



OPEN

Water quality assessment of east Tiaoxi River, China, based on a comprehensive water quality index model and Monte-Carlo simulation

Wei Jin¹, Yuan Li¹, Li Lu^{1,2}✉, Dong Zhang³, Shanying He¹, Jiali Shentu¹, Qiwei Chai¹ & Lei Huang¹

The comprehensive water quality index (*CWQI*) reflects the comprehensive pollution status of rivers through mathematical statistics of several water quality indicators. Using computational mathematical simulations, high-confidence *CWQI* predictions can be obtained based on limited water quality monitoring samples. At present, most of the *CWQI* reported in the literature are based on conventional indicators such as nitrogen and phosphorus levels, and do not include the petroleum hydrocarbons levels. This article takes a typical river in eastern China as an example, based on the 1-year monitoring at 20 sampling sets, a *CWQI* containing five factors, TN, NH₄⁺-N, TP, Σn-Alks, and ΣPAHs was established, and further predicted by a Monte-Carlo model. The predicted *CWQI* for each monitoring section is above 0.7, indicating that most of the monitoring sections are moderately polluted, and some sections are seriously polluted. The Spearman rank correlation coefficient analysis results show that TN, ΣPAHs, and Σn-Alks are the main factors influencing the water quality, especially the petroleum hydrocarbons have a significant impact on the middle and lower reaches due to shipping. In the future, more attention should be paid to petroleum hydrocarbon organic pollutants in the water quality evaluation of similar rivers.

River as important fresh water resources on land, bear important missions in people's daily life, but at the same time, river pollution, as a serious problem worldwide, threatens environmental safety and people's healthy¹. The collection of river water quality related data can be used as the basis for river management². Collection of data on river water pollutants is generally based on several aspects. The first is an indicator system based on eutrophication of water bodies, especially nitrogen, phosphorus, dissolved oxygen and other indicators have important impacts on river ecology. River assessment system based on indicators is widely used³⁻⁵. The second is about trace pollutants including PAHs, antibiotics, etc. These pollutants are extremely low in water, but their hard-to-degrade and cumulative properties will undoubtedly affect the ecological environment at a higher concentration^{6,7}. It has an important impact, and the collection of trace pollutant data has a longer-term direction for the formulation of river ecological management plans. The other is the collection of some related indicators of particulate matter and microorganisms. After a thorough understanding of the concentration level of pollutants in the basin, water quality assessment can be used to simulate the assessment of the pollution situation in the area. As a kind of water quality index (*WQI*), The Comprehensive Water Quality Index (*CWQI*) Method has the advantages of being able to process long series of data and comprehensive water quality for multi-point and multi section water bodies⁸. *CWQI* assigns water quality levels based on pollution classification standards. *CWQI* is in the range of (0, 0.4], which is a clean water body, and the water quality is better; in the range of (0.4, 0.7], the water quality is slightly pollution; in (0.7, 1], the water quality is moderate pollution; (1, 2], the water quality is a serious pollution; at (2, +∞), the water quality pollution is very serious⁹. The selected water quality index will be calculated according to the following formula:

¹Zhejiang Provincial Key Laboratory of Solid Waste Treatment and Recycling, School of Environmental Science and Engineering, Zhejiang Gongshang University, Hangzhou 310018, China. ²Instrumental Analysis Center of Zhejiang, Gongshang University, Hangzhou 310018, China. ³College of Materials and Environmental Engineering, Hangzhou Dianzi University, Hangzhou 310018, China. ✉email: LL0106@zjgsu.edu.cn

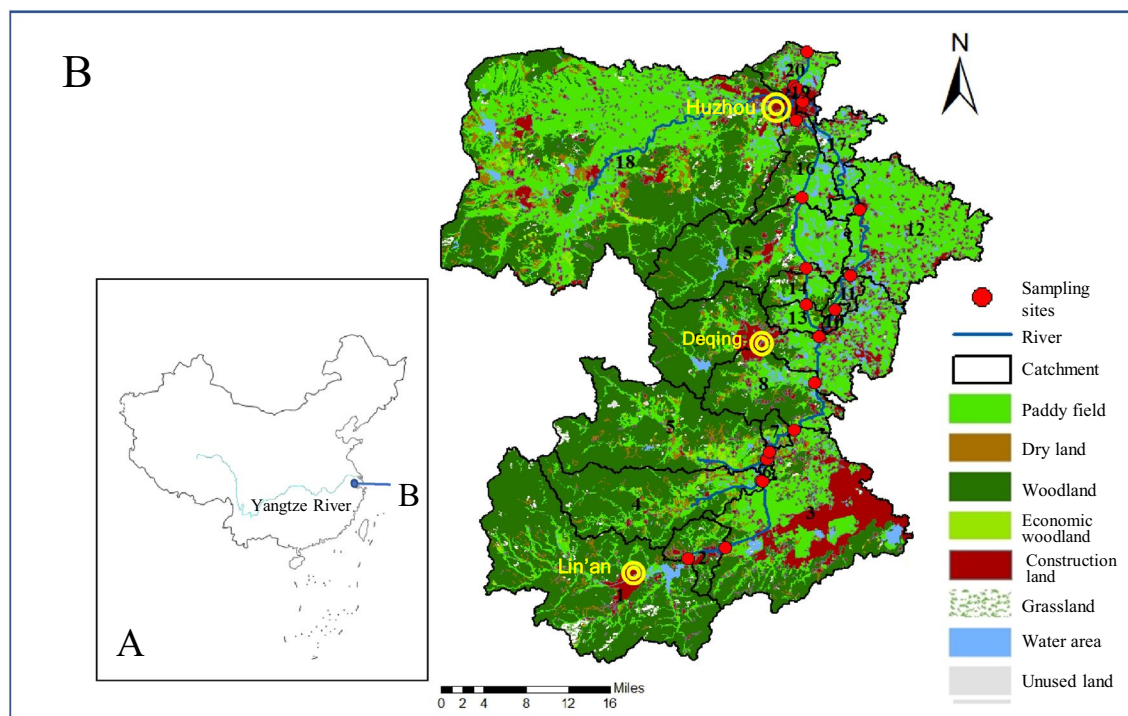


Figure 1. The catchments and sampling points (ArcGIS10.1, ESRI).

