Spatial and temporal characterization of water quality in Bosten Lake, China based on comprehensive water quality index (Postprint)

Authors: GUO Mengjing, BAI Zichen, Bo Yuan, WANG Wen, ZHANG Tiegang, XIANG Ke, ZHANG Jiao, ZHAO Huiyizhe, GUO Mengjing

Date: 2025-09-22T11:52:37+00:00

Abstract

Water quality is a pressing issue affecting the sustainable development of lakes. To elucidate the spatial and temporal characteristics of water quality in Bosten Lake, China, this study constructed a comprehensive water quality index (CWQI) based on key water quality indicators, utilizing water quality data collected from 17 sampling sites spaning from 2011 to 2019. Key water quality indicators were determined using factor analysis, and the spatial and temporal characteristics of key water quality indicators and the CWQI were examined using multivariate statistical analysis. The key water quality indicators included pH, chemical oxygen demand (COD), water transparency (SD), NO3-, total dissolved solids (TDS), Cl-, SO42-, and electrical conductivity (EC). Furthermore, the contribution rates of all water quality indicators to the water quality were quantitatively elucidated using the SHapley Additive exPlanations (SHAP) values, thereby validating the factor analysis outcomes. Among the eight key water quality indicators, the COD had the most significant influence on the water quality of Bosten Lake. The water quality condition of Bosten Lake has remained at Class III from 2011 to 2019 (CWQI ranging from 3.19 to 3.90). The water quality of Bosten Lake was characterized by distinct regional differences that arose from hydrodynamic processes within the lake and upstream water quality. The southwestern region exhibited the best water quality (mean CWQI of 3.47), whereas the northwestern region exhibited the worst (mean CWQI of 3.58). It is crucial to acknowledge that alongside the increase in industrial and agricultural effluent discharge monitoring, a series of ecological restoration projects for the lake basin have been initiated. Over time, the water quality of Bosten Lake showed gradual improvement (improvement rate of CWQI at 0.05/a). This study provides a critical scientific basis for enhancing the understanding and effective management of water quality in

the Bosten Lake Basin through a comprehensive analysis of its spatial and temporal evolution and driving mechanisms.

Full Text

Preamble

J Arid Land (2025) 17(9): 1234-1251
doi: 10.1007/s40333-025-0086-7; CSTR: 32276.14.JAL.02500867
Science Press Springer-Verlag

Spatial and temporal characterization of water quality in Bosten Lake, China based on comprehensive water quality index

GUO Mengjing[1]*, BAI Zichen[1], YUAN Bo[2], WANG Wen[1], ZHANG Tiegang[3], XIANG Ke[1], ZHANG Jiao[1], ZHAO Huiyizhe[1]
[1] State Key Laboratory of Eco-hydraulics in Northwest Arid Region, Xi' an University of Technology, Xi' an 710048, China
[2] College of Geology and Environment, Xi' an University of Science and Technology, Xi' an 710054, China
[3] Institute of Water Resources for Pastoral Area, Ministry of Water Resources, Hohhot 010020, China

Abstract

Water quality is a pressing issue affecting the sustainable development of lakes. To elucidate the spatial and temporal characteristics of water quality in Bosten Lake, China, this study constructed a comprehensive water quality index (CWQI) based on key water quality indicators, utilizing water quality data collected from 17 sampling sites spanning from 2011 to 2019. Key water quality indicators were determined using factor analysis, and the spatial and temporal characteristics of key water quality indicators and the CWQI were examined using multivariate statistical analysis. The key water quality indicators included pH, chemical oxygen demand (COD), water transparency (SD), NO_3^-, total dissolved solids (TDS), Cl^-, SO_4^{2-}, and electrical conductivity (EC). Furthermore, the contribution rates of all water quality indicators to the water quality were quantitatively elucidated using the SHapley Additive exPlanations (SHAP) values, thereby validating the factor analysis outcomes. Among the eight key water quality indicators, COD had the most significant influence on the water quality of Bosten Lake. The water quality condition of Bosten Lake remained at Class III from 2011 to 2019 (CWQI ranging from 3.19 to 3.90). The water quality of Bosten Lake was characterized by distinct regional differences that arose from hydrodynamic processes within the lake and upstream water quality. The southwestern region exhibited the best water quality (mean CWQI of 3.47), whereas the northwestern region exhibited the worst (mean CWQI of 3.58). It is crucial to acknowledge that alongside the increase in industrial and agricultural effluent discharge monitoring, a series of ecological restoration projects for the lake basin have been initiated. Over time, the water quality of Bosten Lake showed gradual improvement (improvement rate of CWQI at 0.05/a). This study provides a critical scientific basis for enhancing the understanding and effective management of water quality in the Bosten Lake Basin through a comprehensive analysis of its spatial and temporal evolution and driving mechanisms.

Keywords: water quality; chemical oxygen demand (COD); comprehensive water quality index (CWQI); multivariate statistical analysis; SHapley Additive exPlanations (SHAP); Bosten Lake

Citation: GUO Mengjing, BAI Zichen, YUAN Bo, WANG Wen, ZHANG Tiegang, XIANG Ke, ZHANG Jiao, ZHAO Huiyizhe. 2025. Spatial and temporal characterization of water quality in Bosten Lake, China based on comprehensive water quality index. *Journal of Arid Land*, 17(9): 1234–1251. https://doi.org/10.1007/s40333-025-0086-7; https://cstr.cn/32276.14.JAL.02500867

1 Introduction

Water quality assessment serves as a critical quantitative tool for characterizing the condition of water bodies and involves the systematic monitoring and analysis of various physical, chemical, and biological indicators tailored to specific objectives and requirements [?, ?]. The primary objectives of such assessments are to determine the extent of contamination, classify pollution levels, and establish a foundation for effective water management and pollution mitigation strategies. One of the difficulties encountered in assessing surface water quality is the need to establish an efficient monitoring system and select representative indicators. Current methodologies include index assessment, fuzzy comprehensive assessment, gray assessment, and artificial neural network assessment [?, ?, ?, ?, ?, ?]. The general approach to water quality assessment is a sequential data evaluation based on time series and independent sections, frequently neglecting the spatial and temporal heterogeneity and similarities of water pollutants. This is particularly problematic in large-scale studies involving multiple sections and numerous samples over extended periods and can lead to unnecessary repetitive calculations and complicated processes [?, ?, ?, ?, ?]. Neural network models and principal component analysis have shown considerable promise for reducing the dimensionality of water quality datasets [?, ?, ?].

