$. The sought solution is the one that corresponds to the minimum variance of $\underline{v}_\tau$ called the least-squares solution. Equivalently, this solution maximizes the explained variance ratio:

$$e := 1 - \frac{\gamma_v}{\gamma_y} \tag{5}$$

where γ_v and γ_y denote the variances of the processes $\underline{v}_\tau$ and $\underline{y}_\tau$, respectively. (This is similar as in correlation at a single time lag.) If the attained maximum e is close to zero, this will mean that the two processes are uncorrelated and thus no causality condition can exist between them (non-causal system). Otherwise, we may assume, without loss of generality, that processes $\underline{x}_\tau$ and $\underline{y}_\tau$ are positively

correlated, i.e., an increase in $\underline{x}_\tau$ would result in an increase in $\underline{y}_\tau$. In the opposite case (if the processes are negatively correlated) we multiply one of the two by -1. Therefore, we impose a nonnegativity constraint for the sought IRF,

$$g_j \geq 0 \tag{6}$$

In the estimation of IRF, we may also impose a roughness constraint,

$$E \leq E_0 \tag{7}$$

where E is the roughness of the IRF determined in terms of the second difference of g_j, $\Delta^2 g_j$:

$$E := \sum_{L}^{U-2} \Delta^2 g_j\,, \qquad \Delta^2 g_j = g_{j+2} - 2g_{j+1} + g_j \tag{8}$$

Further justification for the two constraints is provided in [2]. The roughness E can be standardized as

$$\varepsilon := \frac{E}{8\sum_{j=-J}^{J} g_j^2} \tag{9}$$

where ε ranges in (0,1) for nonnegative g_j. The determination of the IRF ordinates g_j is thus formulated as a constrained optimization problem, whose numerical solution is always possible, simple and fast, and can be attained even by commonly available solvers, e.g. in commercial or open source spreadsheet software. By choosing a proper multiplier $\lambda > 0$ (which can be adjusted in a trial-and-error fashion, until an acceptable smoothness is attained), we can simplify the optimization problem to:

$$\begin{aligned} &\text{find } g_j, j = L, \dots U \text{ that} \\ &\text{maximize } e(g_j) - \lambda\, \varepsilon(g_j) \\ &\text{subject to } g_j \geq 0 \end{aligned} \tag{10}$$

The total number of unknowns g_j is $N := U - L + 1$, where we assume that $L \leq 0 \leq U$. Depending on the results of the estimation procedure, if e is non-negligible, the system is deemed:

- potentially HOE causal if we have $g_j > 0$ for both some positive and some negative lags j,
- potentially causal if $g_j = 0$ for all $j < 0$, and
- potentially anticausal if $g_j = 0$ for all $j > 0$.

These three cases are graphically illustrated in Figure 9. The adverb “potentially” in the above characterizations highlights the fact that the conditions tested provide necessary but not sufficient conditions for causality.

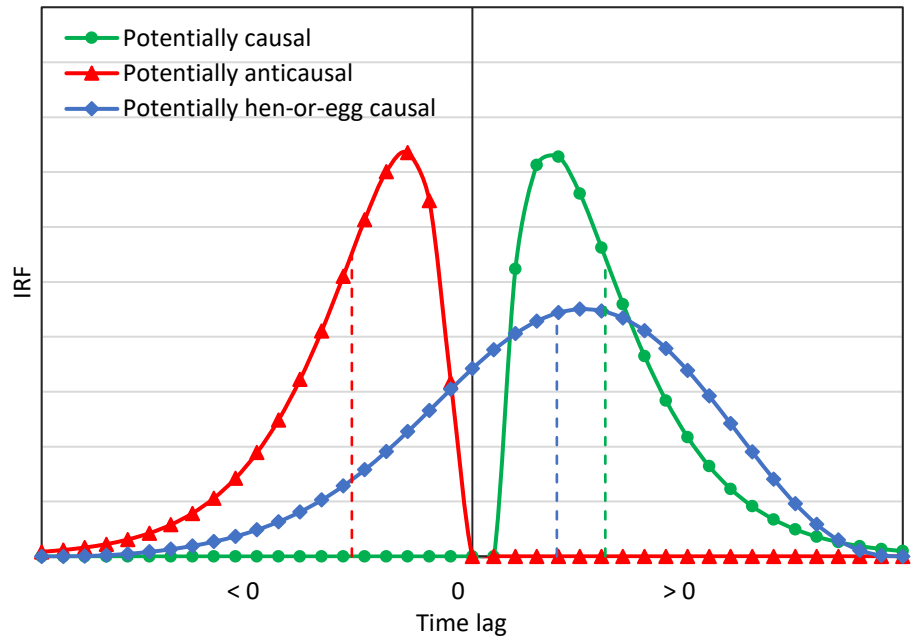

Figure 9. Explanatory sketch for the definition of the different potential causality types. For each graph the mean μ_h is also plotted with dashed line. (Reproduced from [23] licensed under Creative Commons Attribution).

In a potentially causal (or anticausal) system, the time order is explicitly reflected in the above characterizations. In a potentially HOE causal system the time order needs to be clarified by defining the principal direction. The most natural indices for this are:

1. The time lag $h = h_c$ maximizing the (absolute value of) cross-covariance $c_{yx}(j)$.
2. The mean (time average) of the sequence g_j, defined as:

$$\mu_h := \frac{1}{H_0}\sum_{j=L}^{U} j g_j\,, \qquad H_0 := \sum_{j=L}^{U} g_j \tag{11}$$

3. The median $h_{1/2}$ of the sequence g_j, implicitly defined by:

$$\frac{1}{H_0}\sum_{j=L}^{h_{1/2}} g_j = \frac{1}{2} \tag{12}$$

where linear interpolation between consecutive j may be required to specify $h_{1/2}$.

We note that h_c, is completely independent from g_j. The other two, μ_h and $h_{1/2}$, depend on the sequence g_j that is determined. However, extensive analyses in [21] showed that their estimation is quite robust; for example, the use of the roughness constraint, while affecting the resulting sequence g_j, has little impact on the values of μ_h and $h_{1/2}$. Usually, the characteristic lags μ_h and $h_{1/2}$ do not differ substantially with each other and any of them could be chosen for further use. Here we

prefer to note both, as well as h_c, as they all provide useful information about the relationship of the two processes (like in the case of using mean,median and mode, in the characterization of the probability distribution of a single variable).

In estimating the sequence g_j, parsimony requires that the number of unknowns *N* be as small as possible. On the other hand, a small *N* may not be appropriate to properly reconstruct the IRF. To resolve this conflict, we may keep using a small *N* and extend the time length of the IRF by using an integer multiplier $m \geq 1$. In this case we modify Equation (4) as follows:

$$\underline{y_\tau} = \sum_{j=mL}^{mU} g'_j \underline{x}_{\tau-j} + \underline{v_\tau} \tag{13}$$

where g'_j is given by one of the two alternatives:

$$g'_j := g_{\lfloor j/m \rceil}, \qquad g'_j := \{j/m\} g_{\lceil j/m \rceil} + (1 - \{j/m\}) g_{\lfloor j/m \rfloor} \tag{14}$$

with $\lfloor a \rceil, \lceil a \rceil, \lfloor a \rfloor, \{a\}$ denoting, respectively, the standard rounding function, the ceiling, the floor and the fractional part of the real number a. In the first option we round the number j/m and assign g'_j as the g_i for $i = \lfloor j/m \rceil$. In the second option we perform linear interpolation between two g_i values above and below j/m. Again, the multiplier *m* should not be large, otherwise the number of values of $\underline{y_\tau}$ and $\underline{v_\tau}$ remaining to estimate their variances becomes too low. (For example, it can be verified from Equation (13) that if $m > n/U$, where *n* is the length of the available time series, then no item y_τ can be estimated.)

Furthermore, the time scale of analysis, *k*, should not necessarily be identical to the computational time step, Δ. The ratio $l := \Delta/k$ is thus another multiplier referred to the time series, instead of the IRF. For example, if the data values are given on a monthly time step, but the analysis is made for the annual scale, then $l = 1/12$.

In all applications here, we use $N = 21, L = -20, \ldots 0, m \in \{1,4\}, l \leq 1$, where the exact value of l depends on time scale of analysis and the discretization step of the data. When $L = 0$, we assume a potentially causal system as all g_i for $i < 0$ are zero. In contrast, when $L = -20$, we assume a potentially anticausal system. In all other values of *L* we have a potentially HOE causal system. By sliding the window of length $N = 21$ with the lower computational lag *L* moving from –20 to 0, we determine the explained variance in each of the cases and finally we keep as final *L* that which yields the highest explained variance. If that is $L = 0$, the system is deemed potentially causal, if $L = -20$, it is deemed potentially anticausal, and in all other cases it is deemed potentially HOE causal.

We further note that, given the two processes, each of the directions $x \to y$ and $y \to x$ can be investigated separately as there is no symmetry (or antisymmetry) in the produced IRFs in the two directions. When we refer to direction $y \to x$ we mean that we interchange the time series x and y and still estimate the IRF in the same way, as described in our equations in which the direction $x \to y$ is assumed.

As detailed in [22,23], there exist a variety of other methods for estimating IRF and for inferring causality but our method differs conceptually and computationally from them, including from the so-called "Granger causality" [90,91] and the framework proposed by Pearl and collaborators [92-94].

3.3 A synthetic example for assessing the behavior of the method

As an illustration of the method, we use the synthetic example already introduced in Section 3.1, but now using the stochastic method instead of the inconsistent deterministic approach. We note that the true system is causal and the true IRF function extends to lags up to 30, but here we try to recover it with only 21 time lags. The results of the method are shown in Figure 10, for both directions $x \to y$ (the correct one, with the true IRF of a triangular shape shown in the upper left panel of Figure 10) and $y \to x$ (the incorrect one).

For the system $x \to y$ and lower computational lag $L = 0$ (the true one), the method produces a good approximation of the true IRF, but the exact recovering is impossible as the time coverage (20 time steps) is smaller than the true one (30 time steps). The higher estimated g_j ordinates, in comparison to the true ones are due to the reduction of the true time coverage. As seen in the upper right panel, for $L = 0$ the method approximates well the true characteristic lags, mean (μ_h) and median ($h_{1/2}$). The cross-correlation maximizing lag and is not well reproduced, but this is independent of the produced IRF and is related to the high Hurst parameter, which creates high statistical uncertainty. The explained variance for $L = 0$ is very high, about 0.7.

As we decrease the lower computational lag L, the shape of the IRF no longer corresponds to the true one but assumes a high value at the rightmost end. This is expected, because it has to substitute the limited information due to the decrease of the time coverage. Expectedly, the explained variance is reduced, and so are the characteristic time lags. At the lowest computational lag $L = -20$ the performance becomes bad in all aspects, as it should. As a result of the application, we correctly recover the fact that the system we deal with is causal and we keep as final IRF that of the causal system, with $L = 0$.

Figure 10 also shows the case where we incorrectly assumed that the causality direction was opposite, $y \to x$. In this case, all evidence depicted in the lower row of Figure 10, whose behavior is opposite to that of the upper row, is convincing that that system ($y \to x$) is anticausal, as actually is.

One may wonder why even for bad choices, e.g. for $L = -20$ in the direction $x \to y$ or in any L in the direction $y \to x$, we have some positive explained variance. The answer to this question is not difficult. It is the high autocorrelation, which results in some explained variance, as any value in a time series contains some information about its close neighbors due to time dependence.

A further behavior of IRFs, seen Figure 10, that needs to be discussed is the appearance of high tails on the left or the right end of IRFs. Such tails are inconsistent with the (a priori known in this case) true shape of the IRF. Their appearance is an indication that the assumed time frame of the IRF is narrower than required for its faithful representation. Another factor that we should consider in the

explanation of the high tails is the fact that the second derivative is not defined in the end points of the IRF and hence they do not contribute to the roughness, which in the other points is constrained.

Overall, the example provided that the method has excellent performance and a reasonable behavior, as it provides information on how successful the potential causality identification is and what the potential caveats can be.

