Time-Frequency Causal Inference Uncovers Anomalous Events in Environmental Systems

Maha Shadaydeh[1 [0000−0001−6455−2400]⋆], Joachim Denzler[1,4 [0000−0002−3193−3300]], Yanira Guanche García[2,4 [0000−0001−6703−9768]], and Miguel Mahecha[3,4 [0000−0003−3031−613X]]

[1] Computer Vision Group, Friedrich Schiller University, Jena, Germany
[2] Institute of Data Science, German Aerospace Center, DLR, Jena, Germany
[3] Max Planck Institute for Biogeochemistry, Jena, Germany
[4] Michael Stifel Center for Data driven and Simulation Science, Jena, Germany

Abstract. Causal inference in dynamical systems is a challenge for different research areas. So far it is mostly about understanding to what extent the underlying causal mechanisms can be derived from observed time series. Here we investigate whether anomalous events can also be identified based on the observed changes in causal relationships. We use a parametric time-frequency representation of vector autoregressive Granger causality for causal inference. The use of time-frequency approach allows for dealing with the nonstationarity of the time series as well as for defining the time scale on which changes occur. We present two representative examples in environmental systems: land-atmosphere ecosystem and marine climate. We show that an anomalous event can be identified as the event where the causal intensities differ according to a distance measure from the average causal intensities. The driver of the anomalous event can then be identified based on the analysis of changes in the causal effect relationships.

Keywords: Time-Frequency Causality Analysis · Vector Autoregressive Granger Causality · Attribution of anomalous events.

1 Introduction

Understanding causal effect relationships in dynamical systems is a challenging problem in different areas of research such as brain neural connectivity analysis, climatic attribution, psychology, among many others. These relationships are guided by the processes generating them. Hence, monitoring changes in the interaction patterns within the system variables can be used for simultaneous detection and diagnosis of changes in the underlying process. For visual illustration purpose, in Figure 1 we present an example where we show that the time (date of occurrence) and the scale (duration of the event) of three historic

⋆ Corresponding author: maha.shadaydeh@uni-jena.de

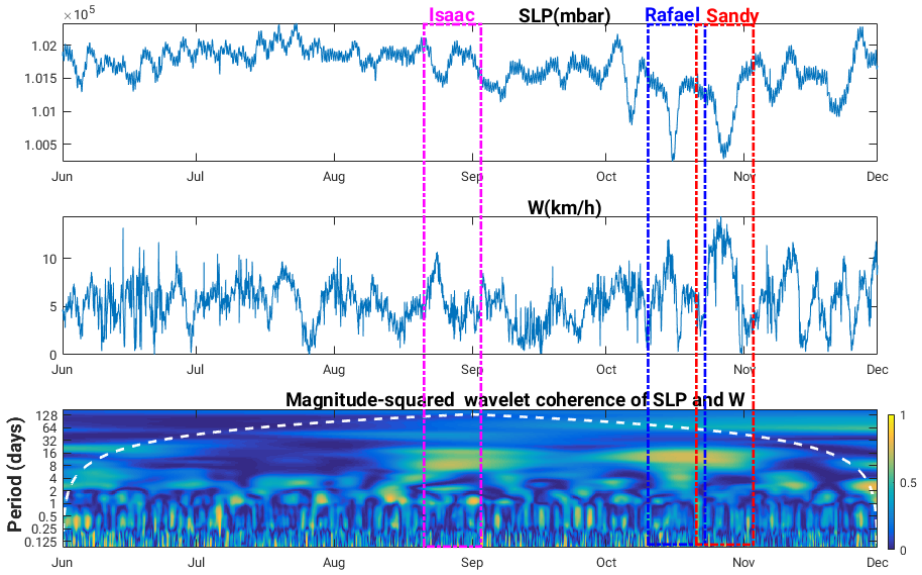

Fig. 1. Plots of sea level pressure (SLP), wind speed (W), and the magnitude-squared wavelet coherence of SLP and W. Data is extracted from the National Data Buoy Center for a buoy located near the Bahamas in the Atlantic Sea (23.838 N, 68.333 W). The high magnitude wavelet coherence correlates well with the time and duration of three historic hurricanes in year 2012: *Isaac*, *Rafael*, and *Sandy*. The white dashed line shows the cone of statistical significance.

hurricanes in the North Sea can be visualized using some correlation measure between two different marine variables namely: sea level pressure (SLP) and wind speed (W). This figure illustrates that, in a normal marine climate conditions, the wavelet coherence (a measure of the correlation between two variables in the time-frequency domain) is low, indicating that these variables are almost independent. However, during certain events, such as hurricanes, these variables start to show higher coherence that can be visualized on the time scale of 8 to 16 days.

To understand events in multivariate environmental time series, however, we need to go beyond monitoring correlation patterns and try to identify the driver of the event based on the analysis of the changes in the causal effect relationships between the variables. In this paper, we aim to build on this core idea for the detection and attribution of changes in environmental time series. To this end, we approach the problem in a way similar to fault detection and diagnosis in plants of automatic control systems [13, 31]. That is, we try to estimate the underlying models of different time intervals, and define change or anomaly based on how far is the estimated model of a certain time interval from an estimation of the true one. The change is then attributed to the variable(s) responsible for this

change and is defined based on the overall change in its causal effects on other variables in the system.