$$CWQI = \frac{1}{n} \sum_{i=1}^n P_i \quad (1)$$

In the formula, n is the number of participating water quality indicators; P_i is the single factor pollution index. The larger the value, the higher the degree of pollution.

For non-dissolved oxygen indicators:

$$P_i = C_i/C_0 \quad (2)$$

In the formula, C_i is the measured value of an evaluation index; C_0 is the standard value of an evaluation index.

Farzadkia et al.¹⁰ used $CWQI$ to model 14 pollutant indicators, including (temperature, electrical conductivity (EC), turbidity, total dissolved solids (TDS), total suspended solids (TSS), nitrate, nitrite, phosphate, dissolved oxygen (DO), chemical oxygen demand (COD), biochemical oxygen demand (BOD), pH, and total coliforms (TCs) and fecal coliforms (FCs). Use this to assess the water quality of the Yamchi Dam basin and identify the main source of water pollution in the basin. Davies et al.¹¹ used $CWQI$ simulation analysis was performed under the conditions of 17 pollutant indicators (including TP, NH_4^+-N , DO, etc.), and it was found that the value of $CWQI$ can reflect the change of the sample. It can be seen that $CWQI$ is mainly targeted at conventional indicators in previous studies, and $CWQI$ can only obtain corresponding water quality conditions from monitoring data during use. Once the number of samples is too small, the significance of its evaluation will be greatly reduced.

The Monte Carlo model can make up for the shortcomings of $CWQI$ when the amount of data is insufficient. Monte Carlo simulation belongs to a branch of computational mathematics. Through random sampling, the average value of a random variable or the probability of this event occurring is used as a solution to the problem. By performing thousands of Monte Carlo sampling simulations on the basic data of river water quality obtained by sampling a limited number of times, all possible comprehensive pollution indexes and their probabilities can be obtained under uncertain conditions of river water quality evaluation parameters, and the credibility and statistics of the simulation results significantly higher than the result obtained from a limited number of samples¹².

In this study, 20 sampling sites were set up in East Tiaoxi River to collect data for one year. And a Monte Carlo- $CWQI$ model with 5 pollutant indicators including TN, NH_4^+-N , TP, Σn -Alks, $\Sigma PAHs$ is established to evaluate the water quality. The aim of this article is to further improve the database of pollutants in East Tiaoxi River and analyze the comprehensive pollution level of the river and obtain the main pollutants and main influencing factors.

Material and methods

Study area. The East Tiaoxi River is a typical basin of the river network in Southeast China. It is one of the largest inflow systems of Lake Taihu. The length of the main stream is 151.4 km (30° 05' ~ 30° 57' N, 119° 28' ~ 120° 08' E), and the basin area reaches 2267 km² (Fig. 1). The upper stream of East Tiaoxi River is a mountain river, which is composed of three tributaries of North Tiaoxi River, Mid Tiaoxi River, and North Tiaoxi

Index	TN	NH ₄ ⁺ -N	TP	ΣPAHs	Σn-Alks
Unit	mg L ⁻¹	mg L ⁻¹	mg L ⁻¹	ng L ⁻¹	μg L ⁻¹
Standard value	1.0 ^a	1.0 ^a	0.2 ^a	280 ^b	50 ^c

Table 1. Calculated standard values of the five single-factor pollution indexes. ^aThe standard value of class III water in environmental quality standards for surface water in China (GB3838-2002). ^bOne hundred times the standard value of 2.8 ng L⁻¹ of benzo(a)pyrene (BaP) in GB3838-2002. ^cThe standard value of class III water in petroleum hydrocarbons in GB3838-2002.

River. The middle reaches of the river are narrow and are flooded areas. The lower reaches of the Hangjiahu Plain, with wide rivers, slow rivers, and advanced shipping. It is an important shipping area with a developed fish breeding industry. As an important water source in this region, the water quality of East Tiaoxi River is related to the ecological security of the basin and people's health.

According to the distribution of water systems in the basin, we use GIS (ArcGIS 10.1, ESRI Company, Redlands, California, USA) to divide the whole basin into 20 catchments, and set up a water quality monitoring section on the main channel in each catchment (Fig. 1). A total of 20 sampling points (set to P1–P20) were set, including 6 tributary monitoring points (P1; P2; P3; P4; P5; P18) and 14 main stream monitoring points. The land type data and population data of each catchment were collected and listed in Supplementary Materials (Table S1).

Sample collection and laboratory analysis. From September 2018 to August 2019, monthly sampling of 20 river water quality monitoring sections of East Tiaoxi River was conducted. Each month, sample collection for all sections was completed in two consecutive daytime hours. The surface water sample is collected in the middle of each monitoring section and 0.5 m away from the water surface, each sample is 10 L. After the samples were collected, they were acidified by adding sulfuric acid to a pH of less than 2 and shaken and stored in a refrigerator.

