

## RESEARCH ARTICLE

# Quantifying variation of non-point source pollution and its impact factors: A study of Nansi Lake Basin

Jiachen Liu<sup>1</sup>, Yuan Tian<sup>2</sup>, Rongqiang Ma<sup>2</sup>, Wenhui Xie<sup>2</sup>, Dongchao Wang<sup>1</sup>, Luoan Yang<sup>1</sup>, Xinyu Wang<sup>1</sup>, Le Yin<sup>1</sup>, Baolei Zhang<sup>1\*</sup>

**1** College of Geography and Environment, Shandong Normal University, Jinan, China, **2** Shandong Provincial Territorial Spatial Ecological Restoration Center, Jinan, China

\* [blzhangsd@163.com](mailto:blzhangsd@163.com)



## Abstract

Agricultural non-point source (NPS) pollution directly affects the quality of soil and water, ecological balance and human health, and is a key challenge to achieve sustainable environmental development and efficient resource management. Taking the Nansi Lake Basin (NLB) as the study area, this study explores the main sources of agricultural NPS pollution and its influencing factors, aiming to provide scientific basis for the management of water resources in the basin. Current studies usually use the runoff pollution partitioning method to estimate agricultural NPS pollution loads in runoff, but the accuracy of the analyses is limited by the incompleteness of water quality monitoring data, especially the lack of complete runoff records in some years. To compensate for this deficiency, this study simulated the river runoff based on the Long-Term Hydrological Impact Assessment (L-THIA) model, and applied the simulation results to the quantitative calculation of agricultural NPS pollution loads after verifying the model reliability through accuracy calibration. Based on L-THIA model, the spatial and temporal distribution data of agricultural NPS pollution in the basin from 2010 to 2020 were obtained, the distribution characteristics of chemical oxygen demand (COD) and ammonia nitrogen (NH<sub>3</sub>-N) were quantitatively assessed, and the impacts of natural and socio-economic factors on them were analyzed. A regression model was developed to simulate future agricultural NPS pollution through multiple regression analysis. The results showed that the total agricultural NPS pollution in the NLB showed an increasing trend during the study period. In particular, among the socio-economic factors, COD and NH<sub>3</sub>-N were significantly correlated with fertilizer application, pesticide use, rural employment and total population. Among the natural factors, topographic index, watershed area and gully density were positively correlated with pollutants, while slope and soil organic matter were negatively correlated. The results of this study raise awareness of the contribution of influencing factors and allow researchers and planners to focus on the most important NPS pollution sources and influencing factors. The study provides an important reference for the prevention and control of agricultural NPS pollution in the NLB, which is of great practical importance.

## OPEN ACCESS

**Citation:** Liu J, Tian Y, Ma R, Xie W, Wang D, Yang L, et al. (2025) Quantifying variation of non-point source pollution and its impact factors: A study of Nansi Lake Basin. PLoS ONE 20(2): e0318691. <https://doi.org/10.1371/journal.pone.0318691>

**Editor:** Wajid Ali, University of Peshawar National Centre of Excellence in Geology, PAKISTAN

**Received:** August 6, 2024

**Accepted:** January 20, 2025

**Published:** February 25, 2025

**Copyright:** © 2025 Liu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** All relevant data are within the manuscript and its [Supporting information](#) files.

**Funding:** This research was funded by the Natural Science Foundation of Shandong Province, China (ZR2021ME203, Baolei Zhang), Natural Science Foundation of Shandong Province, China (ZR2021QD127, Le

Yin), National Natural Science Foundation of China (42201308, Le Yin), Shandong Province Undergraduate Teaching Reform Project (Z20220004, Baolei Zhang), Jinan City-School Integration Project (JNSX2023036, Baolei Zhang).

**Competing interests:** The authors have declared that no competing interests exist.

## 1. Introduction

Non-point source (NPS) pollution is one of the major sources of pollution in the aquatic environment worldwide, especially in regions with intensive agricultural production activities, where its impact on the quality of water bodies and soils becomes more and more significant [1]. As a major source of NPS inputs, agricultural activities release large amounts of nitrogen and phosphorus into aquatic environments, leading to various problems such as toxic algal blooms, oxygen depletion [2] and loss of biodiversity [3]. It is estimated that 30%–50% of the world's land is affected by NPS [4], and about 60% of water body impairments are attributed to NPS pollution [5]. In Denmark, 94% of the nitrogen loads and 52% of the phosphorus loads from 270 rivers are due to NPS processes [6]. Similar trends were observed in the Netherlands, where NPS pollution accounted for 60% of total nitrogen pollutants and 40–50% of total phosphorus pollutants [7]. In China, NPS pollution is estimated to account for 81% of ammonia nitrogen ( $\text{NH}_3\text{-N}$ ) and 93% of phosphorus in total water pollution [8]. More than 60% of lakes are affected by eutrophication, half of which is caused by NPS pollution [9]. Therefore, NPS pollution is a growing global concern [8,10].

In recent years, significant progress has been made in the study of agricultural NPS pollution, which is mainly reflected in the continuous optimization of pollution load estimation methods and the wide application of modelling tools [11]. NPS pollution models can simulate the migration process of NPS pollution, predict the impact of pollutant loads on water quality, and provide effective means for the prevention and management of NPS pollution [12]. At present, the construction of NPS pollution model, NPS pollution spatial and temporal distribution simulation, impact assessment and pollution management, so as to promote the control and management of NPS pollution, is the focus of NPS pollution research [13]. With the development and application of geographic information system (GIS) technology, many hydrological and water quality models have been developed and integrated with GIS, such as the Storm Water Management Model (SWMM) [14], Hydrological Simulation Program-FORTRAN (HSPF) [15], Agricultural non-point Source (AGNPS) [16], and Soil and Water Assessment Tool (SWAT) [17]. These models bring great convenience and high evaluation accuracy to the simulation of spatial and temporal distribution of NPS sources, and provide a feasible method for the prevention and control of NPS pollution in watersheds [18]. However, these models require high input data, not only high-resolution land use, soil and topographic data, but also a large amount of detailed pollution data, meteorological data and agricultural practice data [19]. The difficulty of data collection limits the application of these models to a certain extent, and may even lead to significant deviations in simulation results if the data are incomplete or the model assumptions are consistent with the actual situation [20]. Therefore, there is an urgent need for methods to extract NPS pollution data from existing water quality monitoring data. Long-Term Hydrological Impact Assessment (L-THIA) models can be constructed using land use, soil type, and precipitation data, and their outputs include runoff volumes and NPS pollution loads. Despite the relative simplicity of its mechanism, the model can accurately assess long-term average surface runoff at the basin or urban scale. The L-THIA model was originally developed by Harbor [21] to simulate the impacts of land use changes on groundwater and wetland runoff. With the advancement of technology, the L-THIA model has been integrated with GIS, enabling the GIS-based L-THIA model to first estimate the runoff from smaller watersheds in the study area and then calculate the total runoff for the whole area, thus obtaining more realistic and reliable simulation results [22]. Many researchers have introduced the L-THIA model to simulate pollution in Chinese basins such as Miyun Reservoir area [23], Tao'er River Basin [24], Qingdao City [25], Guanlan River Basin [26], and Shiqiao River Basin [27], and achieved better results.

In addition, existing studies have gradually analyzed the influencing factors of agricultural NPS pollution. NPS pollution is characterized by stochasticity, universality and uncertainty [28], which leads to the complexity and diversity of its influencing factors. Many studies only focus on single or small number of pollutants such as nitrogen and phosphorus [29], ignoring the interaction between multiple pollutants and their comprehensive impact on the ecology of water bodies. Therefore, studying the impacts of NPS pollution involves many aspects, including its generation, transport and transformation mechanisms [30]. Identifying NPS pollution, especially understanding its development trends and determining its influencing factors, is a fundamental scientific challenge that is essential for solving China's deteriorating water environment problems. Unlike point source pollution, which can be traced back to a single source, such as discharges from wastewater treatment plants [21], NPS pollution originates from multiple sources and is influenced by a range of natural environmental and socio-economic factors. There are many studies on the identification of factors affecting agricultural NPS pollution [27,31–33]. For example, some scholars have attempted to analyze how farmers' behavior affects agricultural nitrogen and phosphorus pollution [34,35], while others have used the environmental Lorenz curve to assess the relationship between environmental quality and economic growth [22]. However, this part of the research only cuts from a single dimension and lacks a systematic exploration of the complex interactions between socio-economic factors and natural environmental factors, while ignoring the combined effects of watershed characteristics and socio-economic factors.

Nansi Lake is the largest freshwater lake in northern China and serves as a key hub along the eastern route of the South-North Water Transfer Project. As such, the water quality in the Nansi Lake Basin (NLB) is critical to the health of populations in many cities along the project route and directly influences the success of the overall water transfer initiative. Consequently, many scholars are focused on addressing water pollution prevention and control in the NLB. Based on water quality monitoring data, typical basin characteristics, and socio-economic data from the NLB, this study aims to (1) differentiate between NPS pollution and point source pollution using hydrological and water quality data, (2) quantitatively analyze the changes in chemical oxygen demand (COD) and  $\text{NH}_3\text{-N}$  NPS pollution in Nansi Lake, and (3) identify the key factors influencing COD and  $\text{NH}_3\text{-N}$  NPS pollution. This study will not only provide a scientific foundation for controlling water pollution in Nansi Lake but also help policymakers and researchers better understand the causes of water quality deterioration, particularly the increasing impact of NPS pollution.

## 2. Materials and methods

### 2.1. Study area

Nansi Lake ( $34^{\circ}27'\sim 35^{\circ}20'\text{N}$ ,  $116^{\circ}34'\sim 117^{\circ}21'\text{E}$ ) is the largest freshwater lake in Shandong Province, and even in North China, and one of the ten largest freshwater lakes in China, with a maximum water area of about 1266  $\text{km}^2$ , accounting for 45 per cent of Shandong Province's freshwater area. The NLB covers four cities, Jining, Heze, Zaozhuang and Tai'an, with an area of about 29000  $\text{km}^2$ , of which Tai'an contains only Ningyang County. The study area has a temperate monsoon climate, with average annual rainfall ranging from 590–820 mm. The main land use types are cropland and construction land (the secondary land use classification system is from the Chinese Academy of Sciences), and as of 2020, the total population is about 22.21 million people, of which the agricultural population accounts for 76.03 per cent of the total population.

## 2.2. Datasets

The land use data of Beijing came from Wuhan University (<https://zenodo.org/records/12779975>) and the land use types included 9 categories (cropland, forest, shrub, grassland, wetland, water, impervious, barren, snow/ice). We used the ArcGIS reclassification function to reclassify the land use data, and used the random point selection function to test the accuracy, and the accuracy of the interpretation results was more than 90.8%. Rainfall data were obtained from the Shandong Meteorological Data Monitoring Centre. Hydrological monitoring data were obtained from the Shandong Environmental Monitoring Centre, including water quantity and water flow rate at key cross sections of each river. 2010–2020 data on fertilizer application, poultry farming, investment in environmental protection, value added of the primary industry, value added of the secondary industry, and agricultural population were obtained from the statistical yearbooks of each prefecture-level city. Water quality monitoring data were obtained from the Shandong Environmental Monitoring Center, and the water quality factors include COD and  $\text{NH}_3\text{-N}$ . Rainfall data, hydrological monitoring data, water quality monitoring data, and statistical yearbook data were processed using Excel to remove outliers and then used directly.