The utilization and sustainable development of lake water resources in arid areas represent a global challenge. Bosten Lake is a typical lake in the arid area of Northwest China that not only constitutes a valuable water resource for agricultural development in the Bayingol Mongolian Autonomous Prefecture in Xinjiang Uygur Autonomous Region but also plays a crucial role in protecting the ecological environment in the lower reaches of the Tarim River, Xinjiang [?, ?, ?]. However, under the combined impacts of climate change and human activities, this lake faces severe water environmental challenges. Specifically, agricultural return flows containing salts and nitrogen (and/or phosphorus) compounds contribute to increased water mineralisation and aggravated eutrophication. Additionally, the regional climate, characterized by low precipitation and high evaporation rates, exacerbates these issues, further threatening the stability of the lake ecosystem and its ecological services [**?**, **?**].

Given the increasing pressure on the environment originating from climate change and human activities, it is essential to better understand the complex relationships between various water quality indicators [**?**, **?**]. Strong positive correlations between specific indicators may indicate common pollution sources and pathways. For example, SO_4^{2-} and Cl^- are major constituents of inorganic salts and are typically found at high levels in high-salinity lakes. Similarly, NH_3-N and NO_3^- exhibit a strong correlation, suggesting a shared origin and potential interactions with competing processes or other environmental factors. This inverse relationship may imply a competition between variables or differential responses to various environmental stimuli. For instance, dissolved oxygen (DO) levels tend to decrease as five days' biochemical oxygen demand (BOD_5) concentrations increase. Correlation analysis can elucidate the fundamental relationships between indicators, aiding in a more accurate diagnosis of water quality issues [**?**, **?**]. This enhances our understanding of water-quality dynamics and may inform more effective water management strategies.

The application of SHapley Additive exPlanation (SHAP) values in water quality analysis holds several advantages. Primarily, this approach enhances our understanding of the interrelationships between water quality indicators and the overall quality of water body. By quantifying the contribution of each indicator, SHAP values can help identify the key factors driving changes in water quality, facilitating the focused allocation of management resources based on the relative importance of these indicators [**?**]. Second, by providing a more accurate estimation of the influence of various water quality indicators, SHAP values can aid in formulating targeted interventions and policy decisions [**?**, **?**, **?**].

In this study, we initially calculated the correlations between various water quality indicators and identified key evaluation indicators through factor analysis. Subsequently, the water quality data were grouped by year and sampling site, and the spatial and temporal variation characteristics of water quality indicators in Bosten Lake were analyzed and tested using the Kruskal-Wallis nonparametric test. Finally, contribution rates were calculated to quantify the degree of influence of each indicator on the water quality assessment results. This study sheds light on the spatial and temporal evolution of water quality in Bosten Lake, providing a robust scientific foundation for its improved understanding and effective management.

2.1 Study area

Bosten Lake (41°46 -42°08 N; 86°19 -87°28 E), located in southern Xinjiang, is the largest freshwater lake in Xinjiang. The northern margin of the basin hosts multiple fault zones, including the South Tianshan Piedmont Fault, which regulates the interactions between surface water and groundwater. The lacustrine sediments primarily consist of fine sand, silt, and clay and are rich in evaporite minerals such as carbonates and sulphates. Salt alkalisation is extensive in the surrounding alluvial-pluvial plains, forming localized salt marsh wetlands. Hydrogeologically, the lake basin and its surrounding piedmont alluvial fan groundwater systems are in dynamic equilibrium, with both lateral groundwater recharge to the lake and infiltration of lake water into deep strata occurring simultaneously. The lake is approximately 55 km long and 25 km wide (Fig. 1a [Figure 1: see original paper]), with a surface area exceeding 1000 km^2 and a storage capacity of 7.3 km^3 [?, ?]. Precipitation feeding the Bosten Lake Basin is limited, with an average annual precipitation of 68.2 mm. The average annual evaporation is 1800.0–2000.0 mm [?]. Bosten Lake is fed by the Kaidu River, Huangshuigou River, and Qingshui River, and it provides water to the Konqi River [?]. The lake serves several functions, including water resource regulation and control of the Kaidu River, irrigation of the Konqi River Basin, and watershed ecological protection [?, ?]. Thus, Bosten Lake plays a crucial role in southern Xinjiang.

Bosten Lake, located on the Mongolian Plateau, is the largest brackish lake in Xinjiang [?, ?]. However, its water level and salinity have experienced significant fluctuations owing to climate change and human activities [?]. Specifically, there has been a discernible downward trend in water level, decreasing from 1048.0 m in 1958 to 1045.0 m in 1987. Conversely, a pronounced increase in water level occurred from 1987 to 2002, culminating in a high record of 1048.7 m. Following this period, the water level declined rapidly from 2003 to 2013, before rising again to 1047.9 m between 2014 and 2019. During the period of 1958–2019, opposing trends were observed in salinity and water level. The salinity increased from 0.39 g/L in 1958, peaked at 1.93 g/L in 1989, and subsequently decreased to <1.00 g/L by 2019. This resulted in a transition from freshwater to saline lake water conditions, followed by a return to freshwater conditions.

2.2 Monitoring indicators and detection methods

This study utilized a monitoring strategy spanning 9 years from 2011 to 2019. Monitoring was performed at 17 sampling sites (S1–S17) in Bosten Lake (Fig. 1b). Bosten Lake was divided into three distinct regions with unique hydrological characteristics. The northwestern region (S6, S7, S8, S9, S10, and S15) was predominantly fed by the Huangshuigou River and Qingshui River. The southwestern region (S1, S11, S12, S13, and S14) was mainly influenced by the Kaidu River and its downstream outlet. The central and eastern region of the lake (S2, S3, S4, S5, S16, and S17) was characterized by relatively poor hydrodynamic conditions. The primary objective of this study was to determine the prevailing water quality characteristics. Most samples were collected between May and October with additional samples collected in November in some years. This yielded a comprehensive dataset of 935 samples. Water transparency (SD) at each sampling site was measured using a Secchi disk. The pH, DO, electrical conductivity (EC), total dissolved solids (TDS), and water temperature at each site were measured using a multiparameter instrument (Hydrolab DS5X, HACH, Loveland, USA) and subsequently corrected in the laboratory.

