https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

Decoupling factors influencing spatial and temporal trends of total nitrogen and ammonia nitrogen levels in lakes across China

Chenyi Shi[1], Xihua Wang[1,2*], Y. Jun Xu[3], Chaomeng Dai[1], Nianqing Zhou[1], Rongbing Fu[4], Chengming Luo[1], Boyang Mao[1], Shunqing Jia[1], Qinya Lv[1], Zejun Liu[1], Xuming Ji[1], Yan Dai[1], Yanxin Rong[1]

[1] College of Civil Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, China

[2] Department of Earth and Environmental Sciences, University of Waterloo, ON N2L 3G1, Canada

[3] School of Renewable Natural Resources, Louisiana State University Agricultural Center, Baton Rouge, Louisiana, USA

[4] College of Environmental Science and Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, China

*Corresponding author at: College of Civil Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, China.

Email: 21531@tongji.edu.cn (Xihua Wang)

Tel.: + 0086 0431 13089410676; Fax: + 0086 021 65986809

Abstract

The heterogeneity of lake nitrogen pollution in its spatial distribution and long-term evolution influences ecosystem functioning and the effectiveness of water environment management. However, whether the drivers shaping the spatial patterns of different nitrogen forms differ from those governing their temporal dynamics remains

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

poorly understood at large spatial scales. Here, we constructed monthly time series (2010-2024) of total nitrogen (TN) and ammonia nitrogen (NH_3-N) concentrations for 3,020 lakes across China's five limnological regions using an ensemble machine learning framework and to reveal the difference in factors that influence spatial patterns and long-term changes of the two nitrogen forms at the large scale. The results showed that the factors determining where lakes are nitrogen-enriched are not necessarily those controlling whether nitrogen conditions improve or deteriorate over time and that different nitrogen forms respond to management and environmental change through distinct pathways. For TN, the spatial patterns are jointly shaped by external nitrogen loading, land cover, and hydro-climatic conditions, highlighting strong landscape-scale controls. However, its long-term evolution is governed more strongly by temporal variations in external nitrogen inputs than by static spatial characteristics. In contrast, NH_3-N exhibits a more direct and rapid response to external emission reduction measures, with both its spatial and temporal dynamics demonstrating high sensitivity to changes in anthropogenic nitrogen inputs. By highlighting the decoupled controls on spatial patterns and temporal trends, this study underscores the necessity of shifting lake management from static, location-based regulation toward an integrated 'state-rate' adaptive strategy.

Keywords

Nitrogen dynamics; TN and NH_3-N; Lakes; Long-term trends; Spatial heterogeneity

1. Introduction

Over the past century, human activities have elevated nitrogen to terrestrial and

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

aquatic systems by two to four times that of natural processes, including fertilizer application, urban wastewater discharge, and atmospheric deposition (Green et al., 2004; Fowler et al., 2013). As critical nodes for nitrogen retention and transformation in terrestrial freshwater ecosystems, lakes are highly sensitive to external nitrogen inputs and internal cycling processes (Monsen et al., 2002). Excessive nitrogen inputs have become one of the primary drivers of lake eutrophication, algal blooms, and ecosystem functional degradation (Wang et al., 2024).

Dissolved nitrogen in aquatic systems exists in different forms, which differ substantially in their sources, transport pathways, and ecological effects. TN integrates dissolved and particulate, inorganic and organic nitrogen in lake systems, thereby better reflecting the long-term accumulation and retention of nitrogen at the watershed scale. NH_3-N, as a key component of reactive nitrogen, is characterized by short transformation pathways and high bioavailability, and its concentration dynamics are closely linked to processes such as dissolved oxygen consumption and nutrient release at the sediment water interface (Wetzel, 2001; Beutel, 2006). The intensity of human activities and land-use patterns jointly regulate the dynamic transformation of nitrogen forms through complex interactions. Climate change significantly alters nitrogen transport and transformation pathways in lakes by modifying temperature gradients and the spatiotemporal distribution of precipitation, thereby increasing the uncertainty in the evolution of different nitrogen forms (Adrian et al., 2009; Causse et al., 2015). Previous studies have documented the differential responses of nitrogen forms to climatic, hydrological, and land-use changes at lake or watershed scales (Du et al., 2023;

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

Jiang et al., 2026). However, at large spatial scales, a unified understanding of whether the driving mechanisms of lake TN and NH_3-N diverge in space and time is still lacking.

While the spatial heterogeneity of human activities and their resulting pollutant loads are well-documented at broad scales (Sadayappan et al., 2022; Wang et al., 2025c), the specific drivers of these patterns vary significantly by region. The high nitrogen (N) of surface water in Northeast China and the Huang-Huai-Hai Plain are largely tied to intensive livestock and fertilizer use (Ma et al., 2021). In the transition zone between the Huang-Huai-Hai Plain and the Yangtze River Basin, rapid population growth, urban expansion, and high-intensity intensive agriculture combine to make domestic sewage and industrial wastewater significant sources of lake nutrients (Zhu et al., 2019). Regional precipitation and runoff patterns further amplify spatial differences in nitrogen concentrations among lakes by modulating the transport efficiency of terrestrial nitrogen (Ji et al., 2025; Wang et al., 2025a). Studies have also recognized that the temporal evolution of lake water quality exhibits significant regional differences. Hu et al. found that changes in the trophic status of eastern lakes in China are mainly driven by human activities, whereas western lakes more strongly reflect responses to climate change(Hu et al., 2024).

Recent advances in ecological assessment emphasize that a comprehensive understanding of lake ecosystem health requires the integration of both ecological "states" and "rates" of change; this "state-rate" framework highlights that current pollution levels and their respective recovery or deterioration speeds may be governed by distinct mechanisms, and overlooking either dimension could lead to biases in

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

setting restoration priorities (Zhang et al., 2026). In recent years, a series of adjustments in China regarding fertilizer application patterns, atmospheric nitrogen emissions, and urban wastewater treatment have led to temporal changes in nitrogen input intensity (Gao et al., 2020; Qu et al., 2022; Xu et al., 2025). Despite these shifts, the spatial patterns (states) and long-term temporal dynamics (rates) of lake nitrogen concentrations are often investigated in isolation. At large spatial scales, there remains a lack of robust research into whether the environmental factors explaining the spatial distribution of high nitrogen concentrations also govern their temporal evolution, particularly across different nitrogen forms. This knowledge gap limits the ability to distinguish static landscape controls from dynamic process-driven mechanisms, thereby hindering the application of integrated "state-rate" management strategies for lake nitrogen pollution at a national scale.

Chinese lakes span environmental gradients ranging from near-natural lakes on the Tibetan Plateau to highly eutrophic urban lakes in the eastern plains, and have experienced multi-stage transitions in pollution intensity and water environmental management over recent decades. Thereby integrating the major pathways of lake nitrogen evolution within a single national context and serving as an important microcosm for understanding global lake nitrogen pollution processes. This highly compressed spatiotemporal heterogeneity makes it difficult for analyses based on single catchments to systematically reveal the divergent drivers of different nitrogen forms across spatial states and long-term evolutionary processes. Accordingly, this study integrates multi-source environmental data to develop an ensemble machine learning

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

framework that simultaneously characterizes concentration levels and trends of TN and NH_3-N in Chinese lakes, and quantitatively disentangles the regulatory effects of multiple environmental factors on different nitrogen forms. This study focuses on the following key questions: (1) At a large geographical scale, do lake TN and NH_3-N exhibit consistent patterns or divergence in their spatial distributions, temporal evolution characteristics, and dominant driving factors? (2) Are the spatial states and temporal evolutionary processes of lake nitrogen concentrations governed by the same environmental factors, and (3) does the divergence in these driving mechanisms exhibit consistent regional patterns?

2. Data and methods

2.1 Study area

This study reconstructed monthly concentrations of total nitrogen (TN) and ammonia nitrogen (NH_3-N) for 3,020 lakes across in the past 15 years (2010-2024) across China in five limnological regions (Fig. 1a). These regions included Inner Mongolia-Xinjiang Lakes (IMXL), Qinghai-Tibet Plateau Lakes (TPL), Yunnan-Guizhou Plateau Lakes (YGPL), Eastern Plain Lakes (EPL), and Northeast Plain-Mountain Lakes (NPML). These regions differ significantly in natural environment and socioeconomic development, and their lakes exhibit distinct physical and chemical characteristics (Ma et al., 2010). The study lakes are in size of greater than 2 km^2 and have the following distribution: 483, 690, 153, 1, 206, and 488 in the IMXL, TPL, YGPL, EPL, and NPML regions, respectively. The majority of these lakes is between 2 and 10 km^2 accounting for 69.73% of the total number of lakes (Fig. 1b).

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

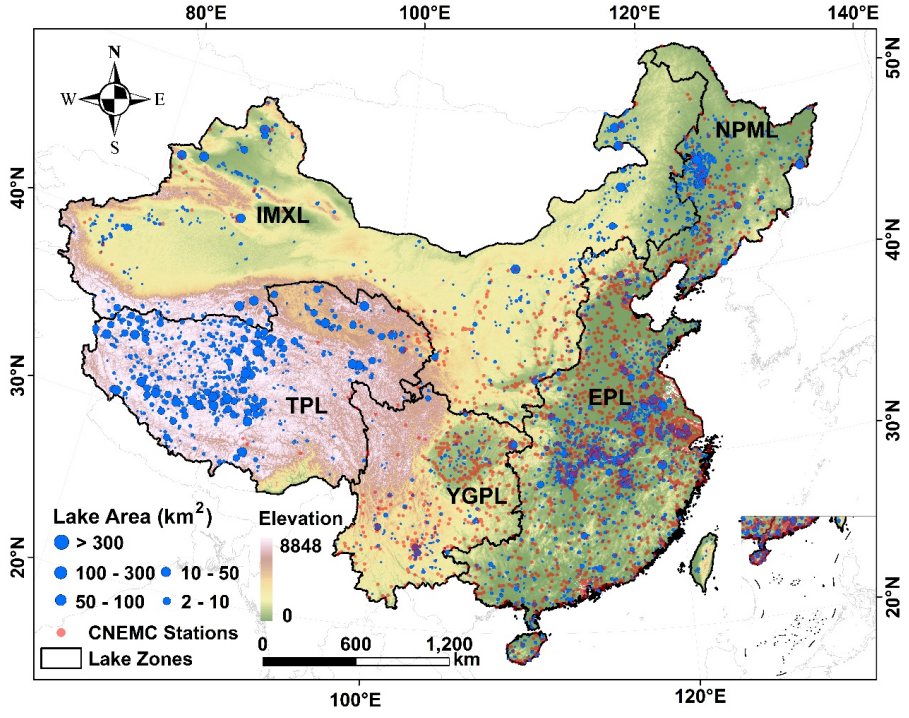

Fig. 1 Spatial distribution of the study lakes and the locations of CNEMC surface water monitoring stations across five limnological zones in China, including Inner Mongolia-Xinjiang Lakes (IMXL), Qinghai-Tibet Plateau Lakes (TPL), Yunnan-Guizhou Plateau Lakes (YGPL), Eastern Plain Lakes (EPL), and Northeast Plain-Mountain Lakes (NPML)

2.2 Datasets

2.2.1 Surface water monitoring data

Water quality records of the study lakes were obtained from the national automatic surface water quality monitoring dataset released by the China National Environmental Monitoring Center (CNEMC, http://www.cnemc.cn/). The dataset covered more than 2,300 automatic water monitoring stations nationwide, spanning from June 2021 to February 2025, with a monitoring frequency of once every 4 hours. The water quality parameters included water temperature (WT, °C), pH, dissolved oxygen (mg/L), electrical conductivity (μS/cm), turbidity (NTU), permanganate index (mg/L),

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

ammonia nitrogen (mg/L), total nitrogen (mg/L), total phosphorus (mg/L), chlorophyll-a (mg/L), and algal density (cells/L).

