

# Teleconnection Patterns of River Water Quality Dynamics Based on Complex Network Analysis

jiping jiang<sup>1</sup>, Sijie Tang<sup>1</sup>, Bellie Sivakumar<sup>2</sup>, Tianrui Pang<sup>3</sup>, Na Wu<sup>4</sup>, and Yi Zheng<sup>1</sup>

<sup>1</sup>Southern University of Science and Technology

<sup>2</sup>Indian Institute of Technology Bombay

<sup>3</sup>Harbin Institute of Technology

<sup>4</sup>Tongji University

January 20, 2023

## Abstract

Water quality in rivers is influenced by natural factors and human activities that interact in complex and nonlinear ways, which make water quality modelling a challenging task. The concepts of complex networks (CN), a recent development in network theory, seem to provide new avenues to unravel the connections and dynamics of water quality phenomenon, including clandestine teleconnections. This study aims to explore the spatial patterns of water quality using the CN concepts, at both catchment scale and larger national scale. Three major water quality parameters, i.e. dissolved oxygen (DO), permanganate index (COD Mn), and ammonia nitrogen (NH 3-N) are considered for analysis. Weekly data over a period of 12 years (since 2006) from 91 monitoring stations across China are analysed. Degree centrality and clustering coefficient methods are employed. The results show that the degree centrality and clustering coefficients values for water quality indicators is  $DO > NH\ 3-N > COD\ Mn$  at both basin scale and national scale. Since COD Mn is more sensitive to the upstream point source pollution, as it depends upon the locality and human activities, it leads to a higher heterogeneity of CN indexes even among spatially closer stations. NH 3-N comes next due to the identical pollution level and degradation process in a certain spatial extension. Meanwhile, DO shows good regional connectivity in line with the strong diffusivity. However, the CN characteristic is relatively inconspicuous in large basins and nationwide scale, which indicates the regional impact on water quality fluctuation and CN analysis. These original findings boost a comprehensive understanding of water quality dynamics and enlighten novel methods for environment system analysis and watershed management.

## Hosted file

essoar.10512718.1.docx available at <https://authorea.com/users/578377/articles/620351-teleconnection-patterns-of-river-water-quality-dynamics-based-on-complex-network-analysis>

## Teleconnection Patterns of River Water Quality Dynamics Based on Complex Network Analysis

Jiping Jiang<sup>a,b\*</sup>, Sijie Tang<sup>a</sup>, Bellie Sivakumar<sup>c</sup>, Tianrui Pang<sup>b</sup>, Na Wu<sup>d</sup>, Yi Zheng<sup>a</sup>

<sup>a</sup> School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

<sup>b</sup> State Key Laboratory of Urban Water Resource and Environment, School of Environment, Harbin Institute of Technology, Harbin 150090, China

<sup>c</sup> Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India

<sup>d</sup> College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China

**Abstract:** Water quality in rivers is influenced by natural factors and human activities that interact in complex and nonlinear ways, which make water quality modelling a challenging task. The concepts of complex networks (CN), a recent development in network theory, seem to provide new avenues to unravel the connections and dynamics of water quality phenomenon, including clandestine teleconnections. This study aims to explore the spatial patterns of water quality using the CN concepts, at both catchment scale and larger national scale. Three major water quality parameters, i.e. dissolved oxygen (DO), permanganate index ( $\text{COD}_{\text{Mn}}$ ), and ammonia nitrogen ( $\text{NH}_3\text{-N}$ ) are considered for analysis. Weekly data over a period of 12 years (since 2006) from 91 monitoring stations across China are analysed. Degree centrality and clustering coefficient methods are employed. The results show that the degree centrality and clustering coefficients values for water quality indicators is  $\text{DO} > \text{NH}_3\text{-N} > \text{COD}_{\text{Mn}}$  at both basin scale and national scale. Since  $\text{COD}_{\text{Mn}}$  is more sensitive to the upstream point source pollution, as it depends upon the locality and human activities, it leads to a higher heterogeneity of CN indexes even among spatially closer stations.  $\text{NH}_3\text{-N}$  comes next due to the identical pollution level and degradation process in a certain spatial extension. Meanwhile, DO shows good regional connectivity in line with the strong diffusivity. However, the CN characteristic is relatively inconspicuous in large basins and nationwide scale, which indicates the regional impact on water quality fluctuation and CN analysis. These original findings boost a comprehensive understanding of water quality dynamics and enlighten novel methods for environment system analysis and watershed management.

**Keywords:** Water quality; Complex network; Degree centrality; Clustering coefficient; Teleconnection; China's rivers

## Introduction

Complexity is an important non-linear characteristic of a river water system (Kirchner and Neal 2013). The hydrological and water quality processes of rivers exhibit complexities in temporal (Jiang et al. 2020) and spatial connection (Fang et al. 2017). Because of the terrain, land use, surface-ground water interaction, regional meteorology, and human activities, the spatial patterns in water quality time series in a watershed is bound to have complex correlations. Such correlations and connections can be due to events that happen far away geographically, which is called teleconnection (Ashok et al. 2007). Analysing the regularity and possible causality of these correlations will help deepen our understanding of the energy and solute transport in river basins from the perspective of system theory, and help significantly advance our ability to model, predict, and manage river water quality dynamics. In this regard, complex network (CN) theory, which has been developing rapidly in recent years, is an effective modelling tool to help us gain knowledge on the teleconnection in water quality dynamics.