Additionally, the collected water samples were transported in a shaded and refrigerated environment (4°C) for immediate analysis. The following indicators were analyzed: BOD_5, chemical oxygen demand (COD), permanganate index (CODMn), total nitrogen (TN), Cl^-, F^-, ${SO_4}^{2-}$, chlorophyll-a (Chl-a), NH_3-N, and ${NO_3}^-$. These analyses followed the water and wastewater monitoring analysis method proposed by the State Environmental Protection Administration of China (2002) and Ministry of Ecology and Environment of the People' s Republic of China (2002). The experimental results were analyzed using SPSS 27.0.1.

2.3 Correlation matrix

A correlation matrix is a mathematical table that represents the correlations between variables and is commonly used in statistical analyses. It presents the correlation coefficients for a set of variables in a matrix form, helping to identify patterns and relationships within the data, thereby enhancing the understanding of variable interdependence. The correlation matrix is calculated as follows:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}$$

where $r_{ij} = \frac{\mathrm{Cov}(X_i,X_j)}{\sqrt{DX_i \cdot DX_j}}$, R indicates the correlation matrix; r_{ij} is the correlation coefficient between variable X_i and variable X_j; $\mathrm{Cov}(X_i, X_j)$ is the covariance between variable X_i and variable X_j; DX_i and DX_j are the variance values of variable X_i and variable X_j; and E indicates the expectation value.

2.4 Statistical analysis

One of the primary challenges in assessing surface water quality is establishing an effective monitoring system that can accurately detect and quantify changes in water quality. This necessitates the identification of appropriate monitoring indicators that accurately reflect the evolving characteristics of assessed water. To achieve this goal, we conducted multivariate statistical analyses using SPSS 27.0.1 software.

2.4.1 Selection of assessment indicators based on factor analysis

The selection of assessment indicators is an important part of water quality assessment. If there are too many assessment indicators and there is a significant correlation between them, it will increase the computation effort and may hide some important characteristics. The objective of factor analysis is to reduce the number of inter-correlated indicators while ensuring the mutual independence of highly correlated indicators. The Kaiser-Meyer-Olkin (KMO) test is essential for evaluating the appropriateness of data for factor analysis. Generally, a KMO value above 0.7 indicates that the data possess good structural validity and are suitable for factor analysis. In this study, the number of significant factors was first determined based on the cumulative percentage of the total variance explained. Next, the Varimax rotation method was employed to derive a rotated factor matrix. Finally, key water quality indicators with significant contributions to each principal component were selected on the basis of a factor-loading matrix, typically considering absolute loadings above 0.5.

2.4.2 Significance test for spatial and temporal distribution differences of each indicator

Considering the heterogeneity and consistency characteristics of spatial and temporal distribution of water pollutants, spatial and temporal distribution differences were tested using a Kruskal-Wallis nonparametric test for each indicator. The data of the 8 key water quality indicators obtained through factor analysis were grouped into 9 temporal groups by year and 17 spatial groups by sampling site. We thus determined whether there were significant differences between the same sampling sites in different years as well as between different sampling sites in the same year.

2.5 Comprehensive water quality index (CWQI) assessment method

Owing to the combined effects of natural and human factors, one or two water quality indicators in some water bodies are more severe than those in water environmental function zones, whereas other indicators are relatively routine. The single-factor assessment method used in the "Environmental Quality Standards for Surface Water" (GB3838-2002) overlooks the influence of other indicators on water quality, which makes it difficult to fully reflect the comprehensive state of the water, potentially leading to inadequate water quality assessment. Based on the results of factor analysis, we selected eight water quality indicators to calculate the CWQI for assessing the water quality status of Bosten Lake. The CWQI consists of two components: the average and maximum of a single water quality index. The following formulae were used:

$$\mathrm{CWQI} = \overline{\mathrm{SWQI}} + \mathrm{SWQI}_{\mathrm{MAX}}$$

where $\overline{\mathrm{SWQI}}$ is the average of the single water quality index; $\mathrm{SWQI}_{\mathrm{MAX}}$ is the maximum value of the single water quality index with n items; SWQI_k and $\mathrm{SWQI}_{\mathrm{DO}}$ are the single water quality index values for the water quality indicator k and DO, respectively; C_k and C_{DO} are the water quality classes for water quality indicator k and DO, respectively, which can be determined by comparison with the "Environmental Quality Standards for Surface Water" (GB3838-2002), with values of 1, 2, ···, 6; ρ_k and ρ_{DO} are the measured concentrations of water quality indicator k and DO, respectively; $\rho_{k,\mathrm{lower}}$ and $\rho_{\mathrm{DO,lower}}$ are the lower limit concentration of C_k class interval for water quality indicator k and DO, respectively; and $\rho_{k,\mathrm{upper}}$ is the upper limit concentration of water quality class C_k for water quality indicator k; $\rho_{\mathrm{DO,upper}}$ is the upper limit concentration of water quality class C_{DO} for water quality indicator DO; and $\rho_{k,5,\mathrm{upper}}$ is the upper limit concentration of Class V for water quality indicator k. When the water quality level was worse than that of Class V for the water quality indicator, SWQI_k and $\mathrm{SWQI}_{\mathrm{DO}}$ were calculated using Equation 8 and Equation 9.

The comprehensive water quality level (CWQL) were determined based on the CWQI (Table 1).

Table 1 Water quality classification

CWQI Range	CWQL
[1, 2]	I
(2, 3]	II
(3, 4]	III
(4, 5]	IV
(5, 6]	V
>6	Inferior V

Note: CWQI is the comprehensive water quality index; CWQL is the comprehensive water quality level.

2.6 Contributions of water quality indicators to the water quality assessment results

SHAP values are derived from Shapley values, which originate from cooperative game theory. The objective is to assign a contribution value to each feature based on its contribution to the model outputs. A positive SHAP value signifies that an increase in a feature' s value correlates with a higher predicted pollution level, whereas a negative value indicates the opposite. This method is applied in the field of water quality prediction, where SHAP is utilized to calculate the contribution of each water quality indicator to the water quality results. By integrating machine-learning techniques, this approach provides more precise insights into the individual contributions of water quality indicators. In practical applications, the eXtreme Gradient Boosting (XGBoost) model exhibits superior performance, delivering increased accuracy and reliability when assessing the influence of indicators on the overall water quality [?]. The formula is as follows:

$$\phi_l = \sum_{S \subseteq N \setminus \{l\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \left[v(S \cup \{l\}) - v(S) \right]$$

where ϕ_l is the contribution of feature l to the prediction result; S is the collection of features; N is all feature columns; $\{l\}$ is a set containing only the single feature l; $N \setminus \{l\}$ is a new set obtained by removing feature l from set N; S is an arbitrary set composed of all features except feature l; $|S|$ is the number of features in S; $v(S)$ is the difference between the model' s predicted value for the feature column and the expectation of the predicted value; and $S \cup \{l\}$ is the new set formed by adding feature l to set S.

