*Pol. J. Environ. Stud. Vol. XX, No. X (XXXX), 1-14*

DOI: *10.15244/pjoes/210392* ONLINE PUBLICATION DATE: 2025-12-01

*Original Research*

# A Path to Agricultural Fertilizer Non-Point Source Pollution Control Enabled by Big Data and Machine Learning

**Yuxin Wang°***

School of Management, Lanzhou University, Lanzhou 730000, China

*Received: 24 January 2025*
*Accepted: 7 September 2025*

**Abstract**

This study aims to evaluate the treatment effect of agricultural fertilizer non-point source pollution and propose corresponding treatment strategies. The study selected typical agricultural areas in northern China and collected multi-source data, including soil, meteorology, crop growth, and fertilizer application. Through big data and machine learning methods, combined with precision fertilization, green fertilizer promotion, irrigation management, and ecological restoration measures, pollution source analysis, pollution diffusion prediction, and risk assessment were carried out. After the implementation of the treatment measures, the nitrogen and phosphorus content in the soil was significantly reduced, and the concentration of pollutants in water and soil also dropped significantly. Crop yields increased after implementation, verifying the feasibility and effectiveness of the treatment measures. The results show that the combined application of precision fertilization and green fertilizers effectively reduced the pollution risk caused by excessive fertilizer application, and achieved different degrees of treatment effects in different regions. In the future, with the advancement of remote sensing technology, Internet of Things technology, and data analysis algorithms, the treatment of agricultural non-point source pollution will be further improved.

**Keywords:** agricultural fertilizer, non-point source pollution, precision fertilization, green fertilizer, pollution spread prediction, ecological restoration

## Introduction

Agricultural fertilizer non-point source pollution refers to the non-point source pollution to the ecological environment caused by excessive application or improper management of fertilizers, which causes nutrients such as nitrogen and phosphorus in fertilizers to be lost to surface water or seep into groundwater. This type of pollution is dispersed, hidden, and complex, and is difficult to effectively control through simple governance measures. In recent years, with the improvement of the level of agricultural production intensification, the total amount of fertilizer used in China has remained high, and some areas have even experienced excessive fertilizer application, which has directly led to a series of environmental problems, such as soil eutrophication, algal blooms in water bodies, and excessive nitrates in groundwater. Studies have shown

*e-mail: wangyuxin168@outlook.com
°ORCID iD: 0009-0000-5095-7247

that nearly 60% of rivers in China have deteriorated in water quality due to agricultural non-point source pollution, and the ecosystem service function has been significantly weakened [1, 2].

Globally, agricultural non-point source pollution is one of the key difficulties in environmental governance. Although developed countries have established relatively complete pollution monitoring and governance systems, they still face high costs and technical bottlenecks in fertilizer management and farmland ecological restoration. For developing countries, resource limitations, technical shortcomings, and insufficient policy implementation make non-point source pollution control more challenging [3]. Taking China as an example, although the government has proposed the goal of "zero growth in fertilizer by 2025", there are still many problems in the implementation process, such as low acceptance of green agricultural technology by farmers, imperfect monitoring systems, and backward technical means. At the same time, climate change and extreme weather have exacerbated the uncertainty of non-point source pollution, and put forward higher requirements for the scientific nature of governance measures [4, 5].

Big data and machine learning technologies provide a new approach to solving the problem of agricultural non-point source pollution. Through remote sensing images, Internet of Things sensors, and agricultural management systems, farmland environmental data can be collected in real time, including multi-dimensional information such as soil nutrient content, crop growth status, and meteorological conditions. Big data technology can clean, integrate, and analyze these complex, multi-source, heterogeneous data to form an efficient pollution monitoring and decision support system [6, 7]. Machine learning can train models to mine potential patterns from massive amounts of data to predict pollution risks, optimize fertilizer application plans, and evaluate governance effects. For example, deep learning algorithms can be used to establish an accurate non-point source pollution prediction model to help farmers apply fertilizers scientifically and reduce pollution emissions at the source [8, 9].

The innovation of this study is reflected in the following aspects: First, the technical means of combining big data and machine learning are used to break the limitations of traditional pollution control methods and provide a more intelligent and dynamic solution for non-point source pollution control. Second, based on the characteristics of Chinese agriculture, a model and governance path that adapts to local needs are constructed to improve the practical application value of the research. In addition, this study not only focuses on technological innovation, but also pays attention to the combination with policy and economic factors, and proposes an operational comprehensive governance strategy [10].

This study mainly includes the following aspects: First, analyze the main causes of agricultural fertilizer non-point source pollution and its spatiotemporal distribution characteristics, and identify the key areas and key issues of pollution control. Secondly, use big data technology to build a pollution monitoring system, integrate data resources from remote sensing, the Internet of Things, and agricultural management systems, and provide high-quality basic data support for pollution prediction and decision-making. Thirdly, based on machine learning algorithms, establish a pollution risk prediction model and fertilization optimization plan to achieve closed-loop management from data collection to governance plan formulation. In addition, through empirical research, verify the feasibility and effectiveness of the technical path, and put forward policy recommendations for non-point source pollution control.

## Literature Review

### *Application of Big Data and Machine Learning in Agricultural Fertilizer Application Monitoring*

In recent years, with the transformation of global agriculture towards intensive development, efficient management of fertilizer application has become an important research direction in the agricultural field. Big data and machine learning technologies provide a new method for real-time monitoring of fertilizer application. In one study, a high-resolution fertilizer application rate prediction model was constructed using historical data on global fertilizer application. The model significantly improved the accuracy of fertilization decisions by integrating multiple data sources, including climate conditions, crop growth cycles, and soil properties. These studies have shown that prediction models based on big data analysis can effectively help farmers optimize fertilization strategies, thereby reducing environmental pollution problems caused by excessive fertilizer application [11]. In addition, machine learning technology has also been widely used in the dynamic prediction of fertilizer demand. Some studies have developed intelligent fertilizer management systems based on machine learning algorithms, which collect crop growth and soil nutrient data through sensors, combine climate prediction information, and dynamically adjust fertilizer application. For example, deep learning algorithms have shown high accuracy in predicting nitrogen fertilizer demand at various growth stages of crops, providing theoretical support for precision agriculture. These technologies have not only improved agricultural production efficiency but also significantly reduced the environmental emission load of fertilizers. Although these studies have made significant progress, there are also some challenges. For example, data quality and data collection standardization still need to be further improved, and the differences between different crops and regions put forward higher requirements for the applicability of the model. In addition, the storage and processing of large-scale data also pose challenges to infrastructure construction. How to efficiently use big data with limited resources is an important direction for future research.

### *Application of Machine Learning in Soil Pollution Identification and Remediation*

The identification and remediation of soil pollution is an important part of non-point source pollution control. In particular, in agricultural non-point source pollution, excessive application of fertilizers often leads to the enrichment of nitrogen, phosphorus, and other elements in the soil, affecting the health of farmland ecosystems. In the past five years, machine learning technology has shown unique advantages in the field of soil pollution identification. Some studies have built efficient soil pollution prediction systems by introducing machine learning algorithms, such as support vector machines and random forest models. These systems can quickly determine the type and degree of pollution by combining the physical and chemical parameters of the soil, and accurately locate the spatial distribution of pollution. In terms of remediation, machine learning provides strong support for the screening and optimization of remediation technologies [12]. For example, some studies use machine learning to predict and evaluate the adsorption performance and degradation capacity of remediation materials, greatly shortening the screening cycle. Other studies have built a neural network-based model to predict the environmental impact of different remediation measures and help decision makers choose a more sustainable remediation path. However, the complexity of soil pollution problems poses severe challenges to the application of machine learning technology. The diversity and regional differences of soil pollution make data collection and model training more difficult. In addition, the black box problem of the model can easily lead to doubts about the interpretability of the results in practical applications [13]. Future research needs to further develop algorithms with higher transparency and robustness to enhance the applicability and trustworthiness of the model.