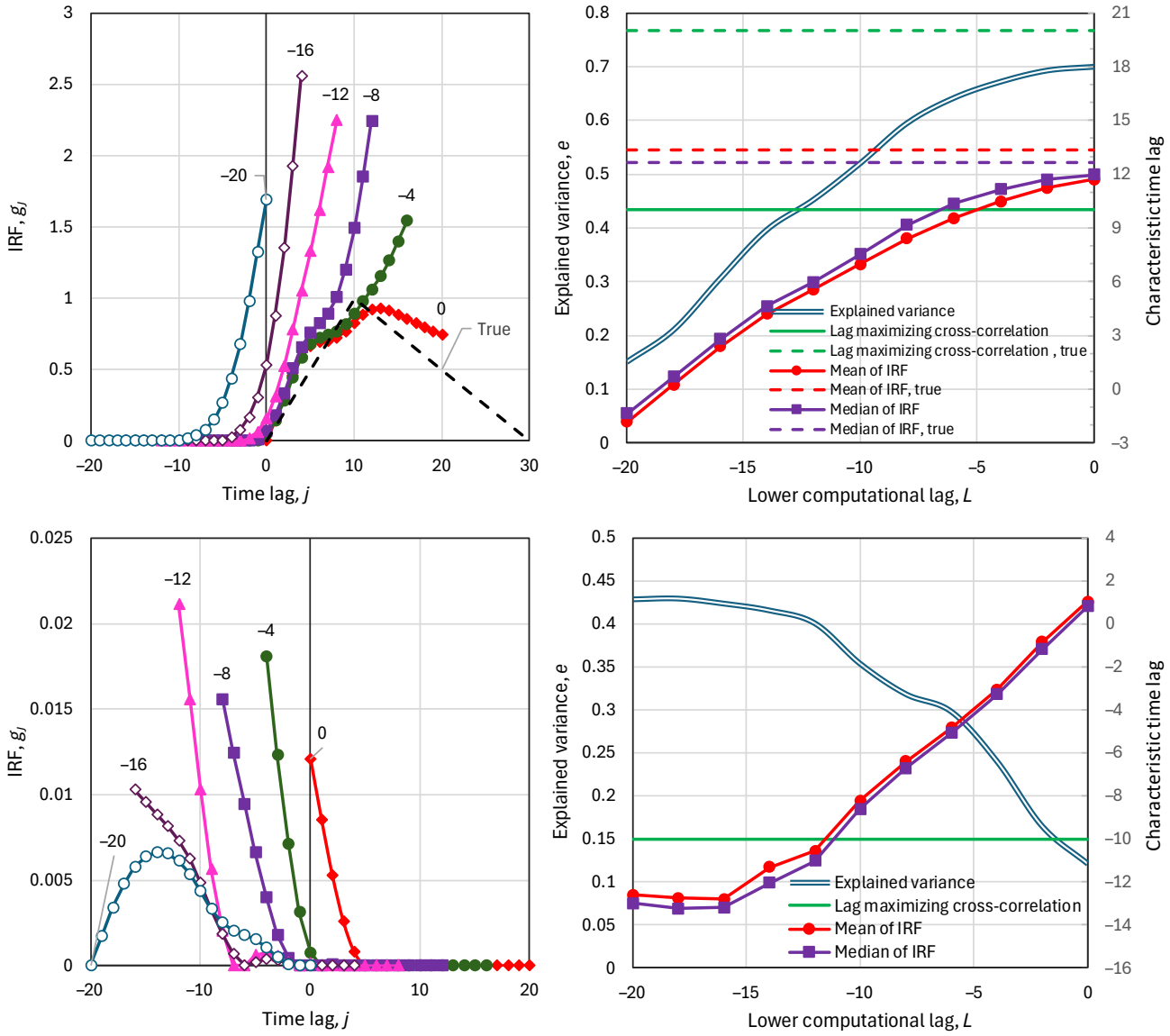

Figure 10 An example of the application of the methodology with a couple of related time series exhibiting Hurst-Kolmogorov dynamics, generated from a causal system $x \rightarrow y$ with true IRF of a triangular shape shown in the upper left panel. (**upper row**) Reconstruction from the data of the system $x \rightarrow y$; (**lower row**) application of the method for the reverse system $y \rightarrow x$. (**left column**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right column**) explained variance and characteristic time lags (mean, median and cross-correlation maximizing) as functions of the lower computational lag, L; the true characteristic time lags are also plotted in the upper-left panel.

4 Results

4.1 General notes for the real-world applications

In Section 4 we present the application of the method to the real-world cases for identifying causality between temperature and [CO_2] using the data described in Section 2. In cases where the time series were given in an irregular time step, this has to be regularized before the application, which is done by linear interpolation. The appropriate time steps have already been discussed in Section 2. As explained in earlier studies [20,22,23], before the application, the [CO_2] time series has to be transformed by taking the logarithms. In several initial applications we tried both directions, $T \rightarrow [CO_2]$ and $[CO_2] \rightarrow T$ but we present the former only, as the latter did not provide useful results. We also present results in which the maximum achieved explained variance is higher than 0.1, as those with lower than this are deemed insignificant. We also omit applications in which the cross-correlations between T and $[CO_2]$ are negative, even for small lags, as these are interpreted as suggesting time series inconsistent to each other. Yet in all families of data presented in Section 2, there was found at least one pair of time series that yielded positive cross-correlations and allowed application of the method.

We present the results graphically for the causality direction $T \rightarrow [CO_2]$ and we occasionally refer to those of the direction $[CO_2] \rightarrow T$. In all cases, we apply the method with gradually increasing lower computational lag L from –20 to 0 and we locate the particular value that yields the highest explained variance, e. For a fair comparison we use the same time frame for all values of L. This means that, if the time spans from 1 to n, the comparison is made for the data points in a time frame spanning $20m$ to $n + 1 - 20m$, where m is the time step multiplier (1 or 4 in our case).

One important issue that should be kept in mind is related to the very high autocorrelations, which appear particularly in $[CO_2]$ series. As high autocorrelation increases uncertainty in the long term, this is a major case leading to false identification of potential causality. This problem was discussed in [22] and illustrated in the electronic supplementary material of this publication, along with the technique to handle such situations and avoid false conclusions. Specifically, the appropriate technique is to difference the time series, so as to investigate the changes of the related processes, rather than the processes per se. Differencing reduces the autocorrelations substantially and thus avoids spurious results but has the disadvantage of reducing the explained variance.

Figure 11 shows the empirical autocorrelation functions of the $[CO_2]$ series, original and differenced. In the instrumental series, the autocorrelations are almost 1, even for lags as high as 100. This prohibits any inference from the original series. However, their differenced series have reasonable positive autocorrelations, which make inference possible. The proxy series have high autocorrelations at small lags, but reasonable ones at large lags. For those we examined both the original and the differenced series, provided that the latter are positive.

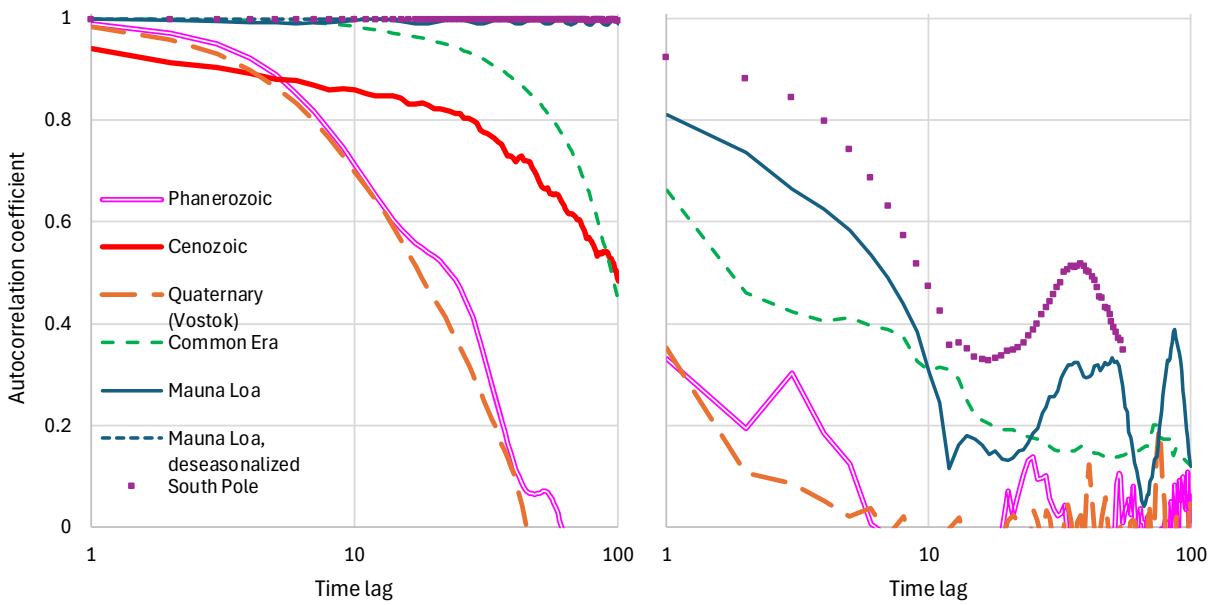

Figure 11 Autocorrelation functions of [CO_2] series: (**left**) original; (**right**) differenced. The differenced series for Cenozoic is not plotted as it is mostly negative.

4.2 Proxy data

For the Phanerozoic series, the pair that gives positive cross-correlations for both the original and the differenced versions consists of the Grossman & Joachimski [36] temperature series and Berner [46] [CO_2] series. The application of the method to these data in the direction $T \rightarrow [CO_2]$, both with original and differenced time series is shown in Figure 12. The original series, while yielding a high explained variance (~0.62), does not help to discern causality as this explained variance is almost independent of the lower lag L. However, the differenced series clearly suggests that we have a potentially causal system, as the explained variance is maximized for $L = 0$, even though it is lower (0.27), as expected. In addition, the more physically consistent shape of the IRF in the case $L = 0$, supports the conclusion of a a potentially causal system. Additional support is provided by the results of the reverse system, $[CO_2] \rightarrow T$, not shown in the figure. In this, the explained variance for the original system is maximized for $L = -20$ with an attained value of 0.55, suggesting an anticausal system. The differenced process results in virtually zero explained variance for all values of L for the system $[CO_2] \rightarrow T$.

It is relevant to note that Davis [45], analyzing different time series, namely those of Prokoph et al. [95] for temperature and Royer [44] for [CO_2], and different methodology, i.e. "detrending" the time series and examining transitions, found different results. Specifically, he found that [CO_2] is correlated weakly but negatively with linearly-detrended T proxies over the last 425 million years. He stated that of 68 correlation coefficients between CO_2 and T proxies encompassing all known major Phanerozoic climate transitions, 77.9% are non-discernible and 60.0% of discernible correlations are negative. In a later publication, Davis [48] also reported discernible and negative correlation ($r = -0.76$) between the same CO_2 and T proxies, but for a small segment of the record, namely the Cretaceous-Paleogene mass extinction period.

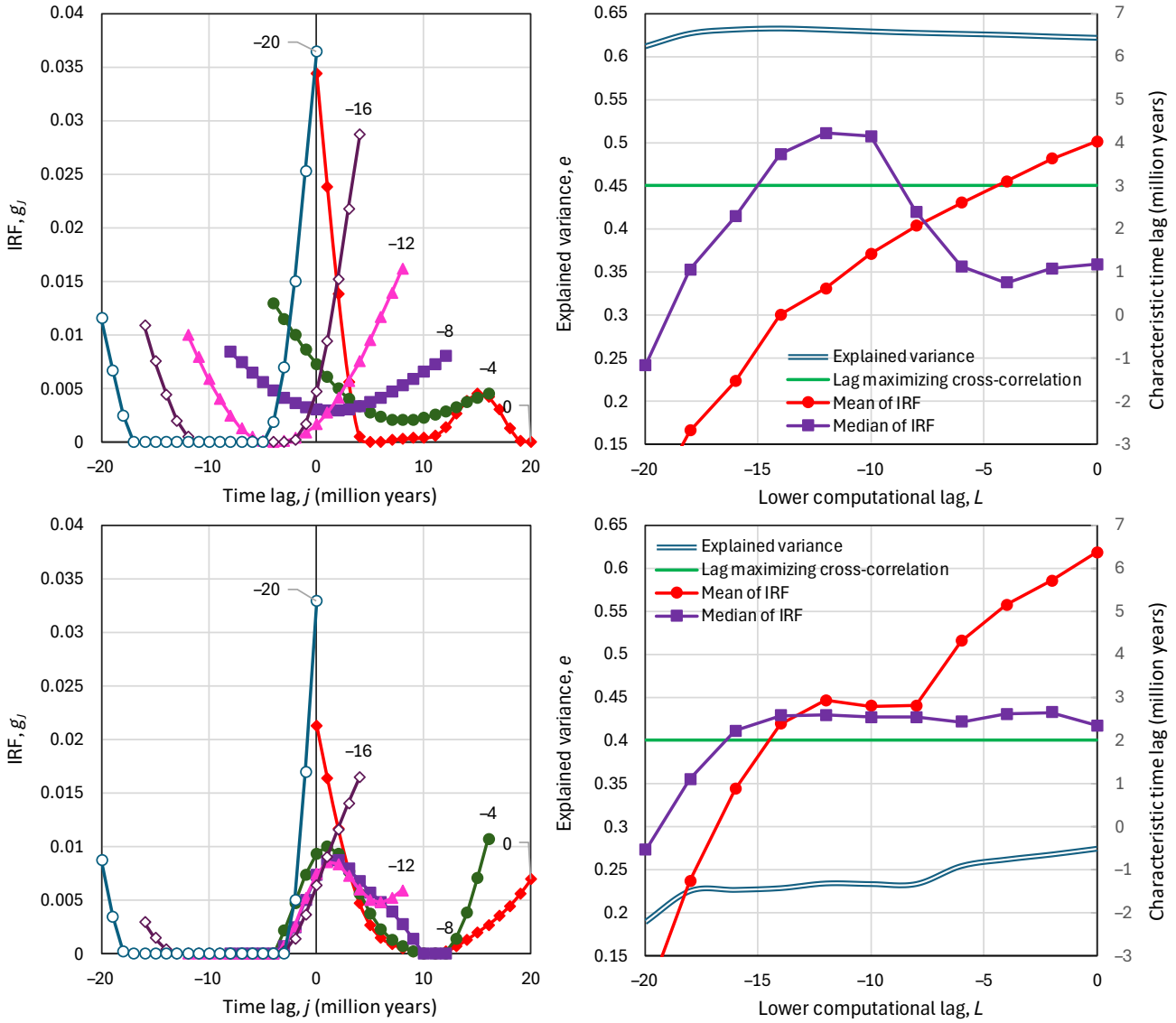

Figure 12 Application of the methodology to the Phanerozoic data (Grossman & Joachimski [36] series for temperature and Berner [46] for [CO_2]). (**upper row**) Original series; (**lower row**) differenced series. (**left column**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right column**) explained variance and characteristic time lags (mean, median and cross-correlation maximizing) as functions of the lower computational lag, L.

