



# Identification of the Spatiotemporal Variability and Pollution Sources for Potential Pollutants of the Malian River Water in Northwest China Using the PCA-APCS-MLR Receptor Model

Dawei Mu<sup>1,2,3</sup> · Jianhua Wu<sup>1,2,3</sup> · Xiaofei Li<sup>1,2,3</sup> · Fei Xu<sup>1,2,3</sup> · Yongqiang Yang<sup>1,2,3</sup>

Received: 24 November 2022 / Revised: 10 January 2023 / Accepted: 13 January 2023 / Published online: 28 January 2023  
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

## Abstract

Identifying potential pollution sources of river water is a basis for effectively improving water pollution and sustainable water quality management. This study presents the analyses of the spatiotemporal variations of 12 water quality parameters from 6 water quality monitoring sections of the Malian River in northwest China. One-way analysis of variance (ANOVA) and cluster analysis (CA) were used to evaluate the changes in the spatiotemporal water quality over the monitoring period spanning from January 2017 to December 2021. The potential pollution sources and their contributions to river water pollution were qualitatively assessed via principal component analysis (PCA), factor analysis (FA), and the absolute principal component score-multiple linear regression (APCS-MLR) analysis. Despite the overall improvement of the water quality from 2017 to 2021, the upstream water quality is generally inferior to the downstream water quality, and is worse in the dry season compared to that in the other seasons. According to the APCS-MLR results, urban domestic-urban runoff and rural sewage contributed 24.08% and 20.10% to the river water pollution, respectively. Agricultural non-point sources, and livestock and poultry breeding, contributed 18.65% and 16.23%, to the river pollution, respectively. This study provides in-depth insights for identifying regional water pollution sources and serves as a reference for water environmental treatment in other similar watersheds in the Chinese Loess Plateau.

**Keywords** Water quality · Receptor model · Source identification · Water pollution · Malian River

## Introduction

Surface water is one of the most significant water resources for the sustainable development of the nature and human society (Bozorg-Haddad et al. 2021; Ebenstein 2010; Ryberg and Chanat 2022). Therefore, its water quality is crucial to maintain the health of the local residents and to ensure

socio-economic development (Ma et al. 2021; Madan and Ankit 2020). The water quality pollution and associated ecological environmental degradation are, however, becoming increasingly severe, particularly in China and India, due to rapid urbanization and agricultural modernization (Ali et al. 2016; Nong et al. 2020; Qin et al. 2020; Li and Wu 2019a; Subba Rao et al. 2022; Sarafraz et al. 2020; Varol 2019, 2020; Xu et al. 2022; Nsabimana and Li 2022; Yang et al. 2022). Deterioration of surface water quality is usually regarded a local or regional problem, but has become the focus of global attention in the context of water resource shortage and water environmental deterioration (Ali et al. 2022; Liu and Diamond 2005; Vörösmarty et al. 2010; Li and Wu 2019b; Downing et al. 2021; Subba Rao et al. 2019; Saha et al. 2020; UNESCO 2021; Velpandian et al. 2018). A scientific clarification of the ways and reasons responsible for the spatiotemporal changes in river water quality is critical for water environmental protection.

The Malian River is the main river in the Qingyang area of the Gansu Province, China, and is also the third biggest

✉ Jianhua Wu  
wujianhua@chd.edu.cn; wjh2005xy@126.com

<sup>1</sup> School of Water and Environment, Chang'an University, No. 126 Yanta Road, Xi'an 710054, Shaanxi, China

<sup>2</sup> Key Laboratory of Subsurface Hydrology and Ecological Effects in Arid Region of the Ministry of Education, Chang'an University, No. 126 Yanta Road, Xi'an 710054, Shaanxi, China

<sup>3</sup> Key Laboratory of Eco-Hydrology and Water Security in Arid and Semi-Arid Regions of the Ministry of Water Resources, Chang'an University, No. 126 Yanta Road, Xi'an 710054, Shaanxi, China

tributary of the Yellow River. It is essential for maintaining the high quality development of the region and even the Yellow River Basin (Li et al. 2022). Nevertheless, water shortage (Ringler et al. 2010), soil erosion (Du et al. 2021), and deterioration of water environment (Ma et al. 2012; Zhao et al. 2020), have plagued the region. The state and local governments have implemented a series of water environmental treatment, ecological protection, and restoration projects in the Malian River. Though progress has been made on these projects, the water quality of some monitored sections does not meet the national surface water quality requirements. Therefore, the spatiotemporal variation of the water quality must be urgently analyzed to identify relevant pollution sources.

Previous water quality studies on the Malian River focus on the hydrochemical characteristics of the Malian River and an associated water quality evaluation (Wang et al. 2018a; Su et al. 2009). Recent assessments of the seasonal water quality variation and the hydrochemistry also provide information on the overall changes in the Malian River (Wang et al. 2018b). Various anthropogenic or environmental factors with certain spatiotemporal variations may influence the water quality of the river during long-distance flow. However, to our knowledge, though information on the pollution sources of the Malian River is very significant for water resource management, they have not yet been identified. The water quality changes of the Malian River at different spatiotemporal scales are often statistically related to the climate, vegetation, hydrological processes, and socioeconomic conditions such as economic development level, industrial structure and population status (Ma et al. 2020; Qian et al. 2015; Tao et al. 2021; Wang et al. 2023; Yang et al. 2022). However, because of insufficient experimental data, comprehensive and systematic research on pollution sources and their contribution to water pollution is still lacking. Current water environmental analysis methods in the river basin fail to fully reflect the water quality status, which results in the incomprehensive use of a large amount of original data information. Therefore, comprehensive and reliable analyses methods should be used to analyze the data before doing correlation and contribution analyses.

In order to fully understand the water quality characteristics and pollution sources of the Malian River, this study used multivariate statistical methods to analyze the monthly water quality monitoring data at 6 monitoring sections during 2017–2021. These methods have been widely used for data analysis in various situations (Li et al. 2019, 2021; Meng et al. 2018; Nong et al. 2020; Subba Rao et al. 2018; Ren et al. 2021; Wu et al. 2014, 2020), indicating the suitability of them in water quality research. The objectives of this study were (1) to evaluate the spatiotemporal variation characteristics of the potential water quality parameters in the Malian River, (2) to identify the pollution sources by

principal component analysis, and (3) to quantify the contribution of the identified sources to the river water quality using the APCS-MLR model. The research can provide a scientific reference for the targeted protection and management of the ecological environment in the Malian River Basin.

## Materials and Methods

### Study Area

The study area focuses at the Malian River Basin, which belongs to the Gansu Province, Northwest China. It is located between the latitudes of 35°14' to 37°23' N and the longitudes of 106°40' to 108°35' E. The main stream of the Malian River is located in Qingyang City, Gansu Province (Fig. 1). It flows from north to south through the Qingyang City, and finally joins the Jing River. The main stream of the Malian River is 374.8 km in length, with a watershed area of 19,086 km<sup>2</sup>, and an average multi-year flow rate of 14.2 m<sup>3</sup>/s (Wang et al. 2018a, 2018b). The runoff of the Malian River is unevenly distributed throughout the hydrological year, with 60% of the annual runoff occurring between July and September. The river has an annual average sand content of 294 kg/m<sup>3</sup> and transports 1.3 × 10<sup>8</sup> tons of sand on average each year. This accounts for more than half of the average annual sand transported by the Jing River. The northern part of the study area belongs to the mid-temperate semi-arid zone, while the southern part belongs to the warm temperate semi-humid zone. The mean annual temperature of the Malian River ranges from 8 to 12 °C. The average annual rainfall ranges from 480 to 660 mm and decreases from southeast to northwest. The water surface evaporation rate ranges from 1380 to 1750 mm (Du et al. 2021). Qingyang City has a long history of oil resources exploitation, and the main towns and arable land are mostly located along the Malian River. The river has been severely polluted by point and non-point sources from industrial and mining enterprises, domestic waste, and agricultural activities.

### Data Collection

Monthly water samples from six water quality monitoring sections were set up and collected by the national and local governments from January 2017 to December 2021. As shown in Fig. 1, six monitoring sections were distributed in the main river channel: Chaijiatai (CJT), Wuliqiao (WLQ), Dianziping (DZP), Tielichuan (TLC), Xinbucun (XBC), and Ningxianqiaotou (NXQT). The detailed information of the monitoring sections is listed in Table 1. Twelve water quality parameters (pH, EC, DO, COD, BOD<sub>5</sub>, F<sup>-</sup>, TP, TN, NH<sub>3</sub>-N, *E. coli*, LAS,

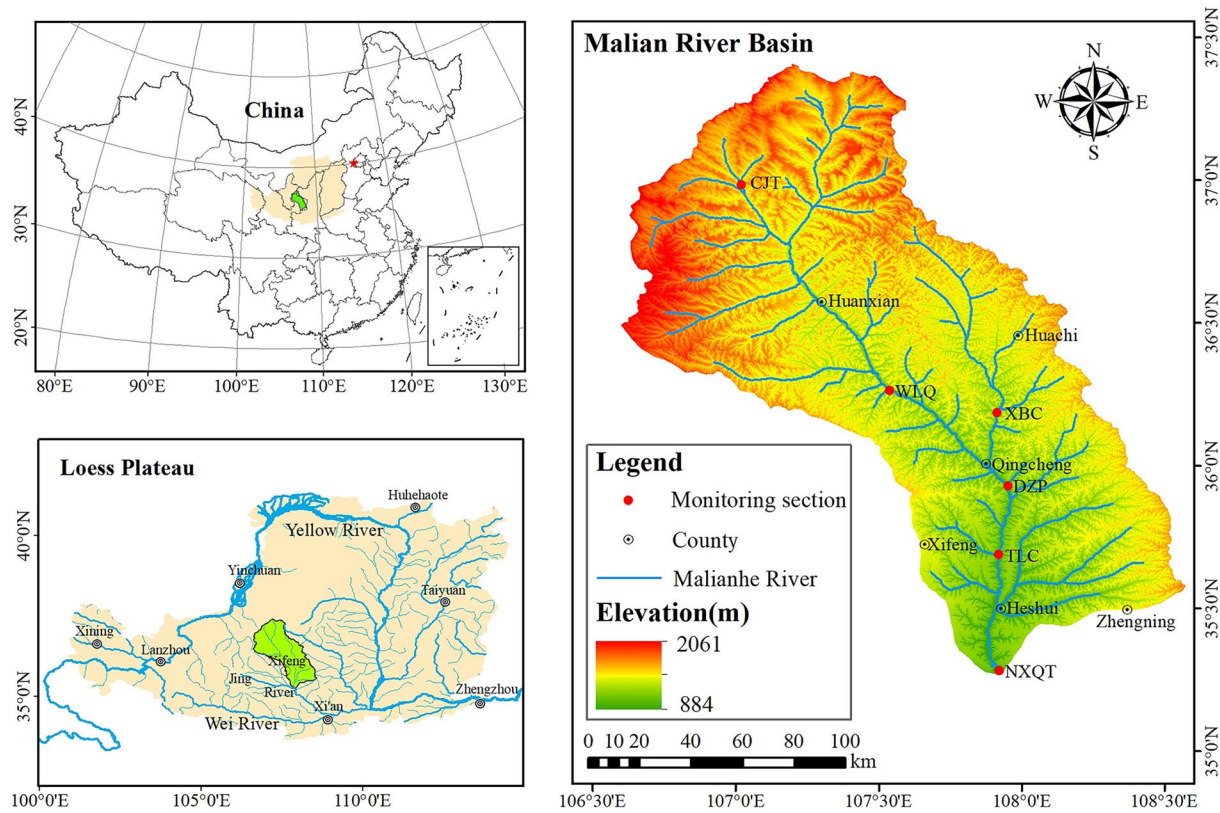


Fig. 1 Geographical locations of the study area and monitoring sections

**Table 1** Locations of water quality monitoring sections of the Malian River

Sections	Code	Latitude	Longitude	County
Chaijiatai	CJT	36°59'03.07"	107°01'08.77"	Huan County
Wuliqiao	WLQ	36°15'53.85"	107°32'02.51"	Huan County
Dianziping	DZP	35°55'25.62"	107°57'08.30"	Qingcheng County
Tielichuan	TLC	35°41'04.81"	107°55'17.03"	Heshui County
Ningxianqiaotou	NXQT	35°16'10.07"	107°55'48.41"	Ning County
Xinbucun	XBC	36°11'07.23"	107°54'46.46"	Huachi County