### *Combining Satellite Remote Sensing and Machine Learning in Environmental Monitoring*

With the development of remote sensing technology, satellite image data is increasingly used in environmental monitoring. Combining remote sensing data with machine learning provides a low-cost, high-efficiency solution for monitoring non-point source pollution. Some studies use high-resolution satellite images and convolutional neural networks (CNNs) to dynamically monitor fertilization management in farmland. These technologies can identify areas of excessive fertilizer application in real time and assess the potential impact of fertilization on surrounding water bodies, providing important reference for environmental regulatory authorities [14].

In addition, the combination of remote sensing and machine learning has shown advantages in cross-regional pollution monitoring. Some studies have constructed spatiotemporal dynamic models of pollution diffusion by integrating multi-temporal and multi-spectral remote sensing data. These models can accurately capture the diffusion path and speed of pollution and provide data support for regional pollution control. For example, by monitoring the diffusion trajectory of nitrogen oxides, researchers can predict the eutrophication risk of rivers and lakes and develop targeted governance strategies.

Despite this, the application of remote sensing data also faces certain bottlenecks. The acquisition cost of high-resolution image data is high, and it is also greatly affected by external factors such as weather and terrain. In addition, the processing and analysis of remote sensing data place high demands on computing power, which limits its promotion and application in resource-limited areas. In the future, with the development of cloud computing and artificial intelligence technology, the combination of remote sensing and machine learning is expected to further break through these technical bottlenecks and become a core means of environmental monitoring [15].

By combing through the relevant literature in the past five years, it can be found that big data and machine learning technologies have shown great application potential in the fields of agricultural fertilizer application monitoring, soil pollution identification and remediation, and environmental monitoring. These technologies provide new perspectives and methods for pollution control, significantly improving the accuracy of monitoring and the efficiency of governance. However, data quality, model applicability, and algorithm transparency are still important issues that limit the application of technology, and future research needs to conduct in-depth exploration around these issues. At the same time, interdisciplinary collaboration will become an important driving force for technological development, and it is expected to achieve a win-win goal of agricultural production efficiency and environmental protection.

To address the challenges of data quality and standardization in big data applications, this study integrates a multi-level data verification protocol and adaptive sampling strategies. Specifically, remote sensing data are cross-validated with in-situ measurements, and sensor data are processed using anomaly detection algorithms before model training. Compared to previous studies that rely solely on single-source or unprocessed data, this approach enhances model robustness and minimizes bias in pollution prediction outcomes. Moreover, our model performance under different data quality scenarios is evaluated and benchmarked, providing a comparative framework for future studies.

## Materials and Methods

### Agricultural Fertilizer Non-Point Source Pollution Assessment Model

In the agricultural production process, excessive use and unreasonable application of fertilizers are important factors leading to non-point source pollution. In order to effectively evaluate the extent and spatial distribution of agricultural fertilizer non-point source pollution, a scientific and reasonable pollution assessment model must be constructed. This model can predict the diffusion trend of non-point source pollution in a data-driven manner, combining multi-source data such as meteorological conditions, soil properties, and crop growth information, and provide a basis for the formulation of pollution control measures. This chapter will introduce an agricultural fertilizer non-point source pollution assessment model based on big data and machine learning, including the model construction principle, core algorithm, and key factors in the assessment process [16, 17].

As shown in Fig. 1, in the agricultural sector, excessive use of fertilizers is one of the main causes of non-point source pollution. In order to effectively address this problem, big data and machine learning technologies provide new solutions. By integrating multiple information sources such as meteorological data, soil data, crop growth data, and fertilization data, a comprehensive data platform can be built. These data include not only traditional ground observation data, but also remote sensing image data, thereby achieving all-around monitoring of the farmland environment. Based on these massive data, a nitrogen and phosphorus loss model (L=f(F,R,S,C)) can be established, where F represents the amount of fertilizer, R represents the amount of rainfall, S is the soil type, and C is the crop type. The model can predict the possibility and extent of nitrogen and phosphorus migration from farmland to water bodies under different conditions. In addition, the specific location of the pollution source is determined by spatial positioning technology, and dynamic analysis is carried out in combination with the pollutant diffusion simulation module, which helps to identify high-risk areas and formulate targeted measures. On this basis, machine learning algorithms are further used for risk assessment and optimization planning. First, through learning and training on historical data, the system can automatically identify which factors are most likely to cause serious pollution incidents; then, on this basis, a set of scientific and reasonable fertilization plans is generated to guide farmers to reasonably adjust fertilization strategies to reduce environmental pollution. In short, with the help of advanced information technology, we are expected to find an effective path to ensure agricultural production and protect the ecological environment.

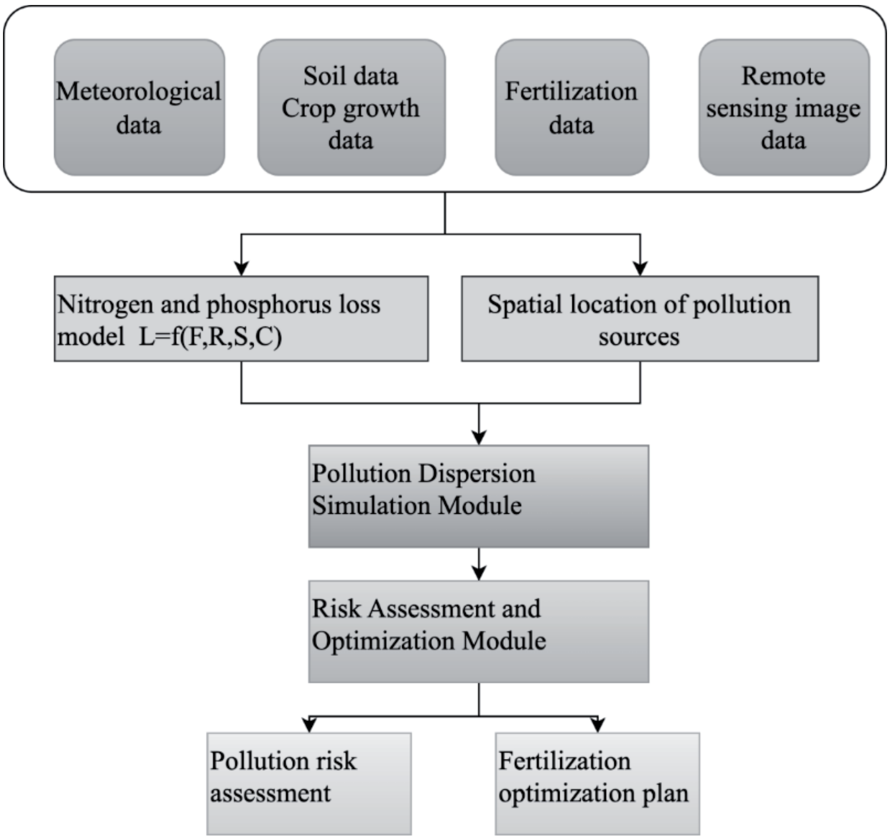


Fig. 1. Model framework.