Various causality measures have been reported in literature [21, 9]. Among many other linear regression-based models, Granger causality (GC) [15] is the most widely known method for causality analysis. GC is based on the idea that causes both precede and help predict their effects. Hence, GC mainly focuses on linear models which assumes that the causes and effects are separable. It has received intense research interest in neuroscience literature aiming to unravel the detailed circuitry underlying perception, cognition, and behaviour [25]. Several attempts in the literature have applied the concept of Granger causality to climatic attribution. The authors of [5] provide a review of the use of GC for the attribution of global warming. To overcome the shortcoming of the separability assumption in Granger causality, several causality analysis methods for nonlinear systems have been proposed so far. Examples include methods developed from transfer entropy, recurrence networks [12] and nonlinear extension of Granger causality [20, 19].

Despite the abundance of environmental data, data-driven causal inference methods remain at its infancy compared to other areas of research [11]. A major challenge particular to environmental data is that causal effect interactions occur on multiple temporal and spatial scales. Recently, the importance of using deep learning approaches for understanding Earth and ecological processes has been addressed in [23]. While deep learning as well as other non-linear methods are continuously under development, linear methods remain of great interest mainly for being strictly connected to the frequency domain representation of multivariate time series [10], which is of great advantage in the analysis of environmental data. This is due to the fact that enviromental time series most often contain trend and periodical components (diurnal and seasonal cycles) that can significantly mask the underling causality structure in time domain. Moreover, the impact of filtering and down sampling (the two necessary steps for time domain multiscale causality analysis in discrete systems) on causal inference has been only recently studied for linear discrete systems [27, 7]; and it still presents a major challenge in nonlinear systems [27]. For the black box of deep learning structure, such impact is still far from being comprehensible.

In this paper we present a method for simultaneous detection and attribution of changes in multivariate environmental systems based on time-frequency causality analysis. The causal effect relationships are extracted using the parametric frequency domain representation of Vector AutoRegressive Granger causality (VAR-GC) [16, 14] applied on a sliding time window. In particular we use the generalized partial directed coherence (gPDC) [6] method for causal inference which allows for causality analysis at different frequency components. After introducing the time-frequency causal inference method, we present two different examples in environmental systems, where we show that an anomalous event can be defined as this event where the causal intensities between the variables differ according to some statistical distance measure from the average dynamical be-

haviour, and such anomalous event can be directly attributed to the variable(s) causing such deviation assuming that there is no hidden drivers.

2 Methodology

2.1 Vector Autoregressive Granger Causality (VAR-GC)

Let $\mathcal{S} = \{X_k, 1 \leq k \leq N\}$ denote a discrete stationary stochastic process which constitutes of N real valued environmental variables. Given the length m time series $x_k(n), n = 1, \ldots, m$ as the realizations of $X_k, k = 1, \ldots, N$, these time series can be represented by a pth order vector autoregressive model (VAR(p)) of the form

$$\begin{bmatrix} x_1(n) \\ \vdots \\ x_N(n) \end{bmatrix} = \sum_{r=1}^{p} A_r \begin{bmatrix} x_1(n-r) \\ \vdots \\ x_N(n-r) \end{bmatrix} + \begin{bmatrix} \epsilon_1(n) \\ \vdots \\ \epsilon_N(n) \end{bmatrix}, \tag{1}$$

The residuals $\epsilon_k, k = 1, \ldots, N$ constitute a white noise stationary process with an $N \times N$ residual covariance matrix Σ. The model parameters at time lags $r = 1, \ldots, p$ are predefined by

$$A_r = \begin{bmatrix} a_{11}(r) & \ldots & a_{1N}(r) \\ \vdots & \ddots & \vdots \\ a_{N1}(r) & \ldots & a_{NN}(r) \end{bmatrix}. \tag{2}$$

The elements of the matrix A_r, $a_{ij}(r)$ quantify the causal link from x_j to x_i at time lag r. The model order p defines the maximum lag used to estimate causal interactions. It can be estimated using either Akaike [3] or Bayesian Criterion [24]. The model parameters $a_{ij}(r), i, j = 1, \ldots, N, r = 1, \ldots, p$ can then be estimated using for example the method of Least Squares (LS) [17].

It should be noted that the use of the VAR(p) model (1) makes no assumption on the mechanism that produced the data (for example whether it is a linear one) except that the model itself exists and is stable [4].

The time domain VAR-GC of x_i on x_j conditioned on all other variables is defined by the likelihood ratio [14, 1]

$$\gamma_{i \to j} = \ln \frac{|\Sigma_{\mathrm{j}}^{\mathrm{i}-}|}{|\Sigma_{\mathrm{j}}|}, \tag{3}$$

where Σ_{j} and $\Sigma_{\mathrm{j}}^{\mathrm{i}-}$ are the covariance matrices of the residual ϵ_j associated to x_j using the full and reduced model (after eliminating x_i) respectively. The conditional MVAR GC (3) thus quantifies the degree to which the past of x_i helps predict x_j, over and above the degree to which x_j is already predicted by its own past and the past of variables other than x_i.

