sustainability

MDPI

Article

Machine Learning—Driven Analysis of Agricultural Nonpoint Source Pollution Losses Under Variable Meteorological Conditions: Insights from 5 Year Site-Specific Tracking

Ran Jing [1,2,*], Yinghui Xie [1], Zheng Hu [1], Xingjian Yang [1], Xueming Lin [1], Wenbin Duan [3], Feifan Zeng [4], Tianyi Chen [5], Xin Wu [1], Xiaoming He [1] and Zhen Zhang [1,*]

[1] College of Natural Resources and Environment, South China Agricultural University, Guangzhou 510642, China; xieyinghui2024@163.com (Y.X.); z.hu@scau.edu.cn (Z.H.); xjyang@scau.edu.cn (X.Y.); xm.lin@scau.edu.cn (X.L.); 17872321738@163.com (X.W.); 20233138020@stu.scau.edu.cn (X.H.)
[2] Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, USA
[3] Anyang Huanshui Park Management Station, Anyang 455002, China; daniel_627@163.com
[4] Quliang Electronics Co., Ltd., Jinjiang 362200, China; feifan.zeng@qleltd.com
[5] Maoming Energy Conservation Center, Maoming 525000, China; lkycty@163.com
* Correspondence: rjingraymond88@gmail.com (R.J.); zhangzhen_23102@163.com (Z.Z.)

Abstract

Agricultural nonpoint source pollution is emerging as one of the increasingly serious environmental concerns all over the world. This study conducted field experiments in Zengcheng District, Guangzhou City, from 2019 to 2023 to explore the mechanisms by which different crop types, fertilization modes, and meteorological conditions affect the loss of nitrogen and phosphorus in agricultural nonpoint source pollution. In rice and corn, the CK and PK treatment groups showed significant fitting advantages, such as the R^2 of rice-CK reaching 0.309. MAE was 0.395, and the R^2 of corn-PK was as high as 0.415. For compound fertilization groups such as NPK and OF, the model fitting ability decreased, such as the R^2 of rice-NPK dropping to 0.193 and the R^2 of corn-OF being only 0.168. In addition, the overall performance of the model was limited in the modeling of total phosphorus. A relatively good fit was achieved in corn (such as NPK group R^2 = 0.272) and in vegetables and citrus. R^2 was mostly below 0.25. The results indicated that fertilization management, crop types, and meteorological conditions affected nitrogen and phosphorus losses in agricultural runoff. Cornfields under conventional nitrogen, phosphorus, and potassium fertilizer (NPK) and conventional nitrogen and potassium fertilizer treatment without phosphorus fertilizer (NK) treatments exhibited the highest nitrogen losses, while citrus fields showed elevated phosphorus concentrations under NPK and PK treatments. Organic fertilizer treatments led to moderate nutrient losses but greater variability. Organic fertilizer treatments resulted in moderate nutrient losses but showed greater interannual variability. Meteorological drivers differed among crop types. Nitrogen enrichment was mainly associated with high temperature and precipitation, whereas phosphorus loss was primarily triggered by short-term extreme weather events. Linear regression models performed well under simple fertilization scenarios but struggled with complex nutrient dynamics. Crop-specific traits such as flooding in rice fields, irrigation in corn, and canopy coverage in citrus significantly influenced nutrient migration. The findings of this study highlight that nutrient losses are jointly regulated by crop systems, fertilization practices, and meteorological variability, particularly under extreme weather conditions. These findings underscore the necessity of crop-specific and climate-adaptive nutrient management strategies to reduce agricultural nonpoint source pollution. By integrating long-term field observations with machine learning–based analysis, this study

Check for updates

Academic Editor: Yuxue Liu

Received: 24 November 2025
Revised: 26 December 2025
Accepted: 29 December 2025
Published: 7 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

provides scientific evidence to support sustainable fertilizer management, protection of water resources, and environmentally responsible agricultural development in subtropical regions. The proposed approaches contribute to sustainable land and water resource utilization and climate-resilient agricultural systems, aligning with the goals of sustainable development in rapidly urbanizing river basins.

Keywords: nonpoint source pollution; machine learning; meteorological variability; K-means clustering; multiple linear regression

1. Introduction

Nonpoint source (NPS) pollution of nitrogen and phosphorus from agricultural land is one of the pressing and persistent environmental challenges globally [1]. Unlike point-source pollution, it can be traced to a single discharge location. NPS pollution is usually diffuse and arises from multiple sources across landscapes, which makes it notoriously difficult to monitor and control [2]. This type of pollution could be one of the significant drivers of water quality degradation, resulting in eutrophication of freshwater and coastal ecosystems and biodiversity loss [3]. Nutrient losses from cropland caused by NPS continue to impact downstream water bodies in developed countries such as the United States, the Netherlands, and Switzerland, even though they usually have sophisticated agricultural systems and regulatory frameworks [4]. In developing countries, especially among the BRICS economies, i.e., Brazil, Russia, India, China, and South Africa, rapid agricultural expansion and growing fertilizer use intensify the threat of nutrient runoff [5,6].

