

RESEARCH LETTER

Open Access



Temporal dynamics of streamflow: application of complex networks

Xudong Han¹, Bellie Sivakumar^{1,2,3*}, Fitsum M. Woldmeskel¹ and Milena Guerra de Aguilar⁴

Abstract

This study employs the concepts of complex networks to study the temporal dynamics of streamflow, with emphasis on annual scale (i.e., year-to-year connections). The study proposes a new approach to construct the streamflow network at the annual scale. It uses the daily streamflow data to construct the annual streamflow network, instead of using the annual (mean or accumulated) streamflow data. With this approach, each year serves as a node in the network, with each node having a time series of daily streamflow values (not a single streamflow value). Streamflow data observed over a period of 151 years (October 1862–September 2013) from the Mississippi River basin at St. Louis, Missouri, USA are considered for implementation of the approach. The properties of the annual streamflow network are investigated using three complex network-based methods: degree centrality, clustering coefficient, and degree distribution. The sensitivity of the results to streamflow correlation threshold is also examined. The results suggest that (1) there are only a few very significant nodes (years) in the annual streamflow network (degree centrality method); (2) the annual streamflow network is not a classical random graph, but may be a small-world network or scale-free network (clustering coefficient method); and (3) the network exhibits a combination of exponential and power-law distribution (degree distribution method). Based on the identification of a significant stretch of period (around the 1950s–1990s) with very weak connections with the rest of the period studied, the results also suggest the influence of dam construction (and other anthropogenic factors) on the evolution of annual streamflow dynamics.

Keywords: Streamflow, Temporal dynamics, Complex networks, Degree centrality, Clustering coefficient, Degree distribution

Background

Identification of patterns in data (e.g., streamflow) serves as a fundamental approach towards modeling and prediction of the underlying systems. Numerous methods have been developed for identification of patterns in data (in space, time, and space–time) and possible connections between the components involved. Such methods can be categorized in different ways depending on their concepts and use of data, such as linear and nonlinear, deterministic and stochastic, parametric and non-parametric, supervised and unsupervised, and their combinations. The methods include those that are based on correlation, trend, spectrum, data distribution, data reconstruction,

dimension, scaling, regression, clustering, and classification, among others. They have been extensively applied to identify patterns in hydrologic data around the world; see, for example, Labat et al. (2011), Sivakumar and Singh (2012), Özger et al. (2013), Tongal and Berndtsson (2014), and Xu et al. (2015) for some recent studies, and Salas et al. (1995) and Sivakumar and Berndtsson (2010) for compilations.

A key aspect in the identification of patterns in data is the search for “connections.” In this context, the concepts of “complex networks” (e.g., Watts and Strogatz 1998; Barabási and Albert 1999; Girvan and Newman 2002; Estrada 2012) seem to provide new avenues—a network is a set of points called “nodes” connected by a set of connections called “links.” Applications of the concepts of complex networks in hydrology have been gaining momentum in the last few years. Thus far, they have included studies of river networks (Rinaldo et al. 2006;

*Correspondence: s.bellie@unsw.edu.au

¹ School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia

Full list of author information is available at the end of the article

Zaliapin et al. 2010; Czuba and Foufoula-Georgiou 2014, 2015; Rinaldo et al. 2014), rainfall monitoring networks (Malik et al. 2012; Boers et al. 2013; Scarsoglio et al. 2013; Sivakumar and Woldemeskel 2015; Jha et al. 2015; Jha and Sivakumar 2017; Naufan et al. 2017), and streamflow monitoring networks (Tang et al. 2010; Sivakumar and Woldemeskel 2014; Halverson and Fleming 2015; Braga et al. 2016; Serinaldi and Kilsby 2016; Fang et al. 2017). Such studies have employed different methods, including degree centrality, clustering coefficient, degree distribution, closeness centrality, shortest path length, and community structure. The outcomes of such applications are encouraging, as they have important implications for the development of hydrologic models, interpolation/extrapolation of hydrologic data, and classification of catchments. The ability of the concepts of complex networks to represent all types of connections also makes them a potential candidate to serve as a generic theory for hydrology (Sivakumar 2015).

Despite their encouraging outcomes, it is important to recognize that most of the above studies have addressed only the spatial connections in hydrologic networks. Since temporal dynamics are an integral part of hydrologic systems, especially from the perspective of time series analysis for modeling and prediction, studying the suitability of complex networks for temporal connections is crucial. To our knowledge, the only studies that have attempted this, in the context of streamflow analysis, are those conducted by Tang et al. (2010), Braga et al. (2016), and Serinaldi and Kilsby (2016). Tang et al. (2010) employed the visibility graph algorithm (Lacasa et al. 2008) to construct networks for daily streamflow series of three rivers: one in China (the Yangtze River) and two in the United States (the Umpqua River and the Ocmulgee River). They then used degree distribution and accumulative degree distribution to identify the type of such streamflow networks. Using daily streamflow data, Braga et al. (2016) employed the horizontal visibility graph (HVG) to construct streamflow networks from 141 gaging stations that cover 53 Brazilian rivers. They further characterized these 141 networks by examining their degree distributions and clustering coefficients. They reported that the river discharges in several stations had evolved to become more or less correlated over the years and attributed that behavior to changes in the climate system and other man-made phenomena. Serinaldi and Kilsby (2016) used the directed horizontal visibility graph (DHVG) to study the dynamics of daily streamflow fluctuations from 699 stations in the continental United States. They explored irreversibility by mapping the time series into ingoing, outgoing, and undirected graphs and comparing the corresponding degree distributions. They showed that the degree distributions do not decay

exponentially, but tend to follow a sub-exponential behavior. The outcomes of these studies have important implications for streamflow modeling, prediction, and catchment classification.

In the present study, we attempt to further advance the applications of the concepts of complex networks for temporal connections in streamflow. Our objective here is to study the year-to-year connections in streamflow, i.e., temporal dynamics at the annual scale. This is motivated by the need to study long-term water management and the influence of large-scale climate patterns as well as anthropogenic effects, including the role of climate change. However, taking advantage of the general availability of daily streamflow time series (for most locations around the world), this study adopts a new approach to construct the streamflow network at the annual scale. The study uses daily streamflow data and constructs the streamflow network corresponding to the annual scale, instead of using the annual (accumulated or average) streamflow and employing the visibility graph. In other words, in this study, each year is considered as a node, with each node consisting of a time series of (365 daily) streamflow values, rather than a single (annual) streamflow value. This approach is different from the one employed in Tang et al. (2010), Braga et al. (2016), and Serinaldi and Kilsby (2016), who considered each day as a node and the entire daily time series/year as a network. The properties of the annual streamflow network are then identified using different methods.

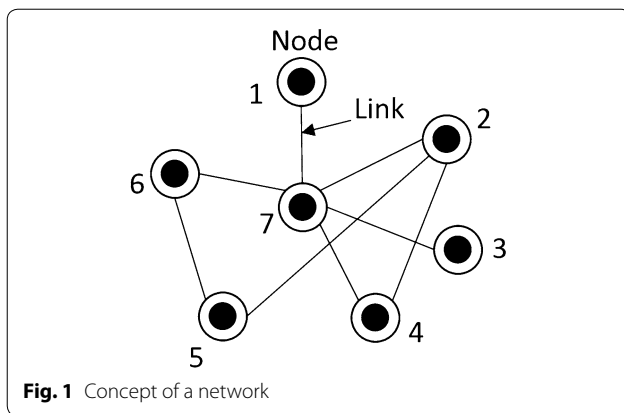
For implementation, streamflow data from the Mississippi River basin in the United States are studied. Specifically, daily streamflow data over a period of as many as 151 years (October 1862–September 2013) observed in the Mississippi River basin at St. Louis, Missouri are used. Considering each year as a node, three different methods are employed to investigate the connections in this annual streamflow network: degree centrality, clustering coefficient, and degree distribution. Different threshold values (i.e., correlations in streamflow between nodes) are also used to study the influence of threshold on the outcomes of degree centrality, clustering coefficient, and degree distribution methods.

The rest of this paper is organized as follows. First, the network construction and the three methods used in this study are described. Next, details of the study area and streamflow data are presented. Then, analysis and results are presented, followed by a discussion. Finally, some closing remarks are made.