Standard Methods recommended by Ministry of Ecology and Environment of China were used for ammonia nitrogen (NH₄-N, HJ 535-2009), total nitrogen (TN, HJ 636-2012), and total phosphorus (TP, GB 11893-89) analysis. Specifically, NH₄-N was determined using a Nessler's reagent spectrophotometric method; TN was determined using an alkaline potassium persulfate digestion UV spectrophotometric method; TP was determined using ammonium molybdate spectrophotometric method^{13,14}. Water samples of 16 kinds of PAHs were passed through a LC-C18 cartridge for solid-phase extraction (SUPELCO Visiprep), eluted with dichloromethane, and after nitrogen blowing, the volume was adjusted with acetonitrile and determined by High Performance Liquid Chromatography (Agilent 1260, USA)¹⁵. 33 kinds of n-Alkanes (C8–C40) water samples were also passed through a LC-C18 cartridge for solid phase extraction and eluted with ethyl acetate. After nitrogen blowing, they were further eluted through a silica gel column with n-hexane and then nitrogen Blow, make up to volume with n-hexane, and measure by Gas Chromatography (Agilent 7890, USA)¹⁶.

Monte Carlo-CWQI Index model and spearman correlation coefficient. In this study, Monte Carlo simulation was used, combined with CWQI, CWQI was used as a predictive variable, and random variables were set as P_i for each evaluation index to establish a Monte Carlo-CWQI water quality evaluation model. The methods of establishing the comprehensive pollution index and Monte Carlo simulation are introduced in "The comprehensive pollution index method" and "Monte Carlo simulation method", respectively.

The comprehensive pollution index method. Based on the monitoring data of the above five pollutants, we established a comprehensive pollution index model for TN, NH₄⁺-N, TP, Σn-Alks and ΣPAHs, and established the standard values of the single factor indicators for the five pollutants according to the formula (2). TN, NH₄⁺-N and TP are based on the standard value of class III water in environmental quality standards for surface water in China (GB3838-2002). Compared with some European standards, the concentration of TN we set is at the same level as that of the category III (0.75 ~ 1.5 mg L⁻¹) of the UNECE (the United Nations Economic Commission for Europe) ambient water quality standard, and the concentration of NH₄⁺-N is between the level of class III and class IV of the ICPDR (International Commission for the Protection of the Danube River) water quality classification (0.6 ~ 1.5 mg L⁻¹). The concentration of TP is at the ICPDR standard class II level (0.2 mg L⁻¹)¹⁷. In the U.S., EPA divided the whole area into 14 distinct eco-regions, and proposed recommended ambient water criteria for each region. The TN standard selected here is comparable to the higher levels in the 14 regions (eg. eco-region XII, 0.9 mg L⁻¹), while the TP standard is significantly lower than those in all 14 regions¹⁸.

The standard value of Σn-Alks here is based on the standard value of class III water in petroleum hydrocarbons in GB3838-2002. However, prior to this study, there is no suitable standard value for ΣPAHs. In this study, the standard value of benzo(a)pyrene (BaP) 2.8 ng L⁻¹ in GB3838-2002 was expanded by a hundred times (280 ng L⁻¹) as the standard value for ΣPAHs. The specific standard values are shown in Table 1.

Monte Carlo simulation method. Based on the standard values of the above indicators, the single factor pollution index of each indicator is calculated by the formula (2). In order to ensure that the data is in a normal distribution, the obtained single-factor pollution index is subjected to natural logarithmic transformation, and SPSS 22.0 is used for Kolmogorov–Smirnov test. The relevant data of the single factor pollution index can be found in