The concepts of CNs (Watts and Strogatz 1998) emerged from earlier developments in network theory or graph theory, topology, trees, and random graph theory ((Erdős and Rényi 1960, König 1936). Complex networks have been applied in many different fields and associated problems, including DNA transcription in molecular biology (Zheng and Flanagan 2017), biodiversity CN for a better understanding of species interactions (Bascompte 2009), predictive power of the behaviour of techno-social systems (Vespignani 2009), sustainability analysis of social-ecological systems (Ostrom 2009), analysis of vaccine efficacy in the vaccination behaviour (Huang et al. 2020), human information processing (Lynn et al. 2020), mathematical epidemic analysis (Wang et al. 2019), and the graph learning (Lynn and Bassett 2020).

In the field of hydrology, CN is a powerful tool to extract information from large high-dimensional datasets (Kurths et al. 2019). It, thus, can be used to unravel the connections in a variety of systems, as has been demonstrated for the global influence of the El Niño–Southern Oscillation (ENSO) on regional rainfall (Agarwal 2019, Ferster et al. 2018), influence of the Atlantic Meridional Overturning Circulation (AMOC) on air surface temperature (Agarwal et al. 2019), dominant climate modes (Agarwal et al. 2018, Halverson and Fleming 2015), catchment classification indicating hydrologic similarity (Fang et al. 2017), short/long range spatial connections of rainfall (Agarwal et al. 2018, Boers et al. 2014, Jha et al. 2015), and spatial and temporal hydrologic patterns (Halverson and Fleming 2015, Konapala and Mishra 2017, Sivakumar and Woldemeskel 2014).

The variability in river water quality parameters is the result of a complex and nonlinear influences of environmental factors (external factors) and regional characteristics (internal factors) and, therefore, connections are the key to understand the water quality dynamics. In this regard, CN provides strong the-

oretical foundations and new ideas and facilitate water environment planning and management. With data-intensive investigations becoming a new scientific discovery paradigm, the concepts of CN can go a long way in our ability to model, predict, and manage water quality and river systems.

This work collected intensive water quality data in plenty water quality stations and utilize network indexes to explore the connections and teleconnections of typical water quality parameters. The Huaihe River basin was selected as a large watershed scale study area and large rivers covering east China were selected as a large region scale study area. Based on the linear correlation between stations, the degree centrality and clustering coefficient were calculated to characterize the statistical characteristics of water quality CNs, and to reveal the potential teleconnection of water quality processes.

1.

## Study Areas and Methods

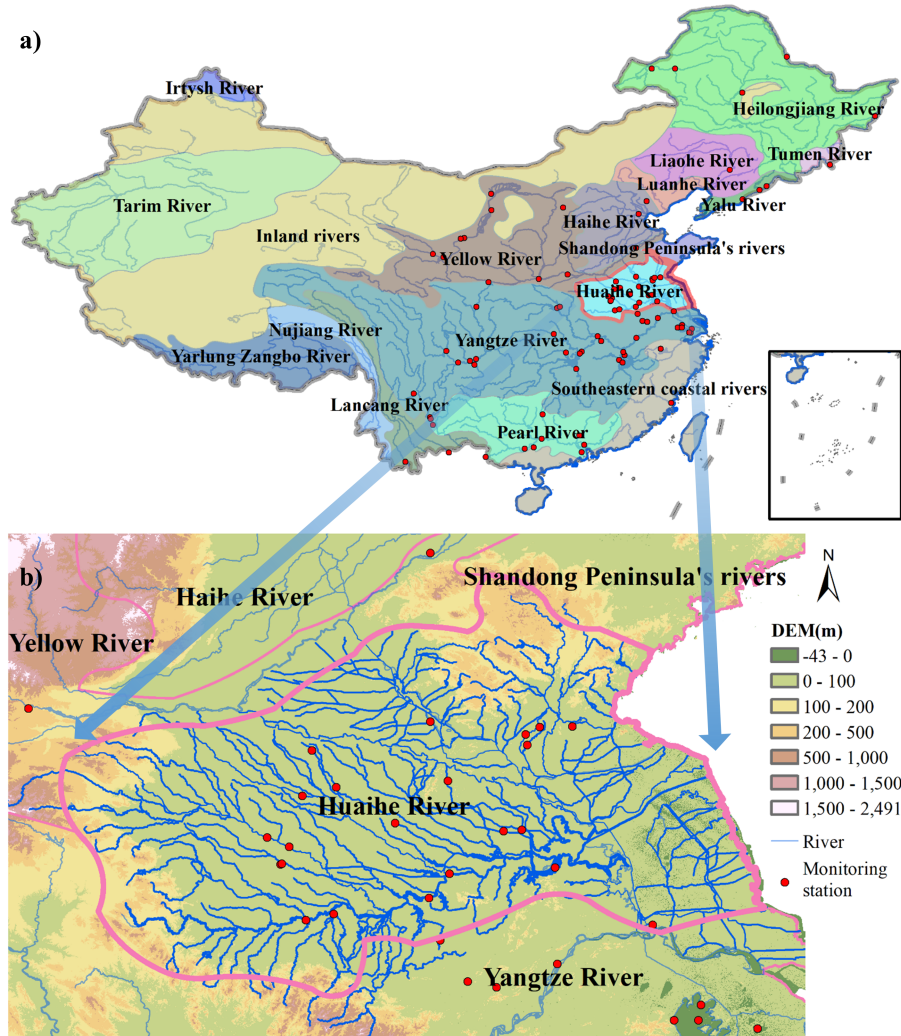
(a)

### Study areas and monitoring campaign

In this study, two scales of water quality system in China are studied: watershed scale and national scale. For the watershed scale, the Huaihe River is considered, and the major rivers in China are considered for the national (or large region) scale. They are both significantly impacted by human activities.