The results for the Cenozoic are shown in Figure 13, which refers to the analysis of the original series in the direction $T \rightarrow [CO_2]$, where the values of $-\delta^{18}O$ were used as a proxy of T. The maximum explained variance (~0.80) is attained for the causal system ($L = 0$), which also gives the most physically plausible shape of IRF. The differenced series lead to practically zero explained variance for all L and thus it does not provide any information. In the reverse system, $[CO_2] \rightarrow T$, not shown in the figure, the explained variance is maximized for $L = -20$ with an attained value of 0.88, while for $L = 0$ the explained variance takes the minimum value. This again suggests an anticausal system $[CO_2] \rightarrow T$ or a causal system $T \rightarrow [CO_2]$.

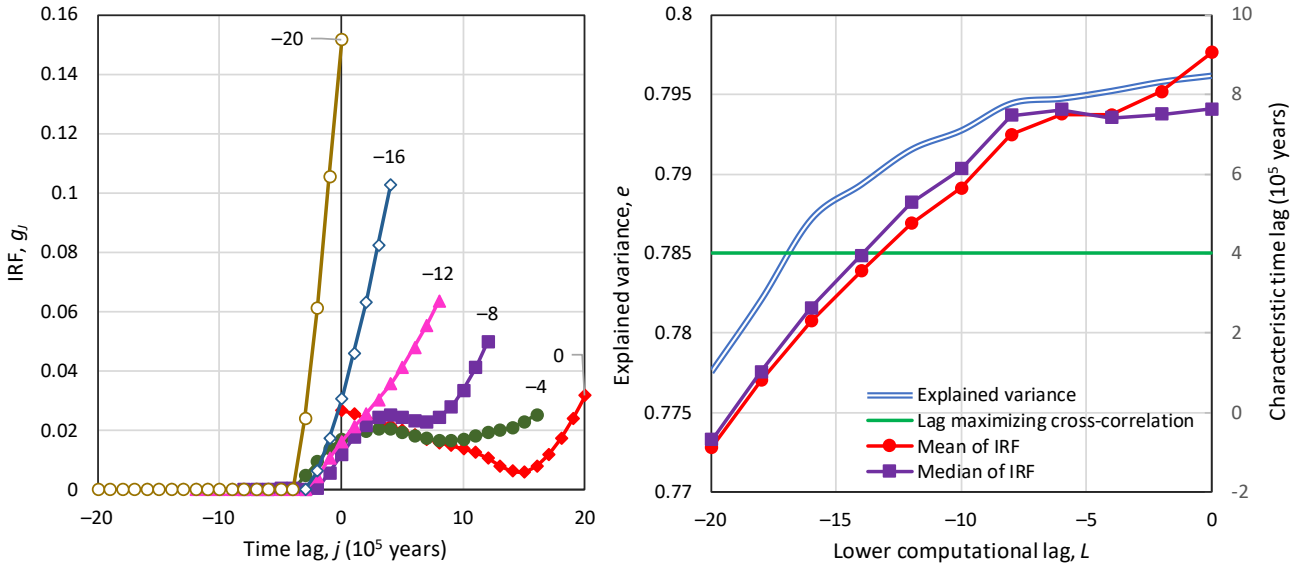

Figure 13 Application of the methodology to the Cenozoic data (original series). (**left**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right**) explained variance and characteristic time lags (mean, median and cross-correlation maximizing) as functions of the lower computational lag, L.

As explained in Section 2.3, for the late Quaternary data, we produced two time series, one with time step of 1000 years extending over 416 000 years and one with time step of 500 years extending over 200 000 years. The results for the first case are shown in Figure 14.

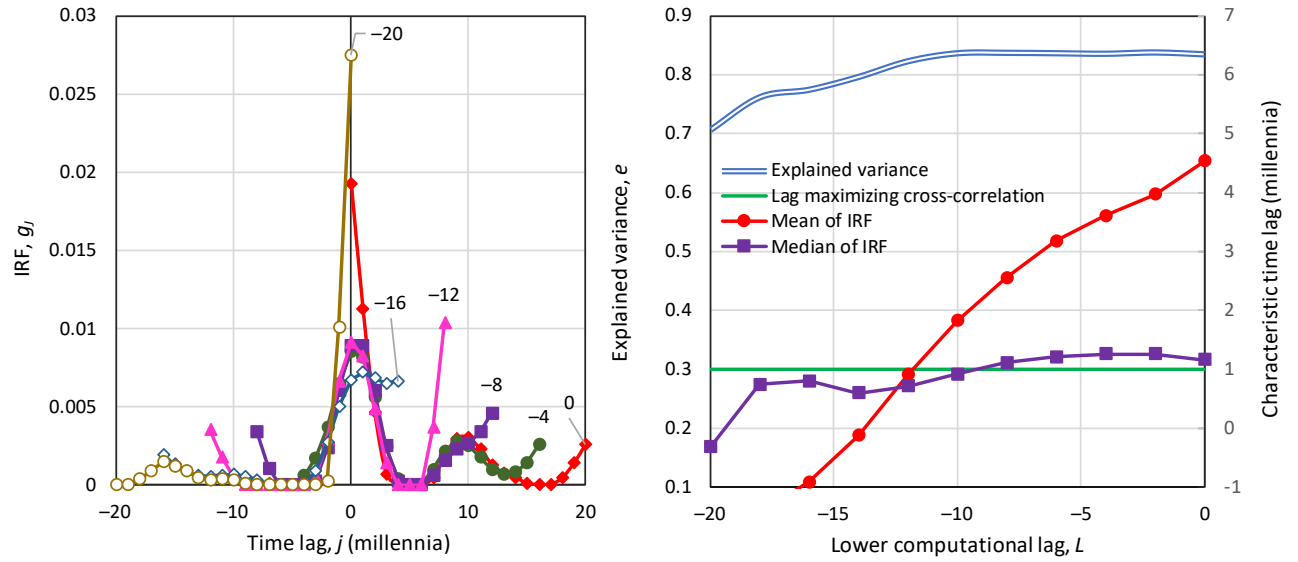

Figure 14 Application of the methodology to the late Quartenary (Vostok) data, namely the original series of 416 000 years with time step of 1000 years; (**left**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right**) explained variance and characteristic time lags as functions of the lower computational lag, L.

The explained variance for $L = 0$ is high, 0.84, and lower for $L = -20$, 0.72. However, the value of 0.84 remains the same for $L = -10$ to 0, which does not provide a clear answer whether the system

is potentially causal or potentially HOE causal. Nonetheless, in either case the characteristic lags are positive, supporting a principal direction $T \rightarrow [CO_2]$ whether this is exclusive (potentially causal system) or not (potentially HOE causal system).

The results for the time series with time step of 500 years are shown in Figure 15, both for the original and the differenced series. The former clearly suggests a potentially causal system, with maximum explained variance at $L = 0$ of $e = 0.89$.

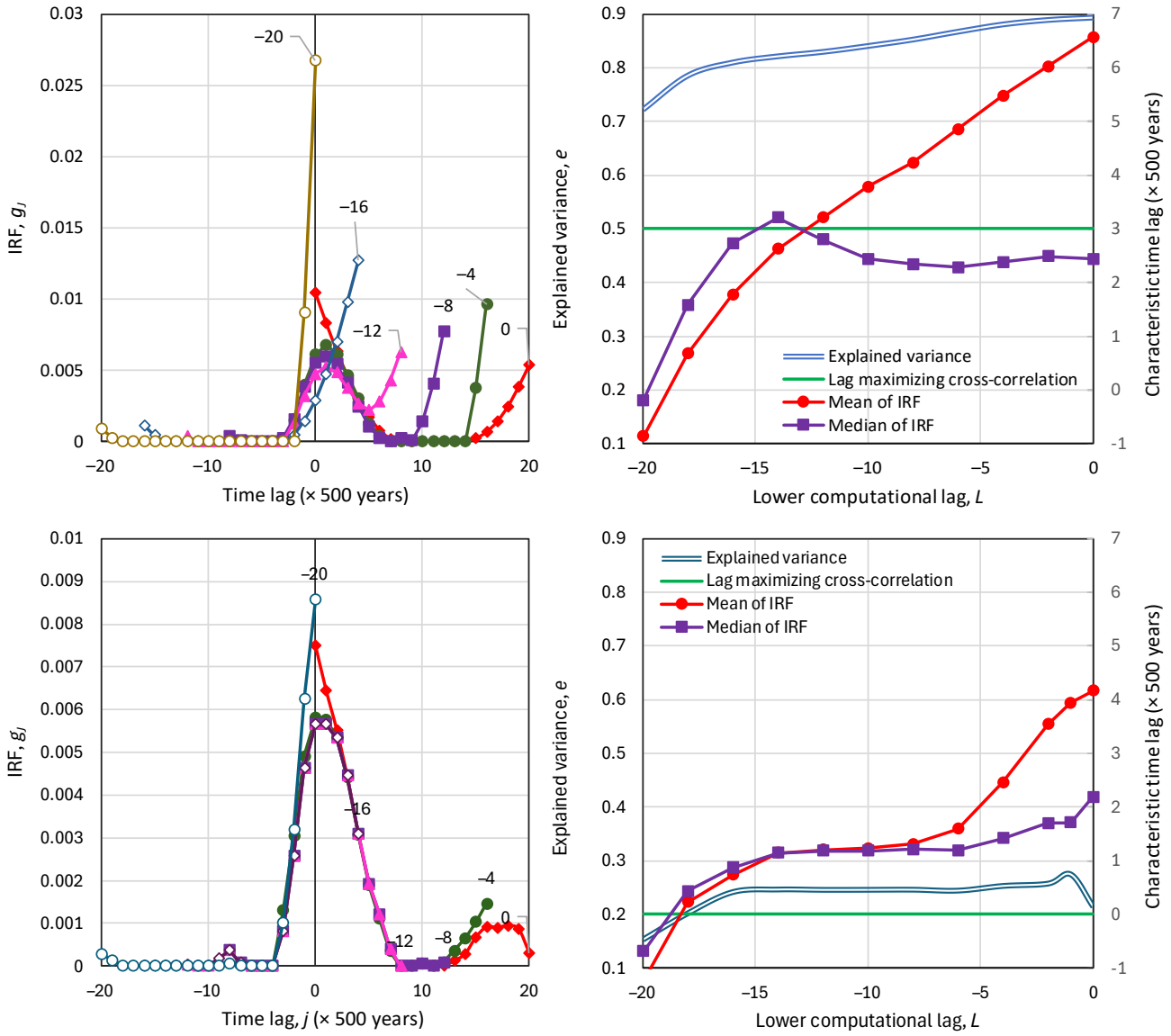

Figure 15 Application of the methodology to the late Quartenary (Vostok) data. (**upper row**) Original series of 200 000 years with time step of 500 years; (**lower row**) differenced series of 200 000 years with time step of 500 years. (**left column**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right column**) explained variance and characteristic time lags as functions of the lower computational lag, L.