Network methodology

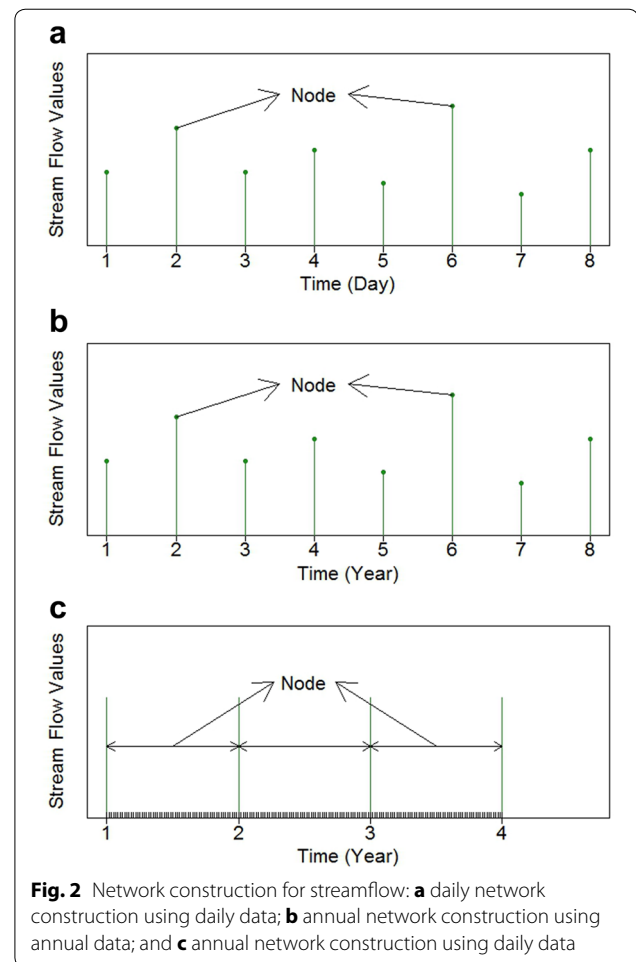
Network construction

A network (or a graph) is a set of points joined together by a set of lines, as shown in Fig. 1. The points are



referred to as *nodes* (or *vertices*) and the lines are referred to as *links* (or *edges*). Mathematically, a network can be represented as $G = \{P, E\}$, where P is a set of N nodes (P_1, P_2, \dots, P_N) and E is a set of n links. The network shown in Fig. 1 has $N = 7$ (nodes) and $n = 8$ (links), with $P = \{1, 2, 3, 4, 5, 6, 7\}$ and $E = \{\{1,7\}, \{2,4\}, \{2,5\}, \{2,7\}, \{3,7\}, \{4,7\}, \{5,6\}, \{6,7\}\}$. Figure 1, consisting of a set of identical type of nodes connected by identical type of links, is perhaps the simplest form of network. This kind of network, however, is rarely seen in nature, since natural (e.g., streamflow) networks are often far more complex. Indeed, there are many ways in which natural networks may be more complex. For instance, networks can (1) have different types of nodes and/or links; (2) contain nodes and links with a variety of properties associated with them (e.g., weights); (3) have links that can be directed; (4) contain multi-links, self-links, and hyperlinks; and (5) contain nodes of two distinct types, with links running only between unlike types (called *bipartite*). For further details, the interested reader is directed to Estrada (2012), among others.

In a network, the existence/non-existence of links is identified based on a measure that represents the strength of the link. The measure used to identify the link and its strength may be different, depending on the network under consideration and the problem of interest. For instance, in the analysis of spatial connections in a streamflow monitoring network (such as the one shown in Fig. 1), a common measure used is the spatial correlation between nodes, and node pairs that have spatial correlation values exceeding a certain threshold value (T) may be assigned links (e.g., Sivakumar and Woldemeskel 2014). However, in the analysis of temporal streamflow connections, the difference in streamflow values between nodes can be used as a measure, and node pairs that have differences below a certain threshold may be assigned links (e.g., Braga et al. 2016). With this basic network concept, construction of the streamflow network, in this



study, to represent the temporal dynamics at the annual scale is described next.

Let us assume that we have daily streamflow data observed over a period of N years at a gaging station. If the objective is to study the day-to-day connections in streamflow, then one can construct the network based on the daily streamflow values using, for example, the visibility graph method (e.g., Lacasa et al. 2008), considering each day as a node in itself, with each node having a single streamflow value (see Fig. 2a), as has been done by, for example, Tang et al. (2010), Braga et al. (2016), and Serinaldi and Kilsby (2016). However, if the objective is to identify the year-to-year connections in streamflow (or connections at any scale coarser than daily), then two different approaches may be adopted:

1. Compute certain statistic (e.g., mean, total) of streamflow for the annual scale, and then use the visibility graph method to construct the network based on such annual streamflow values. In this approach, each year is treated as a node (see Fig. 2b), and a

node has only one streamflow value, i.e., the annual streamflow value; and

2. Use the daily streamflow values to construct the streamflow network at the annual scale. In this approach, again each year is treated as a node, but then each node is made up of a time series of (365 or 366) daily streamflow values (see Fig. 2c).

The present study adopts the latter approach for network construction of streamflow at the annual scale, as it possesses the following advantages over the former: (1) it is simple, as it considers the daily data as they are and eliminates the need for visibility graph (or other methods) for network construction; (2) the construction takes into consideration the within-year streamflow variability to identify connections, rather than simply considering one annual value; and (3) the resulting network is similar to a network in space (i.e., each station as a node with a time series of streamflow and the connections between them as links), and therefore, the analysis becomes fairly straightforward and generic. For the purpose of convenience in the present analysis, each year is considered to contain only 365 days (i.e., February 29th in leap year is excluded). Therefore, the network construction adopted in this study for temporal dynamics is more similar to the construction adopted in Sivakumar and Woldemeskel (2014) and Halverson and Fleming (2015) for spatial dynamics than to the one adopted in Tang et al. (2010), Braga et al. (2016), and Serinaldi and Kilsby (2016) for temporal dynamics.

Network methods

There exist a variety of measures to study the properties of complex networks. These include centrality, clustering, adjacency, distance, community structure, bipartivity, subgraphs, and communicability, among others. Extensive details of these measures are available in Estrada (2012), among others. These measures identify/quantify different properties of networks. For some measures, there are also different definitions, submeasures, and the corresponding methods, as appropriate. In what follows, a brief description of degree centrality (centrality), clustering coefficient (clustering), and degree distribution (adjacency) is provided, as they are employed in this study to examine streamflow connections.

Degree centrality

Centrality is one of the most basic and intuitive measures of a network, as it identifies the significance of the nodes in the network. The concept of centrality goes back to the studies of Bavelas (1948) and Leavitt (1951) for communication networks. However, Jeong et al. (2001) and Newman (2001) were among the first to use the concept in the

context of complex networks. A number of centrality-based measures have been proposed in the network literature, such as degree centrality, centrality beyond nearest neighbors (e.g., Katz centrality, eigenvector centrality, subgraph centrality, PageRank centrality, and vibrational centrality), closeness centrality, betweenness centrality, and information centrality; see Estrada (2012) for details. Among these, the degree centrality has been one of the most widely used measures.

The idea behind the use of degree centrality as a network measure is that it identifies whether a given node, say i in a network, is more significant (or central or influential) than another node in the network. For instance, the node with the highest degree centrality value is considered as the most significant in the network, while the node with the lowest degree centrality value is considered as the least significant. The degree centrality of node i in a network of N nodes is defined as the number of first neighbors (or simply neighbors) of node i divided by the total number of possible neighbors ($N - 1$) in the network. The neighbors of node i are identified through finding the nodes that have links to node i according to an assumed threshold.

Let us consider a selected node i in a network of N nodes. So, the total number of possible direct neighbors for node i is $N - 1$, which means the total number of possible direct links for node i is $N - 1$. Let us assume that node i has only k neighbors (i.e., nodes), denoted as k_i , in the network according to an assumed threshold. This means that node i has k_i direct links (that connect it to k_i other nodes in the network). Therefore, the degree centrality of node i is given by the ratio of the number of direct links for node i (i.e., k_i) to the total number of all possible direct links for node i (i.e., $N - 1$). The procedure is repeated for each and every node of the network. An example of the calculation of the degree centrality is presented in Sivakumar and Woldemeskel (2014).

Clustering coefficient

One of the most basic properties of a network is its tendency to cluster. The concept of clustering has its origin in sociology, *under the name fraction of transitive triples* (Wasserman and Faust 1994). However, Watts and Strogatz (1998) were the first to use this concept in the context of complex networks. The tendency of a network to cluster is quantified by the clustering coefficient. There exist several definitions of clustering coefficient; see Watts and Strogatz (1998), Barrat and Weigt (2000), and Newman (2001) for details. However, the clustering coefficient method proposed by Watts and Strogatz (1998), which measures the local density, is widely used. A brief description of its calculation is presented here, as this method is used in the present study.

Let us consider first a selected node i in the network, having k_i links which connect it to k_i other nodes (i.e., neighbors) according to an assumed threshold, as mentioned earlier. If the neighbors of the original node i were part of a cluster, there would be $k_i(k_i - 1)/2$ links between them. Let us also assume that among the $k_i(k_i - 1)/2$ links, the number of ‘actual links’ that exist (according to the assumed threshold) is only E_i . With these, the clustering coefficient of node i is given by the ratio between the number E_i of links that actually exist between the k_i nodes and the total number of links $k_i(k_i - 1)/2$, i.e.,

$$C_i = \frac{2E_i}{k_i(k_i - 1)}. \tag{1}$$

The procedure is repeated for each and every node of the network. The average of the clustering coefficients of all the individual nodes is the clustering coefficient of the whole network C . An example of the clustering coefficient calculation can be found in Sivakumar and Woldemeskel (2014).