2.2.2 Environmental driving factors

Lake locations and boundary information were gathered from the China Lake Dataset (1960s–2020) (Zhang et al., 2019) and the HydroLAKES subset of the HydroSHEDS database (Lehner et al., 2022). Watershed boundaries associated with each lake were derived from Level-12 sub-basin vector data in the HydroBASINS dataset (Messager et al., 2016). Through spatial overlay analysis, one or more sub-basins corresponding to each lake were identified, and mean values of environmental driving factors were calculated at the lake sub-basin scale.

To generate a temporally continuous and spatially consistent dataset of monthly mean total nitrogen (TN) and ammonia nitrogen (NH_3-N) concentrations across Chinese lakes from 2010 to 2024, this study compiled an environmental driver dataset incorporating multidimensional information on climate conditions, soil properties, human activity intensity, and nitrogen inputs (Table 1). Monthly averages of precipitation, evaporation, total runoff, surface runoff, subsurface runoff, air temperature, and soil temperature and volumetric water across multiple depth layers from 0 to 289 cm were obtained from the ECMWF Reanalysis v5 (ERA5, https://www.ecmwf.int/) (Muñoz-Sabater et al., 2021). Soil property data were derived from the China Soil Dataset v2 (CSDLv2), including soil alkali nitrogen, total nitrogen content, and porosity across soil depths from 0 to 200 cm (Shi et al., 2025). Land-use data were obtained from the China Land Cover Dataset (CLCD, 1990–2024) (Yang and

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

Huang, 2021), including cropland, forest, grassland, water bodies, impervious surfaces, wetlands, and shrubland. NDVI and LAI were derived from the MODIS MOD13A1 and MOD15A2H products, respectively, and monthly datasets were generated using the maximum value compositing (MVC) method on the Google Earth Engine (Zhou et al., 2015). Annual nitrogen input data were obtained from the gridded nitrogen input dataset HaNi (1860–2019) including synthetic fertilizer application, livestock manure application, and atmospheric nitrogen deposition across cropland, pasture, and grassland systems (Tian et al., 2022), with values for 2020–2024 extended using linear regression extrapolation based on the trends. Population density data were sourced from the LandScan Global dataset (2010–2024) developed by the U.S. Oak Ridge National Laboratory (ORNL) (Lebakula et al., 2025). Topographic data were derived from the SRTM 30 m digital elevation model, from which slope information was calculated. River network data were obtained from the HydroRIVERS dataset (www.hydrosheds.org), and river density metrics were subsequently calculated at the sub-basin scale.

Table 1 Sources of the data used in this study.

Category	Variable	Description	Unit	Time
External N Loading	NH_x-N Deposition	Atmospheric NHx-N deposition	g m^{-2} $year^{-1}$	Annual
	NO_y-N Deposition	Atmospheric NOy-N deposition		
	NH_4-FC	NH_4-N fertilizer applied to cropland		
	NO_3-FC	NO_3-N fertilizer applied to cropland		
	NH_4-FP	NH_4-N fertilizer applied to pasture		
	NO_3-FP	NO_3-N fertilizer applied to pasture		
	MNAC	Manure N application on cropland		

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

	MNAP	Manure N application on pasture		
	MNDP	Manure N deposition on pasture		
	MNDR	Manure N deposition on rangeland		
	Population Density	Population density	person km^{-2}	
Hydro-climatic Dynamics	Precipitation	Monthly total precipitation	mm	Monthly
	Evaporation	Monthly total evaporation	mm	
	Total Runoff	Monthly total runoff	m	
	Surface Runoff	Monthly surface runoff	m	
	Subsurface Runoff	Monthly subsurface runoff	m	
	Air Temperature	Monthly mean air temperature at 2 m	K	
	STL	Soil temperature (0-289 cm)	K	
	SWV	Volumetric soil water (0-289 cm)	m^3 m^{-3}	
Soil N and Land Cover	SAN	Soil alkali nitrogen content (0-200cm)	mg/kg	Constant
	STN	Soil total nitrogen content (0-200cm)	g/100g	
	LAI	Leaf area index	Dimensionless	Monthly
	NDVI	Normalized Difference Vegetation Index		
	LU-Cropland	Proportion of cropland area	%	Annual
	LU-Forest	Proportion of forest area		
	LU-Grassland	Proportion of grassland area		
	LU-Water	Proportion of water area		
	LU-Impervious	Proportion of impervious area		
	LU-Wetland	Proportion of wetland area		
Topography Structure	River Density	River density (total river length per unit area)	m	Constant
	Elevation	Elevation	m	
	Slope	Mean slope	°	
	SP	Soil porosity (0-200cm)	% by volume	

187 This study categorized the environmental variables into four groups: External N
188 Loading, Hydro-climatic Dynamics, Soil N and Land Cover, and Topography Structure,

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

based on their physical attributes, biogeochemical functions, and roles in watershed nitrogen cycling. This classification aimed to systematically characterize the potential drivers of lake nitrogen variability from multiple process-oriented perspectives, including external inputs, hydrothermal-driven, substrate accumulation, and topography structural constraints (Table 1).

2.3 Machine learning framework for N concentration reconstruction

2.3.1 Data processing

The input variables were organized into multiple feature branches according to their physical and environmental attributes. To explicitly represent seasonal variability, a cyclic temporal encoding based on the month of the year was applied, allowing the model to capture periodic patterns in lake nitrogen concentrations. To reduce the influence of differing units and numerical ranges among variables (e.g., precipitation, temperature, and human activity indicators) on model convergence and predictive stability, all features within each branch were standardized using Z-score normalization (Kim et al., 2025). Distributions of total nitrogen (TN) and ammonia nitrogen (NH_3-N) concentrations from the national surface water monitoring network exhibited pronounced left-skewness. Accordingly, the target variables were log-transformed during model training to reduce skewness and stabilize variance.

2.3.2 MLP-XGB ensemble learning

In this study, a stacking ensemble combining a Multilayer Perceptron (MLP) and XGBoost was implemented to leverage their complementary strengths. The MLP captured smooth, continuous nonlinear relationships between nitrogen concentrations and environmental drivers across the full feature space, while XGBoost identified

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

localized patterns, threshold effects, and variable interactions. The framework consists of two layers: the first (Level-1) comprises the two base learners, and the second (Level-2) uses ridge regression as a meta-learner to integrate their predictions.

In Level-1, the MLP serves as a feature-aware and high-order nonlinear modeling module, adaptively weighting the relative importance of different environmental feature branches and extracting high-dimensional, nonlinear feature embeddings from the raw input variables. MLP is a typical feedforward artificial neural network that models complex relationships between input variables and target variables through multiple layers of nonlinear mapping (Xie et al., 2026). For the j-th neuron in a hidden layer, the input is first formed by the weighted sum of outputs from the previous layer plus a bias term:

$$z_j = \sum_{i=1}^{n} w_{ij} x_i + b_j \tag{1}$$

Here, $x_i$denotes the output of the i-th neuron in the previous layer, w_{ij}represents the connection weight between neurons, $b_j$is the bias term, and $n$is the number of neurons in the previous layer. The output layer combines the outputs of the neurons in the final hidden layer to produce the final prediction:

$$\hat{y} = g\left(\sum_{j=1}^{m} w_j h_j + b\right) \tag{2}$$

Here, $g(x)$denotes the activation function of the output layer, which is typically linear for regression tasks. During model training, the loss function is minimized using the backpropagation algorithm, iteratively updating the network weights and bias parameters. XGBoost is a gradient boosting-based ensemble learning algorithm that

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

sequentially adds multiple regression trees to correct the prediction errors of previous stages, offering strong capability to capture nonlinear interactions and excellent generalization performance (Niazkar et al., 2024). It is trained on both the original input features and the MLP-derived features, focusing on capturing nonlinear interactions among variables.

In the second layer (Level-2), a ridge regression model with regularization was used to linearly weight and integrate the predictions from the MLP-XGBoost model. Through this stacking ensemble strategy, the MLP focuses on learning global nonlinear response patterns, while XGBoost compensates for local residuals, thereby improving overall predictive accuracy and generalization capability.

The sample data were randomly split into a calibration dataset (70%) and an independent validation dataset (30%). A 5-fold out-of-fold (OOF) strategy was applied during training. Model hyperparameters were optimized using grid search (Table S1). Model performance was quantitatively evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE).

2.4 Trend analysis and mechanism attribution methods

2.4.1 Mann–Kendall (M-K) trend analysis method

This study combined the Theil–Sen trend estimation method with the Mann-Kendall (MK) trend test to quantify the long-term changes in nitrogen concentrations and their potential driving factors over the study period. The Theil-Sen slope estimates the trend by calculating the median of slopes from all possible pairwise combinations of observations in a time series, offering strong robustness and insensitivity to outliers

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

(Zheng et al., 2026).

$$Slope = \text{median}\left(\frac{A_j - A_i}{j - i}\right), \forall j > i \quad (3)$$

In the formula, $A_i$and $A_j$represent the observations of the target variable in the i-th and j-th years, respectively. A slope greater than 0 indicates an increasing trend in the variable over the study period, whereas a slope less than 0 indicates a decreasing trend.

The Mann-Kendall test is a widely used nonparametric method in hydrology and water environment studies for detecting significant monotonic trends in time series (Kendall, 1970; Gilbert, 1987). For a given time series $A = \{A_1, A_2, \dots, A_n\}$, the MK test constructs the statistic S by comparing the relative magnitudes of all sample pairs and subsequently calculates the standardized test statistic Z.

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sgn}(A_j - A_i) \quad (4)$$

$$\text{sgn}(A_j - A_i) = \begin{cases} 1, & A_j - A_i > 0 \\ 0, & A_j - A_i = 0 \\ -1, & A_j - A_i < 0 \end{cases} \quad (5)$$

$$\text{Var}(S) = \frac{n(n-1)(2n+5) - \sum_{m=1}^{k} t_m(t_m - 1)(2t_m + 5)}{18} \quad (6)$$

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{Var}(S)}}, & S > 0 \\ 0, & S = 0 \\ \frac{S+1}{\sqrt{\text{Var}(S)}}, & S < 0 \end{cases} \quad (7)$$

In the formula, $n$is the number of study years; sgn denotes the sign function; Var(S)is the variance; $k$is the number of tied groups in the series, and $t_m$is the number of observations in the m-th tied group. In this study, the Z-value serves a dual purpose: (1) it is used to assess the significance of trends; $|Z|$> 1.96 indicates significance at the 95% confidence level ($p < 0.05$); (2) the magnitude and sign of Z-value reflect the

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

standardized trend strength and direction, which are used as independent and dependent variables in subsequent driver analysis models.