## 2.3. Methods

**2.3.1. L-THIA model.** In order to compensate for the missing runoff volume in the water quality monitoring data, this study simulated the runoff volume of the river based on the L-THIA model, and used it for the calculation of the runoff volume of the agricultural NPS pollution after the accuracy check (Table 1). The L-THIA model was designed by Harbor [20] based on the SCS-CN (Curve Number) model, which is the key parameter of the SCS model, and its value quantifies the subsurface conditions of the basin, and reflects the impact of subsurface conditions on streamflow production and catchment with a quantitative indicator. The L-THIA model is based on the ratio of the actual infiltration volume (F) to the maximum possible soil infiltration volume (S), which is equal to the ratio of the actual surface runoff depth (Q) to the maximum possible runoff depth [36]:

$$\frac{F}{S} = \frac{Q}{P - I_a} \quad (1)$$

The calculation of each parameter is shown in the Table 1 below:

**Table 1. Meaning of each parameter of the L-THIA model.**

Name	Equation	Meaning
Precipitation (P)	$P = I_a + F + Q$	$I_a$ is the initial rainfall loss (mm), F is the actual infiltration (mm).
Antecedent soil moisture conditions (AMC)	$AMC = \sum_{i=1}^5 P_i$	Pre-soil moisture was judged based on the rainfall in the previous 5 days, and the pre-rainfall index AMC was introduced, with $P_i$ being the amount of rainfall (mm) in the last 5 days.
	$AMCI : CN_1 = \frac{4.2 * CN}{10 - 0.058 * CN}$	Changes in CN values are largely influenced by the AMC of soil pre-water conditions, assuming that the AMC values are related to the first five days of rainfall, and dividing the AMC into three classes according to the rainfall situation in the five days before the sub-flood $CN_1$ -soil drought but greater than the wilting point situation; $CN_2$ -soil moderately wet situation; $CN_3$ -soil wet situation with soil saturated with osmotic saturation. In this study, we extracted the first 5 days' rainfall of each land use type from the data of rainfall monitoring stations, and revised the built-in parameter CN of the model with reference [37] to the value of runoff volume of NLB in the study of water resources characteristics of NLB, and the CN values of the basin were taken as shown in Table 2.
	$AMCIII : CN_3 = \frac{23 * CN}{10 + 0.13 * CN}$	
Surface runoff depth (Q)	$Q = \frac{(P - I_a)^2}{P - I_a + S}$	S is the maximum possible soil infiltration (mm).

<https://doi.org/10.1371/journal.pone.0318691.t001>

Table 2. CN values taken in the NLB.

Land use	Soil wetness		
	Dry	Medium wet	Wet
Cropland	64	75	82
Forest	57	73	82
Grassland	68	79	86
Townland	77	85	90
Rural neighborhoods	46	65	77
Transport land	76	85	89
Water	98	98	98
Unused land	77	86	91

<https://doi.org/10.1371/journal.pone.0318691.t002>

**2.3.2. NPS pollution estimation.** In this study, based on the mass balance of pollutants, the amount of agricultural NPS pollution in runoff was obtained using the runoff pollution partitioning method, which identifies the pollution load in the river monitoring section as consisting of point source and NPS pollutants, since the number of pollutants discharged into the river from point source pollution is relatively stable and fixed [20]. Agricultural NPS pollution mainly occurs during periods of abundant water level. During periods of low water level, the average concentration of agricultural NPS pollutants is negligible, and the monitoring values are used only to indicate the concentration of point source pollutants. Based on this assumption, the separation process is described in equations (2) to (5).

The total amount of pollutants in the runoff can be calculated as:

$$L = \sum_{i=1}^{12} C_i \times Q_i \quad (2)$$

where  $L$  is the total amount of pollutants in runoff (kg),  $C_i$  is the monthly monitoring concentration (g/L);  $Q_i$  is the monthly runoff volume (m<sup>3</sup>); and  $i$  denotes the month.

The amount of point source pollution can be calculated as:

$$PS = \sum_j^4 C_j \times Q_j \quad (3)$$

where  $PS$  is the point source pollution load (kg);  $C_j$  the concentration of the pollutant in the dry season (g/L),  $Q_j$  is the dry season runoff volume (m<sup>3</sup>), and  $j$  is the dry season, which in this study is February to May.

$$NPS = L - PS \quad (4)$$

Therefore, the amount (kg) of agricultural NPS pollutants can be calculated as follows:

$$L_{NPS} = \frac{\sum_{i=1}^n UNPS_i}{A_i} \times A \quad (5)$$

The total amount of agricultural NPS pollutants in the NLB was calculated as follows:

$L_{NPS}$  is the total amount of specific NPS pollutants in the entire NLB;  $i$  is an area of a control unit;  $UNPS_i$  is the amount of specific NPS pollutants in subbasin  $i$ ;  $A$  is the NLB; and  $n$  is the number of subbasin control units.

**2.3.3. Selection of influencing factors.** We selected 26 factors as characteristic factors that may affect the basin through expert consultation and based on the literature [38] and previous studies, and then performed the Breusch-Godfrey LM test [39] to determine the autocorrelation between the analyzed factors. Finally, we used a method employing the Pearson correlation test to select basin characteristics that were statistically significant for NPS pollutants. The factors are shown in the Table 3 below:

**2.3.4. Multiple regression analysis and testing.** Based on the physical and socio-economic characteristics of the study area, COD and  $\text{NH}_3\text{-N}$  pollution in 11 control units was investigated using multiple regression analysis. In general, generalization between NPS pollution and basin characteristics may yield a population-like model for the dataset rather than a single best model for stepwise regression analysis. After autocorrelation and correlation tests, several multiple linear regression equations were developed in the software SPSS and expressed as:

$$L_a = a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_nX_n + C \quad (6)$$

where  $L_a$  is the number of COD and  $\text{NH}_3\text{-N}$  NPS pollution (kg),  $a_1, a_2, \dots, a_n$  are regression coefficients,  $X_1, X_2, \dots, X_n$  are basin characteristics and  $C$  is a constant term. The data were

**Table 3. Influence of characteristic factors on agricultural NPS pollution.**

Type	Code	Short name	Name
Socio-economic factors	X1	YTP	Total population at the end of the year
	X2	PD	Population density
	X3	GDP	Gross Domestic Product
	X4	PGDP	Per capita GDP
	X5	PPR	Proportion of primary sector
	X6	PSR	Proportion of secondary sector
	X7	PTR	Proportion of tertiary sector
	X8	RPA	Rural population
	X9	NAPR	Proportion of non-agricultural population
	X10	YAP	Rural practitioners
	X11	FAA	Total fertilizer use
	X12	PAA	Total pesticide use
	X13	RECA	Total rural electricity use
Natural factors	X14	BD	Patch density
	X15	LSI	Landscape shape index
	X16	MPS	Mean plaque size
	X17	SLOPE	Slope
	X18	PRE	Precipitation
	X19	NDVI	Normalized Difference Vegetation Index
	X20	DA	Basin area
	X21	TD	Trench density
	X22	SOMC	Soil organic matter
	X23	PAL	Proportion of cropland
	X24	PGAL	Proportion of garden land
	X25	PFL	Proportion of forest land
	X26	PGRL	Proportion of grassland

<https://doi.org/10.1371/journal.pone.0318691.t003>



analysed as multiple linear regression by applying a logarithmic transformation to all variables to satisfy the normality requirement and establishing a regression equation of the form:

$$\log(L_b) = \log(C) + b_1 \log(X_1) + b_2 \log(X_2) + b_3 \log(X_3) + \cdots + b_n \log(X_n) \quad (7)$$

We used  $R^2$  checks of these equations to select models with better-fitting simple independent variables. Finally, as a result of the analyses, the equations for each pollutant were retained and the significance of the individual p-values for the remaining variables was tested.

The predictive ability of the models was assessed using the relative error (RE) and the Nash efficiency coefficient [40], calculated as:

$$RE = \frac{L_s - L_0}{L_0} \times 100 \quad (8)$$

$$E_{NS} = 1 - \frac{\sum_{i=1}^n (L_{01} - \bar{L}_{s1})^2}{\sum_{i=1}^n (L_{01} - \bar{L}_0)^2} \quad (9)$$

where  $L_{01}$  is the NPS pollution monitoring value,  $L_{s1}$  NPS pollution modelling value, control unit NPS pollution average monitoring value,  $n$  is the total number of years.

## 2.4. Data preprocessing

In this study, the basin analysis tool was used to divide sub-basins by using the river water quality monitoring points in the NLB as outlets, combined with the methods of customizing the major rivers in the plains and determining the catchment basins. Since part of the NLB is dominated by plain terrain, the sub-basin delineation method based on basin analysis tool is not effective. Therefore, we referred to related studies [32] and conducted sub-basin delineation based on the basin control cross sections in Shandong Province delineated by Shandong Provincial Environmental Protection Department, and corrected the results of the sub-basin delineation of the Nansi Lake sub-basin considering the availability of the data to finally delineate 11 sub-basins.

## 3. Results

### 3.1. Average total pollution from NPS in the NLB

The average pollution amount of NPS of COD in the basin of the Nansi Lake from 2010 to 2020 is shown in Table 4. The average pollution amount of NPS of COD in the central part of the basin is relatively good. Overall, the NPS pollution in the central part of the basin is relatively good, and the spatial distribution of pollution is generally characterized as 'low in the center, high on both sides, and more concentrated'. In 2010, the three sub-basins with the largest amount of COD NPS pollution were Dongyu River, Wanfu River and Guangfu River. The three sub-basins with the smallest amount of NPS pollution were Chengguo River, Beisha River, and Shushui River Basin. In 2020, the largest amount of NPS pollution of COD in the NLB will continue to be the Dongyu River, followed by the Zhuzhaoxin River Basin, which has the largest number of COD. In 2020, the largest amount of COD NPS pollution in the NLB is still the Dongyu River, followed by the Zhuzhaoxin River and Xinxue River, showing a substantial increase in the amount of pollution relative to that in 2010, and the smallest amount of COD NPS pollution in 2020 will be in the basins of the Zhushui River, the Baima River, and the Beisha River, and the size of the amount of pollution did not fluctuate too

Table 4. The average amount of COD and NH<sub>3</sub>-N pollution from 2010 to 2020.

Basin control unit	COD pollution/t/m <sup>3</sup>	NPS COD pollution		NH <sub>3</sub> -N pollution/t/m <sup>3</sup>	NPS NH <sub>3</sub> -N pollution	
		Amount/t	Proportion/%		Amount/t	Proportion/%
Baima River	18240.72	14675.32	80.45	44.85	34.94	77.90
Beisha River	20540.16	16144.59	78.60	115.21	68.61	59.55
Chengguo River	20491.29	15373.49	75.02	54.32	39.16	72.09
Dongyu River	188737.27	157556.96	83.48	193.89	135.19	69.73
Guangfu River	43519.39	35739.08	82.12	108.15	83.82	77.51
Liangji River	52004.53	40730.48	78.32	83.67	64.59	77.20
Si River	52443.83	41599.53	79.32	65.64	48.18	73.41
Wanfu River	80628.45	64251.08	79.69	164.75	129.12	78.37
Xinxue River	84368.17	64232.00	76.13	332.13	245.61	73.95
Zhushui River	16267.87	13361.36	82.13	69.16	45.32	65.52
Zhuzhaoxin River	158278.12	115686.72	73.09	236.49	139.74	59.09
Total	735519.79	579350.62	78.77	1468.26	1034.28	70.44

<https://doi.org/10.1371/journal.pone.0318691.t004>

much in comparison with that in 2010. The percentage of COD NPS pollution is maintained at 73.09%–83.48%.

