Ecological Indicators 144 (2022) 109549

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original Articles

Environmental effects and their causes of agricultural production: Evidence from the farming regions of China

Li Ma [a], Yingnan Zhang [b,*], Shuocun Chen [b], Li Yu [b], Yuanli Zhu [c]

[a] *School of Public Policy and Administration, Chongqing University, Chongqing 400044, China*
[b] *School of Public Affairs, Zhejiang University, Hangzhou 310058, China*
[c] *School of Public Policy & Management, China University of Mining and Technology, Xuzhou 221116, China*

ARTICLE INFO

Keywords:
Agricultural sustainability
Agrochemicals use
Water management
Modern agriculture

ABSTRACT

Farming in sustainable ways is rapidly gaining importance, which necessitates a quantitative understanding of the major determinants of agricultural environmental consequences. The main objective of this paper is to investigate the environmental effects and their drivers of agricultural production through the lens of fertilizer application and water quality. An integration of the econometric modeling and machine learning methodologies (multiple regression model and boosted regression trees method) was employed to dissect the relative importance and the non-linearity of each variable. The results showed that (1) Household income level is a primary contributing factor to both fertilizers use and water quality, demonstrating an increasing marginal effect. (2) Proximity factor-distance from provincial capital city positively correlated with these two terms. (3) Water quality is more sensitive to natural conditions, such as precipitation and elevation, which also considerably affect fertilizer use. (4) Agricultural mechanization level is positively related to fertilizer use despite their non-linear effects. In summary, capital, technology, and labor input are the major socio-economic determinants of the environmental consequences, which were also largely associated with regional natural conditions. These findings provide quantitative insights that can be used to improve the agro-environment and achieve sustainable agricultural growth.

1. Introduction

Agriculture remains a major sector of China's economy, while the intense agricultural activities have resulted in extremely negative consequences. To feed the world's continually growing population, farmers use fertilizers, and pesticides, and attempt to expand the agricultural land to maintain the pace of agricultural production (Hofstra and Vermeulen, 2016). Over-fertilization is a major source of agriculture-related environmental deterioration, and a primary contributor to the degradation of water quality, especially eutrophication. It has been estimated that up to 70 % of all surface-water phosphorus (P) imports may come from non-point agricultural sources, and the net P storage of aquatic and terrestrial ecosystems has increased by more than 75 %, primarily due to the application of P fertilizers in agricultural systems (Zhou et al., 2017). There is broad consensus: attention should be paid not only to the occurrence of environmental challenges but also to the socioeconomic drivers of those impacts, which are often neglected (Tan et al., 2022).

Two important themes regarding the environmental effects of agricultural production emerge from the studies discussed so far-fertilizer use and water quality. Investigating the influencing factors of chemical fertilizer overuse serves as a springboard for prospective studies to comprehend the environmental effects of agricultural production (Guo et al., 2021). What causes excessive fertilizer use in China has been the subject of a huge and expanding body of literature, most of which relies on the quantitatively biased approach-econometric method (Chen et al., 2018). A group of related studies examines the factors that influence fertilizer use from a macro perspective, revealing that factors such as grain prices, population growth, industry development, and upgrading, technological advancements in agricultural production, and agricultural and land use policies all demonstrate strong correlations with fertilizer application (Wu, 2011; Takeshima et al., 2016; Olayide, 2009; Tefera et al., 2020). Another group of literature focused on micro-level factors relying on the household survey data, mostly lying within the domain of demographic characteristics (such as education, age, and gender), landholdings, employment, farmland transfer, environmental awareness, etc. (Iticha et al., 2021; Aregay and Minjuan, 2012; Ren

* Corresponding author.
E-mail address: zhangyingnan@zju.edu.cn (Y. Zhang).

https://doi.org/10.1016/j.ecolind.2022.109549
Received 6 July 2022; Received in revised form 12 September 2022; Accepted 5 October 2022
Available online 13 October 2022

et al., 2019). Overall, the technics of identifying the influencing factors are more econometric approach biased and mostly focus on the linear linkages.

Besides, pesticides and fertilizers utilized in agriculture will contaminate both groundwater and surface water. However, water quality (Hereafter in this paper, water quality refers to the quality of surface water) is sensitive to both agricultural activities, industry development, and domestic discharges, furthermore as natural conditions. Thus, identifying the agriculture-related influencing factors and their comparative contribution to water quality is the basis for controlling the negative environmental effects of agricultural production. Given the influencing factors of water quality, it is affected by climatic variations and anthropogenic activities (Deng et al., 2021). The role of climatic variables such as precipitation and water temperature have been extensively examined (Larned et al., 2004). Rainfall and other different kinds of precipitation wash nutrients from human activities like agriculture and fossil fuel combustion into rivers and lakes. While there is a relatively small body of literature that simultaneously consider land use, soil, or geology (O'Donoghue et al., 2021). Interventions incorporating socio-economic, political, and environmental aspects can also lead to negative hydrological and ecological consequences (Wang et al., 2022). Many researchers have suggested that chemical fertilizer application increases the nutrient loading of the rivers, and thus significantly affects water quality (Dean and Mitchell, 2022). Accelerating urbanization with construction land expansion and rapid population growth might also contribute to the deterioration of water quality in terms of the rising nutrient inputs from the solids that accumulate on impervious surfaces and the municipal wastewater discharge (Wu et al., 2021a,b; Liu et al., 2021). Detecting the key drivers influencing water quality is significant for taking effective measures of improving regional water conditions. Traditional multivariate approaches (e.g., multivariate linear regression) (Zorzal-Almeida et al., 2018), spatial statistical modeling approaches (e.g., eigenvector-based spatial filtering, geographically weighted regression) (Detenbeck et al., 2016), Bayesian regression (Lintern et al., 2018) and the partial least squares analysis regression (PLSR) (Ai et al., 2015) has been used to relate various variables to water quality.

Previous analysis trying to clarify the environmental consequences of agricultural production has been centered on different spatial scales, from national, to regional, to water catchments, and even household levels (Byrnes, 1990; Gong et al., 2011). At a macro level, socio-economic data are frequently utilized to identify the determinants of the application of agrochemicals and regional agricultural pollution discharge (Bai et al., 2019; Xu et al., 2021). Analysis at the lower spatial scales attempts to capture details of the drivers, pathways, as well as behavior logic of agricultural environmental effects (Tsiboe et al., 2021). Regression analysis of the econometric approach has been extensively used in examining the influencing factors at different scales (Vasco et al., 2021). For instance, a multiple linear regression model was employed in Segurado et al's (2018) work to examine the associations between a comprehensive set of explanatory variables. However, there are certain disadvantages of this methodology in depicting the nonlinear relations between different drivers and agricultural environmental effects.

Prior studies have examined the effects of agricultural production and the associated drivers at different spatial scales, however, studies have not interrogated the environmental consequences from a comprehensive perspective as agriculture is not the only factor that contributes to deterioration of water quality. Knowledge of coupled fertilizer application and water quality could enhance the understanding of the complex negative environmental externalities of agricultural practices. In addition, a significant body of literature has concentrated on the linear linkages and often adopted the OLS model, logarithmic linear model, and polynomial logit model as the analytical tools, relying on an established theory to formulate a parametric model, collect data for the hypothesized influential predictors, and estimate the parameters of the model (Müller et al., 2013; Ali et al., 2018; Klutse et al., 2018; Ewane, 2020). These "traditional" regression methodologies possess technical merit since they are easily understood, and offer a range of options that relate to the dependent variables and independent variables. However, variables can also have a nonlinear or monotonic relationship, which is difficult to elucidate by regression approach. Thus, accurate and robust findings are difficult to attain because of these limitations, especially when dealing with out-of-sample predictions, which can compromise the model generality. There is an increasing acknowledgment of the complexities in the interactions between environmental effects and the multiple determinants, with ample evidence of nonlinearity (Zhang et al., 2020a; Liu et al., 2016). The data mining approach, specifically, the tree-based models, have the advantages of dealing with the different types of indicators (e.g. categorical and continuous), missing values, and identifying the nonlinear linkages (Hu et al., 2020). Boosted regression tree (BRT) model analyses the contributions of different indicators and marginal effects, which provides insights into the regulatory role of indicators with a nonlinear relationship and the magnitude in a quantitative way (Qin et al., 2020).

Thus, this paper has two main objectives: First, we attempt to gain a comprehensive understanding of the agricultural environmental effects through the lens of both the source-fertilizer application and the consequence-water quality, and reveal the commonalities in determinants of agriculture-related environmental externalities. Second, to obtain more insights into the complexities in the relations between environmental effects and a set of determinants, a BRT algorithm-based variable selection approach is employed to capture the non-linear relationships and the marginal effects between environmental variables and different drivers. Addressing these gaps is meaningful for substantiating the quantitative evidence with comprehensive insights from fertilizer application and water quality changes, and contributes to sustainable agricultural growth.

2. Fertilizer application and water quality

Chemical fertilizers, which supply crops with nitrogen and phosphorus, are applied to the field as nutrients throughout the agricultural production process (Wu et al., 2021a,b). Farmers tend to overuse fertilizer to boost crop yields or accelerate the growth of their plants. However, when fertilizers are not managed properly, they can negatively affect water quality. Runoff and soil erosion allow fertilizer-derived nutrients to reach lakes and streams (Fig. 1) (Havlin et al., 2005). Additionally, erosion transports soil fragments that are nutrient-rich. Eroded soil particles with connected nutrients will accumulate as sediment in water resources (Nazneen et al., 2019).

Nitrogen (N) and phosphorus (P) are the two major nutrients that fertilizer is mostly composed of. In general, excessive N and P moves with water when soil-test N and P levels rise (Carver et al., 2022). These dissolved nutrients may be present in high concentrations in runoff water from fields with high soil-test N and P, increasing the danger of polluting streams, marshes, and lakes. P is the major contributor to algal blooms in surface waters. But, due to its low solubility, phosphate loss in water is negligible, especially when compared to the amounts discharged by industrial waste. Since nitrogen is very soluble, it can seep through the soil (Lundy et al., 2012). Overall, the reduced dissolved oxygen concentrations enrich the water surface with nutrients, which are associated with the growth of organic matter. The increase in the production of biomass and the assemblage in aquatic plants decrease the oxygen level. Besides, the ammonia released by fertilizers causes acidification, which reduces the quality of water sources (Nangia et al., 2010). It has been determined through studies that fertilizers are the source of around 60 % of the nitrogen in rivers and lakes. The amount of nitrogen and phosphorus transported by farm runoff, according to research by the U.S. Department of Environmental Protection, is between 25.0 % and 45.9 % and 29.1 % and 67.5 % of the total N and P that enter rivers annually, respectively (Huang et al., 2004).

