

YSSP Report
Young Scientists Summer Program

Effects of Environmental Protection Policies on Nutrient Export in the Yangtze River Basin

Jingcheng LI
lijincheng@pku.edu.cn

Approved by

Supervisor: Ting Tang, Christian Folberth
Program: Biodiversity and Natural Resources (BNR)
2021.9.28

This report represents the work completed by the author during the IIASA Young Scientists Summer Program (YSSP) with approval from the YSSP supervisor.

It was finished by 28 September 2021 and has not been altered or revised since.

This research was funded by IIASA and its National Member Organizations in Africa, the Americas, Asia, and Europe.



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).
For any commercial use please contact repository@iiasa.ac.at

YSSP Reports on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the institute, its National Member Organizations, or other organizations supporting the work.

ZVR 524808900

Table of ContentsYSSP Report	i
Young Scientists Summer Program	i
Approved by	i
Abstract	iii
Acknowledgments	iv
About the authors	iv
Effects of Environmental Protection Policies on Nutrient Export in the Yangtze River Basin	1
1. Introduction	1
2. Methods	2
3. Case study	10
4. Results and discussion	13
5. Conclusion	18
Reference	20
Support information	22
Nutrient production of 0.1° grid	22

Abstract

High resolution source data are pivotal for the simulation of nutrient production and fluxes in a watershed. This study constructed the CEIN (China emission inventory of nutrient) Model framework based on county data in China. Using the Yangtze River Basin as a case study, this study calculated the nutrient production and flux at the fine sub-basin resolution (level 8, from HydroBasin), analyzed the contribution of each source, and evaluated the impact of environmental protection policies on nutrient flux. The results illustrated that the CEIN Model framework could accurately calculate nutrient production and flux with R^2 greater than 0.8. Compared to the nutrient production, environmental processes reduced TN and TP by more than 85% and 90% respectively for nutrient fluxes. The contributions of diffuse sources were much higher than those of point sources and direct sources. The contribution of cropland and livestock to nutrient production and flux were largest, while the contribution of freshwater aquaculture and urban non-point source couldn't be ignored, especially at the local scale. The effect of increasing livestock manure recycling rate on nutrient flux was about 100 times that of improving discharge water quality of sewage treatment facilities. The environmental protection policies of points sources would require substantial investments for comparatively small gains, which suggests the focus should be put on non-point sources.

Acknowledgments

I would first like to acknowledge the insightful feedback and support provided by my supervisor, Dr. Tang and Dr. Chirs in Biodiversity and Natural Resources Program at IIASA that helped shape this research project over three months. I would also like to thank Dr. Brith for organizing and moderating our virtual calls and presentations and fostering a friendly and welcoming atmosphere for YSSPers.

About the authors

Jincheng LI graduated from Yunnan University with a BSc and MSc in Environmental Science. He is currently a second year PhD student in Water Science Lab of Prof. LIU Yong at Peking University. The title of his PhD thesis is Simulation of Environmental Processes in the Yangtze River Basin. His areas of research interest are simulation of watershed environmental processes, machine learning, climate change and policy evaluation. (Contact: lijincheng@pku.edu.cn)

Effects of Environmental Protection Policies on Nutrient Export in the Yangtze River Basin

Jincheng LI

1. Introduction

Under global warming, increasing extreme rainfall and anthropogenic pollutant sources, the global surface water quality is facing the risk of continuous deterioration (Alcamo et al., 2003). Terrestrial nutrient source management is an effective way to control eutrophication and countries take many actions to control water pollution. Nonpoint source management plays an important role in the control of terrestrial nutrient production (Bouwman et al., 2009). As the most populous developing country, the water quality in China is worsening. Most recent investigations show that the water quality of large natural lakes is mostly in the state of severe water eutrophication, and the concentration of chlorophyll (Chl) a is much higher than the eutrophication standard (Chl a >7 mg/L). Load reduction is effective for water quality improvement, in which accurate pollutant emission inventory plays a key role in facilitating water quality modeling and consequently informed watershed decision making (Diaz et al., 2008).

Many watershed models have been proposed in previous studies, such as MARINA, Global News and NANI/NAPI (Mayorga et al., 2010; Strokal et al., 2014a; Strokal et al., 2014b). The models have been widely used in global, national and watershed scale modelling. However, they are suffering from the challenges of inaccurate data on point and nonpoint sources, especially dealing with the simulations at national or regional level. Ma (2019) calculated the nutrient export from China with a 0.5° grid, equivalent to the area of a county in China. This renders it challenging to distinguish the spatial heterogeneity in one country. Goyette (2018) calculated nitrogen and phosphorus exports by NANI/NAPI models for 76 watersheds of St. Lawrence over 110 years; while only 23 watersheds meet the statistical standards due to the low-resolution data sources. In recent years, with the release of more and more data with high spatial and temporal resolutions on land use, population, industries and sewage treatment facilities, it is possible to develop much more accurate inventories for different pollutants. For example, many research institutions provided global land use data with 30-meter resolution recently. In addition, the wide use of high-performance computing has greatly shortened the time required for complex geographic information data processing. Although various studies exist on sub-basin analyses of nutrient export in large watersheds for different scenarios, many details are still neglected for the nutrient sources. Direct discharges

of aquaculture to waterbodies are not considered in the MARINA Model. Manufacturing nutrient from unconnected sewage treatment facilities was not considered in the WaterQual model. The nutrient production coefficients of human waste were not differentiated among regions with different development levels and did not consider the impact of environmental protection measures, such as China's toilet revolution, in the Global NEWS-2 model (Mayorga et al., 2010). This underpins the need to develop higher resolution nutrient emission inventories and consider these previously mentioned details to support watershed nutrient control.

This study developed a China Emission Inventory of Nutrients (CEIN) at a 0.1° grid resolution based on county level data. The Yangtze River Basin was taken as a case study. The IIASA Water Program had made some studies of nutrient export in the Yangtze River Basin based on GLOBIOM, CWatM and other models in recent years. The experience in developing nutrient budget and water quality models greatly benefited the study. This study aims to answer four main questions:

1. Is the CEIN Model accurate in simulating nutrient production and flux in the watershed with higher resolution of nutrient sources?
2. What are the spatial characteristics of nutrient production under higher resolution of nutrient sources?
3. What are the spatial characteristics of nutrient export under higher resolution of nutrient sources?
4. What are the responses of nutrient export to different environmental policies?

2. Methods

In this study, the method was divided into three parts, as shown in Fig.1. Part I was to define and collect information and data on data sources and environmental policy parameters. Part II was the development of the CEIN Model for nutrient production. Part III addressed environmental nutrient turnover processes which included the calculation of nutrient loss, removal and flux, including parameter calibration.