2.2 Frequency Domain VAR-GC: the Generalized Partial Directed Coherence (gPDC)

As noted above, environmental time series most often contain periodical components (diurnal and seasonal cycles) that can significantly mask the underlying causality structure in time domain. Removal of these periodic components might degrade causal inference [7]. In previous work [26] we have shown through comparison between time and frequency domain causality analysis of environmental time series that time-domain GC might result in several spurious causal links due to the presence of periodic components. Hence, the use of spectral-domain analysis enables change detection in certain frequency bands where for example the influence of trends (low frequency) or daily and seasonal cycle can be excluded.

The causal relation from x_i to x_j is described in the frequency domain via the generalized partial directed coherence (gPDC) [6]:

$$g\pi_{i\to j}(f) = \frac{\frac{1}{\sigma_{jj}}\overline{A}_{ji}(f)}{\sqrt{\sum_{k=1}^{m}\frac{1}{\sigma_{kk}^2}\left|\overline{A}_{ki}(f)\right|^2}}, \tag{4}$$

where $\overline{A}_{ij}(f), i, j = 1 \dots N$ are the elements of the matrix $\overline{A}(f) = I - A(f)$ where $A(f)$ is the Fourier transform of $A(r), r = 1, \dots, p$:

$$A(f) = \sum_{r=1}^{p} \mathbf{A}_r z^{-r}|_{z=e^{i2\pi f}}, \tag{5}$$

and σ_{ii}^2 are the diagonal entries of the residual covariance matrix Σ. The value of $g\pi_{i\to j}(f)$ represents the causality strength of x_i on x_j as compared to all of x_i's interactions to other variables. Nullity of $g\pi_{i\to j}(f)$ indicates absence of Granger causality from x_i to x_j at the normalized frequency f. Note that the normalizing term in the denominator in Eq. 4 is selected such that $0 \leq |g\pi_{i\to j}(f)| \leq 1$.

2.3 Event Detection and Attribution based on Time-Frequency Causality Analysis

A major issue in causal inference in environmental data is that the underlying system is most often a time-varying one. To deal with this issue, one can use for example an adaptive system such as the Kalman filter [2] or a sliding time window approach where the system is assumed to be time-invariant. In this paper we use the latter solution, i.e. we extract changes in the causal intensities by calculating the gPDC on selected time windows where the stationarity assumption is assumed to hold. To this end, for each time window, the model order is first estimated using Bayesian criterion and the model parameters are estimated using the LS method. The spectral causal effect values are then calculated using Eq. 4. The estimated spectral causal effect intensities serve as an approximation of the pairwise sub-models. The accuracy of this approximation depends on several

factors such as the sample size in the selected time window, the method used for model order estimation, as well as the procedure used for the identification of the model parameters.

The authors of [28] presented an extensive study on the asymptotic distribution of the partial directed coherence when applied to multichannel electroencephalographic data. They proved that for a stable stationary Gaussian VAR(p) process, the maximum likelihood estimator $|\pi_{i\to j}(f)|^2$ is consistent and asymptotically normally distributed if $|\pi_{i\to j}(f)|^2 \neq 0$, while $|\pi_{i\to j}(f)|^2 = 0$ indicates an absence of causal link between the time series. Hence the PDC measure provides means of comparing different strengths of connectivity between the observed time series. For non-zero values, the asymptotically normal behaviour degenerates into that of a mixture of χ variables allowing the computation of threshold for connectivity tests [28].

Under similar assumptions, the results of [28] can be directly extended to multivariate environmental data. Accordingly, comparison of the causal intensities in a certain time interval, and in the desired frequency band, with the average causal intensities (the mean value of the gPDC calculated over several realizations) can be utilized in principle to simultaneously detect and attribute anomalous events. The anomalous events here are meant to be those time windows where the causal effect intensities show, according to some distance measure, considerable deviation from the average causal intensities. The detected anomalous event can be then attributed to the variable(s) causing such deviation and is defined based on the change in its causal effect intensities on other variables.

Let us assume that we have L different realizations or different time intervals of the process $\mathcal{S}$ and let $g\pi^l_{i\to j}(f), i, j = 1, \ldots, N$ denote the causal intensities for realizations $l, l = 1, \ldots, L$. We define the average causal intensity of x_j on x_i at frequency f as

$$\overline{g\pi}_{i\to j}(f) = \frac{1}{L}\sum_{l=1}^{L} |g\pi^l_{i\to j}(f)|. \tag{6}$$

An event or time interval l is defined as anomalous if the causal intensities $g\pi^l_{i\to j}(f), i, j = 1, \ldots, N$ over a frequency band $f_1 \leq f \leq f_2$ is significantly higher than the average causal intensities $\overline{g\pi}_{i\to j}(f)$. In this paper we use the difference in the area under the two causal intensity curves for statistically significant values as a distance measure, however, other distance measures can be applied as well.