3.1 Correlation between water quality indicators

Figure 2 [Figure 2: see original paper] shows the correlation matrix for the various water quality indicators and elucidates their correlation coefficients. Correlation analyses revealed significant positive correlations between COD and CODMn and between COD and BOD_5, suggesting a common source of organic pollution in the lake. The correlation between COD and BOD_5 provides relevant information about the type of pollution and the degree of biodegradation related to the organic matter content [**?**]. TDS, Cl^-, $SO_4{}^{2-}$, and EC exhibited strong positive correlations [**?**], suggesting that lake salts may be enriched in $SO_4{}^{2-}$ and Cl^-. An increase in the ion concentration in water leads to higher EC owing to the enhanced ability of ions to conduct electrical currents. Moderate correlations were found between NH_3-N, TN, and other nutrient indicators such as $NO_3{}^-$ and Chl-a [**?**]. Nitric and ammoniacal nitrogen are the two main forms of nitrogen essential for the growth of aquatic plants and algae. The higher correlation of TN with $NO_3{}^-$ compared to that with NH_3-N suggests that more nitrogen is present in water in the form of $NO_3{}^-$ and that changes in $NO_3{}^-$ concentration have a stronger effect on TN. Weak or no correlation was found between pH and most other indicators, suggesting that pH variations were influenced by a complex set of factors or remained relatively stable during the study period.

A strong negative correlation was observed between DO and BOD_5, indicating that DO concentration decreases with increased organic matter consumption. The negative correlation between DO and all other indicators, except for SD, suggests that these indicators contribute to higher organic matter levels in water, which in turn decrease DO. The weak negative correlation between SD and most nutrient indicators indicates that SD is influenced by various factors such as suspended solids and phytoplankton, which can affect light penetration through the water column. Notably, higher SD and DO values indicate better water quality, whereas the opposite is true for the other monitored indicators.

3.2.1 Screening of key water quality indicators based on factor analysis

This study subjected 15 existing basic indicators to factor analysis to ascertain the key indicators for an integrated assessment of water quality. The KMO test value was 0.865, confirming the suitability of the original dataset for factor analysis. Eight indicators were selected based on the percentage of variance explained by each factor. The selected indicators included pH, COD, SD, ${NO_3}^-$, TDS, Cl^-, ${SO_4}^{2-}$, and EC, which were identified as crucial indicators for water quality assessment.

3.2.2 Spatial and temporal heterogeneity analysis

Figures 3 and 4 display the Kruskal-Wallis test results for each key water quality assessment indicator, with a comparison of monitoring values for indicators in different years as well as between different sampling sites. The pH exhibited no discernible alterations, while the remaining indicators (COD, SD, ${NO_3}^-$, TDS, Cl^-, ${SO_4}^{2-}$, and EC) exhibited notable variation. The reason for this phenomenon is that photosynthesis of algae consumes carbon dioxide, thereby increasing the pH of the water body. Concurrently, the respiration of microorganisms produces carbon dioxide, a process that results in a decrease in pH. In the event of a dynamic equilibrium between algae and microorganisms, the effects of these organisms on pH are mutually exclusive, thereby stabilizing the pH over time and space. The presence of specific buffer systems in lacustrine ecosystems, including carbonate-bicarbonate and phosphate buffer systems, confers resistance to acidic and alkaline stress. These buffer systems have demonstrated the ability to resist pH changes caused by the addition of acids and bases [?]. The differences between the samples over time indicate that the concentrations of water quality indicators vary significantly between the years and thus provide empirical support for further understanding of their temporal dynamics. The concentrations of water quality indicators showed high variability at the spatial scale, indicating differences among the various sampling sites in Bosten Lake. Significant differences help identify pollution sources more accurately. An accurate ecological model of lakes can only be constructed on the basis of spatial and temporal differences in water quality indicators. These differences should be considered in further research to accurately reflect the actual situation of lakes.

3.3 Spatial and temporal trends in the concentrations of key water quality indicators

Classifying the samples by time and space provides a multidimensional perspective for subsequent analyses, facilitating the discovery of variable trends across different times and geographical locations. As shown in Figure 3 [Figure 3: see original paper], the 935 samples were grouped into 9 categories by year. Except for SD, the overall trend of other water quality indicators showed a decline. Among these indicators, the pH level showed a slow decrease. In contrast, COD and SD levels showed opposite trends: COD gradually increased from 2012 to 2014, then decreased from 2014 to 2018, before increasing again in 2019; conversely, SD showed an almost opposite trend. Between 2011 and 2013, the surrounding industrial areas likely underwent development, with some small factories having inadequate sewage treatment facilities or illegal discharge. The discharge of industrial wastewater was not effectively controlled and contained many organic pollutants that increased the COD concentration upon entering the lake. After 2014, the water resources of Bosten Lake underwent more rigorous scientific management, and several ecological restoration projects were implemented, including wetland construction and aquatic plant planting. Concurrently, the volume of water entering the lake increased, which enhanced the lake' s renewal rate and helped dilute and remove pollutants, thereby reducing the COD concentration [**?**]. The concentration of NO_3^- gradually decreased over time, and TDS and EC showed similar trends. The correlation coefficient between TDS and EC was 0.80, indicating a strong correlation. Both TDS and EC showed an initial increase, followed by a decrease around 2013. Notably, the TDS concentration in Bosten Lake decreased to below 1000 mg/L in 2019, indicating an effective transition from slightly saline to freshwater conditions [**?**]. The trends in Cl^- and SO_4^{2-} concentrations were similar, with a strong correlation between them (correlation coefficient = 0.95). The concentrations of these two indicators increased between 2011 and 2012, followed by a decrease.

