PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: Jacob R, Harikrishnan KP, Misra R, Ambika G. 2019 Weighted recurrence networks for the analysis of time-series data. *Proc. R. Soc. A* **475**: 20180256. http://dx.doi.org/10.1098/rspa.2018.0256

Received: 17 April 2018
Accepted: 3 January 2019

Subject Areas:
chaos theory, complexity

Keywords:
weighted recurrence networks, node strength distribution, variable star analysis

Author for correspondence:
K. P. Harikrishnan
e-mail: kp.hk05@gmail.com

Weighted recurrence networks for the analysis of time-series data

Rinku Jacob[1], K. P. Harikrishnan[1], R. Misra[2] and G. Ambika[3]

[1]Department of Physics, The Cochin College, Cochin 682 002, India
[2]Inter University Centre for Astronomy and Astrophysics, Pune 411 007, India
[3]Indian Institute of Science Education and Research, Tirupati 517507, India

KPH, 0000-0001-9209-5921; GA, 0000-0001-7303-9692

Recurrence networks (RNs) have become very popular tools for the nonlinear analysis of time-series data. They are unweighted and undirected complex networks constructed with specific criteria from time series. In this work, we propose a method to construct a 'weighted recurrence network' from a time series and show that it can reveal useful information regarding the structure of a chaotic attractor which the usual unweighted RN cannot provide. Especially, a network measure, the node strength distribution, from every chaotic attractor follows a power law (with exponential cut off at the tail) with an index characteristic to the fractal structure of the attractor. This provides a new class among complex networks to which networks from all standard chaotic attractors are found to belong. Two other prominent network measures, clustering coefficient and characteristic path length, are generalized and their utility in discriminating chaotic dynamics from noise is highlighted. As an application of the proposed measure, we present an analysis of variable star light curves whose behaviour has been reported to be strange non-chaotic in a recent study. Our numerical results indicate that the weighted recurrence network and the associated measures can become potentially important tools for the analysis of short and noisy time series from the real world.

1. Introduction

The analysis of time-series data using complex network measures has become an important area of research over

THE ROYAL SOCIETY PUBLISHING

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 475: 20180256

the last two decades [1]. This graph theoretic approach to nonlinear time series analysis has several advantages over the conventional approach based on measures like dimension and Lyapunov exponent, especially in the characterization of the structural properties of the underlying attractor. This approach involves first transforming the time series into a complex network and analysing it using different network measures. Several methods [2–4] have been proposed in the literature to transform a time series into a complex network, with each of them finding application in particular contexts, to address the complementary features of the time series not obtained from the conventional approach based on measures like dimension and entropy.

A simple and direct method to convert a time series to a complex network is using the property of recurrence [5] of every dynamical system and the resulting network is called recurrence network (RN) [6], which is the focus of our study in this paper. To transform the time series into an RN, it is first embedded in a multi-variate state space of dimension M using the time delay coordinates [7]. Every point on the reconstructed attractor is then identified as a node and a recurrence threshold (ϵ) is set to define the connection between two nodes. Two nodes are considered to be connected if the corresponding points on the attractor are separated within the limit of this threshold. From the construction, it is clear that the RN is an unweighted and undirected network with the elements of the adjacency matrix $A_{\imath\jmath}$ either 1 or 0 depending on whether two nodes are connected or not. Once constructed, an array of statistical measures [8] can be defined from the RN that can characterize the structural properties of the attractor underlying the time series [9].

Though the analyses reported so far in the literature have been mostly confined to such unweighted networks, there have also been a few attempts in the past to use weighted network and recurrence plot measures for the analysis of time series for specific practical applications [10,11]. However, none of these works have used the weighted recurrence network (WRN) approach and the related measures for the analysis of time series. Recently, Sun *et al.* [12] have stressed the importance of weighted networks in the analysis of time series from dynamical systems. The authors, however, use a different approach via a sliding window and symbolic scheme to construct the weighted network from time series and illustrate the potential of their approach in characterizing the global properties of the attractor and dynamical transitions, using measures derived from the network. There have also been many recent advances on complex network analysis using time series derived from a wide range of real-world systems [13–15]. A recent review of the existing methods and results can be found in [16]. Moreover, weighted networks with weights assigned to links in different ways have been applied as models to analyse many real systems [17].

In this paper, we show that the method of RN can be generalized to construct WRN from a time series. We introduce a new method of assigning weight to a link between two nodes. Our analysis presented here reveals some novel structural properties associated with chaotic attractors on the one hand, and some potential practical applications of WRN on the other. Especially, we are able to identify the WRNs from different chaotic attractors as a single class with similar node strength distribution which may open up a new window in the study of complexity of dynamical systems through time series. Our paper is organized as follows: The details regarding the selection of recurrence threshold and the scheme for constructing the WRN are presented in the next section. In §3, important network measures derived from the WRN and their relevance in characterizing the structural properties of chaotic attractors are discussed. An analysis of variable star data is presented in §4 and the paper is concluded in §5.

2. Weighted recurrence networks from time series

(a) Selection of recurrence threshold

The recurrence threshold, ϵ, is a crucial parameter since the characteristic properties of the RN depend on its value. In general, for each embedded attractor reconstructed from a time series, the value of ϵ has to be determined separately as it varies with the size of the attractor. (Since RNs tend to random geometric graphs for large ϵ, the percolation threshold depends on the network size, an aspect not much explored in the literature). Two criteria are usually employed [6,8] to select ϵ. The first and the primary one is that there should be a giant component for the resulting RN which sets a lower bound for ϵ. In order to ensure that the network is not overconnected, the upper bound for ϵ is set such that the link density (the ratio of actual connections to all possible connections in a network of N nodes) is only a small fixed fraction of the maximum possible value. This provides a small range $\Delta\epsilon$ of suitable threshold for each system where the resulting network is considered to be a proper network representation of the time series.

Recently, we have proposed a scheme [18] where we tried to fix a small uniform range $\Delta\epsilon$ for choosing the threshold for time series from different chaotic systems. For this, we first transform the time series to a uniform deviate so that the size of the attractor always remains within the unit cube. To find the lower bound of ϵ, we use the standard criterion that the network turns into a single giant component. The upper bound is determined by the condition that the network is not overconnected. However, instead of fixing the link density, we apply a criterion that the characteristic path length (CPL, that describes the global connectivity of the network) of RN from chaotic time series is significantly different from that of white noise. Though this condition appears subjective, we are able to fix an empirical upper bound for $\Delta\epsilon$ with this. We also find that, due to the uniform deviate transformation, the scheme provides (as an empirical result) an approximately identical range of $\Delta\epsilon$ for constructing RN from different chaotic time series for a given embedding dimension M if the number of nodes N in the network is less than 10 000.