#### Huaihe River basin

The Huaihe River basin ( $30^{\circ}55' \sim 36^{\circ}36' \text{ N}$ , and  $111^{\circ}55' \sim 121^{\circ}25' \text{ E}$ ) is located in eastern China (**Fig. 1**), between the Yangtze River basin and the Yellow River basin. It flows through five provinces of China, namely Hubei, Henan, Anhui, Shandong, and Jiangsu. It is the seventh largest river basin in China, with a drainage area of  $270,000 \text{ km}^2$  approximately. The population living in the basin is 165 million. The basin average annual precipitation is about 894 mm, of which more than 70% occurs during the flood season from June to October. The spatial distribution of precipitation is quite uneven, decreasing from 1,400 mm in the southern mountain region to less than 700 mm in the northern region near the Yellow River. Although the basin-average annual runoff is about 230 mm, the temporal and spatial distributions are highly uneven due to climatic variability.



**Fig. 1.** Study area and monitoring campaign in China and Huaihe River basin. a) the major basins in China and the location of Huaihe River basin; b) river network in Huaihe River basin.

### China's major rivers

The large region scale (or national scale) study area includes a total of 63 major river systems in China as shown in **Fig. 1**. Some of these are the Yellow River, Yangtze River, Pearl River, Haihe River, Huaihe River, Heilongjiang River, Liaohe River, Taihu Lake, Chaohu Lake, and Dianchi Lake.

## Monitoring campaign and data processing

The water quality data for this study are obtained from the China National Environmental Monitoring Centre that operates 148 automatic water quality monitoring stations all over China, 91 of which are available and effective displayed as red dots in **Fig.1**. The weekly monitoring values are used for analysis. The water quality parameters monitored consist of pH, dissolved oxygen (DO), total organic carbon (TOC), permanganate index ( $\text{COD}_{\text{Mn}}$ ), and ammonia nitrogen ( $\text{NH}_3\text{-N}$ ). Among these, we mainly focus on DO,  $\text{COD}_{\text{Mn}}$ , and  $\text{NH}_3\text{-N}$  in this study. Most of the stations considered in this study are located on the east side of Heihe-Tengchong line, in economically developed area.

Data from 91 active stations are used for national scale research, and the time span is from Jan 1<sup>st</sup>, 2004 to Jun 2<sup>nd</sup>, 2016. Twenty-five stations are used for the study of the Huaihe River basin, though 27 actually exist. The time span of data considered is from the 29<sup>th</sup> week of 2008 to the 22<sup>nd</sup> week of 2016. For any missing data, the value is estimated by interpolating the continuous monitoring results in the previous and following three weeks. If data for such weeks are not available, then the average value of the same weeks in the previous and following years are taken.

1.

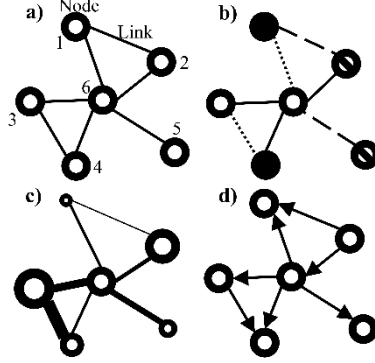
## Theory of complex network

(a)

### Concept of network

Basically, a network is a set of points and lines, in which various types of nodes or vertices may be connected together by all sorts of links or edges. A series of indexes are introduced to characterise the pattern of CN, including degree centrality, clustering coefficient, average shortest path length, degree distribution, and betweenness. These simple statistics can be used to quantify the properties of networks and to define the types, e.g. regular graph, random graph, small-world network, and scale-free network.

The network shown in **Fig. 2**, for example, can be represented as  $G = \{P, E\}$ , where  $P$  is a set of 6 nodes and  $E$  is a set of 7 links, namely  $P = \{1, 2, 3, 4, 5, 6\}$  and  $E = \{\{1, 2\}, \{1, 6\}, \{2, 6\}, \{3, 6\}, \{4, 6\}, \{5, 6\}, \{3, 4\}\}$ . **Fig. 2a** shows the simplest form of network which contains identical nodes and links; however, the network may have different types of nodes and/or links (**Fig. 2b**); the nodes and links may have a variety of properties, such as weight (**Fig. 2c**); and the connection may be directed, pointing in one direction (**Fig. 2d**). Besides, more particular features, such as multilink, self-link, and hyperlink, also exist in more CNs of real systems.



**Fig. 2.** Different networks (Sivakumar 2015): a) an undirected network with identical nodes and links; b) a network with more than one type of nodes and links; c) a network with weighted nodes and links; and d) a directed network.

### Indexes of complex network

The number of links  $k_i$  connected to a node  $i$  is called as **node degree**. Then the **degree centrality** is defined as follow:

which indicates the influence of node  $i$  in the network. The **degree distribution**,  $P(k)$ , represents the probability that a randomly selected node has exactly  $k$  neighbours, i.e. node degree of  $k$ . The common types of degree distribution are exponential distribution, power-law distribution and Poisson distribution (Krapivsky et al. 2001).