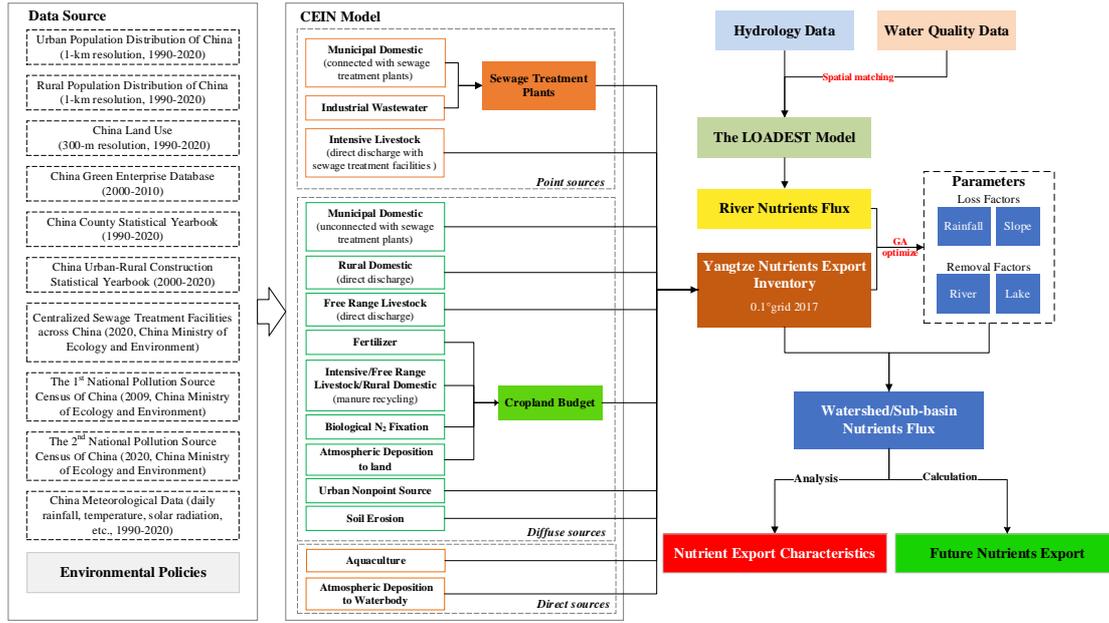


Fig.1 The method framework of the study

2.1 CEIN Model

The CEIN Model is a China emission inventory of nutrients with 0.1° grid which is based on county-level data. The inventory is finer and more accurate than the current 0.5° grid nutrient emission data. It can be coupled with other models such as CWATM and GLOBIOM. In this study, the CEIN Model was coupled with a river nutrient flux model and automatic optimization algorithm. The latter takes river nutrient flux as the optimization objective and optimizes the parameters in the loss function within a reasonable range. In the CEIN Model, the nutrient sources were divided into three categories: diffuse sources, point sources and direct sources. In this study, the nutrient of point and diffuse sources must pass through the landscape to enter rivers. So, we created direct sources to present the nutrient that directly enter waterbodies.

2.1.1 Diffuse sources

$$DFS_n = RD_n + BUG_n + FRL_n + UNS_n + ADL_n \quad (1)$$

where DFS_n is the nutrient emission of diffuse sources. RD_n , BUG_n , FRL_n , UNS_n is the nutrient production of rural domestic, cropland, free range livestock and urban non-point source, respectively. Their unit is t/a. n denotes different nutrient elements, representing TN, and TP respectively.

① Rural Domestic Wastewater

$$RD_n = 3.65 \times POP_r \times [CT_r^n \times R_T + CO_r^n \times (1 - R_T)] \quad (2)$$

where POP_r is the rural population (10^4 per). CT_r^n is the production coefficient of rural domestic population with water closets (g/(per·d)). CO_r^n is the nutrient production coefficient of rural domestic without water closets (g/(per·d)). R_T is the population fraction equipped with water closets (%). In the study, there are 365 days in 2017.

② Cropland Soil Budget

$$BUG_N = N_{fert} + N_{man} + N_{fix} + N_{dep} - N_{withdr} - N_{vol} \quad (3)$$

$$BUG_P = P_{fert} + P_{man} - P_{withdr} \quad (4)$$

where N_{fert} , P_{fert} are the synthetic fertilizers for N and P respectively. N_{man} , P_{man} is the animal manure input to cropland for N, P respectively. N_{fix} is the biological N fixation; N_{dep} is the atmospheric N deposition; N_{withdr} , P_{withdr} is the N, P withdrawal from the field through crop harvesting, hay and grass cutting, and grass consumed by grazing animals respectively. The unit is t/a.

③ Free Range Livestock

$$FRL_n = 3.65 \times 10^{-4} \times FRL_m \times CO_m^n \quad (5)$$

where FRL_m is the raising quantity of m (ind). CO_m^n is the production coefficient of free-range livestock (g/(ind·d)). m is the species of livestock which includes cattle, sheep, horses, broilers, laying hens, ducks and geese.

④ Urban Non-point Source

$$UNS_n = URC_n \times FLOW_u \times 10^{-6} \quad (6)$$

where URC_n is the nutrient concentration of urban runoff which comes from references (mg/L). $FLOW_u$ is the urban runoff (m^3/a) which was sourced from **CWatM (IIASA)**.

⑤ Soil Erosion

$$SE_n = A \times CS_n \times Area \times 10^{-3} \quad (7)$$

$$A = R \times K \times LS \times C \times P \quad (8)$$

where CS_n is the background content of adsorbed nutrient in topsoil (g/kg). $Area$ is the bare land area of grid (km^2). Based on the RUSLE (Han et al., 2019; Chao et al., 2021), the soil erosion was estimated. A represents the annual amount of soil erosion ($t/(hm^2 \cdot a)$). R represents the rainfall erosivity factor ($MJ \cdot mm / (hm^2 \cdot h \cdot a)$). K represents the soil erodibility factor ($t \cdot hm^2 \cdot h / (MJ \cdot mm \cdot hm^2)$). LS represents the topographic factor. C represents the land cover

factor, and P represents the conservation practice factor. The LS , C , and P factors are dimensionless.

The rainfall erosivity factor is calculated by

$$R = \sum_{i=1}^{12} \left(1.735 \times 10^{\left(1.5 \times \log_{10}(p_i^2/p) - 0.8188\right)} \right) \quad (9)$$

where p_i and p represent average monthly precipitation and average annual precipitation. The soil erodibility factor is calculated by

$$K = \left\{ 0.2 + 0.3 \exp \left[-0.0256 SAN (1 - SIL / 100) \right] \right\} \\ \times \left(\frac{SIL}{CLA + SIL} \right)^{0.3} \times \left(1 - \frac{0.25 OC}{OC + \exp(3.72 - 2.95 OC)} \right) \\ \times \left\{ 1 - \frac{0.7 \times (1 - SAN / 100)}{(1 - SAN / 100) + \exp[-5.51 + 22.9 \times (1 - SAN / 100)]} \right\} \quad (10)$$

where SAN , SIL , CLA , and OC denote the percentages of sand, silt, clay, and organic carbon, respectively. The final calculation results need to be multiplied by 0.1317 and converted into international units.

$$L = \left(\frac{l}{22.13} \right)^m \\ m = \frac{\beta}{1 + \beta} \quad (11)$$

$$\beta = \frac{\sin \theta / 0.0896}{3 \times (\sin \theta)^{0.8} + 0.56}$$

$$S = \begin{cases} 10.8 \times \sin \theta + 0.03 & \tan \theta < 0.09 \\ 16.8 \times \sin \theta - 0.50 & \tan \theta \geq 0.09 \end{cases} \quad (12)$$

where λ represents slope length; β is a coefficient of change with slope; and θ is slope angle.

The land cover factor is calculated by

$$C = \begin{cases} 1 & fc = 0 \\ 0.6508 - 0.3436 \log_{10}(fc) & 0 < fc \leq 78.3\% \\ 0 & fc > 0 \end{cases} \quad (13)$$

$$fc = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}$$

where $NDVI_{max}$ is the maximum values of NDVI; $NDVI_{min}$ is the minimum values of NDVI; fc represents normalized vegetation cover degree.

The P factor was set as paddy field 0.15, dryland and orchard 0.40, forest and grassland 1.00, and water and built-up land 0 respectively (Han et al., 2019; Chao et al., 2021).

2.1.2 Point sources

$$PS_n = MD_n + STF_n + IED_n + IL_n + AC_n \quad (14)$$

where PS_n is the nutrient emission of point sources (t/a). MD_n , STF_n , IED_n , IL_n , AC_n is the nutrient production of municipal domestic wastewater not collected or connected by sewage facilities, sewage treatment facilities, industrial enterprises not collected or connected by sewage facilities, intensive livestock, aquaculture. The unit is $t \cdot a^{-1}$. The nutrient emission of sewage treatment facilities includes the connected municipal domestic wastewater and connected industrial enterprises.