For example, for $M = 3$, the value of the recurrence threshold ϵ in our scheme is taken as 0.1. The exact emergence of giant component for some systems is found to be at a slightly lower value, but none greater than 0.1. In other words, the RNs constructed from standard chaotic time series that we have analysed so far possess a giant component for $\epsilon = 0.1$ for $M = 3$. Moreover, the CPL and the clustering coefficient (CC) of RN from chaotic systems significantly differ from that of white noise at this value of ϵ and the difference tends to zero as ϵ increases. Thus we choose a small range $\Delta\epsilon$ applicable to most systems with a lower bound 0.1 which is taken as the threshold value for constructing the RN for $M = 3$. Similar results are obtained for other M values as well.

The scheme has been effectively applied to compare network measures from different chaotic attractors [18], to study the influence of noise on the structure of chaotic attractors [19] and to propose a new heterogeneity index [20] for complex networks which, in turn, provides a unique measure for each chaotic attractor through RN. We stick to the same criteria for the selection of ϵ in this work.

(b) Construction of the weighted recurrence network

The unweighted RN is constructed first using the selected value of ϵ. In order to convert it into a weighted RN, one has to assign weight factors to every link in the network. For weighted networks that model any real-world system or interaction, the weight factor will be specific to the network. For example, in a transportation network, it may depend on the distance between two nodes while for a communication network, the same may be characterized by the rate of information transfer through the link. Here, we introduce a general criterion for assigning the weight factors that can be adopted to any kind of network, but is especially useful for RNs.

Assume that the RN has N nodes and the $\imath$th node has a degree k_i. That is, it is connected to k_i other nodes in the network. The weight factor w_{ij} for the link between two nodes $\imath$ and $\jmath$ in the network is defined as:

$$w_{ij} = \frac{\sqrt{k_i k_j}}{k_{\max}}, \tag{2.1}$$

where $k_{\max}$ is the maximum degree in the network. Note that the maximum possible value of w_{ij} is normalized as 1 and occurs for a link between two nodes which are connected to $k_{\max}$ other nodes in the network. For a reference node i in general, it is connected to k_i other nodes with each link having a different weight factor. Note that the above method of assigning the weight factor is a non-subjective criterion (independent of the details of interaction the network represents) and hence can be adopted in several contexts. In particular, in the case of RNs, the connection between two nodes comes by way of the proximity of the corresponding points on the embedded attractor. Hence, nodes corresponding to more densely populated regions of the attractor will be connected with more weight and vice versa. From the point of view of a network, the higher the number of connections for a node $\imath$, the shorter the path becomes when connected through the node. For example, if a nearly isolated (with $k_i \sim 1$) node is connected to a hub, that connection carries a high weight factor (due to the large degree of the hub) and provides an easy path between the node and any other arbitrary node in the network.

The sum of the weight factors associated with a node as determined by its connections defines the *strength* of the node, s [21,22]. For example, for the node $\imath$, we have:

$$s_i = \sum_{j=1}^{k_i} w_{ij}. \tag{2.2}$$

If all the nodes have equal number of connections $\langle k \rangle$, the strength of each node is the same and the network can be considered as a homogeneous weighted network. As the strengths among the nodes become more diverse, the network becomes more heterogeneous. The average strength associated with the whole network is defined as the weighted link density:

$$\rho_w = \frac{\sum_{i,j} w_{ij}}{N(N-1)}. \tag{2.3}$$

3. Measures from weighted recurrence network

After constructing the WRN, we are now in a position to analyse various time-series data using the characteristic measures of WRN. Here, we generalize three important network measures, the degree distribution, the CC and the CPL for the WRN.

In this paper, we use time series from two standard chaotic systems and two representative noise data as examples to illustrate the potential of WRN and its utility in the analysis of time-series data. The chaotic time series are from the standard Lorenz attractor (parameters $\sigma = 10$, $\rho = 8/3$ and $r = 28$) and the Rössler attractor (parameters $a = 0.2$, $b = 0.2$ and $c = 7.8$). In both cases, the time series is generated using a time step $\Delta t = 0.05$ after discarding the first 20 000 points as transients.

Two different types of noise data are also used for the analysis. One is the white noise and the other one is a candidate from the family of coloured noise. The coloured noises are correlated random processes which are considered to be important in the analysis of chaotic data since they share many characteristic properties common to chaotic data [23]. They are a class of random processes with power spectra scaling as $1/f^{\alpha}$, with α ranging, typically, from 1 to 2. Here, we choose $1/f$ noise as a representative coloured noise which are ubiquitous in the real world and are popularly called Brownian noise, since they behave identical to the Brownian motion. To compute the characteristic measures, 10 different time series are used (by changing initial conditions for chaos and with different simulations for noise) and the average is taken.

(a) Normalized strength distribution

For any unweighted complex network, the degree distribution, denoted by $P(k)$, is a probability distribution representing how many nodes have a given degree k. For random graphs (RG), the degree distribution is Poissonian whereas for scale-free (SF) networks, it obeys a power law [24]. For the RN from chaotic time series, the degree distribution is characteristic to the structure of the attractor [18]. To generalize the degree distribution for the WRN, we first note that the characteristic property of a node that describes its connectivity in the network is not its degree, but its strength s as defined in equation (2.2). In other words, the degree distribution has to be replaced by the *strength distribution* [25] of the weighted network which represents the probability $P(s)$ of nodes having a given strength s in a network of N nodes. Even though s varies discretely, it is not an integer like the degree k.

Instead of using the strength distribution directly, we use a *normalized strength distribution* that reveals the utility of WRN. Since s varies discretely, we can write

$$\sum_s P(s)\Delta s = 1. \tag{3.1}$$

We now find the number of nodes $n(s) \equiv NP(s)$ (rather than the probability of nodes) corresponding to the interval Δs. Then the above equation can be re-written as

$$\frac{1}{N}\sum_s n(s)\Delta s = 1. \tag{3.2}$$

We now consider the normalized strength $s_n = s/N$, which varies in the unit interval $[0, 1]$, and compute the variation of $n(s_n)$ with s_n for WRN from chaotic and random time series. Here, $n(s_n)$ is the number of nodes having strength around the normalized value s_n.

