IEICE Proceeding Series

Identifying nonlinearities by time-reversal asymmetry of vertex properties in visibility graphs

Reik V. Donner, Jonathan F. Donges

Vol. 1 pp. 435-438
Publication Date: 2014/03/17
Online ISSN: 2188-5079

2012 International Symposium on Nonlinear Theory and its Applications
NOLTA2012, Palma, Majorca, Spain, October 22-26, 2012

Identifying nonlinearities by time-reversal asymmetry of vertex properties in visibility graphs

Reik V. Donner[†] and Jonathan F. Donges[†,‡]

†Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany
‡Department of Physics, Humboldt University, Newtonstr. 15. 12489 Berlin, Germany
Email: reik.donner@pik-potsdam.de, donges@pik-potsdam.de

Abstract—The absence of time-reversal symmetry is a fundamental property of nonlinear time series. Here, we propose a novel set of statistical tests for time series reversibility based on visibility graphs. Specifically, we statistically compare the distributions of time-directed variants of some common graph-theoretical measures like degree and local clustering coefficient. Unlike other tests for reversibility, our method has the important advantage of not requiring the construction of surrogate time series. We illustrate its potentials for time series from paradigmatic model systems with known time-reversal properties as well as some real-world paleoclimate data.

1. Introduction

Motivated by the advent of complex network theory during the last decade [1, 2], a number of techniques for network-based time series analysis have been proposed recently [3]. Among other successful approaches [4, 5, 6, 7], the application of visibility graph (VG) methods to time series has recently attracted specific interest [8, 9]. Originally, VGs have been introduced for the analysis of mutual visibility relationships between points and obstacles in two-dimensional landscapes in the framework of computational geometry. Lacasa *et al.* [8] adopted this concept to the analysis of structures in scalar, univariate time series. One important aspect of the resulting network structures is that their degree distributions can be used for obtaining a classification of time series and are in particular closely related with eventual fractal and multifractal properties of the underlying data. This relationships makes VGs and related concepts promising candidates for applications to observational time series from various fields of research [10].

In this work, we demonstrate that VGs can be used for statistically testing the reversibility properties of observational time series. For this purpose, we define time-directed variants of vertex characteristics, the distributions of which can be statistically evaluated and compared. The remainder of this paper is accordingly organized as follows: The basic principles behind the construction of VGs are reviewed in Sect. 2. Subsequently, we introduce time-directed variants of the degree and local clustering coefficient as two particularly prominent vertex characteristics. Conceptually related approaches are briefly discussed. A systematic analysis of the resulting probability distributions for different linear (time-reversible) as well as non-linear (irreversible) processes is presented in Sect. 3 to illustrate the power of the proposed approach. Motivated by the very promising results obtained for these model systems, we furthermore apply our method to a well-studied paleoclimate time series describing the temperature variations over Greenland during the last glacial cycle. Finally, the main conclusions of our research are summarized in Sect. 4.

2. Methodology

2.1. Visibility Graphs

VGs provide a simple mapping from the time series to the network domain by exploiting certain convexity characteristics of scalar (univariate) time series $\{x(t_i)\}_{i=1}^N$. Here, each observation $x(t_i)$ is assigned a vertex i of a complex network, which is uniquely defined by the time of observation t_i. Two vertices i and j are linked by an edge (i, j) iff the convexity condition [8]

$$x_k < x_j + (x_i - x_j)\frac{t_j - t_k}{t_j - t_i} \tag{1}$$

holds for all vertices k with $t_i < t_k < t_j$ (see Fig. 1). That is, the corresponding adjacency matrix completely describing the VG as a simple undirected and unweighted network reads

$$A_{ij}^{(VG)} = A_{ji}^{(VG)} = \prod_{k=i+1}^{j-1} \Theta\left(x_j + (x_i - x_j)\frac{t_j - t_k}{t_j - t_i} - x_k\right), \tag{2}$$

where $\Theta(\cdot)$ is the Heaviside function defined in the usual way. The corresponding construction algorithm implies that VGs are spatial networks with vertices embedded on the one-dimensional time axis, which leads to the emergence of boundary effects on commonly studied network measures at the respective ends of the time series [10]. In turn, we emphasize that VGs have the important advantage that one does not need to require a uniform sampling of observations in time, which is a major problem of many existing (linear and nonlinear) methods of time series analysis when working with real-world observational data, e.g. in paleoclimatology or astrophysics. This fact makes VGs

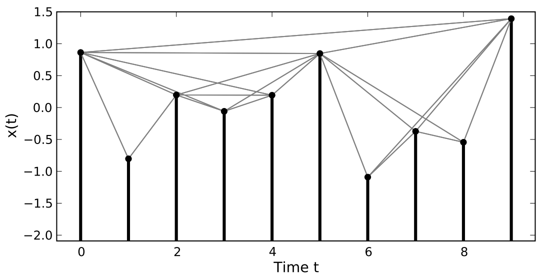

Figure 1: Schematic illustration of the construction principle of a VG.

a promising candidate for analyzing the properties of such records.