and  $Cr^{6+}$ ) were adopted in this study, among which pH, electrical conductivity (EC), dissolved oxygen (DO) were obtained on site using an portable water quality devices (OTT HydrolabDS5X multiparameter water quality monitor). The sampling procedures followed the national guidelines recommended by the Environmental Quality Standards for Surface Water (General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China and State Environmental Protection Administration of the People's Republic of China 2002) and the protocols outlined in the Guidance on Sampling Techniques (Ministry of Environmental Protection of the People's Republic of China 2009). Water samples were collected with pre-washed polyethylene bottles and

stored in refrigerators. After sampling, all bottles were immediately transferred to the laboratory, and analyzed within 24 h. The remaining physicochemical analyses of water quality indicators referenced the Standard Methods for the Examination of Water and Wastewater (State Environmental Protection Administration of the People's Republic of China 2002), and the detailed methods are as follows: chemical oxygen demand (COD), potassium dichromate method; five-day biochemical oxygen demand ( $BOD_5$ ), Winkler's method (titrimetric); fluoride ( $F^-$ ), Ion chromatograph; total nitrogen (TN) and total phosphorus (TP), potassium persulfate oxidation; ammonia nitrogen ( $NH_3-N$ ), Nesslerization method; fecal coliform (*E. coli*), Enzyme substrate method; linear alkylbenzene sulfonates (LAS), Methylene blue spectrophotometry; hexavalent

chromium (Cr<sup>6+</sup>), Diphenylcarbohydrazide spectrophotometric method. Additionally, blank samples and duplicates were used to control the data quality.

### Source Apportionment Using the APCS-MLR Receptor Modeling Technique

APCS-MLR was used to model the receptor for the water pollution source allocation (Cheng et al. 2020; Li et al. 2021; Meng et al. 2018; Zhang et al. 2022). The model assumes that all the possible pollution sources have a linear contribution to the final pollutant concentrations at the receptor site.

First, the data were analyzed using PCA, where the observations of a group of possible-related variables called principal components (PCs) were generated and extracted. To further simplify the PCA data structure, FA was performed on the basis of the PCA. This was obtained by orthogonal transformation of factor loading matrix using maximum variance method and building new variables. The rotated factor loading matrix and the eigenvalue are then obtained and used to calculate the principal factor eigenvector. The eigenvector and APCS that were obtained from standardized physical and chemical data were further subject to APCS-MLR.

In the first step of the APCS-MLR model, the principal components of the water quality indicators were extracted. This forms the foundation for identifying and quantifying pollution sources. The calculation formulae are as follows:

$$(A_s)_k = \sum_{j=1} w_j \cdot z_k \quad (1)$$

$$z_k = \frac{c_k - \bar{c}}{\sigma}, \quad (2)$$

where,  $(A_s)_k$  is the score of the principal component,  $w_j$  is the weight of the  $j$ -th principal component, and  $j$  is the principal component serial number in Eq. 1.  $z_k$  is the standardized value of the parameter concentration at the  $k$ -th monitoring value. In Eq. 2,  $c_k$  is the parameter concentration at the  $k$ -th monitoring value, and  $\bar{c}$  represents the standard deviation of the parameter concentration.

$(A_s)_k$  is a standardized value, and therefore, cannot be directly used to calculate the original contribution of the PCs. The standardized factor score must be transformed into a non-standard APCS following the Eqs. (3, 4, 5).

$$APCS_{ji} = (A_s)_{ji} - (A_0)_j \quad (3)$$

$$(A_0)_j = \sum_{i=1} S_{ij} \cdot (Z_0)_i \quad (4)$$

$$(Z_0)_i = \frac{0 - \bar{c}_i}{\sigma_i}, \quad (5)$$

where,  $(A_0)_j$  is the factor score under the value of 0,  $i$  is the code of chemical index,  $S_{ij}$  is the factor coefficient, and  $(Z_0)_i$  is the standardized value when the observation concentration is set to 0.

MLR takes measured value as the dependent variable, and the APCS is the independent variable of the budgeted pollution concentration. MLR can be expressed as follows:

$$C_i = \sum_m a_{mi} \cdot APCS_{mi} + b_i, \quad (6)$$

where,  $b_i$  is the value of unidentified source,  $C_i$  is the measured value, and  $a_{mi}$  represents the coefficients of the  $m$ -th component to the  $i$ -th parameter.

The formula for calculating the contribution rate of the  $m$ -th pollution source to the  $i$ -th pollution factor ( $PC_{mi}$ ) is as follows (Gholizadeh et al. 2016):

$$PC_{mi} = \frac{|a_{mi} \cdot \overline{APCS_{mi}}|}{|b_i| + \sum_m |a_{mi} \cdot \overline{APCS_{mi}}|} \quad (7)$$

The contribution of the unidentified sources is:

$$PC_{mi} = \frac{|b_i|}{|b_i| + \sum_m |a_{mi} \cdot \overline{APCS_{mi}}|} \quad (8)$$

### Data Processing

In this study, time and space are primarily the controlling variables, which belong to one-way ANOVA. ANOVA was used to verify whether different levels of time and space have any significant impact on the parameters (Varol 2019). The possible pollution sources in the river water and the pollutant characteristics were identified via PCA/FA (Kabir et al. 2020a; Tusher et al. 2020). PCA/FA parses the dataset and compresses the data dimensions to maximize the explanation of the original variables with fewer proxies. This is done by analyzing the relationships between multiple variables (Kabir et al. 2020b). This study also used cluster analysis (CA) to group the monitoring sections according to the similarity of water quality. Here, Ward's correlation approach and squared Euclidean distance were used as the similarity measures, and the clusters were displayed graphically using dendrograms (Varol 2019). This study used the following data processing methods: (1) missing data were estimated as the average value of the corresponding datasets, (2) the goodness of fit was tested using the Kolmogorove-Smirnov (K-S) statistics (Gholizadeh et al. 2016), (3) the applicability



of the results was verified by Kaiser–Meyer–Olkin (KMO) and Bartlett’s sphericity test (Wang et al. 2017), and (4) the percentage error (PE) and the root means square error (RMSE) were used to verify the degree of fitting for the established APCS-MLR model (Castrillo and García 2020). The statistical software package SPSS 25 and Microsoft Office Excel 2019 were used for all the data processing.

## Results and Discussion

### Interannual and Spatial Water Quality Characteristics

From 2017 to 2021, the annually minimum value of pH was 7.56 and the maximum value was 9.95 (Table 2). The annually average value ranged between 8.40 and 8.54, and this indicates that the water quality of Malian River was alkaline during the monitoring period. Irin et al. (2017) found that the river water is slightly alkaline in arid and rainless areas, which is similar to the current research results. The average pH values for the six monitoring sections were all higher than 8.3 (Fig. 2), and there was no noticeable spatial change in the pH from downstream to the upstream sections.

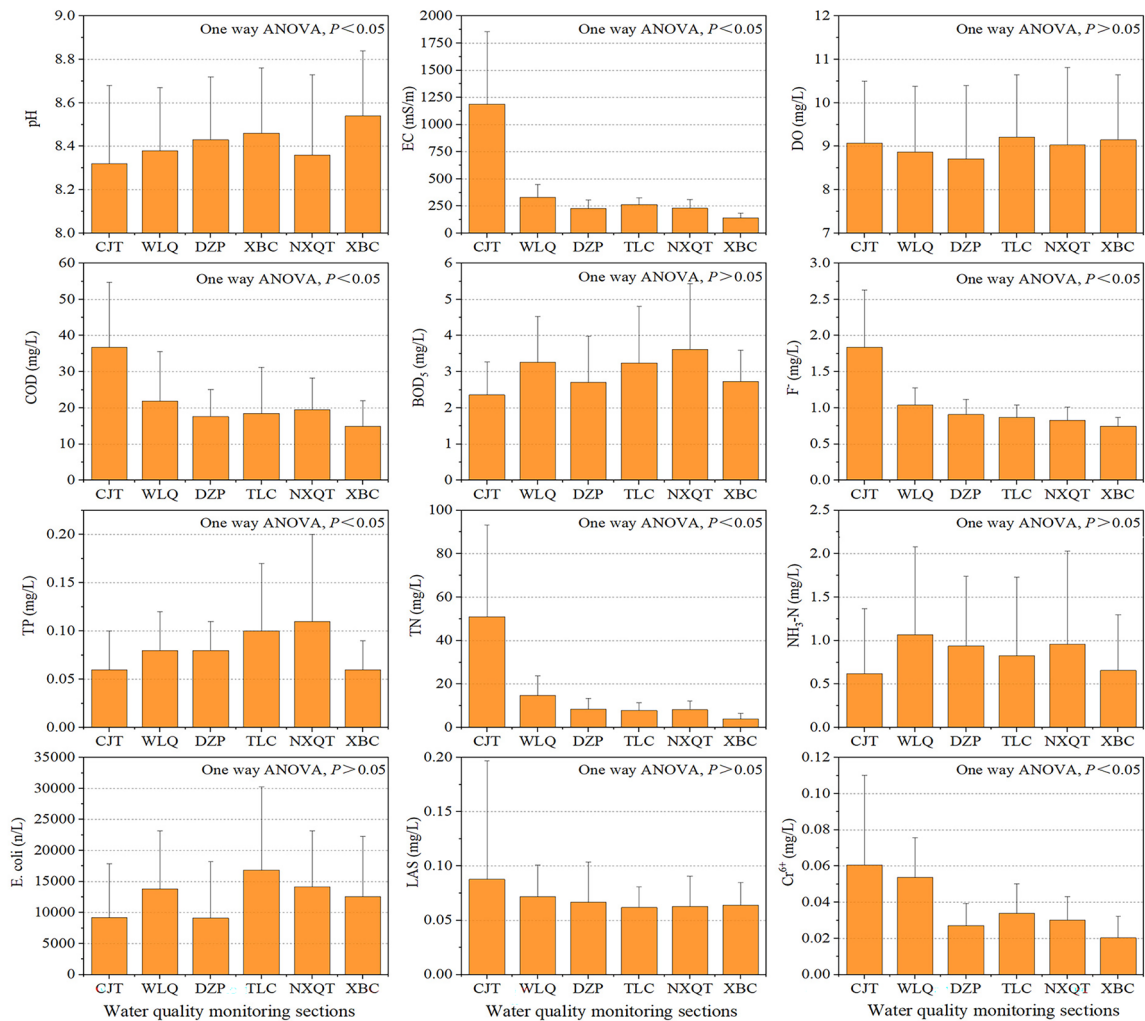
EC, a representation of the cation concentration, can significantly affect water quality. The EC is influenced by both natural weathering of sedimentary rocks and the artificial sources like industrial and sewage pollution (Martinez-Tavera et al. 2017). EC value fluctuated greatly during the year (Table 2). However, there was no interannual change, and the EC remained relatively stable in different years (Fig. 3). The northernmost section, namely CJT, had the highest average EC value. The spatial variation of EC showed a declining trend from the upstream to the downstream sections (Fig. 2). Increasing downstream water input and the associated dilution reduced the EC concentration of the river water in the downstream.