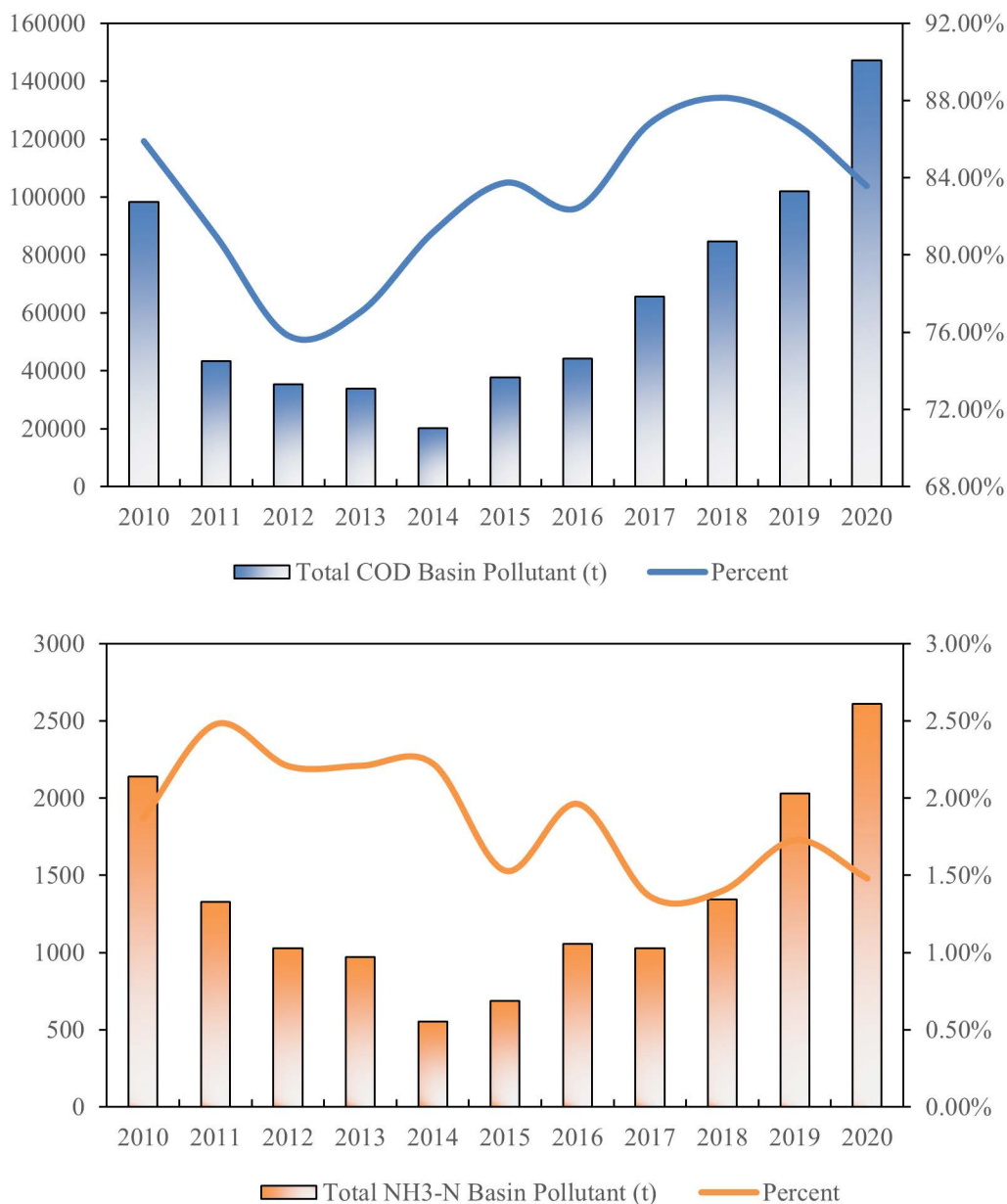
The time series changes of NH<sub>3</sub>-N NPS pollution in the basin of the Nansi Lake from 2010 to 2020 are shown in Table 4. The overall distribution of NH<sub>3</sub>-N pollution in the central and eastern parts of the basin is generally in good condition. Overall, the NH<sub>3</sub>-N pollution in the central and eastern part of the basin is generally in good condition, and the overall distribution of pollution is characterized by 'high in the west, low in the center, locally concentrated and unevenly distributed'. The most obvious change in the amount of NH<sub>3</sub>-N pollution from 2010 to 2020 was in Xinxue River, where the amount of NH<sub>3</sub>-N pollution showed an upward trend, increasing by about 13 times. The three sub-basins with the largest amount of pollution are Xinxue River, Chengguo River, Baima River and Si River Basin in the order of the largest amount of pollution, the amount of pollution is less than 100 t, and the proportion of NH<sub>3</sub>-N pollution from unorganised discharge stays in the range of 59.09% to 78.37%.

### 3.2. Temporal changes in total NPS pollution in the NLB

By analyzing the statistical data of the NLB as well as the monitoring data of the river, this study reveals the spatial and temporal evolutionary characteristics of the pollutants in the basin, which provides a key basis for further analyses. The time span of the study is 2010–2020, and the time series changes of total COD and NH<sub>3</sub>-N runoff pollutants within the NLB are shown in Fig 1 (Temporal changes in total runoff pollution in NLB, 2010–2020). In terms of the overall trend, the changes in the values of these two indicators generally show an upward trend. Specifically, the values of each pollutant change along a trajectory of first decreasing and then increasing, and show significant differences between the magnitude of increase and decrease. Our study found that COD is the main pollutant in the basin. In 2020, the fluctuation of COD emissions was more significant. Compared to COD, the change in NH<sub>3</sub>-N emissions was relatively small, with an overall rate of change of 92%.

According to the time-series change process of COD pollution volume in the NLB from 2010 to 2020 (Fig 2 Temporal changes in total runoff COD pollution in each sub-basin, 2010–2020), the COD pollutant emissions in the central and eastern parts of the basin have the smallest change and are always at the lower value, while the COD pollution volume in the western part is always at the higher value, and the overall change is more uniform. The