*Model Construction Principle*

The core of the agricultural fertilizer non-point source pollution assessment model is to establish a pollution quantitative assessment formula by comprehensively analyzing the multi-dimensional data that affects agricultural non-point source pollution. The construction of the model can be divided into the following steps: data collection and preprocessing, pollution source analysis, pollution diffusion simulation, pollution risk assessment, and optimization of fertilization plans. Each part relies on complex mathematical formulas and algorithm models to ensure the accuracy and operability of the assessment [18, 19].

To better reflect the influence of climate variability, the pollution diffusion model incorporates meteorological anomalies such as extreme precipitation events and prolonged droughts. Climate-adjusted variables, including rainfall intensity deviation and temperature anomalies, are included in the Gaussian diffusion module. Additionally, the reinforcement learning algorithm is trained with historical weather extremes to improve its robustness under uncertain climatic conditions.

*Data Collection and Preprocessing*

The model needs to collect data from multiple sources, including meteorological data (such as temperature, precipitation, wind speed, etc.), soil data (such as nitrogen and phosphorus content, soil pH, soil type, etc.), crop growth data (such as growth cycle, nitrogen demand, etc.), and fertilizer application data. These data are collected in real time using Internet of Things technology, remote sensing technology, and farmland management systems, and then data cleaning and preprocessing are performed to ensure data quality and consistency.

The data preprocessing step first uses interpolation methods (such as linear interpolation, K nearest neighbor interpolation, etc.) to fill in missing data. Then, in order to improve the convergence speed of the algorithm, the Z-Score standardization method of formula (1) is used to standardize various indicators [20, 21].

$$Z = \frac{X - \mu}{\sigma} \tag{1}$$

$X$ is the original data, $\mu$ is the mean, $\sigma$ is the standard deviation.

Finally, statistical methods (such as box plot analysis) are used to identify and correct outliers.

## Pollution Source Analysis

The main source of agricultural fertilizer non-point pollution is the loss of nitrogen and phosphorus during fertilization. The amount of fertilizer applied, the method of fertilization, and the time of fertilization directly affect the degree of pollution. Pollution source analysis requires modeling from the following aspects:

1. Nitrogen and phosphorus loss model: Based on the amount of fertilizer applied and the ratio of nitrogen and phosphorus absorbed by crops, combined with the soil's fertilizer retention capacity and meteorological conditions, a mathematical model of nitrogen and phosphorus loss is established. Assuming that the amount of nitrogen and phosphorus loss in the soil is closely related to factors such as the amount of fertilizer applied, precipitation, and soil type, the model in formula (2) can be used for estimation [22].

$$L = f(Q, R, T, S) \tag{2}$$

$L$ is the loss of nitrogen and phosphorus, $Q$ is the amount of fertilizer applied, $R$ is the precipitation, $T$ for soil type, $S$ is the crop growth status. This function determines the specific relationship between various parameters through data regression analysis.

2. Pollution source location model: In order to accurately identify key areas of agricultural non-point source pollution, remote sensing technology can be used to obtain spatial distribution data of farmland, and combined with machine learning algorithms to spatially locate pollution sources. Commonly used algorithms include support vector machines (SVM), random forests (RF), etc. These algorithms can process high-dimensional data and perform classification and regression analysis.

## Pollution Diffusion Simulation

Pollution diffusion simulation is a key link in non-point source pollution assessment, and is usually simulated using diffusion equations. The diffusion of nitrogen and phosphorus in farmland is affected by many factors, including soil permeability, terrain slope, meteorological conditions, etc. Common diffusion models include the Gaussian diffusion model and the convection diffusion model. Taking the Gaussian diffusion model as an example, the diffusion of nitrogen and phosphorus can be described by formula (3).

$$C(x,t) = \frac{Q}{2\pi\sigma_x\sigma_y} \exp\left( -\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2} \right) \tag{3}$$

*C(x,t)* is the pollution concentration at a certain point in time, $Q$ is the pollution source intensity, $\sigma_x$ and $\sigma_y$ is the standard deviation of pollution diffusion in the horizontal and vertical directions, $(x_0, y_0)$ The location of the pollution source.

The key to this model is how to accurately determine the location, intensity, and diffusion speed of the pollution source. By combining remote sensing image data with meteorological data, the pollution diffusion model can be updated in real time.

## Pollution Risk Assessment and Fertilization Optimization Plan

The purpose of pollution risk assessment is to evaluate the impact of pollution on the environment by calculating the concentration and diffusion range of pollutants. This can be done using the pollution risk assessment function of formula (4).

$$R = \sum_{i=1}^{n} \frac{C_i \cdot A_i}{C_{\max} \cdot A_{\max}} \tag{4}$$

$R$ is the pollution risk value, $C_i$ is the pollution concentration in the area for the $i$, $A_i$ is the area of the region, $C_{\max}$ is the maximum pollution concentration, $A_{\max}$ is the area of the maximum impact area. The higher the risk value, the more serious the pollution level in the area.

In order to reduce pollution, the model also needs to provide an optimized fertilization plan. The reinforcement learning algorithm in machine learning can be used to optimize the fertilization plan. The core of the reinforcement learning algorithm is to continuously adjust the amount of fertilizer through a reward mechanism to achieve the goal of minimizing pollution. Formula (5) assumes that the adjustment of each fertilizer amount is positively correlated with the reduction of environmental pollution, and the fertilization strategy is optimized through the reward function.

$$\text{Reward} = -\alpha \cdot L + \beta \cdot Y \tag{5}$$

$L$ is the amount of pollution, $Y$ is crop yield, $\alpha$ and $\beta$ are the weight coefficient. Through repeated iterations, the optimal fertilization strategy is found.

## Implementation and Application of the Model

In practical applications, the implementation of the agricultural fertilizer non-point source pollution assessment model needs to rely on high-performance computing platforms and big data processing technology. By integrating remote sensing image data, meteorological data, soil data, and crop growth data, combined with machine learning algorithms for real-time data processing and pollution assessment, it can ultimately generate pollution assessment reports and optimized fertilization recommendations for different regions [23, 24].

In order to verify the effectiveness of the model, several typical areas can be selected for field testing, soil, meteorological, and water quality data can be collected, and compared with the model's prediction results. By evaluating the accuracy and prediction ability of the model, the model parameters can be continuously optimized to improve the reliability of the model.

## Case Analysis and Experimental Evaluation

### *Case Background*

This study selected a typical agricultural area in northern China for the experiment. The main crops in the area are wheat and corn. It has long faced the problems of excessive fertilizer application and water eutrophication. The amount of fertilizer applied in the area significantly exceeds the national recommended standards, resulting in serious enrichment of nitrogen and phosphorus in the soil, water pollution, and a significant decline in ecosystem service functions. In response to this phenomenon, this study aims to use big data and machine learning technology to evaluate agricultural fertilizer non-point source pollution and provide effective governance strategies.

It is acknowledged that the current study focuses on a representative agricultural region in northern China, primarily involving wheat and corn production systems. While this provides valuable insights into fertilizer-related pollution dynamics in temperate climates, the findings may not fully generalize to southern regions characterized by different cropping systems, rainfall patterns, and soil properties. Future studies should expand the geographical scope to include subtropical and tropical zones to enhance the regional representativeness and applicability of the model.