First, we show that the distribution is qualitatively different for chaotic and random data. In figure 1*a*,*c*, we show the normalized strength distribution of the WRN from the Lorenz attractor (solid circles) time series and the same for white noise (solid triangles). It is clear that the two distributions are qualitatively different. While the first one appears to be a power law, the second one decreases exponentially. We find that in the case of the Lorenz system, the variation can be represented using the following functional fit:

$$n(s_n) \propto s_n^{-\gamma}\, \mathrm{e}^{-s_n/c}, \tag{3.3}$$

with the parameters γ and c depending on the particular system. On the other hand, the distribution for the WRN from white noise is purely exponential as:

$$n(s_n) \propto \mathrm{e}^{-s_n/c}. \tag{3.4}$$

To show the variation more clearly, the same distributions are plotted in log scale in figure 1*b*,*d* with the functional fit as solid line in both cases. The crucial parameter here is the power law index γ indicating a SF character for the distribution before exponential cut-off. The average value of γ from 20 different simulations is found and turns out to be 0.33 ± 0.04 for WRN from the Lorenz attractor. We have also checked the distribution for a few other standard chaotic attractors, such as, the Rössler attractor, Henon attractor and Duffing attractor, and have found similar behaviour in all cases.

We now show that this distribution is a characteristic property of every chaotic attractor and is independent of changes in algorithmic parameters, such as, embedding dimension M and the number of nodes in the network N. This is illustrated using the Lorenz attractor in figure 2*a*,*b* for two N values with fixed M and vice versa. The result implies that the power law index γ is a characteristic index for a chaotic attractor. The computation is repeated for white noise by changing N and M. The results are also shown in figure 2*c*,*d*. Apart from the absence of power-law scaling, the distribution is found to shift (figure 2*d*) as M changes. This is because, unlike for the chaotic attractor, noise tends to fill the available state space which, in turn, changes the constructed network and the corresponding measures as M changes.

To understand the emergence of a power law in the strength distribution for chaotic systems, we look at the construction of the WRN more closely and find that the SF character follows from the method of defining the strength of a node by summation over the weight factors of its links. For the unweighted RN, the degree k_i of the $\imath$th node represents the local probability density around the corresponding point on the attractor. So the degree distribution $P(k)$ represents, as a discrete distribution, how many nodes (local regions over the attractor) have the same degree (probability density). This will be characteristic of the structure of the chaotic attractor and does not change with either N or M, as we have shown [18].

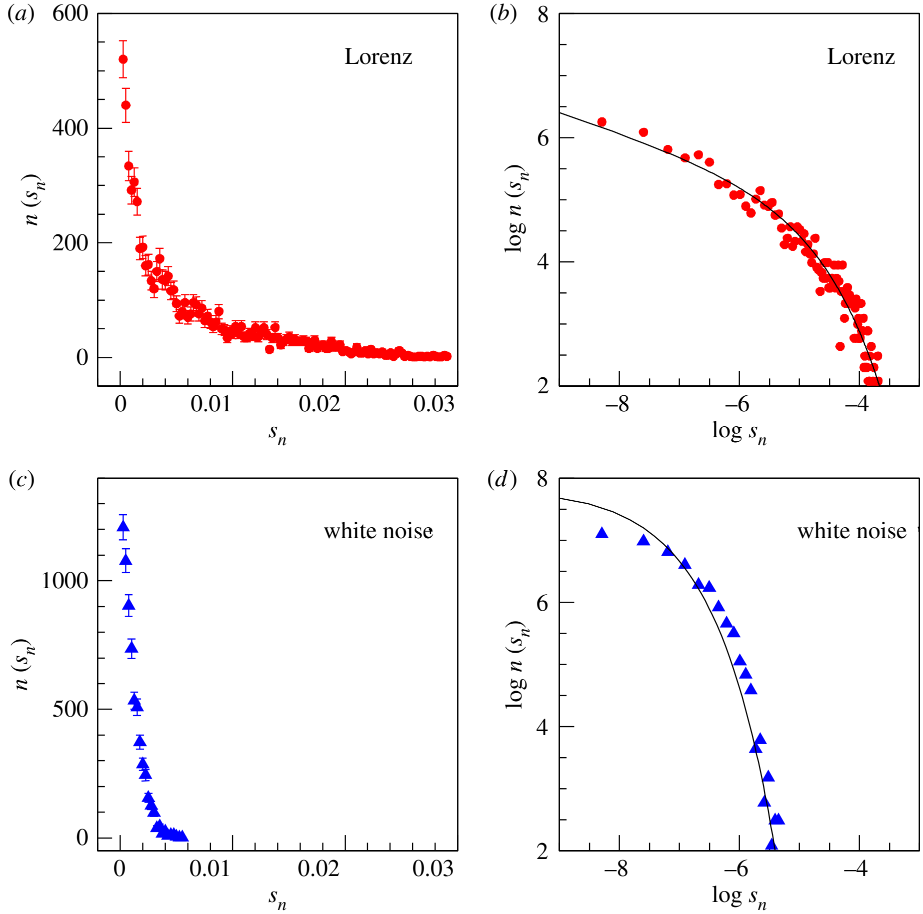

Figure 1. Normalized strength distribution of the weighted RN constructed from the Lorenz attractor time series (red solid circles) and white noise (blue solid triangles) are shown in (*a*,*c*) with errorbar resulting from counting statistics. (For a network with N nodes, the number n of nodes with a given degree k has a standard error of $\sqrt{n}(k)$. For $n(k) \longrightarrow 0$, its value is normalized to 1, the minimum count.) The same distributions drawn in the log scale without errorbar are shown in (*b*,*d*). The solid line in (*b*,*d*) represents the fitting function for the corresponding distributions (see text). The embedding dimension used for constructing the networks is $M = 3$ and the number of nodes in the network, $N = 5000$. (Online version in colour.)

On the other hand, for the WRN with weight given by equation (2.1), each node ι has a strength s_i, which is the sum of the normalized weight factors of its connections. In other words, it represents a blunt measure as it takes into consideration a node's total involvement in the network and not the number of other nodes connected to it [22]. For example, two nodes having totally different number of connections (that is, degree) can have the same node strength as the weights of edges of each connection are different. Thus, unlike the degree distribution, the distribution $P(s)$ cannot be directly correlated to the probability density variations over the attractor. However, we find that this distribution, in general, has a common form independent of the structure of the attractor.