As a noteworthy variant of the basic VG approach, the alternative concept of horizontal visibility graphs (HVGs) has been recently introduced [9]. For a given time series, the vertex sets of VG and HVG are the same, however, the edge set of the HVG is defined by the mutual horizontal visibility of two observations x_i and x_j, i.e., there is an edge (i, j) iff $x_k < \min(x_i, x_j)$ for all k with $t_i < t_k < t_j$:

$$A_{ij}^{(HVG)} = A_{ji}^{(HVG)} = \prod_{k=i+1}^{j-1} \Theta\left(x_i - x_k\right) \Theta\left(x_j - x_k\right). \quad (3)$$

We note that although the HVG algorithm appears somewhat simpler, VG and HVG essentially capture the same properties of the system under study. Specifically, the HVG is a subgraph of the VG with the same vertex set, but only a subset of edges contained in the VG. To put it differently, the VG is invariant with respect to the superposition of linear trends whereas the HVG is not.

2.2. Standard and time-directed vertex properties

Complex networks can be quantitatively characterized and possibly classified by various properties acting on the local (vertex- or edge-based), mesoscopic, or global network scale. In this work, we will focus on two well-studied vertex characteristics.

2.2.1. Vertex degree

The *degree* of a given vertex i,

$$k_i = \sum_{j=1}^{N} A_{ij}, \quad (4)$$

measures the number of edges incident to i. For directed networks (which have, in contrast to the VGs studied here, an asymmetric adjacency matrix, i.e., $\mathbf{A} \neq \mathbf{A}^T$), one has to distinguish the *in-degree* (the number of edges entering i) and *out-degree* (the number of edges leaving i):

$$k_i^{in} = \sum_{j=1}^{N} A_{ji} \quad \text{and} \quad k_i^{out} = \sum_{j=1}^{N} A_{ij} \quad (5)$$

In a recent work, Lacasa *et al.* [11] studied directional HVGs defined as

$$A_{ij}^{(DHVG)} = \prod_{k=i+1}^{j-1} \Theta\left(x_i - x_k\right) \Theta\left(x_j - x_k\right) \Theta(t_j - t_i), \quad (6)$$

which have directed edges for pairs of vertices (i, j) for $t_j > t_i$ only. It has been shown that there are differences between the distributions of in- and out-degrees obtained from the directional HVGs of nonlinear processes that can be used for quantifying the degree of time-reversal asymmetry.

In this paper, we propose a slightly different approach by introducing two related properties based on the standard VG algorithm: the *retarded and advanced degrees*

$$k_i^r = \sum_{j<i} A_{ij} \quad \text{and} \quad k_i^a = \sum_{j>i} A_{ij}. \quad (7)$$

Both characteristics can be considered as the respective in- and out-degrees of a directional VG where all edges follow the arrow of time. In general, we note that this approach explicitly makes use of the fact that VGs are "embedded" on the one-dimensional time axis, such that their local structure is intrinsically interwoven with the direction of time. We furthermore emphasize that considering VGs instead of HVGs may eventuelly lead to a better detectability of time-reversal asymmetries, since VGs incorporate slightly more moderate restrictions to the graph by their construction. However, a detailed comparison of the performance of VGs and HVGs for the considered purpose is beyond the scope of this work.

2.2.2. Local clustering coefficient

The retarded and advanced degrees measure a rather coarse property of the system under study, that basically corresponds to the distance to the next observation, the magnitude of which (corrected for possible linear trends) exceeds that of the considered point of observation in the past and future, respectively. Thus, it is possible that there are subtle manifestations of time-series irreversibility that cannot be statistically detected based on the distributions of both measures. As a possible solution, we propose using higher-order characteristics involving three-point relationships between vertices of a VG.

One such measure is the *local clustering coefficient*

$$C_i = \frac{\sum_{j,k=1}^{n} A_{ij} A_{ik} A_{jk}}{k_i(k_i - 1)/2}, \quad (8)$$

which describes the neighborhood structure of i. Specifically, it is a measure of transitivity characterizing to what extent the linkage structure of a vertex i is transitive (i.e., how often $A_{ij} = A_{ik} = 1$ also imply $A_{jk} = 1$). In order to detect possible signatures of time-reversal asymmetry, we define in analogy to the degree the *retarded and advanced local clustering coefficients* as

$$C_i^r = \frac{\sum_{j,k<i} A_{ij} A_{ik} A_{jk}}{k_i^r(k_i^r - 1)/2} \quad \text{and} \quad C_i^a = \frac{\sum_{j,k>i} A_{ij} A_{ik} A_{jk}}{k_i^a(k_i^a - 1)/2} \quad (9)$$

We note that both properties are equivalent to the in- and out-clustering coefficients of directed networks [12] when applied to directional VGs.