2.4.2 Driver mechanism analysis method

This study employed the SHAP (SHapley Additive exPlanations) feature attribution method based on cooperative game theory(Lundberg and Lee, 2017). SHAP computes the Shapley value for each feature, precisely attributing the predictions of a machine learning model to individual input features, thereby quantifying the strength and direction of each feature's contribution to the target variable at specific spatial and temporal scales (Lundberg et al., 2020).

$$y_i = \varphi_0 + \sum_{j=1}^{m} shap(x_{ij}) \tag{8}$$

Here, $y_{(i)}$ is the predicted nitrogen concentration, ϕ_0 is the mean predicted nitrogen concentration, and $shap(x_{ij})$ represents the contribution of feature variable j to $y_{(i)}$. SHAP was applied across both spatial and temporal dimensions: (1) using the multi-year average nitrogen concentration as the target variable and the spatial background values of drivers as features, to identify the dominant factors driving spatial differentiation of nitrogen concentrations and to explain the geographic origins of background-level differences. (2) using the Z-values of nitrogen concentrations as the target variable and the corresponding Z-values of drivers as features, to identify the marginal contributions of dynamic environmental changes to nitrogen trends.

Considering that different feature categories contain unequal numbers of features, directly summing them could give categories with more features a spurious statistical advantage. This study calculated the mean absolute SHAP value of each feature across

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

all samples and then computed the arithmetic mean of all features' SHAP values within each feature category.

3. Results

3.1 Model performance

By comparing observed nitrogen concentrations with model predictions for both the calibration and independent validation datasets, the predictive accuracy and generalization ability of the ensemble learning based lake nitrogen reconstruction model were evaluated. The results indicate that the TN reconstruction model exhibited overall excellent predictive performance (Fig. 2). During the calibration phase, the model achieved an R^2 of 0.860, with RMSE and MAE of 0.840 mg/L and 0.495 mg/L, respectively. The model performed similarly in validation using the independent dataset, with R^2 increasing to 0.875 and RMSE and MAE decreasing to 0.810 mg/L and 0.464 mg/L, respectively. The high consistency between calibration and validation indicates strong generalization capability and the absence of overfitting.

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

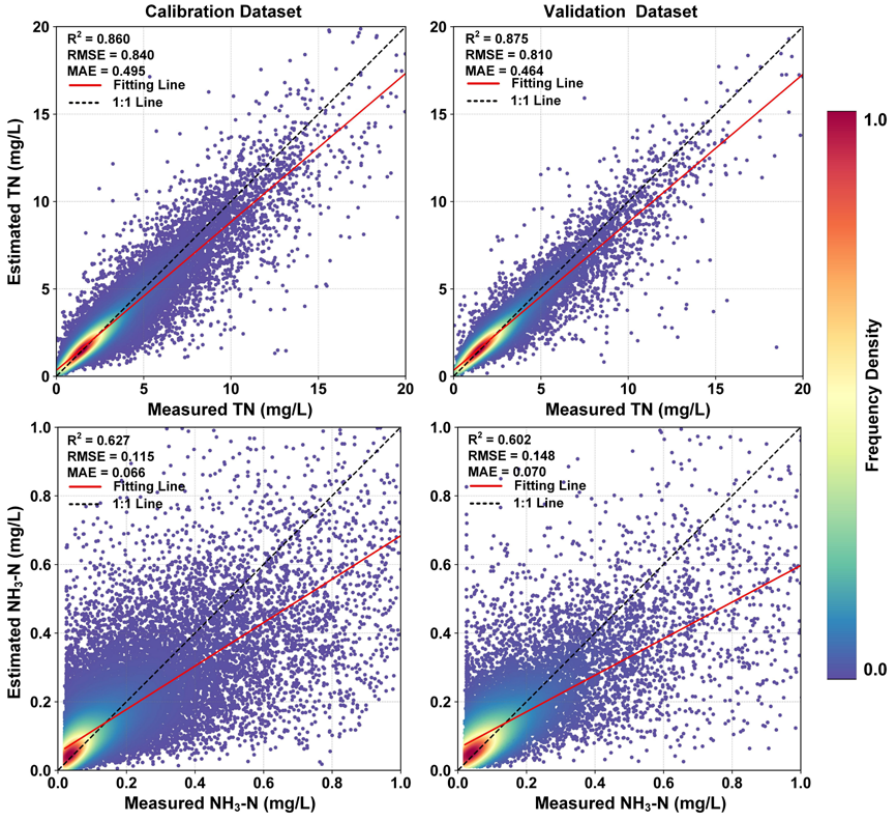

Fig. 2 Calibration and validation of the MLP-XGB model for simulating TN and NH_3-N.

In contrast, NH_3-N is more challenging to predict, with R^2 values of 0.627 and 0.602 for the calibration and validation dataset, respectively. Despite generally low environmental concentrations and high variability of NH_3-N, the model performed well in controlling absolute errors, with RMSE and MAE of only 0.148 mg/L and 0.070 mg/L in the validation dataset, respectively. Frequency distributions show that the majority of observations are concentrated within the 0-0.2 mg/L range. The fitted line closely follows the 1:1 line at low concentrations but shows reduced accuracy at high concentrations (>0.6 mg/L), likely due to the limited number of high-concentration samples and the intrinsic variability of NH_3-N in the environment.

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

3.2 Spatial heterogeneity of lake nitrogen concentrations

The study lakes showed profound spatial heterogeneity of nitrogen concentrations across China (Fig. 3), with a decreasing trend from the lakes in the east to the lakes in the west, and from the lakes in the north to the lakes in the south. The national multi-year mean TN concentration in the lakes was approximately 2.07 mg/L, varying largely from 1.77 mg/L in TPL to 2.56 mg/L in NPML. The EPL and NPML generally exhibited high TN levels nationwide, whereas the TPL showed the lowest TN concentrations. Lakes in the Northern EPL showed the highest TN concentrations (most > 2.5 mg/L), with many lakes around the Shandong Peninsula region often exceeding 4.5 mg/L. In contrast, TN concentrations in the lakes of the southern EPL were mostly between 1.5 and 2.0 mg/L. In the lakes of the TPL, TN concentrations were substantially lower than the national average, with a multi-year mean of approximately 1.77 mg/L. The lakes in the IMXL had a mean TN concentration of 1.98 mg/L, varying largely from < 1.0 mg/L in the western region to 2.0 mg/L in the central and eastern regions.

The national multi-year mean NH_3-N concentration in the lakes was 0.141 mg/L, showing an overall spatial pattern broadly consistent with TN but with stronger localized clustering of high values. High concentrations were mainly observed in the NPML (mean 0.159 mg/L) and EPL (mean 0.149 mg/L). Specifically, continuous high-concentration patches occurred in the middle and northern EPL, including the middle–lower Yangtze Plain and North China, as well as in the central NPML. In contrast, NH_3-N concentrations in the TPL and YGPL were generally low, with most lakes below 0.15 mg/L and only a few lakes exhibiting relatively elevated values. Overall, Lakes

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

exhibiting high NH_3-N concentrations often overlap with TN hotspots, particularly in the eastern plains and northeastern regions, forming the core hotspots of lake nitrogen pollution in China.

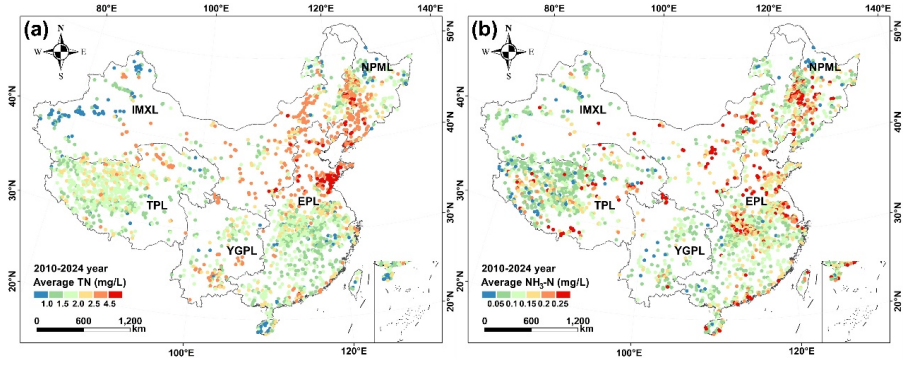

Fig. 3 Spatial patterns of long-term (2010-2024) mean TN and NH_3-N concentrations of lakes in five limnological zones of China, Eastern Plain Lakes (EPL), Inner Mongolia-Xinjiang Lakes (IMXL), Northeast Plain-Mountain Lakes (NPML), Qinghai-Tibet Plateau Lakes (TPL), and Yunnan-Guizhou Plateau Lakes (YGPL).

3.3 Temporal variations and long-term trend analysis

Over the past 15 years, the national average concentration of total nitrogen (TN) in Chinese lakes has remained essentially stable, while the concentration of ammonia nitrogen (NH_3-N) has exhibited a continuous downward trend (Figs. 4a, c). For TN, the NPML and the IMXL displayed upward fluctuations; in contrast. No significant monotonic trends were observed in the TPL, EPL and YGPL during the study period. The interannual variation pattern of NH_3-N differed markedly from that of TN. Except for TPL, NH_3-N concentrations in all other limnological zones exhibited significant declining trends.

Monthly-scale analysis further revealed seasonal variations in lake nitrogen concentrations (Fig. 4b, d). For TN, EPL, IMXL, and NPML exhibited consistent seasonal rhythms, displaying a typical "U-shaped" pattern: concentrations steadily

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

declined from January, reached the annual minimum in July, and gradually increased from August to December. In contrast, TN concentrations in YGPL peaked primarily during August-November, whereas TPL showed relatively weak seasonal fluctuations. Seasonal variations of NH_3-N differed more markedly among limnological zones. NPML exhibited the strongest intra-annual fluctuations, with NH_3-N concentrations remaining high during November to April of the following year, rapidly declining between March and May from annual highs to the lowest nationwide levels, and then remaining relatively stable for the subsequent months. Conversely, NH_3-N concentrations in TPL and YGPL peaked during July-August and remained at low levels from late autumn through the following spring.