Another important property of networks is the tendency to cluster, which is quantified by the **clustering coefficient** (Wasserman and Faust 1994). Considering a selected node  $i$  with the node degree of  $k_i$ , there is supposed to be  $k_i(k_i-1)/2$  links at most between the  $k_i$  neighbours. Then the clustering coefficient of node  $i$  can be given as follows:

where  $E_i$  is the actually existing links in the cluster. Further, the clustering coefficient of a network is the average of the clustering coefficient of all the nodes.

### CNs-based teleconnection analysis

Taking the monitoring stations as the network nodes and the correlations (of their temporal water quality dynamics) as the potential links, a virtual CN is thereby constructed. The hidden teleconnections and network patterns can be revealed by indexes of CNs, including degree centrality and clustering coefficient. The links, meanwhile, are depicted by the Pearson correlation coefficient (Benesty et al. 2009), which is calculated between each pair of stations. The identification of links between nodes is done by assuming different correlation

threshold (CT) values, which are arbitrarily chosen. In general, the threshold values can significantly influence the properties of CNs.

In this study, the difference of network features between several water quality indicators and between basins/regions are analysed. However, such teleconnections may be indirect, and result from similar drivers, such as pollution release, climate pattern, and hydrologic conditions in a basin. Investigations on the details of such mechanisms/drivers are beyond the scope of this study and, therefore, are not attempted here.

1.

## Results and discussion

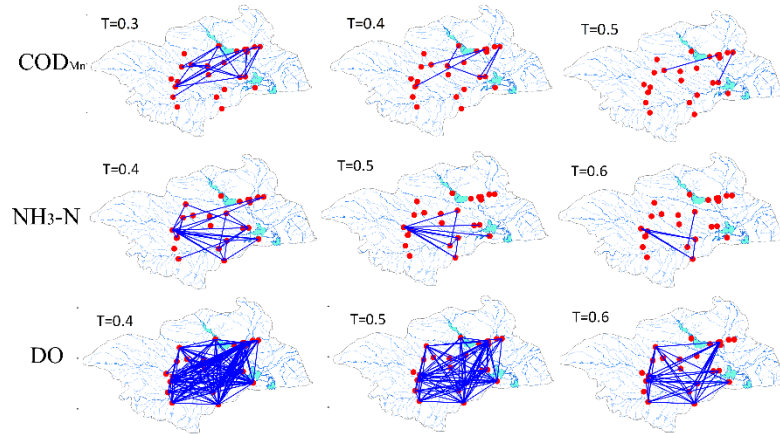
(a)

### Teleconnection pattern on Huaihe River basin

i.

#### Linear correlation analysis

**Fig. 3** presents the results from the linear correlation analysis. As expected, the correlation analysis between pairwise water quality monitoring stations shows the number of links of networks (i.e. the neighbours of nodes) decreases with an increase of CT value. Besides, for different stations, the distribution of the number of neighbours is uneven; some stations have many links to other stations, while some others do not. Specifically, the correlation for  $\text{COD}_{\text{MN}}$  is significantly weaker than that for  $\text{NH}_3\text{-N}$  and  $\text{DO}$ . Even when the CT is as low as 0.3, the teleconnection is still not obvious in the  $\text{COD}_{\text{MN}}$  network, let alone the higher CT values.

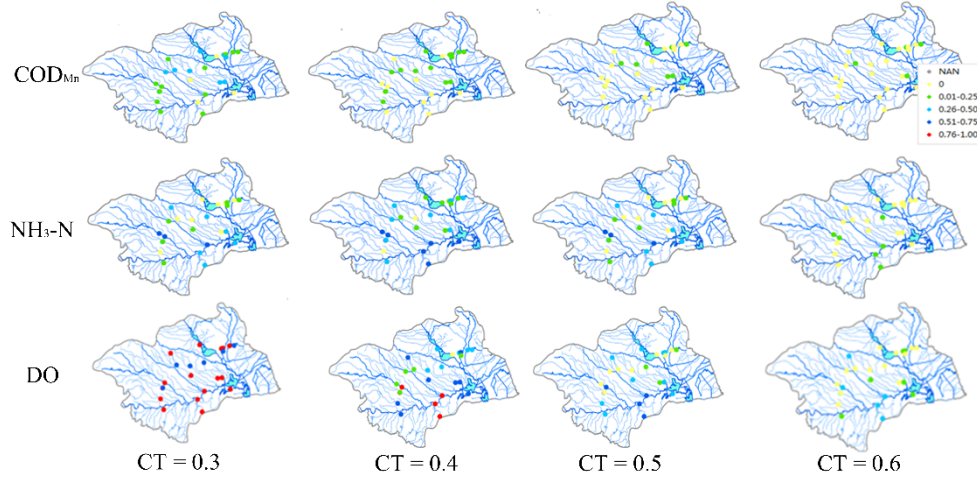




**Fig. 3.** Topological structure of water quality monitoring networks in the Huaihe River basin based on certain assumed CT values of linear regression for three typical water quality parameters.