Figure 4 [Figure 4: see original paper] categorizes the samples into 17 groups according to the sampling sites. Except for S1, which exhibited lower pH levels, the other sampling sites showed minimal variation in pH at approximately 8.6. The SD values exhibited greater variability among different sampling sites, which can be primarily attributed to hydrodynamic factors and water composition. Water supply to the lake area is predominantly derived from the Kaidu River, which exhibits high hydrodynamic characteristics [**?**]. Suspended matter readily remains resuspended and is quickly dispersed, resulting in lower SD. In contrast, the central and eastern region of the lake, which has poor hydrodynamic conditions, exhibited higher SD values. The SD across the entire lake area generally fluctuated within a range of 100-300 cm. Overall, NO_3^- and the other five indicators (COD, TDS, Cl^-, SO_4^{2-}, and EC) exhibited an inverse trend. The variations in TDS, Cl^-, SO_4^{2-}, and EC across different sampling sites displayed remarkable consistency, reflecting strong positive correlations among these four indicators. TDS reflects the total dissolved salt content in water, with SO_4^{2-} and Cl^- being the predominant anions constituting these salts. Their concentrations significantly influence the chemical properties and ecological impacts within water bodies. EC, a measure of the ionic concentration in water, was strongly correlated with TDS and showed coordinated variation with the other three indicators. Furthermore, this study revealed significant fluctuations in most indicators at S7 and substantial changes at S14.

3.4 Spatial and temporal trends in water quality

The SWQI was computed for each key water quality indicator in each year (Fig. 5 [Figure 5: see original paper]), and the annual CWQI was determined based on the SWQI values (Table 2). Except for SD, the SWQI values of all other water quality indicators showed a progressive decrease. The SWQI values of COD were stable within the range of 3.81–4.87, whereas the SWQI values of SO_4^{2-} and EC fluctuated around 4.00. The SWQI values of pH, SD, and Cl^- fluctuated around 3.00. The SWQI values of TDS and NO_3^- were generally below 2.58 and 1.20, respectively. The CWQI showed an increasing trend from 2011 to 2013, followed by a gradual decrease from 2014, before a minor increase in 2019. All nine sets of water samples were classified as Class III. The best water quality, indicated by a CWQI of 3.19, was observed in 2018, whereas the worst water quality, with a CWQI of 3.90, was observed in 2013. The overall trend of CWQI suggests a progressive improvement in water quality over the years (at a change rate of 0.05/a).

The period from 2011 to 2013 witnessed a marked increase in the utilization of pesticides and fertilizers owing to the expansion of agricultural production, resulting in a substantial elevation in nutrient salts transported by surface runoff from agricultural land. During this period, nutrients were introduced into the lake at significant amounts via drains, leading to eutrophication of the water body and deterioration of water quality. In the absence of sufficient water from upstream sources or adequate precipitation levels, the lake' s self-purification capacity diminished because of the lower volume of available water. Hydrodynamic conditions were increasingly undermined, thereby hindering the dilution and dispersion of pollutants in the lake. This phenomenon has led to an accumulation of pollutants in localized areas, which has subsequently caused deterioration in water quality. After 2014, the regulation of industrial enterprises was improved, and ecological agricultural models were implemented, which led to a decrease in the discharge of pollutants. Several ecological restoration projects for lake locales have also been conducted, resulting in improved water quality. This study revealed a significant correlation between water quality dynamics and water-level fluctuations in Bosten Lake. Specifically, during the observation period (2011–2019), these variables exhibited a highly consistent temporal trend: during the period when the water level was continuously dropping (2011–2013), multiple water quality indicators showed synchronous deterioration with receding water levels, whereas during the water-level recovery phase (2014–2019), the water quality indicators showed a corresponding gradual improvement. This pronounced covariation pattern confirms the crucial regulatory role of hydrological processes in lake water quality. Throughout the entire period, COD was the most significant factor affecting CWQI values. SO_4^{2-} was found to be the least influential. NO_3^- and EC were also identified as crucial factors, whereas NO_3^- was found to be the least influential across all sampling sites.

The SWQI was computed for each sampling site (Fig. 6 [Figure 6: see original paper]), and the CWQI for each sampling site was determined based on the

SWQI values (Table 3). The SWQI values for all water quality indicators showed only slight fluctuations, except for SD, which exhibited highly significant variations. Notably, the SWQI values for COD ranged from 3.93 to 4.48, and the SWQI values for SO_4^{2-} ranged from 3.06 to 4.05, and the SWQI values for EC were approximately 4.00. The SWQI values for pH, Cl^-, and SD showed slight fluctuations around a mean of approximately 3.00. The SWQI values for TDS ranged from 1.82 to 2.35, and the SWQI values for NO_3^- were less than 1.20. The water quality of all sampling groups was categorized as Class III. The highest-quality water was observed at S14 (CWQI = 3.24), whereas the lowest-quality water was observed at S7 (CWQI = 3.67). Notably, the CWQI value at S14 showed a significant change, which coincided with the observed trends in the water quality indicators. At the spatial scale, the most influential factor on the CWQI was COD, whereas NO_3^- consistently remained the least influential across all sampling sites.

Sampling sites S7 and S14 are located at the mouths of their respective river systems, i.e., the Huangshuigou River and Kaidu River. The water quality at S7 was primarily influenced by inflow from Huangshuigou River, whereas that at S14 was predominantly affected by inflow from the Kaidu River. These inflows originate from the Tianshan Mountain range, where the vegetation and soil conditions contribute to the preservation of the natural environment. Soil erosion in this region is relatively minimal. As a seasonal river, Huangshuigou River features a wide riverbed and dispersed flow below the mountain outlet. Consequently, a substantial volume of water is lost through seepage, particularly in the pre-mountain alluvial flood fan section, where the water loss rate per kilometer reaches 2.88% [?]. The Huangshuigou River Basin significantly affects the volume of water entering Bosten Lake by several factors, including agricultural activities along the shoreline, cultivation of artificial reeds, and dam construction. Additionally, the riverbed is characterized by siltation and clogging, impairing its connectivity and ultimately preventing floodwaters from directly reaching Bosten Lake [?, ?]. In contrast, the Kaidu River is less influenced by human activities, allowing for a more direct flow into Bosten Lake, positively affecting water quality. Besides, the Kaidu River is less affected by climate change owing to the preservation of its upstream vegetation and the stability of its water supply. The Kaidu River has a relatively lower pH, water temperature, and TDS and TN concentrations [?]. Furthermore, the water quality of Huangshuigou River is influenced by elevated levels of suspended solids and sediments. Thus, it can be concluded that the water quality conditions in Kaidu River are superior to those in Huangshuigou River.