① Unconnected Municipal Domestic Wastewater

$$MD_n = 3.65 \times POP_u \times CO_{md}^n \times (1 - SC) \quad (15)$$

$$MDS = 3.65 \times 10^3 \times POP_u \times CO_{md}^s \times SC \quad (16)$$

where POP_u is the urban population (10^4 person). CO_{md}^n is the production coefficient of N and P in municipal domestic wastewater (g/(person·a)). SC is the collection rate of municipal domestic sewage (%). MDS is the emission of municipal domestic sewage (m^3/a). CO_{md}^s is the emission coefficient of municipal domestic sewage (L/(per·a)).

② Sewage Treatment Facilities

Most of the previous studies took the designed treatment capacity of sewage treatment facilities as the actual treatment volumes, and the discharge water quality standard as the actual nutrient concentration or estimated the two parameters through the food chain, which would bring great errors (Bouwman et al., 2009; Mayorga et al., 2010; Strokal et al., 2014a; Strokal et al., 2014b). This study used sewage treatment facility monitoring data, including actual sewage treatment volumes and nutrient concentrations.

$$STF_n = AST \times CC_n \times 10^{-6} \quad (17)$$

where AST is the actual sewage treatment volume (m^3/a). CC_n is the emission concentration of nutrient (mg/L).

③ Unconnected Industrial Enterprises

$$MGIE_n = f_n(MGIP) \quad (18)$$

$$IEP_n = f_n(GIOV) \quad (19)$$

$$IED_n = IEP_n \cdot \left(1 - \frac{AST - MDS}{IES} \right) \quad (20)$$

where MIE_n is the nutrient production of municipal industrial enterprises from Green Enterprise Database. $MGIP$ is the gross industrial production of green enterprise. In general, $f_n(x)$ is a power function of different nutrients that is fitted by MIE_n and $MGIP$ of different cities. $GIOV$ is the municipal gross industrial production. IEP_n is the nutrient production of municipal industrial enterprises. IED_n is the nutrient production of industrial enterprises not collected or connected by sewage facilities (t/a). IES is the production of industrial enterprises sewage (m^3/a).

④ Intensive Livestock

$$IL_n = 3.65 \times 10^{-4} \times IL_m \times CO_m^n \quad (21)$$

where IL_m is the raising quantity of intensive livestock type m (ind). CO_m^n is the nutrient production coefficient of intensive livestock ($g/(ind \cdot d)$). m is the species of livestock which includes cattle, sheep, horses, broilers, laying hens, ducks and geese.

2.1.3 Direct sources

Nutrient produced from direct sources enters waterbodies directly, including atmospheric deposition to waterbodies and freshwater farming in this study.

① Atmospheric Deposition to Waterbodies

$$ADW_n = APU_n \times Area_w \times 10^{-6} \quad (22)$$

where APU_n is the unit area atmospheric deposition ($kg/(km^2 \cdot a)$) which is determined by references. $Area_w$ is the waterbody area of grid (km^2). This study only calculated the atmospheric deposition for nitrogen.

② Freshwater Aquaculture

$$FA_n = 3.65 \times 10^{-4} \times FA_g \times CO_g^n \quad (23)$$

where FA_g is the raising quantity of aquaculture n (ind). CO_g^n is the nutrient production coefficient of aquaculture g ($g/(ind \cdot d)$).

2.1.4 Nutrient loss function

$$L_n = \sum_{i=1}^o DIS_i^n + \lambda_i^n \left[(\alpha_i \cdot \beta_i^n \cdot DFS_i^n) + SE_i^n + PS_i^n \right] \quad (24)$$

where L_n is the nutrient loss at the subbasin level. λ is the removal coefficient of the landscape. α is the spatial unevenness impact factor of precipitation. β is the terrain impact factor. i is the number of grid cell. o is the amount of sub-basin grids.

$$\alpha = \frac{R_i}{\bar{R}} \quad (25)$$

where R_i is the precipitation of grid cell i . \bar{R} is the average precipitation of sub-basin.

$$\beta = \left(\frac{\theta_i}{\bar{\theta}} \right)^a \quad (26)$$

where θ_i is the slope of grid cell i . $\bar{\theta}$ is the average slope of sub-basin. a is the parameter of regression.

2.1.5 Nutrient removal function

The nutrient removal function of waterbodies was adopted from the MARINA Model and key parameters were optimized in this study. The function is represented as follows:

Removal by Reservoirs (Lakes):

$$\tau_j = V_j / Q_j \quad (27)$$

$$R_{N,j} = b \cdot (H_j / \tau_j)^c \quad (28)$$

$$R_{P,j} = d \cdot [1 - \exp(-e \cdot \tau_j)] \quad (29)$$

$$R_k = \frac{1}{Q} \sum_{j=1}^w (Q_j \cdot R_{N/P,j}) \quad (30)$$

where τ_j , H_j , V_j are the water residence time, depth, volume of reservoir (lake) j respectively. $R_{N,j}$, $R_{P,j}$ is the N, P removal of reservoir (lake) j respectively. Q , R_k are the flow, reservoir (lake) nutrient removal of sub-basin k respectively.

Removal by River Reaches:

$$L_k = f \cdot \ln(\text{Area}_k) - h \quad (31)$$

Where Area_k , L_k are the drainage area, nitrogen removal in the river reaches of sub-basin k respectively.

Removal by Water Withdrawal:

$$FQrem_k = 1 - Q_{act,k} / Q_{nat,k} \quad (32)$$

Where $FQrem_k$, $Q_{act,k}$, $Q_{nat,k}$ are the water consumption nutrient removal, the actual and natural water discharge at the outlet of sub-basin k respectively.

b , c , d , e , f , h are the parameters of regression.

2.2 River nutrient flux

Using hydrological time series data and water quality concentration in CEIN, LOADEST (LOAD ESTimator) can establish a regression model of river component load. Based on this

model, monthly or seasonal assessment loads, standard deviations, and 95% confidence intervals can be estimated. Based on the measured index concentration, discharge and monitoring date, the model first establishes the regression relationship between the index concentration and the other two variables, and then estimates the river load in other periods without water quality monitoring but with discharge. River nutrient flux was the objective of parameters optimization in this study.

Three main LOADEST models as followed:

$$\hat{L}_{MVUE} = \exp\left(a_0 + \sum_{j=1}^M a_j X_j\right) g_m(m, s^2, V) \quad (33)$$

$$\hat{L}_{AMLE} = \exp\left(a_0 + \sum_{j=1}^M a_j X_j\right) H(a, b, s^2, \alpha, \kappa) \quad (34)$$

$$\hat{L}_{LAD} = \exp\left(a_0 + \sum_{j=1}^M a_j X_j\right) \frac{\sum_{k=1}^n \exp(e_k)}{n} \quad (35)$$

where \hat{L}_{MVUE} is the instantaneous load estimated by maximum likelihood estimation model. m is the degree of freedom. s^2 is the residual variance. V is the equation of explanatory variables. The coefficients, a_0 and a_j , are estimated by maximum likelihood method. X_j is the explanatory variable. $g_m(m, s^2, V)$ is the deviation correction factor.

2.3 Parameter calibration

A genetic algorithm was applied to optimize the parameters in the nutrient loss and removal functions in this study. In a genetic algorithm, a population of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem is evolved toward better solutions. Each candidate solution has a set of properties (its chromosomes or genotype) which can be mutated and altered; traditionally, solutions are represented in binary form as strings of 0s and 1s, but other encodings are also possible. R^2 and RMSE were applied to evaluate the accuracy of parameters optimization, as followed:

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \quad (36)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (37)$$

where y_i and \hat{y}_i is the actual and simulated nutrient flux. m is the number of monitoring station.