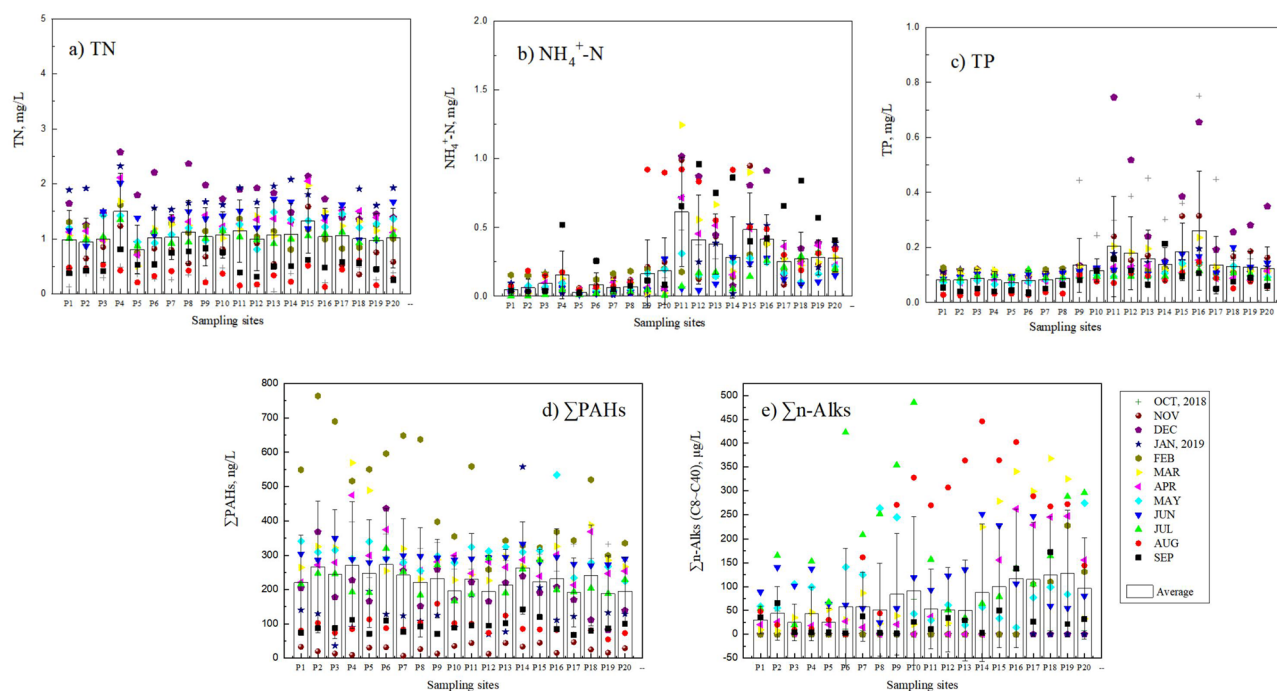


Figure 2. Annual average concentration and distribution of the typical pollutants in each monitoring section along the East Tiaoxi River. (a) TN, (b) NH₄⁺-N, (c) TP, (d) ΣPAHs, (e) Σn-Alks (C8~C40).

the supplementary materials (Table S2). Monte Carlo simulation was performed in Crystal Ball software in the Microsoft Excel environment. Edit the formula (1) in the Monte Carlo model, use *CWQI* as the predictive variable, and use the single factor pollution index as the random variable. Enter the mean and standard values of the single-factor pollution index into the Monte Carlo model and perform 10,000 sampling simulation.

In order to characterize the impact of various water quality indicators on the degree of water pollution, Spearman rank correlation coefficient (*SRCC*) will be calculated according to the following formula:

$$SRCC = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{[\sum_{i=1}^m (x_i - \bar{x})^2 \sum_{i=1}^m (y_i - \bar{y})^2]^{1/2}} \quad (3)$$

In the formula, *m* is the number of simulations; *x_i* is the ranking value of the input parameter; *y_i* is the ranking value of the output result; \bar{x} and \bar{y} are the means of the *x_i*s and the *y_i*s. The value of *SRCC* ranges from -1 to 1. The higher the *SRCC* value, the greater the influence of the input variable on the target variable.

Results and discussion

Concentration levels and distribution characteristics of the typical pollutants. According to the monitoring results, the annual average concentration of TN, NH₄⁺-N are 0.80–1.50 mg L⁻¹ and 0.03–0.61 mg L⁻¹, respectively. The TN concentration in East Tiaoxi River is relatively stable throughout the year, with a high concentration in December and an average concentration of 1.74 mg L⁻¹ in the basin (Fig. 2a). The variation trend of TN concentration in the river within one year is basically consistent with the change of runoff. The highest annual concentration appears in P4 of Mid Tiaoxi River monitoring section, with an average concentration of 1.50 mg L⁻¹ and a maximum concentration of 2.58 mg L⁻¹. There may be a large number of pollution sources of fertilization or discharge in this area. The highest annual concentration of NH₄⁺-N is found in monitoring section P11, with an average concentration of 0.61 mg L⁻¹ and a maximum concentration of 1.24 mg L⁻¹. Since monitoring section P11, the concentration of NH₄⁺-N in East Tiaoxi River began to rise, and the concentration in the middle and lower reaches of East Tiaoxi River was generally higher than that in the middle and upper reaches; the concentration of NH₄⁺-N reached the highest in August and September, with an average concentration of 0.40 mg L⁻¹ and 0.38 mg L⁻¹ (Fig. 2b).

The annual average concentrations of TP are 0.07~0.26 mg L⁻¹. The highest annual concentration of TP is found in monitoring section P17 at the lower reaches of East Tiaoxi River, with an average concentration of 0.13 mg L⁻¹ and a maximum concentration of 0.44 mg L⁻¹. The abnormal value of TP occurred in December 2018. From monitoring section P11 to monitoring section P20 at the inlet of the lake, the concentration increased to varying degrees, with an average concentration of 0.24 mg L⁻¹ (Fig. 2c).

In the monitoring of 16 kinds of PAHs, it is found that the ΣPAHs in each monitoring section of East Tiaoxi River basin is 0.19–0.28 μg L⁻¹, which is lower than other rivers in China^{19–21}, but the analysis on time scale shows that the concentration of PAHs in February is generally high, with an average concentration of 0.47 μg