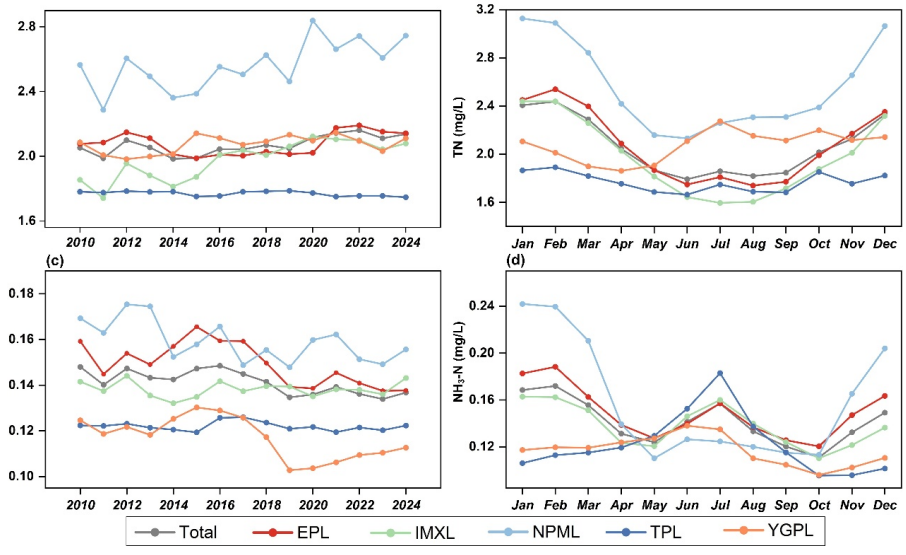

Fig. 4 Interannual and seasonal changes of annual mean TN and NH_3-N concentrations in the 3020 study lakes across five limnological zones in China,

Over the past 15 years, TN concentrations exhibited an overall increasing trend in 21% of the study lakes, especially a significant upward trend in in the NPML and IMXL regions. But, 14% of the study lakes showed a significant decrease and the remaining

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

lakes did not show a clear change trend. In contrast, NH_3-N concentrations generally showed continuous decline; nationwide, 23% of lakes exhibited a significant decrease, while only 9% show an increase, markedly lower than TN (Fig. 5a, 6a). The NPML and IMXL were the regions with the clearest increasing trend of TN, with 36% and 32% of lakes showing significant upward trends, respectively. And IMXL also exhibited the highest proportion of NH_3-N deterioration (15%). YGPL stands out in NH_3-N changes, with 37% of lakes significantly decreasing and only 6% increasing. Lakes in the TPL region demonstrated high stability in changes of both nitrogen forms, with more than half of the lakes showing no significant trends (Fig. 5b, 6b).

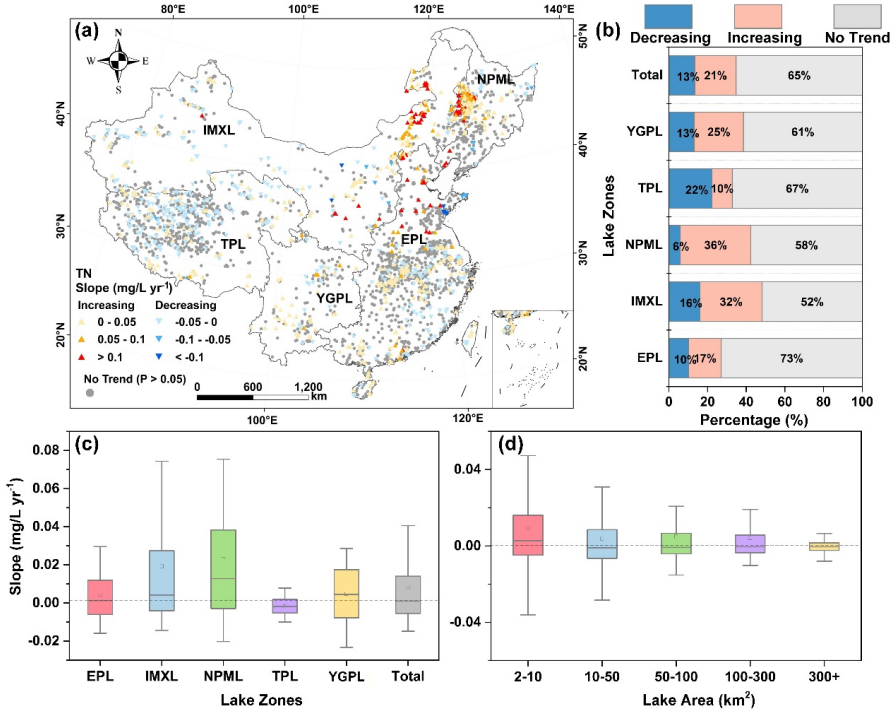

Fig. 5 Mann-Kendall trend analysis of changes in TN concentration from 2000 to 2024. (a) Spatial distribution of TN trends (Slope). (b) Percentage of lakes with increasing, decreasing, or no significant TN trends in different limnological zones. (c) Boxplot of TN slope across different limnological zones. (d) Boxplot of TN slope across lakes of

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

different area.

The distribution of nitrogen concentration changes rates further revealed differences in responses among limnological zones (Fig. 5c, 6c). The slope distributions of TN and NH_3-N in TPL were the most concentrated, showing the smallest variability, whereas the slope distributions in IMXL and YGPL were markedly broader, indicating stronger individual variability in lake responses to environmental disturbances in these regions. Lake area exerted a significant constraining effect on long-term TN and NH_3-N trends (Fig. 5d, 6d): as lake size increases, the dispersion of slope values converged markedly, with median slopes gradually approaching zero. Large lakes (>300 km^2) generally exhibit strong stability, whereas small lakes (2-50 km^2) showed highly dispersed change rates, reflecting their sensitivity to external nitrogen inputs and hydrological fluctuations.

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

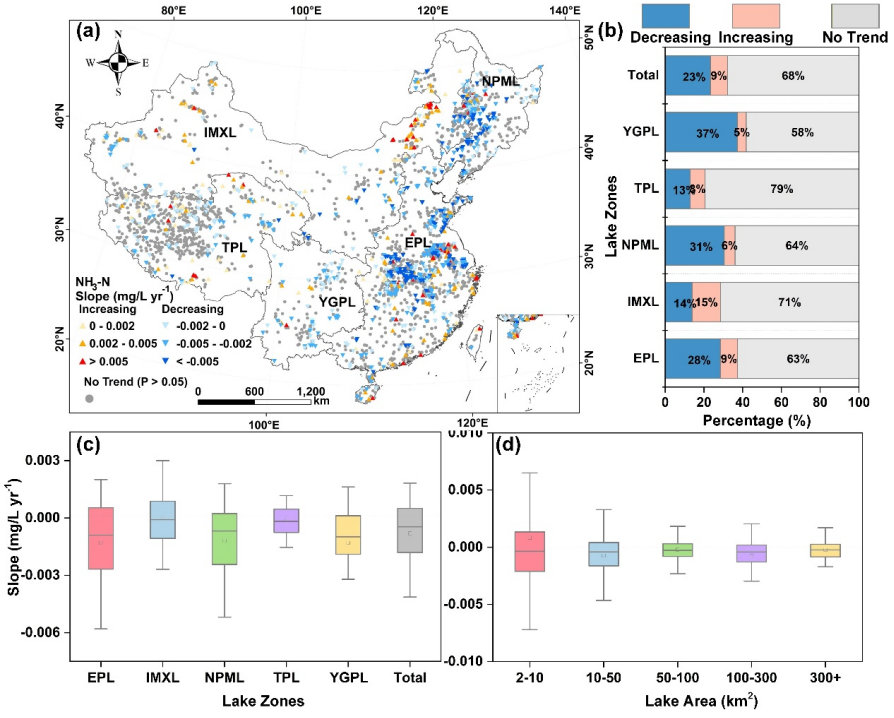

Fig. 6 Mann-Kendall trend analysis of changes in NH_3-N concentration from 2000 to 2024. (a) Spatial distribution of NH_3-N trends (Slope). (b) Percentage of lakes with increasing, decreasing, or no significant NH_3-N trends in different limnological zones. (c) Boxplot of NH_3-N slope across different limnological zones. (d) Boxplot of NH_3-N slope across lakes of different area.

3.4 Factors associated with the spatial patterns and temporal trends

Kruskal–Wallis tests based on absolute SHAP values indicated that the contributions of all feature categories differed significantly both across seasons and among limnological zones ($p < 0.001$) (Table S2). The dominant environmental factors associated with the spatial pattern were primarily related to external nitrogen loading (Fig. 7). This pattern was particularly evident in the IMXL, NPML, and northern EPL regions, whereas hydro-climatic dynamics exhibited higher relative contributions in multiple sub-basins of the middle and lower Yangtze Plain and the YGPL. In contrast, the TPL region did not show a single clearly dominant feature category; instead,

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

multiple natural factors exhibited comparable contributions and were spatially interwoven. For NH_3-N, the spatial pattern exhibited a clearer north–south differentiation in dominant feature categories. In the EPL and YGPL, NH_3-N concentrations were mainly associated with external nitrogen loading, whereas in IMXL, NPML, and TPL, higher relative contributions were generally observed for soil nitrogen and land cover–related factors (Fig. 7b).

The contributions of all feature categories differed significantly across seasons ($p < 0.001$; Fig. 7c; Table S2), with external nitrogen loading showing particularly large seasonal variability during summer. In contrast, the seasonal variation in NH_3-N feature contributions was comparatively more stable, with hydro-climatic dynamics consistently contributing a small proportion (median <12%), and NH_3-N concentrations being primarily associated with external inputs and land cover background over the long term (Fig. 7e). At the limnological zone scale, TN generally exhibited a multi-factor co-dominant contribution structure, whereas NH_3-N showed a more pronounced single-category dominance in several lake regions. For example, in the EPL, 76% of sub-basins were dominated by external nitrogen loading for NH_3-N, markedly higher than the corresponding proportion for TN (22%). In TPL and NPML, NH_3-N was largely associated with land cover, accounting for 82% and 97% of sub-basins, respectively (Fig. 7d,f).

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

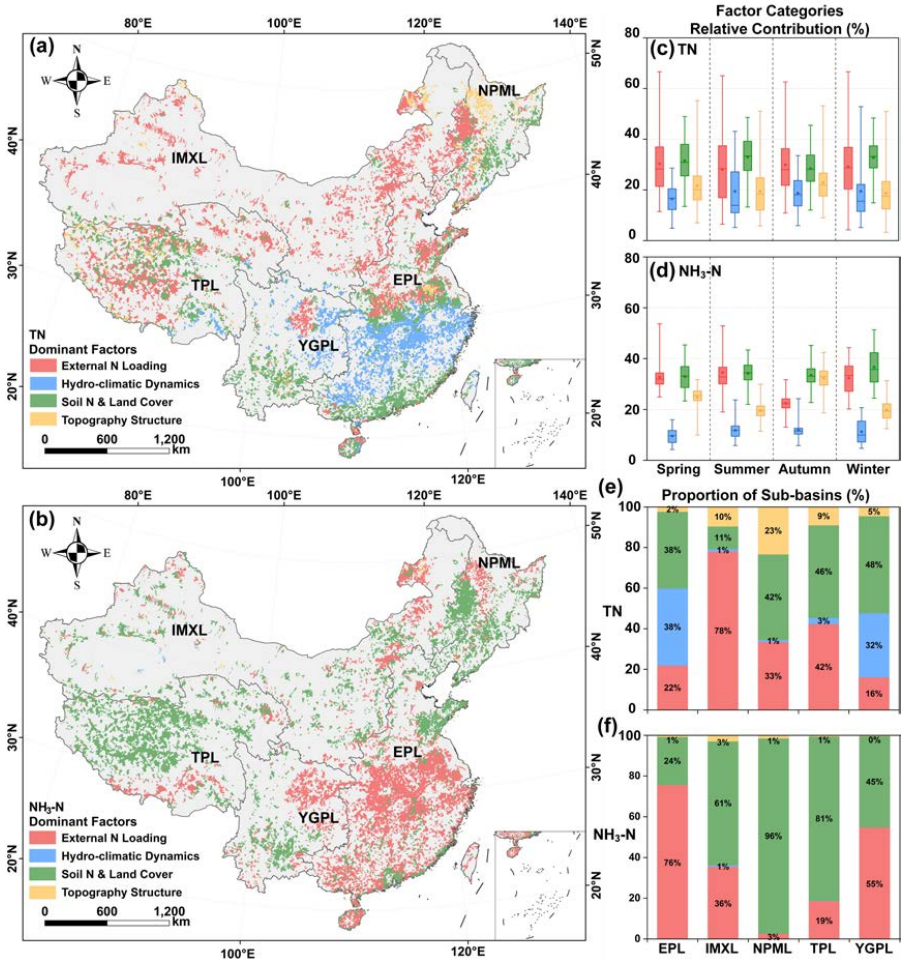

Fig. 7 (a, b) Spatial distribution of dominant feature categories for TN and NH_3-N in lake sub-basins; (c, d) Seasonal dynamics of feature category contributions for TN and NH_3-N; (e, f) Proportions of sub-basins in each lake zone dominated by TN and NH_3-N feature categories.