Table1 The range of parameters optimization

Coefficients	Related Variable	Range
a	—	(0, 1)
b, c	$R_{N,j}$	(0, 0.96)
d, e	$R_{P,j}$	(0, 0.85)
f, h	L_k	(0, 0.65)

2.4 Environmental policy scenarios

Two scenarios were evaluated in the study. Scenario I was to improve the effluent water quality standard of sewage treatment facilities from level II and 1B to level IA. The water quality discharged from sewage treatment facilities is regulated by “Discharge Standards of Pollutants for municipal wastewater treatment plant (GB18918-2002)”. Nutrient concentrations at different levels of standards were shown in Table 2. Most of the sewage treatment facilities built before 2005 comply with Level II and Level 1B. Most of the sewage treatment facilities built after 2005 are more advanced and therefore comply with Level 1A. In recent years, municipal wastewater treatment facilities were retrofitted to reduce nutrient discharges in China.

Scenario II was to improve the recycling of livestock manures to replace mineral fertilizer use in cropland. In 2017, the recycling rate of livestock waste in China was about 50%. The Ministry of Ecological Environment & Ministry of Agriculture and Rural Affairs enacted the “Action Plan to fight pollution in agriculture and Rural areas” in 2018. It requests the comprehensive recycling rate of livestock manure to be above 75 % by 2020. Therefore, the recycling rate of livestock manure was set to be from 50% to 75% to replace the same amount of fertilizer use in cropland in scenario II.

Table2 Effluent nutrient concentrations allowed at different levels of WW treatment (mg/L)

Levels	TN	TP	NH ₃ -N	COD	BOD ₅
Level IA	15	0.5	5	50	10
Level IB	20	1	8	60	20
Level II	-	3	25	100	30

3. Case study

3.1 Study Area

The Yangtze River Basin is the largest basin in China and the third largest in the world, covering 1.8 million km² area, accounting for 18.8% of China's area. The average annual precipitation and temperature is 1067 mm and 16°C respectively. Due to the vast territory and complex terrain, the spatial and temporal distribution of annual mean precipitation and annual mean temperature is highly uneven, while governed by monsoon climate. The Yangtze River is China's most abundant river with 975.5 billion m³ of water resources, accounting for 36% of the country's total river runoff and ranking third in the world.

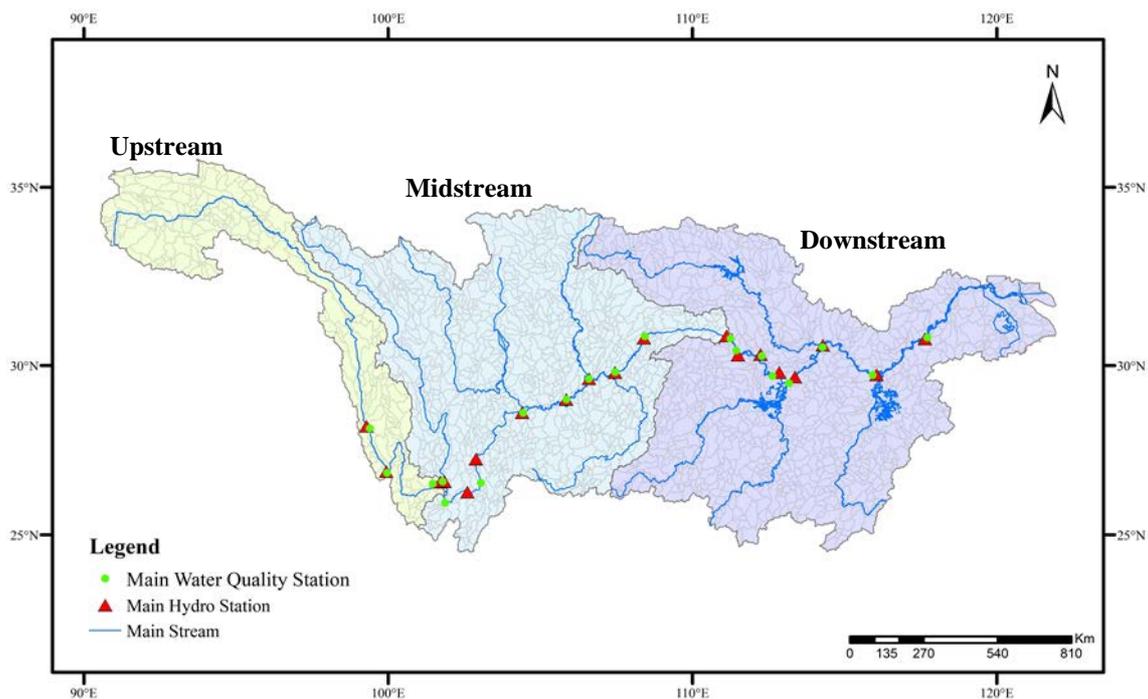


Fig2. The subbasin, main stream, main hydrological stations, main water quality stations, and upstream, midstream and downstream of Yangtze River Basin

3.2 Data Source

The Yangtze River basin delineation is from HydroBASINS Version 1.0 which has 2913 subbasins in level 8. The soil components of sand, silt, clay and organic carbon is from HWDS dataset. The background content of adsorbed TN and TP in topsoil is from Institute of Tibetan Plateau Research. The data of DEM, land use, NDVI is from Resource and Environment Science and Data Center.

The daily runoff and 735 hydrological stations were from China Hydrological Yearbook. The daily weather data and 717 stations information were from National Meteorological Science Data Center. The water quality and 679 monitoring stations information were from China National Environmental Monitoring Center. The industrial and corporate related data is

from China Green Enterprise Database. The sewage treatment facilities related data is from Ministry of Ecology and Environment of the People’s Republic of China. The production and emission coefficients of rural and urban population, livestock and aquaculture were from the 1st and 2nd National Pollution Source Census. The nutrient sources data were from China County Statistical Yearbook.

Table3 The source and related information of data

Category	Year	Resolution Ratio	Source
Digital Elevation Model (DEM)	2017	30 m	Chinese Academy of Sciences
Land cover	2017	30 m	Chinese Academy of Sciences
Normalized Differential Vegetation Index	2017	1 km	Chinese Academy of Sciences
Net Primary Productivity of vegetation	2017	1 km	Chinese Academy of Sciences
Soil physical properties	History	1 km	Nanjing Soil Science Institute
Soil chemical compositions	History	1 km	China Soil Database
China atmospheric nitrogen deposition	2017	1 km	Chinese Academy of Sciences
China water system	History	China	Chinese Academy of Sciences
Hydrological station (HS)	735	flow, depth, area...	Water Yearbook
Weather station (WS)	717	prec, tem...	China Meteorological Bureau
Water quality station (WQS)	679	NH ₃ -N, TN, TP...	China Monitoring Station
Lakes/Reservoirs	7502	volume, depth, Res...	HydroLakes
Administrative boundary	957	province, city, county...	Chinese Academy of Sciences
County statistics data	Millions	boundary, economic...	China Administer

4. Results and discussion

4.1 The nutrient production of Yangtze River Subbasin

In 2017, the production of TN and TP was 11.45 Mt, 4.67 Mt respectively in the Yangtze River Basin. The proportion of each source is shown in Fig.3 a and b. Diffuse sources had the largest proportion accounting for 99.13% of TN and 96.92% of TP. The proportion of point sources was relatively small at 2.52% and 0.73% of TN and TP. Direct sources had the smallest proportion which accounted for 0.56% of TN and 0.14% of TP. Cropland and aquaculture were the largest and smallest source for nutrient production, respectively. The distribution of livestock, soil erosion was second for TN, TP production respectively. The TN production of urban non-point source and sewage treatment facilities was approximately equal and the TP production of urban non-point sources was nearly twice that of sewage treatment facilities. Although the proportion of urban non-point sources was not high, it could not be ignored in watershed nutrient management. The contribution of different sources to nutrient production in the upstream, midstream and downstream of the Yangtze River was greatly different. Overall, cropland remained the largest source. Soil erosion was mainly concentrated in the upstream, because there are higher fractions of bare land, mountainous areas and steep slopes. There was little soil erosion in the midstream and downstream, so the contribution was minimal. Due to the large population and high urbanization rate in the downstream, the source of municipal domestic and urban non-point was relatively large. The livestock in the midstream was larger and less intensive than that in the downstream, so the proportion of nutrient production in the livestock was larger in the midstream. The freshwater aquaculture was developed with flat terrain, large water area, and numerous lakes, so its relative contribution in the downstream was higher than that in the midstream and upstream.