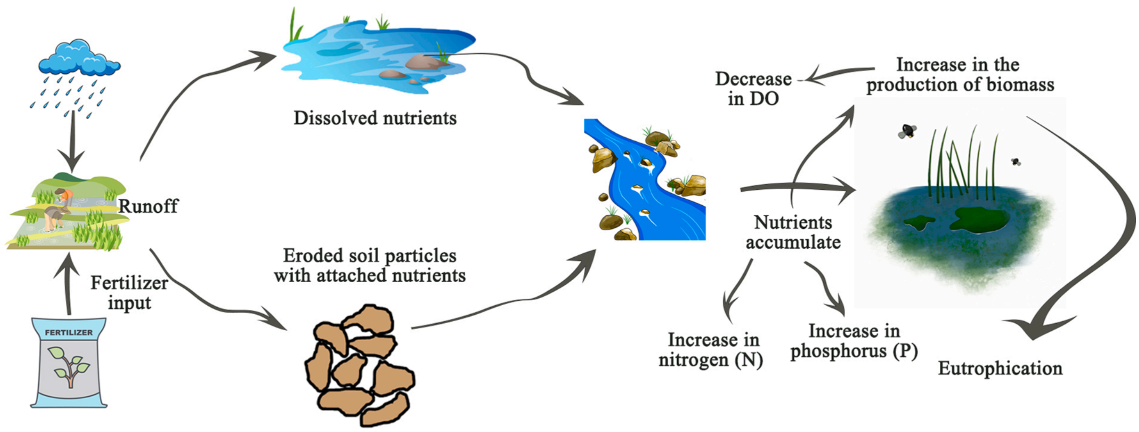


Fig. 1. The interactions between fertilizer application and water quality.

The abovementioned physical, biological, or chemical processes are the direct manifestation of how fertilizer overuse results in water quality deterioration. However, the socio-economic drivers are the fundamental causes. Microeconomic theory and the New Economics of Labor Migration are widely acknowledged when explaining the farmers' fertilizer overuse behavior, which assumes that farmers are rational economic individuals and pursue the maximization of household benefits (Stark and David, 1985). From a macro perspective, the New Growth Theory (Berkum and Meijl, 2000), which holds that investment in human capital, innovation, and knowledge are significant contributors to economic growth, and provides insights into the connections between fertilizer input and agricultural productivity. As an important production factor, increasing fertilizer use is regarded as an effective choice for boosting grain yield driven by the rising conflicts between increasing population and finite land resources.

3. Data and methodology

3.1. Study area

The study area falls within the Huang-Huai-Hai Plain in China, covering the areas of five provinces and two municipalities, i.e., the whole Shandong, most of Beijing, Tianjin, Hebei, and Henan, and the north part of Anhui and Jiangsu (Fig. 2). It is one of the most densely populated regions of China, which has 22.5 % population of the nation's total. The farmland accounts for nearly-one-sixth of China's total, with 50 million hm^2 sowing area, 33 % of the whole country. Relying on the abundant farmland, 35–40 % and 60–80 % of China's wheat and corn are produced. The vast territory and complicated topography demonstrate diverse climate conditions. The precipitation presents a decreasing trend from the southeast coastal region to the northwest region, ranging from 400 mm to 2000 mm. Recent decades have witnessed a rapid growth of fertilizer application in this traditional farming region, with more than 90 % of prefecture-level cities exceeding the international standard (225 kg/hm^2). It has been proven to be one of the seriously affected regions of non-point source pollution (Zhang et al., 2020b) as a result of intensive farming activities.

The reasons why we chose the Huang-Huai-Hai Plain as our study area can be summarized as follows: (1) It is a part of the primary production areas of grain crops and farmers here generally utilize more agrochemicals, which has challenged agricultural sustainability. (2) There are numerous rivers and lakes within this area, and the water degradation has worsened. (3) Mitigating the adverse environmental effects of agricultural production is in desperate need of redirecting farming to sustainable ways and ensuring food security in China. Thus, when concerning the relations between the environmental effects of agricultural production and their determinants, the Huang-Huai-Hai Plain is considered a suitable national representative.

3.2. Data

The data on water quality monitoring is obtained from China National Environmental Monitoring Center. According to the Environmental Quality Standards for Surface Water (GB3838-2002) issued by the Ministry of Ecology and Environment,[1] China (Table 1) and the regional characteristics of the Huang-Huai-Hai Plain, chemical oxygen demand (COD), permanganate index ($KMnO_4$), total phosphorus (TP), ammonia nitrogen (NH_3-N), dissolved oxygen (DO) and five-day biochemical oxygen demand (BOD_5) were selected. The water bodies are divided into five classes according to the utilization purposes and protection objectives, i.e. Class I-V (The higher the ranking is, the worst the water quality is). The assessment of water quality follows the "barrel theory", which means that it is determined by the worst performance of certain evaluation indicators. For example, if COD falls into the category of V, the other indicators are all less than V, and the comprehensive water quality index is classified into V. There are 538 hydrometric stations within the Huang-Huai-Hai Plain. The water environment quality index is determined by the category in which it falls. Kriging interpolation was used to generate the spatial distribution across the whole study area, and on this basis, the mean value of each index for each county was calculated. Kriging has the benefit of giving a measurement of the estimated surface's error or uncertainty. Additionally, it considers the spatial connection between the data collected at various monitoring sites, and its geostatistical framework may also incorporate secondary data to enhance the interpolation results (Krige, 1951).

Fertilizer use information and socio-economic statistics are derived from China County Statistical Yearbook. Geographical data of the Digital Elevation Model (DEM), precipitation, and road networks are collected from the Resources and Environmental Science Data Center, Chinese Academy of Sciences.

3.3. Method

3.3.1. Explanatory variables

The mean value of fertilizer application per unit farmland area and the regional water quality index (The reciprocal value of the water quality class that it belongs to represents the water quality) are used as the explained variables in the model (Table 2). The fertilizer input is an important factor affecting the quality of the regional water environment, its release of nitrogen and phosphorus is the primary cause of the abnormal growth of NH_3-N and TP in water bodies. However, water

[1] https://english.mee.gov.cn/SOE/soechina1997/water/standard.htm.

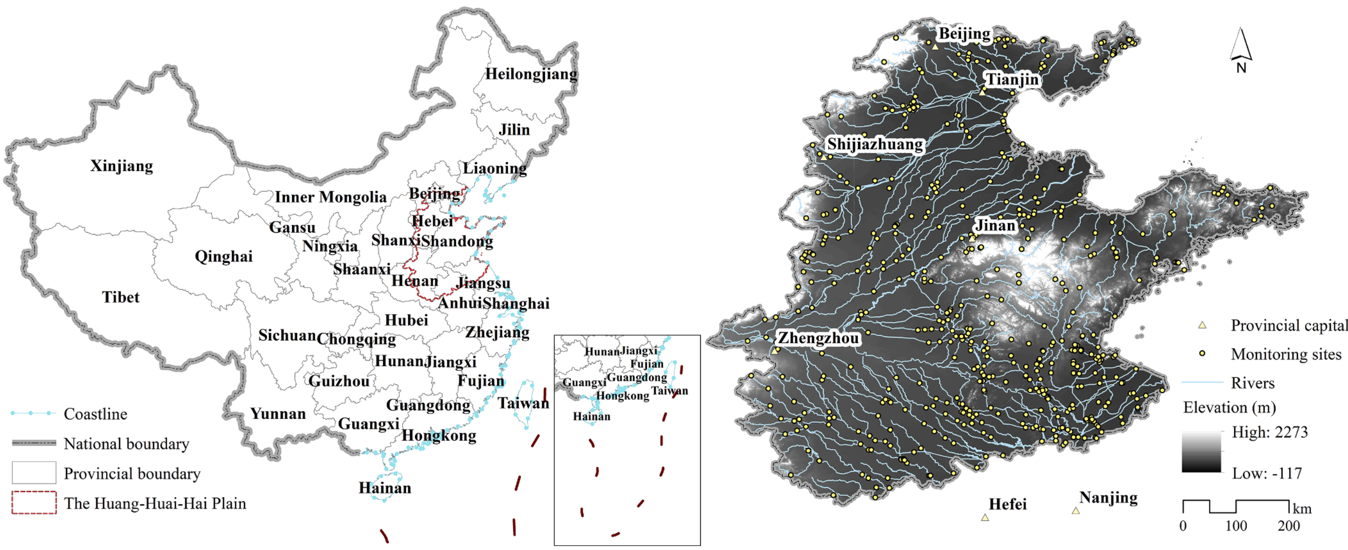


Fig. 2. The spatial distribution of water quality monitoring sites in the Huang-Huai-Hai Plain, China.

Table 1
The criterion of water quality assessment (GB3838-2002).

Contaminants (mg/L)		I	II	III	IV	V
Chemical oxygen demand (COD)	≤	15	15	20	30	40
High manganate index ($KMnO_4$)	≤	2	4	6	10	15
Total phosphorus (TP)	≤	0.02	0.1	0.2	0.3	0.4
Ammonia nitrogen (NH_3-N)	≤	0.15	0.5	1.0	1.5	2.0
Dissolved oxygen (DO)	≥	Saturation rate 90 % (or 7.5)	6	5	3	2
Five-day biochemical oxygen demand (BOD_5)	≤	3	3	4	6	10

environment quality is a comprehensive index, which is sensitive to the unreasonable discharge of pollutants originating from agricultural production activities, industrial development, and the domestic discharge of urban and rural residents. It is hard to fully understand the environmental effects of agricultural production merely depending on the water quality index since identifying the water degradation that just results from unreasonable agricultural activities is tough work. Thus, this study analyzes the determinants of both fertilizer application and water quality, which helps us better understand the environmental consequences of agricultural production.

Agricultural production activity is influenced by both natural and social factors. Besides, the Huang-Huai-Hai Plain is not homogeneous in landform, there is also a certain area of hilly regions, it is necessary to consider the influence of topographical factors on agricultural production inputs and the water environment. Agricultural production is sensitive to rainfall and other meteorological factors. Factors such as rainfall, illumination, and temperature act individually or interactively

Table 2
The descriptive statistics of each variable.