The differenced series suggest a nearly causal system, as the lower computational lag maximizing the explained variance is not zero but $L = -1$. However, given the uncertainty in dating, as discussed

in Section 2.3, and the fact that uncertainty is magnified by differencing, we may regard this small displacement from $L = 0$ to $L = -1$ as a statistical error and remain with the conclusion of a potentially causal system.

For the Common Era, preliminary analysis showed that there is one pair of temperature and [CO_2] series that yields cross-correlations sufficient for further analysis. This consists of the Loehle and McCulloch [58] series for temperature and the ice core data of Indermühle et al. [60], Etheridge et al. [61] and Francey et al. [62] for [CO_2]. However, even this pair gives cross-correlations close to zero or negative (and explained variance zero), after differencing at a differencing time step equal to one. Therefore, we only analyze the original series and we use a multiplier of $m = 4$ to expand the time frame of the IRF to 80 years. The results of this analysis, for time scale of 1 year, are shown in Figure 16 and support the case of a potentially causal system at the direction $T \rightarrow [CO_2]$, with an explained variance of 0.49 at $L = 0$.

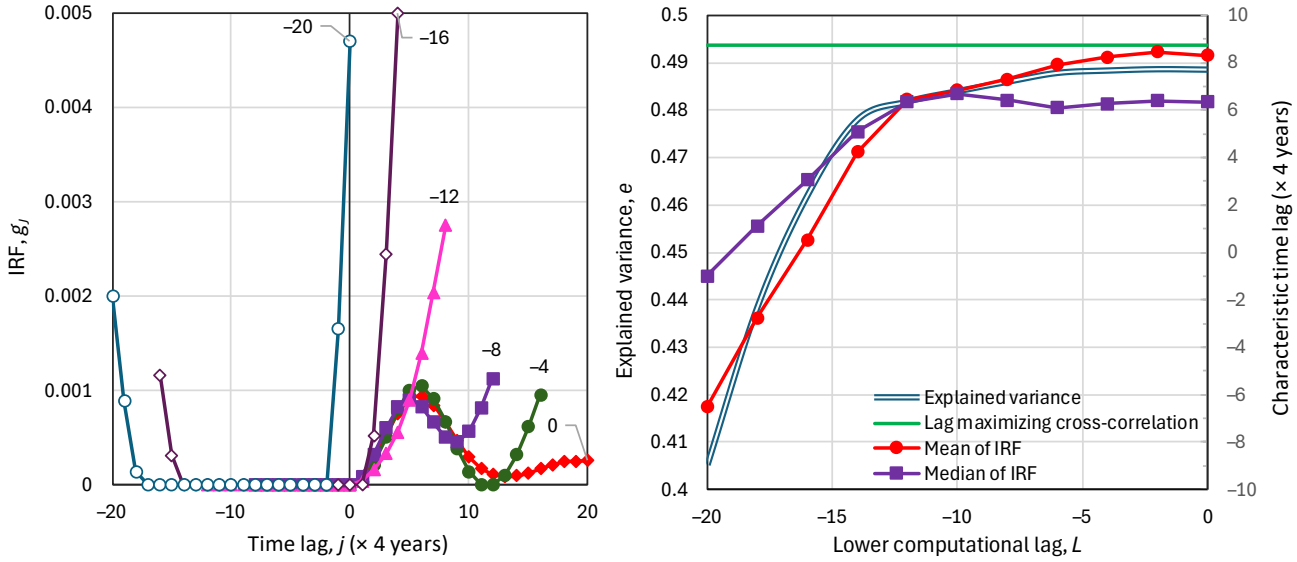

Figure 16 Application of the methodology to the Common Era data, namely the Loehle and McCulloch [58] for temperature and ice core data for [CO_2] (original series at time scale of 1 year with a multiplier $m = 4$). (**left**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right**) explained variance and characteristic time lags as functions of the lower computational lag, L.

We also analyze these data on a decadal time scale, by replacing the original series with moving averages of ten years. Now the differenced time series also provides information as the cross-correlations are not zero. The results of the application of the method for both the original and differenced series are shown in Figure 17 and again support the case of a potentially causal system at the direction $T \rightarrow [CO_2]$, with an explained variance at $L = 0$ equal to 0.49 for the original series but only 0.11 for the differenced series.

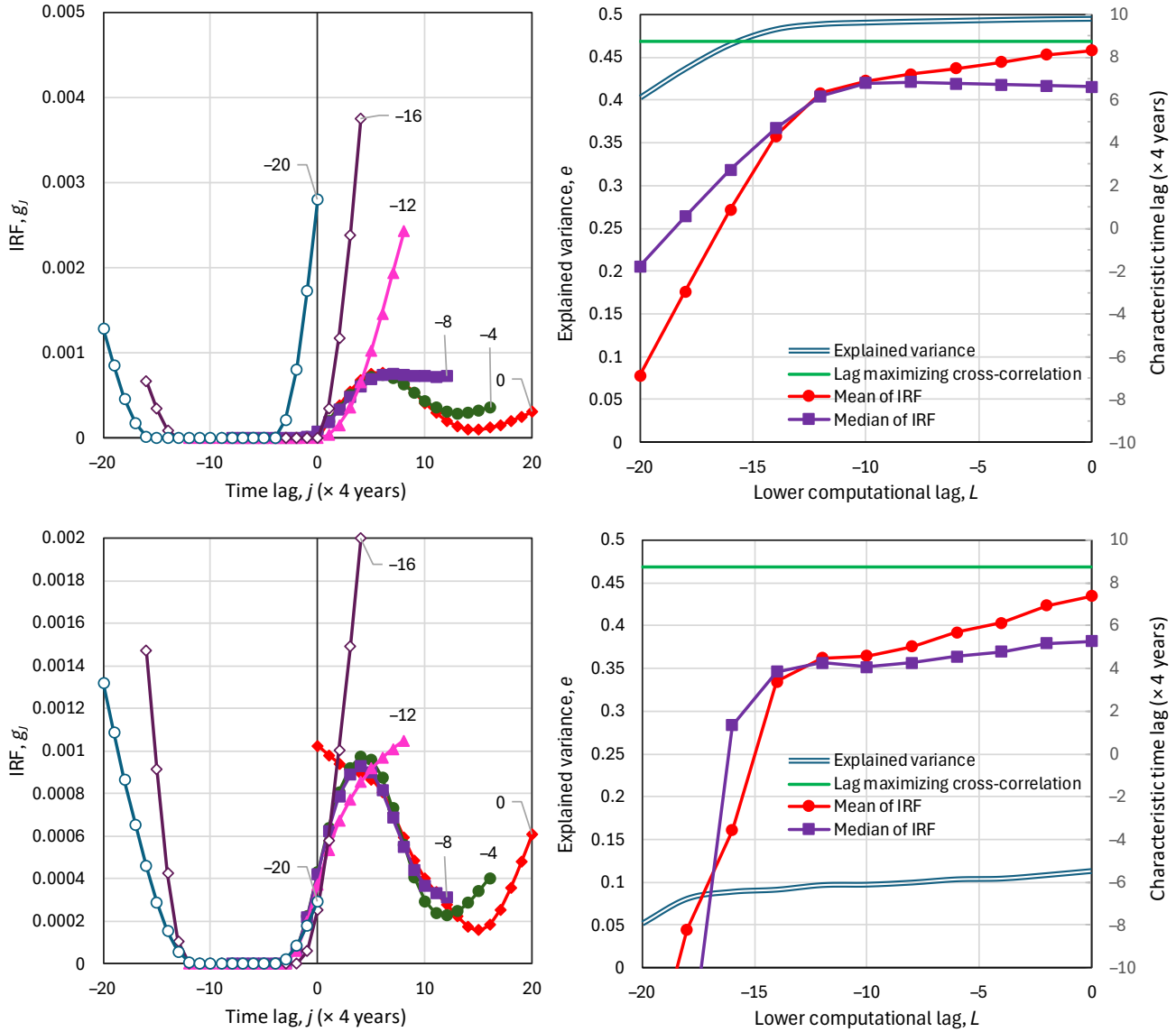

Figure 17 Application of the methodology to the Common Era data, namely the Loehle and McCulloch [58] for temperature and ice core data for [CO_2]; time scale of 10 years, multiplier $m = 4$. (**upper row**) Original series; (**lower row**) differenced series. (**left column**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right column**) explained variance and characteristic time lags as functions of the lower computational lag, L.

4.3 Instrumental data

As already discussed in Section 4.1, for the instrumental data we only use differenced series. The time step in the instrumental data is monthly, but the analyses are made on annual scale, so that the multiplier $l = 1/12$. To this aim, the differencing time step is taken as one year. In the earlier studies [20,22,23] from the original time series. x_i, the differenced time series was constructed as:

$$\tilde{x}_i := x_{i+12} - x_i \tag{15}$$

This has the additional advantage of limiting the influence of seasonality. To completely remove this influence, here we follow a different tactic. First, we take a moving average at a sliding window of one year, forming the time series:

$$x'_i := \frac{x_i + \cdots + x_{i+11}}{12} \tag{16}$$

and then we difference with a step of 12 months, forming the time series:

$$\tilde{x}'_i := x'_{i+12} - x'_i = \frac{(x_{i+12} - x_i) + \cdots (x_{i+23} - x_{i+11})}{12} = \frac{\tilde{x}_i + \cdots + \tilde{x}_{i+11}}{12} \tag{17}$$

The same transformation we make for the time series y_i.

The results of the analyses of $(\tilde{x}'_i \to \tilde{y}'_i)$ are seen in Figure 18 for Mauna Loa and in Figure 19 for the South Pole. In the latter case, two temperature series have been used (even though one is shown in the figure), southern hemisphere and global, and the latter resulted in higher explained variance, 0.30 against 0.21 of the former for $L = 0$. As in the earlier studies, each of the two figures clearly suggests a potentially causal system at the direction $T \to [CO_2]$. We further note that the case $(\tilde{x}_i \to \tilde{y}_i)$, i.e., without averaging, was also tested and the results were very similar, with a slightly smaller explained variance.

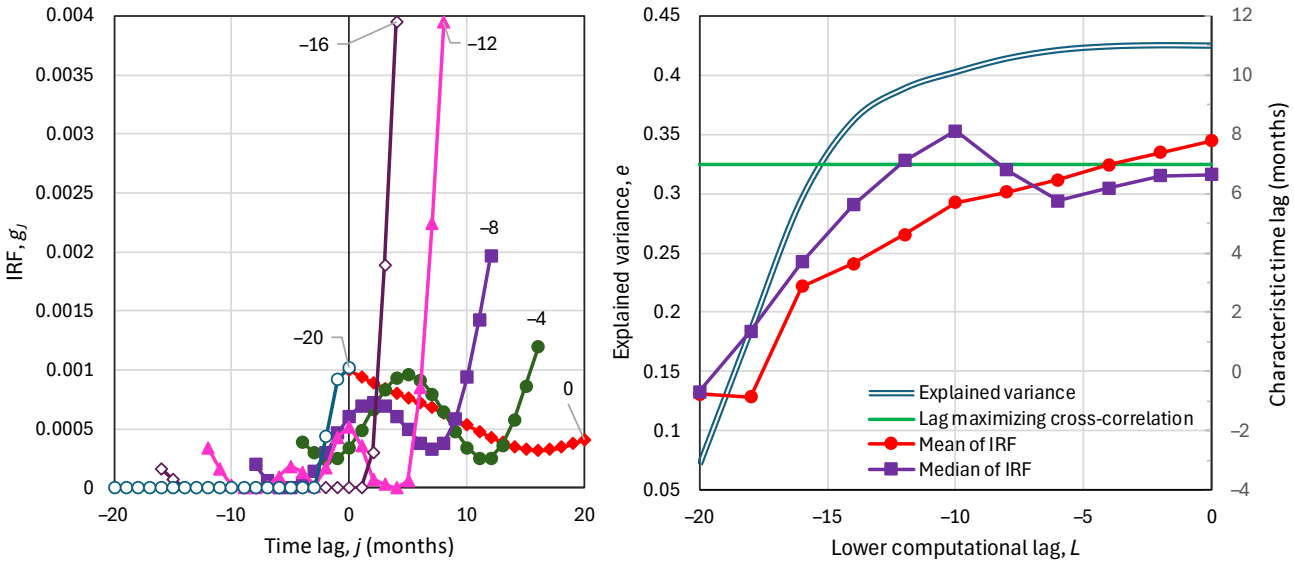

Figure 18 Application of the methodology to the Mauna Loa [CO_2] data with global ERA5 temperature data (differenced series at time annual scale). (**left**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right**) explained variance and characteristic time lags as functions of the lower computational lag, L.