The clustering coefficient of the individual nodes and of the entire network can be used to obtain important information about the type of network, grouping (or classification) of nodes, and identification of the most significant nodes. For instance, a very high clustering coefficient (close to 1.0) indicates a *regular network*, since in a regular network, every node is connected to every other node in the same manner. A very low clustering coefficient (close to zero), with $C=p$ (where p is the probability of any two nodes in the network being connected), indicates a (classical) *random network*, since the connections between the nodes are purely random in nature. For a *small-world network* (e.g., Watts and Strogatz 1998), the clustering coefficient is generally smaller than that of the regular network but also considerably larger than that of a comparable random network (i.e., having the same number of nodes and links). A *scale-free network* (e.g., Barabási and Albert 1999) may also have such a clustering coefficient value. Therefore, it is often not easy to distinguish between small-world networks and scale-free networks based on the clustering coefficient alone (both small-world networks and scale-free networks essentially belong to the category of random networks, but their properties are different from that of classical random networks). However, other network-based measures, such as the shortest path length (e.g., Watts and Strogatz 1998) and the degree distribution (e.g., Barabási and Albert 1999), can provide reliable information to identify/distinguish between small-world networks and scale-free networks, or even some other type. It is relevant to note, at this point, that for a number of real-world networks studied in the literature, including hydrologic networks, the clustering coefficient is reported to be above 0.5 (e.g.,

Watts and Strogatz 1998; Jeong et al. 2000; Newman 2001; Newman et al. 2001; Tsonis and Roebber 2004; Suweis et al. 2011; Scarsoglio et al. 2013; Sivakumar and Woldemeskel 2014, 2015; Halverson and Fleming 2015), suggesting that such networks are not classical random networks, but may be small-world networks or scale-free networks or some other types.

Degree distribution

In a network, different nodes may have different number of links. The number of links (k) of a node is called *node degree*. The degree is an important characteristic of a node, as it allows one to derive many measurements for the network. The spread in the node degrees is characterized by a distribution function $p(k)$, which expresses the fraction of nodes in a network with degree k . This distribution is called *degree distribution* (e.g., Barabási and Albert 1999). The degree distribution is often a reliable indicator of the type of network.

In a random graph, since the links are placed randomly, the majority of nodes have approximately the same degree, and close to the average degree \bar{k} of the network. Therefore, the degree distribution of a completely random graph is a Poisson distribution with a peak at $p(\bar{k})$, and is given by

$$p(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!}. \tag{2}$$

Similarly, depending upon the properties of networks, the degree distribution can also be Gaussian, given by

$$p(k) = \frac{1}{\sqrt{2\pi\sigma_k}} e^{-\left(\frac{(k-\bar{k})^2}{2\sigma_k^2}\right)}, \tag{3}$$

exponential, given by

$$p(k) \sim e^{-k/\bar{k}}, \tag{4}$$

power-law or scale-free, given by

$$p(k) \sim k^{-\gamma}, \tag{5}$$

or other, or their combinations.

Among these distributions, the power-law or scale-free distribution (e.g., Barabási and Albert 1999) has attracted the most attention in the literature on complex networks, since such a distribution has been found in a number of natural and social networks (e.g., Barabási and Albert 1999; Kim et al. 2004; Keller 2005; Clauset et al. 2010). The fractal or scale-free nature of numerous natural systems, including hydrologic systems, and their ability to self-organize themselves, already well-documented in the literature (e.g. Mandelbrot 1983; Bak 1996;

Rodriguez-Iturbe and Rinaldo 1997; Peckham and Gupta 1999; Barnsley 2012), give both credence and motivation to further advance research on scale-free networks. While it is true that some scale-free networks display an exponential tail, the functional form of $p(k)$ still deviates significantly from the Poisson distribution expected for a random graph.

Study area and data

In the present study, streamflow data from the Mississippi River basin are considered to investigate the usefulness of complex networks for temporal streamflow dynamics. The Mississippi River originates at Lake Itasca in northern Minnesota in the United States and flows for about 3770 km (2342 mi) through the mid-continental United States, the Gulf of Mexico Coastal Plain, and its subtropical Louisiana Delta (Fig. 3). The entire river basin measures about 4.76 million km² (1.84 million mi²), of which about 3.22 million km² (1.24 million mi²) is in the continental United States; see Alexander et al. (2012) for further details.

In the Mississippi River basin, streamflow data are measured at thousands of locations. For the present study, daily streamflow data observed in a sub-basin station of the Mississippi River basin at St. Louis, Missouri (USGS station 07010000) are analyzed; see Fig. 3 for the location of St. Louis. The sub-basin is situated between 38°37'03" latitude and 90°10'47" longitude, on downstream side of west pier of Eads Bridge at St. Louis, 24.1 km downstream from the Missouri River, and at 289.6 km above the Ohio River. The drainage area of this sub-basin is 251,230 km² (97,000 mi²). The natural flow of stream in this sub-basin is affected by many reservoirs and navigation dams in the upper Mississippi River basin and by many reservoirs and diversion for irrigation in the Missouri River basin (e.g., Alexander et al. 2012).

For the present analysis, daily streamflow data observed over a period of 151 years (October 1862–September 2013) (i.e., “water year”) are considered. The data are obtained from the USGS National Water Information System website; see <http://nwis.waterdata.usgs.gov/nwis>. Figure 4 shows the variation of this daily streamflow

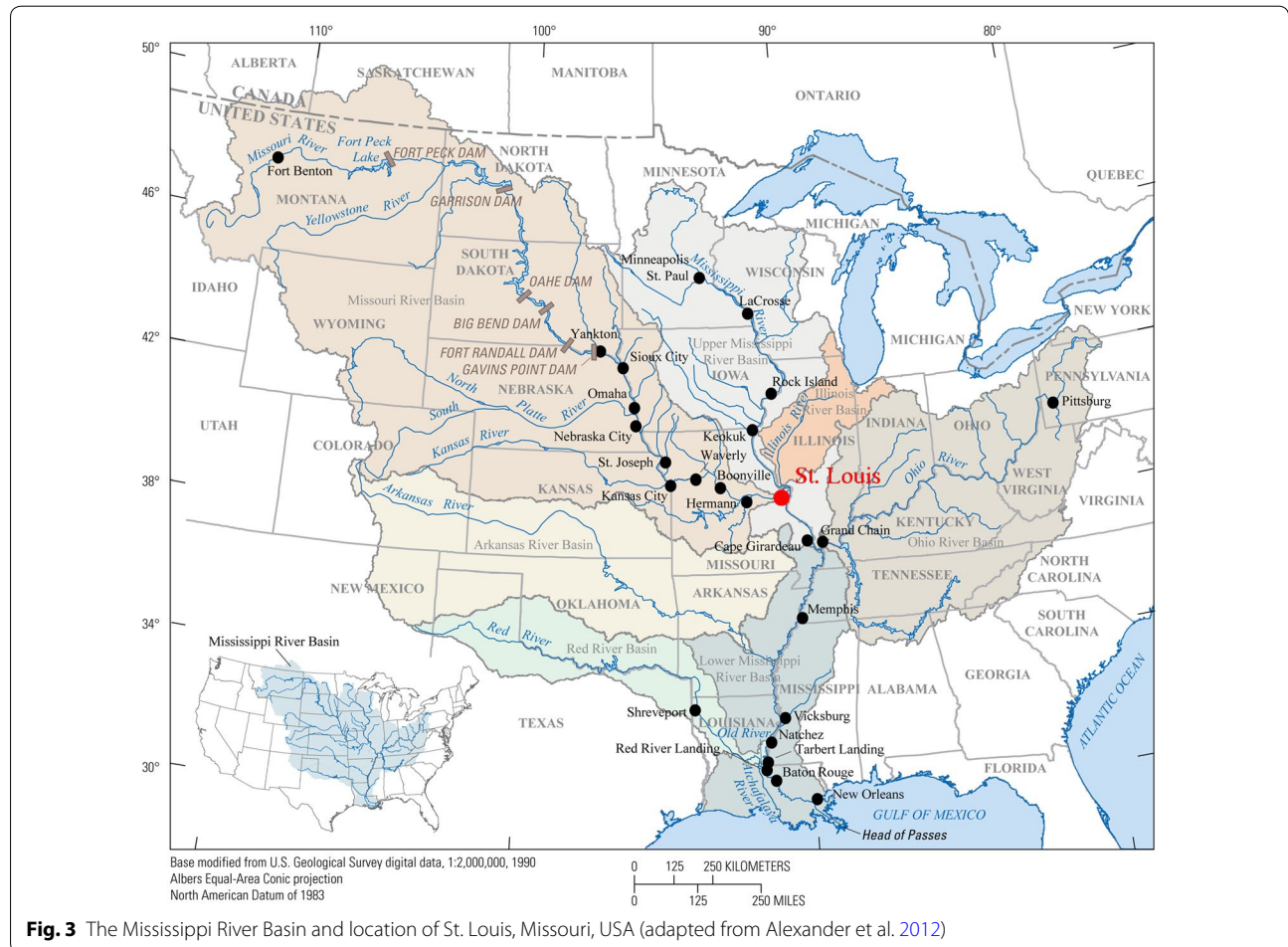


Fig. 3 The Mississippi River Basin and location of St. Louis, Missouri, USA (adapted from Alexander et al. 2012)