Variables	Variable Description	Mean value	Standard deviation	Minimum value	Maximum value	
Explained variables	Fertilizer application intensity	Fertilizer application per unit farmland area (FIPUFA)	442.4	171.6	142.4	1315.8
	Water quality index	According to the water environment quality evaluation standard (GB3838-2002) to give the corresponding level, and take the reciprocal value of the corresponding level as the water quality index (WQI)	0.22	0.04	0.17	0.39
Explanatory variables	Slope	Average regional slope (Slope)	3.17	2.25	0.77	20.3
	Elevation	Average regional elevation (Elevation)	71.10	85.81	0.79	746.56
	Precipitation	Average regional Precipitation (Precipitation)	6553	1419	4185	9711
	Road network density	$D_i = \sum_{j=1}^{6} (w_j \times R_j) D_i$ is the road network density of county $i.w_i$ is the weight of road type $i.R_j$ is the density of road type j; the density is calculated as road distance per unit area; this road network density value is normalized (RND); the weights of a highway, railroad, national road, provincial road, county road, and township road are 0.3, 0.3, 0.2, 0.1, 0.05 and 0.05, respectively (Zhang et al., 2019)	0.12	0.10	0.01	0.55
	Distance from the provincial capital city	Euclidean distance from the county center to the provincial capital city (DPC)	120.4	758.1	0	451.1
	Distance from prefecture-level city	Euclidean distance from the county center to its prefecture-level city (DPLCC)	213.9	165.1	0	716.2
	Agricultural mechanization level	Total power of agricultural machinery per unit farmland area (TPAMFA)	1.27	0.67	0.14	6.65
	Farmland holding	Per capita farmland area (PCFA)	1905.4	908.8	22.43	7480.15
	Farmers' income level	Per capita net income of farmers (PCNIRR)	6363	1850	2521	12,898

on the conditions for crop growth, affecting agricultural productivity. Road network density directly characterizes the condition of regional transportation infrastructure, which affects the freight transport of agricultural products, and production materials, as well as the dissemination of agricultural science and technology. The theoretical model of the agricultural location reveals a "concentric circle" distribution pattern (Kellerman, 1989). The suburban areas are closely dependent on and serve the development of the cities. The distance from the central cities of different levels can help us to understand the influence of location factors on fertilizer inputs and water quality. The level of agricultural mechanization can directly reflect the level of agricultural intensification and large-scale operation, and characterize the level of agricultural modernization. Per capita farmland area indicates the regional farmland endowments and the scale of the farm operation. The per capita net income of farmers reflects the average level of farmers' income and living standards.

3.3.2. Multiple linear regression

Multiple linear regression based on least squares was used to analyze the relationship between fertilizer application and water quality index with the corresponding explanatory variables. The equations were as follows:

$$FIPUFA = \beta_1 Slope + \beta_2 Elevation + \beta_3 Precipitation + \beta_4 RND + \beta_5 DPC + \beta_6 DPLCC + \beta_7 TPAMFA + \beta_8 PCFA + \beta_9 PCNIRR + \varepsilon \tag{1}$$

$$WQI = \theta_1 Slope + \theta_2 Elevation + \theta_3 Precipitation + \theta_4 RND + \theta_5 DPC + \theta_6 DPLCC + \theta_7 TPAMFA + \theta_8 PCFA + \theta_9 PCNIRR + \mu \tag{2}$$

where, β_1 ....β_9 are the coefficients of each explanatory variable in Eq. (1), and ε are the constant terms. θ_1......θ_9 are the coefficients of the explanatory variables in Eq. (2), and μ are constant terms. The abbreviations of the variables are explained in Table 2.

3.3.3. Boosted regression tree model (BRT)

BRT is a non-parametric model based on the machine learning algorithm with no default estimation of the relationship between independent and dependent variables, which combines boosting and regression trees in a single algorithm (Elith et al., 2008). To address the interaction effects among the predictor variables, the BRT uses recursive binary splitting to generate a regression tree algorithm. The advantage of the boosting approach is that it effectively handles the missing data and abnormal values, allowing for more accurate and faster computations through numerical optimization and regularization (Elith et al., 2008). The data is divided into two groups, one is randomly selected data that will be used to identify the associations between dependent variables and the independent variables, and the rest is used to examine the results (Friedman and Meulman, 2003). Compared to other data mining approaches, BRT tends to be more robust against overfitting and has high predictive accuracy and good interpretability of resulting input–output linkages (Friedman, 2001).

The first step in BRT modeling is selecting the appropriate divisor variable that can divide the dependent variable. Following that, the outcomes are divided into two classes using the strongest divisor variable (Saha et al., 2020). The recursive grouping of the aspect points initially divides each node into two domains and eventually further divides those domains (Therneau et al., 2015). The nodes remain in this division until they are uninformed. In regression tree analysis, the decision of when to terminate the divisions is crucial. If the divisions don't come to an end, the tree algorithm will unavoidably segregate every detail and result in over-fitting (Therneau et al., 2015). Trees that are too challenging for fresh insights are volatile, meaning that they make more mistakes. Friedman (Friedman, 2001) described the BRT model as follows:

$$\mathrm{F}\{X(\beta_m \alpha_m) m_0\} = \sum_{m-0}^{m} \beta_m h(x\alpha_m) \tag{3}$$

where h(x,m)is a basic classification process with parameter α and the variables are the vector x, m describing the model's level and β_m is the weighting coefficient in level m.

In our paper, BRT analysis is performed by using R 3.6.3, with gbm package written by Elith. The learning rate is set as 0.005 and the tree complexity is set as 4.

4. Results

4.1. Fertilizer application and water quality in the Huang-Huai-Hai Plain

The spatial distribution of COD is uneven, with high values concentrated in the Beijing-Tianjin-Hebei region, especially in the heavy industry-dominated area adjacent to Shijiazhuang (Fig. 3a). In terms of data distribution, a small number of outliers leads to a "sharp peak and heavy tail" pattern (Fig. 4). There are two major clusters of high values of permanganate, around Shijiazhuang and in northern Jiangsu (Fig. 3b). The TP in the Beijing-Tianjin-Hebei region was significantly higher than other regions, showing a strip-like high-value area along Beijing-Shijiazhuang-Zhengzhou (Fig. 3c). The spatial distribution trends of NH_3-N and TP were similar (Fig. 3d), while the spatial distribution of dissolved oxygen is generally homogeneous (Fig. 3e). The distribution of DO showed an approximately normal distribution with few outliers (Fig. 4). Zhengzhou city exhibited a high BOD_5, significantly higher than other regions.

Based on the comprehensive evaluation of the above indicators, the water quality degree is derived, and its reciprocal value is used as the water environment quality index (Fig. 5a). The areas with lower water quality are mainly located in the Beijing-Tianjin-Hebei region and the traditional farming areas in northern Anhui and eastern Henan. The areas with better water quality are concentrated in the Jiaodong Peninsula, Shandong hilly areas, northern Jiangsu, and the mountainous areas in western Hebei. The spatial distribution of fertilizer application intensity is not consistent with the spatial distribution of water quality (Fig. 5b). For example, Cangzhou, Hengshui, Xingtai, and Dezhou have relatively low fertilizer application intensity, but their water quality index is low. The high-value areas of fertilizer application were mainly distributed along the Zhengzhou-Jinan-Weihai, which did not show an obvious clustering trend in general. Although the low-value area of fertilizer application ($\leq$426 kg/hm^2) occupied 53.7 %, 94.9 % of the county units exceeded the international fertilizer application standard (225 kg/hm^2).

4.2. Factors affecting fertilizer application and water quality

4.2.1. Multiple linear regression

The analysis of the influencing factors of fertilizer application intensity and water quality by multiple linear regression based on OLS shows that (Table 3) there are differences in the factors affecting these two explained variables. Distance from the provincial capital city and farmers' income level was positively correlated with fertilizer application at the 1 % significance level. Elevation and road network density was positively correlated with fertilizer application intensity at 10 % significance level, while per capita farmland area negatively influence fertilizer use with a standardized coefficient of −0.10. Besides, the agricultural mechanization level demonstrates a positive connection with fertilizer input at a 5 % significance level. The effects of slope, precipitation, and distance from the center of prefecture-level cities on fertilizer application were not significant. Overall, the distance from the provincial capital city and elevation have a greater influence on fertilizer application intensity. This increased proximity of farmland to central urban regions will raise the likelihood of being encroached. The

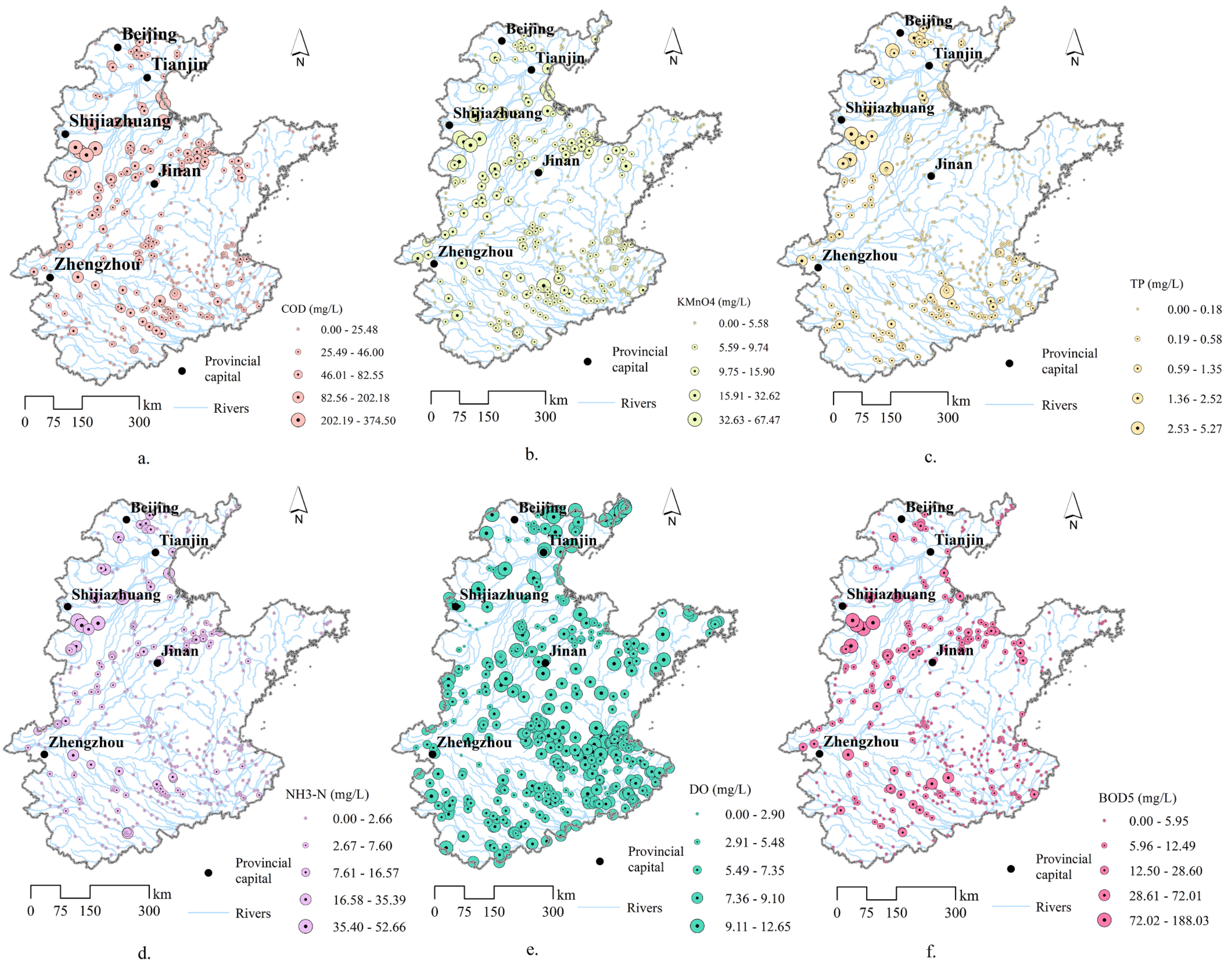


Fig. 3. The spatial distribution of water quality indicators.

growing number of off-farm employees affects the labor allocation, thus intensifying the marginality of farmland. Conversely, in traditional farming areas where far from large cities, agricultural production plays a considerable role in regional economic development and sustaining farmers' livelihood, with a high level of agricultural industrialization and relatively high chemical fertilizer inputs. The significant positive correlation between farmers' per capita net income and fertilizer application intensity shows that capital accumulation helps to increase agricultural production inputs, and the positive feedback from the income-increasing effect of fertilizer application will further increase farmers' motivation for fertilizer inputs. Meanwhile, the growing income level and changing income structure due to household labor reallocation will also increase fertilizer inputs through the remittance effect and labor substitution effect. Fertilizer input can be understood as the farmers' behavioral response taking the natural, social and economic factors into consideration. The positive effect of elevation on fertilizer application is probably related to the fact that high altitude tends to cause soil nutrients to be transported downward through surface runoff, which affects soil quality and requires more fertilizer. The fact that Shandong's mountainous regions, where cash crops like fruit farming are primarily produced, have a higher need for chemical fertilizers, is another potential explanation.