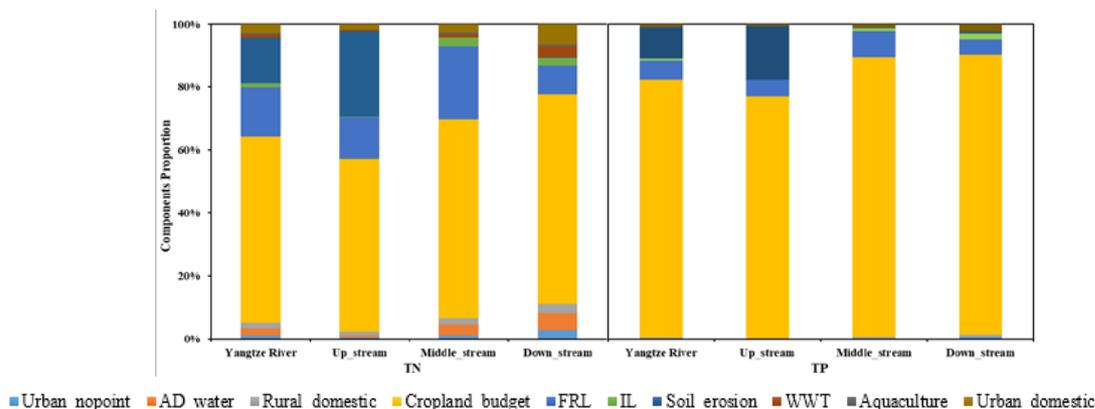


Fig3. The sources contribution of nutrient production for whole, upstream, midstream and downstream of Yangtze River Basin

The TN and TP production per unit area of the 2193 sub-basins in Yangtze River basin is shown in Fig4 which illustrates the hot spots. The spatial distribution of both nutrients was particularly similar. The maximum production of per unit area reached 65.10 and 22.50 t/km² for TN and TP which were located at the Yangtze River Delta. This is one of the most developed regions for economics in China. Another hot spot was the central Yangtze River Basin. With a lowland, flat terrain, fertile soil and large population, it was the most developed region in central China in both economy and agriculture. Compared with the hot spot distribution of TP, the distribution of TN was more discrete and had more secondary hot spots. In the upstream, most sub-basins were cold spots for TP, while some sub-basins were secondary hot spots for TN. Topography was the most important factor affecting the distribution of hot spots. In the upstream and midstream, hot spots concentrated in plain areas because mountainous regions are less used and therefore less influenced by anthropogenic nutrient inputs. In the downstream, hot spots were distributed along the waterbody because of broad plain.

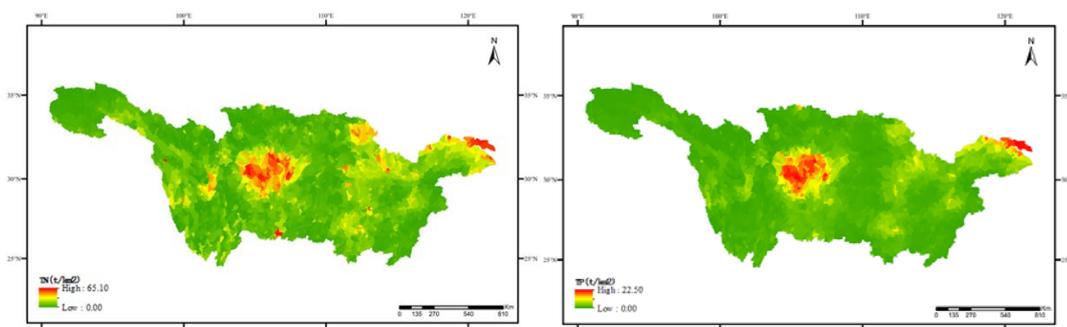


Fig4. Spatial distribution of nutrient production per unit area in 2193 sub-basins

4.2 Nutrient flux and parameter calibration

Based on daily flow and weekly water quality data, the LOADEST model was used to calculate monthly fluxes at the upstream, midstream and downstream boundary points of the Yangtze River Basin, as shown in Fig.5. Nutrient flux was obviously seasonal which was more evident in the upstream than in the midstream and downstream. The annual net TN fluxes in the upstream, midstream and downstream were 35,551, 654,756 and 1,631,732 t/a, respectively. These fluxes refer to the net fluxes of the respective areas. That is, the upstream input flux is extracted from midstream fluxes. The annual TP fluxes in the upstream, midstream

and downstream were 1,205, 39,285 and 91,902 t/a, respectively. The scatter diagram represented the accuracy of equation parameter optimization, and the dotted line was the 1:1 of observed and the simulated values. For TN flux, the R^2 and RMSE were 0.72 and 3150 in upstream, 0.70 and 53764 in midstream, 0.81 and 48597 in the downstream site, respectively. For TP flux, the R^2 and RMSE were 0.84 and 44 in upstream, 0.80 and 2395 in midstream, 0.88 and 2920 in the downstream site, respectively. R^2 was greater than 0.5, which could meet the requirement of parameters optimization accuracy and the optimization results in the downstream were better than those in the upstream and midstream. The optimization results of TP flux were better than that of TN flux.