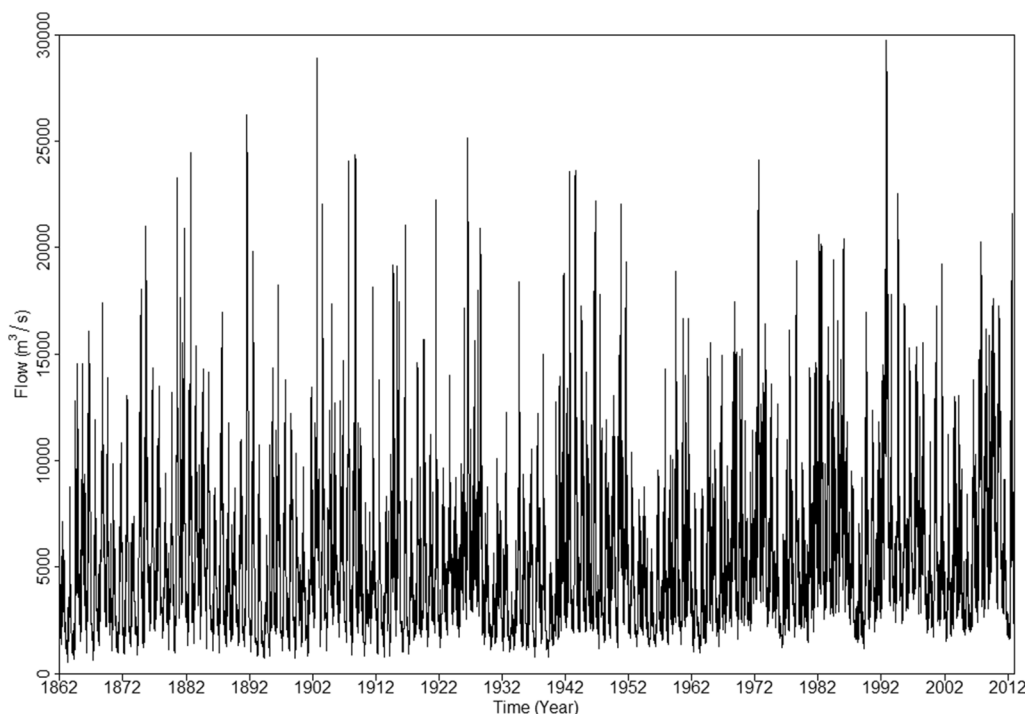


Fig. 4 Variation of daily streamflow time series from the Mississippi River basin at St. Louis, Missouri, USA

series. It is relevant to mention here that the temporal dynamics of streamflow (and other river-related processes) observed at the St. Louis station have been investigated by many studies in recent years. Among such studies, those that have employed nonlinear dynamic and chaos concepts for system identification, prediction, and catchment classification (e.g., Sivakumar and Jayawardena 2002; Sivakumar and Wallender 2005; Sivakumar et al. 2007) may be of particular interest in the context of complex networks, as there is potential to construct networks based on nonlinear data reconstruction (phase space reconstruction). This will be addressed in a future study.

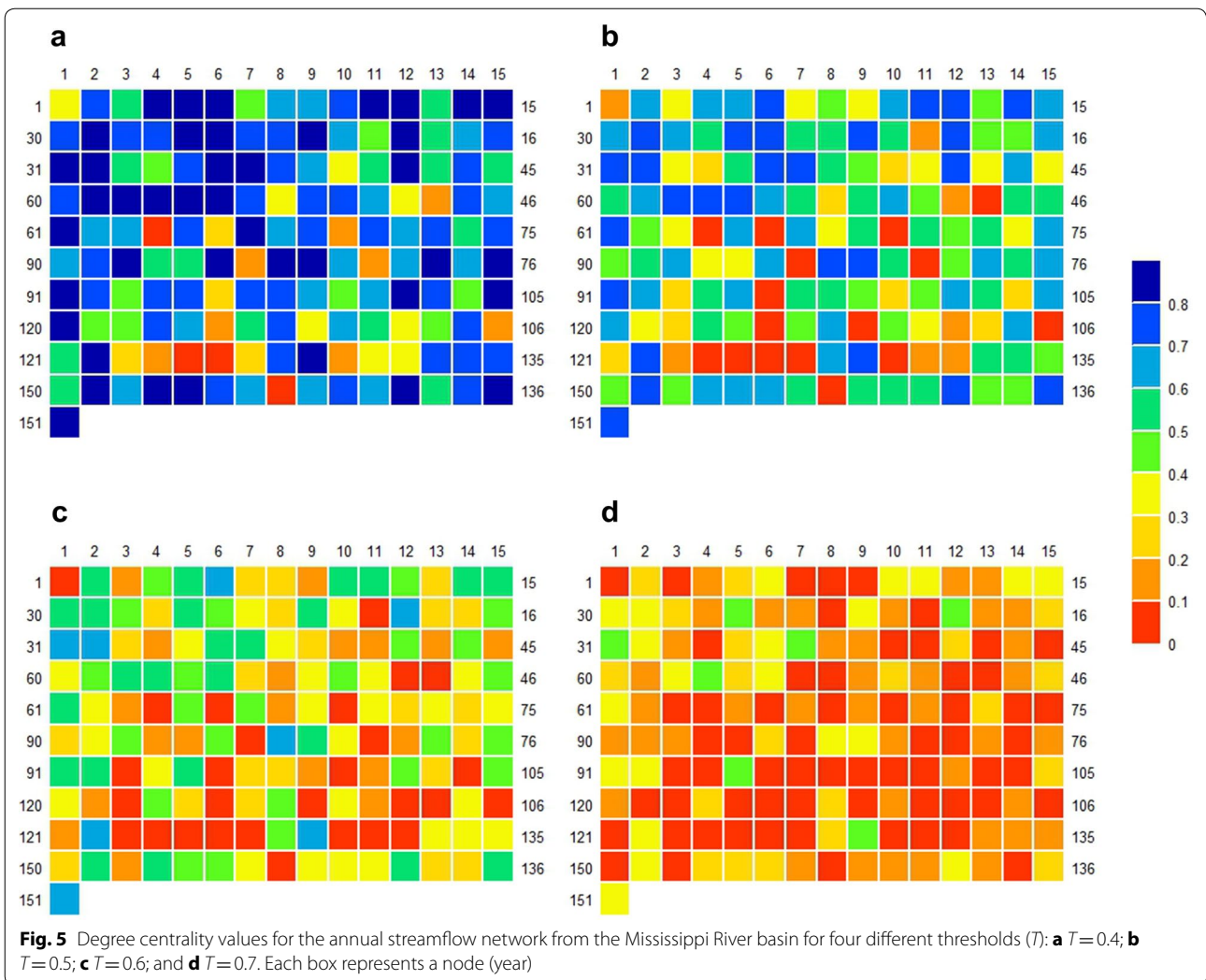
Analysis and results

Using the daily streamflow data of 151 years (October 1862–September 2013), the annual streamflow network for the Mississippi River basin at St. Louis, Missouri is constructed, following the procedure explained earlier. The annual streamflow network thus constructed has 151 nodes, corresponding to 151 years of daily data. Each node consists of 365 daily streamflow values (excluding the data for February 29 in leap years). This allows calculation of correlations in streamflow between each of the 151 nodes (years) with each and every other node in the network. In this study, the Pearson correlation coefficient is used to calculate the correlation. The correlations

in flow between nodes, in turn, allow identification of neighbors (i.e., links) for each and every node in the network, which is the key to the implementation of the degree centrality, clustering coefficient, and degree distribution methods. It is important to note that the correlation threshold (T) may significantly influence the identification of the neighbors (i.e., links), and hence, the outcomes of the methods. However, the optimum correlation threshold is not known a priori. To take this issue into account and examine the influence of threshold, eight different threshold values are considered in the analysis: 0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, and 0.8 (see Sivakumar and Woldemeskel (2014) for some details on the selection of the correlation threshold values). The results are presented next, where different threshold values may be considered for different methods to allow better visualization of the differences in results.

Degree centrality

Figure 5a–d, for instance, shows the results from the degree centrality analysis for the annual streamflow network from the Mississippi River basin at St. Louis, Missouri, for threshold values of 0.4, 0.5, 0.6, and 0.7, respectively. In these plots, a box corresponds to a node (i.e., there are 151 boxes in total), and the boxes are numbered from 1 to 151, corresponding to the year numbers. As normally expected, the degree centrality value (for



any given node) is found to decrease with an increase in the threshold value. However, the plots also indicate the enormous sensitivity of the degree centrality to the threshold level, as significant differences in the centrality values are observed between different thresholds. For instance, while more than 50% of the nodes (80 nodes) have degree centrality values exceeding 0.7 when $T=0.4$, only about 18% of the nodes (27 nodes) have degree centrality values exceeding 0.7 when $T=0.5$, and this number falls to zero when $T=0.6$ and $T=0.7$. This means that more than half the number of nodes (years) have connections with more than 70% of the rest of the network when $T=0.4$, but this number falls to just a quarter when $T=0.5$ and then to zero when $T \geq 0.6$. Indeed, when $T=0.7$, more than 40% of the nodes (63 nodes) have connections with only less than 10% of the other nodes. These observations suggest that the connections are only very little or even none when more stringent conditions are imposed, such as when $T \geq 0.5$ and especially when

$T \geq 0.6$, even considering the streamflow dynamics at the annual scale (where correlations and, thus, connections are normally expected to be much stronger when compared to those at the daily scale, for example, because of the presence of seasonality and “smoothing” at the annual scale).