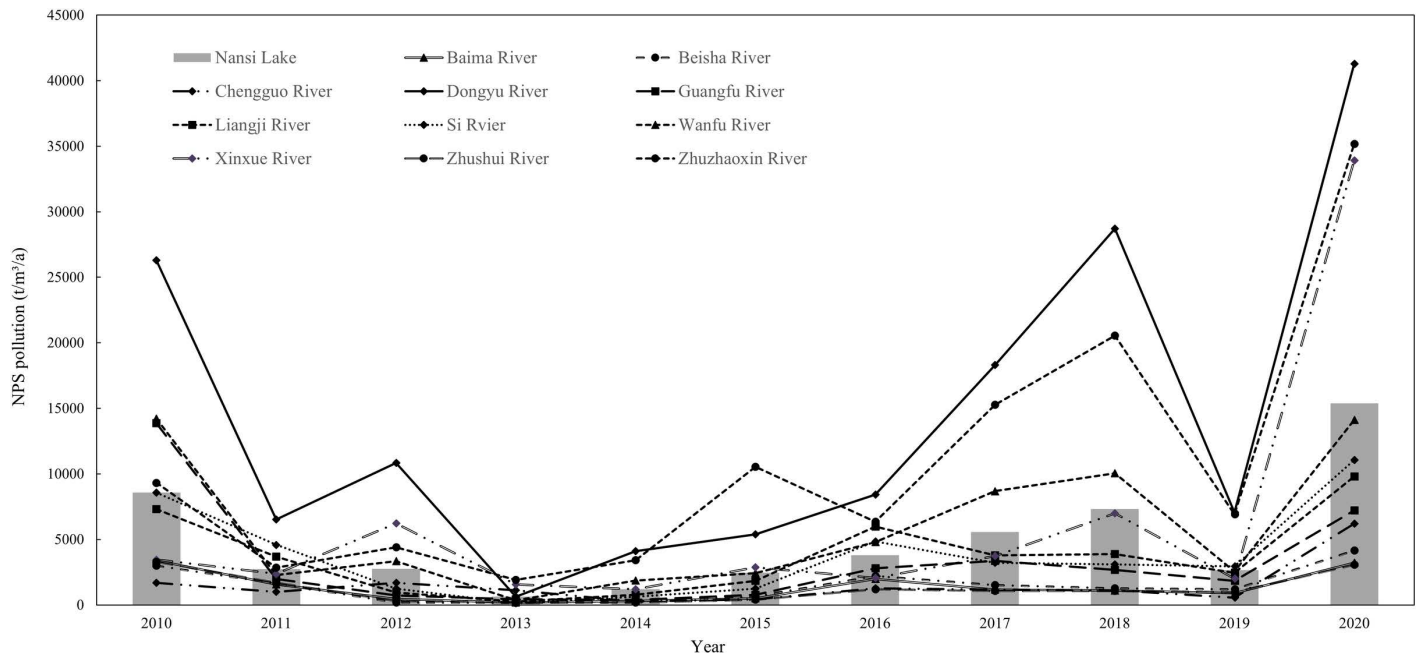


**Fig 1. Temporal changes in total runoff pollution in NLB, 2010–2020.**

<https://doi.org/10.1371/journal.pone.0318691.g001>

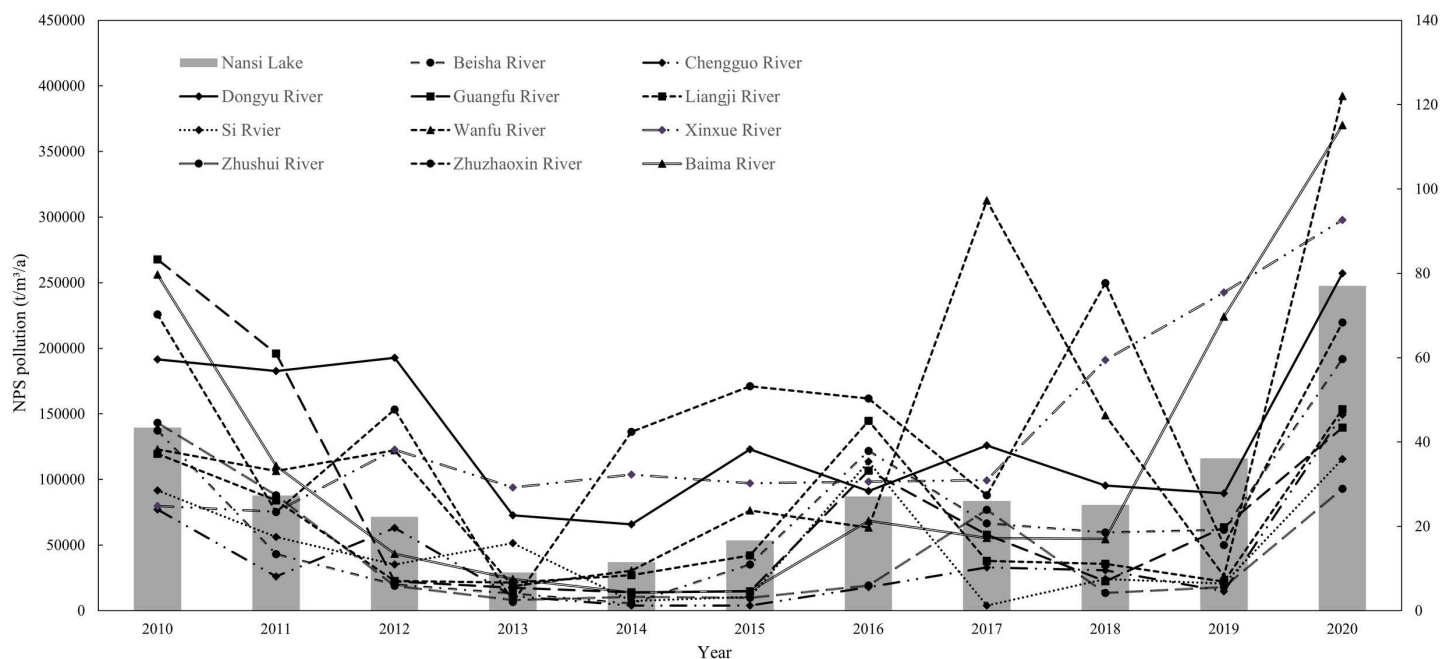
amount of COD pollution in the three major sub-basins in the west, namely, the Zhuzhaoxin River, the Wanfu River, and the Dongyu River, was always at the downstream level from 2010 to 2020, and the results of the pollution treatment were not significant, and further treatment and restoration were needed in the future. The four sub-basins of the Zhushui River, the Baima River, the Beisha River, and the Chengguo River were always at the upper-middle level in the process of the change of the amount of COD pollution during the period of 2010–2020. The Xinxue River basin fluctuated considerably during this 11-year period, improving from the upper-middle level in 2010 to the lower-middle level in 2020.

The time-series course of  $\text{NH}_3\text{-N}$  pollution loads in the NLB from 2010 to 2020 is shown in Fig 3 (Temporal changes in total runoff  $\text{NH}_3\text{-N}$  pollution in each sub-basin, 2010–2020).



**Fig 2.** Temporal changes in total runoff COD pollution in each sub-basin, 2010–2020.

<https://doi.org/10.1371/journal.pone.0318691.g002>



**Fig 3.** Temporal changes in total runoff  $\text{NH}_3\text{-N}$  pollution in each sub-basin, 2010–2020.

<https://doi.org/10.1371/journal.pone.0318691.g003>

The overall change in  $\text{NH}_3\text{-N}$  pollutant discharges was uneven over the study period. For each sub-basin, the Guangfu River Basin shifted from a downstream level in 2010 to an upstream level in 2018, and maintained to an upper-middle level in 2020. Improvement was more pronounced in the Zhushui River Basin, which gradually improved from the downstream level in

2010 and returned to the upstream level by 2020. The overall change in the Xinxue River Basin is not promising, being at the upstream level in 2010 and changing to the midstream level by 2012–2014. Although it recovered in 2016, it gradually receded to the downstream level after 2018, and further management of  $\text{NH}_3\text{-N}$  pollutants in the basin is needed.

### 3.3. Spatial variation of total NPS pollution in the Nansi Lake sub-basin

The spatial changes in the number of COD NPS pollution in the basin of the Nansi Lake from 2010 to 2020 are shown in Table 5. Overall, the NPS pollution in the central part of the basin is relatively good, and the spatial distribution of pollution is generally characterized as ‘low in the middle, high on both sides, and more concentrated’. The NPS pollution in the Zhushui River, Baima River, and Beisha River in the central part of the basin is always at a low value, while the COD pollution in the Zhuzhaoxin River, Wanfu River, and Dongyu River in the western part of the basin is always at the maximum value; the three basins with the largest COD NPS pollution in 2010 are Dongyu River, Wanfu River, and Zhuzhaoxin River in the order of magnitude. In 2020, the largest number of COD NPS pollution in NLB is still Dongyu River, which shows a significant increase compared with the amount of pollution in 2010. However, the basin with the smallest amount of COD NPS pollution in 2020 will still be the Zhushui River, the Baima River, and the Beisha River, and the amount of pollution will not fluctuate too much compared with that in 2010.

In the east-central part of the basin, the amount of  $\text{NH}_3\text{-N}$  pollution was generally good, and the basin with the most obvious change in the amount of  $\text{NH}_3\text{-N}$  pollution between 2010 and 2020 was the Xinxue River, which showed an increasing trend in the amount of  $\text{NH}_3\text{-N}$  pollution. The three sub-basins with the largest amount of  $\text{NH}_3\text{-N}$  NPS pollution in the NLB in 2010 were the Zhuzhaoxin River, the Dongyue River, and the Guangfu River in that order, while the three sub-basins with the smallest amount were the Xinxue River, the Chengguo River, and the Baima River in that order. The three sub-basins with the highest amount of  $\text{NH}_3\text{-N}$  NPS pollution in the NLB in 2010 were Xinxue River, Zhuzhaoxin River and Dongyu River, while the three sub-basins with the lowest amount of pollution were Xinxue River, Chengguo River and Baima River, in order.

From the sensitivity test of each key factor of each subsystem above, the rate of change of population, the rate of change of GDP, the rate of rural sewage treatment, the proportion of the area of soil application of fertilizer, the rate of change of the number of livestock and

**Table 5. Spatial variation in total runoff COD and  $\text{NH}_3\text{-N}$  pollution by sub-basin.**

Basin control unit	NPS COD pollution			NPS $\text{NH}_3\text{-N}$ pollution		
	2010	2015	2020	2010	2015	2020
Baima River	3390241.57	500245.39	3214485.21	79690.97	4662.64	115125.12
Beisha River	3001108.89	663579.59	4157904.82	137308.39	35088.30	191669.76
Chengguo River	1698922.51	431232.56	6200647.34	77136.95	3799.60	149261.23
Dongyu River	26294560.91	5376567.34	41274035.74	191442.92	122970.48	256984.81
Guangfu River	13863814.00	768786.23	7205661.35	267613.17	14666.19	139600.97
Liangji River	7308127.96	1844075.41	9809169.23	119615.14	42070.77	153641.17
Si River	8577944.53	1253949.17	11055870.01	91570.10	10499.30	115678.35
Wanfu River	14203462.48	2418552.14	14116911.26	122892.72	76462.04	392245.91
Xinxue River	3450505.51	2862605.52	33918028.73	79808.10	97099.20	897706.66
Zhushui River	3301641.40	417280.02	3060958.97	143188.85	9878.36	92868.60
Zhuzhaoxin River	9318503.39	10535105.37	35175967.05	225815.73	171097.01	219509.44

<https://doi.org/10.1371/journal.pone.0318691.t005>

poultry farming, and the rate of treatment of livestock and poultry farming on a large scale are the key and sensitive factors.

### 3.4. Analysis of factors affecting NPS pollution in the NLB

To further investigate the correlation between socio-economic factors and pollutants, this study analyzed the correlation between two pollutants,  $\text{NH}_3\text{-N}$  and COD, and the influencing factors from 2010 to 2020 using the Pearson's coefficient model of SPSS. Overall, both  $\text{NH}_3\text{-N}$  and COD pollutants were significantly correlated with PAA, FAA, YAP and YTP from 2010 to 2020, with significant positive correlations with PAA and FAA. This result indicates that as the amount of annual fertilizer and annual pesticide increases, the pollutant content also increases, and the excessive use of fertilizer and pesticide can cause some NPS pollution. Meanwhile, the 2 pollutants were almost significantly negatively correlated with PTR, indicating that the pollutant content tends to decrease with the growth of the proportion of tertiary industry. The correlation coefficients between the characteristic factors such as GDP, YAP and PD were not consistent due to the different degrees of NPS pollution of  $\text{NH}_3\text{-N}$  and COD pollutants in various site types. For COD pollutants, the main positive influencing factors of COD pollutants are YTP, FAA, RECA, LSI, and DA, with DA having the largest correlation coefficient of 0.92. The main negative influencing factors are NAPR, and SOMC. The main positive impact factors of  $\text{NH}_3\text{-N}$  pollutants were YTP, YAP, FAA, RECA, LSI, and DA, and the main negative impact factors were PPR (−0.340) and SOMC (−0.449).

COD is an important indicator for responding to organic matter in water bodies, and most of the influencing factors showed significant positive correlations with COD. Specifically, COD was consistently positively correlated with YTP and RPA, indicating that COD was more associated with demographic factors, showing a significant positive correlation. In 2010, the correlation coefficients between COD pollutants and RPA, YAP, PAA, and YTP characteristic factors were all positive, and all exceeded 0.5. The correlations of the factors that showed negative correlations in 2010 were all less than 0.11. By 2015, PSR and RECA were the most significantly changed influencing factors, which changed from a lower negative correlation to a positive correlation. In addition, the changes of PGDP, PSR and FAA gradually changed from a low correlation to a high correlation. The correlation between the influencing factors and COD pollutants changed less during the period of 2015–2020, which may be due to the fact that the socio-economic situation gradually changed from a period of rapid development to a period of stabilization, and the correlation between the factors has formed a more stable relationship (Fig 4 Correlation coefficients between socio-economic factors and COD pollutants).