The findings of the multiple linear regression show that water quality is influenced by both natural and socio-economic determinants. Elevation and precipitation have a positive effect on water quality at a 1 % significance level, indicating that the higher the elevation and rainfall, the better the water quality is. The high elevation makes it easier for the pollutants to be transported by surface runoff. The heavy precipitation and adequate water resources help dilute pollutants, thus reducing the accumulation of pollutants. The water quality is not only determined by agricultural production activity, but also the industrial development. Water quality is significantly influenced by location. The farther away from province capital cities and prefecture-level center cities, the better the water environment. This is mostly due to the concentration of industrial and residential activity in the region's center city, where more pollutants are released and have a higher impact on the water quality. Although the sampling sites in this study were primarily distributed in non-urban areas, there is a positive correlation between the road network density and the water quality, contradicting the earlier finding that the water quality declines as the distance to the central city decrease. As a result, areas with high road network density show relatively low pollution loads. Agricultural mechanization level and farmers' income showed a positive correlation with water quality, which seems to contradict the traditional perception. There are three possible reasons: (1) The impact of agricultural production on the water quality of the environment is overshadowed by higher industrial pollution emissions. (2) The scale effect of agricultural mechanization contributes to the reduction of fertilizer inputs, thus reducing the pressure on the water environment. (3) The increase in non-farm income leads to the marginalization of agricultural production, and the phenomenon of

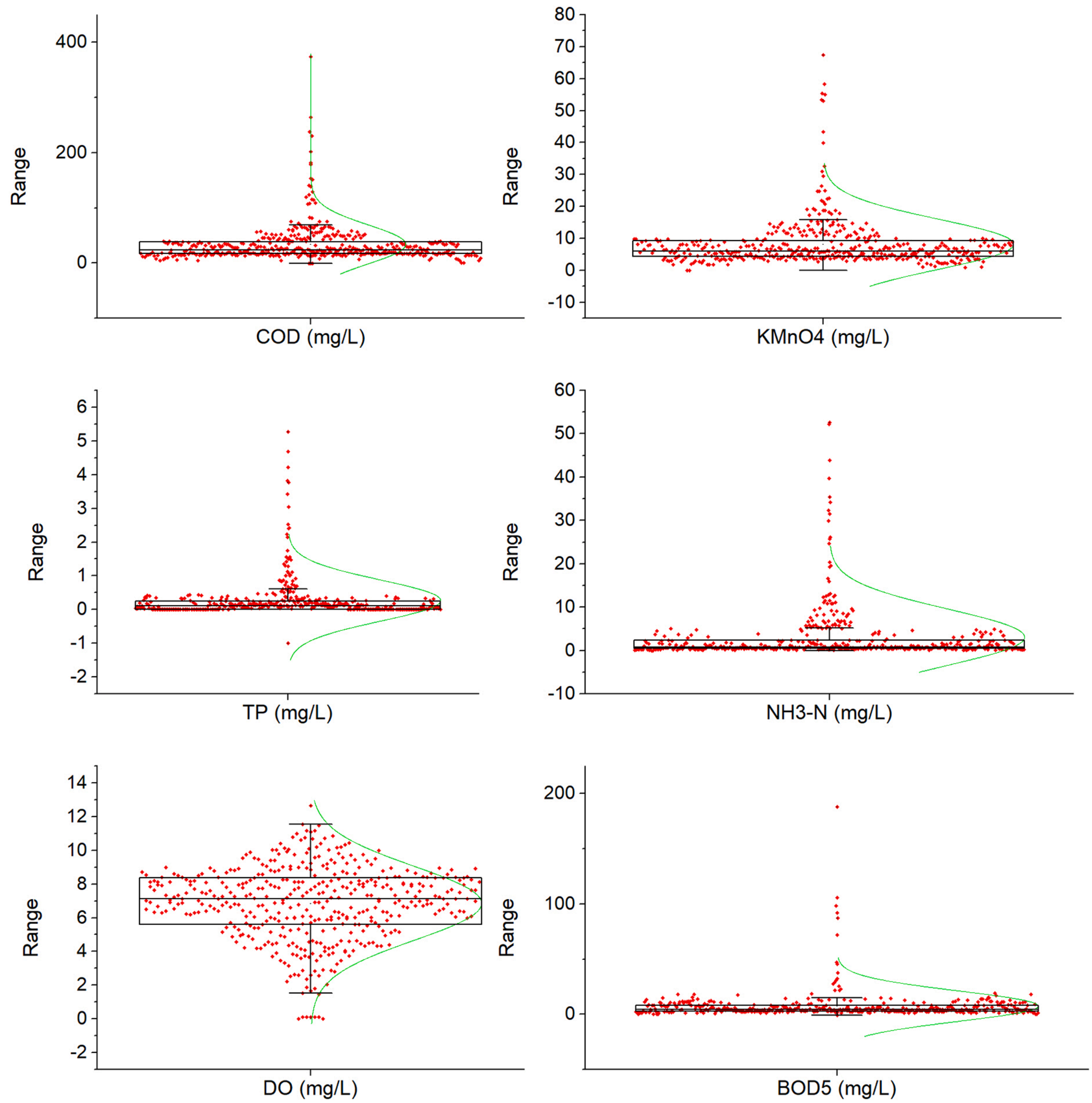


Fig. 4. The box plots of different water quality indicators.

farmland abandonment occurs in some areas, which leads to the decline of chemical fertilizer use. There are some discrepancies and conflicts in the regression results between the water quality index and fertilizer application intensity: Fertilizer application exhibits good relationships with the degree of agricultural mechanization and farmers' income levels, although some of these characteristics also have favorable effects on water quality. The primary cause of this paradox is the fact that different indicators of the water environment have distinct axes of directionality and that agricultural and industrial activities are superimposed to produce the quality of the water environment. TP and NH_3-N, for instance, are frequently impacted by agricultural production operations, whereas COD is frequently used to denote pollution brought by industrial wastewater. Fertilizer application and water quality should therefore both be considered in the analysis of the environmental effects of agricultural production on water quality.

4.2.2. BRT model

When examining the relations between fertilizer application intensity and the corresponding variables, the model was fitted using a learning rate (lr) of 0.005 and a total of 800 decision trees, with a cross-validation deviance of 117.9 and a correlation of 0.64 for the training data (Fig. 6a). As for the water quality index, with the lowest cross-validation deviance (0.015), the optimal model was found (or determined) to have a lr of 0.005, and a total of 4150 decision trees, and showed a correlation of 0.89 in the training data (Fig. 6b).

Fig. 7 depicts the relative influence of each predictor variable on the fertilizer application intensity and water quality. In terms of fertilizer application intensity, the per capita net income of farmers is the most critical predictor variable in the model (relative influence equal to 18.27 %), followed by one natural condition-related variables-the precipitation (14.04 %). In terms of the socio-economic drivers, agricultural mechanization level, and per capita farmland area also have a higher relative contribution to the fertilizer application. When considering water quality, we observed that natural factors substantially contributed to the variation of water quality, and the top two natural condition-related contributors are precipitation (19.8 %) and elevation (13.1 %). Meanwhile, the geographical proximity factors have significant importance for water quality. It has been found that the distance from the provincial capital city considerably contributed to the models with a relative importance of 13.7 %, indicating that the spatial distribution of anthropic activities is the main factor affecting the water quality. Agricultural production activities are also the main contributors of

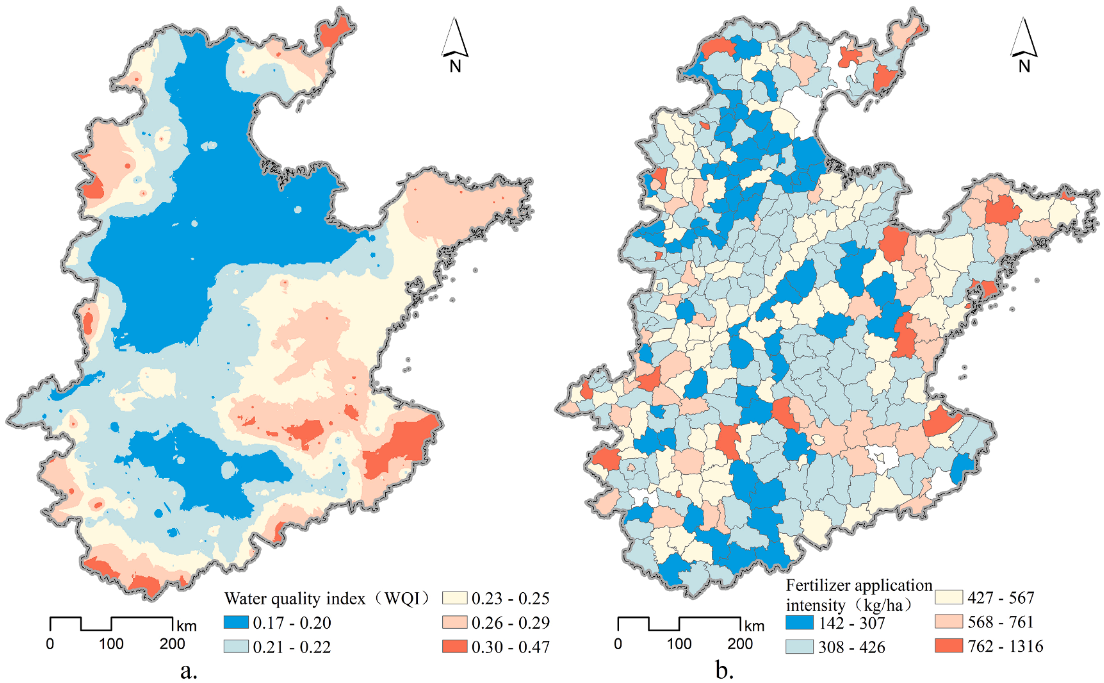


Fig. 5. The spatial distribution of water quality index and fertilizer application intensity.