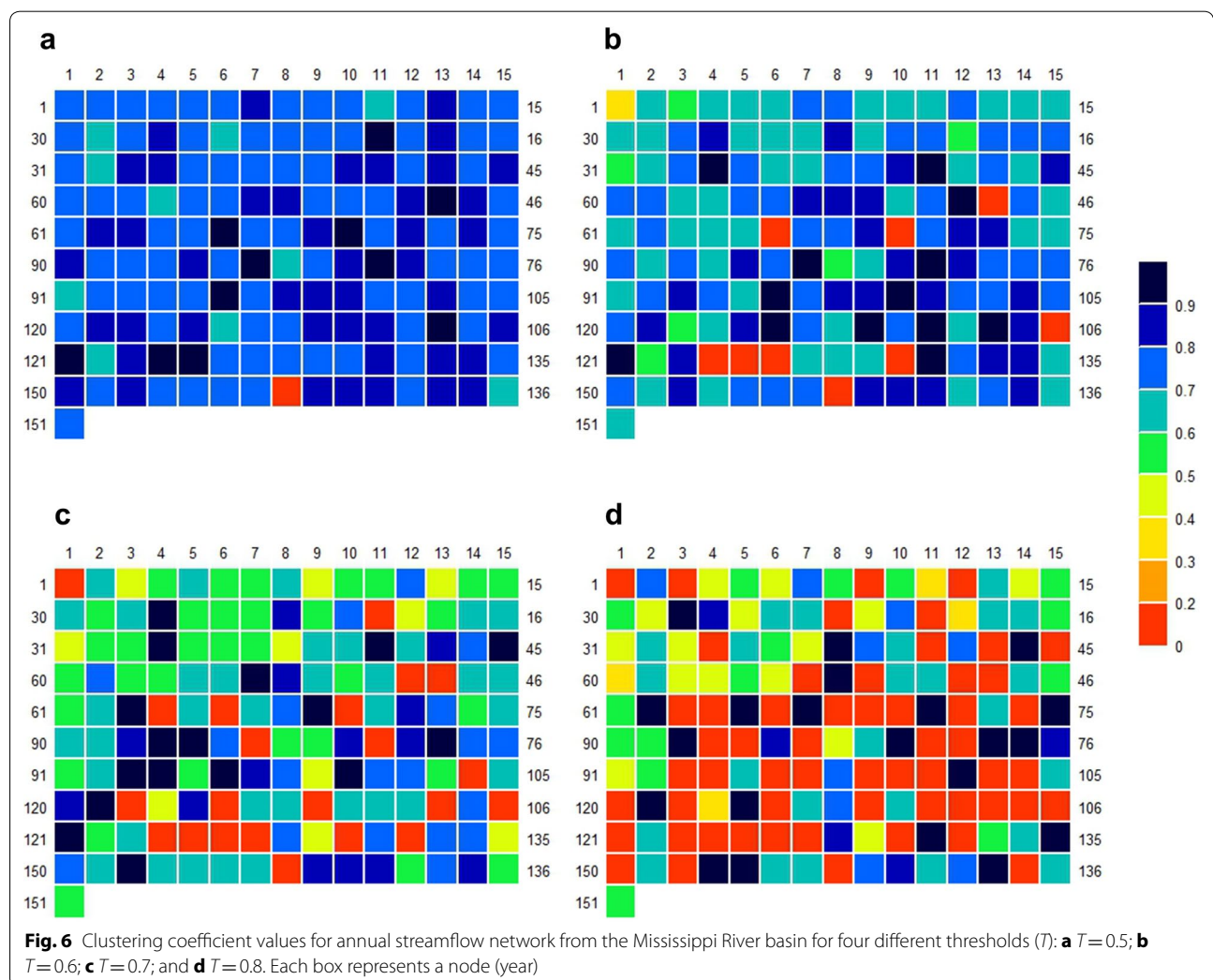
Overall, the results suggest that only a very few nodes (years), with very high degree centrality values, have great significance in terms of connections in the network especially when $T \geq 0.5$ (see the boxes colored in dark blue). Similarly, only a very few nodes, with very low degree centrality values, are found to have almost no significance in terms of connections, even for very low threshold values, such as $T=0.4$ and $T=0.5$ (see the boxes colored in red in Fig. 5a, b). It is also important to note that not all of the years that a given year has connection with are ‘closer’ in time (e.g., successive years), and some are very much apart in time. In other words, ‘proximity’ in time does not necessarily mean similarity in behavior,

at least when it is considered as part of a network as a whole. However, the results also indicate some kind of order, since at least some successive years show similar degree centrality values; see, for instance, nodes 55–59 (1916–1920) when $T=0.4$, nodes 56–58 (1917–1919) or nodes 124–127 (1985–1988) when $T=0.5$, nodes 123–127 (1984–1988) when $T=0.6$, and a number of stretches of nodes for $T=0.7$ (see the boxes colored in red). It is not clear why only a few nodes have great significance in terms of connections, why only a few other nodes have almost no significance, and why the rest of the nodes fall in between these two extremes—similar questions are also relevant for the clustering coefficient results (see below). An insight into the time series and some basic statistical characteristics (e.g., mean, standard deviation) of the daily flow series for the 151 years also does not offer any convincing explanation to these questions. Despite these questions (and indeed because of them), one can clearly recognize that the above results and

observations have important implications for long-term streamflow predictions (including in the use of methods that are based on temporal dependence) and potentially indicate the influence of large-scale climate patterns (and perhaps anthropogenic effects) on streamflow.

Clustering coefficient

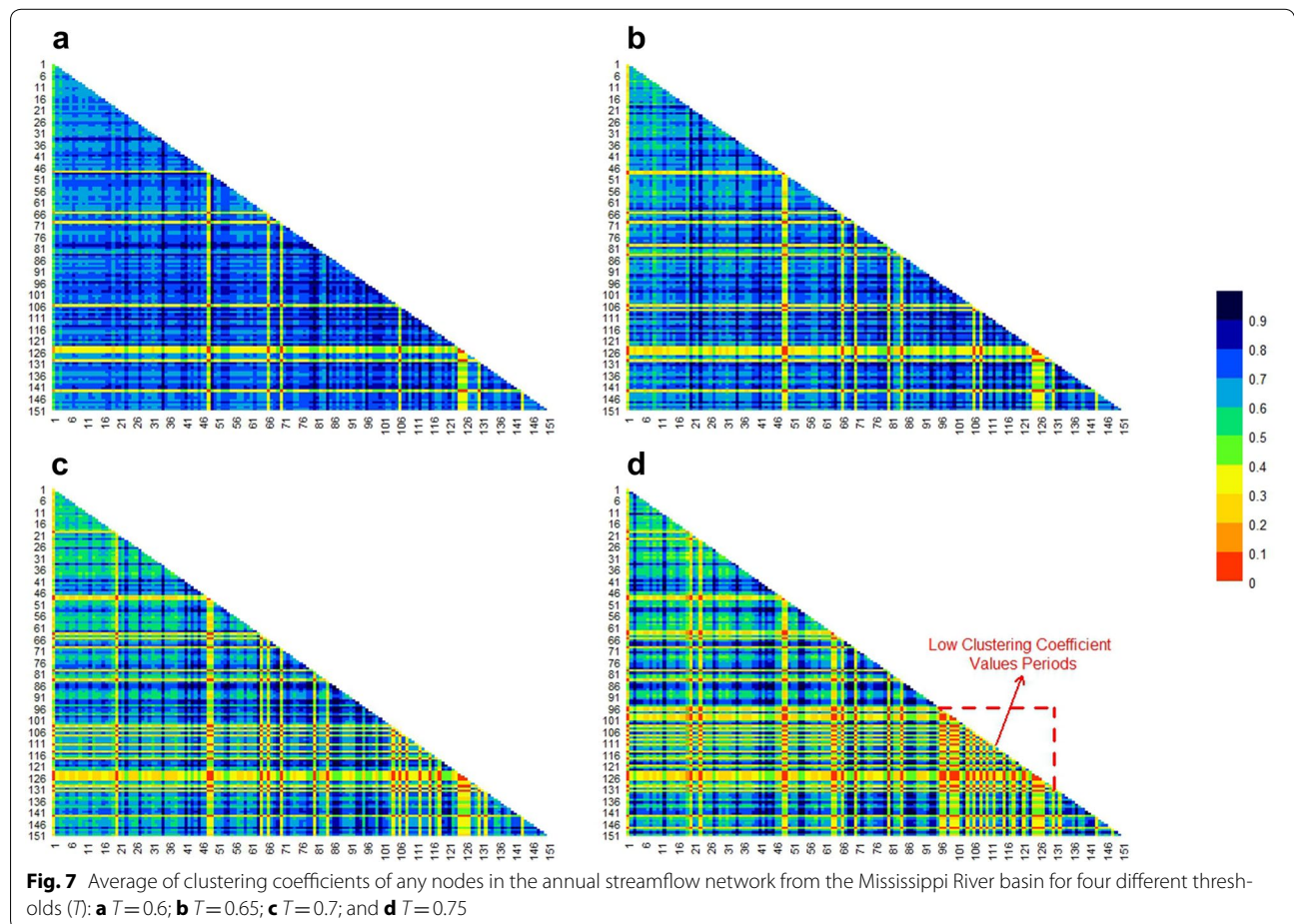
Figure 6a–d, for instance, shows the clustering coefficient values for the annual streamflow network from the Mississippi River basin at St. Louis, Missouri for threshold values of 0.5, 0.6, 0.7, and 0.8, with each box representing a node. Similar to the degree centrality, and as expected, the clustering coefficient value (for any given node) is found to decrease with an increase in the threshold and also shows significant sensitivity. When $T=0.5$, almost 90% of the nodes (137 nodes) have clustering coefficient values above 0.7, and about 52% of the nodes (79 nodes) have clustering coefficient values above 0.7 when $T=0.6$. This number becomes as low as 28% (43 nodes)