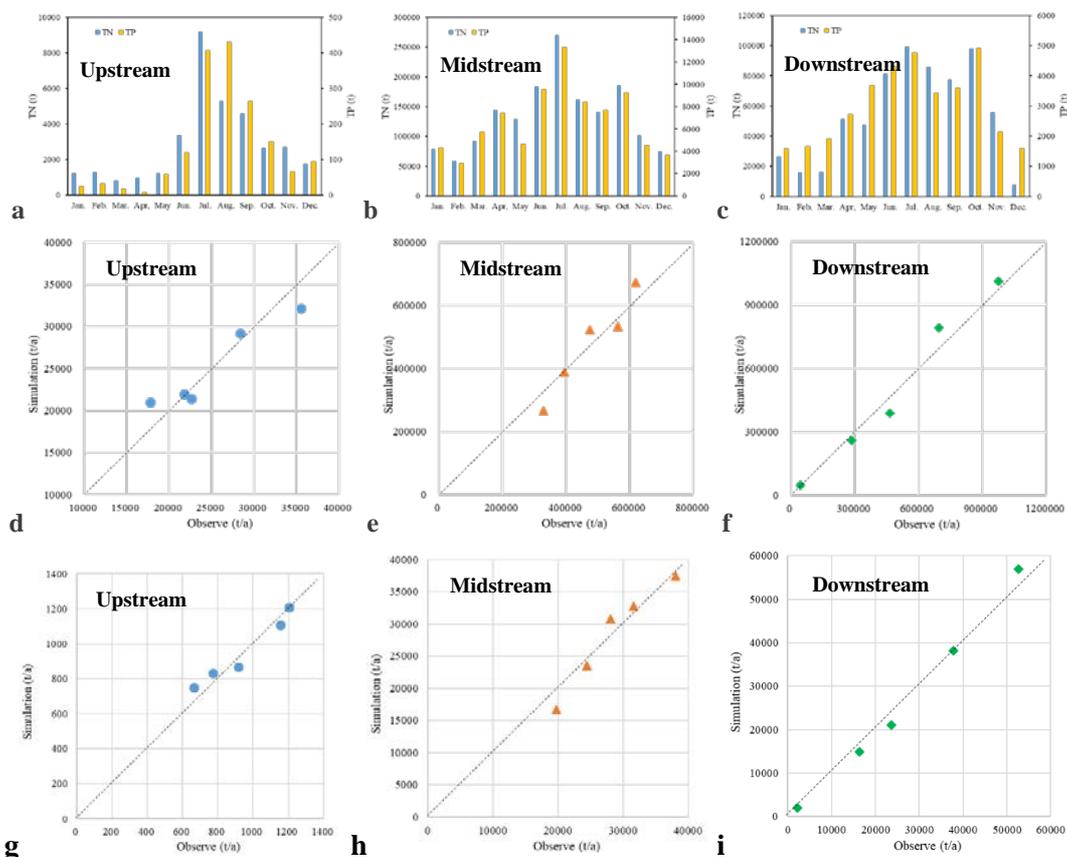


Fig5. Nutrient flux of monitoring sites for the upstream, midstream and downstream of the Yangtze River Basin. a, b, c was the nutrient flux at boundary points. d, e, f was the simulated and observed TN flux of upstream, midstream and downstream. g, h, i was the simulated and observed TP flux of upstream, midstream and downstream.

4.3 The source contribution to nutrient fluxes

The contribution proportion of each source for nutrient flux is shown in Fig.6 a and b. The contributions of diffuse, point and direct sources to TN and TP flux were similar to TN and TP production, about 95%, 4% and 1%. Cropland was the largest source contribution, followed by aquaculture and soil erosion, which was similar to the source contribution of nutrient

production but not in the same proportion. Due to freshwater aquaculture and part of atmospheric nitrogen deposition directly into the waterbody, their contribution was increased compared to that of nutrient production. With the soil particle sedimentation, the contribution of soil erosion was reduced and the contribution decreased on TN flux was considerably higher than that of TP flux. The contribution of livestock to TN flux was nearly twice that of TP flux. With the environmental process, the contribution of urban non-point sources to nutrient flux decreased and became the smallest contribution source. The source contributions in the upstream, midstream and downstream to nutrient production and flux were vastly different, especially in the upstream. The contributions of each source for nutrient flux of upstream, midstream and downstream are shown in Fig.6 c and d. The contribution of cropland to nutrient flux in upstream was smaller than that in midstream and downstream because there are mainly mountainous and grassland and little cropland area. The livestock in downstream was highly intensive, and contributed the most to nutrient flux.

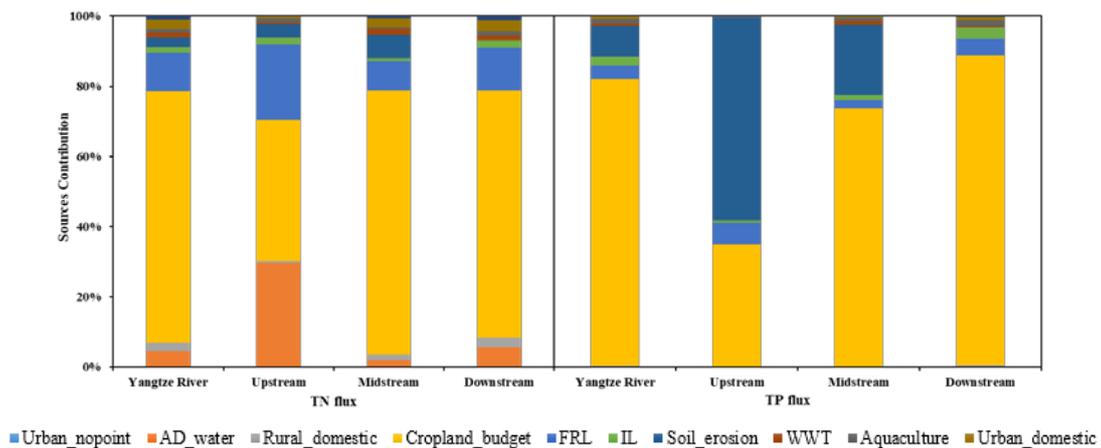


Fig6. The sources contribution of net nutrient flux for whole, upstream, midstream and downstream of Yangtze River Basin. AD_water represents the atmospheric deposition to waterbodies. FRL represents the free-range livestock. IL represents the intensive livestock. WWT represents the wastewater treatment.

The TN and TP net flux per unit area of 2193 sub-basin is shown in Fig7, which again illustrates the hot spots. Their spatial distribution was similar to the nutrient production but the central Yangtze River Basin changed from hot spots to secondary hot spot. The maximum flux of per unit area reached 12.02 and 1.12 t/km² for the TN and TP at the Yangtze River Delta. The amount of TN and TP decreased by 81.54% and 95.02% respectively due to environmental processes, which means the environmental removal of TP was generally greater than that of TN in Yangtze River Basin.

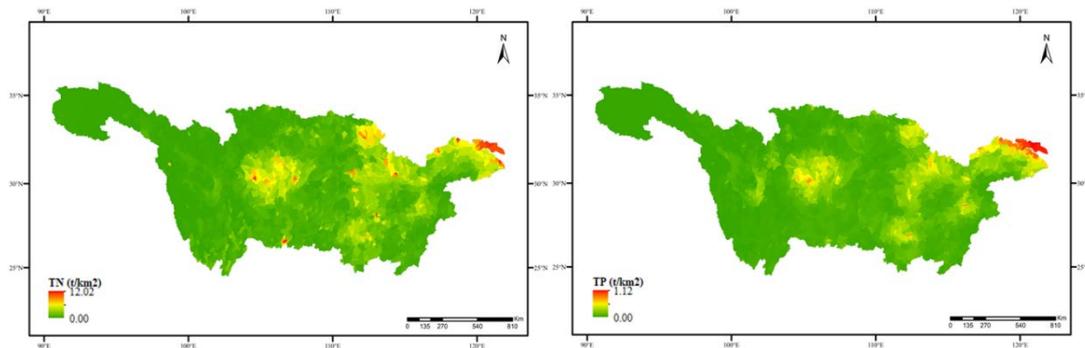


Fig7. Spatial distribution of nutrient net flux per unit area in 2193 sub-basins

4.4 The nutrient flux sources contribution under different scenarios

For scenario I with improvement in wastewater treatment and effluent standard, the TN and TP fluxes were reduced by 1,197 t and 63.19 t, accounting for 0.07% and 0.067% in the Yangtze River Basin. The TN and TP flux were only reduced by 4 t and 0.04 t, accounting for 0.01% and 0.003% in the sparsely populated upstream. In the midstream, the TN and TP flux were reduced by 552 t and 37.47 t which accounted for 0.08% and 0.101%. In the downstream, the TN and TP flux were reduced by 641 t and 25.68 t which accounted for 0.06% and 0.046%. Compared with the midstream, although there were more sewage treatment facilities in the downstream, the discharge standards were higher, so there was little difference in nutrient reduction between them. Because of the large downstream pollutant flux, the reduction accounted for less. Overall, scenario I entails substantial costs, but has weak effects on nutrient flux reduction for the Yangtze River Basin.