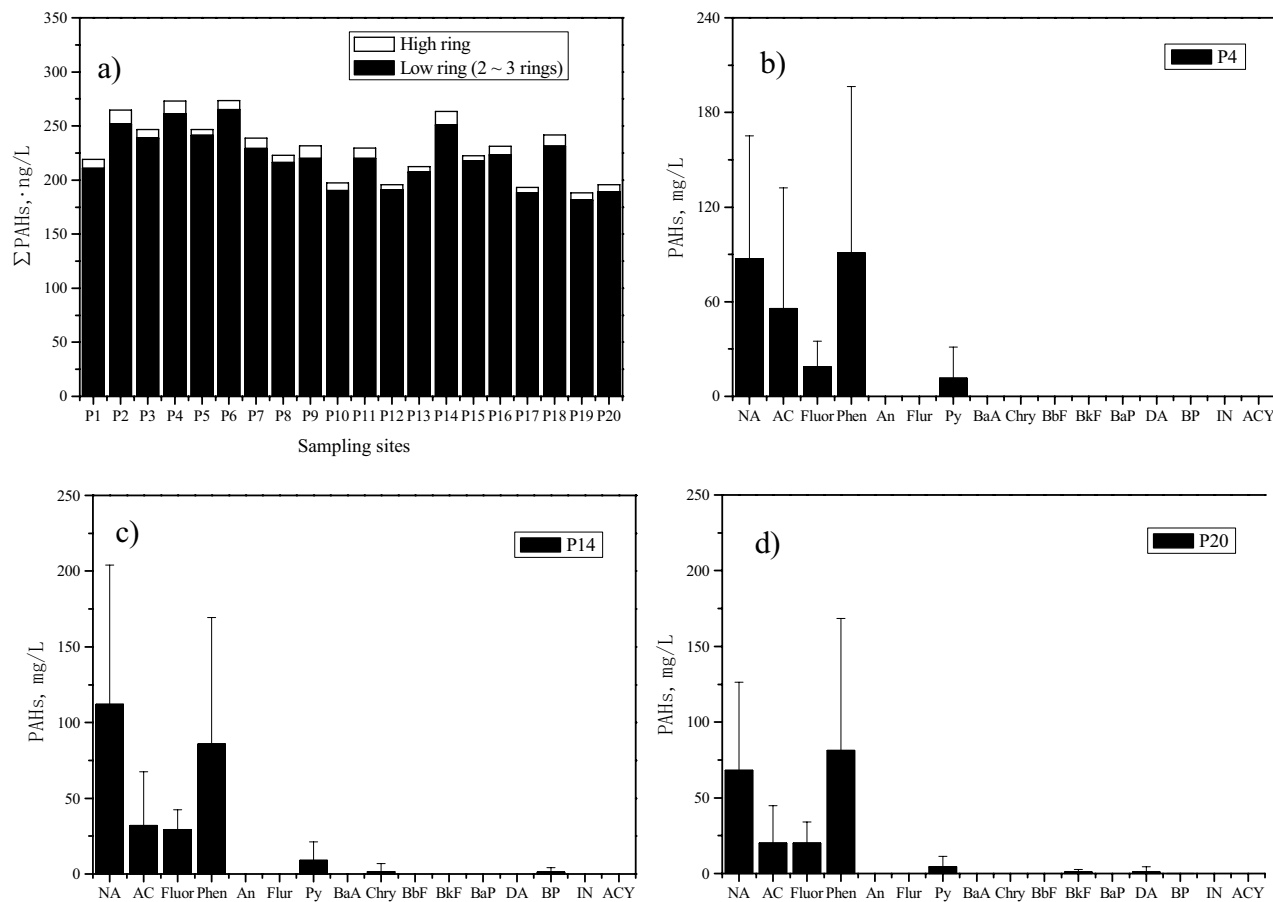


Figure 3. Distribution characteristics of the 16 PAH in the monitoring sections of P4, P14 and P20. (a) Σ PAHs, (b) P4, (c) P14, (d) P20.

L^{-1} , and some sections also have high concentrations in January and March. To determine the source of PAHs, more long-term monitoring is needed (Fig. 2d).

As the middle and lower reaches of East Tiaoxi River are widely used in shipping, n-alkanes (the n-alkanes with carbon chain lengths from 8 to 40, which are denoted by C8 to C40) was monitored to explore the impact of shipping on water quality. In the monitoring of n-alkanes (C8~C40), it was found that Σ n-Alks concentration of 20 monitoring sections in East Tiaoxi River was 24.92–127.10 $\mu g L^{-1}$. In terms of average concentration, the areas with higher concentration are mainly concentrated in the middle and lower reaches of East Tiaoxi River, which coincides with the shipping area. In terms of time, the high concentration points are mainly in July and August, with the average concentrations of 178.02 $\mu g L^{-1}$ and 202.42 $\mu g L^{-1}$. At this time, the rainfall is large, the water level rises, and the shipping is relatively frequent (Fig. 2e).

Analysis of types and characteristics of the organic pollutants. The characteristics of 16 types of PAHs were analyzed. Overall, the low loops (2 to 3 loops) accounted for about 98% of the total PAHs (Fig. 3a). Three monitoring sections (P4, P14 and P20) were selected in the upstream, middle and lower reaches of East Tiaoxi River for more detailed analysis. NA and Phen are the PAHs that mainly exist in East Tiaoxi River. AC, Fluor, and Py also exist. The concentration of Phen in the river basin as a whole did not change much. The concentration of NA in the upper and middle reaches is higher, which may be caused by the disturbance of the river in the upper and middle reaches. The concentration of AC is higher in the upper reaches of East Tiaoxi River (Fig. 3b–d).

Three monitoring sections (P2, P15 and P20) were also selected for the analysis of 33 types of n-alkanes in the upper, middle and lower reaches of East Tiaoxi River. C22 is the main pollutant in the upstream region of East Tiaoxi River, and C39 is the main pollutant in the middle and lower reaches (Supplementary Fig. S1). From the perspective of monthly distribution, the production of high-density n-alkanes in some months is the cause of excessively high n-alkanes concentration in the East Tiaoxi River.

The carbon preference index (CPI) was selected to conduct a qualitative analysis of the origin of n-alkanes in the East Tiaoxi River. CPI is the concentration ratio of the sum of odd-numbered alkanes to the sum of even-numbered alkanes. When $CPI > 5$, n-alkanes comes from plant wax source; when $CPI \approx 1$, n-alkanes is considered to come from fossil fuel^{22,23}.