when $T=0.7$ and only 9% (13 nodes) when $T=0.8$. These results indicate that almost 90% of the nodes have reasonably good connections with the rest of the network (i.e., correlation ≥ 0.5), but only less than one-tenth of the nodes have strong connections (i.e., correlation ≥ 0.8), even at the annual scale. Similar observations can also be made in terms of very low clustering coefficient values. For instance, only one node has a clustering coefficient value below 0.2 when $T=0.5$, and only nine nodes have a clustering coefficient value below 0.2 when $T=0.6$ (see the boxes colored in red in Fig. 6a, b). The results also indicate that even some distant nodes (i.e., years far apart), with similar clustering coefficient values, may have strong connections in the overall network, even when they may or may not be connected between themselves. That is, they are ‘similar’ in some way, in the long-term evolution of streamflow dynamic system. In a similar vein, even ‘closer’ nodes (successive years) may behave very differently when considered as part of a network. Again, the reasons for these are unclear, and an insight into the time series and basic statistical characteristics (e.g., mean, standard deviation) of the flow series

does not offer any convincing explanation either. Nevertheless, it is clear that the clustering coefficient results have implications for streamflow predictions, especially when using methods that are based on temporal dependence, and also highlight the potential role of long-term climate change/variability, thus providing support to the results from the degree centrality method.

Although Fig. 6 provides useful information on the extent of connection of each node (year) with the rest of the 150 nodes of the network collectively, comparing the clustering coefficient value of each node with respect to each and every other node in the network on an individual basis may offer additional information. A simple way to do this may be to present the average of clustering coefficients of any two nodes for the entire network. This is done in Fig. 7, which shows the results for $T=0.6$, 0.65, 0.7, and 0.75—these four thresholds are presented for better visualization and discussion. The results generally show very high connections (i.e., average clustering coefficient >0.7) of each node with respect to each and every other node (light blue, dark blue, and black boxes) for $T=0.6$ (Fig. 7a), and to a certain extent, for $T=0.65$



(Fig. 7b). The connections become considerably weaker (yellow, orange, and red boxes) for $T=0.7$ (Fig. 7c) and more so for $T=0.75$ (Fig. 7d). The results also seem to indicate that a particular stretch of nodes, i.e., nodes 95–130 (1957–1992) (see the glaring yellow–orange–red color part, marked in Fig. 7d), have very poor connections with the rest of the network. Further discussion on this is made in the next section.

While the clustering coefficient values for each of the 151 nodes (Fig. 6) and their comparison with each and every other node (Fig. 7) indeed provide useful information about individual connections in the network, an even broader interest in this network-based study is the identification of the nature of the entire network, for development of an appropriate model. To this end, the clustering coefficient of the entire network, calculated as the average of the clustering coefficients for all the 151 nodes, is useful. The clustering coefficient values of the entire network for the eight different thresholds considered in this study (i.e., 0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, and 0.8) are 0.883, 0.835, 0.763, 0.656, 0.612, 0.560, 0.431, and 0.288, respectively. As normally expected, the clustering coefficient value decreases with an increase in the threshold value. The generally high clustering coefficient values (including for $T \geq 0.7$) seem to suggest that the network is not a purely random graph, as the clustering coefficient values for classical random networks are typically very low (close to zero, essentially due to random distribution of links), as mentioned in the methodology section earlier; see also, for example, Watts and Strogatz (1998). As the clustering coefficient for the annual streamflow

network is much higher than that for the classical random network but lower than the ones expected for fully connected networks (for which the clustering coefficient should be equal to 1.0), one may interpret that the network is a small-world network (e.g., Watts and Strogatz 1998) or a scale-free network (e.g., Barabási and Albert 1999) or some other type, as highlighted in the methodology section earlier. In the identification of the network type, the results from the degree distribution method could also offer some clues, and are presented next.

Degree distribution

Figure 8 presents the results from the degree distribution analysis of the annual streamflow network from the Mississippi River basin at St. Louis, Missouri for all the eight threshold levels considered in this study. The results are shown both in the normal scale (Fig. 8a) and in the log–log scale (Fig. 8b). The values are the complementary cumulative distribution, defined as the fraction of nodes with degree at least k and denoted as $p(K \geq k)$.

The results in Fig. 8 clearly show that the degree distribution for the annual streamflow network changes with respect to the correlation thresholds. For instance, when $T=0.3$, there are over 80% of the nodes with at least 100 neighbors. This number becomes over 60% when $T=0.4$, and less than 30% when $T=0.5$. For $T \geq 0.6$, the number of nodes with at least 100 neighbors is zero, indicating very poor connections in the network.

The shape of the degree distribution curves in Fig. 8 also offers some interesting observations. For low thresholds (say $T=0.3$, $T=0.4$, and also perhaps $T=0.5$), the curves

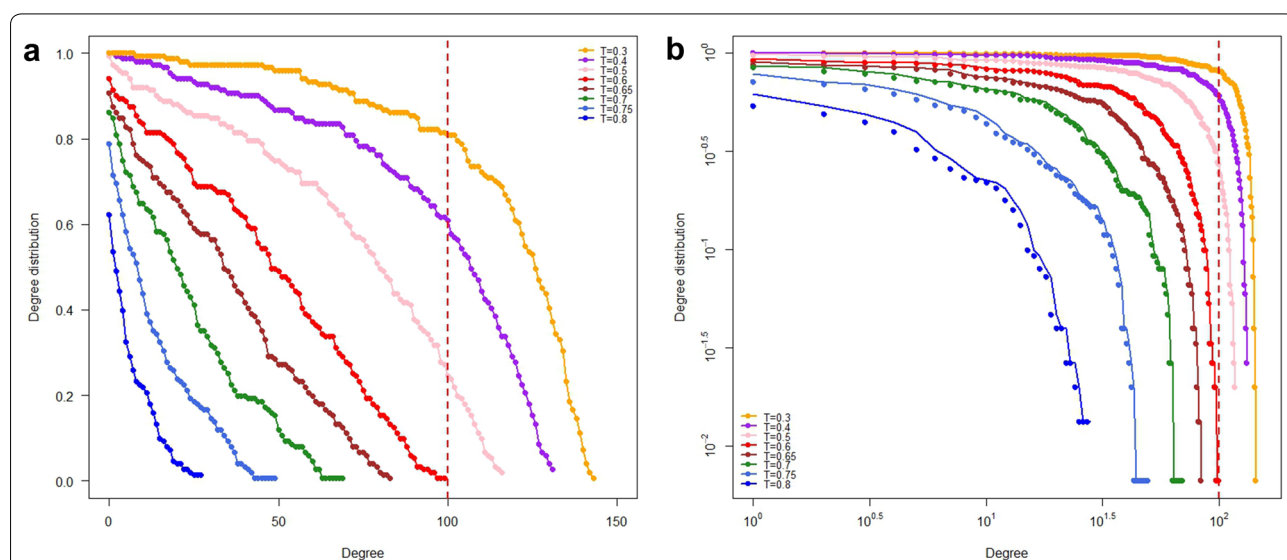


Fig. 8 Degree distribution for the annual streamflow network from the Mississippi River basin for eight different thresholds (T) (0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, and 0.8): **a** normal scale; and **b** log–log scale

seem to resemble exponential distribution. For high thresholds (say $T=0.8$, and $T=0.75$), the curves seem to resemble power-law distribution, especially at the tail. For medium thresholds (say $T=0.6$, 0.65 , and 0.7), the curves seem to resemble a distribution that is somewhere in between exponential and power-law, and perhaps a combination. With these observations, the annual streamflow network may be considered as a combination of exponential distribution and power-law distribution, with clear dependence on the correlation threshold level. This result has important implication for the selection of the type of model for annual streamflow dynamics.

Discussion of results

The results from the construction of annual streamflow network based on daily streamflow data and application of the degree centrality, clustering coefficient, and degree distribution methods to such a network are useful and interesting in several ways. A few important aspects are highlighted here.