Table 3
The results of multiple linear regression.

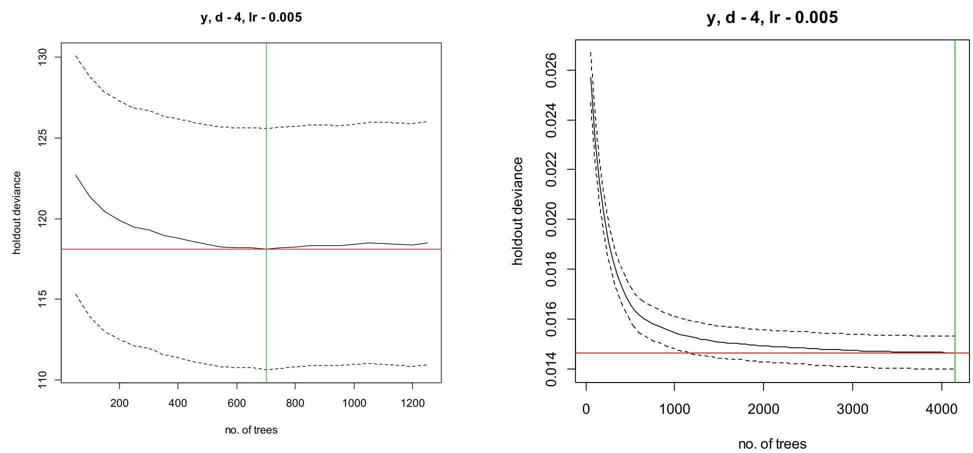
Variable Name	Dependent variable: Fertilizer application intensity (FIPUFA) (Eq. (1))		Dependent variable: Water quality index (WQI) (Eq. (2))	
	Coefficient	Standardized coefficients	Coefficient	Standardized coefficients
Slope	−11.24 (-1.19)	**−0.15**	−0.002 (-0.87)	**−0.096**
Elevation	0.43* (2.05)	**0.22**	0.0002*** (3.43)	**0.37**
Precipitation	0.005 (0.64)	**0.04**	7.00e-06*** (5.23)	**0.28**
Road network density (RND)	201.3* (1.83)	**0.11**	0.09*** (4.42)	**0.24**
Distance from the provincial capital city (DPC)	0.001*** (4.50)	**0.27**	1.59e-07*** (6.28)	**0.34**
Distance to prefectural City (DPLCC)	−0.0001 (-0.16)	**−0.16**	2.88e-07*** (2.75)	**0.13**
Agricultural mechanization level (TPAMFA)	35.12** (2.51)	**0.14**	0.006** (2.24)	**0.11**
Per capita farmland area (PCFA)	−0.018* (-1.70)	**−0.10**	−3.14e-07 (-0.16)	**−0.01**
Farmers' income level (PCNIRR)	0.018*** (3.57)	**0.20**	1.79e-06 * (1.91)	**0.09**
Constant term	192.43*** (3.57)		0.11*** (7.83)	
Observations	353		353	
F	8.22		21.42	
Prob. > F	0.000		0.000	
R^2	0.18		0.36	

pollutants in water bodies, which can be reflected in the contribution of per capita net income of farmers, the total power of agricultural machinery, and per capita farmland area with the combined contribution of the three reaching 28.4 %. Road network density show relatively insignificant relative influence on both fertilizer application and water environment quality, mainly because the overall road network density in the Huang-Huai-Hai Plain is more homogeneous, except for some central cities.

The regression results of these two models show that fertilizer application is not only a technical issue but also a socio-economic issue, which can be conceptualized into both technical and socio-economic dimensions. Similarly, water quality is the combined result of natural factors and regional agricultural and industrial activities. The fact that farmers' income and the mechanization level contribute positively to fertilizer inputs suggests that factor inputs have a favorable impact on the economic output they produce. In contrast, due to the minor intra-regional variability in factor endowments, location considerations have little influence on the application of fertilizer in farming areas. The relative contribution of natural variables to water quality is higher than the impact of agricultural production activities. While agricultural activity continues to be the primary source of pollutants in water bodies, natural factors are primarily reflected in the capacity for self-regulation. Therefore, it is important to separate the effects of natural conditions from those of agricultural activity when evaluating the contribution of agricultural activity to the influence on water quality.

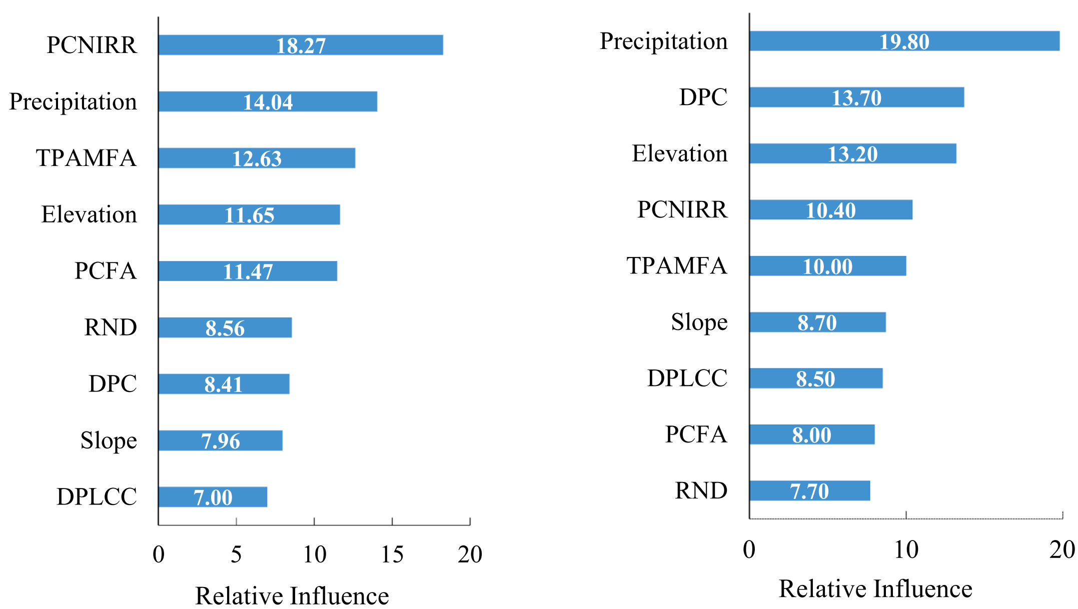
4.2.3. Partial dependence of the influencing factors

The influence of explanatory variables on the probability generated by BRT is shown in Fig. 8 and Fig. 9. The relative influence greater than zero indicates that the driver is positively correlated with the dependent variable, whereas less than zero indicates a negative correlation, and zero values indicate no correlation. The farmers' per capita net income showed an increasing non-linear association with fertilizer application. When farmers' per capita income exceeds a certain level (approximately 10,000 yuan), the marginal benefit of non-farm income is much greater than that of agricultural production, and the labor substitution effect of fertilizer is manifested. With the increase of precipitation, its marginal effect becomes more and more obvious, and its influence on fertilizer application alternates between positive and negative effects, but generally an increasing trend. The influence of agricultural mechanization level on fertilizer application in the low-value area fluctuates to



a. Fertilizer application intensity (FAI) b. Water quality index (WQI)

Fig. 6. The deviance of optimal lr and tc.



a. Relative effect of each influence factor on fertilizer application intensity

b. The relative impact of each influence factor on the water quality index

Fig. 7. The relative influence of different indicators.

some extent, but the overall trend is upward and tends to be stable when it reaches a certain degree. When the elevation is less than 80 m, an increasing non-linear association with fertilizer application was observed. The higher the elevation, the more complex the terrain conditions are, and the more difficult it is to cultivate. If the elevation is over 200 m, the amount of farmland is scarce and there is no obvious influence on fertilizer input. Per capita farmland area showed a non-linear negative correlation with fertilizer application when it is less than 2000 m^2. The marginal effect of road network density on fertilizer input first increases, then decrease, and finally, increases. As the distance from the provincial capital city grows, the influence of fertilizer application increases until it ultimately stabilizes. Since the metropolitan area is dominated by non-agricultural industries and fertilizer consumption is naturally lower than in other areas, these exhibit a negative correlation when the distance from the provincial capital city is close. However, as the distance increases to a certain point, the central city's radiation-driven effect becomes apparent. Application of fertilizer first has a decreasing effect before becoming more influential. At the prefecture level, the topographic slope and distance from the city center have a little impact on fertilizer application, and the response curve is comparatively smooth.

The findings show that different variables' marginal effect curves on water quality exhibit various trajectories. The relationship between precipitation and water quality was shown to be "U" shaped. The distance from the provincial capital city presents an increasing non-linear linkage with water quality. When the elevation is below 190 m, there is a rising non-linear link between elevation and water quality. However, as the elevation rises, the effect becomes less significant and ultimately stabilizes. The marginal effect of farmers' per capita net income on the water quality gradually increases and then maintains a stable level. Within a specific range, the overall power of agricultural machinery is inversely connected with the quality of the water, and after a