For scenario II with improved manure reuse on cropland, the TN and TP flux were reduced by 193,343 t and 4,969 t, accounting for 11.31% and 5.27%. The reduction of TN and TP fluxes in scenario II was 162 and 79 times of that in scenario I, respectively. In the upstream, the reduction of TN and TP flux were 7,517 t and 77 t, accounted for 21.16% and 6.39%. The economy of upstream is based on agriculture, so nutrient flux reduction accounted for the largest proportion. The reduction of TN and TP flux were 59,496 t and 1,295 t, accounted for 8.87% and 3.49% in the midstream. In the downstream, the reduction of TN and TP flux were 126,330 t and 3,597 t, accounted for 12.54% and 6.44%. With the highest degree of agricultural intensification, the nutrient production of livestock and fertilizer application and nutrient flux reduction in downstream were several times that in the midstream. Overall, the nutrient reduction in scenario II was nearly hundredfold that in scenario I, but the cost was relatively low. This illustrates that the current nutrient control measures for non-point sources, especially for cropland, were better than point sources from the perspective of cost-benefit.

The contribution of various sources to nutrient fluxes under scenario II is shown in Figure 8. There was little change in the proportion of nutrient contributions from different sources. The replacement of fertilizer by livestock manure resulted in a decrease in the total nutrient production, so the proportion of cropland contribution increased, and the livestock decreased. The nutrient production of other sources remained unchanged, so the contribution proportion of nutrient flux increased slightly. The hot spots of nutrient flux in sub-basin are shown in Fig9. The maximum TN and TP flux per unit area decreased by 0.12 and 0.01kg/km², respectively. Scenario II had weak effects on the spatial distribution of nutrient flux per unit area, and hot spots were still mainly located at the Yangtze River Delta. The number of secondary hot spots decreased slightly.

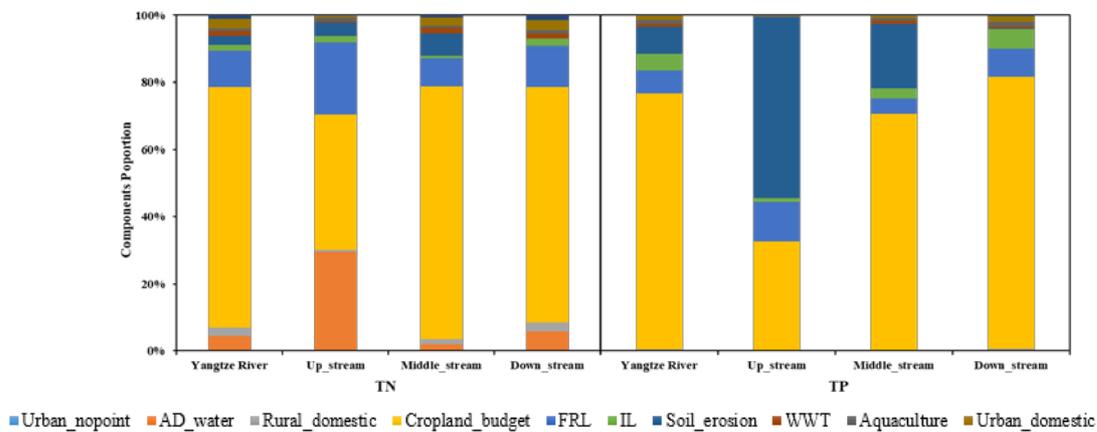


Fig8. The sources contribution of nutrient net flux under scenario II for whole, upstream, midstream and downstream of Yangtze River Basin

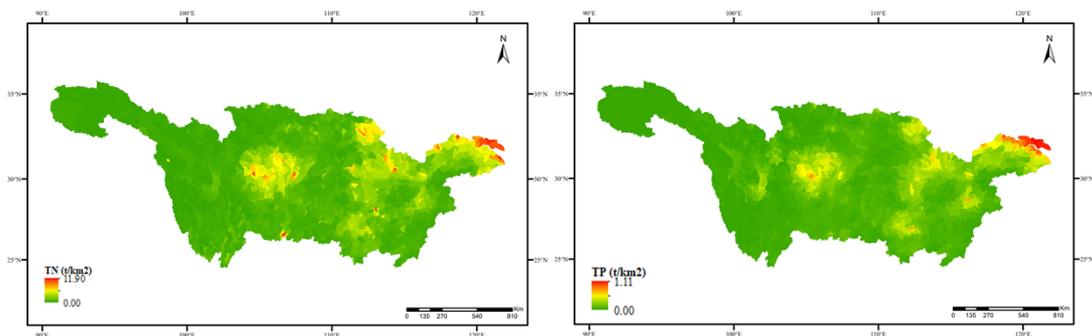


Fig9. Spatial distribution of nutrient net flux per unit area under scenario II in 2193 sub-basins

5. Conclusion

In this study, the CEIN Model was developed as a China emission inventory of nutrient with 0.1° grid, based on county-level data. The model combines nutrient sources, loss functions and removal functions. Coupled with a GA optimization algorithm and LOADEST Model, the model could accurately calculate nutrient fluxes in the watershed. Using the Yangtze River

Basin as a case study, the nutrient production and fluxes in 2193 sub-basins were calculated, and the spatial distribution and contribution of each source was analyzed. According to the nutrient control policies in the Yangtze River Basin, different scenarios were set up to evaluate their impact on nutrient fluxes. This study aimed to provide a reference for nutrient control. Conclusions can be summarized as follows:

(1) The framework of the CEIN Model could successfully simulate nutrient production and flux for smaller sub-basins in a larger watershed, such as the level 8 sub-basins of Yangtze River Basin. In this study, the R^2 of nutrient flux optimization simulation at the target sites was greater than 0.8, which can be considered satisfactory. The division of diffuse sources, point sources and direct sources reflects the real situation of sources more accurately. The nutrient loss function could better reflect the influence of natural conditions, such as rainfall and terrain, on the nutrient environmental processes. The model also provided survey-based, more detailed local aquaculture nutrient production coefficients.

(2) The contribution of cropland and livestock for nutrient production and flux was largest in the Yangtze River Basin. For the less populated and underdeveloped upstream, soil erosion was one of the most important sources. The urbanization process in the midstream and downstream was fast, and the nutrient production contribution of urban non-point source was comparable to that of sewage treatment facilities. The nutrients of freshwater aquaculture directly enter the waterbodies and cannot be ignored in their magnitude. The hot spots of nutrient production and flux were principally located at the central and delta parts of Yangtze River Basin and there were more hot spots of TN than for TP. Model results indicate that reducing nutrient discharge from sewage treatment facilities is not an effective choice for nutrient control. The nutrient management measures for cropland had obvious effects on nutrient control. Therefore, the environmental policies of Yangtze River Basin should be paid more attention to non-point sources, such as cropland.

(3) The CEIN Model leaves much room for further development. With data limitations, manufacturing nutrient production was assumed to be all collected by sewage treatment facilities and the nutrient production of sewage treatment facilities was calculated according to the designed treatment capacity and discharge standard. The soil budget was used roughly to estimate nutrient production of cropland which didn't consider the nutrient environmental processes in the surface flow, interflow and ground flow. With enough data and improvement of model conceptualization, these problems will be solved. In the future, the CEIN Model will also be improved with respect to the temporal resolution to simulate nutrient production and flux at monthly and up to daily scales.