It should be noted that, by construction, the WRN and the associated measures are related to the structure of the embedded attractor and not directly to its dynamics. In order to support this, we have checked the distributions derived from two different types of attractors. The first is a strange non-chaotic attractor (SNA) [26,27] generated from a quasi-periodically forced pendulum and the second one is periodic attractor. In the first case, the strength distribution was similar to that from a chaotic attractor with power-law scaling and exponential cut off, while in the second

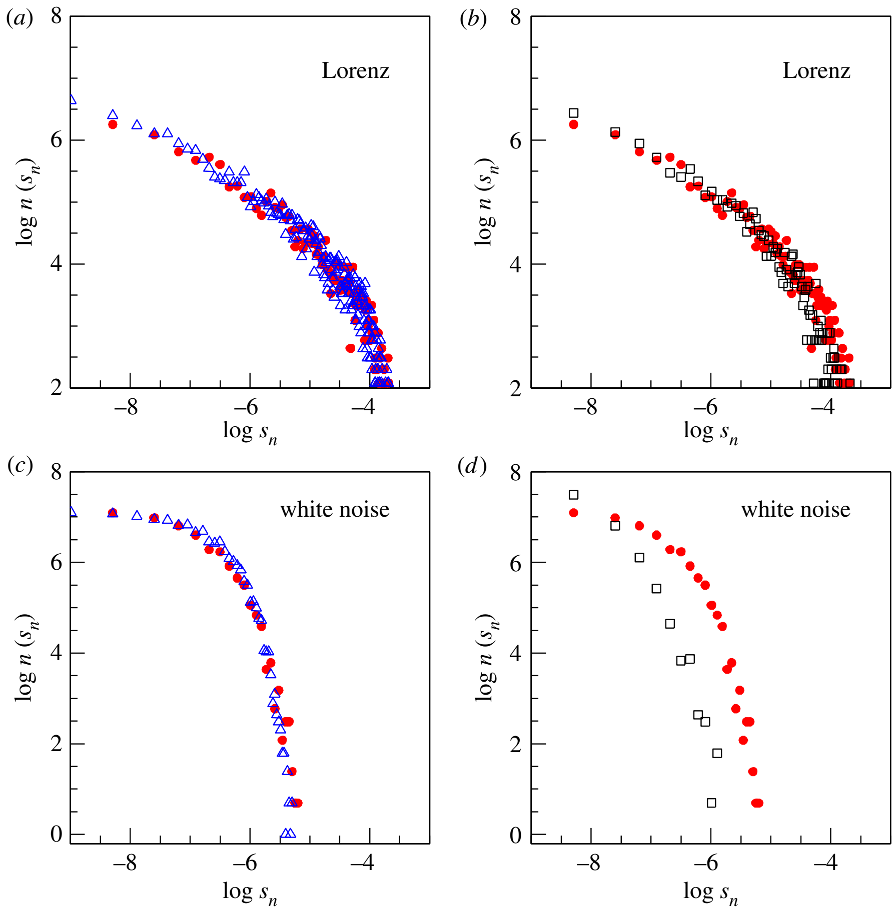

Figure 2. (*a,b*) The normalized strength distributions of WRN constructed from the Lorenz attractor (without errorbar) for two different values of *N* with fixed *M* (*a*) and vice versa (*b*). It is clear that the distribution remains unchanged with respect to changes in *M* and *N*. (*a,c*) $M = 3$ with $N = 5000$ (red solid circles) and $N = 10\,000$ (blue open triangles) while (*b,d*) $N = 5000$ with $M = 3$ (red solid circles) and $M = 4$ (black open squares). (*c,d*) The same results for WRN constructed from white noise. Note the shift in the distribution as M changes (*d*). (Online version in colour.)

case, the distribution was qualitatively different with strength distributed over a small range of s_n and the number of nodes having strength within this range approximately remaining constant. The range shrinks as N increases. From the above results, it follows that the power law scaling is a consequence of the fractal structure of the attractor and has nothing to do with whether the system is chaotic or not. The connection between the SF character of the distribution and the fractal structure is a matter that requires a more detailed investigation.

(b) Weighted clustering coefficient

We now compute another primary network measure, the global CC [28] for WRN. To compute the CC for the unweighted complex network, we first consider the local CC of a node $\imath$ as

$$C_i = \frac{2f_i}{k_i(k_i - 1)}, \tag{3.5}$$

where k_i is the degree of the node and f_i are the number of basic non-trivial motifs (triangles) attached to the node [28]. The value of C_i measures how many of the nodes connected to the node $\imath$ are also mutually interconnected and its value is normalized in the range from 0 to 1. By averaging C_i for all the nodes over the entire network, we get the global CC of the network.

To generalize this, the weighted local CC of a node $\imath$, denoted by C_i^W, is first determined following Onnela *et al.* [29]. For this, we replace the number of triangles in equation (3.5) with the sum of triangle intensities as

$$C_i^W = \frac{2}{k_i(k_i-1)} \sum_{j,l} (w_{ij}^n w_{jl}^n w_{li}^n)^{1/3}, \tag{3.6}$$

where the weight factors of the links are scaled by the largest weight factor in the network:

$$w_{ij}^n = \frac{w_{ij}}{\max(w_{ij})}. \tag{3.7}$$

This definition also fulfills the requirement that $C_i^W \to C_i$ as the weights become binary. The global weighted CC of the network is now given by

$$\mathrm{CC}^W = \frac{1}{N} \sum_{i=1}^{N} C_i^W. \tag{3.8}$$

We now compute the CC^W for WRN from several standard chaotic attractors and random data by changing N and M. The results are shown in figure 3*a*,*b* for the Lorenz and the Rössler attractors as well as the white and the $1/f$ noise. Figure 3*a* shows the results when fixing $M=3$ and changing N and figure 3*b* for fixed $N=5000$ and changing M. The important result here is that, just like the strength distribution, CC^W is also a characteristic measure for a given chaotic attractor independent of both N and M. On the other hand, CC^W for both white noise and $1/f$ noise show a decreasing trend as M increases, though tends to remain constant when M is fixed and N is increased. This is because, for the random data, the trajectory tends to fill the available dimension even as $M \to \infty$, unlike for the case of a chaotic attractor. Thus, its CC tends to decrease with M for a fixed N.