The dominant factors driving the long-term trends of TN and NH_3-N were quantified using trend significance indicators (Z-values) from the Mann-Kendall test (Fig. 8). At the national scale, the dominant drivers of TN trends exhibited distinct spatial zonation (Fig. 8a). External nitrogen loading dominated TN trends in the central EPL, YGPL, and eastern IMXL, while TN trends in the majority of sub-basins in the NPML and IMXL were primarily attributed to hydro-climatic dynamics. In the TPL,

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

TN trends were governed by soil nitrogen and land cover. Atmospheric nitrogen deposition (both NH_x-N and NO_y-N) was the most explanatory positive driver, with variations in deposition intensity recorded as a primary factor affecting TN trend changes. Shallow soil moisture (0-7 cm) showed a positive correlation with TN trends, and increases in subsurface runoff were associated with upward TN trajectories. Conversely, the expansion of water area exhibited an inhibitory effect on TN trends, while increases in impervious surfaces and cropland were associated with positive feedback on TN trends (Fig. 8b).

The driving factors of long-term NH_3-N trends exhibited higher spatial consistency at the national scale (Fig. 8c). Except for localized sub-basins in the TPL, the long-term changes in NH_3-N concentrations were primarily governed by external nitrogen loading. Features including nitrogen deposition, fertilizer application, and manure application showed strong positive driving effects (Fig. 8d). Hydro-climatic factors, such as precipitation and subsurface runoff, contributed less to NH_3-N trends compared to TN. Overall, TN trends were associated with the combined influence of external loading and hydrological processes, whereas long-term NH_3-N trends were predominantly driven by external loading intensity.

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

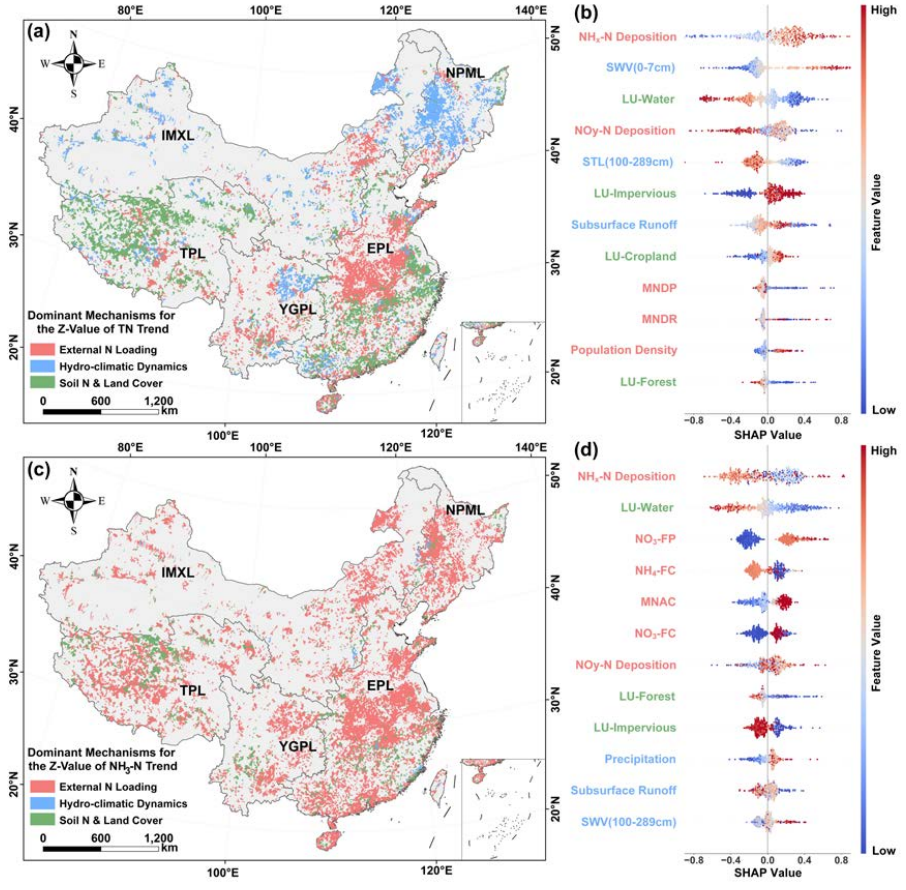

Fig. 8 (a-b) Dominant feature categories for long-term significant trends (Z-Value) of TN and NH_3-N concentrations in lake sub-basins. (c-d) Global explanation of feature importance for long-term trends of TN and NH_3-N.

4. Discussion

4.1 Decoupling between spatial patterns and temporal trends

This study decouples the spatial pattern and long-term trend of TN and NH_3-N concentrations in 3020 lakes across five limnological zones in China. The study found that the dominant factors affecting the spatial difference in lake nitrogen concentration are not necessarily the key factors that influence their long-term trends. The spatial distribution of TN concentrations is jointly regulated by external nitrogen loading,

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

hydro-climatic dynamics, and background environmental conditions (Fig. 7b, d), highlighting its close coupling with solute transport processes at the watershed scale. In the southern part of the EPL, the spatial distribution of the TN region dominated by climatic and hydrological dynamics exhibits a high degree of consistency with the annual precipitation gradient. With their boundaries roughly aligning with the ~800 mm precipitation isohyet in China, highlighting the critical role of the hydro-climatic background in shaping TN spatial patterns. Interactions among hydro-climatic factors can give a rise to a complex concentration-load relationship for TN, whereby under wet conditions, the combined effects of increased runoff inputs and dilution can result in an asynchrony between concentration peaks and precipitation maxima, as reported by Jiang et al. (2025).

By contrast, the spatial distribution of NH_3-N more directly reflects the regional difference in external nitrogen input (Fig. 7b, d). In densely populated regions with intensive industrial and urban development, domestic and industrial wastewater discharges can be a major factor affecting dissolved organic matter in receiving water bodies (Cao et al., 2025), potentially playing a major role in shaping NH_3-N spatial patterns. In agriculturally dominated areas, fertilizer application and atmospheric deposition act as diffuse sources controlling its distribution at broader spatial scales (Ye et al., 2021; Vincent et al., 2025; Zhao et al., 2024a). Such contrasts are particularly evident among major river basins: in the Yellow River Basin, TN is more readily driven by the combined effects of agricultural activities and rainfall events, whereas in the Yangtze River Basin, NH_3-N exhibits a more pronounced control by urban and

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

industrial emissions (Wang et al., 2016; Cui et al., 2021).

At the level of long-term trends (Fig. 8), the evolutionary directions of TN and NH_3-N exhibit a more pronounced divergence. The trends in TN vary markedly across regions, being jointly influenced by changes in external nitrogen loading and significantly modulated by interannual hydro-climatic variability (Fig. 8a). In semi-arid and semi-humid regions characterized by strong climatic variability, high-runoff events can deliver substantial nitrogen loads over short periods, allowing TN to remain highly variable or even continue increasing despite emission reduction efforts (Tian et al., 2025). By contrast, NH_3-N exhibits a highly consistent declining trend at the national scale (Fig. 8b), with its long-term evolution being almost entirely governed by changes in the intensity of external nitrogen inputs, in close agreement with recent improvements in wastewater treatment capacity and the implementation of emission control policies (Meng et al., 2021; Zhou et al., 2021). This study finds that the North China Plain shows the most pronounced declining trends in both TN and NH_3-N, consistent with the findings of Deng et al (Deng et al., 2024). During China's urbanization process, improvements in wastewater management and reductions in agricultural emissions have substantially decreased nitrogen loads exported to inland waters and the ocean (from 6.8 to 3.5 million tons), with the most pronounced reductions occurring in the North China Plain. The trend in external nitrogen input synchronizes with the long-term evolution of lake nitrogen concentrations, confirming the dominant role of external nitrogen loading.

These spatiotemporal differences collectively indicate a decoupling between the

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

spatial patterns and temporal trends of lake nitrogen dynamics. The spatial distribution of TN is jointly controlled by multiple categories of drivers (Fig. 7a,c), whereas at the long-term trend level, the dominant controls shift markedly toward changes in external nitrogen inputs and climatic variability (Fig. 8a). NH_3-N concentrations respond rapidly to changes in external inputs, rendering their trend dynamics relatively independent of spatial heterogeneity. Moreover, differences in the behavior of driving factors across spatial and temporal dimensions constitute an important cause of this separation. Factors such as land use, soil properties, and hydrological configurations are highly heterogeneous in space but exhibit limited interannual variability, and therefore primarily shape the spatial distribution of nitrogen concentrations (Zhao et al., 2024b; Wang et al., 2025b). In contrast, external inputs including nitrogen deposition, fertilization intensity, and emission controls show clear temporal trend and are more likely to govern long-term evolutionary directions (Shen et al., 2024; Li et al., 2025).

4.2 Nonlinear responses of nitrogen concentration trends to environmental drivers

This study further reveals nonlinear responses of the long-term TN and NH_3–N trends to changes in key environmental drivers (Fig. 9). Changes in external nitrogen loading represent a core factor on nitrogen condition in the Chinese lakes; once the trends of atmospheric NH_x-N deposition and ammonium fertilizer application (NH_4-FC) shift from negative to positive (Z>0), the corresponding SHAP values rapidly switch from inhibitory to strongly promotive and accelerate with increasing Z values, indicating a highly sensitive and rapid response of NH_3-N to enhanced external reactive nitrogen inputs. Agricultural nitrogen inputs make a clear difference in trend responses

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

across nitrogen forms, with inorganic fertilizers and manure-related factors significantly amplifying TN increases during prolonged periods of high-intensity growth, reflecting a time-lagged accumulation and release of agricultural nitrogen in the form of "legacy nitrogen" as reported by Basu(Basu et al., 2022). Many studies have shown that despite the implementation of nutrient management, legacy nitrogen stored in soils and groundwater can continue to affect surface water quality over decades(Gao et al., 2023; Nguyen et al., 2022). In contrast, the effect of agricultural input on NH_3-N trends is more direct, exhibiting an approximately monotonic positive response overall, further confirming that NH_3-N variations can rapidly reflect an increase or a decrease in external input intensity, as it was also reported by Liu(Liu et al., 2022). This is likely because NH_3-N is a gas, whose concentration in water bodies changes quickly.