DO, a measure of the grade of the water quality ensures the survival of organisms in water (Rajendran et al. 2018). The annually average DO concentrations exceeded the standard of grade I (7.5 mg/L) from 2017 to 2021 (Fig. 3). In 2019 the DO concentration was at its lowest level (4.69 mg/L), while the DO concentrations at all the monitoring sections in other years were > 5.0 mg/L (Fig. 2). This indicates that DO level in the river water is generally within the expected control limit (Table 2). There was no spatial difference in the average DO concentration for each section. DO did not cause the water quality deterioration in the Malian River. The high DO concentrations and the spatial characteristics of this index can be explained by the fact that high-flow water can be replenished more adequate than other natural water bodies through long-distance open channels, such as lakes (Kangabam and Govindaraju 2017; Yang et al. 2019). The

**Table 2** Variations of the water quality parameters in the Malian River from 2017 to 2021 (Avg average, S.D standard deviation)

Parameters	Expected control limits <sup>a</sup>	2017			2018			2019			2020			2021		
		Avg.±S.D	Min	Max	Avg.±S.D	Min	Max	Avg.±S.D	Min	Max	Avg.±S.D	Min	Max	Avg.±S.D	Min	Max
pH	6–9	8.31±0.43	7.56	9.95	8.40±0.26	7.71	8.91	8.54±0.25	7.83	8.94	8.42±0.21	7.98	8.97	8.44±0.38	7.93	9.00
EC (mS/m)	N/A	378±510	55	2060	392±481	49	2170	312±283	71	1663	340±328	71	1867	396±416	103	2780
DO (mg/L)	≥5	9.39±1.53	6.50	14.3	9.06±1.57	6.55	13.40	8.53±1.46	4.69	12.16	9.07±1.37	6.54	11.62	8.91±1.76	5.14	13.16
COD (mg/L)	≤20	29.5±17.5	7.0	100.0	17.4±9.9	5.0	50.0	19.0±10.9	6.0	64.0	20.29±13.2	8.0	87.0	17.9±9.0	9.0	58.0
BOD <sub>5</sub> (mg/L)	≤4	3.06±0.58	1.92	4.80	2.98±1.28	1.00	8.10	3.59±1.80	1.20	10.90	2.83±1.546	1.00	8.60	2.63±1.34	1.00	9.00
F <sup>-</sup> (mg/L)	≤1.0	1.08±0.62	0.49	4.74	1.01±0.33	0.60	1.99	1.02±0.35	0.57	2.14	0.99±0.55	0.42	3.01	0.95±0.44	0.46	2.81
TP (mg/L)	≤0.2	0.106±0.065	0.010	0.340	0.081±0.093	0.020	0.610	0.072±0.034	0.020	0.170	0.077±0.031	0.020	0.140	0.070±0.029	0.030	0.160
TN (mg/L)	≤1.0	14.29±17.07	3.09	86.68	23.70±35.54	0.84	184.60	13.94±17.31	1.32	94.00	10.68±15.59	1.02	120.40	9.17±13.36	0.43	73.60
NH <sub>3</sub> -N (mg/L)	≤10,000	1.13±1.11	0.12	5.68	1.16±1.02	0.08	4.46	1.04±1.01	0.10	4.88	0.52±0.34	0.07	1.53	0.42±0.27	0.04	1.28
E.coli (n/L)	≤1,000	14,892±10,069	630	24,000	18,353±13,284	230	92,000	10,650±10,003	20	24,000	13,598±6930	1533	29,800	6024±5477	840	24,000
L-AS (mg/L)	≤0.2	0.103±0.088	0.050	0.580	0.064±0.022	0.010	0.200	0.052±0.006	0.050	0.080	0.071±0.028	0.035	0.197	0.049±0.011	0.020	0.070
C <sub>1</sub> 6+ (mg/L)	≤0.05	0.032±0.019	0.004	0.089	0.031±0.017	0.006	0.089	0.036±0.015	0.008	0.076	0.040±0.020	0.010	0.106	0.045±0.047	0.002	0.232

<sup>a</sup>Expected control limits are adopted from the Chinese Environmental Quality Standards for Surface Water. The expected control limits equal to the grade III limits prescribed by the Chinese Environmental Quality Standards for Surface Water



**Fig. 2** Spatial variations of the means and SD values for the potential water quality parameters at each water quality monitoring section in the Malian River

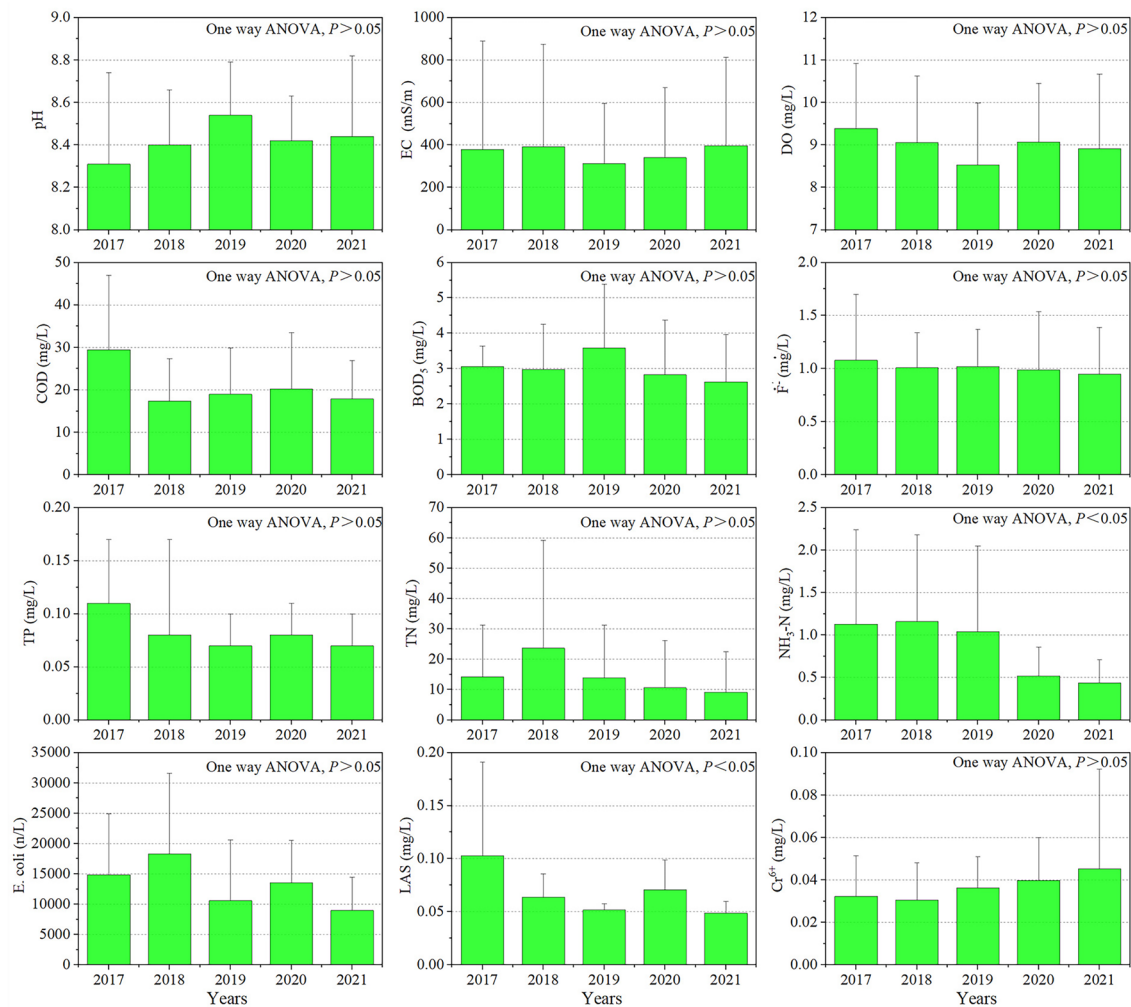
correlation between dissolved oxygen and other parameters needs further study since dissolved oxygen simulates the common physical and biological processes in water.

COD levels indicated that organic pollution and nutrient concentrations are present in the river (Nong et al. 2019). Overall, the annually average values of COD had decreased (Fig. 3). The highest detected COD concentration was detected in 2017 (29.48 mg/L). In the two northernmost monitoring sections, the mean COD concentrations exceeded the expected control limit (20 mg/L), while COD concentration in the other sections was below this threshold (Fig. 2). BOD<sub>5</sub> can also be used to describe the levels of pollution by organic matter (Lee et al. 2016). The average BOD<sub>5</sub> content ranged from 2.63 (2021) to 3.59 mg/L (2019) (Table 2). The annually mean BOD<sub>5</sub> values in the NXQT, the WLQ, and the TLC sections were greater than 3 mg/L (Fig. 2). The high BOD<sub>5</sub> concentration could be caused

by the discharge of a considerable amount of organic-rich domestic sewage and industrial wastewater (Lin et al. 2021).

There was a slight change in the annual value of soluble ion F<sup>-</sup> from 1.077 to 0.946 mg/L (Table 2). The average F<sup>-</sup> concentration of the monitoring section showed spatial difference. Here, the northernmost section had the highest F<sup>-</sup> concentration, while the remaining 5 sections showed weak or no variation (Fig. 2). These results might be due to the input of atmospheric pollutants with precipitation (Huang et al. 2017). Apart from the impact of the water source itself, some external inputs that may be closely related to the effects of human activities also affected the F<sup>-</sup> concentration (Ali et al. 2018; Kimambo et al. 2019; Li et al. 2017).

The annually average TP concentration reduced from 0.106 to 0.070 mg/L from 2017 to 2021 (Fig. 3). The maximum TP concentration in 2017 and 2018 exceeded the standard of the grade III (0.2 mg/L). The inter-annual



**Fig. 3** Temporal variations of the means and SD values for the 12 water quality parameters of the Malian River from 2017 to 2021

variation of the TP parameters was in line with the variation of the NH<sub>3</sub>-N concentration. This further confirmed that the efforts of the management department had improved the surface water status. Spatially, the TP concentration showed an upward trend, and fluctuations in the north and south sections (Fig. 2). The TP concentration of 6 sections was greater than 0.06 mg/L. This means that pollution sources in the area (such as human and livestock releases, and agricultural planting), have an impact on the surface state (Mao et al. 2019).