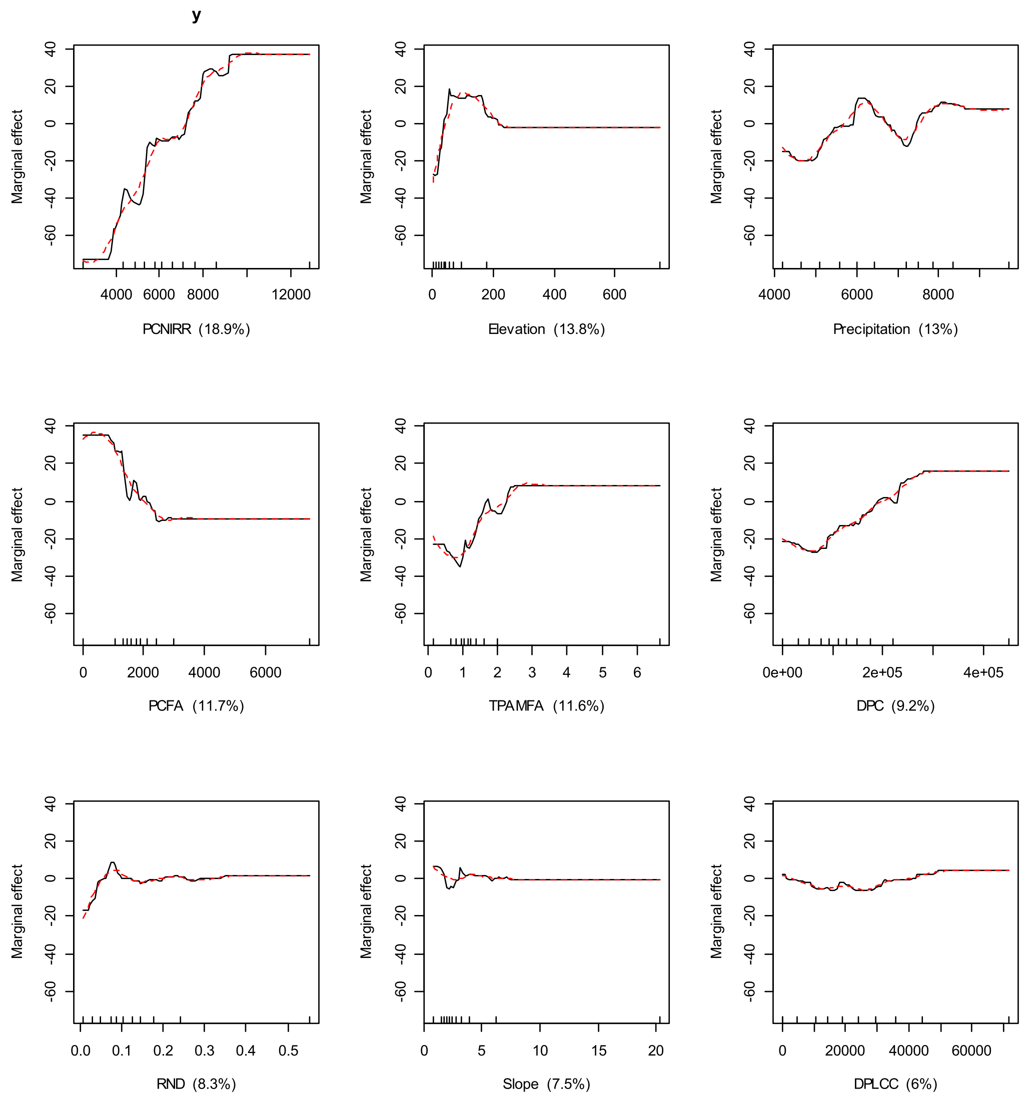


Fig. 8. Partial dependence plots for the most influential factors of fertilizer application.

particular value is reached, the impact is greatly reduced. The impact on the water environment is only marginally influenced by the slope, separation from the prefecture-level center city, area of cropland per unit, and density of the road network.

Comparative analysis of the findings derived from BRT and multiple linear regression results (Table 4). There are some differences between BRT and multiple linear regression results, but the dominant factors obtained have some commonality. For fertilizer application intensity, multiple linear regression identified six significant influencing factors, among which distance from the provincial capital city was the most influential factor, while the analysis results of BRT showed that the contribution of farmers' per capita net income has the largest importer, and the contribution of distance from the provincial capital city was not large. In general, except for the location factor, which differed significantly between the two models, all other dominant factors were the same, probably due to the small variability of location factors. For the water quality index, the dominant factors have high similarity, with rainfall, elevation, and distance from the provincial capital city ranking in the top three. In the BRT, the contribution of per capita net income of farmers ranks 4th, while its ranking is 7th in the multiple linear regression. The two parameters that have the least impact on BRT are the distance to prefecture-level cities and the density of the road network, but they show a stronger link in the results of multiple linear regression.

5. Discussion

5.1. *The macro view of the mechanism of the agro-environmental effects in farming regions*

Our findings have revealed that the negative environmental responses are attributed to both socio-economic disturbances and regional natural conditions. In modern agriculture, maximizing and sustaining crop yields are the main objectives, thus achieving an economically successful farming practice. The New Growth Theory provides a useful theoretical lens for understanding the unreasonable agrochemical inputs, which posits that humans' desires and unlimited wants foster ever-increasing productivity and economic growth (Berkum and Meijl, 2000). Lewis's model explains economic growth in terms of the labor transition between the agricultural and industrial sectors and considers that the subsistence agricultural sector is typically characterized by low wages, an abundance of labor, and low productivity (Lewis, 1954). But, if the marginal output of the migrated labor is more than zero, which requires a growth in labor productivity to counteract the negative effects on the agricultural sector. Previous studies have revealed that capital and money have become an important source of growth engine in modernized agriculture, especially in this transitional stage (Yang and Zhu, 2013). This also confirms that capital, technology, and labor are the core factors of economic growth that neoclassical economics acknowledges

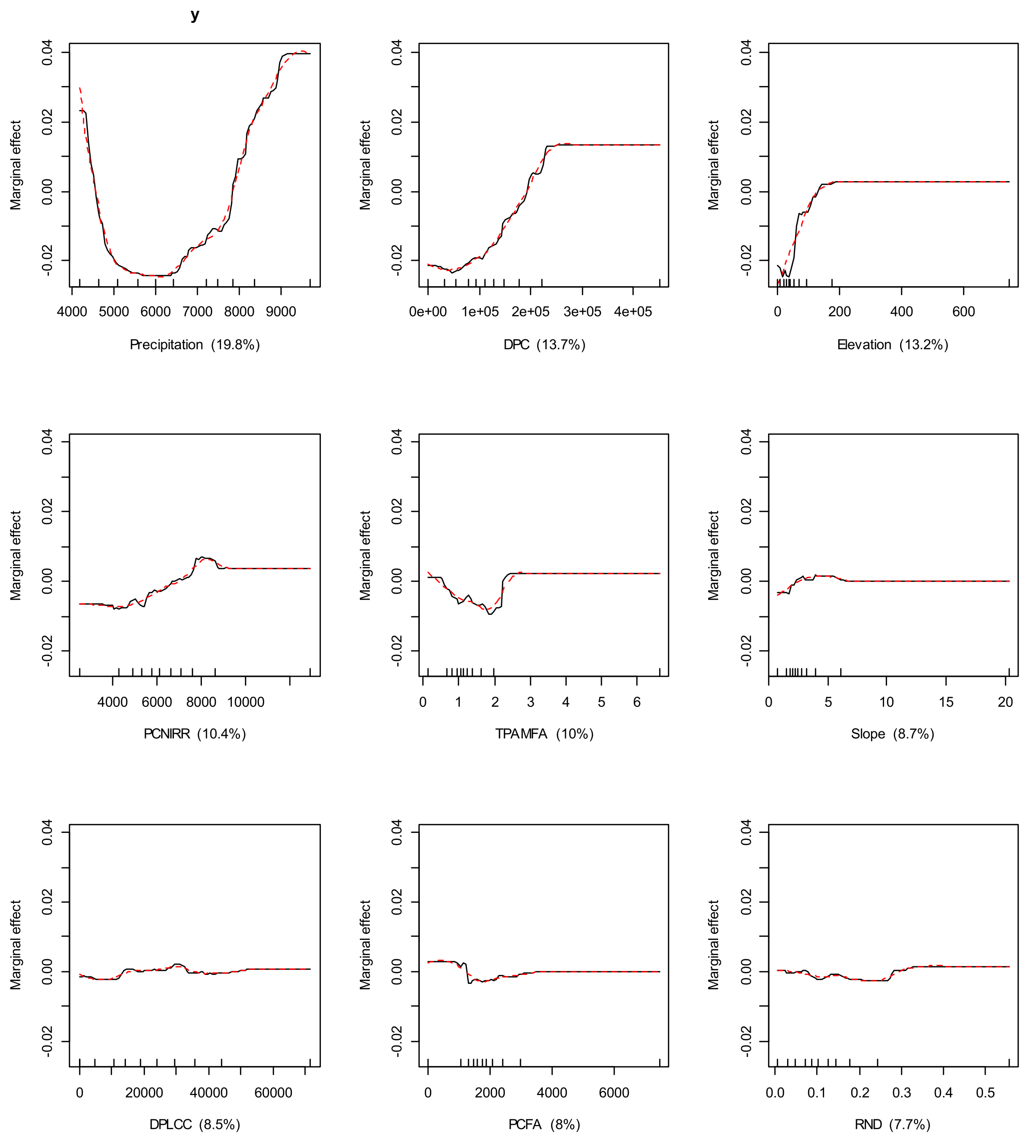


Fig. 9. Partial dependence plots for the most influential factors of water quality.

(Solow, 1956). Household income level is closely related to capital input. That is to say, a household with a high-income level has a larger probability to invest more in agricultural production in the process of household capital reallocation. Our results have substantiated this by disclosing the positive relations between per capita net income of farmers and fertilizer application, as well as water quality. Agricultural mechanization arises as a response to limited agricultural labor and complements mutually with fertilizers use. Our findings indicate that the fertilizer application intensity and water quality all initially decrease and then increase as the agricultural mechanization level increases. This finding is inconsistent with the prior work, which discovers that agricultural mechanization would reduce fertilizer application through the scale effect (Lu and Xie, 2018). However, we found that, when agricultural mechanization increases to a certain level, fertilizer use would increases. The possible reason may be that the exodus of the able-bodied rural population to cities driven by the attraction of profitable jobs in cities increases the demand for agricultural mechanics, and, at the same time, trigger the increasing substitution effect between rural labor and fertilizer. In reality, some major agricultural technologies are not commonly used in China, and the farm structure is not diversified. Agricultural growth still highly depends on fertilizer application. The productivity-oriented agricultural policy guides the flow of technology and capital into the agricultural sector through institutional supply. China's No.1 Central Document successively focuses on the vital role of grain production and food security (Ye, 2015), However, China is growing more food on less land. To maximize food production, agrochemicals become the priority of farmers under the condition that the contribution rate of agricultural technology to agricultural growth is low. Besides, regional characteristics such as natural conditions, resource endowments, and infrastructure also affect the input of agrochemicals and water quality through different pathways. Our analysis uncovers that elevation and precipitation are positively related to fertilizer use and water quality while showing pronounced negative effects on water quality. This can be explained by the fact that precipitation-induced surface runoff can hasten the loss of nutrients like nitrogen, necessitating a greater input of fertilizer to sustain production. Areas with high rainfall have more surface water resources, which can dilute the pollutant concentration to a certain extent. From the perspective of new structural economics, the cultivated land endowment is the basis for the development of the planting industry, and it is difficult to expand quantitatively within a certain period, which determined the agricultural productivity. Abundant farmland resources endow the plain agricultural areas with potential comparative advantage in developing agriculture, while infrastructure determines whether the potential

Table 4
Comparative analysis of the major driving factors identified by BRT and multiple linear regression.