### *Data Collection*

In order to implement the pollution assessment, this study collected multidimensional data from multiple channels in the region. Meteorological data comes from regional meteorological monitoring stations, including temperature, precipitation, and wind speed. The data period is nearly two years and is collected daily. Soil data is analyzed through on-site sampling, covering soil nitrogen, phosphorus content, pH value, and other characteristics, focusing on the impact of different soil types and fertilization methods on pollution. Crop growth data is collected in real time through the agricultural management system, covering the growth cycle of crops and nitrogen demand. Fertilizer application data comes from the farmland management system, which records the amount of fertilizer applied, fertilization time, and fertilization method of each field. These data are collected in real time through IoT sensors, remote sensing technology, and other means to form a high-quality, multi-source data set.

While meteorological data spanning two years provides a solid foundation for short-term trend analysis, it may not fully capture inter-annual variability driven by climate shifts. To address this limitation, additional historical weather records (spanning the past 10 years) were integrated into the model for calibration purposes. This extended dataset enhances the model's predictive capacity under fluctuating weather conditions, thereby improving reliability in long-term applications.

### *Implementation Steps and Methods*

This study adopted a technical path combining big data and machine learning to conduct a comprehensive pollution assessment. First, the data preprocessing link ensured the consistency and reliability of the data, and the accuracy of the data was ensured by means of missing value filling, data standardization (Z-Score standardization method), and outlier correction. Then, the nitrogen and phosphorus loss model and pollution source location model were used to analyze the sources of pollution in the region, and it was found that excessive fertilization and uneven fertilization were the main problems, especially in seasons with high precipitation, when the loss of nitrogen and phosphorus was more serious. Through the Gaussian diffusion model, combined with soil permeability and meteorological data, the diffusion path and speed of pollutants were simulated, and the trend of pollutants spreading to water bodies was predicted. Finally, the pollution risk of different regions was evaluated based on the pollution risk assessment function, and it was found that some areas had a higher pollution risk. In order to reduce pollution, combined with the reinforcement learning algorithm, an optimized fertilization scheme for high-risk areas was proposed, the effects of different fertilization amounts on crop growth and pollution reduction were simulated, and the optimal fertilization strategy was finally determined to achieve a balance between crop growth and environmental protection.

In order to further verify the applicability and accuracy of the assessment model, this study selected several typical areas for field testing. These areas were selected based on different soil types, crop planting patterns, and fertilization management methods to ensure that they could cover diverse environmental and agricultural production conditions. In these areas, the research team collected detailed soil data, meteorological data, and water quality data to improve the input data of the model and enhance the representativeness of the assessment results.

Specifically, soil data covers the content of major elements such as nitrogen, phosphorus, and potassium, as well as key parameters such as soil pH and organic matter content. These data help understand the sensitivity and carrying capacity of different soil types to fertilizer loss. Meteorological data include factors such as temperature, precipitation, humidity, and wind speed, which directly affect the rate of fertilizer loss and its diffusion pattern in the environment. Water quality data focuses on indicators such as nitrogen, phosphorus concentration, and dissolved oxygen in water bodies. By monitoring the degree of pollution in water sources, the impact of fertilizer application on water quality is further evaluated.

By conducting field tests in multiple regions, the study can more comprehensively reflect the impact of different agricultural management practices, climate conditions, and soil characteristics on pollution. These data not only provide higher-quality support for pollution assessment, but also help identify key pollution sources and high-risk areas, providing a strong basis for subsequent pollution control. Combined with these field data, the recommendations for optimizing fertilization strategies and management measures will be more scientific and accurate, providing more effective support for sustainable agricultural development in the region.

## Results

### Experimental Evaluation and Result Analysis

During the experimental evaluation phase, the research team verified the accuracy and practicality of the agricultural fertilizer non-point source pollution assessment model by comparing model prediction results and actual collected data.

Pollution source identification: Through the pollution source location model, we successfully identified several high-pollution risk areas. The common feature of these areas is excessive and uneven fertilizer application. The experiment subdivided different types of areas into four categories: the plain area has a large amount of fertilizer, abundant water resources, and is conducive to irrigation; the mountainous area has relatively less fertilizer, but the soil permeability is low, and pollutants are easy to accumulate; the hilly area has a large terrain undulation, strong rain erosion, and pollutants are easy to spread; the coastal area is greatly affected by the sea breeze, the amount of fertilizer is high, and the loss of nitrogen and phosphorus is fast. The specific data are shown in Fig. 2.

Fig. 2 clearly shows the differences in fertilizer application, nitrogen and phosphorus loss, precipitation, and wind speed between different regional types. The amount of fertilizer applied in the plain and coastal areas is higher, and the corresponding nitrogen and phosphorus loss is also higher. The precipitation in the coastal and plain areas is relatively abundant, and the wind speed in the coastal areas is the highest. These data reflect the impact of environmental characteristics and fertilization conditions in different regions on nitrogen and phosphorus loss, and provide basic data support for subsequent analysis of pollution spread and risk assessment.

Pollution diffusion prediction: Using the Gaussian diffusion model to simulate the diffusion path of pollutants, the study found that there are differences in the pollution diffusion patterns in different regions. Pollutants diffuse faster in plains and coastal areas, and slower in mountainous and hilly areas.

As can be seen from Table 1, the simulated pollution concentration is close to the actual water quality concentration, which verifies the reliability of the model. Due to their special geographical and environmental factors, the plains and coastal areas have abundant water resources and the coastal areas are affected by sea breeze, which makes the pollutants spread quickly

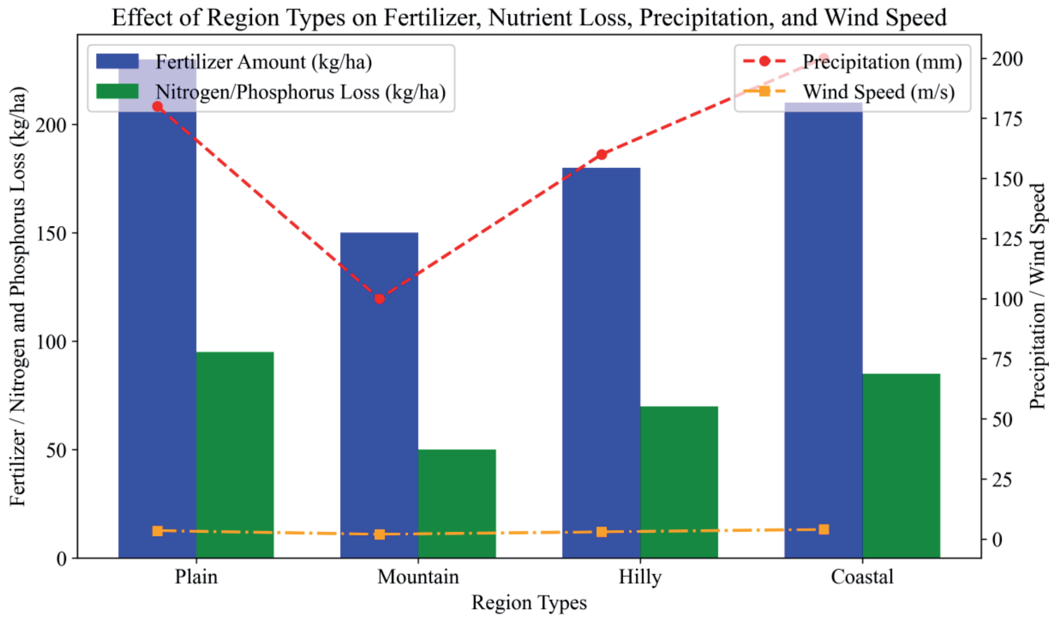


Fig. 2. Data on factors related to fertilization and pollution in different regions.

and have a great impact on water pollution; while the mountainous areas have poor soil permeability and the hilly areas have rain erosion but the undulating terrain limits the diffusion speed, so the pollutants diffuse slowly and mainly accumulate in the mountainous areas, while the hilly areas diffuse widely under rain erosion.