Apart from the CJT section (annually average TN concentration of 51.09 mg/L), which showed a statistically significant increase in TN concentrations compared to the others (with the annually average TN concentration fluctuations of about 10 mg/L), there were no conspicuous spatial changes (Fig. 2). The annually mean concentration of TN from 2018 to 2021 generally decreased year by year, but was still above 10 mg/L (Fig. 3). 30% of the animal waste

in China is returned to the farmland as fertilizer. Previous studies found that livestock and poultry manure is one of the main causes to nitrogen pollution (Li et al. 2016, 2020). However, inadequate governance and poor supervision have caused nitrogen loss in surface water and serious environmental pollution (Bai et al. 2017).

NH<sub>3</sub>-N characterizes the water pollution level of nutrients (Mao et al. 2019). The mean NH<sub>3</sub>-N concentration of each section fluctuated at the threshold of the grade III standard limit (1.0 mg/L), with no spatial differences (Fig. 2). The annually mean NH<sub>3</sub>-N concentration decreased significantly to 0.422 and 0.521 mg/L in 2021 and 2022, compared with the TN concentration from 2017 to 2019 (Table 2).

The maximum measured *E. coli* concentration was detected in 2018, with values as high as 92,000 n/L (Spring 2018). The annually mean concentrations of *E. coli* reduced from 2017 to 2021 (Fig. 3). The average concentrations of *E. coli* were high, especially in the TLC section where the

average concentration was 16,877 n/L (Fig. 2). These results illustrate the impact of pollution sources within the basin (e.g., human and livestock releases) (Tong et al. 2016).

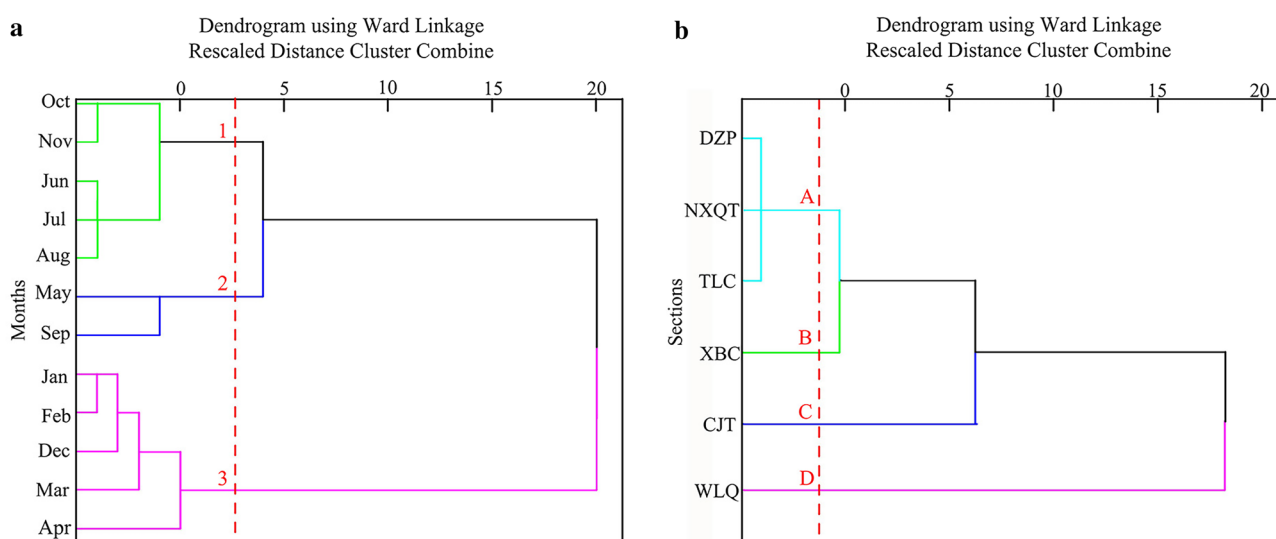
LAS is an active component of ordinary synthetic detergents. The LAS concentrations declined from 2017 to 2021 from 0.103 to 0.049 mg/L (Table 2). It has maintained a relatively stable level in space, and the maximum concentration of LAS (0.58 mg/L) was measured in 2017 (Fig. 2). LAS assists in the production of foam that adheres to the water surface. Because of these properties, it reduces dissolved oxygen in the water, which means the water quality and the survival of aquatic organisms may be adversely affected by this component (Ding et al. 2020).

For  $\text{Cr}^{6+}$ , there was no significant change in the annually average concentration, which ranged between 0.0306 and 0.0453 mg/L (Fig. 3). However, the mean concentration showed significant spatial differences, of which the CJT and WLQ sections in the north had much higher  $\text{Cr}^{6+}$  concentrations than the other sections and exceeded the standard of the grade IV prescribed in the Chinese standards for surface water (0.5 mg/L) (Fig. 2). The environmental background value of  $\text{Cr}^{6+}$  for the strata in the Huan County area is high, and the groundwater and the river are frequently transformed alternately (Cao 2003). The geology of the area causes a slow groundwater flow speed and strong dissolution and filtration. This causes relatively high groundwater trace element concentrations (Li et al. 2017). The high  $\text{Cr}^{6+}$  concentration in surface water may be due to its flow over the local geology and the interaction between surface water and groundwater.

## Seasonal and Sectional Water Quality Grouping

This study clustered the monthly monitoring data using cluster analysis (CA) by calculating the Euclidean square distance between the samples, and then using the Ward algorithm to generate the dendrogram. According to the CA results, the monitoring period can be divided into three periods (Fig. 4a): period 1 (January–May), period 2 (June–September), and period 3 (October–December). The grouped temporal characteristics of the river were to a large degree consistent with those of the four seasons (March–May for spring, June–August for summer, September–November for autumn, and December–February for winter). Even though it is autumn in September, the Malian River Basin is in the rainy season (Wang et al. 2018b), with large river runoff and water quality characteristics which are similar to those in summer. These characteristics are typical in period 2, revealing the delayed response of water quality after the rainy season. March and April are dry seasons, with little river runoff, but here the water quality characteristics are similar to those in winter where the water is polluted. The time clustering results for the Malian River show that obvious seasonal changes have been detected, and that the water quality depends on hydrological conditions, such as river runoff (Xu et al. 2019). The runoff in the rainy season dilutes the river pollutants. This is also the time when river levels and flow change, which significantly affects atmospheric reoxygenation and algal growth (Kong et al. 2021).

The three periods of the year showed obvious differences in the concentrations of  $\text{NH}_3\text{-N}$  and TN in the Malian River water. These water quality indexes in periods 2 and 3 were significantly lower than in period 1. This means that nitrogen-containing sewage discharged from industrial and



**Fig. 4** Dendrogram showing temporal (a) and spatial (b) similarities of monitoring periods or sections



domestic sources significantly affects the Malian River. The NH<sub>3</sub>-N and TN concentrations in period 2 were the lowest, indicating that the runoff in the rainy season diluted the nitrogen nutrients discharged from the point sources (Yuan et al. 2020). The accumulation of pollutants in the river channels and sewage pipelines will enter the river water during this period (Ding et al. 2020). These pollutants lead to deterioration of water quality, which also adequately explains the cause of COD and BOD<sub>5</sub> exceeding the standard in the rainy season.

The Ward algorithm was used to generate a dendrogram for the spatial CA, similar to the temporal CA. The water samples of the Malian River can be divided into 4 groups (Fig. 4b). Samples in group A belong to the CJT section and group B belongs to the XBC section, group C includes the DZP, TLC, and WLQ sections, and group D is the NXQT section. The sections in these groups had similar water conservancy conditions and were likely affected by similar sources, which caused the classifications to vary with significance level. The first group (CJT) is the initial section of the Malian River, where the COD, BOD<sub>5</sub>, TN, and Cr<sup>6+</sup> exceed the water quality standards. This causes the degraded natural conditions of the Malian River in this section. Group B is relatively far from anthropogenic influences. This group has a class III water quality grade, which is relatively good. Group C consists of three sections, which are each located in a large populous area (Huan County, Qingcheng County, Heshui County). This group shows high COD, NH<sub>3</sub>-N, F<sup>-</sup>, and LAS concentrations. Water pollution in the study area, including industrial wastewater, agricultural runoff, and urban sewage, affects this group. Group D (NXQT) is the section where the Malian River flows into the Jing River. The wastewater from the Xifeng County and Ning County domestic and industrial sewage treatment plants in this section, as well as rural domestic sewage and agricultural

species along the coast, causes a grade of IV–V water quality in this section.

Spatial clustering shows that water quality characteristics of the main stream of Malian River are spatially distributed. The corresponding pollution sources of the cluster are related to the land use, and changes of the clusters in different seasons are caused by hydrological changes in the basin (Xu et al. 2019). This is also consistent with the hydrological conditions of the Malian River and the distribution of coastal pollution sources. For example, industrial, domestic and urban sewage discharges are the main sources of pollution in the WLQ and TLC sections, which explains why the two sections are classified into the same cluster (Fig. 3b). The NXQT at the last section of the river belongs to the same cluster, which may be caused by diluted pollutants that concentrate in the tributary and that degrade downstream.

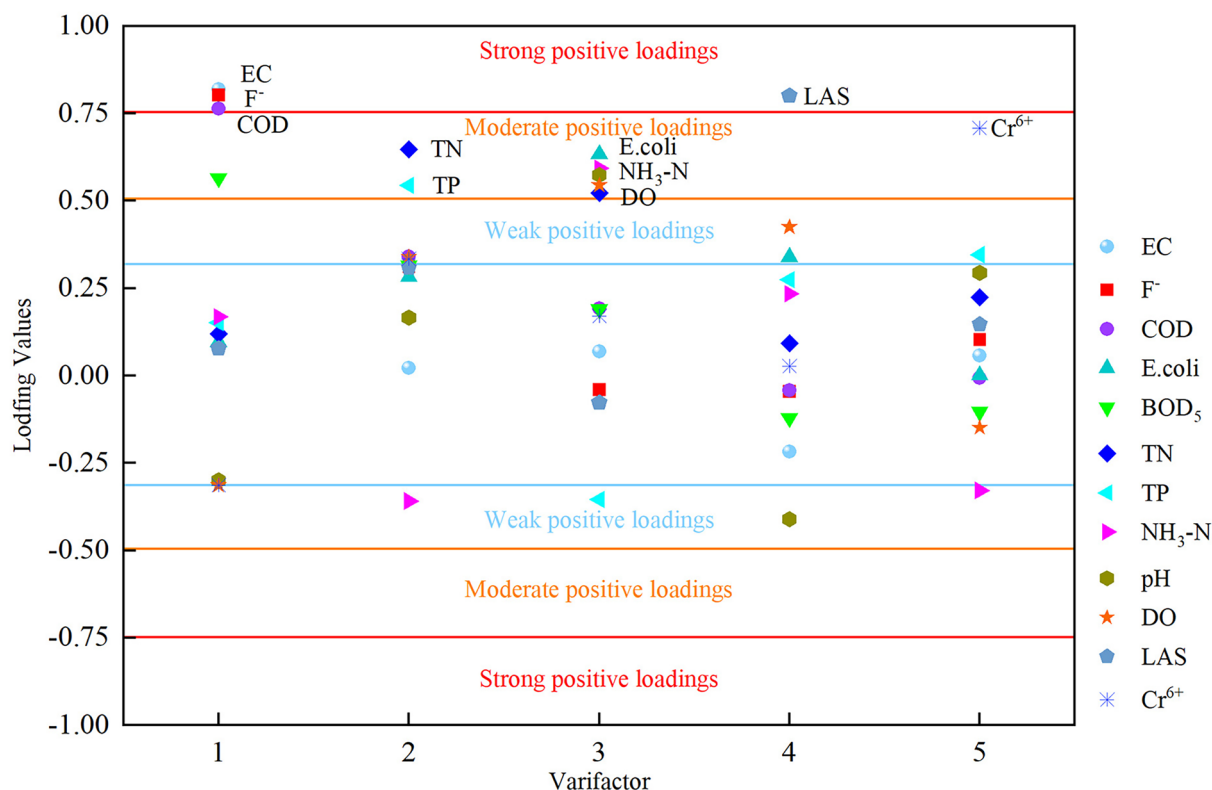
### Identification of Main Pollution Factors

The PCA/FA method was used to identify the natural and anthropogenic sources of the indicators with the largest contribution. The results aided in further understanding their distribution characteristics. To correlate the different parameters, KMO was conducted. The results (KMO=0.68,  $p < 0.001$ ) indicate a significant relationship between the different parameters and confirm the suitability of PCA/FA (Gholizadeh et al. 2016). Based on the Kaiser criteria (eigenvalue > 1), a total of 5 PCs were extracted, and these PCs explain 71% of the total variance (Table 3). The low total variance may be due to the inability to sample that was caused by the river ice in winter. Other relevant studies have also shown a similar low value of cumulative total variance (Liu et al. 2019; Huang et al. 2020).