During the period 2010–2020,  $\text{NH}_3\text{-N}$  pollutants are positively correlated with YTP and the correlation coefficients are large, indicating that the total population is an important factor affecting the emission of  $\text{NH}_3\text{-N}$  pollutants, in addition to this, the positive correlation between  $\text{NH}_3\text{-N}$  pollutants and RECA is more obvious, indicating that as the total amount of electricity consumption in the rural areas increases, the production of  $\text{NH}_3\text{-N}$  pollutants is also higher, and the electricity consumption in the rural areas has a negative impact on  $\text{NH}_3\text{-N}$  has an adverse effect, while the correlation coefficients of  $\text{NH}_3\text{-N}$  pollutants with PPR and NAPR are mostly negative, indicating that the increase in the proportion of non-agricultural population and the development of primary industry are favorable to  $\text{NH}_3\text{-N}$  emissions, which may be that the more developed the plantation and forestry industry is, the stronger the fixation of ammonia nitrogen is, which makes the pollutants in the air less. In 2010, the positive correlation coefficient between  $\text{NH}_3\text{-N}$  and YTP and RPA characteristic factors was the largest, reaching more than 0.7. In 2015, PGDP was significantly negatively correlated with  $\text{NH}_3\text{-N}$  (−0.4230), indicating that as GDP per capita increases, it leads to a decrease in  $\text{NH}_3\text{-N}$  emissions. In 2020, there was a negative correlation between  $\text{NH}_3\text{-N}$  pollutants and

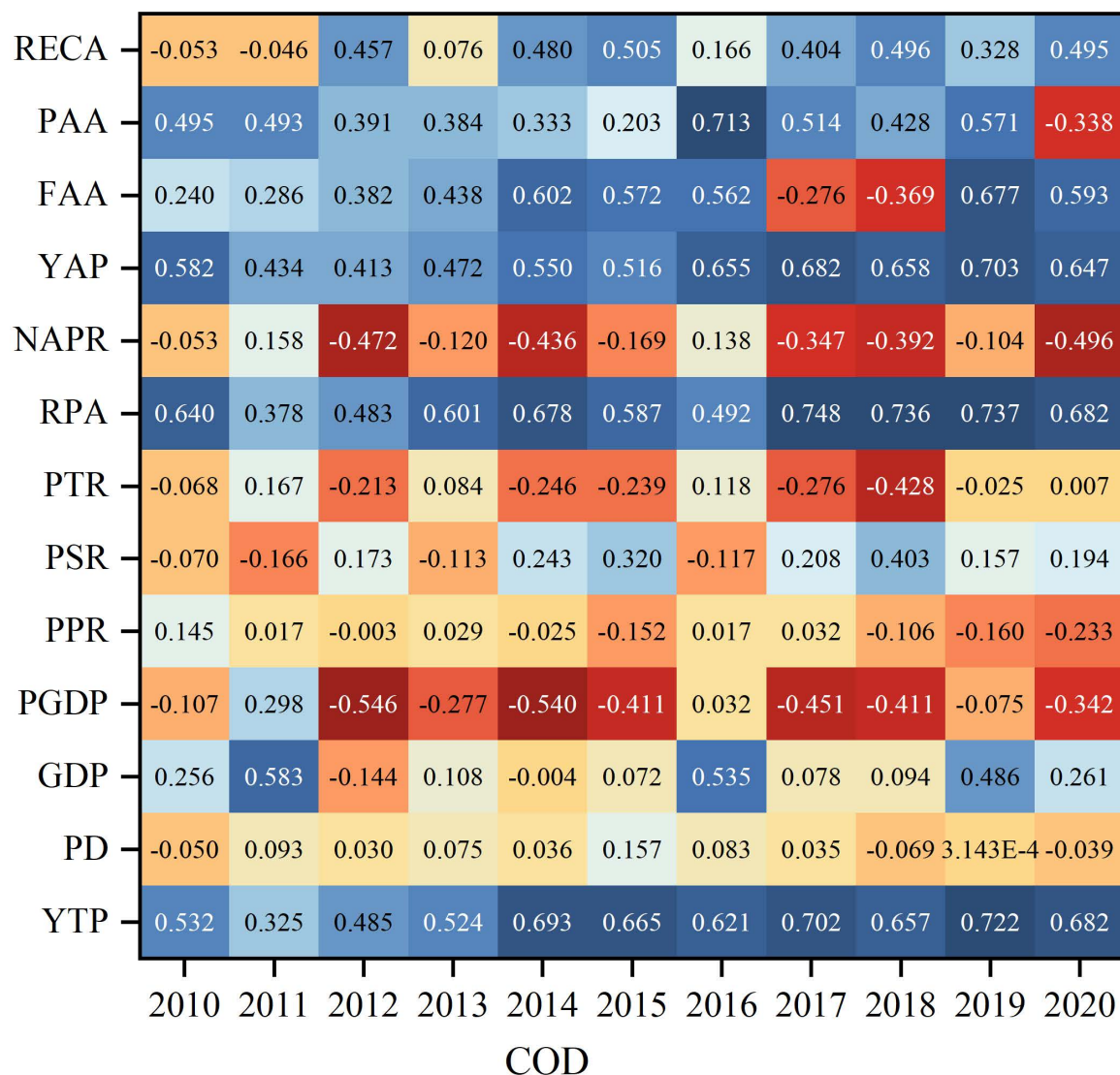


Fig 4. Correlation coefficients between socio-economic factors and COD pollutants.

<https://doi.org/10.1371/journal.pone.0318691.g004>

three characterization factors, NAPR, PPR and PGDP, and a positive correlation with all other characterization factors (Fig 5 Correlation coefficients between socio-economic factors and  $\text{NH}_3\text{-N}$  pollutants).

Most of the natural factors were positively correlated with COD, and only 5 out of 13 natural factors were negatively correlated with it. LSI had the most obvious positive correlation, with a Pearson coefficient as high as 0.5747, which is due to the fact that when the complexity of the local landscape increases, it increases the possibility of contact between the landforms and the water body and the possibility of pollutants flowing into the water body. PRE and PFL also have the highest positive correlation with COD, and their Pearson coefficients exceeded 0.5. PRE and PFL are also the natural factors with the highest positive correlation with COD, with Pearson's coefficients exceeding 0.5, as the scouring effect of precipitation and runoff transport lead to increased pollutant levels in water bodies (Fig 6 Correlation coefficients between natural factors and COD pollutants).



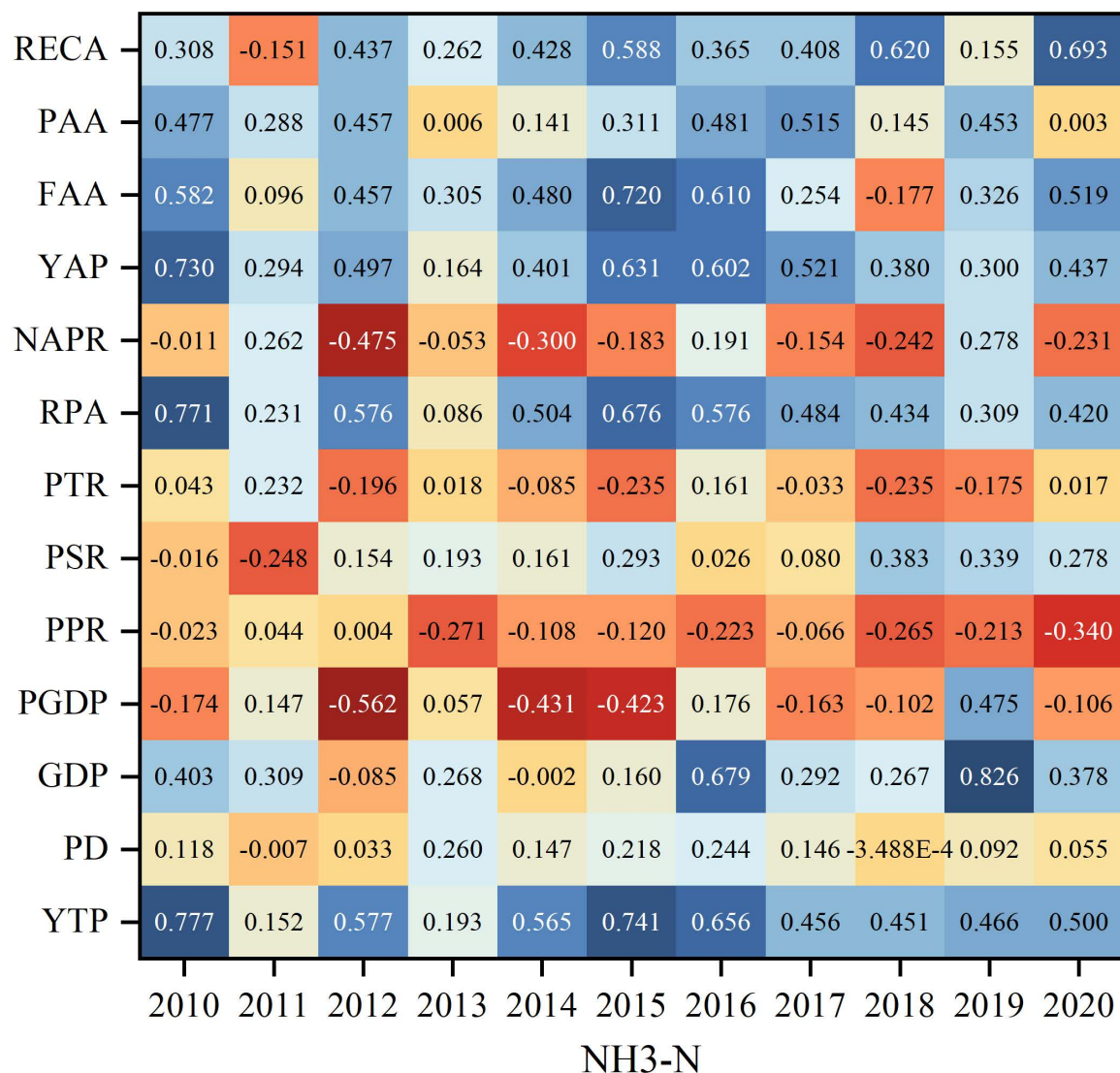


Fig 5. Correlation coefficients between socio-economic factors and NH<sub>3</sub>-N pollutants.

<https://doi.org/10.1371/journal.pone.0318691.g005>

Agricultural NPS pollutants, particularly NH<sub>3</sub>-N, show a mostly positive correlation with natural factors. Among these, NH<sub>3</sub>-N is significantly positively correlated with the natural factors LSI and DA, with correlation coefficients exceeding 0.5, reaching 0.7050 and 0.6630, respectively. Pollutant emissions are also closely related to DA, pollutant transport distance, and the retention time within the basin. The larger the area, the greater the pollutant retention and degradation, which in turn leads to increased NH<sub>3</sub>-N pollutant production. Furthermore, TD is another natural factor affecting pollutant discharge. A higher density of trenches results in organic matter, NH<sub>3</sub>-N, and other pollutants from farmland drainage being directly transported into nearby water bodies, leading to a higher subsequent generation of pollutants. NH<sub>3</sub>-N pollutants are significantly negatively correlated with natural factors such as SOMC and PGRL. Specifically, higher organic matter content in the soil is associated with lower NH<sub>3</sub>-N pollution levels. Additionally, as polar organic polymers increase, pollutant degradation becomes stronger, leading to reduced pollutant production. Moreover, a higher PGRL is associated with lower