Sampling sites	Average predicted CWQI value	Clean water (%)	Slight pollution (%)	Moderate pollution (%)	Serious pollution (%)	Highly severe pollution (%)
P1	0.76	1	47	38	13	1
P2	0.89	0	28	44	26	2
P3	0.84	1	40	38	19	2
P4	0.98	0	25	39	33	3
P5	0.72	1	54	35	9	1
P6	0.81	0	40	42	17	1
P7	0.82	0	45	37	16	2
P8	0.78	0	43	44	12	1
P9	0.88	0	30	44	24	2
P10	0.86	0	26	52	22	1
P11	1.08	0	10	39	48	3
P12	1.00	0	12	46	40	2
P13	0.98	0	22	44	31	3
P14	1.03	0	17	38	42	3
P15	1.14	0	6	34	57	3
P16	1.13	0	10	34	29	4
P17	0.95	0	20	45	34	1
P18	0.94	0	18	49	32	1
P19	0.90	0	25	47	27	1
P20	0.84	0	29	52	18	1

Table 2. Probability distributions of the CWQI for the different sampling locations in the East Tiaoxi River.

$$CPI = (C9 + C11 + C13 + C15 + C17 + C19 + C21 + C23 + C25 + C27 + C29 + C31 + C33 + C35 + C37 + C39) / (C8 + C10 + C12 + C14 + C16 + C18 + C20 + C22 + C24 + C26 + C28 + C30 + C32 + C34 + C36 + C38 + C40) \quad (4)$$

According to the formula (4), the *CPI* values of the monitoring sections P2, P15, and P20 are 1.37, 2.43, and 1.94, respectively. The *CPI* is higher than 1 but much lower than 5, and the n-alkanes in the East Tiaoxi River mainly come from fossil fuels.

Monte Carlo-CWQI for water quality evaluation. According to the water quality monitoring data of each section, a series of single-factor pollution index values were calculated by formula (2) and then the natural logarithmic transformation was performed. After the Kolmogorov–Smirnov test, the transformed single-factor pollution indices all obeyed normal distribution, and their mean values and probability distribution parameters are shown in Table S2.

Although the number of sampling times for water quality monitoring of any river is always limited, Monte Carlo simulation can obtain all possible comprehensive pollution indices and parameters under uncertain conditions of river water quality evaluation parameters by randomly sampling the limited basic data^{24–26}. In this study, the above-mentioned single-factor pollution index was set as a random variable, and the probability distribution parameters of each random variable were input. For each section, randomly sample 10,000 times within the range of the probability distribution, and the cumulative frequency distributions of CWQI for 20 monitoring sections, the CWQI prediction average value, and the probability of different degrees of pollution in each monitoring section can be obtained (Table. 2).

The results show that the average predicted CWQI of each monitoring section in East Tiaoxi River is above 0.7 (Table. 2). The water quality in North Tiaoxi River (monitoring section P5) is the best. Monte Carlo simulation shows that the probability of water pollution in monitoring section P5 is 54%, and the average CWQI prediction is 0.72. The water quality in the middle and lower reaches of the East Tiaoxi River is significantly inferior to that in the upper reaches. Judging from the predicted average values, the predicted CWQI average values of monitoring sections P11–P12 and P14–P16 are all greater than 1, which is a serious pollution. Among them, the probability of severe pollution in monitoring section P15 reached 57%, and the predicted average value of CWQI was 1.14. The water quality in this area was the worst. In summary, the overall water quality of the East Tiaoxi River is relatively poor, mainly with moderate pollution, and some monitoring sections in the lower reaches of the East Tiaoxi River are serious pollution, which requires attention (Fig. 4).

The results of SRCC analysis are shown in Fig. 5. The higher the SRCC value, the greater the influence of the input variable on the target variable. Results show that the water quality parameters that have a greater impact on the comprehensive pollution index in the middle and upper reaches of the East Tiaoxi River are ΣPAHs and TN. The middle and lower reaches of the East Tiaoxi River are mainly affected by TN and Σn-Alks. Compared to the upper reaches, the impact of Σn-Alks on the comprehensive pollution index of lower reaches has increased significantly, which indicates that the impact of shipping on n-alkanes pollution is obvious. Organic pollution in the East Tiaoxi River needs more attention.

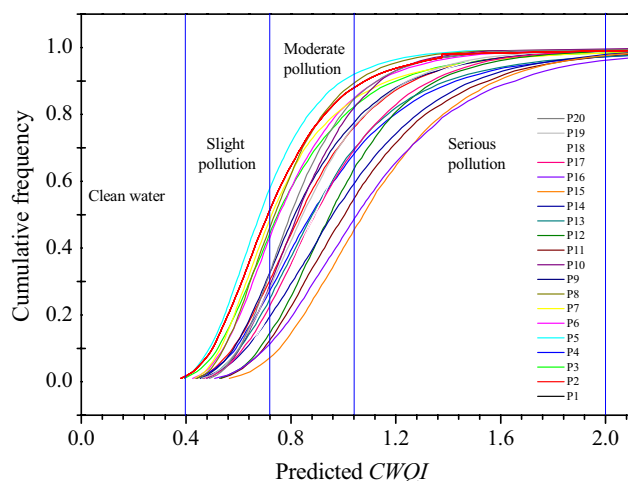


Figure 4. Cumulative probability curve of the CWQI for the different sampling locations in the East Tiaoxi River.