In general, we note that there are many other vertex characteristics one could study in a similar way. In this work, we restrict our attention to the two aforementioned characteristics for two reasons: First, as it has been shown recently [10], vertex properties suffer from boundary effects implying that there is a systematic bias of their distribution towards smaller values particularly for short time series. This bias is still quite moderate for degree and local clustering coefficient, but becomes stronger for other (in particular, path-based) network measures such as closeness and betweenness. Second, we emphasize that the two considered characteristics are those that are most easily computable and have still a rather intuitive interpretation in terms of VGs (see [10] for a detailed discussion).

2.3. Comparison of distribution functions

As we have already stated above, we conjecture that differences in the distributions of retarded and advanced vertex properties are a manifestation of a statistical time-reversal asymmetry of the investigated time series with respect to the respective properties. There are different ways for characterizing such differences. Lacasa *et al.* [11] used the Kullback-Leibler distance between both distribution functions as a measure for the degree of time-reversal asymmetry. In the sense of a statistical test for time-reversal asymmetry, we alternatively propose here using the Kolmogorov-Smirnov statistics (i.e., the maximum difference between the two cumulative distribution functions). The main advantage of the latter approach that it allows directly giving P-values for possibly rejecting the hypothesis of equal distributions. These values are universal and distribution-free in the limit of $N \to \infty$. Hence, we can use them directly as P-values for testing reversibility and do *not* need to construct surrogate time series as in other reversibility tests.

In summary, the proposed strategy allows constructing simple tests for time-reversal asymmetry of certain local VG properties. Specifically, if we can statistically reject the null hypothesis of equal distributions of retarded and advanced properties for a given record with sufficiently high confidence, we can conclude that the underlying time series is irreversible, which necessarily implies the presence of a nonlinear process (since linear processes are reversible with respect to their statistical properties).

3. Examples

We first illustrate the potentials of the proposed method for two simple model systems. On the one hand, we will consider a linear-stochastic first-order autoregressive process

$$x_t = \alpha x_{t-1} + \xi_t \qquad (10)$$

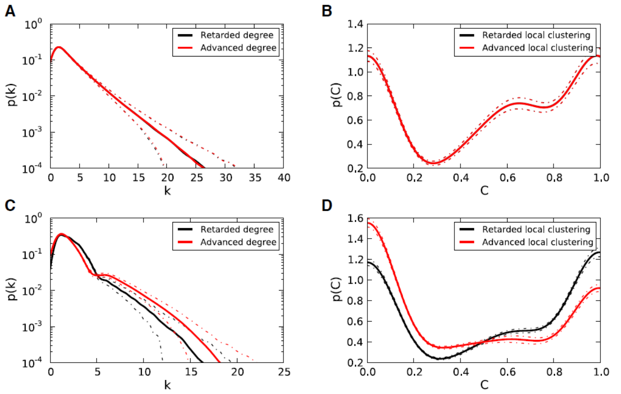

Figure 2: Distributions of retarded/advanced (A,C) degree $k_i^{r,a}$ and (B,D) local clustering coefficient $C_i^{r,a}$ for model systems: (A,B) linear first-order autoregressive process and (C,D) nonlinear Hénon map (first component). Time series of length $N = 500$ have been used for estimating the probability density functions (PDF) with a kernel density estimator. The mean (solid lines) and standard deviation (dashed lines) of the PDFs have been computed based on an ensemble of $M = 1,000$ realizations with random initial conditions for both model systems.

with $\alpha = 0.5$ and the additive noise term ξ_t taken as independent realizations of a Gaussian random variable with zero mean and unit variance. On the other hand, we consider the first (x) component of the nonlinear-deterministic Hénon map

$$x_t = A - x_{t-1}^2 + B y_{t-1}, \quad y_t = x_{t-1} \qquad (11)$$

with $A = 1.4$ and $B = 0.3$. In both cases, we use ensembles of realizations of the respective processes with random initial conditions and discard the first 1,000 points of each resulting time series to avoid possible transients.

As expected, for the linear (reversible) process, the empirical distributions of retarded/advanced degree and local clustering coefficient collapse onto each other (Fig. 2A,B). Consequently, the null hypothesis of reversibility is never rejected by the test based on degree (Fig. 2A) and only rarely rejected by the clustering-based test well below the expected false rejection rate of 5% (Fig. 2B). In contrast, for the irreversible Hénon map the distributions of retarded and advanced VG measures appear distinct already by visual inspection (Fig. 2C,D). In accordance with this observation, the null hypothesis of reversibility is nearly always (Fig. 2C) or always (Fig. 2D) rejected.