(c) Characteristic weight

In this section, we focus on another important measure of any complex network, namely, the CPL and try to generalize this measure for the WRN that we consider here. We first briefly review the basic ideas for the unweighted case. The CPL is a measure of the global connectivity of a network and is defined through the shortest path length l_{ij} between any pair of nodes $(\imath, \jmath)$ in the network. Here, l_{ij} represents the minimum number of nodes to be covered starting from a reference node $\imath$ to reach any other node $\jmath$ in the network. To calculate CPL, we first compute l_{ij} for all the nodes $\jmath$ for a given $\imath$ and the average is found. This is repeated by changing $\imath$ for all the nodes in the network and the global average is found:

$$\mathrm{CPL} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{N-1} \sum_{j(\neq i)=1}^{N-1} l_{ij} \right). \tag{3.9}$$

To generalize this for the weighted network, one should first note that the *shortest path* has to be replaced by *the path with the maximum weight*, since the connectivity increases with the weight factor for a link. For example, suppose we consider the shortest path from a reference node $\imath$ to any other node $\jmath$, where $\imath$ and $\jmath$ are not directly connected. There will be different paths to reach $\jmath$ from $\imath$. The effective weight Ω_{ij} for different *paths* will be different. Note that we use a different notation for the weight for a path which is, in general, a combination of several links. We have to choose the path with the maximum of Ω_{ij}. To calculate the effective weight for a path between nodes $\imath$ and $\jmath$, we use *the harmonic mean of the weight factors of all the links intermediate between $\imath$ and $\jmath$ for the path* [28]. This is repeated for all possible paths between $\imath$ and $\jmath$ and the highest effective weight is chosen as the characteristic weight (CW) between the nodes $\imath$ and $\jmath$ denoted as Ω_{ij}^f.

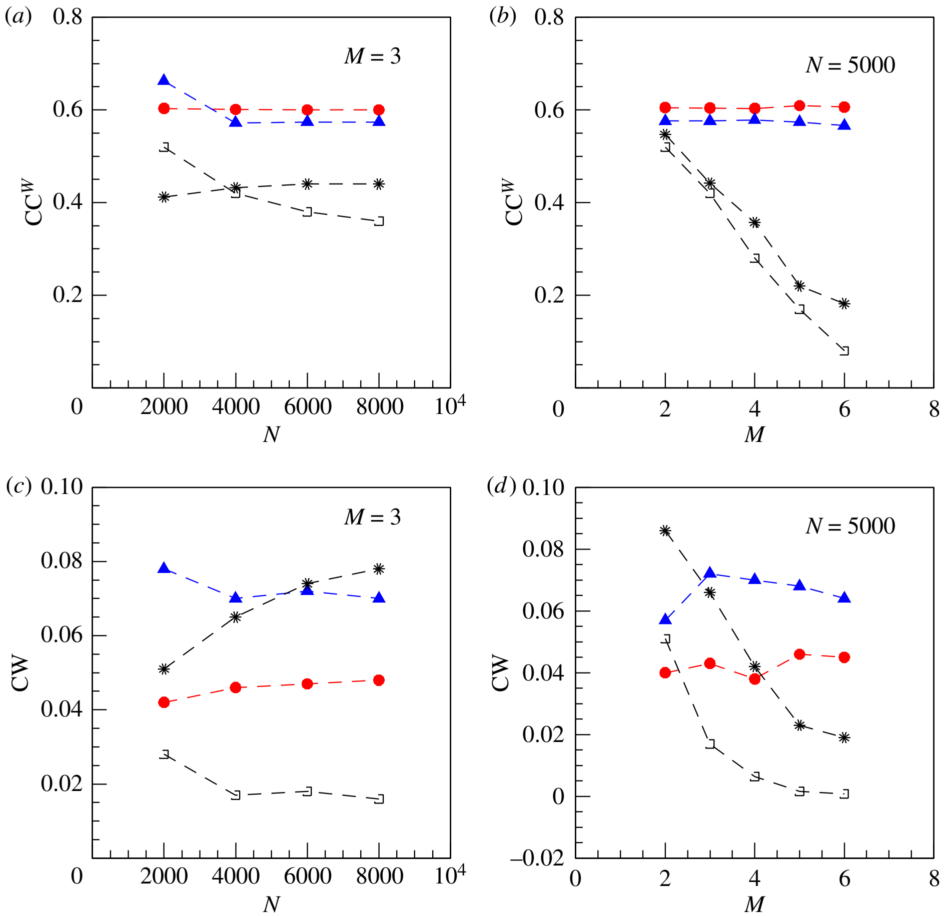

Figure 3. (*a*,*b*) The variation of the clustering coefficient (CC^W) of WRN, as a function of *N*, constructed from Lorenz attractor (red solid circles), Rössler attractor (blue solid triangles), white noise (asterisk) and $1/f$ noise (black open square). The value of *M* is fixed at 3. Panel (*b*) shows the same, but as a function of *M* with *N* fixed at 5000. (*c*,*d*) The results for the same systems (symbols same), but the variation of CW as a function of N (*a*,*c*) with M fixed and vice versa (*b*,*d*). (Online version in colour.)

To illustrate the idea, consider a network of 5 nodes and assume that there are two possible paths between nodes 1 and 2, namely, $1 \to 3 \to 2$ and $1 \to 4 \to 5 \to 2$. The effective weight for the former is given by

$$\frac{1}{\Omega_{12}^{(1)}} = \frac{1}{w_{13}} + \frac{1}{w_{23}}. \tag{3.10}$$

Or,

$$\Omega_{12}^{(1)} = \frac{w_{13}w_{23}}{w_{13} + w_{23}}. \tag{3.11}$$

Similarly, for the latter, the effective weight is

$$\Omega_{12}^{(2)} = \frac{w_{14}w_{45}w_{52}}{w_{14}w_{45} + w_{45}w_{52} + w_{14}w_{52}}. \tag{3.12}$$

The largest effective weight Ω_{ij}^{f} is averaged by changing j for a given i and for the whole network by changing i from 1 to N. We call it the *CW* of the WRN denoted by CW:

$$\mathrm{CW} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{N-1}\sum_{j(\neq i)=1}^{N-1}\Omega_{ij}^{f}\right). \tag{3.13}$$

It gives a measure for the global connectivity of the weighted network analogous to the CPL. However, a crucial difference between CPL and CW should be noted. When the connections are binary, a higher value of CPL indicates that more nodes are to be visited, on the average, to reach a node j from node i. In other words, the global connectivity depends inversely on CPL, which is always greater than 1. On the other hand, the weight Ω_{ij} for a path is defined in such a way that a higher value of CW (which is always less than 1) indicates a better global connectivity for the whole network.