#### Degree centrality analysis

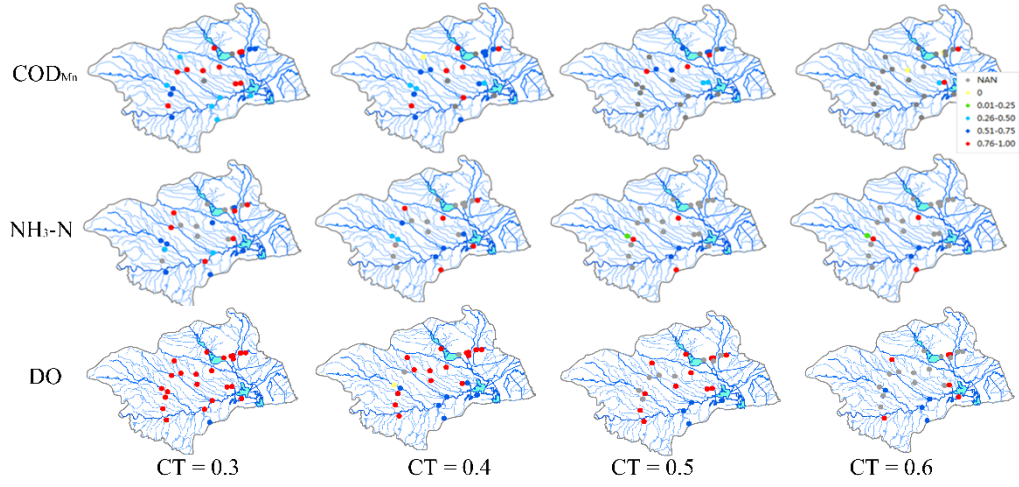
**Fig. 4** shows the degree centrality results for the three water quality networks. It is clear that for all the three water quality parameters, the degree centrality of nodes is not necessarily related to the spatial proximity; even the stations with close geographical location can have completely different degree centrality and vice versa. Though some stations with high degree centrality are near to the secondary tributaries, there is no evidence of a convincing distribution characteristic. Furthermore, the degree centrality of the DO networks is relatively higher than  $COD_{MN}$  and  $NH_3-H$  in line with the linear correlation results.



**Fig. 4.** Degree centrality pattern of water quality monitoring networks in the Huaihe River basin based on certain CT values of linear correlation for three typical water quality parameters.

#### Clustering coefficient analysis

Similar with the results of degree centrality analysis, clustering coefficients of stations help to understand the teleconnection of water quality dynamics as well. The spatial locations of stations are not considered to be relative with the clustering coefficient, as shown in **Fig. 5**. For the three water quality parameters, the average clustering coefficient decreases with an increase in the CT values, except for a few specific stations whose clustering coefficients slightly increase. Moreover, the highest clustering coefficient is still for the DO network, and the next is for  $NH_3-H$ .



**Fig. 5.** Clustering coefficient pattern of water quality monitoring networks in the Huaihe River basin based on certain CT values of linear correlation for three typical water quality parameters.

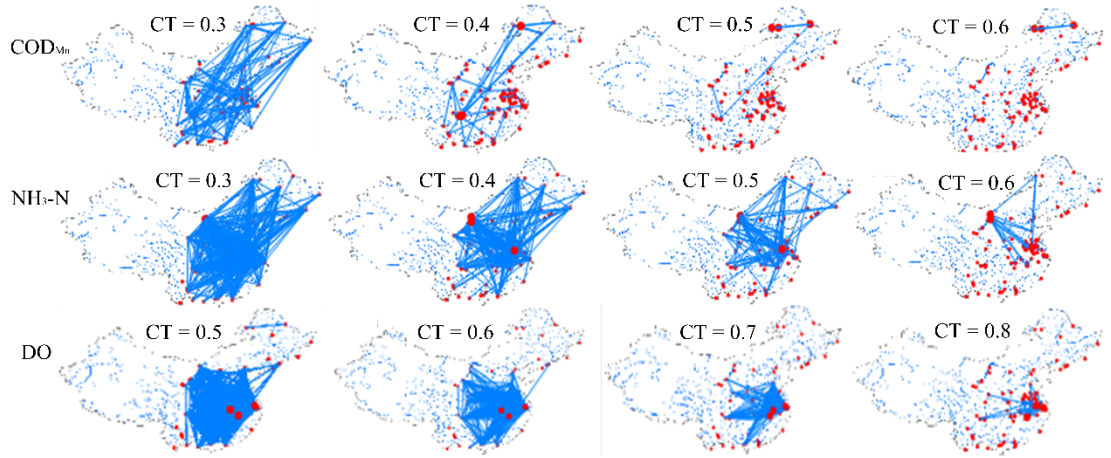
1.

## Teleconnection analysis on China's major rivers

(a)

### Linear correlation analysis

The same procedure, which was carried out for the Huaihe River data, is implemented to look into the teleconnection within the national scale water quality networks. The results vary greatly for different water quality indicators as shown in **Fig. 6**.