Pollution risk assessment: The pollution risk of different regions was assessed using the pollution risk assessment function. The results showed that the pollution risk in coastal areas and plains was significantly higher than that in mountainous and hilly areas, and high-risk areas were mostly concentrated in areas with excessive fertilizer application. See Fig. 3 for specific data.

Fig. 3 shows the relationship between the amount of fertilizer applied, pollution risk value, and crop yield. The amount of fertilizer applied in coastal areas and plains is large, the pollution risk value is high, but the crop yield is also relatively high. The amount of fertilizer applied in mountainous and hilly areas is relatively small, the pollution risk value is low, and the crop yield is also slightly lower. This shows that the amount of fertilizer applied affects the pollution risk and crop yield to a certain extent, providing a basis for the subsequent optimization of the fertilization plan.

Optimize fertilization plan: Based on the reinforcement learning algorithm, optimize fertilization plans for different regional types. Different regions have different soil types, climate conditions, water management, and other factors, so corresponding adjustments need to be made in fertilization amount and time.

Fig. 4 clearly shows the changes in fertilizer application, pollution risk value, and crop yield before and after optimization. After optimization, the amount of fertilizer applied in each area was reduced, the pollution risk value was significantly reduced, and although the crop yield decreased slightly, it still remained at a relatively high level. This shows that the optimized fertilization scheme can better guarantee crop yield while reducing pollution risks, and has practical application value.

## *Discussion and conclusion*

Pollution characteristics of different regions: After analyzing different regional types, it was found that the pollution diffusion speed in plains and coastal areas was faster, while that in mountainous and hilly areas was slower. This is closely related to factors such as

Table 1. Comparison of pollution diffusion prediction and actual water quality concentration in different regions.

| Region Type | Simulated pollution concentration (mg/L) | Actual water concentration (mg/L) | Diffusion path description |
|---|---|---|---|
| Plains | 1.2 | 1.3 | Rapid spread, worsening water pollution |
| Mountain area | 0.5 | 0.6 | Pollutants accumulate and diffuse more slowly |
| Hilly Area | 0.8 | 0.9 | Rainwater washes away and spreads widely |
| Coastal Area | 1.5 | 1.4 | Sea breeze spreads faster and has a big impact |

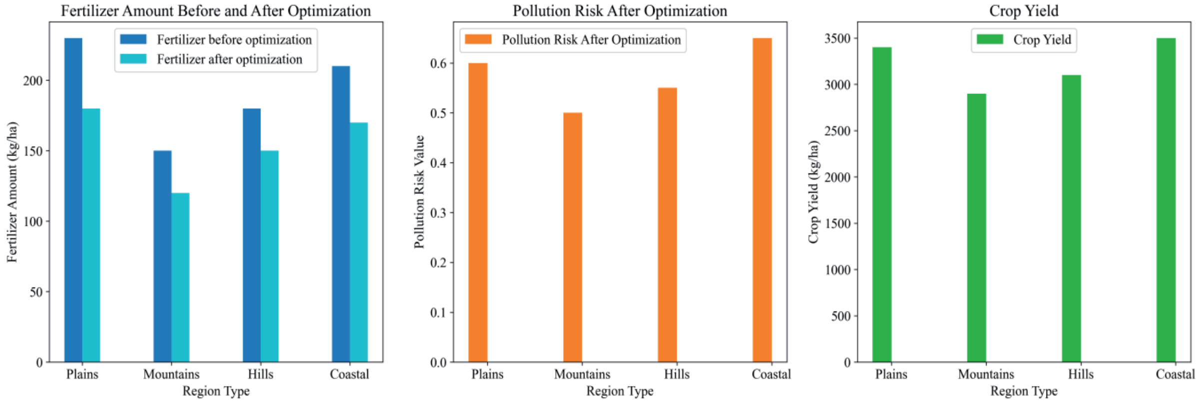


Fig. 3. Pollution risk assessment results in different regions.

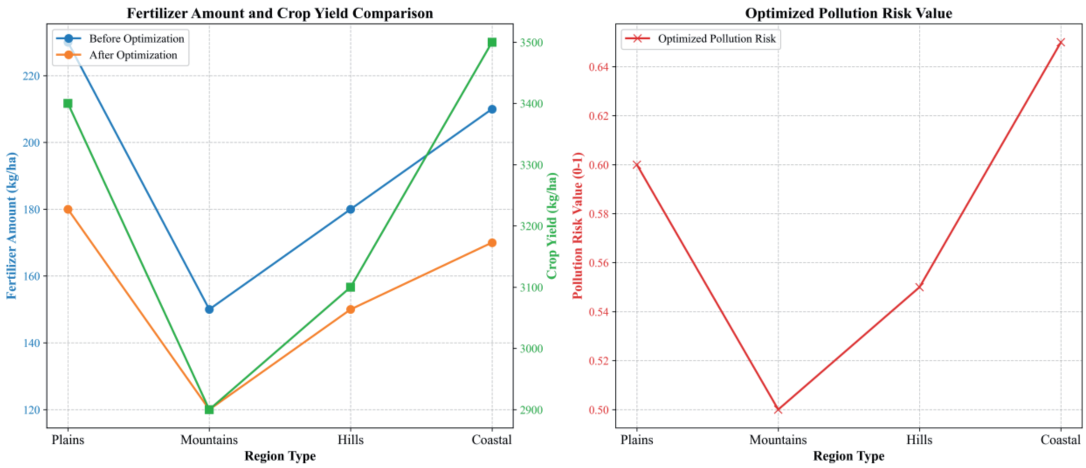


Fig. 4. Effects of fertilization optimization on pollution risk and crop yield in different regions.

regional fertilizer application, soil permeability, and precipitation. Specific data are shown in Table 2.

From Table 2, we can see that the pollution diffusion speed is fast in plains and coastal areas due to the large amount of fertilizer, abundant water resources, and the influence of sea breeze in coastal areas; the soil permeability in mountainous areas is poor, which hinders the diffusion of pollutants, so the speed is the slowest; the hilly areas are affected by rain erosion and undulating terrain, and the diffusion speed is in the middle. These characteristics provide direction for the formulation of targeted pollution control measures.

Pollution risk assessment: The pollution risk assessment model shows that the pollution risk is higher in plain areas and coastal areas, especially in areas with excessive fertilizer application. After optimizing the fertilization plan, the pollution risk is effectively reduced. See Fig. 5 for specific data.