This measure is now computed for the WRN from various chaotic time series and noise. The results are shown in figure 3*c*,*d* for Lorenz, Rössler, white noise and $1/f$ noise. It is evident that the behaviour of CW is similar to that of CC^W. While it is a characteristic measure for a chaotic attractor, the same is not true for white noise and $1/f$ noise. For both, the global connectivity of the WRN decreases systematically as M is increased. This also implies that the WRN provides an effective tool to distinguish a chaotic time series from white noise and $1/f$ coloured noise.

4. Analysis of variable star light curves

In this section, we check how the proposed measures are relevant in a real-world scenario by analysing light curves from variable stars. We consider data from RR Lyrae stars from the Kepler Archive where period doubling and strange non-chaotic behaviour have been reported in previous analyses using conventional methods [30–32]. We use two sample light curves for analysis from this group of stars, namely, KIC 4484128 and KIC 7505345. Continuous data segments without gaps sampled at 0.0204 days involving 32 000 data points are used in both cases. More details regarding the data can be found elsewhere [32].

We divide each light curve into four equal segments of 8000 data points each and the analysis is done separately on each segment. We compute the quantifiers—normalized strength distribution, CC^W and CW—using each segment of the light curve. We find that the distributions of all the segments of the light curves behave in an identical manner and the values of CC^W and CW converge to constant values as a function of M, within the numerical error. In figure 4, we show the typical normalized node strength distribution for the two stars. The presence of a power law with exponential cut off at tail is evident in both, which is in agreement with the strange non-chaotic behaviour obtained in earlier analysis.

In figure 5, we show a combined plot of CW and CC^W of all the data, both synthetic and real world, analysed in this manuscript. Note that the values of both CW and CC^W of the two noise data are clearly separated from the rest of the data showing either chaotic or strange non-chaotic behaviour. Also, the values for the two variable star data once again indicate their fractal nature. The SNA shown in the figure was generated from:

$$\frac{d^2x}{dt^2} + a\frac{dx}{dt} + b\sin x = d + c(\cos\omega t + \cos\Omega t), \tag{4.1}$$

with $a = 3.0$, $b = 1.0$, $c = 1.1$, $d = 1.33$, $\omega = (3 - \sqrt{5})/4$ and $\Omega = (1 + \sqrt{5})/4$.

5. Discussion and conclusion

Nonlinear time series analysis makes use of many quantifiers based on chaos theory to understand the nature of the temporal behaviour by reconstructing the dynamics of the underlying system from a scalar time series. Over the last one decade, the complex network approach to nonlinear time series analysis has gained importance due to several advantages over the conventional approach. The method envisages the use of complex network measures in the analysis of time-series data by first converting it into a complex network using a suitable scheme. A popular scheme to convert time series to networks is by using the property of recurrence of dynamical systems. In all such analyses done so far, the resulting RNs have been un-weighted and un-directed. Here, we propose a method to construct WRN from a time series and discuss the

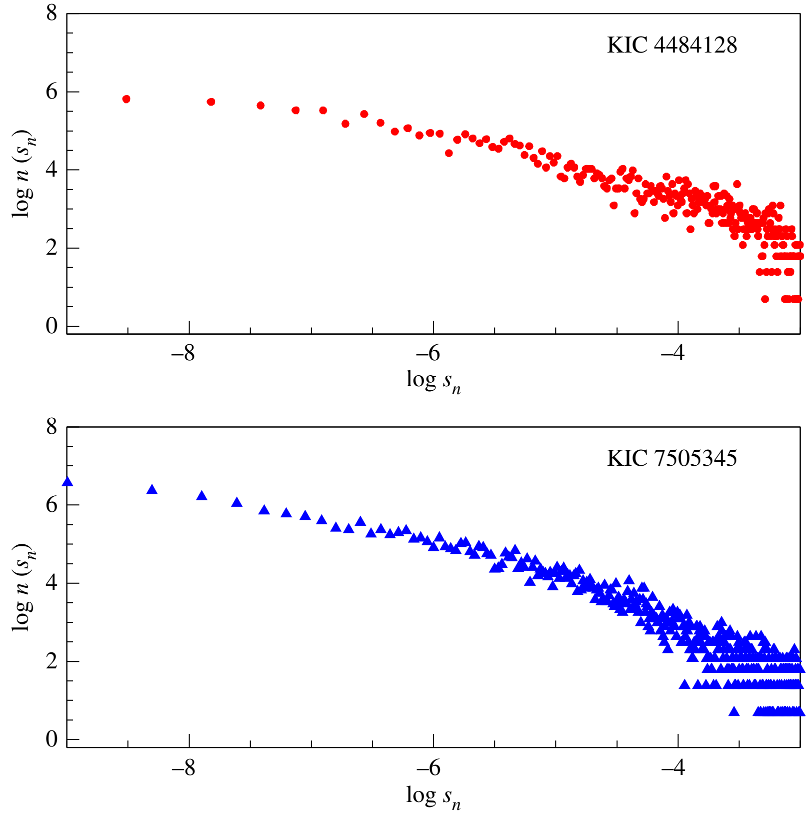

Figure 4. Node strength distributions of the weighted RN constructed from the light curves of two variable stars. (Online version in colour.)

relevance of associated measures and their practical applications in the analysis of both synthetic and real-world data.

We show that the WRN and the associated measures have a lot of potential in the nonlinear analysis of time-series data. The novel aspect in this work is the method of assigning weights to the links which provides useful computational tools for the analysis of time-series data. However, we do not claim that the method proposed here is the only one to construct WRN from a time series. As already mentioned above, this is just one method of assigning weight factor to a link whereby the nodes corresponding to higher probability density regions on the attractor are connected with higher weight. Another possible method of assigning weights is based on the distance function, as is done in some practical cases, such as traffic or communication networks. Different method of assigning weights will be applicable for different practical problems and the way the network is constructed to analyse a particular problem. The utility of the present scheme is that it is suited to extract some useful information regarding the structure of chaotic attractors, such as, the common power law in the node strength distribution.