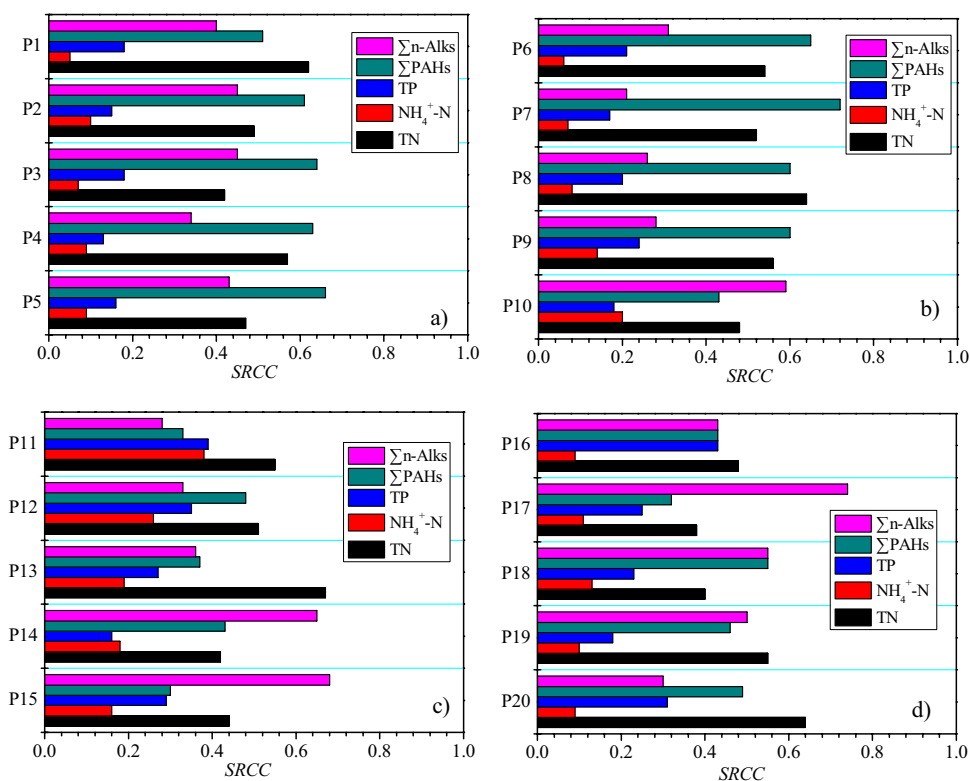


Figure 5. SRCC of the 5 evaluation indicators for the 20 monitoring sections in the East Tiaoxi River. (a) P1 ~ P5, (b) P6 ~ 10, (c) P11 ~ P15, (d) P16 ~ P20.

Correlation analysis of the CWQI prediction values of 20 sections with the paddy field, dry land, economic forest land, construction land, water area, grassland and other areas of the catchments did not show obvious correlation with a certain land type (correlation coefficients Less than 0.5). Obviously, the pollution sources and pollutant distribution trends of the 20 sections have been affected by multiple factors, and there is no consistent pattern. However, for the cross-sections (P3, P5, P6, P7, P9 and P18) with PAHs as main pollution factor (SRCC > 0.5), the CWQI prediction values of these sections have a certain positive correlation with the water area of the catchments, and the correlation coefficient was 0.71 (Fig. 6). This shows that the PAHs pollution in East Tiaoxi River may also have a certain relationship with water navigation.

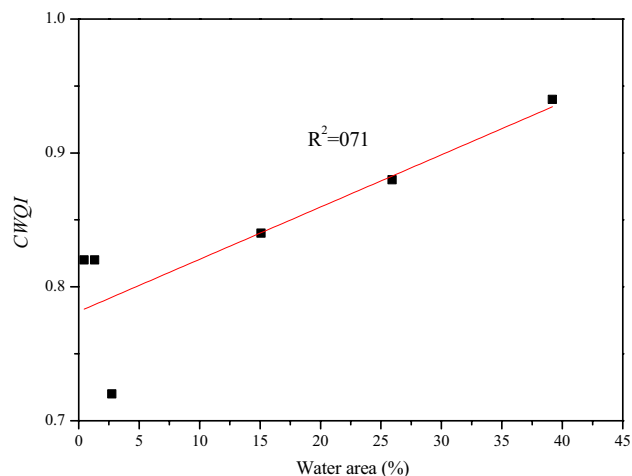


Figure 6. Correlation between the water area and CWQI for the catchments with the PAHs as the main pollution factor.

Conclusions

The concentration levels and spatial–temporal distributions of TN, $\text{NH}_4^+\text{-N}$, TP, $\Sigma\text{n-Alks}$ and ΣPAHs in the East Tiaoxi River were obtained, which supplemented and improved the database of pollutants in the river and provides some support for East Tiaoxi River water quality management and further improvement plans.

A Monte Carlo-CWQI model was established to predict the comprehensive pollution level of the river, two organic pollution indicators, ΣPAHs and $\Sigma\text{n-Alks}$ were incorporate into CWQI, and the weight of each pollution indicator in the comprehensive pollution assessment was analyzed by SRCC analysis. The established model makes up for the limited data in the sampling study and the insufficient attention to organic pollution in the previous comprehensive pollution assessment methods. The prediction and evaluation results of the model for East Tiaoxi River also provide data and theoretical support for the formulation of PAHs and n-alkanes concentration standards in surface water.