Trends in hydrological and underlying surface factors exert steady yet distinct nonlinear regulation on the evolution of TN and NH_3-N by modifying transport pathways and dilution processes (Fig. 9d-f). Increases in the proportion of water area exhibit a consistent linear inhibitory effect on both TN and NH_3-N, highlighting a stable buffering mechanism whereby water body expansion reduces nitrogen concentrations through dilution and prolonged hydraulic residence time for different nitrogen forms(Wang et al., 2025d). Forested land intercepts pollutant transport through plant uptake, but its nitrogen removal capacity tends to saturate under high nitrogen loading and may even exhibit promotive effects over certain ranges, potentially associated with litter decomposition and enhanced nitrogen transport through ditch and channel networks (Xu et al., 2023; Cui et al., 2025). Under conditions of markedly reduced

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

subsurface runoff or rapidly increasing shallow soil water content, upward trends in TN are substantially intensified (Fig. 9f,k), suggesting a key role of soil moisture condition in regulating the relative contribution of surface runoff versus deep leaching. When deep leaching is constrained or surface soils approach saturation, surface runoff becomes more effective in mobilizing and exporting accumulated nitrogen (Cui et al., 2024; Zhao et al., 2025). This may have been the cause for the changes occurred in the lake nitrogen concentrations in this study.

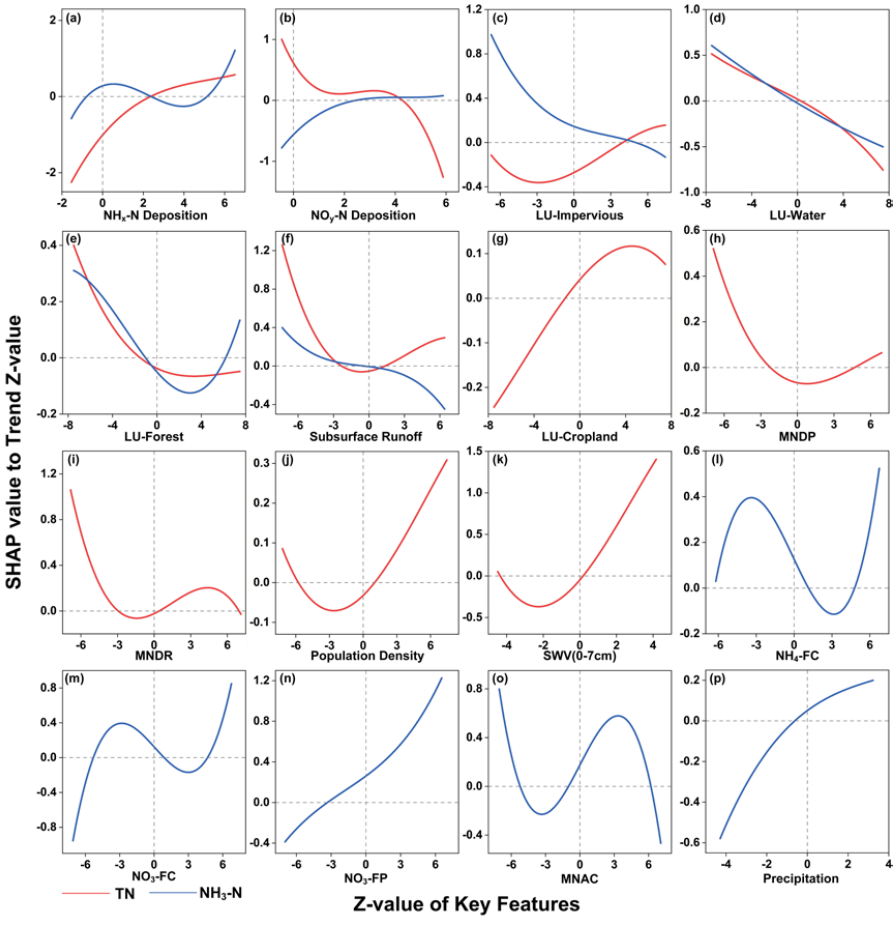

Fig. 9 Response relationships between trends in key factors and trends in TN and NH_3-N. (a-f) Key driving factors that are important for both TN and NH_3-N trends. (g-k)

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

Specific driving factors exerting high influence exclusively on TN trends. (l-p) Specific driving factors exerting high influence exclusively on NH_3-N trends. The y-axis (SHAP value) quantifies the influence intensity of each driving factor on the significance of TN/NH_3-N concentration trends (Z-values). When the curve lies above the zero line, the current state of the feature (represented by its own Z-value on the x-axis) exerts a positive effect on concentration increases, either intensifying an upward trend or alleviating a declining trend. Conversely, values below the zero line indicate a suppressive effect on concentration trends. A positive feature Z-value indicates a long-term increasing trend of the driver, whereas a negative value indicates a decreasing trend. The slope of the curve reflects the sensitivity of nutrient evolution trends to changes in the corresponding driving factor.

4.3 China as a Representative Microcosm of Lake Nitrogen Evolution

This study identifies a clear divergence between the rapid response of NH_3-N and the more persistent fluctuations of TN, reflecting the widespread inertia inherent in watershed- and lake-scale nitrogen pollution and recovery processes at the global scale. For example, a study in Germany (Nguyen et al., 2022) showed that approximately 58% of nitrogen surpluses accumulated through long-term agricultural development can be removed via denitrification, while the remaining fraction is persistently stored in soils and groundwater, exerting long-term impacts on surface water nitrogen concentrations. Similarly, in the Mississippi River Basin of the United States, despite the reduction in point-source nitrogen inputs, the decline rates of nitrate and TN remain slow(Sadayappan et al., 2022).

Socioeconomic gradient analyses further reveal an “environmental Kuznets curve” pattern, whereby middle-income areas tend to experience more severe nitrogen pollution. This stage typically corresponds to periods of accelerated urban expansion, during which wastewater treatment infrastructures lag behind pollution growth (Tong et al., 2022; Chai et al., 2025). For instance, in India, nitrogen fertilizer application increased by approximately 11.3-fold between 1970 and 2022, whereas grain

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

production rose by only about threefold, indicating a clear transition toward nitrogen saturation (Sapkota and Bijay-Singh, 2025). Meanwhile, driven by variations in climate, cropping systems, soil properties, and management policies, nitrogen loss pathways exhibit regional heterogeneity across India, closely mirroring the nitrogen pollution processes experienced by China during its phase of rapid agricultural intensification (Elsayed et al., 2025). At present, China, India, the United States, Brazil, and Pakistan together account for more than 60% of global synthetic fertilizer nitrogen consumption, and when combined with manure nitrogen inputs from livestock production, constitute major sources of nitrogen pollution in global aquatic systems (Lu and Tian, 2017). In addition, in sub-Saharan Africa, more than 80% of sewage remains untreated, with approximately 20-60 million cubic meters of nitrogen-containing wastewater (about 1×10^6 kg N) discharged directly into the environment each year (Nyenje et al., 2010). Without coordinated catchment-scale source control and hydrological pathway interventions, rapidly developing lake systems worldwide are likely to face diverging nitrogen evolutionary trajectories and prolonged delays in ecological recovery.

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

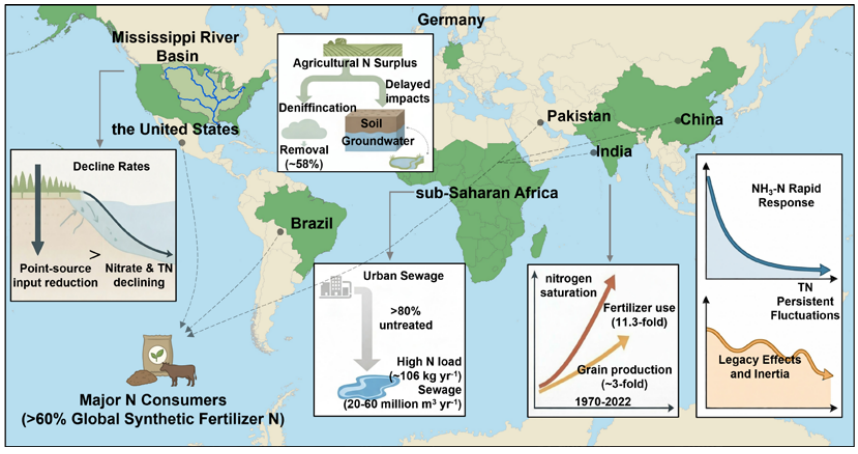

Fig. 10 China as a microcosm of global nitrogen evolution: From direct discharge to legacy-driven inertia. The figure illustrates the contrasting management challenges across major nitrogen consumers, highlighting the shift from point-source control to long-term legacy management (Data sources gathered from Nyenje et al., 2010; Lu and Tian, 2017; Nguyen et al., 2022; Elsayed et al., 2025; Sadayappan et al., 2022; Sapkota and Bijay-Singh, 2025; Elsayed et al., 2025).

5. Conclusions

This study constructed long-term monthly time series (2010-2024) of total nitrogen (TN) and ammonia nitrogen (NH_3-N) concentrations for 3,020 lakes across China's five limnological regions using an ensemble machine learning framework (MLP-XGB). The study revealed the difference in factors that influence spatial patterns and long-term changes of the two nitrogen forms at the large scale. Both the spatial distribution and long-term trend of TN concentration in these lakes were jointly influenced by external nitrogen loading, hydro-climatic dynamics, and land use / land cover types, and soil conditions. On the other hand, the long-term declining trend of NH_3-N concentrations observed in the Chinese lakes appeared to be more directly affected by the external nitrogen loading. Based on these findings, we postulate that both the spatial distribution and long-term trend of lake TN across a large geographical

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

region are reflected by watershed-scale cumulation of nitrogen, and that the long-term trend of lake NH_3-N is primarily affected by external nitrogen loading. The different responses of lake nitrogen forms in spatial and temporal characteristic across China highlight the need for nitrogen-form-specific management strategies. Source control should be prioritized to mitigate NH_3-N pollution, whereas long-term reductions in TN require integrated watershed nutrient management.