Using the proposed scheme, we generalize three prominent network measures so as to make them suitable for WRN and apply them to analyse time series from standard chaotic attractors, white noise and $1/f$ coloured noise. For example, the degree distribution is generalized to the node strength distribution, the CC and the CPL are converted to their weighted counterparts. It is specifically shown that the node strength distributions of WRN from chaotic attractors follow a power law with exponential cut-off while those from white and $1/f$ noise are purely exponential. As the percentage of noise in the data increases, the power law part in the distribution vanishes systematically and the distribution tends to an exponential with $\gamma \to 0$. Moreover, the

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 475: 20180256

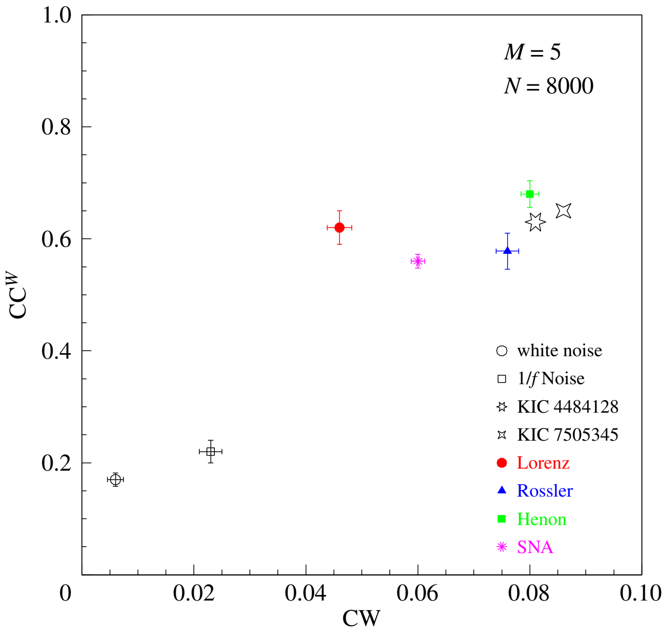

Figure 5. A combined plot of two WRN measures to distinguish chaos and strange non-chaos from white noise and $1/f$ noise. For chaotic systems, the average value from 10 different time series with changing the initial conditions is plotted. For noise, the average value from 10 different simulations is taken. The values for 3 standard chaotic attractors and 2 types of noise are shown as indicated in the figure. SNA is a strange non-chaotic attractor from a quasi-periodically forced pendulum (see text). To get a comparison with the real-world data, the values for two variable stars are also shown. The parameters used are $M = 5$ and $N = 8000$ in all cases. (Online version in colour.)

weighted clustering coefficient (CC^W) and the CW of WRN from a chaotic attractor are invariants independent of all parameters involved in the construction of the network other than recurrence threshold which has to be chosen properly. On the other hand, both these measures systematically get reduced with embedding dimension for white and $1/f$ noise, which implies that the local clustering and the global connectivity keeps on reducing with M for noise. We foresee how the measures proposed here therefore can become important and powerful tools for the analysis of real world data. To show this explicitly, we analyse sample light curves from variable stars which confirm their fractal nature as suspected from the previous analysis.

An important aspect of the whole analysis is that it is possible to identify the WRNs from different chaotic attractors as a single class with qualitatively similar strength distribution having power-law scaling and exponential tail. Such a classification was not possible so far with the existing methods involving undirected RNs. Thus, WRNs can be added as a separate class to the general set of complex networks. The power law index γ is a manifestation of the fractal structure of the reconstructed attractor and can serve as a single unique index characterizing its geometric complexity.

Finally, the method presented here has the advantage that accurate results can be obtained with much lower size of the data since the measures are defined not based on the size of the network (number of nodes) but on the connections between the nodes which will be much higher by construction. Also, the proposed method of assigning weights is general and can be adopted to generate weighted networks in other contexts like airport data, internet etc. with similar characterizations. We hope the methods and measures presented here will open up a new window in the pursuit of complexity in real-world systems, such as, climate and earth sciences [33,34] where weighted networks play an important role, apart from quantifying structural complexity of attractors through network-based approach.

Ethics. The work did not involve any active collection of human data, but only computer simulations using synthetic data.

Data accessibility. The work does not involve any experimental data. The synthetic data used are available in our nonlinear dynamics web page mentioned above. The variable star data are publicly available in the Kepler Data Archive as mentioned in the text.

Computations. All the computations have been done by Fortran codes developed by the authors. The codes and the data used are publicly available at our nonlinear dynamics web page: https://www.sites.google.com/site/kphk11/home.

Author contribution. The initial idea for the paper started from the discussions of R.J. and K.P.H. The codes for computations of the weighted network measures were developed in association with R.M. The interpretation of the results and the expert guidance at various stages are from G.A. and the manuscript was finalized after a series of discussions with G.A. All the authors gave the final approval of the manuscript.

Competing interests. The authors have no competing interests.

Funding. The work is not supported by any funding agencies.

Acknowledgements. We thank one of the anonymous referees for several suggestions to improve our manuscript. R.J and K.P.H. acknowledge the computing facilities in IUCAA, Pune.