This study further investigated the monthly water quality of the three regions in Bosten Lake during May–November (Fig. 7 [Figure 7: see original paper]). The southwestern region of the lake exhibited the highest water quality (mean CWQI of 3.47), the central and eastern region of the lake demonstrated relatively poor water quality (mean CWQI of 3.53), and the northwestern region displayed the lowest quality (mean CWQI of 3.58). This distribution of water quality is consistent with the findings of Tohti et al. (2021). The Huangshuigou River, with its numerous agricultural irrigation drains, contributes a significant volume of industrial, agricultural, and domestic wastewater to Bosten Lake. Despite the beneficial filtration and adsorption of wetlands in lakeshore areas, the lake still contains a significant amount of highly concentrated pollutants [?]. The water level in the western part of Bosten Lake is higher than that in the eastern part, and the prevailing wind direction in the study area is from the northwest. The combined effects of wind direction and water-level differences lead to a gradual eastward flow of water from the Huangshuigou River, which lacks any downstream river connection within the Bosten Lake area. The central and eastern region of the lake has poor hydrodynamic conditions and limited water exchange capacity, contributing to high pollutant concentrations in this region. The southwestern region located at the mouth of Kaidu River introduces a substantial amount of high-quality freshwater into the lake area. Additionally, the two pumping stations outlet for Bosten Lake enhanced the water exchange capacity, thereby improving the water quality in the northwestern region.

Regarding the monthly distribution of water quality, the three regions displayed similar patterns, marked by gradual deterioration from May to August, minor fluctuations from August to October, and rapid recovery from October to November. September and October were identified as the most problematic months in terms of water quality, whereas November was the most favourable month. The water quality began to deteriorate in May. Over 95% of the contaminants from human activities that impact the water quality of Bosten Lake are attributed to livestock and poultry farming, fertilizer application, and agricultural practices [?]. Most agricultural activities in the Bosten Lake Basin occur between April and September and have a significant impact on water quality. This can be confirmed in the observed variations in TN, TP, and COD concentrations. The primary source of these pollutants is the drainage canals, which contribute to the deterioration of lake water quality [?]. The precipitation patterns in the basin show a concentration of precipitation in June, July, and August, which may result in fertilizer loss from agricultural fields. After precipitation ceases, an improvement in water quality can be observed.

3.5 Contributions of water quality indicators to water quality assessment results

Understanding the control mechanisms of various water quality indicators is essential for a comprehensive assessment of water quality. The pH exerts multifaceted effects on aquatic systems by affecting the physiological functions of aquatic organisms [?]. DO is a significant factor that can trigger catastrophic biological events in freshwater lakes [?]. Additionally, DO is a critical factor for aquatic life and directly participates in biochemical processes. Adequate levels of DO can maintain aerobic conditions that are favourable for pollutant degradation, whereas oxygen depletion creates anaerobic environments in which microbial activity produces toxic gases, such as hydrogen sulphide and methane. COD is a key indicator of organic matter content and reflects the oxygen consumption potential [?], with excessive organic loads leading to hypoxia and potential toxicity in aquatic ecosystems. Nutrient indicators such as NH_3-N and NO_3^- drive eutrophication when concentrations exceed certain thresholds, stimulating algal blooms that further deplete oxygen through the decomposition of organic matter [?, ?]. SD indicates the content of suspended and colloidal matter and serves as an effective control indicator for enhancing water clarity and ecological conditions. SO_4^{2-} increases water hardness through ionic interactions and may be converted to sulphides under anaerobic conditions, resulting in the generation of toxic hydrogen sulphide. Cl^- is a common anion that maintains osmotic pressure and electrolyte balance at appropriate levels, whereas F^- exhibits dual effects, which are beneficial at low concentrations but harmful at excessive levels. TDS represents the sum of inorganic and organic substances, with high levels altering the osmotic pressure and impairing biological functions. EC is closely correlated with ionic composition and concentration and serves as a sensitive indicator of variations in the dissolved salt content within water bodies [?].

The SHAP value analysis was performed to determine the impact of each indicator on the water quality assessment model (Fig. 8 [Figure 8: see original paper]). The top eight water quality indicators were COD, SD, SO_4^{2-}, BOD_5, EC, CODMn, DO, and NO_3^-, listed in order of significance. This suggests that these indicators play a significant role in determining the water quality outcomes. The results demonstrate a high degree of similarity with those obtained from the factor analysis. The high positive SHAP values for COD, SD, and SO_4^{2-} highlight the significant role of these indicators in the model' s predictions, emphasizing the serious implications of water quality pollution. Conversely, the negative SHAP values for NH_3-N and pH suggest that alterations in these indicators may correlate with improved water quality.

A comprehensive analysis of the impact of water quality indicators on model predictions was achieved by synthesizing the findings from the two plots in Figure 8. The COD concentration, a pivotal indicator of organic pollution, showed a high SHAP value, signifying the substantial impact of organic matter on water quality pollution levels. SD reflects the optical properties of the water body and exhibited a high SHAP value, indicating a robust correlation between water transparency and pollution levels. The elevated SHAP value of SO_4^{2-} may be linked to specific geological conditions or human-induced pollution sources, highlighting the role of sulphate in water quality alterations. Furthermore, the negative SHAP values for NH_3-N and pH suggest that, although these indicators may not always serve as prominent pollution indicators, changes in them may positively contribute to water quality improvement. For instance, a decrease in NH_3-N may indicate a reduction in nitrogen input to the water body, and pH stabilization may reflect the preservation of the water body' s self-purification potential.

4 Conclusions

This study analyzed the spatial and temporal variations in water quality of Bosten Lake and the main factors driving the variations. During the period of 2011-2019, the water quality of Bosten Lake was classified as Class III but exhibited an improving trend. Temporal variation analysis found that the pH of Bosten Lake, which is regulated by a water buffering system, exhibited no significant fluctuations, whereas all other water quality indicators exhibited gradual improvement. Water quality deterioration occurred during 2011–2013 due to agricultural non-point source pollution, but significant improvement was observed after 2014 due to the implementation of ecological restoration projects and enhanced industrial pollution control. Monthly water quality analysis found that water quality gradually deteriorated from May to August due to agricultural irrigation and precipitation effects, exhibited minor variations from August to October due to hydrological fluctuations, and rapidly recovered from October to November through pollutant sedimentation and natural water purification. Spatial heterogeneity analysis revealed that the southwestern region of the lake, primarily recharged by the Kaidu River with its high-quality water, maintained excellent water quality (mean CWQI of 3.47), whereas the northwestern region (mean CWQI of 3.58) exhibited significantly poorer water quality because of higher pollution loads from Huangshuigou River inflow. This study pioneered the application of SHAP values to water quality assessment in arid-area lakes, revealing that COD was the most critical indicator affecting water quality, with its contribution rate being significantly higher than that of the other indicators. Additionally, ${SO_4}^{2-}$ and SD were also identified as significant influencing factors. A long-term retention effect of pollutants has been formed owing to the input of large amounts of recalcitrant organic matter from agricultural non-point source pollution and industrial point source pollution in the upper reaches, combined with the slow water flow and poor hydrodynamic conditions in the lake area. This was the primary cause of persistently high COD concentrations in the lake. Further research is needed to fully explore the potential of SHAP values in water quality analysis and develop more effective water management strategies to promote the sustainable development and utilization of Bosten Lake and other similar inland lakes.