The value of the first principal component (PC1) is 2.442, and it explains 22.2% of the variance. Figure 5 shows that the main load variables include COD, BOD<sub>5</sub>, EC, and F<sup>-</sup>. The

**Table 3** Variance explained by the main components for the Malian River

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
PC1	2.442	22.202	22.202	2.442	22.202	22.202	2.247	20.628	20.428
PC2	1.523	14.641	36.843	1.523	14.641	36.843	1.51	13.93	34.157
PC3	1.495	13.791	50.634	1.495	13.791	50.634	1.448	13.267	48.324
PC4	1.148	10.541	61.175	1.148	10.541	61.175	1.226	11.85	59.474
PC5	1.011	9.892	71.067	1.011	9.892	71.067	1.187	10.792	71.067
PC6	0.89	8.088	79.155	–	–	–	–	–	–
PC7	0.705	6.31	85.465	–	–	–	–	–	–
PC8	0.592	5.083	90.548	–	–	–	–	–	–
PC9	0.528	4.158	94.706	–	–	–	–	–	–
PC10	0.36	3.112	97.818	–	–	–	–	–	–
PC11	0.306	2.182	100	–	–	–	–	–	–



**Fig. 5** Component loadings for 12 parameters after varimax rotation

source of this pollution may be industrial sewage, sanitary sewage, aquaculture wastewater (Matiatos 2016; Varol 2020; Zhang et al. 2020). Responding to the increasing pressure for environmental protection since 2017, the local government has made great efforts to renovate the Malian River Basin. The renovation includes the construction and improvement of the urban sewage pipe network, the relocation of the industrial discharge outlets, and the relocation and closure of the livestock and poultry breeding industry within a 500 m range of both sides river. Other efforts include the restoration of the embankment vegetation, the treatment of the polluted water bodies in the urban areas, the collection, storage, and transportation of crop straws, the dredging of some rivers that have serious sediment deposition, and the removal of toilets near the rivers in rural areas. However, upgrading and renovating the sewage treatment plant and constructing the sewage pipelines, have taken a long time and are still to be fully completed. Rain is currently still mixed with sewage flow since large volumes of urban domestic sewage and urban runoff overflow into the Malian River area during flood events (Ding et al. 2020). Research show that urban sewage and urban runoff are important organic pollutants sources for pollutants such as COD (Lin et al. 2021). The average concentrations of COD in the roof and road runoff in Chinese urban sites are, for example, 125 and 284 mg/L, respectively (Zhang et al. 2020). These values exceed the

grade V standard limit of surface water by 2.1 and 6.1 times. Urban runoff is the second largest non-point source of pollution after agricultural source pollution. Therefore, PC1 can be defined as having an urban sewage-urban runoff source.

PC2 explains 14.64% of the total variance. TN and TP have a strong positive load (Fig. 5). This principal component has high nitrogen and phosphorus nutrient concentrations. High nutrient concentrations are a typical feature of chemical fertilizer. The sources of these nutrients are the non-point source pollution from orchards and farmlands. The proportion of land use type area affects the concentration of nitrogen and phosphorus in rivers (Yang et al. 2019). About 29% of the cultivated land in the Malian River Basin was treated with pesticides. Since fertilizers and residual pesticides can fairly easily enter the gully through leaching, infiltration, and soil erosion. PC2 is characterized as an agricultural non-point source.

PC3 accounts for 13.79% of the variance. The main load variables of PC3 include *E. coli*, NH<sub>3</sub>-N, and DO, at a moderate positive loading (0.593, 0.574, 0.545, respectively). *Escherichia coli* which grows in human and animal intestines represents the degree of fecal pollution in river. *Escherichia coli* can come either directly from the discharge of feces or from the release of pollutants that have accumulated in the river sediment. The temperature and nutrients in the sediment and the light in the deep water provide suitable

conditions for the survival of *E. coli* (Taoufik et al. 2017). By 2019, the Malian River Basin has been cleared from livestock and poultry farms, and slaughterhouses that were located within 500 m of the main river. However, pollutants that have accumulated from livestock and poultry breeding in the past have built up in the river sediment and contaminated the river. The scale of the centralized and decentralized livestock and poultry breeding industry in the basin has gradually expanded in recent years, with a total of 365 breeding enterprises and professional breeding cooperatives according to our survey. Many farms and slaughterhouses discharge wastewater without treatment. This wastewater leaches into the river during rainfall events or is directly discharged into rural rivers by simple pipeline devices which pollute the water source. Statistics on China's animal husbandry indicate that the output of livestock manure in China has increased in the past decade (Bao et al. 2019). As much as one-third of animal waste is used as fertilizer for farmland (Bai et al. 2017). Therefore, PC3 can be characterized as the source of livestock and poultry breeding.

PC4 explains 10.54% of the variance. The main load variable is LAS. LAS generally comes from three aspects: (1) the emulsifiers, spreading agents, and detergents in pesticides that are released during agricultural production; (2) the detergent, cosmetics, and other articles used by residents; and (3) the industrial wastewater discharged by industrial enterprises producing and applying surfactants (Katam et al. 2020). The water quality of the Malian River has greatly improved over recent years because industrial sewage outlets have moved outside. Each county in the basin has also built a sewage treatment plant. TN and TP are also not the primary factors of PC4. Domestic sewage in villages and towns accounts for more than one-fifth of the total domestic sewage discharges in China. Qingyang City predominantly practices agriculture production and is located on the Loess Plateau. It has many rural residents and high volumes of rural domestic sewage discharges. More than half of the total discharge of rural domestic sewage comes from washing wastewater which contains a large amount of LAS. The PC4 excludes industrial sources, urban living sources, and plantation sources, and is considered the main rural sewage source.

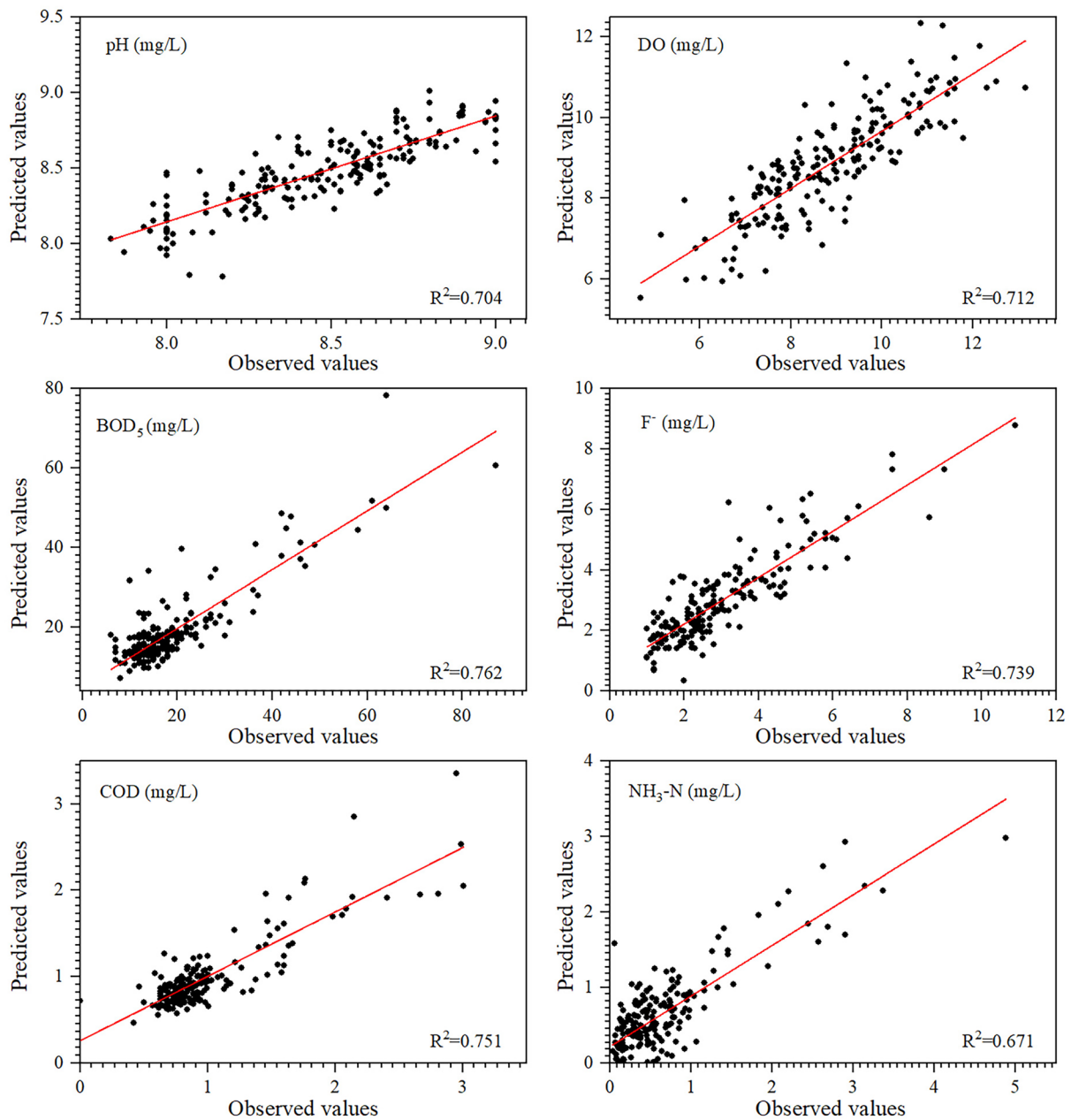
PC5 accounts for 9.89% of the variance, while the main load variable is  $\text{Cr}^{6+}$ . The high concentration of  $\text{Cr}^{6+}$  in this area is caused by the high environmental background value. Firstly, the geological mineral components in the study area form a geological environment rich in  $\text{Cr}^{6+}$ . The  $\text{Cr}^{6+}$  is weathered and dissolves into groundwater and surface water (Cao 2003). Secondly, the groundwater hydrological regime is slow, leading to the slow diffusion of  $\text{Cr}^{6+}$  in local areas. This increases the concentration of  $\text{Cr}^{6+}$  (Zhang et al. 2020). Therefore, PC5 can be defined as the geogenic source.