Code and data availability

Lake locations and boundaries from the China Lake Dataset (1960s-2020) available online (https://doi.org/10.11888/Hydro.tpdc.270302) and the HydroLAKES database (https://www.hydrosheds.org/products/hydrolakes);
Watershed and river network data from HydroBASINS and HydroRIVERS (https://www.hydrosheds.org/products/hydrobasins);
Climate and soil moisture data from the ERA5-Land monthly averaged reanalysis (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land-monthly-means);
Soil property data from the China Soil Dataset v2 (CSDLv2) available online (https://www.scidb.cn/s/ZZJzAz);
Land-cover data (CLCD) from the China Land Cover Dataset (https://doi.org/10.5194/essd-13-3907-2021);
NDVI and LAI data from MODIS MOD13A1 and MOD15A2H products available from NASA (https://ladsweb.modaps.eosdis.nasa.gov/);
History of anthropogenic Nitrogen inputs (HaNi) from Earth System Science Data. (https://doi.org/10.5194/essd-14-4551-2022);
Population density data from the LandScan Global dataset (https://landscan.ornl.gov/);
Topographic data from the SRTM 30 m digital elevation model available from the USGS (https://earthexplorer.usgs.gov/).

Financial support

This research was supported by the National Key Research and Development Program of China (2023YFC3707700), the Fundamental Research Funds for the Central Universities, and Overseas High-level Talents Program of Shanghai, and the Leading Talents (Overseas) Program of Shanghai.

Author contributions

Chenyi Shi: Conceptualization, Methodology, Software, Formal analysis, Data

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

curation, Visualization, Writing - original draft, Writing - review & editing; Xihua Wang: Conceptualization, Supervision, Project administration, Writing - review & editing; Y. Jun Xu: Conceptualization, Supervision, Writing - review & editing; Chaomeng Dai: Methodology, Supervision; Nianqing Zhou: Methodology, Supervision; Rongbing Fu: Validation, Investigation; Chengming Luo: Validation, Investigation; Boyang Mao: Data curation, Investigation; Shunqing Jia: Data curation, Investigation; Qinya Lv: Data curation, Investigation; Zejun Liu: Data curation, Investigation; Xuming Ji: Validation, Resources; Yan Dai: Resources; Yanxin Rong: Resources. All authors discussed the results, contributed to the interpretation of the findings, and approved the final manuscript.

Competing interests

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

Reference

Adrian, R., O'Reilly, C.M., Zagarese, H., Baines, S.B., Hessen, D.O., Keller, W., Livingstone, D.M., Sommaruga, R., Straile, D., Van Donk, E., Weyhenmeyer, G.A., Winder, M., 2009. Lakes as sentinels of climate change. Limnol. Oceanogr. 54, 2283–2297. https://doi.org/10.4319/lo.2009.54.6_part_2.2283

Basu, N.B., Van Meter, K.J., Byrnes, D.K., Van Cappellen, P., Brouwer, R., Jacobsen, B.H., Jarsjö, J., Rudolph, D.L., Cunha, M.C., Nelson, N., Bhattacharya, R., Destouni, G., Olsen, S.B., 2022. Managing nitrogen legacies to accelerate water quality improvement. Nat. Geosci. 15, 97–105. https://doi.org/10.1038/s41561-021-00889-9

Beutel, M.W., 2006. Inhibition of ammonia release from anoxic profundal sediments in lakes using hypolimnetic oxygenation. Ecol. Eng. 28, 271–279. https://doi.org/10.1016/j.ecoleng.2006.05.009

Cao, X., Chen, S., Liu, Y., Long, G., Xu, Y.J., 2025. Domestic wastewater is an overlooked source and quantity in global river dissolved carbon. Nat. Commun. 16, 7522. https://doi.org/10.1038/s41467-025-62920-6

Causse, J., BaurèS, E., Mery, Y., Jung, A.-V., Thomas, O., 2015. Variability of N export in water: a review. Crit. Rev. Environ. Sci. Technol. 45, 2245–2281. https://doi.org/10.1080/10643389.2015.1010432

Chai, G., Jiang, T., Guo, Y., Peng, H., Huang, M., Cui, H., Cui, X., Tong, Y., 2025. Machine learning reveals global drivers of dissolved nitrogen dynamics in aquatic ecosystems: implications for N_2O emissions and regional management. Water Res. 287, 124450. https://doi.org/10.1016/j.watres.2025.124450

Cui, M., Guo, Q., Wei, R., Chen, T., 2021. Temporal-spatial dynamics of anthropogenic nitrogen

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

inputs and hotspots in a large river basin. Chemosphere 269, 129411. https://doi.org/10.1016/j.chemosphere.2020.129411

Cui, X., Ouyang, W., Wang, J., Tulcan, R.X.S., Zhu, W., 2025. The neglected role of forest eco-hydrological process representation in regulating watershed nitrogen loss. Water Res. 282, 123735. https://doi.org/10.1016/j.watres.2025.123735

Cui, Z., Chen, C., Chen, Q., Huang, J., 2024. Difference in the contribution of driving factors to nitrogen loss with surface runoff between the hill and plain agricultural watersheds. J. Geophys. Res.: Biogeosci. 129. https://doi.org/10.1029/2023JG007931

Deng, O., Wang, S., Ran, J., Huang, S., Zhang, X., Duan, J., Zhang, L., Xia, Y., Reis, S., Xu, J., Xu, J., de Vries, W., Sutton, M.A., Gu, B., 2024. Managing urban development could halve nitrogen pollution in China. Nat. Commun. 15, 401. https://doi.org/10.1038/s41467-023-44685-y

Du, Y., Chen, F., Zhang, Y., He, H., Wen, S., Huang, X., Song, C., Li, K., Wang, J., Keellings, D., Lu, Y., 2023. Human activity coupled with climate change strengthens the role of lakes as an active pipe of dissolved organic matter. Earth's Future 11, e2022EF003412. https://doi.org/10.1029/2022EF003412

Elsayed, H., Beusen, A., Prusty, A.K., Bouwman, L., 2025. Long-term variations (1970-2020) and spatial patterns of nitrogen and phosphorus soil budgets and fates in indian agriculture. Nutr. Cycling Agroecosyst. 130, 17–32. https://doi.org/10.1007/s10705-024-10375-4

Fowler, D., Coyle, M., Skiba, U., Sutton, M.A., Cape, J.N., Reis, S., Sheppard, L.J., Jenkins, A., Grizzetti, B., Galloway, J.N., Vitousek, P., Leach, A., Bouwman, A.F., Butterbach-Bahl, K., Dentener, F., Stevenson, D., Amann, M., Voss, M., 2013. The global nitrogen cycle in the twenty-first century. Philos. Trans. R. Soc. B: Biol. Sci. 368, 20130164. https://doi.org/10.1098/rstb.2013.0164

Gao, Y., Tian, Y., Zhan, W., Li, L., Sun, H., Zhao, T., Zhang, H., Meng, Y., Li, Y., Liu, T., Ding, J., 2023. Characterizing legacy nitrogen-induced time lags in riverine nitrogen reduction for the songhuajiang river basin: source analysis, spatio-seasonal patterns, and impacts on future water quality improvement. Water Res. 242, 120292. https://doi.org/10.1016/j.watres.2023.120292

Gao, Y., Zhou, F., Ciais, P., Miao, C., Yang, T., Jia, Y., Zhou, X., Klaus, B.-B., Yang, T., Yu, G., 2020. Human activities aggravate nitrogen-deposition pollution to inland water over china. Natl. Sci. Rev. 7, 430–440. https://doi.org/10.1093/nsr/nwz073

Gilbert, R.O., 1987. Statistical methods for environmental pollution monitoring (No. PNL--4660). Pacific Northwest National Lab. (PNNL), Richland, WA (United States).

Green, P.A., Vörösmarty, C.J., Meybeck, M., Galloway, J.N., Peterson, B.J., Boyer, E.W., 2004. Pre-industrial and contemporary fluxes of nitrogen through rivers: a global assessment based on typology. Biogeochemistry 68, 71–105. https://doi.org/10.1023/B:BIOG.0000025742.82155.92

Hu, M., Ma, R., Xue, K., Cao, Z., Xiong, J., Loiselle, S.A., Shen, M., Hou, X., 2024. Eutrophication evolution of lakes in China: four decades of observations from space. J. Hazard. Mater. 470, 134225. https://doi.org/10.1016/j.jhazmat.2024.134225

Ji, K., Ouyang, W., Hao, X., Hu, J., Li, D., Chen, T., 2025. Hydrological regime mediates stoichiometry regime and source–sink alternation in shallow lakes. Water Res. 287, 124523. https://doi.org/10.1016/j.watres.2025.124523

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

Jiang, S., Werner, A.D., Gao, L., Rode, M., 2026. Climate change expected to amplify land-use change impacts on nitrogen export from a subtropical catchment in China. J. Hydrol. 667, 134888. https://doi.org/10.1016/j.jhydrol.2025.134888

Jiang, X., Zhao, X., Qiang, M., Li, Z., Yang, Z., Yang, E., Zhang, Y., Wang, J., Han, J., 2025. Combined effects of human activity and rainfall shape the differentiated relationships of concentration–discharge in the Yangtze and Yellow River basin, china. Land Degrad. Dev. https://doi.org/10.1002/ldr.70188

Kendall, M.G., 1970. Rank correlation methods, 4th ed. Griffin, London.

Kim, Y.-S., Kim, M.K., Fu, N., Liu, J., Wang, J., Srebric, J., 2025. Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models. Sustainable Cities Soc. 118. https://doi.org/10.1016/j.scs.2024.105570

Lebakula, V., Sims, K., Reith, A., Rose, A., McKee, J., Coleman, P., Kaufman, J., Urban, M., Jochem, C., Whitlock, C., Ogden, M., Pyle, J., Roddy, D., Epting, J., Bright, E., 2025. LandScan global 30 arcsecond annual global gridded population datasets from 2000 to 2022. Sci. Data 12, 495. https://doi.org/10.1038/s41597-025-04817-z

Lehner, B., Messager, M.L., Korver, M.C., Linke, S., 2022. Global hydro-environmental lake characteristics at high spatial resolution. Sci. Data 9, 351. https://doi.org/10.1038/s41597-022-01425-z

Li, W., Wang, X., Song, W., Zhang, Z., Wang, X., Liu, X., Ma, T., Wang, Q., Zhang, Y., Wang, X., Geng, L., 2025. On the contribution of atmospheric reactive nitrogen deposition to nitrogen burden in a eutrophic lake in eastern China. Water Res. 268, 122597. https://doi.org/10.1016/j.watres.2024.122597

Liu, L., Xu, W., Lu, X., Zhong, B., Guo, Y., Lu, X., Zhao, Y., He, W., Wang, S., Zhang, X., Liu, X., Vitousek, P., 2022. Exploring global changes in agricultural ammonia emissions and their contribution to nitrogen deposition since 1980. Proc. Natl. Acad. Sci. U. S. A. 119. https://doi.org/10.1073/pnas.2121998119

Lu, C., Tian, H., 2017. Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: shifted hot spots and nutrient imbalance. Earth Syst. Sci. Data 9, 181–192. https://doi.org/10.5194/essd-9-181-2017

Lundberg, S., Lee, S.-I., 2017. A unified approach to interpreting model predictions. https://doi.org/10.48550/arXiv.1705.07874

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9

Ma, E., Cai, J., Lin, J., Liao, L., Guo, H., Han, Y., 2021. Spatial pattern and water environmental impact of nitrogen and phosphorus emissions from agricultural sources in China in recent 30 years. J. OF NAT. RESOUR. 36, 752–770. https://doi.org/10.31497/zrzyxb.20210316

Meng, X., Feifei, F., Lifeng, W., 2021. Prediction of major pollutants discharge from wastewater in 31 cities of China. Sustainable Prod. Consumption 26, 54–64. https://doi.org/10.1016/j.spc.2020.09.010

Messager, M.L., Lehner, B., Grill, G., Nedeva, I., Schmitt, O., 2016. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. Nat. Commun. 7, 13603.