For statistical significance test against the null hypothesis of absence of a causal link, i.e. $|g\pi_{i\to j}(f)|^2 = 0$, we used two different tests. The first is the permutation test where we first generate permutations of the time series, calculate the gPDC values, and then take the maximum of the gPDC values over all permutations at each frequency. The second statistical significance test is the the Fourier Transform (FT) surrogate method [29] where time series surrogates can be generated by substituting the phase of the FT of the time series with a realization of uniformly distributed random variable in the range $[0, \pi)$ while keeping the amplitude of the FT the same. The FT surrogate method gives a

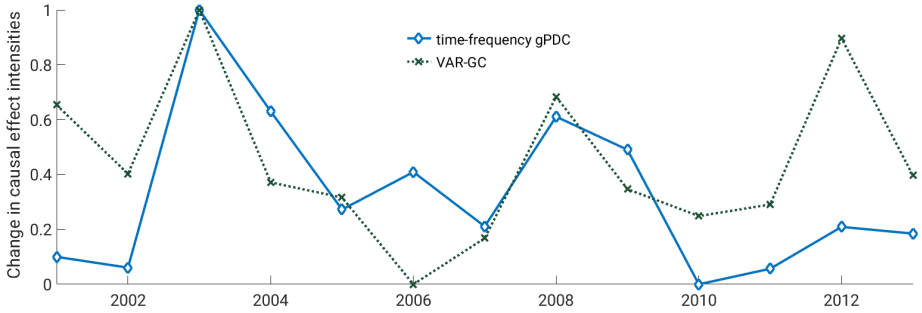

Fig. 2. The deviation of the causal effect intensities from the average ones during the months of August in years 2001-2013 summed over all variables and normalized to the [0-1] range. The peak in year 2003 correlates with the summer heatwave in France in August 2003.

more strict confidence interval when compared to other statistical significance tests of gPDC [10].

Different anomaly detection methods can then be utilized to classify events as normal or anomalous. For example, a simple approach for anomalous event detection would be to use a fixed percentile of the χ^2 distribution of the distance measure as a threshold to detect anomalous events.

3 Experimental Results

3.1 Event Detection and Attribution in Ecological Time Series

Understanding causal effect relationships are essential to understand ecosystem behaviour under climate change conditions. A particular cause of concern is the question of how ecosystem functioning (e.g. land-atmosphere exchange processes of CO2, water, and energy) are affected during unusual hydro-meteorological conditions [18]. This experiment aims to investigate the causal effect relationships between air temperature (T), vapour pressure deficit (VPD), latent energy (LE) and net ecosystem exchange (NEE). Experiments are performed on the real half-hourly meteorological observations and land flux eddy covariance data measured at the flux tower site of Puechabon-France spanning years 2001-2013 [22].

For the gPDC time-frequency analysis, we adopted a sliding time window approach of length 1440 (30 days by 48 samples/day) and followed the steps in Section 2.3: the model order and parameters are estimated for each time window, then the spectral causal effect values are calculated using Eq. 4. The summation of the absolute difference between the causal effect intensities of the system within the month of August over years 2001-2013 and the average causal intensities (average of the causal intensities in the month of August over 13

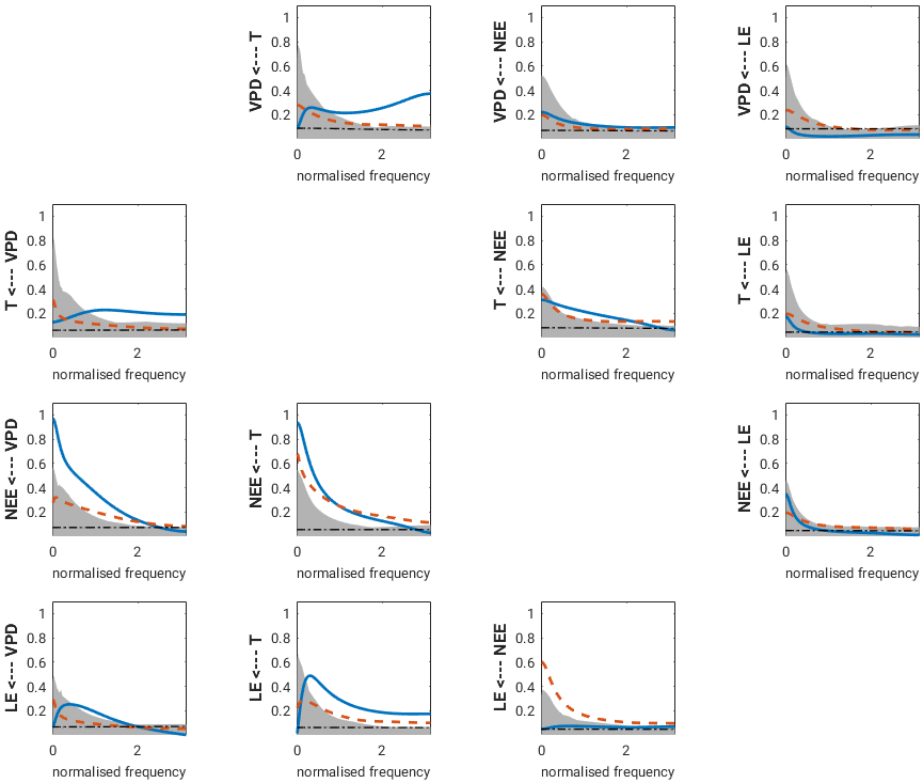

Fig. 3. Plots of the causal effect intensities $|g\pi_{i\to j}(f)|$ of the four variables T, VPD, NEE, LE measured at the flux tower site of Puechabon-France during the heatwave in August 2003 (solid blue line) when compared to the average causal intensities $\overline{g\pi}_{i\to j}(f)$ of similar summer period within years 2001-2013 (red dashed line). The threshold for statistical significance estimated using permutation test and the FT surrogate test is shown in the dashed-dotted line and gray area respectively.