## Contribution of Pollution Sources

The multiple linear regression model was constructed and linearly fitted to the measured results based on the determined sources of the Malian River basin. Correlation coefficient greater than 0.5 means that the model has a high degree of prediction (Gholizadeh et al. 2016; Shen et al. 2021; Zhang et al. 2022). The  $R^2$  between the predicted concentration and the observed concentration of the water quality indices ranges from 0.671 to 0.762 (Fig. 6). This indicates that the prediction results of the APCS-MLR model in this study perform well.

Figure 7 shows the proportions of influence of different pollution sources to water chemical parameters. The urban sewage-urban runoff source (PC1), which has an average contribution of 23.67%, was the main source of COD (57.40%). It also contributed  $\text{BOD}_5$  at 36.47%. The contribution of PC1 to the EC, pH,  $\text{NH}_3\text{-N}$ , DO and LAS, ranges from 20 to 25%, while its contribution to other water quality indicators is relatively low. Agricultural non-point sources accounted for 18.65%, as indicated by the analysis of PC2. This is explained by the high proportions of TP (52.61%) and TN (41.84%). For PC3, the contributions of livestock and poultry breeding sources were 16.03%. For the 12 water quality indices, the contributions ranged from 0.50% ( $\text{Cr}^{6+}$ ) to 43.93% (*E. coli*). *Escherichia coli* (43.93%),  $\text{NH}_3\text{-N}$  (43.54%), and TN (23%) were largely affected by the pollution sources within PC3. For PC4, rural domestic sources accounted for 19.81% of all the pollution sources. These are the main sources of the LAS (48.11%). The contribution of PC3 to COD, *E. coli*, and  $\text{NH}_3\text{-N}$  is nearly 30%. This could be the case since rural domestic pollution has a certain pollution impact on the nutrient elements and the organic matter of the surface water quality. The geogenic source (PC5) accounted for 11.12% of all pollution and is associated with the unique environmental geological and geomorphic characteristics of the study area. The change in concentration of  $\text{Cr}^{6+}$  mainly reflects the impact on the surface water quality, where  $\text{Cr}^{6+}$  contributes 68.09%. In addition, PC5 contributes more than 15% to  $\text{F}^-$  and pH. The original geological environment of the study area may also possibly affect the water quality in terms of hardness, nutrients, and other indicators.

Compared with previous studies, the unknown contribution rate in this study is generally low at 5.39% (Cheng et al. 2020; Fu et al. 2020; Liu et al. 2020). This means that the selected water quality parameters for source identification are reasonable, and the distribution results are reliable and accurate. Based on the results of the contribution of the five common factors to the surface water quality, the urban sewage-urban runoff is still the main pollution source of the Malian River. The rural living sources and the agricultural non-point sources are secondary pollution sources.



**Fig. 6** Scatter plots of predicted and observed normalized concentrations of species using the APCS-MLR model

Livestock and poultry breeding accounts for significant river water pollution. The impact of geogenic sources on the water quality can also not be ignored.

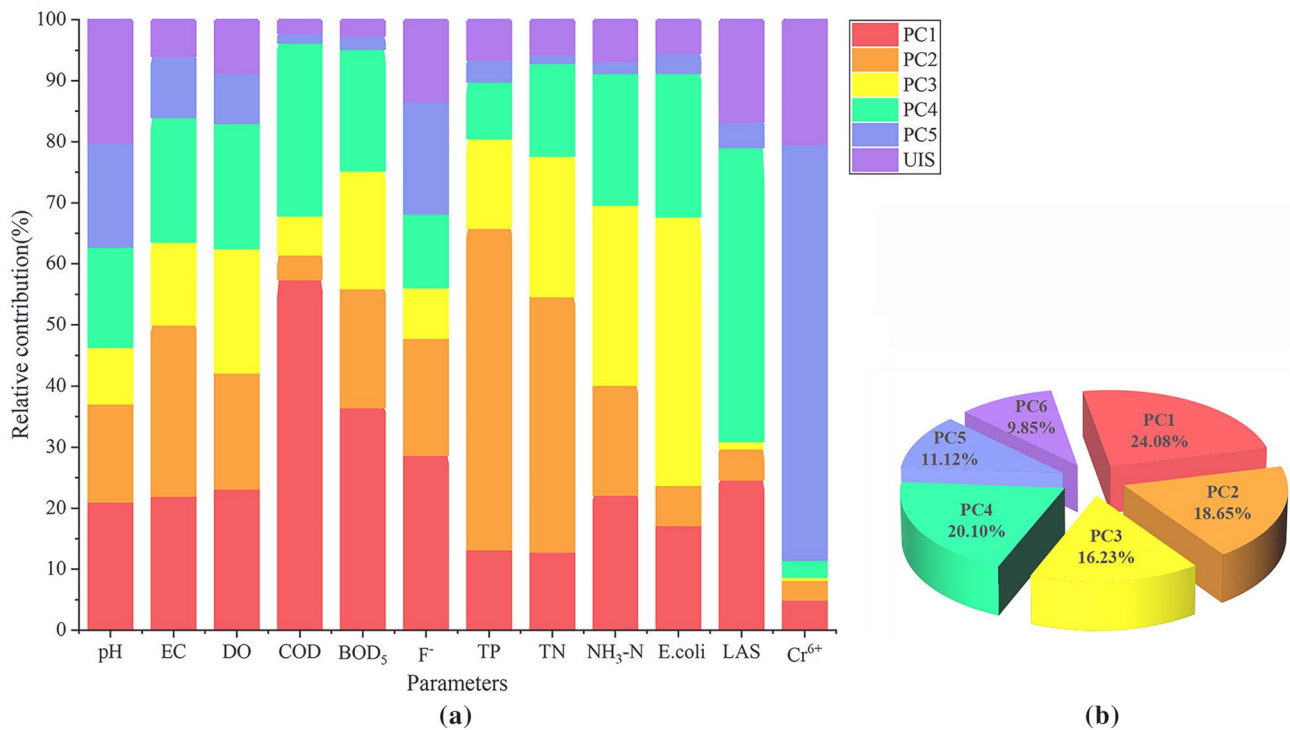
## Conclusions

This work analyzed the temporal and spatial changes of 12 parameters of the Malian River from January 2017 to December 2021. The pollution sources and their main contributions to groundwater pollution were analyzed by

running the APCS-MLR receptor model. A summary of the results follows:

- (1) The water quality of the Malian River shows an obvious spatiotemporal distribution. The EC, *E. coli*, and  $\text{Cr}^{6+}$  concentrations in the two northern sections (CJT and WLQ) are higher than in the other sections. This means that more frequent monitoring must be done in the two sections. The water environment has improved from 2017 to 2021, but is still dominated by excessive nutrient elements (TP and TN), and a high organic index





**Fig. 7** Contributions of different pollution sources on water quality variables (a) and the average contributions in the study area (b) using APCS-MLR model

(COD and BOD<sub>5</sub>). The pollution characteristics of the river are largely consistent with the four seasons. Water quality issues in the dry season should specifically be addressed.

- (2) Five potential pollution sources were determined based on the PCA/FA results. These include an urban sewage-urban runoff source, a rural sewage source, an agricultural non-point source, the livestock and poultry breeding, and a geogenic source.
- (3) The impacting ratios of the five pollution sources on surface water quality calculated by the APCS-MLR model are 24.08%, 20.10%, 18.65%, 16.23%, and 11.12%, respectively. The five sources have different contributions to COD, TP, TN, and NH<sub>3</sub>-N. The main sources of COD and NH<sub>3</sub>-N pollution in the river include urban and rural domestic sewage, while farmland chemical fertilizers are the main sources of TN and TP pollution. High levels of *E. coli* were also detected in this river. The *E. coli* pollution could be caused by feces, and further study on this aspect is needed. The local geological background causes Cr<sup>6+</sup> concentrations that exceed the water quality standard. Terminal treatment should be considered to reduce Cr<sup>6+</sup> pollution.

**Acknowledgements** The authors would like to thank Qingyang Municipal Ecological Environment Bureau of Gansu Province and The Second Geological Mineral Exploration Institute, Gansu Provincial Geology and Mineral Bureau for providing the climatic and water quality data. Finally, the authors thank the reviewers and the editors for their careful and valuable editing and suggestions, which greatly improved the quality of the manuscript.

**Author Contributions** DM: formal analysis, software, data curation, writing—original draft, writing—review & editing. JW: conceptualization, methodology, writing—original draft, writing—review & editing, supervision, funding acquisition. XL: conceptualization, writing—original draft, writing—review & editing, supervision. FX: methodology, formal analysis, writing—original draft, writing—review & editing. YY: resources, writing—review & editing.

**Funding** The research was supported by the National Natural Science Foundation of China (Grant Nos. 42272302 and 42090053), the Qinchuangyuan “Scientist + Engineer” Team Development Program of the Shaanxi Provincial Department of Science and Technology (Grant No. 2022KXJ-005), the Fok Ying Tong Education Foundation (Grant No. 161098), and the National Ten Thousand Talent Program (Grant No. W03070125).