Motivated by these results for simple model systems, we finally apply our method to detecting signatures of time-irreversibility in real-world paleoclimate data. Note again that such data are typically characterized by non-uniform sampling, which is a challenge for most methods of time series analysis. In turn, the use of VG-based methods does not explicitly require uniform sampling of observations, making this method a promising tool for the explorative analysis of such data. As a particular example, we consider the well-studied δ^{18}O isotope record from the GISP2 ice core from Greenland [13], which can be interpreted as a proxy for paleo-temperatures and covers the present interglacial period (Holocene) as well as the vast part of the last glacial. Since warm and cold periods likely display different dynamical patterns, we apply our method separately to both respective parts of the record, i.e., $N = 824$ data points for the Holocene (11.65 kyr before present (BP) until today) and $N = 566$ data points for the last glacial (110.98 – 11.65 kyr BP, note the different sampling rates in both periods due to the compactification of older ice deposits).

As a result, we find that the null hypothesis of reversibility can be rejected for the last glacial with high confidence using both degrees and local clustering coefficients as discriminators. In order to further support this result, we applied an leave-K-out cross-validation procedure by randomly removing 20% of the data from the record and repeating the computations afterwards. It turned out that that for the thus bootstrapped data, the null hypothesis could be rejected in 93 (degree) and 82 (local clustering coefficient) out of 100 realizations. The apparent irreversibility of the temperature variability during the glacial is most probably due to the marked presence of multiple Dansgaard-Øschger events that are characterized by a fast warming followed by a slow cooling. In turn, such asymmetric events (indicating the presence of a strongly nonlinear dynamics) are largely missing during the Holocene. As a consequence, the null hypothesis of time-reversibility could not be rejected by our tests for this time interval (for all realizations when doing cross-validation).

4. Conclusions

We have proposed a new test for reversibility of scalar time series based on visibility graphs. Our approach has two important advantages in comparison with existing tests: it can be directly applied to time series with non-uniform sampling (such as paleoclimate records) and does not require the construction of surrogate data, but directly supplies a P-value for the associated null hypothesis. Time-directed versions of the degree and local clustering coefficient have been shown to serve as powerful discriminators between reversible linear, and irreversible nonlinear dynamics. Our results suggest potentials for future applications to many fields of research. Methodological questions that will be studied in future work include a comparison of the performance of our tests based on standard as well as horizontal visibility graphs, the use of different network characteristics, and a detailed study of the possible impact of sampling on the outcomes of our method.

Acknowledgments

This work has been financially supported by the Leibniz Society (project ECONS), the German National Academic Foundation, and the joint German-Portuguese project "Temporal changes in sea-level variability: From local to regional scales" (DAAD project ID 50750929).

References

[1] R. Albert, A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol.74, pp.47–97, 2002.

[2] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol.45, pp.167–256, 2003.

[3] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang, J. Kurths, "Recurrence-based time series analysis by means of complex network methods," *Int. J. Bifurcation Chaos,* vol.21, pp.1019–1048, 2011.

[4] J. Zhang, M. Small, "Complex network from pseudoperiodic time series: Topology versus dynamics," *Phys. Rev. Lett.,* vol.96, 238701, 2006.

[5] X. Xu, J. Zhang, M. Small, "Superfamily phenomena and motifs of networks induced from time series," *Proc. Natl. Acad. Sci. USA*, vol.105, pp.19601–19605, 2008.

[6] Y. Yang, H. Yang, "Complex network-based time series analysis," *Physica A*, vol.387, 1381–1386, 2008.

[7] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, J. Kurths, "Recurrence networks - a novel paradigm for nonlinear time series analysis," *New J. Phys.*, vol.12, 033025, 2010.

[8] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J. C. Nuno, "From time series to complex networks: The visibility graph," *Proc. Natl. Acad. Sci. USA*, vol.105, pp.4972–4975, 2008.

[9] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, "Horizontal visibility graphs: Exact results for random time series," *Phys. Rev. E*, vol.80, 046103, 2009.

[10] R. V. Donner, J. F. Donges, "Visibility graph analysis of geophysical time series: Potentials and possible pitfalls," *Acta Geophys.*, in press.

[11] L. Lacasa, A. Nuñez, E. Roldán, J. M. R. Parrondo, B. Luque, "Time series irreversibility: a visibility graph approach," arXiv:1108.1691, 2011.

[12] G. Fagiolo, "Clustering in Complex Directed Networks," *Phys. Rev. E*, vol.76, 026107, 2007.

[13] P. M. Grootes, M. Stuiver, "Oxygen 18/16 variability in Greenland snow and ice with 10^{-3}- to 10^{5}-year time resolution," *J. Geophys. Res.*, vol.102(C12), pp.26455-26470, 1997.