years) over all variables is calculated. Note that in the calculation of this change we excluded the values of gPDC that are not statistically significant.

Similar experiment is performed using time domain VAR-GC as defined in Eq. 3 . Figure 2 shows the comparison between the proposed time-frequency gPDC-based method and the time domain VAR-GC. Both approaches shows clear peak in August 2003 which correlates with the historic heatwave in France in August 2003. For the VAR-GC, however, we can notice another clear peak in year 2012. Although we have no ground truth to present quantitative comparison between the two methods, to the best of our knowledge, no historical event has been recorded in August 2012.

In Figure 3, we compare the pairwise causal intensities of August 2003 with the average pairwise causal intensities of the months of August in years 2001-

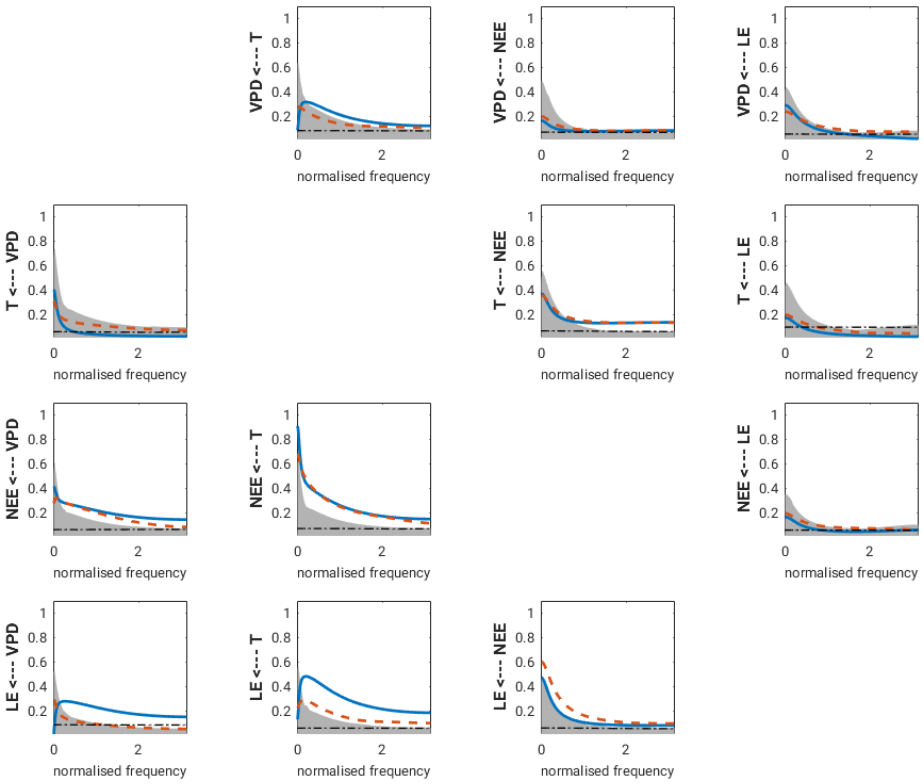

Fig. 4. Plots of the causal effect intensities $|g\pi_{i\to j}(f)|$ of the four variables T, VPD, NEE, LE measured at the flux tower site of Puechabon-France during during August 2002 (blue solid line) when compared to the average causal intensities $\overline{g\pi}_{i\to j}(f)$ of similar summer period within years 2001-2013 (red dashed line). The threshold for statistical significance estimated using permutation test and the FT surrogate test is shown in the dashed-dotted line and gray area respectively.

2013. It can be observed that there is considerable change in the causal intensity of T $\to$ VPD in the high frequency range which corresponds to a short term change (half an hour up to two hours); there is also clear increase in the causal intensity of VPD $\to$ NEE at the low frequency range (long term change) pointing towards an increase in water stress on ecosystem functioning; other causal effect intensities however remain the same. Comparing similar results for year 2002 (Figure 4), shows that while in a normal summer, such as in year 2002, the causal intensities match well with the average behaviour of the system, the ones in 2003 show clear deviation in the system dynamics from the average behaviour with T being the main driving variable.