**Data Availability** Data used in this research may be provided upon reasonable requests.

**Declarations**

**Conflict of interest** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- Ali S, Thakur S, Sarkar A, Shekhar S (2016) Worldwide contamination of water by fluoride. *Environ Chem Lett* 14:291–315. <https://doi.org/10.1007/s10311-016-0563-5>
- Ali S, Shekhar S, Bhattacharya P, Verma G, Chandrasekhar T, Chandrashekar A (2018) Elevated fluoride in groundwater of Siwani Block, Western Haryana, India: a potential concern for sustainable water supplies for drinking and irrigation. *Groundw Sustain Dev* 7:410–420. <https://doi.org/10.1016/j.gsd.2018.05.008>
- Ali S, Shekhar S, Chandrasekhar T (2022) Non-carcinogenic health risk assessment of fluoride in groundwater of the River Yamuna flood plain, Delhi, India. *Comput Earth Environ Sci*. <https://doi.org/10.1016/B978-0-323-89861-4.00046-4>
- Bai Z, Li X, Lu J, Wang X, Velthof GL, Chadwick D, Luo J, Ledgard S, Wu Z, Jin S, Oenema O, Ma L, Hu C (2017) Livestock housing and manure storage need to be improved in China. *Environ Sci Technol* 51(15):8212–8214. <https://doi.org/10.1021/acs.est.7b02672>
- Bao W, Yang Y, Fu T, Xie G (2019) Estimation of livestock excrement and its biogas production potential in China. *J Clean Prod* 229:1158–1166. <https://doi.org/10.1016/j.jclepro.2019.05.059>
- Bozorg-Haddad O, Delpasand M, Loáiciga H (2021) Water quality, hygiene, and health, Economical, Political, and Social Issues in Water. *Resources* 10:217–257. <https://doi.org/10.1016/B978-0-323-90567-1.00008-5>
- Cao T (2003) A study on source of Cr<sup>6+</sup> in Malian River water. *China Environ Monit* 19(5):43–46. <https://doi.org/10.19316/j.issn.1002-6002.2003.06.015> (in Chinese)
- Castrillo M, García A (2020) Estimation of high frequency nutrient concentrations from water quality surrogates using machine learning methods. *Water Res* 172:115490. <https://doi.org/10.1016/j.watres.2020.115490>
- Cheng G, Wang M, Chen Y, Gao W (2020) Source apportionment of water pollutants in the upstream of Yangtze River using APCS-MLR. *Environ Geochem Health* 42:3795–3810. <https://doi.org/10.1007/s10653-020-00641-z>
- Ding F, Liu Y, Wang L, Liu H, Ji C, Zhang L, Wu D (2020) Analysis of the palladium response relationship of a receiving water body under multiple scenario changes in rainfall-runoff pollution. *Environ Sci Pollut Res* 28:26684–26696. <https://doi.org/10.1007/s11356-021-12597-3>
- Downing J, Polasky S, Olmstead S, Newbold S (2021) Protecting local water quality has global benefits. *Nat Commun* 12:2709. <https://doi.org/10.1038/s41467-021-22836-3>
- Du M, Mu X, Zhao G, Gao P, Sun W (2021) Changes in runoff and sediment load and potential causes in the Malian River Basin on the Loess Plateau. *Sustainability* 13:443. <https://doi.org/10.3390/su13020443>
- Ebenstein A (2010) The consequences of industrialization: evidence from water pollution and digestive cancers in China. *Rev Econ Stat* 94(1):186–201. [https://doi.org/10.1162/REST\\_a\\_00150](https://doi.org/10.1162/REST_a_00150)
- Fu D, Wu X, Chen Y, Yi Z (2020) Spatial variation and source apportionment of surface water pollution in the Tuo River, China, using multivariate statistical techniques. *Environ Monit Assess* 192:745. <https://doi.org/10.1007/s10661-020-08706-3>
- General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, State Environmental Protection Administration of the People's Republic of China (2002) *Environmental Quality Standards for Surface Water* (GB 3838–2002). China Environmental Science Press, Beijing (in Chinese)
- Gholizadeh M, Melesse A, Reddi L (2016) Water quality assessment and apportionment of pollution sources using APCS-MLR and PMF receptor modeling techniques in three major rivers of South Florida. *Sci Total Environ* 565–567:1552–1567. <https://doi.org/10.1016/j.scitotenv.2016.06.046>
- Huang J, Gao J (2017) An improved Ensemble Kalman Filter for optimizing parameters in a coupled phosphorus model for low-land polders in Lake Taihu Basin, China. *Ecol Model* 357:14–22. <https://doi.org/10.1016/j.ecolmodel.2017.04.019>
- Huang G, Liu C, Li L, Zhang F, Chen Z (2020) Spatial distribution and origin of shallow groundwater iodide in a rapidly urbanized delta: a case study of the Pearl River Delta. *J Hydrol* 585:1260. <https://doi.org/10.1016/j.jhydrol.2020.124860>
- Irin A, Islam M, Kabir M, Hoq M (2017) Heavy metal contamination in water and fishes from the Shitalakhyariver at Narayanganj, Bangladesh. *Bangladesh J Zool* 44(2):267–273. <https://doi.org/10.3329/bjz.v44i2.32766>
- Kabir M, Tusher T, Hossain M, Islam M, Shanmi R, Kormoker T, Proshad R, Islam M (2020a) Evaluation of spatiotemporal variations in water quality and suitability of an ecologically critical urban river employing water quality index and multivariate statistical approaches: A study on Shitalakhya river, Bangladesh. *Hum Ecol Risk Assess* 27(5):1388–1415. <https://doi.org/10.1080/10807039.2020.1848415>
- Kabir M, Islam M, Hoq M, Tusher T, Islam M (2020b) Appraisal of heavy metal contamination in sediments of the Shitalakhya River in Bangladesh using pollution indices, geo-spatial, and multivariate statistical analysis. *Arab J Geosci* 13(21):1135. <https://doi.org/10.1007/s12517-020-06072-5>
- Kangabam R, Govindaraju M (2017) Anthropogenic activity-induced water quality degradation in the loktak lake, a Ramsar site in the Indo-Burma biodiversity hotspot. *Environ Technol* 40(17):2232–2241. <https://doi.org/10.1080/09593330.2017.1378267>
- Katam K, Shimizu T, Soda S, Bhattacharyya D (2020) Performance evaluation of two trickling filters removing LAS and caffeine from wastewater: Light reactor (algal-bacterial consortium) vs dark reactor (bacterial consortium). *Sci Total Environ* 707:135987. <https://doi.org/10.1016/j.scitotenv.2019.135987>
- Kimambo V, Bhattacharya P, Mtalo F, Mtamba J, Ahmad A (2019) Fluoride occurrence in groundwater systems at global scale and status of defluoridation—state of the art. *Groundw Sustain Dev* 9:100223. <https://doi.org/10.1016/j.gsd.2019.100223>
- Kong Z, Song Y, Shao Z, Chai H (2021) Biochar-pyrite bi-layer biore-tention system for dissolved nutrient treatment and by-product generation control under various stormwater conditions. *Water Res* 206:117737. <https://doi.org/10.1016/j.watres.2021.117737>
- Lee J, Lee S, Yu S, Rhew D (2016) Relationships between water quality parameters in rivers and lakes: BOD<sub>5</sub>, COD, NBOPs, and TOC. *Environ Monit Assess* 188(4):252. <https://doi.org/10.1007/s10661-016-5251-1>
- Li P, Wu J (2019a) Sustainable living with risks: meeting the challenges. *Hum Ecol Risk Assess* 25(1–2):1–10. <https://doi.org/10.1080/10807039.2019.1584030>
- Li P, Wu J (2019b) Drinking water quality and public health. *Expo Health* 11(2):73–79. <https://doi.org/10.1007/s12403-019-00299-8>
- Li F, Cheng S, Yu H, Yang D (2016) Waste from livestock and poultry breeding and its potential assessment of biogas energy in rural China. *J Clean Prod* 126:451–460. <https://doi.org/10.1016/j.jclepro.2016.02.104>
- Li P, Feng W, Xue C, Tian R, Wang S (2017) Spatiotemporal variability of contaminants in lake water and their risks to human health: a case study of the Shahu Lake tourist area, northwest China. *Expo Health* 9:213–225. <https://doi.org/10.1007/s12403-016-0237-3>
- Li P, Tian R, Liu R (2019) Solute geochemistry and multivariate analysis of water quality in the Guohua Phosphorite Mine, Guizhou Province, China. *Expo Health* 11(2):81–94. <https://doi.org/10.1007/s12403-018-0277-y>
- Li W, Lei Q, Yen H, Wollheim W, Zhai L, Hu W, Zhang L, Qiu W, Luo J, Wang H, Ren T, Liu H (2020) The overlooked role of

- diffuse household livestock production in nitrogen pollution at the watershed scale. *J Clean Prod* 272:122758. <https://doi.org/10.1016/j.jclepro.2020.122758>
- Li W, Wu J, Zhou C, Nsabimana A (2021) Groundwater pollution source identification and apportionment using PMF and PCA-APCS-MLR receptor models in Tongchuan City, China. *Arch Environ Contam Toxicol* 81(3):397–413. <https://doi.org/10.1007/s00244-021-00877-5>
- Li P, Wang D, Li W, Liu L (2022) Sustainable water resources development and management in large river basins: an introduction. *Environ Earth Sci* 81(6):179. <https://doi.org/10.1007/s12665-022-10298-9>
- Liu T, Yu H, Wang Q, Hu L, Yin J (2021) Surface water quality assessment based on the Integrated Water Quality Index in the Maozhou River basin, Guangdong, China. *Environ Earth Sci* 80:368. <https://doi.org/10.1007/s12665-021-09670-y>
- Liu J, Diamond J (2005) China's environment in a globalizing world. *Nature* 435:1179–1186. <https://doi.org/10.1038/4351179a>
- Liu L, Tang Z, Kong M, Chen X, Zhou C, Huang K, Wang Z (2019) Tracing the potential pollution sources of the coastal water in Hong Kong with statistical models combining APCS-MLR. *J Environ Manage* 245:143–150. <https://doi.org/10.1016/j.jenvman.2019.05.066>
- Liu L, Dong Y, Kong M, Zhou J, Zhao H, Tang Z, Zhang M, Wang Z (2020) Insights into the long-term pollution trends and sources contributions in Lake Taihu, China using multi-statistic analyses models. *Chemosphere* 242:125272. <https://doi.org/10.1016/j.chemosphere.2019.125272>
- Ma J, Pan F, He J, Chen L, Fu S, Jia B (2012) Petroleum pollution and evolution of water quality in the Malian River Basin of the Longdong Loess Plateau, Northwestern China. *Environ Earth Sci* 66:1769–1782. <https://doi.org/10.1007/s12665-011-1399-8>
- Ma T, Sun S, Fu G, Hall J, Ni Y, He L, Yi J, Zhao N, Du Y, Pei T, Cheng W, Song C, Fang C, Zhou C (2020) Pollution exacerbates China's water scarcity and its regional inequality. *Nat Commun* 11:650. <https://doi.org/10.1038/s41467-020-14532-5>
- Ma W, Meng L, Wei F, Opp C, Yang D (2021) Spatiotemporal variations of agricultural water footprint and socioeconomic matching evaluation from the perspective of ecological function zone. *Agric Water Manag* 249:106803. <https://doi.org/10.1016/j.agwat.2021.106803>
- Madan K, Ankit S (2020) Assessing groundwater quality for drinking water supply using hybrid fuzzy-GIS-based water quality index. *Water Res* 179:115867. <https://doi.org/10.1016/j.watres.2020.115867>
- Mao Q, Xia D, Wen Q, Shi J (2019) A Revised Method of surface water quality evaluation based on background values and its application to samples collected in Heilongjiang Province, China. *Water* 11:1057. <https://doi.org/10.3390/w11051057>
- Martinez-Tavera E, Rodriguez-Espinosa P, Shruti V, Sujitha S, Morales-Garcia S, MunozSevilla N (2017) Monitoring the seasonal dynamics of physicochemical parameters from Atoyac River basin (Puebla), Central Mexico: multivariate approach. *Environ Earth Sci* 76(2):95. <https://doi.org/10.1007/s12665-017-6406-2>
- Matiatos I (2016) Nitrate source identification in groundwater of multiple land-use areas by combining isotopes and multivariate statistical analysis: a case study of Asopos basin (Central Greece). *Sci Total Environ* 541:802–814. <https://doi.org/10.1016/j.scitotenv.2015.09.134>
- Meng L, Zuo R, Wang J, Yang J, Teng Y, Shi R (2018) Apportionment and evolution of pollution sources in a typical riverside groundwater resource area using PCA-APCS-MLR model. *J Contam Hydrol* 218:0169–7722. <https://doi.org/10.1016/j.jconhyd.2018.10.005>
- Ministry of Environmental Protection of the People's Republic of China (2009) Water Quality Guidance on Sampling Techniques (HJ 494–2009). China Environmental Science Press, Beijing (in Chinese)
- Nong X, Shao D, Xiao Y, Zhong H (2019) spatiotemporal Characterization Analysis and Water Quality Assessment of the South-to-North Water Diversion Project of China. *Int J Environ Res Public Health* 16(12):2227. <https://doi.org/10.3390/ijerph16122227>
- Nong X, Shao D, Zhong H, Liang J (2020) Evaluation of water quality in the South-to-North Water Diversion Project of China using the water quality index (WQI) method. *Water Res* 178:115781. <https://doi.org/10.1016/j.watres.2020.115781>
- Nsabimana A, Li P (2022) Hydrogeochemical characterization and appraisal of groundwater quality for industrial purpose using a novel industrial water quality index (IndWQI) in the Guanzhong Basin, China. *Geochemistry*. <https://doi.org/10.1016/j.chemer.2022.125922>
- Qian Y, Shao J, Zhang Z, Fei Y, Wang C, Cui X, Meng S (2015) Influence of rain on tetrachloroethylene's multiphase migration in soil. *J Earth Sci* 26(3):453–460. <https://doi.org/10.1007/s12583-015-0547-6>
- Qin G, Liu J, Xu S, Wang T (2020) Water quality assessment and pollution source apportionment in a highly regulated river of Northeast China. *Environ Monit Assess* 192:446. <https://doi.org/10.1007/s10661-020-08404-0>
- Rajendran V, Shrinithiviahshini N, Srinivasan B, Rengaraj C, Mariyaselvam S (2018) Quality assessment of pollution indicators in marine water at critical locations of the Gulf of Mannar Biosphere Reserve, Tuticorin. *Mar Pollut Bull* 126:236–240. <https://doi.org/10.1016/j.marpolbul.2017.10.091>
- Ren X, Li P, He X, Su F, Elumalai V (2021) Hydrogeochemical processes affecting groundwater chemistry in the central part of the Guanzhong Basin, China. *Arch Environ Contam Toxicol* 80(1):74–91. <https://doi.org/10.1007/s00244-020-00772-5>
- Ringler C, Cai X, Wang J, Ahmed A, Xue Y, Xu Z, Yang E, Jianshi Z, Zhu T, Cheng L, Yongfeng F, Xinfeng F, Xiaowei G, You L (2010) Yellow River basin: living with scarcity. *Water Int* 35:681–701. <https://doi.org/10.1080/02508060.2010.509857>
- Ryberg K, Chanut J (2022) Climate extremes as drivers of surface-water-quality trends in the United States. *Sci Total Environ* 809:152165. <https://doi.org/10.1016/j.scitotenv.2021.152165>
- Saha D, Shekhar S, Ali S, Elango L, Vittala S (2020) Recent scientific perspectives on the Indian hydrogeology. *Proc Indian National Sci Acad* 86(1):459–478. <https://doi.org/10.16943/ptinsa/2020/49790>
- Sarafraz M, Ali S, Sadani M, Heidarinejad Z, Bay A, Fakhri Y, Mousavi A (2020) A global systematic, review-meta analysis and ecological risk assessment of ciprofloxacin in river water. *Int J Environ Anal Chem* 102(17):5106–5121. <https://doi.org/10.1080/03067319.2020.1791330>
- Shen D, Huang S, Zhang Y, Zhou Y (2021) The source apportionment of N and P pollution in the surface waters of lowland urban area based on EEM-PARAFAC and PCA-APCS-MLR. *Environ Res* 197:111022. <https://doi.org/10.1016/j.envres.2021.111022>
- State Environmental Protection Administration of the People's Republic of China (2002) Standard Methods for the Examination of Water and Wastewater (Version 4). China Environmental Science Press, Beijing (in Chinese)
- Su X, Wan Y, Dong W, Hou G (2009) Hydraulic Relationship between Malianhe River and Groundwater: Hydrogeochemical and Isotopic Evidences. *J Jilin Univ (Earth Science Edition)* 39(6):1087–1094. [https://doi.org/10.13278/j.cnki.jjuese.2009.06.002\(inChinese\)](https://doi.org/10.13278/j.cnki.jjuese.2009.06.002(inChinese))
- Subba Rao N, Sunitha B, Rambabu R, Nageswara Rao PV, Surya Rao P, Spandana B, Sravanthi M, Marghade D (2018) Quality and degree of pollution of groundwater, using PIG from a rural part of Telangana State, India. *Appl Water Sci* 8(8):1–13. <https://doi.org/10.1007/s13201-018-0864-x>