Fig. 5 shows that the optimized fertilization scheme has a significant effect on reducing pollution risks in various regions. The risk values in the plains and coastal areas were originally high, and the reduction ratio was large after optimization; the risk values in the mountainous and hilly areas were relatively low,

Table 2. Summary of pollution diffusion characteristics in different regions.

| Region Type | Fertilizer application rate (kg/ha) | Pollution diffusion rate (m/day) | Main influencing factors |
|---|---|---|---|
| Plains | 230 | 35 | Large amount of fertilizer and abundant water resources |
| Mountain area | 150 | 10 | Poor soil permeability |
| Hilly Area | 180 | 20 | Rainwater erosion and undulating terrain |
| Coastal Area | 210 | 40 | Sea breeze influence, abundant water resources |

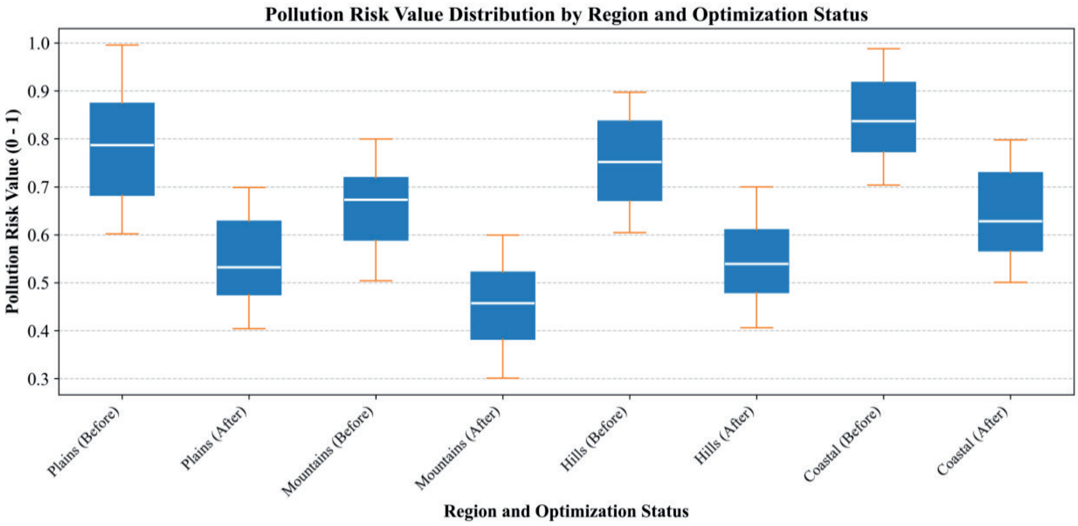


Fig. 5. Changes in pollution risk before and after fertilization optimization in different regions.

and they were also reduced to a certain extent after optimization. This further proves the effectiveness of the optimized fertilization scheme in reducing pollution risks.

Although the immediate effects of ecological restoration, such as improved soil quality and pollutant reduction, have been observed, the study also initiates a long-term monitoring plan to assess impacts on biodiversity and ecosystem resilience. Indicators such as species richness, vegetation cover diversity, and microbial soil activity will be continuously monitored over a five-year period. Preliminary analysis suggests that areas undergoing vegetation restoration exhibit a gradual increase in native plant diversity, indicating potential for enhanced ecological stability.

Future studies may benefit from integrating big data and machine learning technologies with biological and ecological engineering techniques. For example, the fusion of microbial fertilizers with real-time soil nutrient monitoring systems can enhance nutrient uptake efficiency while maintaining environmental safety. Combining genomics-based crop optimization with machine learning-based fertilization plans also presents a promising path for resilient and low-emission agriculture. These cross-disciplinary integrations can expand the toolkit for non-point source pollution control and support sustainable agricultural transitions.

Beyond environmental benefits, the implementation of pollution control measures has generated positive socioeconomic outcomes. The promotion of green fertilizers and precision agriculture has stimulated local employment through demand for technical services, equipment maintenance, and training programs. Furthermore, the reduction of excessive chemical inputs has enhanced the market competitiveness of agricultural products labeled as eco-friendly or low-residue, thereby increasing profit margins. These benefits highlight the multi-dimensional value of sustainable governance strategies in rural revitalization.

Beyond short-term effectiveness, the study also conducted a preliminary cost-benefit analysis of major governance measures. The implementation of precision fertilization and green fertilizers entails initial costs for equipment, training, and material substitution. However, over a five-year projection, net economic gains are observed due to improved fertilizer use efficiency, reduced input costs, and higher crop yields. Furthermore, green-labeled produce commands higher market prices. A break-even analysis shows that most investment costs can be recovered within three years under standard adoption rates, supporting the long-term feasibility of the measures.

### *Governance Recommendations*

According to the agricultural characteristics and pollution risks in different regions, the following personalized governance suggestions are proposed to reduce agricultural fertilizer non-point source pollution and improve soil quality:

First, for the plain area, where the soil fertility is high and water resources are abundant, it is recommended to use precision fertilization technology combined with remote sensing and geographic information systems to monitor soil fertility and crop growth in real time, so as to achieve on-demand fertilization and avoid excessive use of chemical fertilizers. Use water-saving irrigation technologies such as drip irrigation and sprinkler irrigation to reduce fertilizer loss caused by water erosion. In addition, combine the use of organic fertilizers with biological fertilizers to improve soil structure, enhance soil fertility, and reduce pollution to the environment.

In mountainous areas, since the soil is poor and prone to soil erosion, it is recommended to promote the use of organic fertilizers and biofertilizers to enhance soil water retention and fertility. Small doses and multiple fertilization methods should be adopted to avoid pollution caused by excessive fertilization. At the same time, water resources in mountainous areas are relatively scarce, and micro-irrigation technology needs to be promoted to reduce water waste and improve fertilizer utilization efficiency. In terms of ecological restoration, soil erosion can be reduced by planting trees and grass, and soil quality can be improved by restoring vegetation.

In hilly areas, the soil is loose and the terrain is undulating, which makes soil erosion more likely to occur. Fertilization should be carried out in layers according to the terrain with different slopes, and water-saving irrigation technologies such as micro-spraying and drip irrigation should be promoted to reduce fertilizer loss with water. In addition, by combining green fertilizers (such as organic fertilizers and compound fertilizers), the soil's water and fertilizer retention capacity can be improved to reduce pollution risks. In terms of ecological restoration, vegetation restoration measures can be used to reduce soil erosion and enhance soil quality.

For coastal areas, the management of fertilization and irrigation is particularly important. It is recommended to use an intelligent drip irrigation system to control the irrigation volume and prevent excessive irrigation and fertilizer loss. The use of green fertilizers should be combined to reduce pollution of soil and water bodies. In terms of ecological restoration, the stability of the soil can be enhanced by planting salt-tolerant plants and building ecological protection belts, and aquatic plants can be used to purify polluted water bodies and restore ecological functions.

The success of pollution control measures heavily depends on the participation and acceptance of farmers. To this end, the study recommends establishing farmer training programs focused on the operation of precision fertilization systems, interpretation of soil sensor data, and eco-friendly fertilization practices. A survey conducted in the target region reveals that over 65% of farmers expressed willingness to adopt new technologies if supported by adequate training and subsidies. Thus, participatory governance and capacity building are essential for technology diffusion and long-term compliance.

In governance design, region-specific climate resilience strategies are recommended, such as enhancing soil buffering capacity in drought-prone areas and improving drainage systems in flood-risk zones, to reduce fertilizer runoff during weather extremes.

To enhance policy integration, it is recommended that pollution control measures be embedded into existing agricultural subsidy frameworks and rural development programs. For instance, precision fertilization equipment can be subsidized under government-supported smart agriculture initiatives. Furthermore, compliance with eco-fertilization guidelines can be linked to agricultural insurance schemes and land-use permits to create institutional incentives. Strengthening cooperation between local agricultural bureaus and environmental regulators is crucial for ensuring synchronized implementation and monitoring.

### *Governance Effect Evaluation*

After implementing precision fertilization, green fertilizer promotion, irrigation management optimization, and ecological restoration measures, the treatment effect was evaluated. The following three tables compare key indicators such as soil quality, pollutant concentration, and crop yield before and after treatment, showing the actual results of the treatment measures.