Streamflow dynamics at the annual scale often exhibit a certain level of temporal correlation. However, the results from the present analysis do not readily indicate strong connections in streamflow dynamics between successive/different years (as a result of “annual cycle”) or between distant years (as a result of the influence of large-scale climate patterns and long-term evolution, including decadal cycles). The degree centrality results (Fig. 5) indicate that the streamflow dynamics in only a few years have great significance (or almost no significance) in terms of connections in the network of 151 years of data considered. Similarly, the clustering coefficient results (Fig. 6) indicate that the streamflow dynamics in only a very few years are very strongly (or very weakly) connected to the streamflow dynamics in all the other years of the 151-year period of study. Considering that there are also some differences between the few years identified in the degree centrality method and those identified in the clustering coefficient method, what makes such years highly significant (or almost insignificant) in the network or very strongly (or very weakly) connected in the network is unclear. However, the existence of these years seems to suggest the need to focus on such years in streamflow modeling (both for high flows and for low flows), especially in the long-term perspective. Whether these years reflect the influence of large-scale climate patterns and long-term climate change/variability (including decadal changes) is an important question to ask. The answer remains unknown, and this will be an important future investigation. What is clear, however, is that these results have important implications for studies on the use of methods based on temporal dependence for long-term streamflow modeling and prediction.

The clustering coefficient results (Fig. 6) suggest that the annual streamflow network is neither a purely random graph nor a regular network but something in between, such as a small-world network or a scale-free network or other. The degree distribution results (Fig. 8) suggest that the annual streamflow network exhibits exponential distribution or power-law (scale-free) distribution or a combination of both, depending on the correlation threshold level considered for studying connections in the network. Therefore, identification of the exact type of the network is still not complete and requires additional evidence for confirmation.

Another interesting observation comes from the clustering coefficient results, especially from the average of clustering coefficients of any two nodes for the entire network (Fig. 7). As can be seen from Fig. 7, when the average of clustering coefficients of any two nodes is considered, there is a certain stretch of nodes that exhibit very low connections (the yellow–orange–red colored part) with the rest of the network, depending upon the correlation threshold level. This is particularly clear for high threshold levels, such as the very low connections observed for nodes 95–130 (1957–1992) for $T=0.75$ (marked in Fig. 7d). What makes this stretch of nodes (i.e., period of time) to very weakly connect with the rest of the network is not clear. It is relevant to note, however, that the period 1950s–1990s corresponds to the period when a large number of dams were constructed across the Mississippi River. The natural flow of stream in the sub-basin of the St. Louis gaging station has and continues to be affected by many reservoirs and navigation dams in the upper Mississippi River basin and by many reservoirs and diversion for irrigation in the Missouri River basin (e.g., Alexander et al. 2012). The construction of most of the dams started in the 1950s and construction of dams ended in the 1990s.

It may be premature to associate the very weak connections in the annual streamflow network for the period 1950s–1990s with the influence of dam construction during the 1950s–1990s. However, the possible existence of such an association cannot be dismissed altogether. On the other hand, it may also be argued that, if the construction of dams was indeed a reason for very weak connections in the network, very weak connections should also be observed for the period after the 1990s. However, such is not the case in the clustering coefficient results, as the period after the 1990s exhibits better connections with the rest of the years compared to the period 1950s–1990s. One reason for this may be that there has been better regulation of flows since the 1990s, and only the period 1950s–1990s was severely influenced. These observations seem to suggest that the concepts of complex networks and their outcomes can offer physical explanations about the system dynamics.

Finally, it is important to remember that the streamflow dynamics examined in this study are only at the annual scale. Since streamflow dynamic properties can, and often, change with temporal scale, whether the results obtained in this study for the annual scale would still hold true for any other temporal scale is an obvious question to ask. Such a question still remains to be answered, and will be investigated in a future study. Nevertheless, our opinion, for the moment, especially based on nonlinear dynamic studies on streamflow (and other hydrologic data) and complex network studies on rainfall at different temporal scales, is that the streamflow network properties (including degree centrality, clustering coefficient, and degree distribution) may change for other temporal scales, despite the possible presence of scaling (or fractal) behavior in streamflow; see Sivakumar (2001), Sivakumar et al. (2001, 2004, 2007), Regonda et al. (2004), Salas et al. (2005), Jha and Sivakumar (2017), and Naufan et al. (2017) for some details. We hope to provide more reliable and convincing answers to this question in a future study, as we are currently conducting additional research on network properties in terms of scale and network size.

Conclusions

Understanding the temporal dynamics of streamflow (and other hydrologic processes) continues to be challenging. This study employed modern concepts of network theory, i.e., complex networks, for studying the temporal dynamics of streamflow, with particular focus on the annual scale, i.e., year-to-year connections. It adopted a new approach to construct the streamflow network at the annual scale. Instead of using the annual streamflow data (mean or accumulated) and considering each year as a node with just one streamflow value, the study proposed to use the daily streamflow data, with each year serving as a node in the network and with each node having a time series of (365) daily streamflow values. The approach was implemented on the streamflow data observed over a long period of 151 years from the Mississippi River basin at St. Louis, Missouri. The properties of the network were examined using degree centrality, clustering coefficient, and degree distribution methods.

The results from the present analysis regarding the temporal connections in annual streamflow are useful and interesting in many ways. The degree centrality results suggest the presence of a very few significant (or almost insignificant), but not necessarily consecutive, years in the studied period of 151 years. The clustering coefficient results suggest the presence of a few years that are connected very strongly (or very weakly) to the rest of the years and that the annual streamflow network is neither a purely random network nor a regular network,

but something in between (e.g., small-world or scale-free or other). The degree distribution results also seem to support this, to a certain extent, indicating exponential behavior or power-law behavior or their combination in the distribution of links in the network. The clustering coefficient results also seem to suggest the influence of dam construction (and other anthropogenic influences) on the annual streamflow dynamics, especially through identifying a stretch of period (around the 1950s–1990s) with very weak connections when compared to the rest of the period of data.

All these results have important implications for studies on the temporal dynamics of streamflow at the annual scale (and at other scales), and hence, for streamflow modeling and prediction. Among these are (1) use of models that particularly assume temporal dependence; (2) identification of appropriate model for studying connections in streamflow; (3) long-term predictability of streamflow; (4) influence of large-scale climate patterns and long-term climate change/variability; and (5) influence of anthropogenic factors.

The outcomes of the present study lead to several potential future directions. In addition to studying the issues associated with the implications above, one particularly useful area of research may be to improve the construction of the streamflow network based on the available data. To this end, nonlinear data reconstruction and related concepts that use a single-variable (or multi-variable) time series to reconstruct a multi-dimensional phase space, such as phase space reconstruction (e.g., Packard et al. 1980; Takens 1981) and dimensionality (e.g., Grassberger and Procaccia 1983; Kennel et al. 1992), could provide new avenues. For instance, instead of using the HVG or the approach proposed in the present study, one may reconstruct the streamflow data in a multi-dimensional phase space and then construct the network based on the points (vectors) in the reconstructed phase space. This way, each point in the phase space can serve as a node in the network and the distances between the points can serve to identify the links. Such a phase space reconstruction approach for network construction is certainly appealing, especially considering that it has already proved useful for representing the temporal dynamics of streamflow (and other hydrologic processes), both in the Mississippi River basin (e.g., Sivakumar and Jayawardena 2002; Sivakumar and Wallender 2005; Sivakumar et al. 2007) and in many other basins around the world (e.g., Regonda et al. 2004; Salas et al. 2005; Sivakumar and Singh 2012; Jothiprakash and Fathima 2013; Tongal et al. 2013). Research in this direction is currently underway. Indeed, whether, and how, the temporal connections identified from the combination of phase space reconstruction and complex networks can be useful for

streamflow prediction and catchment classification is also being studied. We hope to report the details of such studies in the near future.

Abbreviations

DHVG: directed horizontal visibility graph; HVG: horizontal visibility graph; USGS: United States Geological Survey.

Authors' contributions

All authors contributed extensively to the work presented in this manuscript. BS designed the methodology for network construction of streamflow data and guided the research. XH carried out the analysis, with the assistance of FMW in developing the codes and MGA in compiling the results. All authors discussed the results and implications, and contributed to manuscript preparation/modifications, including revisions. All authors read and approved the final manuscript.

Author details

¹ School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia. ² Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA. ³ Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai 400 076, India. ⁴ Department of Civil Engineering, University Federal de Minas Gerais, Belo Horizonte, Brazil.

Acknowledgements

Bellie Sivakumar acknowledges the financial support from the Australian Research Council through the Future Fellowship Grant (FT110100328). The authors are thankful to the two anonymous reviewers for their constructive comments and useful suggestions on an earlier version of the manuscript, which helped improve the quality and presentation of our work.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The Mississippi River streamflow data used in this study are available in the public domain at the USGS National Water Information System website (<http://nwis.waterdata.usgs.gov/nwis>) and can be downloaded for free. The data and related materials may also be requested from the corresponding author.

Consent for publication

The manuscript does not include any personal data.

Ethics approval and consent to participate

The study does not involve any human subjects.