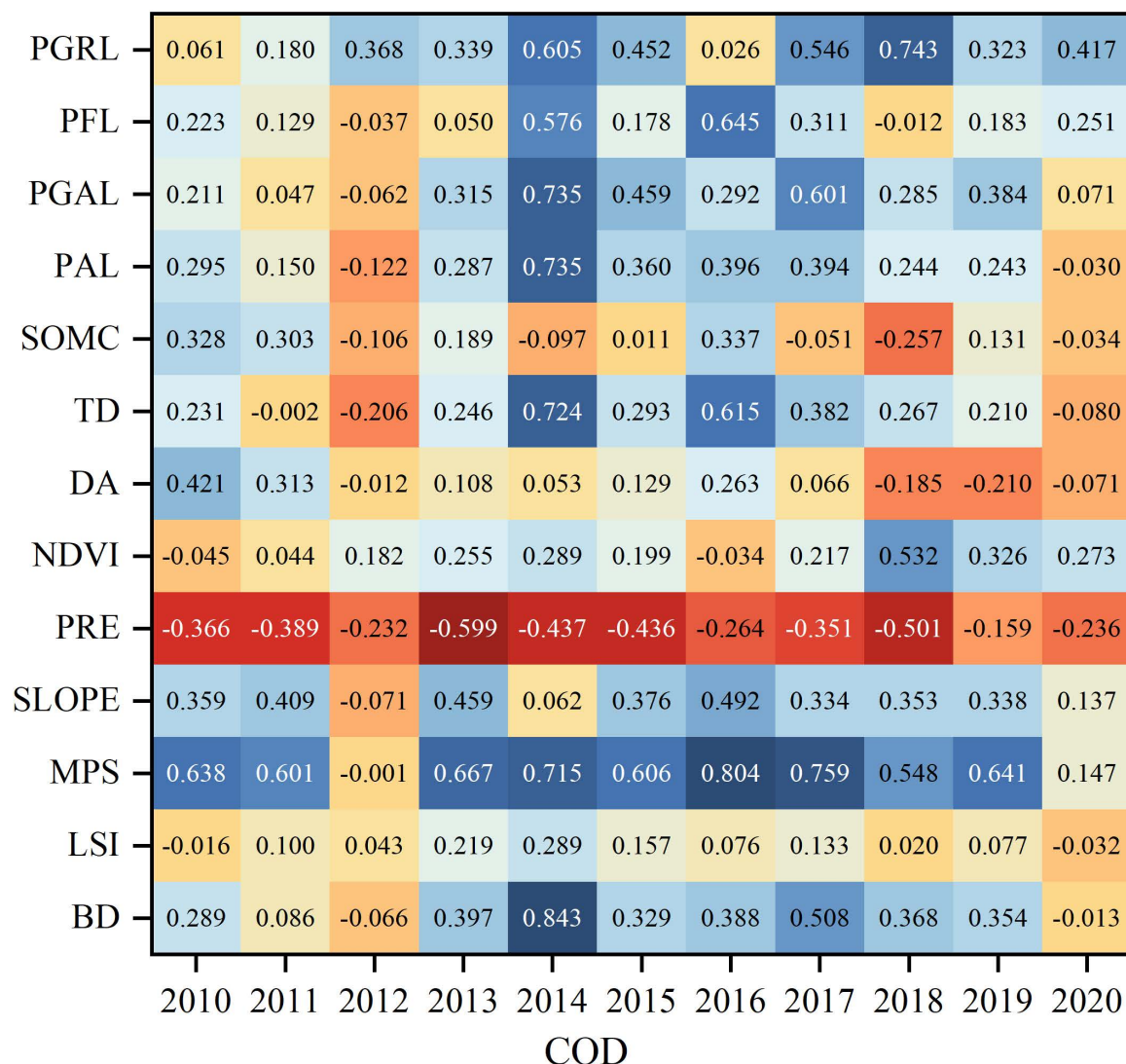


Fig 6. Correlation coefficients between natural factors and COD pollutants.

<https://doi.org/10.1371/journal.pone.0318691.g006>

NH<sub>3</sub>-N concentrations, indicating that grasslands have a stronger pollutant degradation effect (Fig 7 Correlation coefficients between natural factors and NH<sub>3</sub>-N pollutants).

### 3.5. Analysis of regression modelling

Using two regression models (Equations 6 and 7), eight basin characteristics were selected as independent variables, while COD and NH<sub>3</sub>-N NPS pollution amounts from 2010 to 2020 were chosen as the dependent variables. Prediction models were established, and multiple linear regression analyses were conducted, resulting in the regression equations shown in Table 6. After testing for significance, the individual regression equations passed the significance level test ( $P < 0.05$ ).

Models 1a and 1b were developed based on NH<sub>3</sub>-N and COD NPS pollution amounts in Equation 6, while models 2a and 2b were based on NH<sub>3</sub>-N and COD NPS pollution amounts in Equation 7 (Table 6).

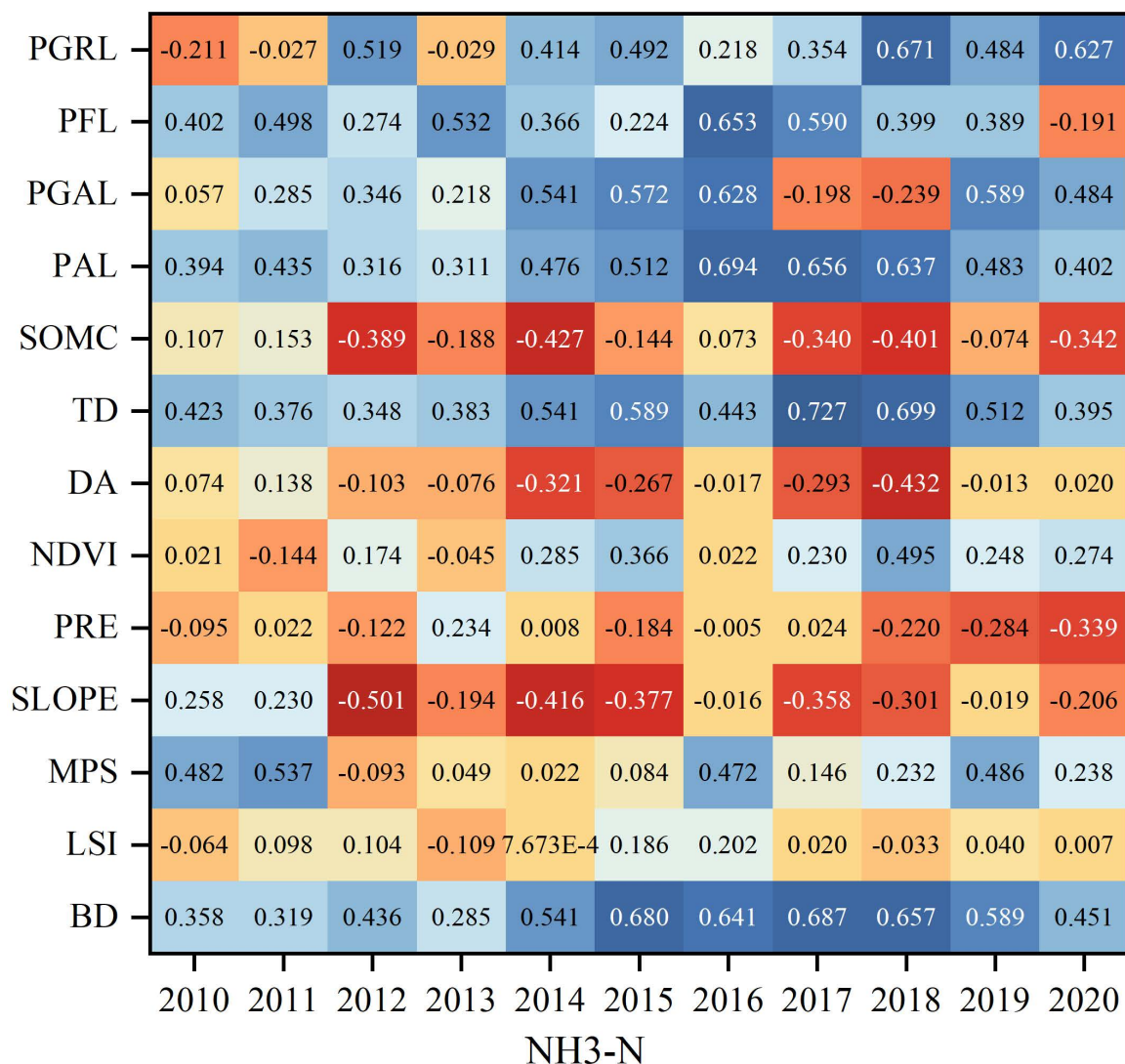


Fig 7. Correlation coefficients between natural factors and  $\text{NH}_3\text{-N}$  pollutants.

<https://doi.org/10.1371/journal.pone.0318691.g007>

Table 6. Regression equations for NPS pollutants.

Name	Equation	R <sup>2</sup>
$\text{NH}_3\text{-N}$ (1a)	$L_a = 1.070X_1 - 0.835X_8 - 0.138X_9 + 0.199X_{10} + 0.101X_{13} + 0.342X_{15} + 0.454X_{20} + 0.058X_{21} + 0.141X_{22}$	0.811
COD (1b)	$L_a = 0.572X_1 + 1.139X_8 - 0.074X_9 - 1.024X_{10} + 0.038X_{13} - 0.309X_{15} + 1.400X_{20} - 0.209X_{21} + 0.229X_{22}$	0.698
$\text{NH}_3\text{-N}$ (2a)	$\log(L_b) = 0.023\log(X_1) - 0.138\log(X_8) + 0.121\log(X_{10}) + 0.198\log(X_{13}) - 1.215\log(X_{15}) - 0.831\log(X_{20}) - 0.616\log(X_{21}) + 0.457\log(X_{22})$	0.763
COD (2b)	$\log(L_b) = 0.275\log(X_1) - 0.074\log(X_8) - 0.136\log(X_{10}) + 0.197\log(X_{13}) - 1.263\log(X_{15}) - 0.953\log(X_{20}) - 0.445\log(X_{21}) + 0.513\log(X_{22})$	0.821

<https://doi.org/10.1371/journal.pone.0318691.t006>

## 4. Discussion

### 4.1. Validation of model results

The results of NPS pollution division were tested by local government statistics and related studies, and the average of the total COD and  $\text{NH}_3\text{-N}$  pollution and NPS pollution share of the 11 control units were taken as the total pollution and NPS pollution share of the NLB. In terms of individual control units, the average NPS pollution amount of  $\text{NH}_3\text{-N}$  in Dongyu River was 135.19 t, with Congmin [41] 70.14 t, with an error of 48.12%, the  $\text{NH}_3\text{-N}$  pollution amount of Zhuzhaixin River was 139.74 t, with Xiao [42] 86.81 t, with an error of 37.88%, the average NPS pollution amount of  $\text{NH}_3\text{-N}$  in Xinxue River was 245.61 t, with Zhang et al. [43] 337.34 t, the error is 27.19%.

NLB simulation 2010–2020 average COD amount of 52668.24 t, with Wang et al. [44] 2005–2012 average NPS pollution amount of 78215.69 t, the error is 32.66%, with Ni et al. [42] 2007 COD NPS pollution reference 53442.47 t, the relative error is 1.45%, with Zhang et al. [43] 46803.15 t with a relative error of 11.14%, and with Wang [44] 64092.9 t with a relative error of 17.83%.

The simulated NPS pollution of COD and  $\text{NH}_3\text{-N}$  in Nansi Lake accounted for 78.77% and 70.44%, respectively, with Congmin [41] 67.90% and 60.29% observations with errors of 10.87% and 10.15%, respectively. The proportion of NPS pollution in Nansi Lake is 65.05% and the proportion of NPS pollution discharge from living agriculture in Corwin [45] is 67.22%, with an error of 2.17%; therefore, the reliability of the simulation results meets the requirements of the analysis.

### 4.2. Validation of multiple regression equations

The regression equations were tested for prediction by simulating the amount of NPS pollution in Dongyu River, Chengguo River, Baima River and Beisha River from 2010 to 2020. As shown in Tables 7 and 8, in general, RE values between 0–0.3,  $R^2$  values above 0.5, and ENS values between 0–1 indicate that the fitness test of the model was passed [43]. The test results showed that models 1a and 2a simulated  $\text{NH}_3\text{-N}$  NPS pollution in the Dongyu River with a large RE, and model 1a simulated the Baima River Basin with an RE also greater than 0.3 and an  $R^2$  less than 0.5, which is a large error, but in testing the rest of the control units. Models 1a and 2a had  $R^2$  values between 0.62–0.89 and 0.51–0.67, respectively, and ENS was also between 0 and 1, simulating well. Except for model 1b which simulated COD NPS pollution in Dongyu River with RE greater than 0.3 and  $R^2$  value of 0.48. In testing the rest of the control units, RE,  $R^2$ , and ENS passed the suitability test, which indicated that models 1a, 2a, and 2b were better suited to simulate the amount of NPS pollution in the NLB.