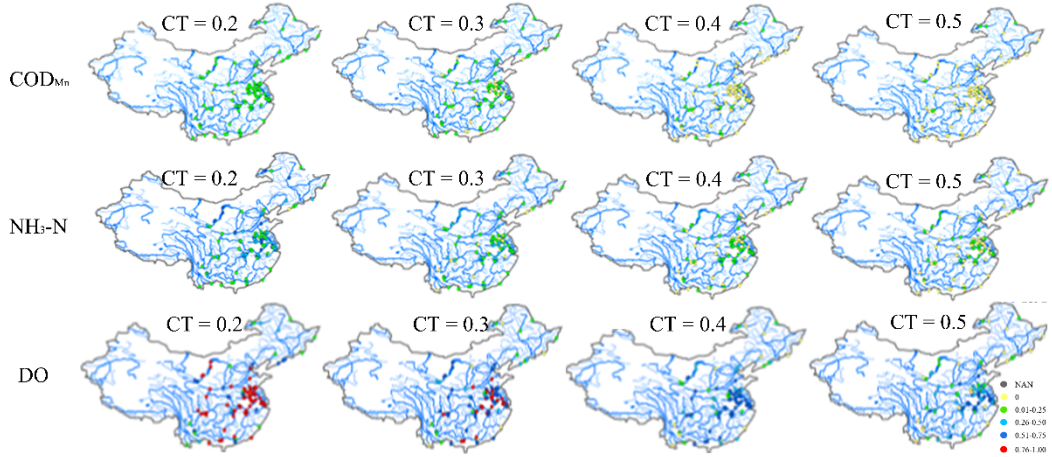


**Fig. 6.** Topological structure of water quality monitoring networks in China's major rivers based on certain CT values of linear correlation for three typical water quality parameters.

For  $\text{COD}_{\text{MN}}$ , there are only 30 pairs of stations (0.73% of total pairs) whose correlation is deemed to have a link with a  $\text{CT} = 0.4$ . Most of them are from the same or homologous basin, where the climatic, hydrologic, and  $\text{COD}_{\text{MN}}$  pollution status are close. For  $\text{NH}_3\text{-H}$ , the correlation is more apparent with 264 pairs of stations (6.45%) having values greater than 0.4 and 23 pairs (0.56%) having values greater than 0.6. It should be noted, however, that 10 out of the 23 pairs are trans-basin pairs. The high correlation values, therefore, can be attributed to their similar emission and degradation conditions. The correlation coefficients based on the DO data are most significant, averagely with 1184 pairs of stations (28.91%) having greater than values of 0.5 and 36 pairs (0.88%) having values greater than 0.8. Among the 36 pairs, 20 are located in the Taihu Lake basin and Poyang Lake basin, meaning that large lakes may play an important role in the DO dynamics of surface water system. Nevertheless, overall, the teleconnection in national scale water quality monitoring network is not obvious, and closer stations usually have stronger correlations.

#### Degree centrality analysis

Since simple networks with too few links, which usually occur for high threshold values, will undermine the validity of the CN analysis, CT value is as 0.2, 0.3, 0.4, and 0.5 to perform degree centrality analysis. Again, the average degree centrality for the DO monitoring data is greater than that for  $\text{COD}_{\text{MN}}$  and  $\text{NH}_3\text{-H}$ , as shown in **Fig. 7**. When  $\text{CT} = 0.2$ , most of the nodes in the  $\text{COD}_{\text{MN}}$  network is less than 0.5, while in the DO network, it is greater than 0.75. Even when CT is up to 0.5, the latter can be around 0.25~0.75 as degree centrality of the  $\text{COD}_{\text{MN}}$  network almost declines to 0. The  $\text{NH}_3\text{-H}$  network falls in between.

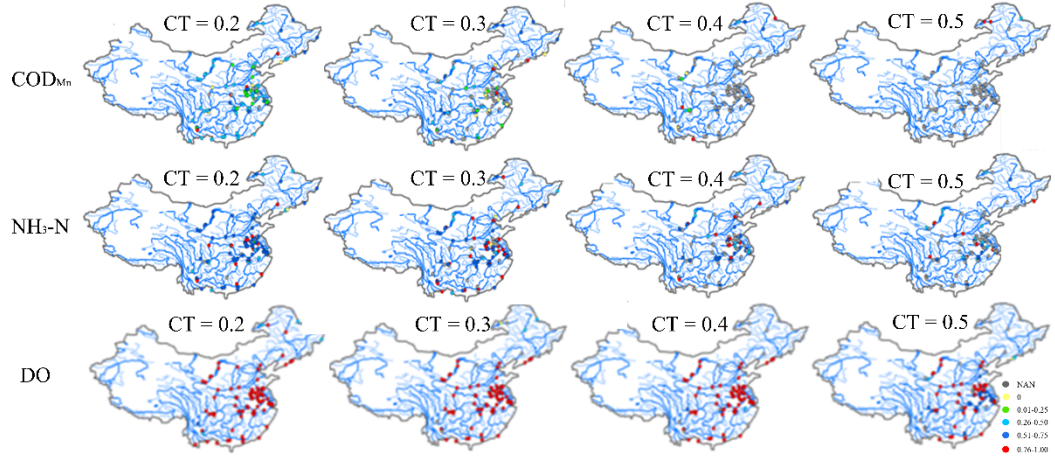


**Fig. 7.** Degree centrality pattern of water quality monitoring networks in

China's major rivers based on certain CT values of linear correlation for three typical water quality parameters.

### Clustering coefficient analysis

As can be seen from **Fig. 8**, the clustering coefficient pattern of the national scale water quality monitoring network is similar to that of the Huaihe River basin. The teleconnection of  $\text{COD}_{\text{MN}}$  is not apparent, and only three stations in the Songhua River basin have clustering coefficient values when  $\text{CT} = 0.5$ , because of the regional differences in upstream pollution emission. Meanwhile, 18 stations' clustering coefficients are greater than 0.5 for  $\text{NH}_3\text{-H}$  network when  $\text{CT} = 0.5$ . There are 2 stations each in the Yellow River basin and the Yangtze River basin, and 1 station each in the Songhua River basin and Huaihe River basin, which have a clustering coefficient greater than 0.75, and can be seen as key stations of  $\text{NH}_3\text{-H}$  control. For DO, approximately 90% of nodes have a clustering coefficient greater than 0.75 even when  $\text{CT} = 0.5$ .