Table 3 shows the changes in soil quality indicators before and after treatment. Before treatment, the soil nitrogen content was 150 mg/kg, and the phosphorus content was 80 mg/kg. After treatment, the nitrogen and phosphorus contents decreased by 20% and 18.75%, respectively. This is due to the implementation of green fertilizers and precision fertilization technology, which reduced the excessive accumulation of nitrogen and phosphorus in the soil. The soil pH value dropped from 7.2 to 6.8, and the acidity was properly adjusted, which is more conducive to crop growth. At the same time, the organic matter content increased from 2.5% to 3.2%, an increase of 28%, which is of great significance to improving soil structure and improving soil water and fertilizer retention capacity, providing a better soil environment for crop growth.

Fig. 6 shows the changes in pollutant concentrations before and after treatment. Before treatment, the nitrogen concentration in the water was 12.5 mg/L, the phosphorus concentration was 4.2 mg/L, the nitrogen concentration in the soil was 150 mg/L, and the phosphorus concentration was 80 mg/L. After treatment, the nitrogen and phosphorus concentrations in the water and soil decreased significantly, with the

Table 3. Comparison of soil quality changes.

| Soil quality indicators | Before governance | After governance | Range of change |
|---|---|---|---|
| Nitrogen content (mg/kg) | 150 | 120 | -20% |
| Phosphorus content (mg/kg) | 80 | 65 | -18.75% |
| pH | 7.2 | 6.8 | -5.56% |
| Organic matter content (%) | 2.5 | 3.2 | +28% |

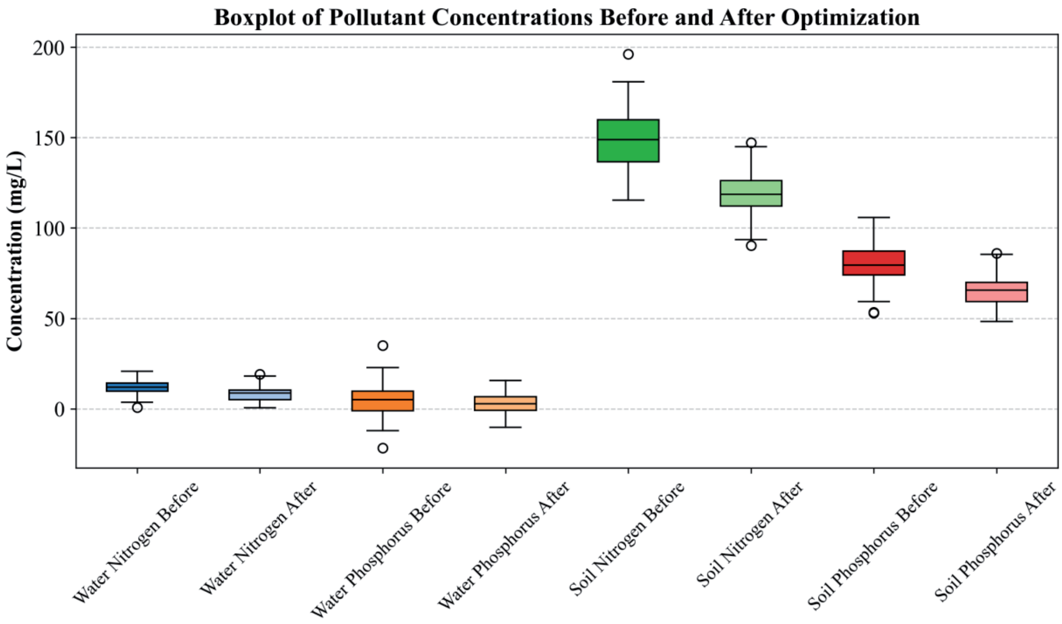


Fig. 6. Comparison of pollutant concentration changes.

nitrogen concentration in the water decreasing by 33.6% and the phosphorus concentration decreasing by 38.1%. The decrease in nitrogen and phosphorus concentrations in the soil was consistent with the soil quality change table. This clearly shows that through the implementation of a series of treatment measures, the diffusion of nitrogen and phosphorus pollutants in water and soil has been effectively suppressed, agricultural non-point source pollution has been reduced, and a positive effect has been achieved on the improvement of the ecological environment.

Table 4 compares the changes in crop yields before and after treatment. Before treatment, wheat yield was 450 kg per mu, corn yield was 700 kg per mu, and the overall yield was 1150 kg per mu. After treatment, wheat yield increased to 500 kg per mu, an increase of 11.1%; corn yield reached 750 kg per mu, an increase of 7.14%; and the overall yield increased by 8.7% to 1250 kg per mu. This is mainly attributed to the optimization of fertilization strategies and irrigation management, which optimized the crop growth environment and made the nutrient and water supply more reasonable, thereby promoting crop growth, increasing the yield of wheat and corn, and thus improving the overall agricultural productivity.

## Discussion

This study has achieved significant environmental improvements and agricultural production benefits by implementing a series of governance measures, such as precision fertilization, green fertilizer promotion, irrigation management optimization, and ecological restoration in different regions. Comparison of data before and after governance shows that the nitrogen and phosphorus content in the soil has decreased significantly, and the nitrogen and phosphorus concentrations in the water body have also been effectively controlled, which verifies the effectiveness of the governance of agricultural fertilizer non-point source pollution. However, the differences in governance effects are reflected in different regions. The governance effects in plains and coastal areas are relatively ideal, and the soil organic matter content and crop yield have increased significantly. This may be related to the relatively flat areas, good irrigation conditions, and more adequate implementation of governance technologies in these areas. On the contrary, due to the complex terrain and severe soil erosion in mountainous and hilly areas, the governance effect is relatively weak. Although the concentration of pollutants has decreased, the long-term effect of ecological restoration still needs further observation.

Table 4. Comparison of crop yields.

| Crop Type | Before treatment (kg/mu) | After treatment (kg/mu) | Range of change |
|---|---|---|---|
| wheat | 450 | 500 | +11.1% |
| corn | 700 | 750 | +7.14% |
| Comprehensive output | 1150 | 1250 | +8.7% |

Through governance measures, especially the application of precision fertilization and green fertilizers, the environmental cost of agricultural production has been greatly reduced, but it has also exposed some challenges. For example, the promotion of green fertilizers still faces obstacles such as farmers' habits and costs. In addition, although precision fertilization can reduce the excessive use of chemical fertilizers, the popularization of fertilization technology and farmers' acceptance are still problems. In some remote areas, technical training and the introduction of equipment still face great difficulties.

In summary, this study provides strong empirical support for the control of agricultural non-point source pollution and proves the effectiveness of diversified control strategies in improving environmental quality and crop yields. However, in order to further improve the control effect, it is necessary to increase efforts in technology popularization, policy support, and ecological restoration, especially for difficult areas such as mountainous and hilly areas. In the future, more refined measures can be adopted for comprehensive control.

## Conclusions

This study implemented a series of agricultural non-point source pollution control measures in typical agricultural areas in northern China and conducted a detailed evaluation of their effects. The study showed that control measures such as precision fertilization, green fertilizer promotion, irrigation management optimization, and ecological restoration have achieved significant results in reducing soil and water pollution and increasing crop yields. By comparing model predictions with field data, the nitrogen and phosphorus content in the soil decreased significantly after treatment, the concentration of water pollutants decreased significantly, and crop yields generally increased, indicating that these measures can effectively alleviate agricultural non-point source pollution and improve agricultural productivity.

In terms of the implementation effects in different regions, the plains and coastal areas benefited the most, and the governance measures were able to achieve significant ecological and economic benefits in a relatively short period of time. However, the governance effects in mountainous and hilly areas were relatively weak, mainly due to factors such as complex terrain and soil erosion. Nevertheless, through the adoption of targeted ecological restoration and improvement measures, the governance effects in mountainous and hilly areas have still improved.