References

1. Albert R, Barabasi AL. 2002 Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97. (doi:10.1103/RevModPhys.74.47)
2. Zhang J, Small M. 2006 Complex network from pseudoperiodic time series: topology versus dynamics. *Phys. Rev. Lett.* **96**, 238701. (doi:10.1103/PhysRevLett.96.238701)
3. Xu X, Zhang J, Small M. 2008 Superfamily phenomena and motifs of networks induced from time series. *Proc. Natl Acad. Sci. USA* **105**, 19601. (doi:10.1073/pnas.0806082105)
4. Marwan N, Donges JF, Zou Y, Donner RV, Kurths J. 2009 Complex network approach for recurrence analysis of time series. *Phys. Lett. A* **373**, 4246–4254. (doi:10.1016/j.physleta.2009.09.042)
5. Eckmann JP, Kamphorst SO, Ruelle D. 1987 Recurrence plots of dynamical systems. *Europhys. Lett.* **4**, 973–977. (doi:10.1209/0295-5075/4/9/004)
6. Donner RV, Zou Y, Donges JF, Marwan N, Kurths J. 2010 Recurrence networks—a novel paradigm for nonlinear time series analysis. *New J. Phys.* **12**, 033025. (doi:10.1088/1367-2630/12/3/033025)
7. Grassberger P, Procaccia I. 1983 Measuring the strangeness of strange attractors. *Physica D* **9**, 189–208. (doi:10.1016/0167-2789(83)90298-1)
8. Donner RV, Small M, Donges JF, Marwan N, Zou Y, Xiang R, Kurths J. 2011 Recurrence-based time series analysis by means of complex network methods. *Int. J. Bif. Chaos* **21**, 1019–1046. (doi:10.1142/S0218127411029021)
9. Donner RV, Heitzig J, Donges JF, Zou Y, Marwan N, Kurths J. 2011 The geometry of chaotic dynamics—a complex network perspective. *Eur. Phys. J. B* **84**, 653–672. (doi:10.1140/epjb/e2011-10899-1)
10. Gao Z, Yang Y-X, Cai Q, Zhang S-S, Jin N-D. 2016 Multivariate weighted recurrence network inference for uncovering oil-water transitional flow behavior in a vertical pipe. *Chaos* **26**, 063117. (doi:10.1063/1.4954271)
11. Eroglu D, Thomas K, Perin DM, Marwan N, Rodrigues FA, Costa LF, Sebek M, Kiss IZ, Kurths J. 2014 Entropy of weighted recurrence plots. *Phys. Rev. E* **90**, 042919. (doi:10.1103/PhysRevE.90.042919)
12. Sun X, Small M, Zhao Y, Xue X. 2014 Characterizing system dynamics with a weighted and directed network constructed from time series data. *Chaos* **24**, 024402. (doi:10.1063/1.4868261)
13. Gao ZK, Li S, Dang WD, Yang YX, Do Y, Grebogi C. 2017 Wavelet multiresolution complex network for analyzing multivariate nonlinear time series. *Int. J. Bif. Chaos* **27**, 1750123. (doi:10.1142/S0218127417501231)
14. Gao ZK, Dang WD, Mu C, Yang Y, Li S, Grebogi C. 2018 A novel multiplex network-based sensor information fusion model and its application to industrial multiphase flow system. *IEEE Trans. Ind. Informatics* **14**, 3982–3988. (doi:10.1109/TII.2017.2785384)
15. Gao ZK, Liu Y, Yang YX, Cai Q, Dang WD, Du XL, Jia HX. 2018 Multivariate weighted recurrence network analysis of EEG signals from ERP-based smart home system. *Chaos* **28**, 085713. (doi:10.1063/1.5018824)

16. Gao ZK, Small M, Kurths J. 2016 Complex network analysis of time series. *Europhys. Lett.* **116**, 50001. (doi:10.1209/0295-5075/116/50001)
17. Barrat A, Barthelemy M, Pastor-Satoras R, Vespignani A. 2004 The architecture of complex weighted networks. *Proc. Natl Acad. Sci. USA* **101**, 3747–3752. (doi:10.1073/pnas.0400087101)
18. Jacob R, Harikrishnan KP, Misra R, Ambika G. 2016 Uniform framework for the recurrence-network analysis of chaotic time series. *Phys. Rev. E* **93**, 012202. (doi:10.1103/PhysRevE.93.012202)
19. Jacob R, Harikrishnan KP, Misra R, Ambika G. 2016 Characterization of chaotic attractors under noise: a recurrence network perspective. *Commun. Nonlinear Sci. Num. Simul.* **41**, 32–47. (doi:10.1016/j.cnsns.2016.04.028)
20. Jacob R, Harikrishnan KP, Misra R, Ambika G. 2017 Measure for degree heterogeneity in complex networks and its application to recurrence network analysis. *R. Soc. Open Sci.* **4**, 160757. (doi:10.1098/rsos.160757)
21. Newman MEJ. 2004 Analysis of weighted networks. *Phys. Rev. E* **70**, 056131. (doi:10.1103/PhysRevE.70.056131)
22. Opsahl T, Agneessens F, Skvoretz J. 2010 Node centrality in weighted networks: generalizing degree and shortest paths. *Soc. Netw.* **32**, 245–251. (doi:10.1016/j.socnet.2010.03.006)
23. Osborne AR, Provenzale A. 1989 Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* **35**, 357–381. (doi:10.1016/0167-2789(89)90075-4)
24. Newman MEJ. 2003 The structure and function of complex networks. *SIAM Rev.* **45**, 167–256. (doi:10.1137/S003614450342480)
25. Noh JD, Rieger H. 2002 Stability of shortest paths in complex networks with random edge weights. *Phys. Rev. E* **66**, 066127. (doi:10.1103/PhysRevE.66.066127)
26. Grebogi C, Ott E, Pelikan S, Yorke JA. 1984 Strange attractors that are not chaotic. *Physica D* **13**, 261–268. (doi:10.1016/0167-2789(84)90282-3)
27. Kapitaniak T. 1991 On strange nonchaotic attractors and their dimensions. *Chaos Solitons Fractals* **1**, 67–77. (doi:10.1016/0960-0779(91)90056-F)
28. Opsahl T, Panzaraza T. 2009 Clustering in weighted networks. *Soc. Netw.* **31**, 155–163. (doi:10.1016/j.socnet.2009.02.002)
29. Onnela JP, Saramaki J, Kertesz J, Kaski K. 2005 Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E* **71**, 065103. (doi:10.1103/PhysRevE.71.065103)
30. Lindner JF, Kohar V, Kia B, Hippke M, Learned JG, Ditto WL. 2015 Strange nonchaotic stars. *Phys. Rev. Lett.* **114**, 054101. (doi:10.1103/PhysRevLett.114.054101)
31. Moskalik P *et al.* 2015 Kepler photometry of RRc stars: peculiar double-mode pulsations and period doubling. *Mon. Not. R. Astr. Soc.* **447**, 2348–2366 (doi:10.1093/mnras/stu2561)
32. George SV, Ambika G, Misra R. 2017 Detecting dynamical states from noisy time series using bicoherence. *Nonlinear Dyn.* **89**, 465–479. (doi:10.1007/s11071-017-3465-6)
33. Boers N, Rheinwalt A, Bookhagen B, Barbosa HMJ, Marwan N, Marengo J, Kurths J. 2014 The South American rainfall dipole: a complex network analysis of extreme events. *Geophys. Res. Lett.* **41**, 7397–7405. (doi:10.1002/2014GL061829)
34. Donges JF, Donner RV, Trauth MH, Marwan N, Schellnhuber HJ, Kurths J. 2011 Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. *Proc. Natl Acad. Sci. USA* **108**, 20 422–20 427. (doi:10.1073/pnas.1117052108)