### 4.3. NPS pollution sources

Through the analysis of the  $\text{NH}_3\text{-N}$  emission accounting process, it was found that the average annual total  $\text{NH}_3\text{-N}$  discharged into the water environment from pollution sources in the

Table 7. Regression model 1a and 2a tests.

Basin	1a			2a		
	RE	$R^2$	ENS	RE	$R^2$	ENS
Dongyu River	−0.39	0.74	0.13	−0.34	0.62	0.89
Chengguo River	0.15	0.62	0.56	−0.26	0.51	0.22
Baima River	−0.47	0.81	0.06	−0.21	0.38	0.31
Beisha River	0.13	0.89	0.95	0.33	0.67	0.20

<https://doi.org/10.1371/journal.pone.0318691.t007>

Table 8. Regression model 1b and 2b tests.

Basin	1b			2b		
	RE	R <sup>2</sup>	ENS	RE	R <sup>2</sup>	ENS
Dongyu River	−0.75	0.48	0.05	−0.11	0.64	0.86
Chengguo River	−0.17	0.62	0.27	−0.01	0.75	0.85
Baima River	−0.02	0.50	0.04	0.02	0.55	0.72
Beisha River	0.25	0.73	0.69	0.01	0.63	0.87

<https://doi.org/10.1371/journal.pone.0318691.t008>

NLB from 2010 to 2020 was 1,468.26 t, while the average annual total COD discharge was 735,519.79 t. Therefore, COD is the primary pollutant in the NLB. Table 6 illustrate the spatial distribution of COD and NH<sub>3</sub>-N pollution across the sub-basins of Nansi Lake. The spatial variation of NPS pollutants across regions was significant, with the highest pollution levels found in the Dongyu River Basin, where land use is predominantly arable. Most of the COD and NH<sub>3</sub>-N entered the water through sediment adsorption, aligning with the results of previous studies [46,47]. The contribution rates of NH<sub>3</sub>-N and COD pollution from NPS sources flowing into rivers were calculated for different areas, showing that Heze had the highest total NH<sub>3</sub>-N pollution, while Tai'an had the lowest, consistent with Zhang et al. [32]. In the north-eastern part of the NLB, where land use is dominated by forests and grasslands, vegetation effectively retains water, resulting in lower pollutant flow [48]. Excess fertilizers, rural households, and dispersed livestock farms make up a significant portion of the basin [49].

#### 4.4. NPS pollution influences and basin planning

Exploring the similarities between anthropogenic and intrinsic basin factors (such as soil type, climate, topography, and hydrology) is critical. Zoning based on similar characteristics can be used to assess pollutant loss [50,51], evaluate NPS pollution risks [34,52], and implement effective management practices for NPS pollution control [53,54]. The 11 sub-basins show high levels of NH<sub>3</sub>-N and COD pollution with notable seasonality, consistent with previous findings [27]. Pearson's correlation analysis revealed a significant correlation between NH<sub>3</sub>-N and COD pollution and the PAL. NH<sub>3</sub>-N displayed a positive correlation in all studied years, while COD exhibited a negative correlation in 2012 and 2020 but was positively correlated in other years, differing from earlier studies [32]. This suggests that NH<sub>3</sub>-N and COD pollution is higher in cultivation-dominated basins, a result comparable to findings reported in the literature [55,56]. The correlation between NH<sub>3</sub>-N and COD pollution and the PFL was also significant, with a positive correlation in most years. Studies have shown that forests play an important role in reducing nitrogen loads [57]. Prolonged flooding processes, due to the hydrological buffering effects of forests, can reduce NH<sub>3</sub>-N export by slowing down runoff and extending the time for pollutant transport to streams. Seitzinger et al. [58] found that 37% to 76% of NH<sub>3</sub>-N input into streams is removed during transport, which contrasts with our findings. Moreover, the correlations between COD and NH<sub>3</sub>-N loads with factors such as YTP, RPA, LSI, and DA were significantly linear, confirming that the primary source of pollution stems from human activities, a conclusion that differs from previous studies [59–61].

#### 4.5. Study limitations and future directions

For the management of NPS pollution in the NLB, the L-THIA model was used to calculate the total amount of NPS pollution and to analyze the influencing factors of NPS pollutants, and finally a linear equation was established to fit a multiple regression model, with a view to balancing the good interaction between agricultural production and the ecological

environment, and to laying a good theoretical foundation for the sustainable development of the basin. However, the following limitations should be addressed in order to better provide new avenues for future research. The period of high  $\text{NH}_3\text{-N}$  loss mainly occurs in the rainy season (March ~ June) and the period of high  $\text{NH}_3\text{-N}$  enrichment mainly occurs in the dry season (January ~ February). This temporal difference in monthly variation shows a clear transport characteristic of NPS pollution, as pollutant transport in NPS occurs mainly in the rainy season [47], our analyses are mainly year-based and the effects of monthly variations in the rainy versus dry seasons should be considered later. In addition the effects of heavy rainfall processes on COD and  $\text{NH}_3\text{-N}$  [50] should also be considered in future studies.

## 5. Conclusions

Based on the runoff partitioning method, a data series of agricultural NPS pollution distribution in the basin was obtained to quantitatively assess the spatial and temporal distribution of agricultural NPS pollution in the basin and quantitatively analyze the natural and socio-economic influences affecting them. Finally, multiple regression analysis was conducted to construct regression equations and simulate the future agricultural NPS pollution. Agricultural pollution in the NLB takes 2014 as the dividing line, showing a trajectory change of decreasing and then increasing, and the magnitude of increasing and decreasing varies greatly. In terms of socio-economic influencing factors, both COD and  $\text{NH}_3\text{-N}$  were significantly correlated with PAA, FAA, YAP and YTP. In terms of natural influences, LSI, DA and TD had higher positive correlation correlations with various pollutants, and slope and SOMC had higher negative correlations with pollutants. The study found that agricultural surface pollution in the NLB mainly originates from livestock and poultry farming, so it is necessary to carry out a comprehensive promotion of the cleanliness of farming production and the ecology of the industrial model. The relevant departments should guide farm households to carry out appropriate scale farming, promote the creation and enhancement of standardized farming demonstrations, and lead the green and sustainable development of the livestock industry. The government should also increase its support for the animal husbandry and veterinary industry, and improve the construction of the animal disease prevention and control system. In addition, there is a need to optimize and adjust the agricultural structure to achieve a balance between industrial returns and the ecological environment. As a large agricultural province in Shandong Province, rich in agricultural resources, the degree and direction of agricultural development in various regions of the province are also different, so the adjustment of the agricultural structure needs to be fully considered to achieve a combination of regional characteristics and the actual development of a reasonable allocation of resource elements, optimize the layout of productivity according to local conditions.

## Supporting information

**S1 Data. Data of this study.**  
(XLSX)

## Author contributions

**Conceptualization:** Jiachen Liu, Yuan Tian.

**Data curation:** Jiachen Liu, Rongqiang Ma, Wenhui Xie, Xinyu Wang, Baolei Zhang.

**Formal analysis:** Jiachen Liu, Rongqiang Ma, Wenhui Xie, Xinyu Wang, Baolei Zhang.

**Funding acquisition:** Le Yin, Baolei Zhang.

**Investigation:** Jiachen Liu, Le Yin.



**Methodology:** Jiachen Liu, Wenhui Xie, Dongchao Wang, Luoan Yang, Xinyu Wang, Le Yin.

**Project administration:** Jiachen Liu.

**Software:** Wenhui Xie, Dongchao Wang, Luoan Yang.

**Supervision:** Luoan Yang.

**Writing – original draft:** Jiachen Liu, Rongqiang Ma, Dongchao Wang, Baolei Zhang.

**Writing – review & editing:** Jiachen Liu, Baolei Zhang.

## References

1. Jabbar FK, Grote K. Statistical assessment of nonpoint source pollution in agricultural watersheds in the Lower Grand River watershed, MO, USA. *Environ Sci Pollut Res*. 2018; 26(2): 1487–506. <https://doi.org/10.1007/s11356-018-3682-7> PMID: 30430446
2. Guo W, Fu Y, Ruan B, Ge H, Zhao N. Agricultural non-point source pollution in the Yongding River Basin. *Ecol Indic*. 2014;36:254–61. <https://doi.org/10.1016/j.ecolind.2013.07.012>
3. Xue J, Wang Q, Zhang M. A review of non-point source water pollution modeling for the urban–rural transitional areas of China: research status and prospect. *Sci Total Environ*. 2022;826:154146. <https://doi.org/10.1016/j.scitotenv.2022.154146>
4. Agency. Uniform Singapore energy price; 2013.
5. Kbgplselsma E. Diffuse nutrient losses in Denmark. *Water Sci Technol*. 1996;33(4–5):81–8.
6. Boers F. Spatial prepositions and metaphor: a cognitive semantic journey along the up-down and the front-back dimensions. Tübingen: Gunter Narr Verlag; 1996.
7. Edwin D, Zhang X, Tao Y. Current status of agricultural and rural non-point source pollution assessment in China. *Environ Pollut*. 2010;158:1159–68. <https://doi.org/10.1016/j.envpol.2009.10.047>
8. Rudra RP, Mekonnen BA, Shukla R, Shrestha NK, Goel PK, Daggupati P, et al. Currents status, challenges, and future directions in identifying critical source areas for non-point source pollution in Canadian conditions. *Agriculture*. 2020;10(10):468. <https://doi.org/10.3390/agriculture10100468>
9. Guan D, Zhao Z, Tan J. Dynamic simulation of land use change based on logistic-CA-Markov and WLC-CA-Markov models: a case study in three gorges reservoir area of Chongqing, China. *Environ Sci Pollut Res Int*. 2019;26(20):20669–88. <https://doi.org/10.1007/s11356-019-05127-9> PMID: 31104240
10. Ruan S, Zhuang Y, Hong S, Zhang L, Wang Z, Tang X, et al. Cooperative identification for critical periods and critical source areas of nonpoint source pollution in a typical watershed in China. *Environ Sci Pollut Res Int*. 2020;27(10):10472–83. <https://doi.org/10.1007/s11356-020-07630-w> PMID: 31939022
11. Beasley DB, Huggins LF, Monke EJ. ANSWERS: a model for watershed planning. *Am Soc Agric Eng*. 1980;23(4):938–44.
12. Knisel WG. CREAMS: a field-scale model for chemicals, runoff and erosion from agricultural management systems. *USDA Conserv Res Rep*. 1980;(26).
13. Lager JA, Metcalf EE, et al. Storm water management model. Environmental protection Agency; 1971.
14. Crawford NH, Linsley RK. Digital simulation in hydrology: Stanford watershed model IV. Stanford University Technical Report no.39. Palo Alto, CA: Stanford University; 1966.
15. Weng Q H. Modeling urban growth effects on surface runoff with the integration of remote sensing and GIS. *Environ Manag*. 2001;28(6):737–748.
16. Arnold JG, Allen M. Validation of automated methods for estimating base flow and groundwater recharge from stream flow records. *J Am Water Resour Assoc*. 1999;35(2):411–24.
17. Arnold JG, Srinivasan R, Muttiah RS, Williams JR. Large area hydrologic modeling and assessment Part I: model development1. *JAWRA J Am Water Resour Assoc*. 2007;34(1):73–89. <https://doi.org/10.1111/j.1752-1688.1998.tb05961.x>
18. Shen J, Parker A, Riverson J. A new approach for a Windows-based watershed modeling system based on a database-supporting architecture. *Environ Model Softw*. 2005;20(9):1127–38. <https://doi.org/10.1016/j.envsoft.2004.07.004>
19. Kim U, Kaluarachchi JJ. Application of parameter estimation and regionalization methodologies to ungauged basins of the Upper Blue Nile River Basin, Ethiopia. *J Hydrol*. 2008;362(1-2):39–56. <https://doi.org/10.1016/j.jhydrol.2008.08.016>
20. Harbor JM. A practical method for estimating the impact of land-use change on surface runoff, groundwater recharge and wetland hydrology. *J Am Plan Assoc*. 1994;60(1):95–108. <https://doi.org/10.1080/01944369408975555>