Although this study has achieved relatively ideal governance results, it still faces several challenges in the implementation process, especially in the promotion of green fertilizers, the popularization of precision fertilization technology, and the long-term effects of ecological restoration. The high cost of green fertilizers and farmers' usage habits are still the main obstacles to promotion. The popularity of precision fertilization technology in some remote areas is low, and strong technical support is needed during implementation.

Future research should focus more on the popularization and application of governance technologies, especially the optimization of governance in complex terrain areas such as mountainous and hilly areas. In addition, the government can promote the popularization of green fertilizers and advanced irrigation technologies through policy support and subsidies, and improve farmers' environmental awareness and technology acceptance. With the support of remote sensing, the Internet of Things, and big data technologies, agricultural non-point source pollution control is expected to be more widely used and more accurately implemented in the future, thereby promoting a win-win situation for agricultural production and environmental protection.

The pollution assessment model developed in this study exhibits strong adaptability and can be extended to other agricultural settings beyond the test region. With appropriate retraining, the model can be calibrated to suit various cropping systems such as rice paddies, tea plantations, or horticultural zones. The core structure allows integration of region-specific variables, such as crop types, soil features, and irrigation methods, making it a versatile tool for broader non-point source pollution management across diverse agroecological zones.

## Conflict of Interest

The authors declare no conflict of interest.

## References


1. HOU D., BOLA N., TSANG D., KIRKHAM M., O'CONNOR D. Sustainable soil use and management: An interdisciplinary and systematic approach. Science of the Total Environment. **729** (12), 138961, **2020**.
2. ZHANG S., ZHANG L., MENG Q., WANG C., MA J., LI H., MA K. Evaluating agricultural non-point source pollution with high-resolution remote sensing technology and SWAT model: A case study in Ningxia Yellow River Irrigation District, China. Ecological Indicators. **166** (12), 112578, **2024**.
3. POPESCU S., MANSOOR S., WANI O., KUMAR S., SHARMA V., SHARMA A., ARYA V., KIRKHAM M.B., HOU D., BOLAN N., CHUNG Y.S. Artificial intelligence and IoT driven technologies for environmental pollution monitoring and management. Frontiers in Environmental Science. **12**, 19, **2024**.
4. LEI K., LI Y., ZHANG Y., WANG S., YU E., LI F., XIAO F., SHI Z., XIA F. Machine learning combined with Geodetector quantifies the synergistic effect of environmental factors on soil heavy metal pollution. Environmental Science and Pollution Research. **30** (60), 126195, **2023**.

5. LI C., JIANG Z., LI W., YU T., WU X., HU Z., YANG Y., YANG Z., XU H., ZHANG W., ZHANG W., YE Z. Machine learning-based prediction of cadmium pollution in topsoil and identification of critical driving factors in a mining area. Environmental Geochemistry and Health. **46** (9), 19, **2024**.
6. XU Y., SU B., WANG H. Development of a runoff pollution empirical model and pollution machine learning models of the paddy field in the Taihu Lake Basin based on the paddy in situ observation method. Water. **14** (20), 21, **2022**.
7. LU Y., LIU L., QIN F., WANG J., LIU J., LI Y., WAN L. Total nitrogen and total phosphorus pollution reshaped the relationship between water supply and demand in the Huaihe River Watershed, China. Chinese Geographical Science. **33** (3), 512, **2023**.
8. XU Y., LI P., ZHANG M., XIAO L., WANG B., ZHANG X., WANG Y., SHI P. Quantifying seasonal variations in pollution sources with machine learning-enhanced positive matrix factorization. Ecological Indicators. **166**, 14, **2024**.
9. MCGILL T., FORD W. Extreme learning machine predicts high-frequency stream flow and nitrate-n concentrations in a karst agricultural watershed. Journal of the Asabe. **67** (2), 305, **2024**.
10. MALUSÁ E., TARTANUS M., DANELSKI W., MISZCZAK A., SZUSTAKOWSKA E., KICINSKA J., FURMANCZYK E. Monitoring of DDT in agricultural soils under organic farming in Poland and the risk of crop contamination. Environmental Management. **66** (5), 916, **2020**.
11. WU J., ZHAO F. Machine learning: An effective technical method for future use in assessing the effectiveness of phosphorus-dissolving microbial agroremediation. Frontiers in Bioengineering and Biotechnology. **11**, 5, **2023**.
12. DENG Y., YE X., DU X. Predictive modeling and analysis of key drivers of groundwater nitrate pollution based on machine learning. Journal of Hydrology. **624**, 12, **2023**.
13. JIA X., CAO Y., O'CONNOR D., ZHU J., TSANG D., ZOU B., HOU D. Mapping soil pollution by using drone image recognition and machine learning at an arsenic-contaminated agricultural field. Environmental Pollution, **270**, 10, **2021.**
14. ZHAO W., MA J., LIU Q., QU Y., DOU L., SHI H., SUN Y., CHEN H., TIAN Y., WU F. Accurate prediction of soil heavy metal pollution using an improved machine learning method: a case study in the Pearl River Delta, China. Environmental Science & Technology. 11, **2023**.
15. SONG Y., WU D., JU X., DORSCH P., WANG M., WANG R., SONG X., DENG L., WANG R., GAO Z., HAIDER H., HOU L., LIU M., YU Y. Nitrite stimulates HONO and NOx but not $N_2O$ emissions in Chinese agricultural soils during nitrification. Science of the Total Environment. **902**, 12, **2023**.
16. JIANG Z., YANG S., CHEN X., PANG Q., XU Y., QI S., YU W., DAI H. Controlled release urea improves rice production and reduces environmental pollution: a research based on meta-analysis and machine learning. Environmental Science and Pollution Research. **29** (3), 3587, **2022**.
17. PAN B., LEI J., PAN B., TIAN H., HUANG L. Dialogue between algorithms and soil: Machine learning unravels the mystery of phthalates pollution in soil. Journal of Hazardous Materials. **482**, 13, **2025**.
18. MA X., GUAN D., ZHANG C., YU T., LI C., WU Z., B. LI, GENG W., WU T., YANG Z. Improved mapping of heavy metals in agricultural soils using machine learning augmented with spatial regionalization indices. Journal of Hazardous Materials. **478**, 11, **2024**.
19. ZHANG Z., HUANG Y., HUANG J. A spatially explicit interpretable machine-learning method to track dissolved inorganic nitrogen pollution in a coastal watershed. Ecological Indicators. **158**, 13, **2024**.
20. LIU D., YAO Z., YANG X., XIONG C., NIE Q. Research progress and trend of agricultural non-point source pollution from non-irrigated farming based on bibliometrics. Water. **15** (8), 11, **2023**.
21. CHEN F., HU Y. Agricultural and rural ecological management system based on big data in complex system. Environmental Technology & Innovation. **22**, 10, **2021**.
22. HUO Z., TIAN J., WU Y., MA F. a soil environmental quality assessment model based on data fusion and its application in Hebei Province. Sustainability. **12** (17), 15, **2020**.
23. WANG X., YU D., MA L., LU X., SONG J., LEI M. Using big data searching and machine learning to predict human health risk probability from pesticide site soils in China. Journal of Environmental Management. **320**, 9, **2022**.
24. LI Y., SHA Z., TANG A., GOULDING K., LIU X. The application of machine learning to air pollution research: A bibliometric analysis. Ecotoxicology and Environmental Safety. **257**, 10, **2023.**