**Fig. 8.** Clustering coefficient pattern of water quality monitoring networks in China's major rivers based on certain CT values of linear correlation for three typical water quality parameters.

### Comparison of teleconnection features between scales and parameters

Although the above results indicate that the average degree centrality and clustering coefficient values are much smaller than those for streamflow system of a watershed or across large regions (Sivakumar and Woldemeskel 2014; Halverson and Fleming 2015; Fang et al. 2017), some interesting observations are still worthy to be reported at this stage. The basin-scale analysis shows that the degree centrality and clustering coefficient rank for water quality indicator is  $\text{DO} > \text{NH}_3\text{-N} > \text{COD}_{\text{MN}}$ , and that for  $\text{COD}_{\text{MN}}$  is very low. This is reasonable,

since COD is more impacted by human activity (especially in urban areas) and more localized leading to lower correlation among those distant stations. As expected, it presents similar degree and clustering characteristics for some distant stations and quite a few differences for some nearby stations. However,  $\text{NH}_3\text{-N}$  is more relative to stormwater-induced non-point pollution, leading to stronger upstream-downstream relationship. The high teleconnection in DO network may be attributed to the overall connectivity of surface water systems and strong diffusion of DO. A similar distribution pattern of degree centrality and clustering coefficient is also found at the national scale for all parameters. With a detailed inspection on the range of Huaihe River in the national scale results (Fig.3 vs Fig.6),  $\text{NH}_3\text{-N}$  presents that the important nodes with higher degree centrality and clustering coefficient in Huaihe River basin still tend to play a vital role in national-scale network. But the pattern for DO presents difference between the two scales even in the same river. It reflects the different drivers of teleconnection among different parameters.

## Study limitations and future works

We conducted the CN methods to analyse the teleconnections in water quality monitoring networks. However, the number of stations considered in this study may be insufficient to provide strong interpretations and conclusions, from the perspective of CN. For instance, the small number of stations, yielding small number of nodes and links, cannot produce accurate CN indexes, especially in the context of teleconnection patterns. It should be noted, however, that the water environment is getting more attention from government and the public. Concepts such as Smart City requires a large on-line monitoring network for real-time management. The rapid development of water quality monitoring technology and instruments make the big data of water quality possible, leading to more widespread applications of the CN method.

Besides, there should be a time lag between the water quality dynamics at upstream and downstream stations. In this study, however, we ignored the drivers from upstream water quality process, which may hide the local correlation, and exaggerate the teleconnection crossing the basin/region.

The teleconnection analysis can certainly help us understand the intuitive rules of water quality dynamics. However, more fundamental and intrinsic laws need to be brought in to enlighten the dependence on attribution analysis, for further improving the reasonability of this work. The physical causes of the CN results, especially for their spatial difference should be more discussed. Usually, water quality is mainly controlled by the geographical conditions in a basin, including the climatic conditions, soil characteristics, geological conditions, and-use/land-cover and many others. How they directly influence the spatial pattern of CN and if these factors will be identified as important as well at national scale should be focused on. Furthermore, human activity has achieved a dominant position in environmental changes. Research on to what extent artificial factors (e.g. population density and point source pollution) affects the features of water

quality network and teleconnection by reshaping the geographical environment can be promising. Besides, some natural factors such as climate type, longitude and latitude may also influence the teleconnection of water quality dynamics in certain important ways on a larger scale. The insight into water quality process can enhance our ability to more properly manage the river systems.

## Conclusion

In this study, the concept of CN was used to look for potential spatial patterns in water quality dynamics at the basin and national scales. Weekly monitoring data of three major water quality indicators (i.e. DO,  $\text{NH}_3\text{-N}$ , and  $\text{COD}_{\text{Mn}}$ ), with a time span longer than 10 years, were used for the calculation of CN indexes, including degree centrality and clustering coefficient. The following key findings have emerged, which will help understand the teleconnection of water quality dynamics and to guide relevant works in the future.

1. Though the degree centrality and average clustering coefficient values at the national scale are relatively smaller than that at the basin scale, and much smaller than that for the runoff system, the existence of teleconnection of water quality dynamics is worthy of consideration.
2. The correlations in water quality dynamics are not related to the spatial location. Even close stations can have completely different water quality features, and distant stations may have strong correlations, i.e. teleconnection, of water quality.
3. In general, the degree centrality and clustering coefficient rank for water quality indicators is  $\text{DO} > \text{NH}_3\text{-N} > \text{COD}_{\text{MN}}$ , which is reasonable because  $\text{COD}_{\text{MN}}$  is more impacted by human activity in urban areas,  $\text{NH}_3\text{-N}$  is relative to non-point pollution, and DO has stronger diffusion trend.
4. The explosive growth of water quality monitoring data is conducive to explore the applications of CN methods further, among which attribution analysis may deepen our understanding of the teleconnection of water quality dynamics.