- Subba Rao N, Srihari C, Spandana B, Sravanthi M, Kamalesh T, Jayadeep V (2019) Comprehensive understanding of groundwater quality and hydrogeochemistry for the sustainable development of suburban area of Visakhapatnam, Andhra Pradesh, India. *Hum Ecol Risk Assess* 25(1–2):52–80. <https://doi.org/10.1080/10807039.2019.1571403>
- Subba Rao N, Das R, Gugulothu S (2022) Understanding the factors contributing to groundwater salinity in the coastal region of Andhra Pradesh, India. *J Contam Hydrol* 250:104053. <https://doi.org/10.1016/j.jconhyd.2022.104053>
- Tao H, Song K, Liu G, Wen Z, Wang Q, Du Y, Lyu L, Du J, Shang Y (2021) Songhua River basin's improving water quality since 2005 based on Landsat observation of water clarity. *Environ Res* 199:111299. <https://doi.org/10.1016/j.envres.2021.111299>
- Taoufik G, Khouni I, Ghrabi A (2017) Assessment of physico-chemical and microbiological surface water quality using multivariate statistical techniques: a case study of the Wadi El-Bey River, Tunisia. *Arab J Geosci* 10:181. <https://doi.org/10.1007/s12517-017-2898-z>
- Tong Y, Yao R, He W, Zhou F, Chen C, Liu X, Lu Y, Zhang W, Wang X, Lin Y, Zhou M (2016) Impacts of sanitation upgrading to the decrease of fecal coliforms entering into the environment in China. *Environ Res* 149:57–65. <https://doi.org/10.1016/j.envres.2016.05.009>
- Tusher T, Sarker M, Nasrin S, Kormoker T, Proshad R, Islam M, Mamun S, Tareq A (2020) Contamination of toxic metals and polycyclic aromatic hydrocarbons (PAHs) in rooftop vegetables and human health risks in Bangladesh. *Toxin Rev* 40(4):736–751. <https://doi.org/10.1080/15569543.2020.1767650>
- UNESCO (2021) The global water quality challenge & SDGs. Available at <https://en.unesco.org/waterquality-iiwq/wq-challenge>. Accessed 31 July 2022
- Varol M (2019) Arsenic and trace metals in a large reservoir: seasonal and spatial variations, source identification and risk assessment for both residential and recreational users. *Chemosphere* 228:1–8. <https://doi.org/10.1016/j.chemosphere.2019.04.126>
- Varol M (2020) Use of water quality index and multivariate statistical methods for the evaluation of water quality of a stream affected by multiple stressors: a case study. *Environ Pollut* 266:115417. <https://doi.org/10.1016/j.envpol.2020.115417>
- Velupandian T, Halder N, Nath M, Das U, Moksha L, Gowtham L, Batta S (2018) Un-segregated waste disposal: an alarming threat of antimicrobials in surface and ground water sources in Delhi. *Environ Sci Pollut Res* 25:29518–29528. <https://doi.org/10.1007/s11356-018-2927-9>
- Vörösmarty C, McIntyre P, Gessner M, Dudgeon D, Prusevich A, Green P, Glidden S, Bunn S, Sullivan C, Liermann C, Davies P (2010) Global threats to human water security and river biodiversity. *Nature* 467:555–561. <https://doi.org/10.1038/nature09440>
- Wang J, Liu G, Liu H, Lam P (2017) Multivariate statistical evaluation of dissolved trace elements and a water quality assessment in the middle reaches of Huaihe River, Anhui, China. *Sci Total Environ* 583:421–431. <https://doi.org/10.1016/j.scitotenv.2017.01.088>
- Wang Y, Cheng X, Zhang M, Qi X (2018a) Hydrochemical characteristics and formation mechanisms of Malian River in Yellow River basin during dry season. *Environ Chem* 37(1):164–172. <https://doi.org/10.7524/j.issn.0254-6108.2017052602> (in Chinese)
- Wang Y, Han S, Deng Q, Qi X (2018b) Seasonal variations in river water chemical weathering and its influence factors in the Malian River Basin. *Environ Sci* 39(9):4132–4141. [https://doi.org/10.13227/j.hjck.201801096\(inChinese\)](https://doi.org/10.13227/j.hjck.201801096(inChinese))
- Wang D, Li P, He X, He S (2023) Exploring the response of shallow groundwater to precipitation in the northern piedmont of the Qinling Mountains, China. *Urban Clim* 47:101379. <https://doi.org/10.1016/j.uclim.2022.101379>
- Wu J, Li P, Qian H, Duan Z, Zhang X (2014) Using correlation and multivariate statistical analysis to identify hydrogeochemical processes affecting the major ion chemistry of waters: Case study in Laoheba phosphorite mine in Sichuan, China. *Arab J Geosci* 7(10):3973–3982. <https://doi.org/10.1007/s12517-013-1057-4>
- Wu J, Li P, Wang D, Ren X, Wei M (2020) Statistical and multivariate statistical techniques to trace the sources and affecting factors of groundwater pollution in a rapidly growing city on the Chinese Loess Plateau. *Hum Ecol Risk Assess* 26(6):1603–1621. <https://doi.org/10.1080/10807039.2019.1594156>
- Xu G, Li P, Lu K, Tantai Z, Zhang J, Ren Z, Wang X, Yu K, Shi P, Cheng Y (2019) Seasonal changes in water quality and its main influencing factors in the Dan River basin. *CATENA* 173:131–140. <https://doi.org/10.1016/j.catena.2018.10.014>
- Xu F, Li P, Chen W, He S, Li F, Mu D, Elumalai V (2022) Impacts of land use/land cover patterns on groundwater quality in the Guanzhong Basin of northwest China. *Geocarto Int*. <https://doi.org/10.1080/10106049.2022.2115153>
- Yang J, Wang F, Lv J, Liu Q, Nan F, Xie SL, Feng J (2019) Responses of freshwater algal cell density to hydrochemical variables in an urban aquatic ecosystem, northern China. *Environ Monit Assess* 191(1):29. <https://doi.org/10.1007/s10661-018-7177-2>
- Yang Y, Li P, Elumalai V, Ning J, Xu F, Mu D (2022) Groundwater quality assessment using EWQI with updated water quality classification criteria: a case study in and around Zhouzhi County, Guanzhong Basin (China). *Expo Health*. <https://doi.org/10.1007/s12403-022-00526-9>
- Yuan H, Yin H, Yang Z, Yu J, Liu E, Li Q, Tai Z, Ca Y (2020) Diffusion kinetic process of heavy metals in lacustrine sediment assessed under different redox conditions by DGT and DIFS model. *Sci Total Environ* 741:140418. <https://doi.org/10.1016/j.scitotenv.2020.140418>
- Zhang H, Li H, Yu H, Cheng S (2020) Water quality assessment and pollution source apportionment using multi-statistic and APCS-MLR modeling techniques in Min River Basin, China. *Environ Sci Pollut Res* 27:41987–42000. <https://doi.org/10.1007/s11356-020-10219-y>
- Zhang H, Li H, Gao D, Yu H (2022) Source identification of surface water pollution using multivariate statistics combined with physicochemical and socioeconomic parameters. *Sci Total Environ* 806(3):151274. <https://doi.org/10.1016/j.scitotenv.2021.151274>
- Zhao M, Wang S, Chen Y, Wu J, Xue L, Fan T (2020) Pollution status of the Yellow River tributaries in middle and lower reaches. *Sci Total Environ* 722:137861. <https://doi.org/10.1016/j.scitotenv.2020.137861>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.