In China, with ongoing agricultural modernization and increasing fertilizer input intensity, nitrogen and phosphorus NPS pollution has become an increasingly serious issue [7]. Studies have shown that over 55 million tons of nitrogen and phosphorus fertilizers, accounting for over 30% of global fertilizer use, were applied annually in Chinese farmland [8]. Over half of the fertilizer is not absorbed by crops and is lost to the environment through surface runoff [9]. This problem is particularly acute in southern regions, where high rainfall intensity, complex terrain, and widespread paddy fields exacerbate nutrient losses. These losses not only waste fertilizer resources but also have destructive and widespread impacts on ecological systems. As a result, nitrogen and phosphorus losses can pollute the environment through various pathways [10]. Surface runoff can wash residual nitrogen and phosphorus into rivers and lakes, triggering eutrophication and algal blooms [11]. In addition, long-term over-fertilization and nutrient runoff can reduce the biodiversity and disrupt nitrogen cycling, ultimately affecting the sustainability of agricultural production [12,13]. These impacts extend beyond agriculture, affecting aquatic ecosystems and public health. Therefore, monitoring and precise management of NPS pollution is a critical direction for agricultural environmental governance [14]. Among various monitoring methods, the integration of machine learning (ML) and traditional environmental monitoring technologies offers an efficient way to keep track of NPS pollution [15]. In recent years, ML methods based on environmental monitoring data have been widely used in studies on pollution monitoring, nutrient load estimation, and source tracking [16–18]. However, most existing studies focus on large watershed scales, with limited detailed investigations of nitrogen and phosphorus loss mechanisms at regional or field scales. Meanwhile, high-resolution data-driven models are needed to predict nutrient runoff under varying climate and land use conditions. In particular, related research in the Pearl River Delta is insufficient. The water quality of the Dongjiang River in this region is important as one of the three main tributaries of the Pearl River and a crucial drinking

water source for Guangdong Province and Hong Kong [19]. Agricultural activities such as rice and corn cultivation with high fertilizer inputs are intensive in Zengcheng District, Guangzhou, which is located within the Dongjiang watershed [20]. Therefore, this region serves as a representative area for studying agricultural NPS pollution. In addition, one critical factor that could potentially influence nutrient losses is weather variability, particularly precipitation patterns. Rainfall amount and intensity are major drivers of surface runoff, soil erosion, and the transport of dissolved nutrients [21]. For instance, light rainfall may not generate enough to carry fertilizer particles, while intense storms can cause flash floods that transport large volumes of nutrients off-site [22]. The antecedent soil moisture condition, timing of fertilizer application relative to rainfall events, and temperature fluctuations all interact to modulate nutrient loss [23]. Moreover, extreme weather events, such as typhoons and prolonged droughts followed by intense rain, could rapidly mobilize fertilizers into water systems. Despite these risks, few studies have investigated how short-term and long-term weather variations interact with farming practices to influence nitrogen and phosphorus runoff in real agricultural landscapes.

To better understand the spatial–temporal characteristics and influencing factors of nitrogen and phosphorus fertilizer loss in Zengcheng, four types of crops (rice, vegetables, corn, and citrus) are involved in this study to integrate field plot experiments, in situ measurements, and meteorological data to develop ML models for simulating and analyzing nutrient runoff. The objectives of this study include the following: (1) analyzing the concentration and total load of nitrogen and phosphorus losses under different crop types; (2) investigating the effects of weather variability, especially rainfall, on the concentrations of nitrogen and phosphorus fertilizer; and (3) exploring how different planting patterns affect nutrient runoff characteristics. This study aims to provide data support and technical solutions for regional-scale pollution prevention and control, contributing to the integration of smart agriculture and ecological protection. This study uses a data-driven approach to determine the field-scale nutrient responses to meteorological variability, although process-based models such as SWAT are widely used for watershed-scale simulations.

2. Methodology

2.1. Characteristics of Experimental Sites

This study was conducted in Ningxi Town, Zengcheng District, Guangzhou City, Guangdong Province, China (23°14′ N, 113°37′ E) from 2019 to 2023. It is located along the west side of the lower reaches of the Dongjiang River, as shown in Figure 1. The average annual temperature is 22.8 °C, and the average annual precipitation is between 1500 and 2200 mm. The precipitation is unevenly distributed throughout the year, and approximately 70% of the precipitation occurs in summer (June–September). The physical and chemical properties of surface soil (0–20 cm) at the experimental site were summarized in Table 1.

Table 1. Physical and chemical properties of surface soil (0–20 cm) at the experimental site.