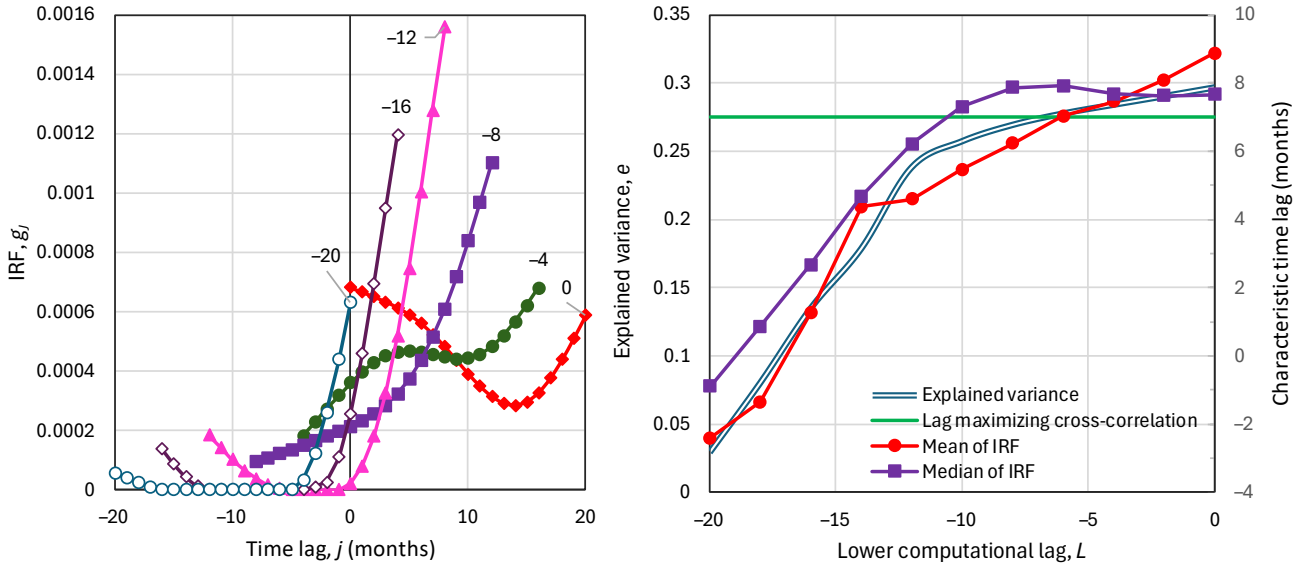

Figure 19 Application of the methodology for differenced series at time annual scale to the South Pole [CO_2] data with global ERA5 temperature data. (**left**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right**) explained variance and characteristic time lags as functions of the lower computational lag, L.

As the instrumental series cover a period of nearly 70 years, it is possible to investigate the causality at time scales higher than the annual. As in the case of the Common Era data, we also analyze the system at a time scale an order of magnitude higher, i.e. 10 years. We first note that the averaged differenced series for a moving window of 10 years is

$$\begin{aligned}\frac{\tilde{x}_i + \tilde{x}_{i+1} + \cdots + \tilde{x}_{i+119}}{120} &= \frac{(x_{i+12} - x_i) + (x_{i+13} - x_{i+1}) + \cdots + (x_{i+131} - x_{i+119})}{120} \\ &= \frac{(x_{i+131} + \cdots + x_{i+120}) - (x_{i+11} + \cdots + x_i)}{120} = \frac{12x'_{i+120} - 12x'_i}{120} \\ &= \frac{x'_{i+120} - x'_i}{10}\end{aligned} \tag{18}$$

Hence, to assess causality at the decadal scale it suffices to analyze the differenced series:

$$\tilde{x}''_i := x'_{i+120} - x'_i \tag{19}$$

i.e. to change the differencing time step from one year (as in $\tilde{x}'_i$) to ten years (as in $\tilde{x}''_i$).

The results are shown graphically in Figure 20 for Mauna Loa and Figure 21 for South Pole. In both cases, the IRFs for multiplier $m = 1$ do not seem reasonable (in their convex shape) and suggest that an IRF time frame of 20 months does not suffice to properly recover the IRF. Therefore, in both cases we also constructed an IRF for multiplier $m = 4$, corresponding to an IRF time frame of 80 months (6.7 years). Inspection of all cases shows no ambiguity that we have a potentially causal system at the decadal scale as the explained variance, particularly for $m = 4$, is high at $L = 0$ and clearly lower at all other L values. It can be said that the results are even clearer at the decadal scale, compared to those at annual scale.

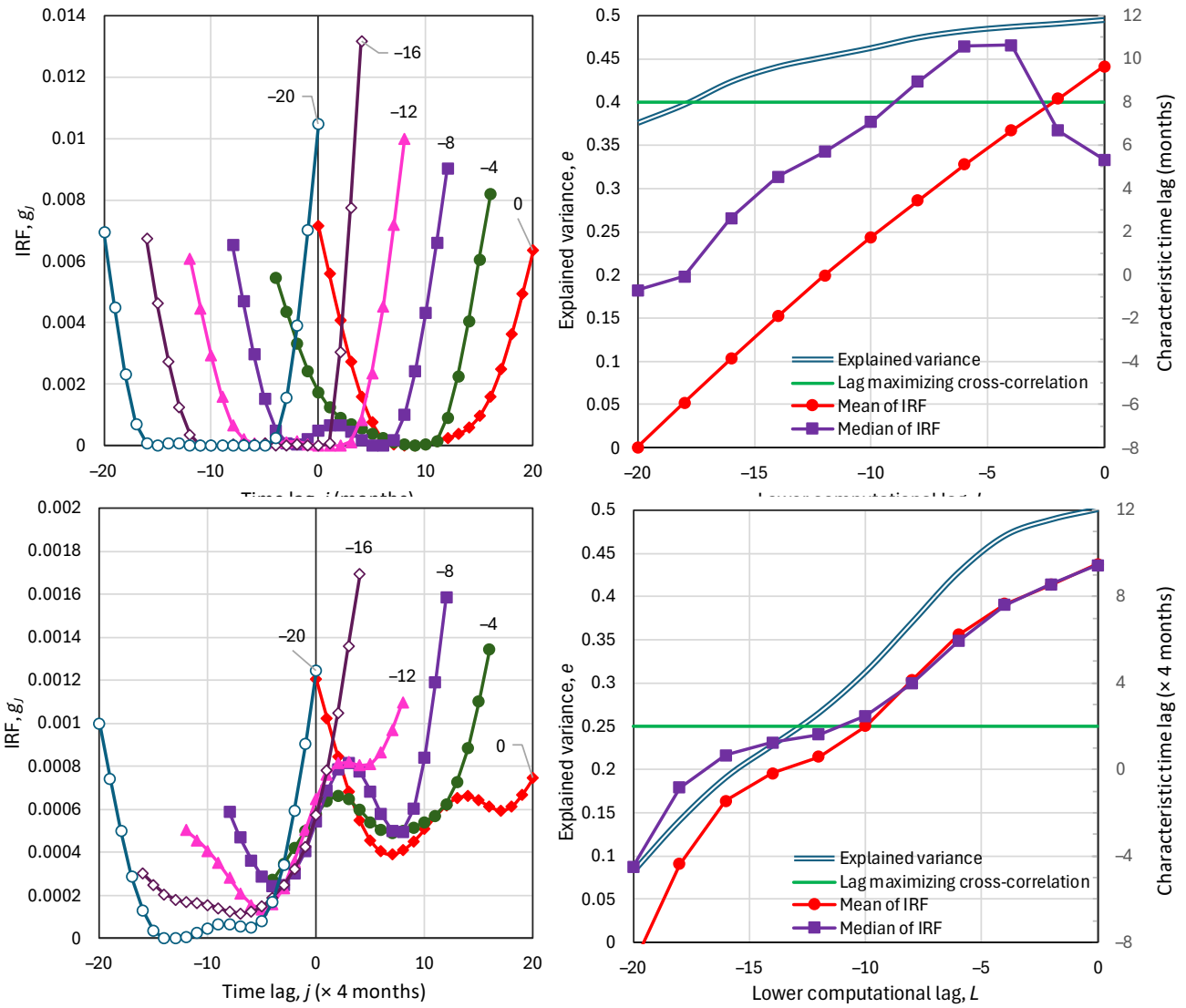

Figure 20 Application of the methodology to differenced series at decadal time scale to the Mauna Loa [CO_2] data with global ERA5 temperature data and monthly time step with (**upper row**) multiplier m = 1; (**lower row**) multiplier m = 4. (**left column**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right column**) explained variance and characteristic time lags as functions of the lower computational lag, L.

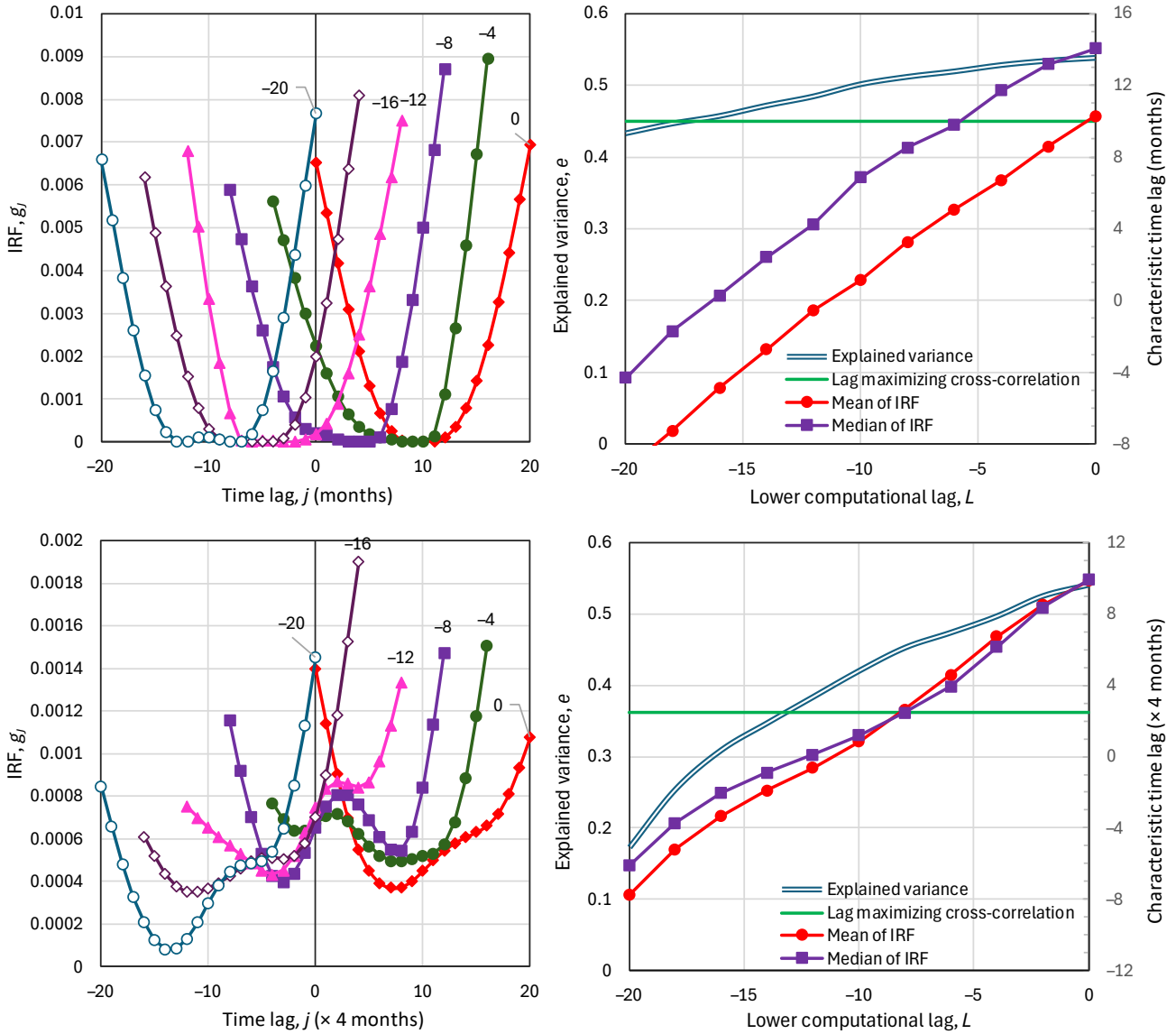

Figure 21 Application of the methodology to differenced series at decadal time scale to the South Pole [CO_2] data with global ERA5 temperature data and monthly time step with (**upper row**) multiplier m = 1; (**lower row**) multiplier m = 4. (**left column**) Estimated IRFs for the indicated lower computational lag, L (marked at the high end of each curve); (**right column**) explained variance and characteristic time lags as functions of the lower computational lag, L.