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 51979136), the Open Project of State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology (Grant No. ESK201901), and Science, Technology and Innovation Commission of Shenzhen Municipality (Grant No. JSGG20180508151852303). We thank the help of Baoyu Wang and Meng Yuan from Harbin Institute of Technology for the raw data collection and treatment, and anonymous reviewers for constructive suggestions.

## References

- Agarwal, A. (2019) Unraveling spatio-temporal climatic patterns via multi-scale complex networks, Universität Potsdam Potsdam.
- Agarwal, A., Caesar, L., Marwan, N., Maheswaran, R., Merz, B. and Kurths, J. (2019) Network-based identification and characterization of teleconnections on different scales. *Scientific Reports* 9(1), 1-12.
- Agarwal, A., Marwan, N., Maheswaran, R., Merz, B. and Kurths, J. (2018) Quantifying the roles of single stations within homogeneous regions using complex network analysis. *Journal of Hydrology* 563, 802-810.
- Ashok, K., Behera, S.K., Rao, S.A., Weng, H. and Yamagata, T. (2007) El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research: Oceans* 112(C11).
- Bascompte, J. (2009) Disentangling the web of life. *science* 325(5939), 416-419.
- Benesty, J., Chen, J., Huang, Y. and Cohen, I. (2009) Noise reduction in speech processing, pp. 1-4, Springer.
- Boers, N., Rheinwalt, A., Bookhagen, B., Barbosa, H.M., Marwan, N., Marengo, J. and Kurths, J. (2014) The South American rainfall dipole: A complex network analysis of extreme events. *Geophysical Research Letters* 41(20), 7397-7405.
- Erdős, P. and Rényi, A. (1960) On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci* 5(1), 17-60.
- Fang, K., Sivakumar, B. and Woldemeskel, F.M. (2017) Complex networks, community structure, and catchment classification in a large-scale river basin. *Journal of Hydrology* 545, 478-493.
- Ferster, B.S., Subrahmanyam, B. and Macdonald, A.M. (2018) Confirmation of ENSO-Southern Ocean teleconnections using satellite-derived SST. *Remote Sensing* 10(2), 331.
- Halverson, M. and Fleming, S. (2015) Complex network theory, streamflow, and hydrometric monitoring system design. *Hydrology and Earth System Sciences* 19(7).
- Huang, J., Wang, J. and Xia, C. (2020) Role of vaccine efficacy in the vaccination behavior under myopic update rule on complex networks. *Chaos, Solitons & Fractals* 130, 109425.
- Jha, S.K., Zhao, H., Woldemeskel, F.M. and Sivakumar, B. (2015) Network theory and spatial rainfall connections: An interpretation. *Journal of Hydrology* 527, 13-19.
- Jiang, J., Zheng, Y., Pang, T., Wang, B., Chachan, R. and Tian, Y. (2020) A comprehensive study on spectral analysis and anomaly detection of river water quality dynamics with high time resolution measurements. *Journal of Hydrology*, 125175.
- Kirchner, J.W. and Neal, C. (2013) Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proceedings of the National Academy of Sciences* 110(30), 12213-12218.
- Konapala, G. and Mishra, A. (2017) Review of complex networks application in hydroclimatic extremes with an implementation to characterize spatio-temporal drought propagation in continental USA. *Journal of Hydrology* 555, 600-620.
- König, D. (1936) *Theorie der endlichen und unendlichen Graphen: Kombinatorische Topologie der Streckenkomplexe*, Akademische Verlagsgesellschaft mbH.
- Krapivsky, P.L., Rodgers, G.J. and Redner, S., . (2001) Degree distributions of growing networks. *Physical Review Letters* 86(23), 5401-5404.
- Kurths, J., Agarwal, A., Shukla, R., Marwan, N., Rathinasamy, M., Caesar, L., Krishnan, R. and Merz, B. (2019) Unraveling the spatial diversity of Indian precipitation teleconnections via a non-linear multi-scale approach. *Nonlinear Processes in Geophysics* 26, 251-266.
- Lynn,

C.W. and Bassett, D.S. (2020) How humans learn and represent networks. *Proceedings of the National Academy of Sciences* 117(47), 29407-29415.

Lynn, C.W., Papadopoulos, L., Kahn, A.E. and Bassett, D.S. (2020) Human information processing in complex networks. *Nature Physics* 16(9), 965-973.

Ostrom, E. (2009) A general framework for analyzing sustainability of social-ecological systems. *science* 325(5939), 419-422.

Sivakumar, B. (2015) Networks: a generic theory for hydrology? *Stochastic environmental research and risk assessment* 29(3), 761-771.

Sivakumar, B. and Woldemeskel, F.M. (2014) Complex networks for streamflow dynamics. *Hydrology and Earth System Sciences* 18(11), 4565-4578.

Vespignani, A. (2009) Predicting the behavior of techno-social systems. *science* 325(5939), 425-428.

Wang, X., Wang, Z. and Shen, H. (2019) Dynamical analysis of a discrete-time SIS epidemic model on complex networks. *Applied Mathematics Letters* 94, 292-299.

Wasserman, S. and Faust, K. (1994) *Social network analysis: Methods and applications*, Cambridge university press.

Watts, D.J. and Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks. *nature* 393(6684), 440-442.

Zheng, Y. and Flanagan, S.V. (2017) The Case for Universal Screening of Private Well Water Quality in the U.S. and Testing Requirements to Achieve It: Evidence from Arsenic. *Environ Health Perspect* 125(8), 085002.