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.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (42377072, 52409105).

Author contributions

Conceptualization: YUAN Bo, GUO Mengjing
Data curation: GUO Mengjing, YUAN Bo
Formal analysis: BAI Zichen, YUAN Bo
Funding acquisition: GUO Mengjing
Investigation: BAI Zichen, XIANG Ke, ZHAO Huiyizhe
Methodology: YUAN Bo
Project administration: GUO Mengjing
Resources: WANG Wen, ZHANG Tiegang
Software: BAI Zichen, ZHAO Huiyizhe
Supervision: WANG Wen, ZHANG Tiegang, ZHANG Jiao
Validation: BAI Zichen, YUAN Bo
Visualization: BAI Zichen, YUAN Bo, GUO Mengjing
Writing - original draft: BAI Zichen, YUAN Bo
Writing - review & editing: GUO Mengjing, YUAN Bo

All authors approved the manuscript.

References

Carpenter S R, Caraco N F, Correll D L, et al. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, 8(3): 559–568.

Chen S F. 2021. Water quality change and its influencing factors in Bosten Lake. MSc Thesis. Urumqi: Xinjiang Normal University. (in Chinese)

Cheng C Y, Zhang F, Li X Y, et al. 2023. Variations in water storage of Bosten Lake, China, over the last two decades based on multi-source satellite data. *Journal of Hydrology: Regional Studies*, 49: 101496, doi: 10.1016/j.ejrh.2023.101496.

Dudgeon D, Arthington A H, Gessner M O, et al. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, 81(2): 163–182.

Gholizadeh M H, Melesse A M, 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. *Science of the Total Environment*, 566–567: 1552–1567.

Guo M J, Wu W, Zhou X D, et al. 2015. Investigation of the dramatic changes in lake level of the Bosten Lake in northwestern China. *Theoretical and Applied Climatology*, 119: 341–351.

Han J W, Kim T H, Lee S, et al. 2024. Machine learning and explainable AI for chlorophyll-a prediction in Namhan River Watershed, South Korea. *Ecological Indicators*, 166: 112361, doi: 10.1016/j.ecolind.2024.112361.

Kothari V, Vij S, Sharma S K, et al. 2021. Correlation of various water quality parameters and water quality index of districts of Uttarakhand. *Environmental and Sustainability Indicators*, 9: 100093, doi: 10.1016/j.indic.2020.100093.

Lan W H, Abiti, An H Y. 2003. Conservation and control of aquatic environment of Bosten Lake watershed, Xinjiang. *Journal of Lake Sciences*, 15(2): 147–152. (in Chinese)

Leng J C. 2023. Analysis of trend of water environment evolution in Bosten lake. *Technical Supervision in Water Resources*, (12): 85–88. (in Chinese)

Li D L, Xu X B, Li Z, et al. 2020. Detection methods of ammonia nitrogen in water: A review. *TrAC Trends in Analytical Chemistry*, 127: 115890, doi: 10.1016/j.trac.2020.115890.

Li J, Xu Y B, Niu Y, et al. 2023. Analysis of ecological flow threshold of seasonal river—Huangshuigou in the southern slope of the Tianshan Mountains. *Journal of Water Resources Research*, 12(5): 519–529. (in Chinese)

Liang L. 2021. The study on pH changes and algae growth response model in alkaline lakes under water diversion conditions. PhD Thesis. Chengdu: Sichuan University. (in Chinese)

Liao Y S, Xiao Q T, Li Y M, et al. 2024. Salinity is an important factor in carbon emissions from an inland lake in arid region. *Science of the Total Environment*, 906: 167721, doi: 10.1016/j.scitotenv.2023.167721.

Ling Y N, Liu Y, Peng J B, et al. 2020. Analysis of minimum ecological water level and water surplus in XiaoHu Lake District in Bosten Lake. *Environmental Engineering*, 38(10): 26–32, 60. (in Chinese)

Liu L, You X Y. 2023. Water quality assessment and contribution rates of main pollution sources in Baiyangdian Lake, northern China. *Environmental Impact Assessment Review*, 98: 106965, doi: 10.1016/j.eiar.2022.106965.

Lundberg S M, Lee S I. 2017. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30: 4768–4777.

Luo C M, Wang X H, Xu Y J, et al. 2025. Combining POA-VMD for multi-machine learning methods to predict ammonia nitrogen in the largest freshwater lake in China (Poyang Lake). *Journal of Water Process Engineering*, 72: 107511, doi: 10.1016/j.jwpe.2025.107511.

Luo H Q, Nong X Z, Xia H J, et al. 2024. Integrating Water Quality Index (WQI) and multivariate statistics for regional surface water quality evaluation: Key parameter identification and human health risk assessment. *Water*, 16(23): 3412, doi: 10.3390/w16233412.

Luo K, Hu X B, He Q, et al. 2017. Using multivariate techniques to assess the effects of urbanization on surface water quality: a case study the Liangjiang New Area, China. *Environmental Monitoring and Assessment*, 189: 174, doi: 10.1007/s10661-017-5884-8.

Lyu N, Guo M J, Zhao X, et al. 2024. Remote sensing inversion of water quality and spatiotemporal evolution characteristics of the Bosten Inland Freshwater Lake. *Arid Land Geography*, 47(6): 953-966. (in Chinese)

Mahmud R, Inoue N, Sen R. 2007. Assessment of irrigation water quality by using principal component analysis in an arsenic affected area of Bangladesh. *Journal of Soil and Nature*, 1(2): 8-17.

Mankin K R, Koelliker J K, Kalita P K. 1999. Watershed and lake water quality assessment: An integrated modeling approach. *Jawra Journal of the American Water Resources Association*, 35(5): 1069-1080.