5 Summary and discussion

The extensive analyses made can easily be summarized as their results converge to the single inference that change in temperature leads and that in carbon dioxide concertation lags. This is supported both by proxy and instrumental data in all time scales and time spans. In quantified terms these results are summarized in Table 1. The causal system at the direction $T \rightarrow [CO_2]$ is always superior, in terms of explained variance, to the anticausal system, which would imply a causality direction $[CO_2] \rightarrow T$. The fact that in an anticausal system the explained variance is mostly nonzero is

expected, because of the high autocorrelation values appearing in both processes (particularly in [CO_2]) as discussed in Sections 3.3 and 4.1.

The direction $[CO_2] \rightarrow T$ was also directly investigated, even though no detailed graphs are provided, and the results are consistent with the above, i.e., in this case the explained variance for the anticausal system is greater than that of the causal system. Even a HOE causality is generally excluded, with a couple of exceptions noted in the Table, where again the time lags are positive. This means that, even if the exceptions are not statistical errors, temperature change leads and change in carbon dioxide concertation lags.

Table 1 Summary of results of all causality analyses for direction $T \rightarrow [CO_2]$ and for the causal ($L = 0$) and the anticausal ($L = -20$) systems; the time unit is in years.

Case	O/D*	Time span	Time scale	Multipliers l, m	Explained variance		Time lags†		
					Causal	Anticausal	h_c	$h_{1/2}$	μ_h
Phanerozoic	O	491×10^6	10^6	1, 1	0.62	0.61	3×10^6	1.2×10^6	4.0×10^6
	D	490×10^6	10^6	1, 1	0.27	0.19	2×10^6	2.3×10^6	6.4×10^6
Cenozoic	O	66.8×10^6	10^5	1, 1	0.80	0.78	4×10^5	7.6×10^5	9.1×10^5
Late Quaternary	O	416 000	1000	1, 1	0.84	0.71	1000	1160	4540
	O	200 500	500	1, 1	0.89	0.72	1500	1220	3290
	D	200 000	500	1, 1	0.22	0.15	0	1210	2410
Common Era	O	1711	1	1, 4	0.49	0.41	35	25	33
	O	1701	10	1/10, 4	0.49	0.40	35	26	33
	D	1700	10	1/10, 4	0.11	0.05	35	21	30
Modern/Mauna Loa	D	63	1	1/12, 1	0.43	0.07	0.58	0.55	0.65
	D	54	10	1/120, 1	0.49	0.38	0.58	0.44	0.80
	D	54	10	1/120, 4	0.50	0.08	0.67	3.15	3.17
Modern/South Pole	D	65	1	1/12, 1	0.30	0.03	0.58	0.64	0.74
	D	56	10	1/120, 1	0.54	0.43	0.83	1.16	0.86
	D	56	10	1/120, 4	0.54	0.17	0.83	3.31	3.28

* O: original, C: differenced.

† The time lags of the causal system are noted, which generally yield the highest explained variance. There are two exceptions, in which a HOE system maximizes the explained variance. These are: (a) Phanerozoic/ Original, with lower computational lag L = –14, in which $h_{1/2} = 3.7 \times 10^6, \mu_h = 0$ (years); (b) Late Quaternary/Differenced, with lower computational lag L = –1, in which $h_{1/2} = 850, \mu_h = 1970$ (years).

A remarkable result is that, as the time scale increases, so do the characteristic time lags. This sounds paradoxical, as one would expect an IRF independent of the data resolution and time span. However, this could be explained both on physical and mathematical grounds. On physical grounds, one may think that several mechanisms exist, which act on different time scales, that produce increase of atmospheric [CO_2] when temperature rises. Usually Henry's law is invoked, according to which increase of water temperature results in decreasing solubility of CO_2 in water and hence CO_2 degassing from the oceans [96]. However, this may not be responsible for changes on a multitude of time scales. As explained in [23], the biosphere processes offer the key mechanisms to this causal relationship.

CO_2 is produced during degradation of biological matter, by bacterial and thermophilic processes. During cool periods, degradation slows more than photosynthesis, and this traps CO_2 into soil. During warm periods, carbon trapped in soils is released faster than photosynthesis can absorb it, and atmospheric CO_2 increases. Respiration of all aerobic organisms also emits CO_2 and hence changes in the species occurring at creation and extinction events, which enter the scene as time scales increase, affect the T – [CO_2] relationship. Thus, as time span increases, more mechanisms of this relationship manifest themselves in the data.

On mathematical grounds, it is explained in Appendix A that the increase of time span may result in increasing characteristic lags. This is particularly the case for the mean, which may diverge to infinity if the IRF has a heavy tail with tail index higher than 1. The median is finite, but again its convergence to its true value can be very slow. Thus, even the median lag could increase as the time span of the data increases.

6 Conclusions

The premise of this study is that the climatic system is very complex and subject to perpetual change due to numerous processes, either internal or external to it. The fact that one of them, namely the relationship of climate with atmospheric CO_2, is highlighted in the last decades does not correspond to its actual importance as a climate driver. The promoted importance is a non-scientific issue, related to the narrative that humans, through their emissions by fossil fuel burning, are responsible for the changes we see in climate.

However, changes occurred in the entire history of Earth and the data sets used here, going back to the entire Phanerozoic eon (the last 541 million years), clearly show the perpetual nature of change. Furthermore, the consistent stochastic methodology for causality identification further developed in this study, as opposite to inconsistent deterministic methods, clearly supports a single conclusion: while there is a close relationship between temperature and atmospheric carbon dioxide concertation, it is temperature that leads, while [CO_2] lags. That is, the causality direction is $T \rightarrow [CO_2]$ and not the opposite, as emphatically promoted in the last decades. This is the case not only in the distant past and on large time scales, but also in the recent period covered by reliable instrumental records. And it happens not only at the monthly or annual time scale, but also at the decadal time scale, which can be resolved by the existing data sets. Notably, the results of the method application on the decadal time scale with instrumental data are even more conclusive than those on the annual time scale. And this is similar to what the proxy data for the Common Era suggest.

In conclusion, both the paleoclimatic and the instrumental data confirm the same causality direction $T \rightarrow [CO_2]$ for all time scales, from very large to very short, and all periods. The mechanisms responsible to this causality direction are physical, related to the interaction of oceans and atmosphere, but mostly are related to the biosphere processes. Is there a mechanism that, at the recent period has, due to human or presumed ‘unnatural’ actions, reversed directionality, as the popular claim is?

Perhaps, but no analysis based on observational data has shown that. Rather, such claims are based on imagination and climatic models full of assumptions. However, as shown in [23] the causality direction in time series produced by climatic models is opposite to that of the real-world data.

Use of AI tools declaration

The author declares he has not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research received no funding but was conducted out of scientific curiosity. I would like to thank Judith Curry and the discussers in her blog (https://judithcurry.com/2023/09/26/causality-and-climate/) on an earlier paper [23] coauthored by me. I particularly thank Robert Cutler for providing independent confirmation of the findings of that paper, using a spectral method, as well as an anonymous blogger nicknamed "Agnostic" for providing rich information, which also supported our results. I also thank John A Marcou (Jakarta, Indonesia) for the encouragement to make the present research and Richard Cina (St. Louis, Missouri, USA) for his constructive suggestions on a first draft of the paper.

Conflict of interest

The author declares no conflict of interest.

Data availability

This research uses no new data. The data sets used have been retrieved from the sources described in detail in the paper.

Appendix A: Theoretical investigation on characteristic time lags

To illustrate the effect of the observation time window on the estimated characteristic lags, we assume that the true IRF is of Pareto type,

$$g(h) := (1 + \xi h)^{-1-\frac{1}{\xi}} \tag{A1}$$

with tail index $\xi = 2$. We assume that, while in reality the lag h spans from 0 to ∞, we only have an observational window of length W and we truncate Equation (A1) to the interval [0, W]. By integration it is found that the truncated mean will be

$$\mu_h^W = \frac{(1+W)(1+\xi W)^{-1/\xi} - 1}{\xi - 1} \tag{A2}$$

while the truncated median can be found from

$$\int_0^{h_{1/2}^W} g(x)\mathrm{d}x = \frac{1}{2}\int_0^W g(x)\mathrm{d}x \tag{A3}$$

which, when solved, yields

$$h_{1/2}^W = \frac{2^{\xi}\left(1+(1+\xi W)^{-1/\xi}\right)^{-\xi} - 1}{\xi} \tag{A4}$$

These are shown graphically in Figure A1.

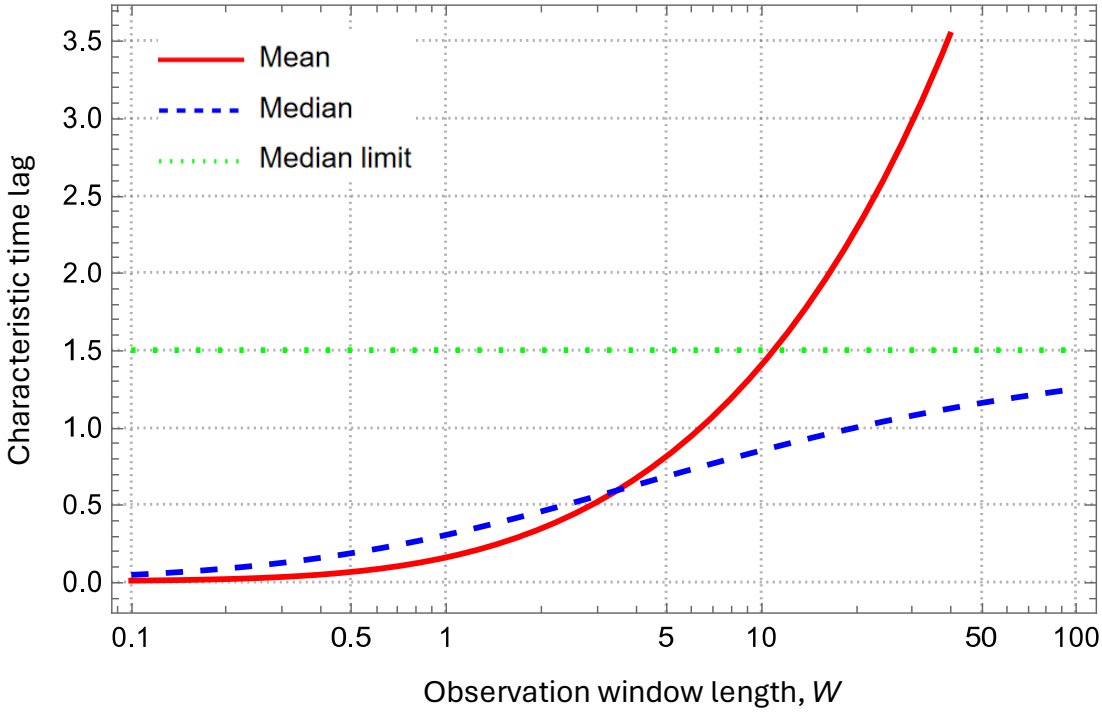

Figure A1 Characteristic time lags for a truncated Pareto IRF as functions of the window length.

To find the true characteristic time lags we take the limits as $W \to \infty$ and obtain:

$$\lim_{W\to\infty} \mu_h^W = \infty, \qquad \lim_{W\to\infty} h_{1/2}^W = \frac{2^{\xi}-1}{\xi} \tag{A5}$$

and for $\xi = 2$:

$$\lim_{W\to\infty} \mu_h^W = \infty, \qquad \lim_{W\to\infty} h_{1/2}^W = \frac{3}{2} \tag{A6}$$

Apparently, the infinite mean cannot be recovered by any observation window length W, while even the median converges slowly to its theoretical value of 1.5. These observations justify why the characteristic time lags estimated by our stochastic method can increase as the observation window length increases.