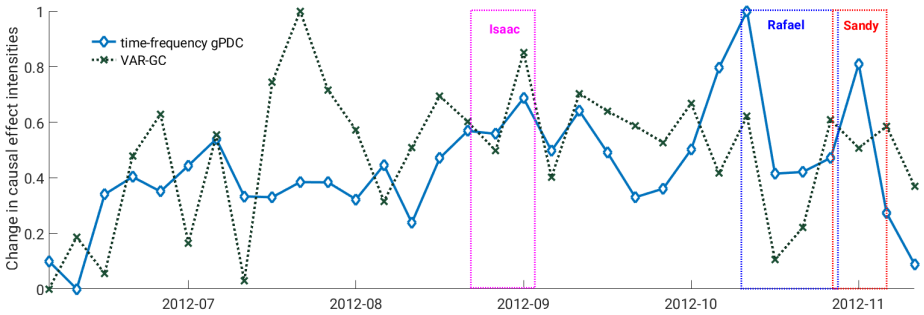

Fig. 5. Changes in the causal effect intensities of all the extracted segments within the six months period of the data (June 2012 until November 2012) when compared to the average causal intensities summed over all variables and normalized to the [0-1] range. For the time-frequency analysis using the generalized partial directed coherence (gPDC), the high change values correlate highly with the three hurricanes: Issac (from Aug. 21 to Sept. 03, 2012), Rafael (from Oct. 12 to Oct. 26, 2012), and Sandy (from Oct.22 to Nov. 02., 2012). The peak on the first week of July is due to a tropical storm. The rectangles show the start and end dates of the hurricanes. The dates of the markers correspond to the dates of the center of the segments, each of length 20 days.

3.2 Event Detection and Attribution in Marine Climate Time Series

In this example we study the causal effect relationships within three climate marine variables: sea level pressure (SLP), wind speed (W), and wave height (Hs). Time series data of a buoy located near the Bahamas in the Atlantic Sea (23.838 N, 68.333 W) were extracted from the National Data Buoy Center (http://www.ndbc.noaa.gov/) and used in this experiment (same data used in Figure 1). The time series comprises six months of hourly data, from June 2012 until November 2012. This period corresponds to the Atlantic hurricane season, which in that year was especially active [8].

For the gPDC spectral analysis, we used a sliding hamming window to extract segments of length 480 samples (20 days by 24 sample/day) with 25% overlap. Figure 5 shows the changes in the causal effect intensities of all extracted segments within the six months period of the data when compared to the average causal intensities for both the VAR-GC and the gPDC methods. The high change values of the proposed gPDC-based method correlate highly with the three historic hurricanes *Issac*, *Rafael* and *Sandy* of the year 2012 as shown in Figure 1. The peak on the first week of July is due to a tropical storm.

It can be even visually verified that applying any fixed percentile threshold to the VAR-GC and gPDC-based causal intensities-change measures (Figure 5) for the detection of anomalous events would reveal that the proposed method has higher detection accuracy of the hurricanes and less false positive detection.

Figure 6 shows the plots of the pairwise causal effect relationships at different time windows when compared to the average causal intensities. For a time