Marandi A, Polikarpus M, Jõeleht A. 2013. A new approach for describing the relationship between electrical conductivity and major anion concentration in natural waters. *Applied Geochemistry*, 38: 103-109.

Ministry of Ecology and Environment of the People' s Republic of China. 2002. Environmental quality standards for surface water (GB3838-2002). [2024-08-18]. https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/shjbh/shjzlbz/200206/W020061027509896672057.pdf.

Misaghi F, Delgosha F, Razzaghmanesh M, et al. 2017. Introducing a water quality index for assessing water for irrigation purposes: A case study of the Ghezel Ozan River. *Science of the Total Environment*, 589: 107-116.

Muangthong S, Shrestha S. 2015. Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. *Environmental Monitoring and Assessment*, 187: 548, doi: 10.1007/s10661-015-4774-1.

Muchandi S S, Raikar R V, Virupakshi A S, et al. 2017. Assessment of lake water quality using factor analysis: A case study of North Belgaum City, India. *Asian Journal of Chemistry*, 29(1): 213-220.

Park J, Lee W H, Kim K T, et al. 2022. Interpretation of ensemble learning to predict water quality using explainable artificial intelligence. *Science of the Total Environment*, 832: 155070, doi: 10.1016/j.scitotenv.2022.155070.

Pham T L. 2017. Comparison between Water Quality Index (WQI) and biological indices, based on planktonic diatom for water quality assessment in the Dong Nai River, Vietnam. *Pollution*, 3(2): 311-323.

Rezaei A, Hassani H, Tziritis E, et al. 2020. Hydrochemical characterization and evaluation of groundwater quality in Dalgan basin, SE Iran. *Groundwater for Sustainable Development*, 10: 100353, doi: 10.1016/j.gsd.2020.100353.

Rudaru D G, Lucaciu I E, Fulgheci A M. 2022. Correlation between BOD_5 and COD-biodegradability indicator of wastewater. *Romanian Journal of Ecology & Environmental Chemistry*, 4(2): 80-86.

Rusuli Y, Li L H, Ahmad S, et al. 2015. Dynamics model to simulate water and salt balance of Bosten Lake in Xinjiang, China. *Environmental Earth Sciences*, 74: 2499-2510.

Saad A S, Massoud M A, Amer R A, et al. 2017. Assessment of the physico-chemical characteristics and water quality analysis of Mariout Lake, Southern of Alexandria, Egypt. *Journal of Environmental & Analytical Toxicology*, 7(1): 1–19.

Saimire T, Dilinuer A, Zhang M, et al. 2024. Characteristics and evaluation of spatial variation of water quality in Bosten Lake. *Transactions of Oceanology and Limnology*, 46(2): 143–150. (in Chinese)

Singh C K, Kumar A, Shashtri S, et al. 2017. Multivariate statistical analysis and geochemical modeling for geochemical assessment of groundwater of Delhi, India. *Journal of Geochemical Exploration*, 175: 59–71.

Šolcová A, Alahuhta J, Gałka M, et al. 2024. Developing a European aquatic macrophyte transfer function for reconstructing past lake-water chemistry. *Science of the Total Environment*, 954: 176613, doi: 10.1016/j.scitotenv.2024.176613.

State Environmental Protection Administration of China. 2002. *Water and Wastewater Monitoring and Analysis Methods* (4th ed.). Beijing: China Environmental Science Press, 210–213.

Sudheer K P, Chaubey I, Garg V. 2006. Lake water quality assessment from Landsat thematic mapper data using neural network: an approach to optimal band combination selection. *Journal of the American Water Resources Association*, 42(6): 1681–1695.

Sutadian A D, Muttil N, Yilmaz A G, et al. 2017. Using the Analytic Hierarchy Process to identify parameter weights for developing a water quality index. *Ecological Indicators*, 75: 220–233.

Tan Y F, Zhang Q, Zhu S, et al. 2015. Study on lake water quality assessment based on simple modeling and factor analysis. *Journal of Ecology and Rural Environment*, 31(3): 432–439. (in Chinese)

Tang X M, Xie G J, Deng J M, et al. 2022. Effects of climate change and anthropogenic activities on lake environmental dynamics: A case study in Lake Bosten Catchment, NW China. *Journal of Environmental Management*, 319: 115764, doi: 10.1016/j.jenvman.2022.115764.

Tohti G, Sai B, Zhang J P, et al. 2021. Study on water environmental characteristics of River Kaidu Catchment in Xinjiang. *Journal of Environment Engineering Technology*, 11(6): 1102–1109. (in Chinese)

Wang X Y, Tang X Y, Zhu M, et al. 2024. Predicting abrupt depletion of dissolved oxygen in Chaohu lake using CNN-BiLSTM with improved attention mechanism. *Water Research*, 261: 122027, doi: 10.1016/j.watres.2024.122027.

Wang Y M, Zhou X D, Engel B. 2018. Water environment carrying capacity in Bosten Lake basin. *Journal of Cleaner Production*, 199: 574–583.

Wu J L, Ma L, Zeng H A. 2013. Water quality and quantity characteristics and its evolution in lake Bosten, Xinjiang over the past 50 years. *Scientia Geographica Sinica*, 33(2): 231-237. (in Chinese)

Xu B, Zhou T, Kuang C Y, et al. 2024. Water quality assessment in a large plateau lake in China from 2014 to 2021 with machine learning models: Implications for future water quality management. *Science of the Total Environment*, 946: 174212, doi: 10.1016/j.scitotenv.2024.174212.

Zhang L, Fang W K, Li X C, et al. 2020. Water quality evaluation based on the water quality index in Lake Bosten: a large brackish inland lake in arid northwest China. *Desalination and Water Treatment*, 182: 68-76.

Zhang X, Wang Q S, Liu Y F, et al. 2011. Application of multivariate statistical techniques in the assessment of water quality in the Southwest New Territories and Kowloon, Hong Kong. *Environmental Monitoring and Assessment*, 173(1-4): 17-27.

Zhou H H, Chen Y N, Perry L, et al. 2015. Implications of climate change for water management of an arid inland lake in Northwest China. *Lake and Reservoir Management*, 31(3): 202-213.

Zhu M F, Yu X X, Chen K, et al. 2024. Spatiotemporal characteristics and driving factors of chemical oxygen demand emissions in China’s wastewater: An analysis based on spatial autocorrelation and geodetector. *Ecological Indicators*, 166: 112308, doi: 10.1016/j.ecolind.2024.112308.

Note: Figure translations are in progress. See original paper for figures.