As already mentioned, the lag that maximizes the cross-correlation does not depend on the IRF if a time series is given. However, in a theoretical level, if both the IRF, $g(h)$, and the autocorrelation

function, $c_{xx}(h)$, are identified, then the cross-correlation function is [2]:

$$c_{yx}(h) = \int_0^\infty g(a)c_{xx}(h-a)\mathrm{d}a = \int_{-\infty}^{h} c_{xx}(b)g(h-b)\,\mathrm{d}b \tag{A7}$$

and by maximizing it we can determine the lag in question. By virtue of the symmetry of $c_{xx}(h)$ we can write

$$c_{yx}(h) = \int_0^\infty c_{xx}(y)g(h+y)\,dy + \int_0^h c_{xx}(y)g(h-y)\,dy \tag{A8}$$

As an illustration, we again use the example of Equation (A1) with a power-type autocorrelation function,

$$c_{xx}(h) = \left(1+\frac{h}{\lambda}\right)^{2H-2} \tag{A9}$$

where H is the Hurst parameter and λ is a scale parameter. As an additional case, we also examine the exponential (Markov) autocorrelation,

$$c_{xx}(h) = e^{-\frac{h}{\lambda}} \tag{A10}$$

For particular values of H, analytical solution of Equation (A8), are possible. For example, for $H = 3/4$ (and $\xi = 2$) we get:

$$c_{yx}(h) = \frac{\sqrt{2}(1+2h)\sqrt{\lambda} + 2\sqrt{2}\lambda^{3/2} + 2(1+2h)\sqrt{\lambda(h+\lambda)} - 4\lambda(\sqrt{1+2h} + \sqrt{\lambda(h+\lambda)})}{(1+2h)^2 - 4\lambda^2} \tag{A11}$$

However, numerical integration and maximization to find the maximizing lag h_c is always possible. Figure A2 exemplifies several cases, and also shows the true median, which, as found above is 1.5. The figure shows that h_c can substantially differ from $h_{1/2}$ and, hence, the differences found in real-world applications should not need to trouble us.

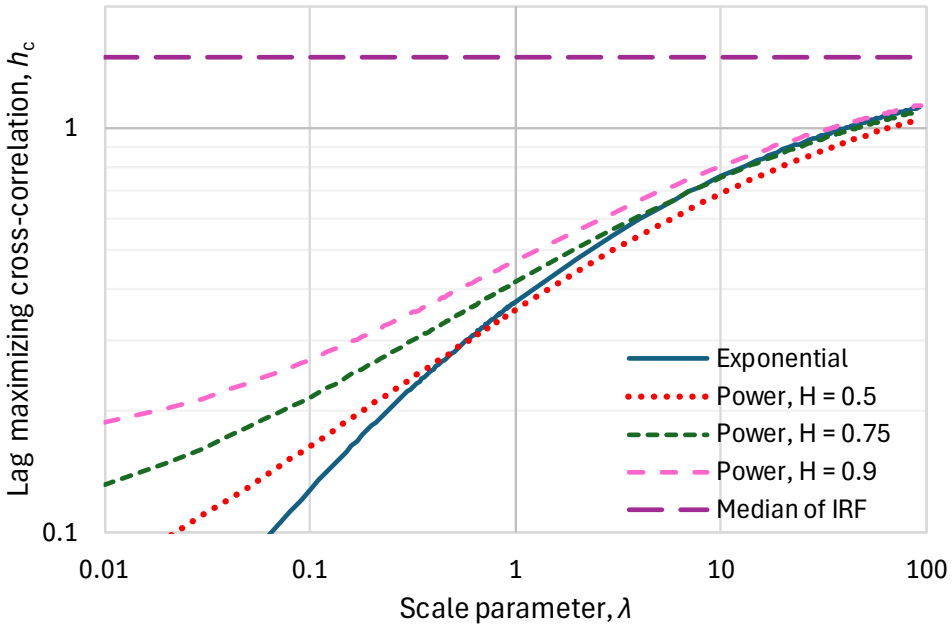

Figure A2 An example of the variation of the time lag maximizing the cross-correlation coefficient as a function of the scale parameter of the autocorrelation function.

References

1. L. Goldensohn, Hans Fritzsche interview, in *The Nuremberg Interviews*, Ed. by R. Gellately, Vintage Books, New York, 2005.
2. D. Koutsoyiannis, C. Onof, A. Christofides, Z.W. Kundzewicz, Revisiting causality using stochastics: 1. Theory, *Proc. R. Soc. A*, **478** (2022), 20210836. https://doi.org/10.1098/rspa.2021.0836.
3. J. Veizer, Celestial climate driver: a perspective from four billion years of the carbon cycle, *Geoscience Canada*, **32** (2005), 13-28.
4. J. Veizer, The role of water in the fate of carbon dioxide: implications for the climate system, in: *43rd Int. Seminar on Nuclear War and Planetary Emergencies* (2011) R. Ragaini (Ed.), World Scientific, 313-327, doi.org/10.1142/8232.
5. J. Veizer, Planetary temperatures/climate across geological time scales, in *International Seminar on Nuclear War and Planetary Emergencies—44th Session: The Role of Science in the Third Millennium* (2012), 287-288.
6. J. Laskar, A numerical experiment on the chaotic behaviour of the Solar System, *Nature*, **338** (1989), 237–238.
7. J. Laskar, The limits of Earth orbital calculations for geological time-scale use, *Philos. Tr. R. Soc. Lond. A*, **357** (1999), 1735–1759.
8. M. J. Duncan, Orbital stability and the structure of the solar system, in 1994, in: *Circumstellar*

Dust Disks and Planet Formation, Proceedings of the 10th IAP Astrophysics Meeting, Institut d'Astrophysique de Paris (1994), edited by: R. Ferlet, and A.Vidal-Madjar, Editions Frontiers, Gif sur Yvette, France, 245–256.

9. J. L. Lissauer, Chaotic motion in the Solar System, *Rev. Mod. Phys.*, **71** (1999), 835–845.
10. D. Koutsoyiannis, A random walk on water, *Hydrology and Earth System Sciences*, **14** (2010), 585–601. doi:10.5194/hess-14-585-2010.
11. M. Milanković, *Nebeska Mehanika*, Beograd, 1935.
12. M.Milanković, *Kanon der Erdbestrahlung und seine Anwendung auf das Eiszeitenproblem*, Ko niglich Serbische Akademice, Beograd, 1941.
13. M. Milanković, *Canon of Insolation and the Ice-Age Problem*, Agency for Textbooks, Belgrade, 1998.
14. J. D. Hays, J. Imbrie, J., N. J. Shackleton, Variations in the Earth's Orbit: Pacemaker of the Ice Ages, *Science*, **194** (1976), 1121-1132.
15. G. Roe, In defense of Milankovitch. *Geophys. Res. Lett.* **33** (2006), doi:10.1029/2006GL027817.
16. N. J. Shaviv, H. Svensmark, J. Veizer, The phanerozoic climate. *Annals of the New York Academy of Sciences*, **1519** (2023), 7-19.
17. D. Koutsoyiannis, Relative importance of carbon dioxide and water in the greenhouse effect: Does the tail wag the dog? (in review).
18. D. Koutsoyiannis, Time's arrow in stochastic characterization and simulation of atmospheric and hydrological processes, *Hydrological Sciences Journal*, **64** (2019), 1013–1037. doi: 10.1080/02626667.2019.1600700
19. D. Koutsoyiannis, Revisiting the global hydrological cycle: is it intensifying?, *Hydrology and Earth System Sciences*, **24** (2020), 3899–3932. doi:10.5194/hess-24-3899-2020
20. D. Koutsoyiannis, Z. W. Kundzewicz, Atmospheric temperature and CO_2: Hen-or-egg causality?, *Sci*, **2** (2020), 83. doi:10.3390/sci2040083.
21. D. Koutsoyiannis, Rethinking climate, climate change, and their relationship with water, *Water*, **13** (2021), 849. doi:10.3390/w13060849
22. D. Koutsoyiannis, C. Onof, A. Christofides, Z. W. Kundzewicz, Revisiting causality using stochastics: 2. Applications, *Proc. R. Soc. A*, **478** (2022), 20210835. doi:10.1098/rspa.2021.0835
23. D. Koutsoyiannis, C. Onof, Z. W. Kundzewicz, A. Christofides, On hens, eggs, temperatures and CO_2: Causal links in Earth's atmosphere, *Sci*, **5** (2023), 35. doi:10.3390/sci5030035.
24. D. Koutsoyiannis, C. Vournas, Revisiting the greenhouse effect—a hydrological perspective, *Hydrological Sciences Journal*, **69** (2024), 151–164. doi:10.1080/02626667.2023.2287047
25. D. Koutsoyiannis, Net isotopic signature of atmospheric CO_2 sources and sinks: No change since the Little Ice Age, *Sci*, **6** (2024), 17. doi:10.3390/sci6010017
26. J. D. Shakun, P. U. Clark, F. He, S. A. Marcott, A. C. Mix, Z. Liu, B. Otto-Bliesner, A. Schmittner, E. Bard, Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, *Nature*, 484 (2012), 49-54.
27. J. Chowdhry Beeman, L. Gest, F. Parrenin, D. Raynaud, T. J. Fudge, C. Buizert, E. J. Brook,

Antarctic temperature and CO_2: near-synchrony yet variable phasing during the last deglaciation, *Clim. Past*, 15 (2019), 913–926. doi: 10.5194/cp-15-913-2019

28. F. Parrenin, V. Masson-Delmotte, P. Köhler, D. Raynaud, D. Paillard, J. Schwander, C. Barbante, A. Landais, A. Wegner, J. Jouzel, Synchronous change of atmospheric CO_2 and Antarctic temperature during the last deglacial warming, *Science* **339**, (2013), 1060-1063.
29. W. Soon, Implications of the secondary role of carbon dioxide and methane forcing in climate change: past, present, and future. *Physical Geography* **28** (2007), 97-125.
30. R. A. Berner, Z. Kothavala, GEOCARB III: a revised model of atmospheric CO_2 over Phanerozoic time, *American Journal of Science*, **301** (2001), 182-204. doi: 10.2475/ajs.301.2.182
31. J. Veizer, Y. Godderis, L. M. François, Evidence for decoupling of atmospheric CO_2 and global climate during the Phanerozoic eon, *Nature*, **408** (2000), 698-701.
32. N. Caillon, J. P. Severinghaus, J. Jouzel, J. M. Barnola, J. Kang, V. Y. Lipenkov, Timing of atmospheric CO_2 and Antarctic temperature changes across Termination III, *Science* **299** (2003), 1728-1731.
33. J. B. Pedro, S. O. Rasmussen, T. D. van Ommen, Tightened constraints on the time-lag between Antarctic temperature and CO_2 during the last deglaciation, *Climate of the Past*, **8** (2012), 1213-1221.
34. O. Humlum, K. Stordahl, J. E. Solheim, The phase relation between atmospheric carbon dioxide and global temperature, *Global and Planetary Change*, **100** (2013), 51-69.
35. C. R. Scotese, H. Song, B. J. Milels, D. G. van der Meer, Phanerozoic paleotemperatures: The earth's changing climate during the last 540 million years, *Earth-Science Reviews*, **215** (2021), 103503.
36. E. L. Grossman, M. M. Joachimski, Ocean temperatures through the Phanerozoic reassessed, *Sci. Rep.*, **12** (2022), 8938.
37. H. Song, P. B. Wignall, H. Song, X. Dai, D. Chu, Seawater temperature and dissolved oxygen over the past 500 million years, *Journal of Earth Science*, **30** (2019), 236-243.
38. J. P. Klages, U. Salzmann, T. Bickert, et al. Temperate rainforests near the South Pole during peak Cretaceous warmth, *Nature* **580** (2020), 81–86. https://doi.org/10.1038/s41586-020-2148-5
39. J. M. Schaefer, R. C. Finkel, G. Balco, R. B. Alley, M. W. Caffee, J. P. Briner, N. E. Young, A. J. Gow, R. Schwartz, Greenland was nearly ice-free for extended periods during the Pleistocene, *Nature*, **540** (2016), 252-255.
40. K. H. Kjær, M. Winther Pedersen, B. De Sanctis, B. De Cahsan, T. S. Korneliussen, C. S. Michelsen, K. K. Sand, S. Jelavić, A. H. Ruter, A. M. Schmidt, K. K. Kjeldsen, A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA, *Nature*, **612** (2022), 283-291.
41. R. A.Berner, GEOCARBSULF: A combined model for Phanerozoic atmospheric O_2 and CO_2, *Geochimica Et Cosmochimica Acta*, **70** (2006), 5653–5664.
42. D. L. Royer, Y. Donnadieu, J. Park, J. Kowalczyk, Y. Godderis, Error analysis of CO_2 and O_2 estimates from the long-term geochemical model GEOCARBSULF, *American Journal of Science*, **314** (2014), 1259–1283.