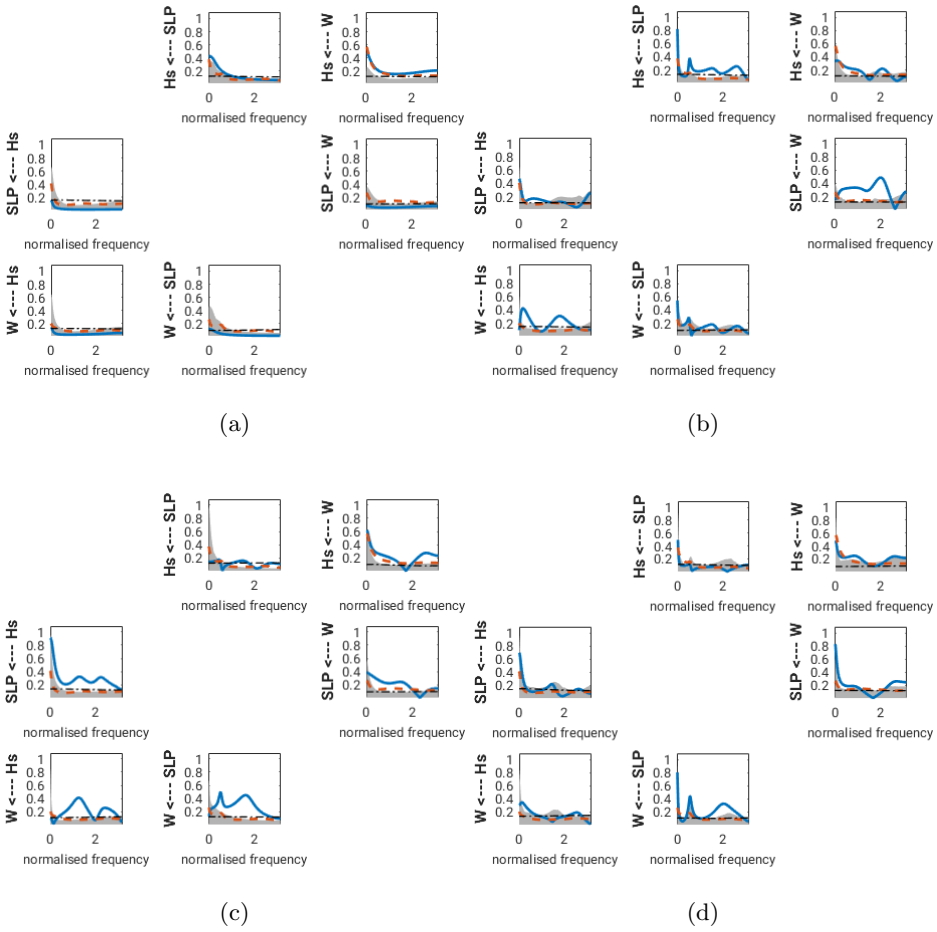

Fig. 6. Plots of the causal effect intensities $|g\pi_{i\to j}(f)|$ (solid blue line) of the three marine climate variables: sea level pressure (SLP), wave height (Hs), and wind speed (W) during normal sea conditions (a), hurricane *Issac* (from Aug. 21 to Sept. 03, 2012) (b), hurricane *Rafael* (from Oct. 12 to Oct. 26, 2012)(c), and hurricane *Sandy* (from Oct.22 to Nov. 02., 2012) (d). The red dashed line shows the average causal intensities $\overline{g\pi}_{i\to j}(f)$ over all segments within the six months period of the data (June 2012 until November 2012). The threshold for statistical significance estimated using permutation test and the FT surrogate test is shown in the dashed-dotted line and gray area respectively.

window with normal sea conditions, the causal effect relationships are either statistically insignificant or match well with the average causal intensities. For hurricane *Isaac* however, there is an increase in the causal effect intensities, particularly the causal effect of W on other variables. During hurricanes *Rafael* and *Sandy*, we can notice an increase in the causal effect of SLP on other variables. Interestingly, this increase is particularly high at the frequency that corresponds to the semi-diurnal (twice-daily) cycle of the atmospheric pressure. The differences in the causal effect patterns of these three hurricanes can be related to the trajectories of the hurricanes compared to the location of the buoy where the data were recorded: *Rafael* and *Sandy* passed much closer to the Bahamas than *Isaac*.

It should be noted that the proposed attribution method is based on the assumption that there are no unobserved drivers. To account for hidden drivers, the presented approach can be extended by using a state-space model with latent variables instead of the VAR model. For highly nonlinear causal effect relationship with hidden confounding, the concept of detecting anomalous event based on the changes in the causal effect relationships is currently being explored by the authors using deep learning approach along with domain knowledge integration [30].

4 Conclusions and Future Work

In this paper, we have presented an attribution scheme for changes in environmental data based on the analysis of the causal effect relationships in multivariate environmental time series. The coupling between the used variables is assumed to be well represented by a vector autoregressive (VAR) model. The causal effect relationships are extracted using the generalized partial directed coherence. Through some representative examples in environmental systems, we have shown that an anomalous event can be detected as the one where the causal intensities between the variables differ according to some statistical measure from the average causal intensities. Moreover, the analysis of the causal effect patterns allows for understanding these events and defining the time scale on which changes occur. Current research work is directed towards the development of methods that are able to integrate the presented approach with domain knowledge for improved anomalous event detection and classification in a spatiotemporal context as well as towards using adaptive time window selection instead of the fixed-size window used in this study.

Acknowledgments

The authors thank the Carl Zeiss Foundation for the financial support within the scope of the program line "Breakthroughs: Exploring Intelligent Systems" for "Digitization — explore the basics, use applications". This work used eddy covariance data acquired and shared by the FLUXNET community.