21. Liu Y, Chaubey I, Bowling LC, Bralts VF, Engel BA. Sensitivity and uncertainty analysis of the L-THIA-LID 2.1 Model. *Water Resour Manage*. 2016;30(13):4927–49. <https://doi.org/10.1007/s11269-016-1462-z>
22. Shen T, Liu LY, Wang JH, Li CJ. Spatial distribution characteristics of non-point pollution of Miyun reservoir areas based on L-THIA model. *Trans Chin Soc Agric Eng*. 2007.
23. Liu YM, Li LM, Liang LQ, Zeng HW. Effects of Landuse change on non-point source pollution in upper and middle reaches of Taoer River Basin. *Bull Soil Water Conserv*. 2011:117–21.
24. Zhang JH, Shen T, Liu MH, Wan Y, Liu J, Li J. Research on non-point source pollution spatial distribution of Qingdao based on L-THIA model. *Math Comput Model*. 2011;54:1151–9.
25. Bai FJ, Li TH. GIS and L-THIA based analysis on variations of non-point pollution in the Guanlan River Watershed, Shenzhen. *Environ Sci*. 2012;08:2667–73.
26. Li K, Zeng FT, Fang HY, Lin S. Analysis on nitrogen and phosphorus loading of non-point sources in Shiqiao River Watershed based on L-THIA Model. *Environ Sci*. 2013;11:4218–25.
27. Li Y, Jiao J, Wang Y, Yang W, Meng C, Li B, et al. Characteristics of nitrogen loading and its influencing factors in several typical agricultural watersheds of subtropical China. *Environ Sci Pollut Res*. 2014;22(3):1831–40. <https://doi.org/10.1007/s11356-014-3446-y> PMID: 25172462
28. Lehning DW, Corradini KJ, Petersen GW, Nizeyimana E, Hamlett JM, Robillard PD, et al. A comprehensive GIS-based modeling approach for predicting nutrient loads in watersheds. *J Spat Hydrol*. 2002;2(2).
29. Wang A, Yang D, Tang L. Spatiotemporal variation in nitrogen loads and their impacts on river water quality in the upper Yangtze River basin. *J Hydrol*. 2020;590:125487. <https://doi.org/10.1016/j.jhydrol.2020.125487>
30. Rozos E, Dimitriadis P, Bellos V. Machine learning in assessing the performance of hydrological models. *Hydrology*. 2021;9(1):5. <https://doi.org/10.3390/hydrology9010005>
31. Ouyang W, Huang H, Hao F, Shan Y, Guo B. Evaluating spatial interaction of soil property with non-point source pollution at watershed scale: the phosphorus indicator in Northeast China. *Sci Total Environ*. 2012;432:412–21. <https://doi.org/10.1016/j.scitotenv.2012.06.017>
32. Zhang B, Cui B, Zhang S, Wu Q, Yao L. Source apportionment of nitrogen and phosphorus from non-point source pollution in Nansi Lake Basin, China. *Environ Sci Pollut Res Int*. 2018;25(19):19101–13. <https://doi.org/10.1007/s11356-018-1956-8> PMID: 29725920
33. Islam F, Lian Q, Ahmad ZU, Zappi ME, Yao L, Gang DD. Nonpoint source pollution. *Water Environ Res*. 2018;90(10):1872–98. <https://doi.org/10.2175/106143017X15131012188033> PMID: 30180923
34. Ouyang W, Wu Y, Hao Z, Zhang Q, Bu Q, Gao X. Combined impacts of land use and soil property changes on soil erosion in a mollisol area under long-term agricultural development. *Sci Total Environ*. 2018;613–614:798–809. <https://doi.org/10.1016/j.scitotenv.2017.09.173> PMID: 28950672
35. Huang J, Lin X, Wang J, Wang H. The precipitation driven correlation based mapping method (PCM) for identifying the critical source areas of non-point source pollution. *J Hydrol*. 2015;524:100–10. <https://doi.org/10.1016/j.jhydrol.2015.02.011>
36. Zhong S, Chen F, Xie D, Shao J, Yong Y, Zhang S, et al. A three-dimensional and multi-source integrated technology system for controlling rural non-point source pollution in the Three Gorges Reservoir Area, China. *J Clean Prod*. 2020;272:122579. <https://doi.org/10.1016/j.jclepro.2020.122579>
37. Hydrology national engineering handbook. United States; 2019.
38. Guo H, Zhu J, Yang Y. Research status and development of technologies for controlling agricultural non-point source pollution. *Trans Chin Soc Agric Eng*. 2008;24(4):290–5.
39. Breusch TS. Testing for autocorrelation in dynamic linear models. *Australian Econ Pap*. 1978;17(31):334–55. <https://doi.org/10.1111/j.1467-8454.1978.tb00635.x>
40. Nash JE, Sutcliffe JV. River flow forecasting through conceptual models. Part 1: a discussion of principles. *J Hydrol*. 1970;10(3):282–90. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
41. Congmin G. A thesis submitted to the university of Jinan in partial fulfillment of the requirements for the Degree of Master of Engineering. Jinan University; 2015.
42. Xiao N. Study on total amount control of water pollutants and water quality comprehensive improvement projects in Nansi Lake Basin. Shandong Normal University; 2013.
43. Zhang B, Yin L, Zhang S, et al. GIS and L-THIA Based Analysis on Variations of Non-point Pollution in Nansi Lake Basin, China. *J Donghua Univ (Engl ed)*. 2016;33(6):851–8.
44. Jiarong W. Research on water quality and quantity response and Multi-drive factor evaluation of Nansi Lake Basin in Shandong Province. Shandong University; 2017.
45. Corwin LD. Non-point pollution modeling based on GIS. *Soil Water Conserv*. 1998;(1):75–88.

46. Shen J, Li Y, Liu X, Luo X, Tang H, Zhang Y, et al. Atmospheric dry and wet nitrogen deposition on three contrasting land use types of an agricultural catchment in subtropical central China. *Atmos Environ*. 2013;67:415–24. <https://doi.org/10.1016/j.atmosenv.2012.10.068>
47. Wang W, Ju T, Dong W, Liu X, Yang C, Wang Y, et al. Analysis of nonpoint source pollution and water environmental quality variation trends in the Nansi Lake Basin from 2002 to 2012. *J Chem*. 2015;2015:1–11. <https://doi.org/10.1155/2015/967165>
48. Zhu Q, Sun J, Hua G, Wang JH, Wang H. Runoff characteristics and non-point source pollution analysis in the Taihu Lake Basin: a case study of the town of Xueyan, China. *Environ Sci Pollut Res*. 2015;22(19):15029–36. <https://doi.org/10.1007/s11356-015-4709-y>
49. Wang S, Wang Y, Wang Y, Wang Z. Assessment of influencing factors on non-point source pollution critical source areas in an agricultural watershed. *Ecol Indic*. 2022;141:109084. <https://doi.org/10.1016/j.ecolind.2022.109084>
50. Cheng X, Chen L, Sun R, Jing Y. An improved export coefficient model to estimate non-point source phosphorus pollution risks under complex precipitation and terrain conditions. *Environ Sci Pollut Res Int*. 2018;25(21):20946–55. <https://doi.org/10.1007/s11356-018-2191-z> PMID: 29766427
51. Zou L, Liu Y, Wang Y, Hu X. Assessment and analysis of agricultural non-point source pollution loads in China: 1978–2017. *J Environ Manage*. 2020;263:110400. <https://doi.org/10.1016/j.jenvman.2020.110400>
52. Li Y, Wang H, Deng Y, Liang D, Li Y, Gu Q. Applying water environment capacity to assess the non-point source pollution risks in watersheds. *Water Res*. 2023;240:120092. <https://doi.org/10.1016/j.watres.2023.120092>
53. Hui D, Kuai P, Li W, Liu N. Evaluating the effects of land use planning for non-point source pollution based on a system dynamics approach in China. *PLoS One*. 2015;10(8). <https://doi.org/10.1371/journal.pone.0135572>
54. Liu R, Xu F, Zhang P, Yu W, Men C. Identifying non-point source critical source areas based on multi-factors at a basin scale with SWAT. *J Hydrol*. 2016;533:379–88. <https://doi.org/10.1016/j.jhydrol.2015.12.024>
55. Griffin JR. Introducing NPS water pollution. *EPA J*. 1991;6–9.
56. Wang W, Chen L, Shen Z. Dynamic export coefficient model for evaluating the effects of environmental changes on non-point source pollution. *Sci Total Environ*. 2020;747:141164. <https://doi.org/10.1016/j.scitotenv.2020.141164> PMID: 32771783
57. Wu J, Lu J. Landscape patterns regulate non-point source nutrient pollution in an agricultural watershed. *Sci Total Environ*. 2019;669:377–88. <https://doi.org/10.1016/j.scitotenv.2019.03.014> PMID: 30884262
58. Seitzinger SP, Styles RV, Boyer EW, et al. Nitrogen retention in rivers: model development and application to watersheds in the northeastern U.S.A. *Biogeochemistry*. 2002;57(58):199–237.
59. Johnes PJ. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. *J Hydrol*. 1996;183(3-4):323–49. [https://doi.org/10.1016/0022-1694\(95\)02951-6](https://doi.org/10.1016/0022-1694(95)02951-6)
60. Li C, Wang Y, Ye C, Wei W, Zheng B, Xu B. A proposed delineation method for lake buffer zones in watersheds dominated by non-point source pollution. *Sci Total Environ*. 2019;660:32–9. <https://doi.org/10.1016/j.scitotenv.2018.12.468> PMID: 30639716
61. Harrison JA, Maranger RJ, Alexander RB, Giblin AE, Jacinthe P-A, Mayorga E, et al. The regional and global significance of nitrogen removal in lakes and reservoirs. *Biogeochemistry*. 2008;93(1-2):143–57. <https://doi.org/10.1007/s10533-008-9272-x>