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

https://doi.org/10.1038/ncomms13603

Monsen, N., Cloern, J., Lucas, L., Monismith, S., 2002. A comment on the use of flushing time, residence time, and age as transport time scales. Limnol. Oceanogr. - LIMNOL OCEAN. 47, 1545–1553. https://doi.org/10.4319/lo.2002.47.5.1545

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nguyen, T.V., Sarrazin, F.J., Ebeling, P., Musolff, A., Fleckenstein, J.H., Kumar, R., 2022. Toward understanding of long-term nitrogen transport and retention dynamics across german catchments. Geophys. Res. Lett. 49, e2022GL100278. https://doi.org/10.1029/2022GL100278

Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., Righetti, M., 2024. Applications of XGBoost in water resources engineering: a systematic literature review (dec 2018–may 2023). Environ. Modell. Software 174, 105971. https://doi.org/10.1016/j.envsoft.2024.105971

Nyenje, P.M., Foppen, J.W., Uhlenbrook, S., Kulabako, R., Muwanga, A., 2010. Eutrophication and nutrient release in urban areas of sub-saharan africa — a review. Sci. Total Environ. 408, 447–455. https://doi.org/10.1016/j.scitotenv.2009.10.020

Qu, J., Dai, X., Hu, H.-Y., Huang, X., Chen, Z., Li, T., Cao, Y., Daigger, G.T., 2022. Emerging trends and prospects for municipal wastewater management in China. ACS ES&T Eng. 2, 323–336. https://doi.org/10.1021/acsestengg.1c00345

Sadayappan, K., Kerins, D., Shen, C., Li, L., 2022. Nitrate concentrations predominantly driven by human, climate, and soil properties in US rivers. Water Res. 226, 119295. https://doi.org/10.1016/j.watres.2022.119295

Sapkota, T.B., Bijay-Singh, 2025. India's fertilizer policies: implications for food security, environmental sustainability, and climate change. Reg. Environ. Change 25. https://doi.org/10.1007/s10113-025-02395-9

Shen, W., Zhang, L., Li, S., Zhuang, Y., Li, W., Du, Y., Qin, B., 2024. Paradox of lake nitrogen concentration change response to watershed management: case study of China. J. Hydrol. 642, 131900. https://doi.org/10.1016/j.jhydrol.2024.131900

Shi, G., Sun, W., Shangguan, W., Wei, Z., Yuan, H., Li, L., Sun, X., Zhang, Y., Liang, H., Li, D., Huang, F., Li, Q., Dai, Y., 2025. A China dataset of soil properties for land surface modelling (version 2, CSDLv2). Earth Syst. Sci. Data 17, 517–543. https://doi.org/10.5194/essd-17-517-2025

Tian, H., Bian, Z., Shi, H., Qin, X., Pan, N., Lu, C., Pan, S., Tubiello, F.N., Chang, J., Conchedda, G., Liu, J., Mueller, N., Nishina, K., Xu, R., Yang, J., You, L., Zhang, B., 2022. History of anthropogenic nitrogen inputs (HaNi) to the terrestrial biosphere: a 5 arcmin resolution annual dataset from 1860 to 2019. Earth Syst. Sci. Data 14, 4551–4568. https://doi.org/10.5194/essd-14-4551-2022

Tian, J., Yuan, Z., Mao, X., Ma, T., 2025. Quantifying natural and anthropogenic impacts on riverine total nitrogen concentration and load in the Yellow River basin. Environ. Pollut. 382, 126641. https://doi.org/10.1016/j.envpol.2025.126641

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

Tong, Y., Huang, Z., Janssen, A.B.G., Wishart, M., He, W., Wang, X., Zhao, Y., 2022. Influence of social and environmental drivers on nutrient concentrations and ratios in lakes: a comparison between china and Europe. Water Res. 227, 119347. https://doi.org/10.1016/j.watres.2022.119347

Vincent, A.E.S., Tank, J.L., Mahl, U.H., 2025. Seasonal patterns in sediment nitrification rates and their linkages to ammonium cycling in three agricultural streams. Biogeochemistry 168. https://doi.org/10.1007/s10533-024-01196-2

Wang, A., Tang, L., Yang, D., Lei, H., 2016. Spatio-temporal variation of net anthropogenic nitrogen inputs in the upper yangtze river basin from 1990 to 2012. Sci. China Earth Sci. 59, 2189–2201. https://doi.org/10.1007/s11430-016-0014-6

Wang, C., Wang, X., Xu, Y.J., Lv, Q., Ji, X., Jia, S., Liu, Z., Mao, B., 2024. Multi-evidences investigation into spatiotemporal variety, sources tracing, and health risk assessment of surface water nitrogen contamination in China. Environmental Research 262, 119906. https://doi.org/10.1016/j.envres.2024.119906

Wang, C., Xu, Y.J., Lv, Q., Ji, X., Jia, S., Liu, Z., Mao, B., 2025a. Nitrogen and phosphorus evolution process and driving mechanisms of three major freshwater lakes with different river-lake connectivity (DRLC) in the lower reaches of the Yangtze River, the largest river in Asia. J. Cleaner Prod. 486. https://doi.org/10.1016/j.jclepro.2024.144471

Wang, J., Uwiragiye, Y., Cao, M., Chen, M., Fallah, N., Huang, Y., Huang, Y., Cheng, Y., Cai, Z., Xu, M., Chang, S.X., Müller, C., 2025b. Global land use change impacts on soil nitrogen availability and environmental losses. Environ. Sci. Technol. 59, 17595–17605. https://doi.org/10.1021/acs.est.5c03285

Wang, X., Dai, Y., Xu, Y.J., Lv, Q., Ji, X., Mao, B., Jia, S., Liu, Z., Luo, C., Rong, Y., 2025c. Dual-quantification of the different contributions of climate change and anthropogenic activities to eutrophication of rivers and lakes in Asia's largest river basin (yangtze river). J. Hazard. Mater. 496. https://doi.org/10.1016/j.jhazmat.2025.139205

Wang, X., Liu, Z., Xu, Y.J., Mao, B., Jia, S., 2025d. Extreme drought affects lake water quality, quantity, morphometry: evidence from China's largest fresh water lake under the 2022 global drought. Geosci. Front. 16. https://doi.org/10.1016/j.gsf.2025.102146

Wetzel, R.G., 2001. Limnology: lake and river ecosystems. Gulf Professional Publishing.

Xie, Y., Liu, Q., Liu, T., Lin, L., Li, R., Du, C., Bian, W., Wang, D., 2026. Predicting the occurrence of the emerging disinfection byproducts halocyclopentadienes and analyzing their driving factors in nationwide tap water using machine learning. Water Res. 288, 124624. https://doi.org/10.1016/j.watres.2025.124624

Xu, M., Wang, X., Chen, K., 2025. Leveraging agricultural production organizations to reduce fertilizer use: evidence from China. Food Policy 133. https://doi.org/10.1016/j.foodpol.2025.102891

Xu, Q., Yan, T., Wang, C., Hua, L., Zhai, L., 2023. Managing landscape patterns at the riparian zone and sub-basin scale is equally important for water quality protection. Water Research 229, 119280. https://doi.org/10.1016/j.watres.2022.119280

Yang, J., Huang, X., 2021. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. Earth Syst. Sci. Data 13, 3907–3925. https://doi.org/10.5194/essd-13-3907-2021

Ye, H., Tang, C., Cao, Y., 2021. Sources and transformation mechanisms of inorganic nitrogen:

https://doi.org/10.5194/egusphere-2026-3574
Preprint. Discussion started: 6 July 2026

evidence from multi-isotopes in a rural-urban river area. Sci. Total Environ. 794. https://doi.org/10.1016/j.scitotenv.2021.148615

Zhang, G., Yao, T., Chen, W., Zheng, G., Shum, C.K., Yang, K., Piao, S., Sheng, Y., Yi, S., Li, J., O'Reilly, C.M., Qi, S., Shen, S.S.P., Zhang, H., Jia, Y., 2019. Regional differences of lake evolution across china during 1960s–2015 and its natural and anthropogenic causes. Remote Sens. Environ. 221, 386–404. https://doi.org/10.1016/j.rse.2018.11.038

Zhang, K., Gillson, L., Jin, Y., McGowan, S., Finch, J., Liu, Z., Han, Y., Shen, J., Meadows, M.E., Taylor, D., 2026. Shifting paradigms: towards dynamic approaches to sustain anthropocene lake ecosystems. Sci. Bull. 71, 196–204. https://doi.org/10.1016/j.scib.2025.11.029

Zhao, L., Fu, D., Li, T., Yuan, X., Wang, S., Liu, C., Duan, C., 2025. Temporal variations of N and P losses via surface runoff from chinese farmland after fertilisation. Soil Tillage Res. 246. https://doi.org/10.1016/j.still.2024.106338

Zhao, Shi, J., Xue, L., Li, W., Zamanian, K., Han, J.-G., Chen, S., 2024a. Water point and non-point nitrogen pollution due to land-use change and nitrate deposition in China from 2000 to 2020. Water 16, 1396. https://doi.org/10.3390/w16101396

Zhao, X., Xu, H., Kang, L., Zhu, G., Paerl, H.W., Li, H., Liu, M., Zhu, M., Zou, W., Qin, B., Zhang, Y., 2024b. Nitrate sources and transformations in a river-reservoir system: response to extreme flooding and various land use. J. Hydrol. 638, 131491. https://doi.org/10.1016/j.jhydrol.2024.131491

Zheng, Y., Jiang, L., Pan, F., Zhang, C., Pan, J., Shi, J., 2026. Spatiotemporal variation of snow cover over the Tibetan Plateau based on the MODIS fractional snow cover product: 2000–2023. J. Hydrol. 664, 134474. https://doi.org/10.1016/j.jhydrol.2025.134474

Zhou, J., Jia, L., Menenti, M., 2015. Reconstruction of global MODIS NDVI time series: performance of harmonic ANalysis of time series (HANTS). Remote Sens. Environ. 163, 217–228. https://doi.org/10.1016/j.rse.2015.03.018

Zhou, K., Wu, J., Liu, H., 2021. Spatiotemporal variations and determinants of water pollutant discharge in the Yangtze River economic belt, china: a spatial econometric analysis. Environ. Pollut. 271, 116320. https://doi.org/10.1016/j.envpol.2020.116320

Zhu, G., Xu, H., Zhu, M., 2019. Changing characteristics and driving factors of trophic state of lakes in the middle and lower reaches of Yangtze River in the past 30. Journal of Lake Sciences 31, 1510–1524.