Data availability

All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

Received: 23 August 2021; Accepted: 3 June 2022

Published online: 16 June 2022

References

- Vorosmarty, C. J. *et al.* Global threats to human water security and river biodiversity. *Nature* **467**, 555–561 (2010).
- Behmel, S., Damour, M., Ludwig, R. & Rodriguez, M. J. Water quality monitoring strategies: A review and future perspectives. *Sci. Total. Environ.* **571**, 1312–1329 (2016).
- Akkoyunlu, A. & Akiner, M. E. Pollution evaluation in streams using water quality indices: A case study from Turkey's Sapanca Lake Basin. *Ecol. Indic.* **18**, 501–511 (2012).
- Bian, B., Zhou, Y. & Fang, B. B. Distribution of heavy metals and benthic macroinvertebrates: Impacts from typical inflow river sediments in the Taihu Basin, China. *Ecol. Indic.* **69**, 348–359 (2016).
- Romero, E. *et al.* Longterm water quality in the lower Seine: Lessons learned over 4 decades of monitoring. *Environ. Sci. Policy* **58**, 141–154 (2016).
- Meng, Y. *et al.* A review on occurrence and risk of polycyclic aromatic hydrocarbons (PAHs) in lakes of China. *Sci. Total. Environ.* **651**, 2497–2506 (2018).
- Chen, Y. H. *et al.* High-throughput profiling of antibiotic resistance gene dynamic in a drinking water river-reservoir system. *Water Res.* **149**, 179–189 (2019).
- Rickwood, C. J. & Carr, G. M. Development and sensitivity analysis of a global drinking water quality index. *Environ. Monit. Assess.* **156**, 73–90 (2009).
- Christensen, E. R., Li, A., Razak, I. A., Rachdawong, P. & Karls, J. F. Source of polycyclic aromatic hydrocarbons in sediments of the Kinnickinnic River, Wisconsin. *J. Great Lakes Res.* **23**, 61–73 (1997).
- Farzadkia, M., Djahed, B., Shahsavani, E. & Poureshg, Y. Spatio-temporal evaluation of Yamchi Dam basin water quality using Canadian water quality index. *Environ. Monit. Assess.* **187**, 168 (2015).
- Davies, J. Application and tests of the Canadian Water Quality Index for assessing changes in water quality in lakes and rivers of Central North America. *Lake Reserv. Manage.* **22**, 308–320 (2006).
- Signorino, C. S. Structure and uncertainty in discrete choice models. *Polit. Anal.* **11**, 316–344 (2003).
- Xu, J. L. & Shen, G. X. Growing duckweed in swine wastewater for nutrient recovery and biomass production. *Bioresour. Technol.* **102**, 848–853 (2011).
- Shen, Y., Lu, J., Chen, D. J. & Shi, Y. M. Response of stream pollution characteristics to catchment land cover in Cao-E River Basin, China. *Pedosphere* **21**, 115–123 (2011).
- Ramirez, C. E., Wang, C. T. & Gardinali, P. R. Fully automated trace level determination of parent and alkylated PAHs in environmental waters by online SPE-LC-APPI-MS/MS. *Anal. Bioanal. Chem.* **406**, 329–344 (2014).

16. Ranjit, G. *et al.* Degradation of n-alkanes and PAHs from the heavy crude oil using salt-tolerant bacterial consortia and analysis of their catabolic genes. *Environ. Sci. Pollut. Res.* **24**, 11392–11403 (2017).
17. OECD. *Proposed System of Surface Water Quality Standards for Moldova: Technical Report*. <https://www.oecd.org/env/outreach/38205662.pdf> (2007).
18. USEPA. *Ecoregional Nutrient Criteria for Rivers and Streams*. <https://www.epa.gov/nutrient-policy-data/ecoregional-nutrient-criteria-rivers-and-streams> (2020).
19. Liang, T., Wang, H., Rung, H. T. & Zhang, C. S. Agriculture land-use effects on nutrient losses in west Tiaoxi watershed, China. *J. Am. Water Resour. Assoc.* **40**, 1499–1510 (2004).
20. Wu, J. S., Jiang, P. K., Chang, S. X., Xu, Q. F. & Lin, Y. Dissolved soil organic carbon and nitrogen were affected by conversion of native forests to plantations in subtropical China. *Can. J. Soil Sci.* **90**, 27–36 (2010).
21. Zhang, Z. L., Hong, H. S., Zhou, J. L. & Yu, G. Phase association of polycyclic aromatic hydrocarbons in the Minjiang River Estuary, China. *Sci. Total. Environ.* **323**, 71–86 (2004).
22. Chen, B. L. *et al.* Distributions of polycyclic aromatic hydrocarbons in surface waters, sediments and soils of Hangzhou City, China. *Water Res.* **38**, 3558–3568 (2004).
23. Simoneit, B. R. T. *et al.* Composition and major sources of organic compounds of aerosol particulate matter sampled during the ACE-Asia campaign. *J. Geophys. Res.* **109**, D19S09 (2004).
24. Soleimani, H. *et al.* Groundwater quality evaluation and risk assessment of nitrate using monte carlo simulation and sensitivity analysis in rural areas of Divandarreh County, Kurdistan province, Iran. *Int. J. Environ. Anal. Chem.* <https://doi.org/10.1080/03067319.2020.1751147> (2020).
25. Bazeli, J. *et al.* Health risk assessment techniques to evaluate non-carcinogenic human health risk due to fluoride, nitrite and nitrate using Monte Carlo simulation and sensitivity analysis in Groundwater of Khaf County, Iran. *Int. J. Environ. Anal. Chem.* <https://doi.org/10.1080/03067319.2020.1743280> (2020).
26. Soleimani, H. *et al.* Probabilistic and deterministic approaches to estimation of non-carcinogenic human health risk due to heavy metals in groundwater resources of torbat heydariyeh, southeastern of Iran. *Int. J. Environ. Anal. Chem.* <https://doi.org/10.1080/03067319.2020.1757086> (2020).

Acknowledgements

This work was supported by the National Key R&D Program of China (2018YFC1802100) and the National Natural Science Foundation of China (41977006).

Author contributions

W.J.: writing-revised version, probability statistics of the model. Y.L.: investigation, writing—original draft. L.L.: conceptualization, project administration, writing, review & editing. D.Z. and S.Y.H.: resources. J. L.S.T.: review & editing-revised version. Q.W.C. and L.H.: investigation.

Competing interests

The authors declare no competing interests.


Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-14293-9>.

Correspondence and requests for materials should be addressed to L.L.

Reprints and permissions information is available at www.nature.com/reprints.

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

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022