Influence Ranking	Fertilizer application intensity (FIPUFA)		Water Quality (WQI)	
	BRT	Multiple linear regression	BRT	Multiple linear regression
1	Per capita net income of farmers (PCNIRR)	Distance from the provincial capital city (DPC)	Rainfall (Precipitation)	Elevation (Elevation)
2	Rainfall (Precipitation)	Elevation (Elevation)	Distance from the provincial capital city (DPC)	Distance from the provincial capital city (DPC)
3	Total power of agricultural machinery per unit farmland area (TPAMFA)	Per capita net income of farmers (PCNIRR)	Elevation (Elevation)	Rainfall (Precipitation)
4	Elevation (Elevation)	Total power of agricultural machinery per unit farmland area (TPAMFA)	Per capita net income of farmers (PCNIRR)	Road network density (RND)
5	Per capita farmland area (PCFA)	Road network density (RND)	Total power of agricultural machinery per unit farmland area (TPAMFA)	Distance to prefectural City (DPLCC)
6	Road network density (RND)	Per capita farmland area (PCFA)	Slope (Slope)	Total power of agricultural machinery per unit farmland area (TPAMFA)
7	Distance from the provincial capital city (DPC)		Distance to prefectural City (DPLCC)	Per capita net income of farmers (PCNIRR)
8	Slope (Slope)		Per capita farmland area (PCFA)	
9	Distance to prefectural City (DPLCC)		Road network density (RND)	

comparative advantage can be transformed into economic benefits. Good traffic conditions also shape its regional advantages, and the proximity to the market affects the demand for agricultural products, which is naturally associated with the input of agrochemicals. Distance from the provincial capital city and agricultural mechanization level positively correlates with fertilizer use with an increasing marginal effect.

5.2. Policy implications

This research also provides policy recommendations for sustainable agricultural growth. Firstly, location factors and the corresponding climate factors are the prerequisites for arranging agricultural production. The spatial distribution of precipitation should be considered when controlling fertilizer use and water quality management. In terms of proximity, farmland abandonment and extensive use are prominent in suburban regions due to the profitable gains from off-farm jobs. Although this reduces the application of fertilizer, it is harmful to food security. Thus, more innovative agricultural technology should be encouraged to boost the incentives of farmers to farm, especially in near urban areas. Meanwhile, agricultural subsidies and incentives should focus on pro-environmental behaviors. Water quality in suburban regions is highly influenced by industry and domestic discharge of pollutants, which doesn't imply that agricultural pollution can be ignored. As for the remote rural areas, it is necessary to disseminate more environmental knowledge and provide scientific and technical training in fertilization. Besides, promoting the social service of agricultural mechanics is a better choice for reducing fertilizer use and sustaining farmers' livelihood.

5.3. Contributions and limitations

This paper contributes to the extant literature in two aspects. First, our analysis from an integrated perspective that considers the macro drivers of both fertilizer application and water quality, extend our understanding of the relative contribution and marginal effect of different influencing factors on the agricultural environmental consequences. Our results imply that there exist nonlinear associations between the explained variables and explanatory variables, especially the proximity factors and agricultural mechanization level. This helps comprehend the regularities of the interactions between agricultural environmental consequences and their macro causes and is significant in generating policy implications for regulating agricultural activities. Second, an attempt was undertaken to integrate both econometric modeling and machine learning methodology to scrutinize the nexus between multiple potential drivers and the corresponding environmental indicators. This research is probably a pioneering quantitative exploration of complex connections with dependent variables, which surmounts the disadvantage of traditional econometric models and guarantees the robustness of our findings.

Although this paper adds great significance to existing knowledge on understanding the agricultural environmental effects by adopting an innovative approach, room for improvement cannot be concealed. The investigation of what contributes to the environmental responses of agricultural production is a complicated effect. A single work can only shed a light on a limited set of aspects. The key goal of this paper was to provide an integrated approach to identify the primary socio-economics and natural influencing factors. Due to the data availability, some indicators are not included in explanatory variables, such as regional planting structure, and the price of fertilizer. Besides, policy interventions should also be emphasized in future work, as policies regarding fertilizer use, environmental protection, and water governance also greatly affects fertilizer use and water quality.

6. Conclusions

Understanding the macro drivers of agricultural environmental consequences in farming regions is essential. An approach integrating multiple linear regression and BRT was employed to examine the nonlinear and complex interactions between the multiple factors and two environmental effects indicators-fertilizer application intensity and water quality. It has been proved that data mining approaches are particularly appropriate for analyzing agricultural environmental effects due to the complexities inherent in the patterns and processes that result in environmental responses. Our findings revealed that fertilizer use was largely determined by farmers' income levels, which presents a positive correlation. The inclusion of elevation and precipitation captured 25.7 % of the variation in the outcome of fertilizer use. It was also observed that agricultural mechanization level considerably contributes to fertilizer application and shows an abrupt increasing marginal effects curve. Water quality is largely determined by natural factors including precipitation and elevation, despite their non-linear effects, manifested as inverted U shapes with precipitation. The threshold value of this factor's marginal effects can provide empirical references and implications for the spatial arrangements of farming activities. Distance from the provincial capital city and the income level became increasingly critical in determining water quality variation, which indicates the growing importance of proximity factor and capital accumulation in affecting environmental responses. In summary, the driving mechanism of the agricultural environmental consequences can be summarized as negative feedback of the local eco-environmental system caused by the excessive disturbance of the agroecosystem, which is resulted from the capital, technology, and labor input and is constrained by regional resources endowments and natural conditions.

To sum up, this paper draws from a multi-disciplinary theoretical lens and leverages the strength of multiple research methods to deepen the understanding of the driving factors of agricultural environmental consequences, which are of great significance for sustainable agricultural development. The pertinent policy recommendations generated from this work also contribute to the sustainable transformation of farming activities and improving the man-land relations in plain areas.

CRediT authorship contribution statement

Li Ma: Conceptualization, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Yingnan Zhang:** Supervision, Conceptualization, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Shuocun Chen:** Data curation. **Li Yu:** Visualization. **Yuanli Zhu:** Data curation.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

We would like to recognize China National Science Foundation (Grant No. 41731286 & 42101203&42201204), Natural Science Foundation of Chongqing (Grant No. 2018GXNSFDA281032), China Postdoctoral Science Foundation (No. 2021M693709), Fundamental Research Funds for the Central Universities (No. 2022CDSKXYGG006), and Social Science Planning Research Doctoral Program of Chongqing Social Science Planning Office (Grant No. 2020BS43) for financial support. We also thank all the reviewers for their valuable comments that greatly enhanced this paper.

References

Ai, L., Shi, Z.H., Yin, W., Huang, X., 2015. Spatial and seasonal patterns in stream water contamination across mountainous watersheds: linkage with landscape characteristics. J. Hydrol. 523, 398–408. https://doi.org/10.1016/j.jhydrol.2015.01.082.

Ali, E.B., Awuni, J.A., Danso-Abbeam, G., 2018. Determinants of fertilizer adoption among smallholder cocoa farmers in the Western Region of Ghana. Cogent Food Agric. 4, 1538589. https://doi.org/10.1080/23311932.2018.1538589.

Aregay, F.A., Minjuan, Z., 2012. Impact of irrigation on fertilizer use decision of farmers in China: A case study in Weihe River Basin. J. Sustain. Dev. 5 (4), 74–82. https://doi.org/10.5539/jsd.v5n4p74.

Bai, X., Wang, Y., Huo, X., Salim, R., Bloch, H., Zhang, H., 2019. Assessing fertilizer use efficiency and its determinants for apple production in China. Ecol. Ind. 104, 268–278. https://doi.org/10.1016/j.ecolind.2019.05.006.

Berkum, S., Meijl, H., 2000. The application of trade and growth theories to agriculture: a survey. Austr. J. Agric. Resour. Econ. 44 (4), 505–542. https://doi.org/10.1111/1467-8489.00123.

Byrnes, B.H.I.F., 1990. Environmental effects of N fertilizer use - an overview. Fertil. Res. 26, 209–215. https://doi.org/10.1007/BF01048758.

Carver, R., Nelson, N., Roozeboom, K., et al., 2022. Cover crop and phosphorus fertilizer management impacts on surface water quality from a no-till corn-soybean rotation. J. Environ. Manage. 301, 113818. https://doi.org/10.1016/j.jenvman.2021.113818.

Chen, X., Ma, L., Ma, W., Wu, Z., Cui, Z., Hou, Y., Zhang, F., 2018. What has caused the use of fertilizers to skyrocket in China? Nutr. Cycl. Agroecosys. 110, 241–255. https://doi.org/10.1007/s10705-017-9895-1.

Dean, K., Mitchell, J., 2022. Identifying water quality and environmental factors that influence indicator and pathogen decay in natural surface waters. Water Res. 211, 118051. https://doi.org/10.1016/j.watres.2022.118051.

Deng, C., Liu, L., Li, H., Peng, D., Wu, Y., Xia, H., Zhang, Z., Zhu, Q., 2021. A data-driven framework for spatiotemporal characteristics, complexity dynamics, and environmental risk evaluation of river water quality. Sci. Total Environ. 785, 147134. https://doi.org/10.1016/j.scitotenv.2021.147134.

Detenbeck, N.E., Morrison, A.C., Abele, R.W., Kopp, D.A., 2016. Spatial statistical network models forstream and river temperature in New England, USA. Water Resour. Res. 52, 6018–6040. https://doi.org/10.1002/2015WR018349.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

Environmental Quality Standards for Surface Water, (2002). The National Standards of the People's Republic of China. https://english.mee.gov.cn/SOE/soechina1997/water/standard.htm/ (accessed 10th, October, 2021).

Ewane, E.B., 2020. Assessing land use and landscape factors as determinants of water quality trends in Nyong River basin, Cameroon. Environ. Monit. Assess. 192 https://doi.org/10.1007/s10661-020-08448-2.

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 29 (5), 1189–1232. https://doi.org/10.2307/2699986.

Friedman, J.H., Meulman, J.J., 2003. Multiple additive regression trees with application in epidemiology. Stat. Med. 22, 1365–1381. https://doi.org/10.1002/sim.1501.

Gong, P., Liang, L., Zhang, Q., 2011. China must reduce fertilizer use too. Nature 473 (7347), 284–285. https://doi.org/10.1038/473284e.

Guo, L., Li, H., Cao, X., Cao, A., Huang, M., 2021. Effect of agricultural subsidies on the use of chemical fertilizer. J. Environ. Manage. 299, 113621 https://doi.org/10.1016/j.jenvman.2021.113621.

Havlin, J.L., Beaton, J.D., Tisdale, S.L., Nelson, W.L., 2005. Soil Fertility and Fertilizers: An Introduction to Nutrient Management. Pearson Education, Upper Saddle River, New Jersey.

Hofstra, N., Vermeulen, L.C., 2016. Impacts of population growth, urbanisation and sanitation changes on global human Cryptosporidium emissions to surface water. Int. J. Hyg. Environ. Health 219 (7), 599–605. https://doi.org/10.1016/j.ijheh.2016.06.005.