References

1. The mvgc multivariate granger causality toolbox: A new approach to granger-causal inference. Journal of Neuroscience Methods **223**, 50 – 68 (2014)
2. A., W.E., T., N.A.: Dual Extended Kalman Filter Methods, chap. 5, pp. 123–173. Wiley-Blackwell (2002). https://doi.org/10.1002/0471221546.ch5
3. Akaike, H.: A new look at the statistical model identification. IEEE Transactions on Automatic Control **19**(6), 716–723 (December 1974). https://doi.org/10.1109/TAC.1974.1100705
4. Anderson, T.: The Statistical Analysis of Time Series. Wiley Classics Library, Wiley (1994)
5. Attanasio, A. Pasini, A., Triacca, U.: Granger causality analyses for climatic attribution. Atmospheric and Climate Sciences **3**(4), 515–522 (2013). https://doi.org/10.4236/acs.2013.34054
6. Baccalá, L.A., Sameshima, K., Takahashi, D.: Generalized partial directed coherence. In: Digital Signal Processing, 2007 15th International Conference on. pp. 163–166. IEEE (2007)
7. Barnett, L., Seth, A.K.: Behaviour of granger causality under filtering: Theoretical invariance and practical application. Journal of Neuroscience Methods **201**(2), 404–419 (2011). https://doi.org/https://doi.org/10.1016/j.jneumeth.2011.08.010
8. Barz, B., Guanche, Y., Rodner, E., Denzler, J.: Maximally divergent intervals for extreme weather event detection. In: MTS/IEEE OCEANS Conference Aberdeen. pp. 1–9 (2017). https://doi.org/10.1109/OCEANSE.2017.8084569
9. Eichler, M.: Graphical modelling of multivariate time series. Probability Theory and Related Fields **153**(1), 233–268 (Jun 2012). https://doi.org/10.1007/s00440-011-0345-8
10. Faes, L., Porta, A., Nollo, G.: Testing frequency-domain causality in multivariate time series. IEEE Transactions on Biomedical Engineering **57**(8), 1897–1906 (2010)
11. Faghmous, J.H., Kumar, V.: A big data guide to understanding climate change: The case for theory-guided data science. Big data **2 3**, 155–163 (2014)
12. Feldhoff, J., Donner, R.V., Donges, J.F., Marwan, N., Kurths, J.: Detection of coupling directions by means of inter-system recurrence networks. Phys. Lett. A **376**, 3504–3513 (2012)
13. Frank, P.: Analytical and qualitative model-based fault diagnosis a survey and some new results. European Journal of Control **2**(1), 6 – 28 (1996). https://doi.org/https://doi.org/10.1016/S0947-3580(96)70024-9
14. Geweke, J.: Measurement of linear dependence and feedback between multiple time series. Journal of the American statistical association **77**(378), 304–313 (1982)
15. Granger, C.W.J.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica **37**(3), 424–438 (1969), http://www.jstor.org/stable/1912791
16. Granger, C.W.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society pp. 424–438 (1969)
17. Haykin, S.: Adaptive Filter Theory (3rd Ed.). Prentice-Hall, Inc., Upper Saddle River, NJ, USA (1996)
18. Mahecha, M.D., Gans, F., Sippel, S., Donges, J.F., Kaminski, T., Metzger, S., Migliavacca, M., Papale, D., Rammig, A., Zscheischler, J.: Detecting impacts of extreme events with ecological in situ monitoring networks. Biogeosciences **14**(18), 4255–4277 (2017). https://doi.org/10.5194/bg-14-4255-2017

19. Marinazzo, D., Liao, W., Chen, H., Stramaglia, S.: Nonlinear connectivity by granger causality. NeuroImage **58**(2), 330 – 338 (2011). https://doi.org/https://doi.org/10.1016/j.neuroimage.2010.01.099
20. Papagiannopoulou, C., Miralles, D.G., Decubber, S., Demuzere, M., Verhoest, N.E.C., Dorigo, W.A., Waegeman, W.: A non-linear granger-causality framework to investigate climate–vegetation dynamics. Geoscientific Model Development **10**(5), 1945–1960 (2017). https://doi.org/10.5194/gmd-10-1945-2017
21. Peters, J., Janzing, D., Schölkopf, B.: Elements of Causal Inference - Foundations and Learning Algorithms. Adaptive Computation and Machine Learning Series, The MIT Press, Cambridge, MA, USA (2017)
22. Rambal, S., Joffre, R., Ourcival, J.M., Cavender-Bares, J., Rocheteau, A.: The growth respiration component in eddy co2 flux from a quercusilex mediterranean forest. Global Change Biology **10**(9), 1460–1469 (2004). https://doi.org/10.1111/j.1365-2486.2004.00819.x
23. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat: Deep learning and process understanding for data-driven earth system science. Nature pp. 195–204 (2019). https://doi.org/10.1038/s41586-019-0912-1
24. Schwarz, G.: Estimating the dimension of a model. Ann. Statist. **6**(2), 461–464 (03 1978). https://doi.org/10.1214/aos/1176344136
25. Seth, A.K., Barrett, A.B., Barnett, L.: Granger causality analysis in neuroscience and neuroimaging. Journal of Neuroscience **35**(8), 3293–3297 (2015). https://doi.org/10.1523/JNEUROSCI.4399-14.2015
26. Shadaydeh, M., Garcia, Y.G., Mahecha, M., Reichstein, M., Denzler, J.: Causality analysis of ecological time series: a time-frequency approach. In: In C. Chen, D. Cooley, J. Runge, E. Szekely (Eds.), Climate Informatics Workshop 2018. pp. 111–114 (2018)
27. Solo, V.: State-space analysis of granger-geweke causality measures with application to fmri. Neural Computation **28**(5), 914–949 (2016). https://doi.org/10.1162/NECO_a_00828, pMID: 26942749
28. Takahashi, D.Y., Baccal, L.A., Sameshima, K.: Connectivity inference between neural structures via partial directed coherence. Journal of Applied Statistics **34**(10), 1259–1273 (2007). https://doi.org/10.1080/02664760701593065
29. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., Farmer, J.D.: Testing for nonlinearity in time series: the method of surrogate data. Physica D: Nonlinear Phenomena **58**(1), 77 – 94 (1992). https://doi.org/https://doi.org/10.1016/0167-2789(92)90102-S
30. Trifunov, V.T., Shadaydeh, M., Runge, J., Eyring, V., Reichstein, M., Denzler, J.: Nonlinear causal link estimation under hidden confounding with an application to time series anomaly detection. In: German Conference on Pattern Recognition (2019)
31. Zhong, M., Xue, T., Ding, S.X.: A survey on model-based fault diagnosis for linear discrete time-varying systems. Neurocomputing **306**, 51 – 60 (2018). https://doi.org/https://doi.org/10.1016/j.neucom.2018.04.037