Reference

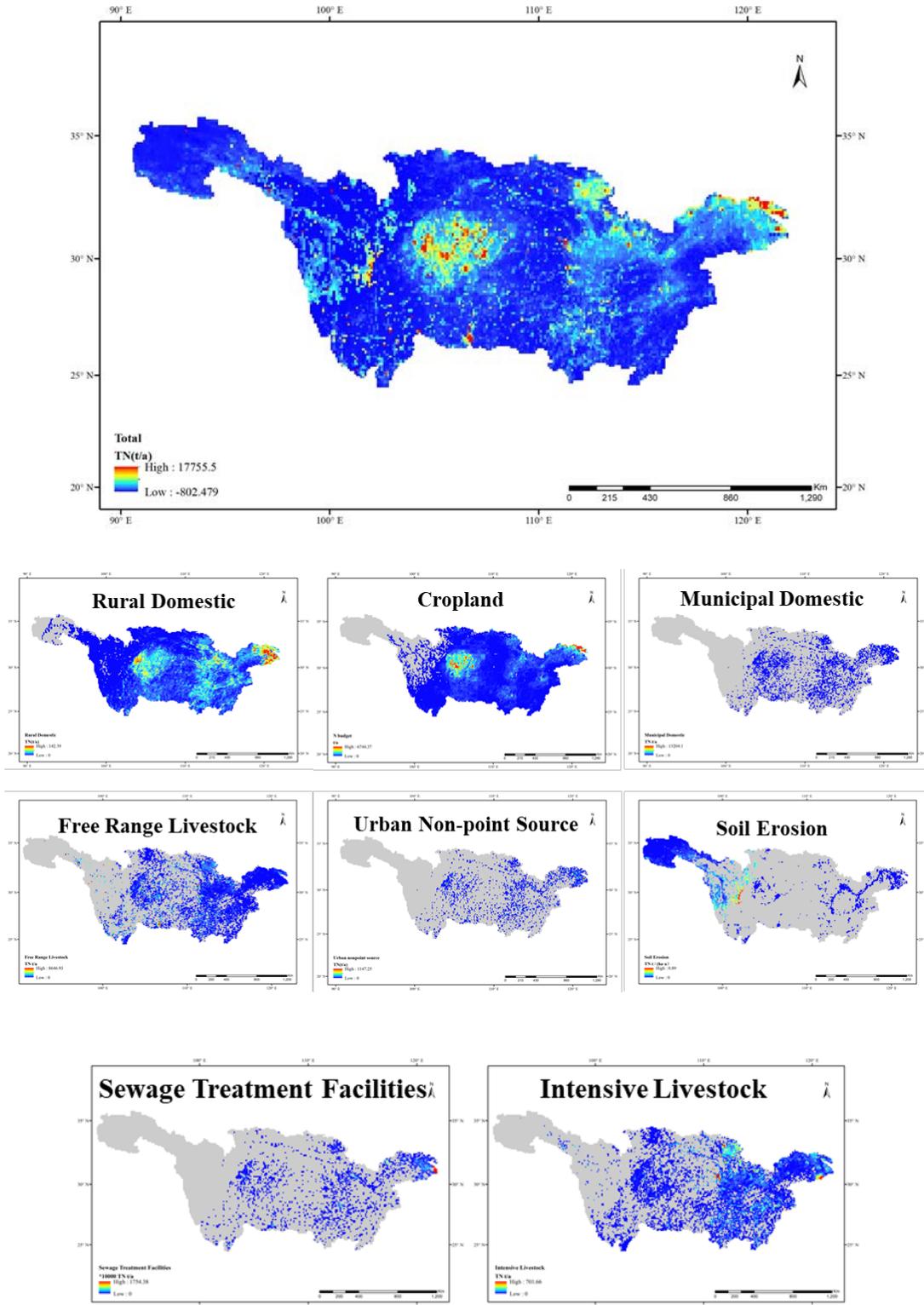
- Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., & Siebert, S. (2003). Development and testing of the WaterGAP 2 global model of water use and availability. *Hydrological Sciences-Journal-des Sciences Hydrologiques*, 48(3), 317-337.
- Bouwman, A. F., Beusen, A. H., & Billen, G. (2009). Human alteration of the global nitrogen and phosphorus soil balances for the period 1970–2050. *Global Biogeochemical Cycles*, 23, GB0A04.
- Chao L, Minghui Y, Yuting H, Xiongzhi X. (2021). Ecosystem service multifunctionality assessment and coupling coordination analysis with land use and land cover change in China's coastal zones, *Science of The Total Environment*, 797, 10048-9697,
- Diaz, R.J., Rosenberg, R., 2008. Spreading dead zones and consequences for marine ecosystems. *Science* 321, 926-929.
- Han Y, Guo X, Jiang Y, et al. (2019). Environmental factors influencing spatial variability of soil total phosphorus content in a small watershed in Poyang Lake Plain under different levels of soil erosion. *Catena*, 187:104357.
- Hartmann, J., Moosdorf, N., Lauerwald, R., Hinderer, M., & West, A. J. (2014). Global chemical weathering and associated P-release-The role of lithology, temperature and soil properties. *Chemical Geology*, 363, 145-163.
- Hou, Y., Ma, L., Gao, Z., Wang, F., Sims, J., Ma, W., et al., 2013. The driving forces for nitrogen and phosphorus flows in the food chain of China, 1980 to 2010. *J. Environ. Qual.* 42, 962-971.
- Mayorga, E., Seitzinger, S.P., Harrison, J.A., Dumont, E., Beusen, A.H.W., Bouwman, A.F., et al., 2010. Global nutrient export from WaterSheds 2 (NEWS 2): model development and implementation. *Environ. Model. Softw.* 25, 837-853
- Qu, H.J., Kroeze, C., 2010. Past and future trends in nutrients export by rivers to the coastal waters of China. *Sci. Total Environ.* 408, 2075-2086.
- Stokal, M., Yang, H., Zhang, Y., Kroeze, C., Li, L., Luan, S., et al., 2014a. Increasing eutrophication in the coastal seas of China from 1970 to 2050. *Mar. Pollut. Bull.* 85, 123-140.
- Stokal, M.P., Kroeze, C., Kopilevych, V.A., Voytenko, L.V., 2014b. Reducing future nutrient inputs to the Black Sea. *Sci. Total Environ.* 466-467, 253-264.

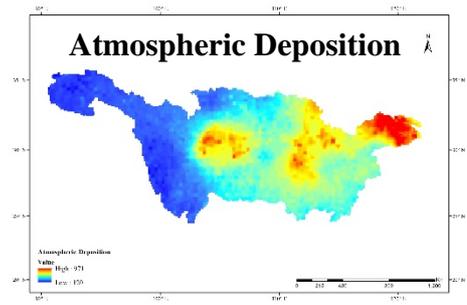
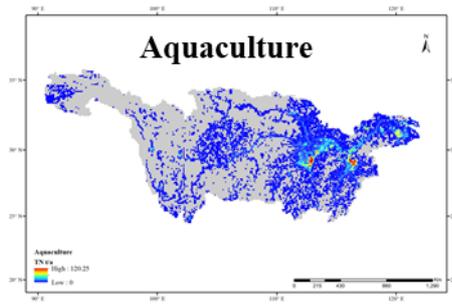
- Strokal, M., Kroeze, C., Li, L., Luan, S., Wang, H., Yang, S., et al., 2015. Increasing dissolved nitrogen and phosphorus export by the Pearl River (Zhujiang): a modeling approach at the sub-basin scale to assess effective nutrient management. *Biogeochemistry* 125, 221-242.
- Tong, Y., Wang, X., Zhen, G., Li, Y., Zhang, W., He, W., 2016. Agricultural water consumption decreasing nutrient burden at Bohai Sea, China. *Estuar. Coast. Shelf Sci.* 169, 85-94.
- Vörösmarty, C., Fekete, B., Meybeck, M., Lammers, R., 2000a. Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution. *J. Hydrol.* 237, 17-39.
- Vörösmarty, C., Fekete, B., Meybeck, M., Lammers, R., 2000b. Global system of rivers: its role in organizing continental land mass and defining land-to-ocean linkages. *Glob. Biogeochem. Cycles* 14, 599-621.
- Williams, R., Keller, V., Voß, A., Bärlund, I., Malve, O., Riihimäki, J., et al. (2012). Assessment of current water pollution loads in Europe: Estimation of grided loads for use in global water quality models. *Hydrological Processes*, 26(16), 2395-2410.

Support information

Nutrient production of 0.1° grid

For TN





For TP