Hu, Y., Dai, Z., Guldmann, J., 2020. Modeling the impact of 2D/3D urban indicators on the urban heat island over different seasons: A boosted regression tree approach. J. Environ. Manage. 266, 110424. https://doi.org/10.1016/j.jenvman.2020.110424.

Huang, G., Wang, X., Qian, H., et al., 2004. Negative impact of inorganic fertilizes application on agricultural environment and its countermeasures. Ecology and Environment 13 (4), 656–660. https://doi.org/10.16258/j.cnki.1674-5906.2004.04.055.

Iticha, M.D., Jaleta, M., Mitiku, F., Tejada Moral, M., 2021. Determinants and profitability of inorganic fertilizer use in smallholder maize production in Ethiopia. Cogent Food Agric. 7 https://doi.org/10.1080/23311932.2021.1911046.

Kellerman, A., 1989. Agricultural location theory 1: Basic models. Environ. Plann. A 21, 1381–1396. https://doi.org/10.1068/a211381.

Klutse, A.R., Bationo, A., Mando, A., 2018. Socio-economic determinants and trends on fertilizer use in West Africa. Springer International Publishing 14, 253–274. https://doi.org/10.1007/978-3-319-58789-9_14.

Krige, D., 1951. A Statistical Approach to Some Mine Valuation and Allied Problems on the Witwatersrand. Masters Thesis. University of Witwatersrand, South Africa.

Larned, S., Scarsbrook, M., Snelder, T.H., Norton, N., Biggs, B.J.F., 2004. Water quality in low-elevation streams and rives of New Zealand. N. Z. J. Mar. Freshwater Res. 38 (2), 347–366. https://doi.org/10.1080/00288330.2004.9517243.

Lewis, W.A., 1954. Economic development with unlimited supplies of labour. The Manchester School 22 (2), 139–191. https://doi.org/10.1111/j.1467-9957.1954.tb00021.x.

Lintern, A., Webb, J.A., Ryu, D., Liu, S., Waters, D., Leahy, P., et al., 2018. What are the key catchment characteristics affecting spatial differences in riverine water quality? Water Resour. Res. 54, 7252–7272. https://doi.org/10.1029/2017WR022172.

Liu, G., Wang, H., Cheng, Y., Zheng, B., Lu, Z., 2016. The impact of rural out-migration on arable land use intensity: Evidence from mountain areas in Guangdong, China. Land Use Policy 59, 569–579. https://doi.org/10.1016/j.landusepol.2016.10.005.

Liu, J., Yan, T., Shen, Z., 2021. Sources, transformations of suspended particulate organic matter and their linkage with landscape patterns in the urbanized Beiyun river Watershed of Beijing. China. Sci. Total. Environ. 791, 148309 https://doi.org/10.1016/j.scitotenv.2021.148309.

Lu, H., Xie, H., 2018. Impact of changes in labor resources and transfers of land use rights on agricultural non-point source pollution in Jiangsu Province, China. J. Environ. Manage. 207, 134–140. https://doi.org/10.1016/j.jenvman.2017.11.033.

Lundy M. E., Spencer D. F., Kessel C. V., Hill J. E., Linquist B. A., Managing phosphorus fertilizer to reduce algae, maintain water quality, and sustain yields in water-seeded rice. Field Crops Res., 131, 81-87. 10.1016/j.fcr.2012.03.005.

Müller, D., Leitão, P.J., Sikor, T., 2013. Comparing the determinants of cropland abandonment in Albania and Romania using boosted regression trees. Agr. Syst. 117, 66–77. https://doi.org/10.1016/j.agsy.2012.12.010.

Nangia V., Gowda P. H., Mulla D. J. et al., Effects of changes in N-fertilizer management on water quality trends at the watershed scale. Agric. Water Manage. 97, 1855-1860. 10.1016/j.agwat.2010.06.023.

Nazneen, S., Raju, N.J., Madhav, S., et al., 2019. Spatial and temporal dynamics of dissolved nutrients and factors affecting water quality of Chilika lagoon. Arab. J. Geosci. 12, 243. https://doi.org/10.1007/s12517-019-4417-x.

O'Donoghue, C., Buckley, C., Chyzheuskaya, A., Green, S., Howley, P., Hynes, S., Upton, V., Ryan, M., 2021. The spatial impact of rural economic change on river water quality. Land Use Policy 103, 105322. https://doi.org/10.1016/j.landusepol.2021.105322.
Olayide, O., 2009. Determinants of fertilizer use in Northern Nigeria. Pak. J. Soc. Sci. 2009 (91), 98.
Qin, X., Li, Y., Lu, Z., Pan, W., 2020. What makes better village economic development in traditional agricultural areas of China? Evidence from 338 villages. Habitat Int. 106, 102286 https://doi.org/10.1016/j.habitatint.2020.102286.
Ren, C., Liu, S., van Grinsven, H., Reis, S., Jin, S., Liu, H., Gu, B., 2019. The impact of farm size on agricultural sustainability. J. Clean Prod. 220, 357–367. https://doi.org/10.1016/j.jclepro.2019.02.151.
Saha, S., Arabameri, A., Saha, A., et al., 2020. Prediction of landslide susceptibility in rudraprayag, india using novel ensemble of conditional probability and boosted regression tree-based on cross-validation method. Sci. Total Environ. 764, 42928. https://doi.org/10.1016/j.scitotenv.2020.142928.
Segurado, P., Almeida, C., Neves, R., Ferreira, M.T., Branco, P., 2018. Understanding multiple stressors in a Mediterranean basin: Combined effects of land use, water scarcity and nutrient enrichment. Sci. Total Environ. 624, 1221–1233. https://doi.org/10.1016/j.scitotenv.2017.12.201.
Solow, R., 1956. A contribution to the theory of economic growth. Quart. J. Econ. 70 (1), 65–94. https://doi.org/10.2307/1884513.
Stark, Oded, David, E. Bloom, 1985. The New Economics of Labor Migration. Am. Econ. Rev. 75 (2), 173–178.
Takeshima, H., Adhikari, R.P., Kaphle, B.D., Shivakoti, S., Kumar, A., 2016. Determinants of chemical fertilizer use in Nepal: Insights based on price responsiveness and income effects. Social Science Electronic Publishing. https://doi.org/10.2139/ssrn.2740540.
Tan, D., Adedoyin, F.F., Alvarado, R., Ramzan, M., Kayesh, M.S., Shah, M.I., 2022. The effects of environmental degradation on agriculture: Evidence from European countries. Gondwana Res. 106, 92–104. https://doi.org/10.1016/j.gr.2021.12.009.
Tefera, T., Elias, E., Beek, C., 2020. Determinants of smallholder farmers' decisions on fertilizer use for cereal crops in the Ethiopian highlands. Exp. Agr. 56, 677–687. https://doi.org/10.1017/s001447972000023x.
Therneau, T., Atkinson, B., Ripley, B., 2015. rpart: recursive partitioning and regression trees. R package version 4, 1–9.
Tsiboe, F., Egyir, I.S., Anaman, G., 2021. Effect of fertilizer subsidy on household level cereal production in Ghana. Sci. Afr. 13 https://doi.org/10.1016/j.sciaf.2021.e00916.
Vasco, C., Torres, B., Jácome, E., Torres, A., Eche, D., Velasco, C., 2021. Use of chemical fertilizers and pesticides in frontier areas: A case study in the Northern Ecuadorian Amazon. Land Use Policy 107, 105490. https://doi.org/10.1016/J.LANDUSEPOL.2021.105490.
Wang, C., Feng, B., Wang, P., Guo, W., Li, X., Gao, H., Zhang, B., Chen, J., 2022. Revealing factors influencing spatial variation in the quantity and quality of rural domestic sewage discharge across China. Process Saf. Environ. Prot. 162, 200–210. https://doi.org/10.1016/j.psep.2022.03.071.
Wu, Y., 2011. Chemical fertilizer use efficiency and its determinants in China's farming sector. China Agric. Econ. Rev. 3, 117–130. https://doi.org/10.1108/17561371111131272.
Wu, J., Zeng, S., Yang, L., Ren, Y., Xia, J., 2021a. Spatiotemporal characteristics of the water quality and its multiscale relationship with land use in the Yangtze River Basin. Remote Sens. 13, 3309. https://doi.org/10.3390/rs1316330.
Wu, J., Jin, Y., Hao, Y., et al., 2021b. Identification of the control factors affecting water quality variation at multi-spatial scales in a headwater watershed. Environ. Sci. Pollut. Res. 28, 11129–11141. https://doi.org/10.1007/s11356-020-11352-4.
Xu, G., Xiao, J., Oliver, D.M., Yang, Z., Xiong, K., Zhao, Z., Zheng, L., Fan, H., Zhang, F., 2021. Spatio-temporal characteristics and determinants of anthropogenic nitrogen and phosphorus inputs in an ecologically fragile karst basin: Environmental responses and management strategies. Ecol. Ind. 133, 108453. https://doi.org/10.1016/j.ecolind.2021.108453.
Yang, D.T., Zhu, X., 2013. Modernization of agriculture and long-term growth. J. Monetary Econ. 60, 367–382. https://doi.org/10.1016/j.jmoneco.2013.01.002.
Ye, J., 2015. Land transfer and the pursuit of agricultural modernization in China. J. Agrar. Change 15, 314–337. https://doi.org/10.1111/joac.12117.
Zhang, Y., Long, H., Tu, S., et al., 2019. Spatial identification of land use functions and their tradeoffs/synergies in China: Implications for sustainable land management. Ecol. Ind. 107, 105550. https://doi.org/10.1016/j.ecolind.2019.105550.
Zhang, Y., Long, H., Li, Y., Ge, D., Tu, S., 2020a. How does off-farm work affect chemical fertilizer application? Evidence from China's mountainous and plain areas. Land Use Policy 99, 104848. https://doi.org/10.1016/j.landusepol.2020.104848.
Zhang, Y., Long, H., Wang, M.Y., Li, Y., Ma, L., Chen, K., Zheng, Y., Jiang, T., 2020b. The hidden mechanism of chemical fertiliser overuse in rural China. Habitat Int. 102, 102210. https://doi.org/10.1016/j.habitatint.2020.102210.
Zhou, Q., Daryanto, S., Xin, Z., Liu, Z., Liu, M., Cui, X., Wang, L., 2017. Soil phosphorus budget in global grasslands and implications for management. J. Arid Environ. 144, 224–235. https://doi.org/10.1016/j.jaridenv.2017.04.008.
Zorzal-Almeida, S., Salim, A., Andrade, M.R.M., de Novaes Nascimento, M., Bini, L.M., Bicudo, D.C., 2018. Effects of land use and spatial processes in water and surface sediment of tropical reservoirs at local and regional scales. Sci. Total Environ. 644, 237–246. https://doi.org/10.1016/j.scitotenv.2018.06.361.