Funding

This study was supported by the Australian Research Council (ARC) Future Fellowship Grant (FT110100328).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 1 September 2017 Accepted: 3 March 2018

Published online: 22 March 2018

References

Alexander JS, Wilson RC, Green WR (2012) A brief history and summary of the effects of river engineering and dams on the Mississippi River system and delta. US geological survey circular, vol 1375

Bak P (1996) How nature works: the science of self-organized criticality. Springer-Verlag, New York

Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512

Barnsley FM (2012) *Fractals everywhere*. Dover, New York

Barrat A, Weigt M (2000) On the properties of small-world networks. *Eur Phys J B* 13:547–560

Bavelas A (1948) A mathematical model for group structure. *Hum Org* 7:16–30

Boers N, Bookhagen B, Marwan N, Kurths J, Marengo J (2013) Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System. *Geophys Res Lett* 40:1–7. <https://doi.org/10.1002/grl.50681>

Braga AC, Alves LGA, Costa LS, Ribeiro AA, de Jesus MMA, Tateishi AA, Ribeiro HV (2016) Characterization of river flow fluctuations via horizontal visibility graphs. *Phys A* 444:1003–1011

Clauset A, Rohilla Shalizi C, Newman MEJ (2010) Power-law distribution in empirical data. *SIAM Rev* 51:661–703

Czuba JA, Foufoula-Georgiou E (2014) A network-based framework for identifying potential synchronizations and amplifications of sediment delivery in river basins. *Water Resour Res* 50:3826–3851

Czuba JA, Foufoula-Georgiou E (2015) Dynamic connectivity in a fluvial network for identifying hotspots of geomorphic change. *Water Resour Res* 51:1401–1421

Estrada E (2012) *The structure of complex networks: theory and applications*. Oxford University Press, Oxford

Fang K, Sivakumar B, Woldemeskel FM (2017) Complex networks, community structure, and catchment classification. *J Hydrol* 545:478–493

Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99:7821–7826

Grassberger P, Procaccia I (1983) Measuring the strangeness of strange attractors. *Physica D* 9:189–208

Halverson MJ, Fleming SW (2015) Complex network theory, streamflow, and hydrometric monitoring system design. *Hydrol Earth Syst Sci* 19:3301–3318

Jeong H, Tomber B, Albert R, Oltvai ZN, Barabási A-L (2000) The large-scale organization of metabolic networks. *Nature* 407:651–654

Jeong H, Mason S, Barabási A-L, Oltvai ZN (2001) Lethality and centrality in protein networks. *Nature* 411:41–42

Jha SK, Sivakumar B (2017) Complex networks for rainfall modeling: spatial connections, temporal scale, and network size. *J Hydrol* 554:482–489

Jha SK, Zhao H, Woldemeskel FM, Sivakumar B (2015) Network theory and spatial rainfall connections: an interpretation. *J Hydrol* 527:13–19

Jothiprakash V, Fathima TA (2013) Chaotic analysis of daily rainfall series in the Koyana Reservoir Catchment Area. *Stoch Environ Res Risk Assess* 27(6):1371–1381

Keller EF (2005) Revisiting 'scale-free' networks. *BioEssays* 27:1060–1068

Kennel MB, Brown R, Abarbanel HDI (1992) Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys Rev A* 45(6):3403–3411

Kim D-H, Noh JD, Jeong H (2004) Scale-free trees: the skeletons of complex networks. *Phys Rev E* 70:046126

Labat D, Masbou J, Beaulieu E, Mangin A (2011) Scaling behavior of the fluctuations in stream flow at the outlet of karstic watersheds. *France J Hydrol* 410(3):162–168

Lacasa L, Luque B, Ballesteros F, Luque J, Nuño J (2008) From time series to complex networks: the visibility graph. *Proc Natl Acad Sci USA* 105:4972–4975

Leavitt HJ (1951) Some effects of certain communication patterns on group performance. *J Abnor Soc Psych* 46:38–50

Malik N, Bookhagen B, Marwan N, Kurths J (2012) Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Clim Dyn* 39:971–987

Mandelbrot BB (1983) *The fractal geometry of nature*. W. H. Freeman, New York

Naufan I, Sivakumar B, Woldemeskel FM, Raghavan SV, Vu MT, Liong S-Y (2018) Spatial connections in regional climate model rainfall outputs at different temporal scales: application of network theory. *J Hydrol* 556:1232–1243. <https://doi.org/10.1016/j.jhydrol.2017.05.029>

Newman MEJ (2001) The structure of scientific collaboration networks. *Proc Nat Acad Sci USA* 98:404–409

Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E* 64:026118

Özger M, Mishra AK, Singh VP (2013) Seasonal and spatial variations in the scaling and correlation structure of streamflow data. *Hydrol Process* 27(12):1681–1690

- Packard NB, Crutchfield JP, Farmer JD, Shaw RS (1980) Geometry from a time series. *Phys Rev Lett* 45(9):712–716
- Peckham S, Gupta V (1999) A reformulation of Horton's laws for large river networks in terms of statistical self-similarity. *Water Resour Res* 35:2763–2777
- Regonda S, Sivakumar B, Jain A (2004) Temporal scaling in river flow: can it be chaotic? *Hydrol Sci J* 49(3):373–385
- Rinaldo A, Banavar JR, Maritan A (2006) Trees, networks, and hydrology. *Water Resour Res* 42:W06D07. <https://doi.org/10.1029/2005wr004108>
- Rinaldo A, Rigon R, Banavar JR, Maritan A, Rodriguez-Iturbe I (2014) Evolution and selection of river networks: statics, dynamics, and complexity. *Proc Nat Acad Sci USA* 111(7):2417–2424
- Rodriguez-Iturbe I, Rinaldo A (1997) *Fractal river networks: chance and self-organization*. Cambridge University Press, New York
- Salas JD, Delleur JW, Yevjevich V, Lane WL (1995) *Applied modeling of hydrologic time series*. Water Resources Publications, Littleton
- Salas JD, Kim HS, Eykholt R, Burlando P, Green TR (2005) Aggregation and sampling in deterministic chaos: implications for chaos identification in hydrological processes. *Nonlinear Process Geophys* 12:557–567
- Scarsoglio S, Laio F, Ridolfi L (2013) Climate dynamics: a network-based approach for the analysis of global precipitation. *PLoS ONE* 8(8):e71129. <https://doi.org/10.1371/journal.pone.0071129>
- Serinaldi F, Kilsby CG (2016) Irreversibility and complex network behavior of stream flow fluctuations. *Phys A* 450:585–600
- Sivakumar B (2001) Rainfall dynamics at different temporal scales: a chaotic perspective. *Hydrol Earth Syst Sci* 5(4):645–651
- Sivakumar B (2015) Networks: a generic theory for hydrology? *Stoch Environ Res Risk Assess* 29:761–771
- Sivakumar B, Berndtsson R (2010) *Advances in data-based approaches for hydrologic modeling and forecasting*. World Scientific Publishing Company, Singapore
- Sivakumar B, Jayawardena AW (2002) An investigation of the presence of low-dimensional chaotic behavior in the sediment transport phenomenon. *Hydrol Sci J* 47(3):405–416
- Sivakumar B, Singh VP (2012) Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework. *Hydrol Earth Syst Sci* 16:4119–4131
- Sivakumar B, Wallender WW (2005) Predictability of river flow and sediment transport in the Mississippi River basin: a nonlinear deterministic approach. *Earth Surf Process Landf* 30:665–677
- Sivakumar B, Woldemeskel FM (2014) Complex networks for streamflow dynamics. *Hydrol Earth Syst Sci* 18:4565–4578
- Sivakumar B, Woldemeskel FM (2015) A network-based analysis of spatial rainfall connections. *Environ Modell Softw* 69:55–62
- Sivakumar B, Sorooshian S, Gupta HV, Gao X (2001) A chaotic approach to rainfall disaggregation. *Water Resour Res* 37(1):61–72
- Sivakumar B, Wallender WW, Puente CE, Islam MN (2004) Streamflow disaggregation: a nonlinear deterministic approach. *Nonlinear Process Geophys* 11:383–392
- Sivakumar B, Jayawardena AW, Li WK (2007) Hydrologic complexity and classification: a simple data reconstruction approach. *Hydrol Process* 21(20):2713–2728
- Suweis S, Konar M, Dalin C, Hanasaki N, Rinaldo A, Rodriguez-Iturbe I (2011) Structure and controls of the global virtual water trade network. *Geophys Res Lett* 38:L10403. <https://doi.org/10.1029/2011GL046837>
- Takens F (1981) Detecting strange attractors in turbulence. In: Rand DA, Young LS (eds) *Dynamical systems and turbulence*. Lecture notes in mathematics, vol 898. Springer, Berlin, pp 366–381
- Tang Q, Liu J, Liu H (2010) Comparison of different daily streamflow series in US and China, under a viewpoint of complex networks. *Mod Phys Lett B* 24(14):1541–1547
- Tongal H, Berndtsson R (2014) Phase-space reconstruction and self-exciting threshold modeling approach to forecast water levels. *Stoch Environ Res Risk Assess* 28:955–971
- Tongal H, Demirel MC, Boojj MJ (2013) Seasonality of low flows and dominant processes in the Rhine River. *Stoch Environ Res Risk Assess* 27:489–503
- Tsonis AA, Roebber PJ (2004) The architecture of the climate network. *Phys A* 333:497–504
- Wasserman S, Faust K (1994) *Social network analysis*. Cambridge University Press, Cambridge
- Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393:440–442
- Xu K, Yang D, Yang H, Li Z, Qin Y, Shen Y (2015) Spatio-temporal variation of drought in China during 1961–2012: a climatic perspective. *J Hydrol* 526:253–264
- Zaliapin I, Foufoula-Georgiou F, Ghil M (2010) Transport on river networks: a dynamic tree approach. *J Geophys Res* 115:F00A15. <https://doi.org/10.1029/2009jf001281>

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