43. G. L. Foster, D. L. Royer, D. J. Lunt, Future climate forcing potentially without precedent in the last 420 million years, *Nature Communications*, **8** (2017), 14845.
44. D. L. Royer, Atmospheric CO_2 and O_2 during the Phanerozoic: Tools, patterns and impacts. In *Treatise on Geochemistry*, 2nd ed.; H. Holland, K. Turekian, Eds., Elselvier, Amsterdam, The Netherlands, (2014), 251–267.
45. W. J. Davis, The relationship between atmospheric carbon dioxide concentration and global temperature for the last 425 million years, *Climate*, **5** (2017), 76.
46. R. A. Berner, Addendum to "Inclusion of the Weathering of Volcanic Rocks in the GEOCARBSULF Model" (R. A, Berner, 2006, V. 306, p. 295–302), *American Journal of Science*, **308** (2008), 100-103. https://doi.org/10.2475/01.2008.04
47. A. L. Melott, R. K. Bambach, Analysis of periodicity of extinction using the 2012 geological timescale. *Paleobiology*, **40** (2014), 177-196.
48. W. J. Davis, Mass extinctions and their relationship with atmospheric carbon dioxide concentration: Implications for Earth's future, *Earth's Future*, **11** (2023), e2022EF003336.
49. J. J. Sepkoski, A factor analytic description of the Phanerozoic marine fossil record, *Paleobiology*, **7** (1981), 36-53.
50. T. Westerhold, N. Marwan, A. J. Drury, D. Liebrand, C. Agnini, E. Anagnostou, J. S .Barnet, S. M. Bohaty, D. De Vleeschouwer, F. Florindo, T. Frederichs, An astronomically dated record of Earth's climate and its predictability over the last 66 million years, *Science*, **369** (2020), 1383-1387.
51. J. W. Rae, Y. G. Zhang, X. Liu, G. L. Foster, H. M. Stoll, R. D. Whiteford, Atmospheric CO_2 over the past 66 million years from marine archives, *Annual Review of Earth and Planetary Sciences*, **49** (2021), 609-641.
52. S. Epstein, R. Buchsbaum, H. A. Lowenstam, H. C. Urey, Revised carbonate-water isotopic temperature scale, *Geological Society of America Bulletin*, 64 (1953), 1315-1326.
53. J. Jouzel, C. Lorius, J. R. Petit, C. Genthon, N. I. Barkov, V. M. Kotlyakov, V. M. Petrov, Vostok ice core: a continuous isotope temperature record over the last climatic cycle (160 000 years), *Nature* **329** (1987), 403-408.
54. J. R. Petit, J. Jouzel, D. Raynaud, N. I. Barkov, J.-M. Barnola, I. Basile, M. Bender, J. Chappellaz, M. Davis, G. Delayque, M. Delmotte, V. M. Kotlyakov, M. Legrand, V. Y. Lipenkov, C. Lorius, L. Pepin, C. Ritz, E. Saltzman, M. Stievenard, Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, *Nature* **399** (1999), 429-436.
55. J.-M. Barnola, P. Pimienta, D. Raynaud, Y. S. Korotkevich. CO_2 - climate relationship as deduced from the Vostok ice core: a re-examination based on new measurements and on a re-evaluation of the air dating, *Tellus B: Chemical and Physical Meteorology*, **43** (1991), 83–90. doi: 10.3402/tellusb.v43i2.15249.
56. B. Christiansen, F. C. Ljungqvist, Challenges and perspectives for large-scale temperature reconstructions of the past two millennia, *Reviews of Geophysics*, **55** (2017), 40-96.
57. A. Moberg, D. M. Sonechkin, K. Holmgren, N. M. Datsenko, W. Karlén, Highly variable Northern Hemisphere temperatures reconstructed from low-and high-resolution proxy data, *Nature*, 433

(2005), 613-617.

58. C. Loehle, J. H. McCulloch, Correction to: A 2000-year global temperature reconstruction based on non-tree ring proxies, *Energy & Environment*, 19 (2008), 93-100.
59. B. Christiansen, and F. C. Ljungqvist, The extra-tropical Northern Hemisphere temperature in the last two millennia: reconstructions of low-frequency variability, *Climate of the Past*, **8** (2012), 765-786.
60. A. Indermühle, T. F. Stocker, F. Joos, H. Fischer, H. J. Smith, M. Wahlen, B. Deck et al. Holocene carbon-cycle dynamics based on CO_2 trapped in ice at Taylor Dome, Antarctica, *Nature*, **398** (1999), 121-126.
61. D. M. Etheridge, L. P. Steele, R. L. Langenfelds, R. J. Francey, J.-M. Barnola, V. I. Morgan, Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn, *Journal of Geophysical Research: Atmospheres* **101**, (1996), 4115-4128.
62. R. J. Francey, C. E. Allison, D. M. Etheridge, C. M. Trudinger, , I. G. Enting, M. Leuenberger, R. L. Langenfelds, E. Michel, L. P. Steele, A 1000-year high precision record of $\delta^{13}C$ in atmospheric CO_2, *Tellus B*, **51** (1999), 170-193.
63. F. Böhm, A. Haase-Schramm, A. Eisenhauer, W. C. Dullo, M. M. Joachimski, H. Lehnert, J. Reitner, Evidence for preindustrial variations in the marine surface water carbonate system from coralline sponges, Geochem. *Geophys. Geosyst.*, **3** (2002), 1 – 13.
64. L. L. R. Kouwenberg, *Application of conifer needles in the reconstruction of Holocene CO_2 levels*, Ph.D. thesis, Utrecht, Laboratory of Paleobotany and Palynology (LPP) Foundation, LPP Contributions series 16, 130 p., 2004. https://dspace.library.uu.nl/bitstream/handle/1874/243/full.pdf (accessed 2024-03-10).
65. L. Kouwenberg, R. Wagner, W. Kürschner, H. Visscher, Atmospheric CO_2 fluctuations during the last millennium reconstructed by stomatal frequency analysis of Tsuga heterophylla needles, *Geology*, **33** (2005), 33-36.
66. T. B. van Hoof, F. Wagner-Cremer, W. M. Kürschner, H. Visscher, A role for atmospheric CO_2 in preindustrial climate forcing, *Proc. Nat. Acad. Sci.*, **105** (2008), 15815-15818.
67. J. M. Barnola, M. Anklin, J. Porcheron, D. Raynaud, J. Schwander, B. Stauffer, CO_2 evolution during the last millennium as recorded by Antarctic and Greenland ice, *Tellus B: Chemical and Physical Meteorology*, **47** (1995), 264-272.
68. M. Anklin, J. Schwander, B. Stauffer, J. Tschumi, A. Fuchs, J. M. Barnola, D. Raynaud, CO_2 record between 40 and 8 kyr BP from the Greenland Ice Core Project ice core, *Journal of Geophysical Research: Oceans*, **102** (1997), 26539-26545.
69. J. C. McElwain, F. E. Mayle, D. J. Beerling, Stomatal evidence for a decline in atmospheric CO_2 concentration during the Younger Dryas stadial: a comparison with Antarctic ice core records, *Journal of Quaternary Science*, **17** (2002), 21-29.
70. F. Wagner, L. L. Kouwenberg, T. B. van Hoof, H. Visscher, Reproducibility of Holocene atmospheric CO_2 records based on stomatal frequency, *Quaternary Sci. Rev.*, **23** (2004), 1947-1954.

71. C. A. Jessen, M. Rundgren, S. Björck, D. Hammarlund, Abrupt climatic changes and an unstable transition into a late Holocene Thermal Decline: a multiproxy lacustrine record from southern Sweden, *J. Quaternary Sci.*, **20** (2005) 349–362.
72. M. Steinthorsdottir, B. Wohlfarth, M. E. Kylander, M. Blaauw, P. J. Reimer, Stomatal proxy record of CO_2 concentrations from the last termination suggests an important role for CO_2 at climate change transitions, *Quaternary Science Reviews*, **68** (2013), 43-58.
73. Y. Wang, A. Momohara, N. Wakamatsu, T. Omori, M. Yoneda, M. Yang, Middle and Late Holocene altitudinal distribution limit changes of Fagus crenata forest, Mt. Kurikoma, Japan indicated by stomatal evidence, *Boreas*, **49** (2020), 718-729.
74. S. Azharuddin, P. Govil, T. B. Chalk, M. Shekhar, G. L. Foster, R. Mishra, Abrupt upwelling and CO_2 outgassing episodes in the north-eastern Arabian Sea since mid-Holocene, *Scientific Reports*, **12** (2022), 3830.
75. B. Christiansen, F. C. Ljungqvist, *Northern Hemisphere extratropical 2000 year temperature reconstruction*, IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series #2012-049, NOAA/NCDC Paleoclimatology Program, Boulder CO, USA, 2012. https://www.ncdc.noaa.gov/paleo/study/12902 (accessed on 16 March 2024).
76. ERA5: data documentation - Copernicus Knowledge Base - ECMWF Confluence Wiki. Available online: https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation #heading-Relatedarticles (accessed on 25 March 2023).
77. C. D. Keeling, S. C. Piper, T. P. Whorf, R. F. Keeling, Evolution of natural and anthropogenic fluxes of atmospheric CO_2 from 1957 to 2003, *Tellus B: Chemical and Physical Meteorology*, **63** (2011), 1-22.
78. Climate Explorer, Dutch Royal Netherlands Meteorological Institute (KNMI), http://climexp.knmi.nl/ (accessed on 25 January 2024).
79. Scripps CO_2 Program, Sampling Station Records, Available online, https://scrippsco2.ucsd.edu/data/atmospheric_co2/sampling_stations.html (accessed on 15 November 2023).
80. World Economic Forum, CO_2 levels in atmosphere are at their highest in 800,000 years https://www.weforum.org/agenda/2018/05/earth-just-hit-a-terrifying-milestone-for-the-first-time-in-more-than-800-000-years/ (accessed on 16 March 2024).
81. The Conversation, The three-minute story of 800,000 years of climate change with a sting in the tail, https://theconversation.com/the-three-minute-story-of-800-000-years-of-climate-change-with-a-sting-in-the-tail-73368 (accessed on 16 March 2024).
82. NASA, Graphic: The relentless rise of carbon dioxide, https://science.nasa.gov/resource/graphic-the-relentless-rise-of-carbon-dioxide (accessed on 16 March 2024).
83. M. Shan, Analysis on the influence of doubled carbon dioxide on the extreme weather, *Proceedings of the 3rd International Conference on Materials Chemistry and Environmental Engineering, Applied and Computational Engineering*, 3 (2023), 368-373. doi: 10.54254/2755-2721/3/20230554

84. H. S. Kang, S. Chester, C. Meneveau, Decaying turbulence in an active-grid-generated flow and comparisons with large-eddy simulation, *J. Fluid Mech.*, **480** (2003), 129–160.
85. Center for Environmental and Applied Fluid Mechanics, 30 data files at x/M=20 for probe separations corresponding to filter scale 5 mm, http://pages.jh.edu/~cmeneve1/datasets/Activegrid/M20/H1 (accessed on 16 March 2024)
86. D. Koutsoyiannis, Hydrology and Change, *Hydrological Sciences Journal*, **58** (2013), 1177–1197. doi: 10.1080/02626667.2013.804626.
87. D. Koutsoyiannis, Entropy production in stochastics, *Entropy*, **19** (2017), 581. doi: 10.3390/e19110581
88. D. Koutsoyiannis, *Stochastics of Hydroclimatic Extremes - A Cool Look at Risk*, Edition 3, 2023, ISBN: 978-618-85370-0-2, 391 pages, doi: 10.57713/kallipos-1, Kallipos Open Academic Editions, Athens.
89. J. Ahn, M. Headly, M. Wahlen, E. J. Brook, P. A. Mayewski, K. C. Taylor, CO_2 diffusion in polar ice: observations from naturally formed CO_2 spikes in the Siple Dome (Antarctica) ice core, *Journal of Glaciology*, **54** (2008), 685-695.
90. C. W. Granger, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, **37** (1969), 424-438.
91. C. W. Granger, Testing for causality: a personal viewpoint, *J. Econ. Dynamics and Control*, **2** (1980), 329-352.
92. J. Pearl, Causal inference in statistics: An overview, *Statistics Surveys*, **3** (2009), 96-146. doi: 10.1214/09-SS057
93. J. Pearl, M. Glymour, N. P. Jewell, *Causal Inference in Statistics: A Primer*, Wiley, Chichester, UK, 2016.
94. Hannart, A.; Pearl, J.; Otto, F.E.L.; Naveau, P.; Ghil, M. Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Amer. Met. Soc.*, 97 (2016), 99-110.
95. A. Prokoph, G. A. Shields, J. Veizer, Compilation and time-series analysis of a marine carbonate $\delta^{18}O$, $\delta^{13}C$, $^{87}Sr/^{86}Sr$ and $\delta^{34}S$ database through Earth history, *Earth Sci. Rev.*, **87** (2008),113–133.
96. R. Weiss, Carbon dioxide in water and seawater: The solubility of a non-ideal gas. *Mar. Chem.* **2** (1974), 203–215.