Property	Value
Soil type	Acidic red soil
Texture	Sandy loam
Clay (%)	14.6
Silt (%)	15.9
Sand (%)	69.5
pH	5.83
Bulk density (g/cm^3)	1.26
Organic matter (g/kg)	15.49
Total nitrogen (g/kg)	1.2
Total phosphorus (g/kg)	0.6
Alkaline nitrogen (mg/kg)	60.97
Available phosphorus (mg/kg)	19.47
Total potassium (g/kg)	4.14

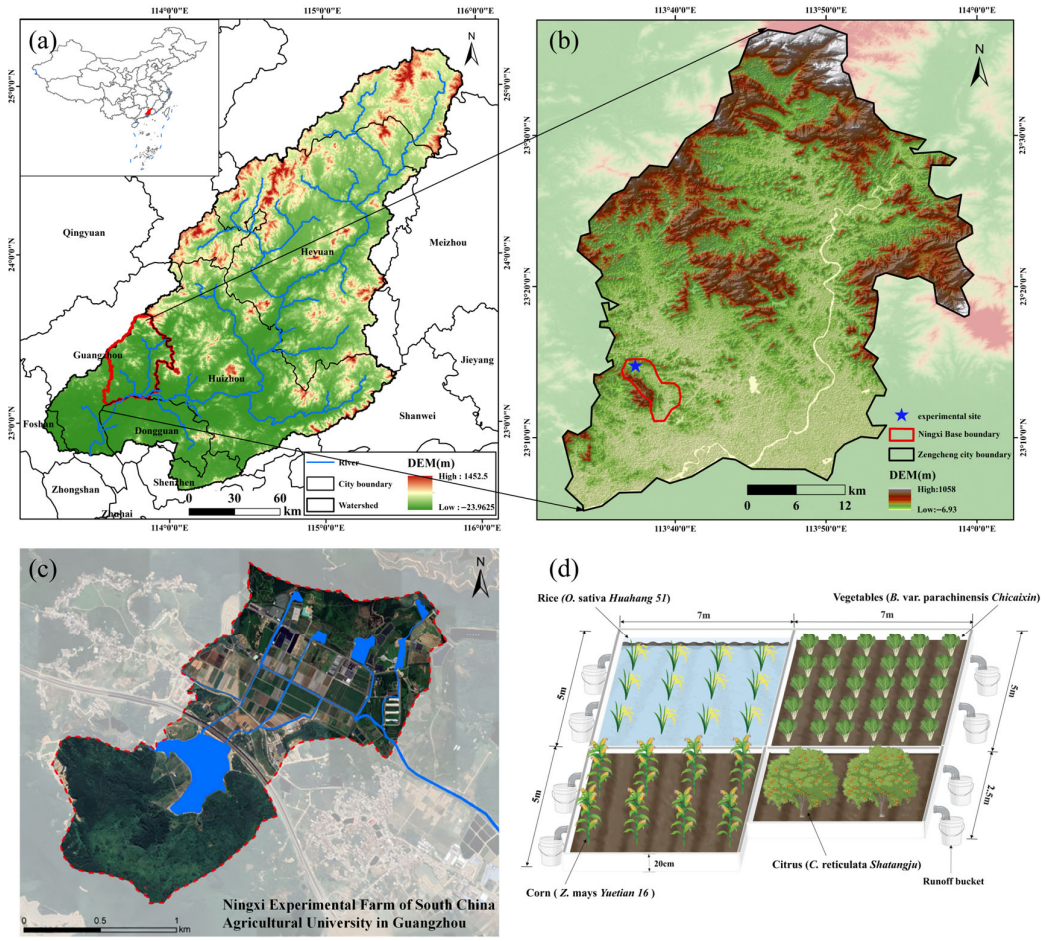


Figure 1. (**a**) Hydrological analysis and geographic information system map of the geographical location of the Dongjiang River Basin, (**b**) elevation information of Zengcheng district in Guangzhou and the geographical location of the experimental farm site, (**c**) aerial photographs of the Ningxi experimental farm site of South China Agricultural University in Guangzhou in this study, and (**d**) layout of the experimental plots for four crops in this study.

2.2. Field Experiment Design and Setup

The crops involved in this study are as follows: *Oryza sativa* Huahang 51 (rice), *Brassica rapa* var. *parachinensis* Late Choy Sum (vegetables), *Zea mays* Yuetian 16 (corn), and *Citrus reticulata* Shatangju (citrus) [18]. These crops are designated as rice, vegetables, corn, and citrus throughout this paper. From July 2019 to December 2023, a total of 9 seasons of rice, 9 seasons of corn, 9 seasons of vegetables, and 2 seasons of citrus were planted. Five different fertilizer management modes were set for each crop as fertilizer practice groups. They are as follows: conventional nitrogen, phosphorus, and potassium fertilizer treatment as the NPK group; conventional nitrogen and potassium fertilizer treatment without phosphorus fertilizer treatment as the NK group; conventional phosphorus and potassium fertilizer treatment without nitrogen fertilizer treatment as the PK group; no fertilizer treatment as the CK group; and organic fertilizer substitution treatment as the OF group. These fertilizer practice groups (i.e., NPK, NK, PK, CK, and OF) were performed with conventional tillage and conventional irrigation modes. The four crops selected in this study are not only representative but also account for a large proportion of the regional planting structure in South China. According to statistics, rice accounts for approximately 14.07%, corn 2.94%, vegetables 43.00%, and citrus 1.17%, with a combined 61.19% in total. This indicates that the crop types in this study encompass the majority of typical crops grown in South China and have strong regional representativeness. More importantly,

the experimental design integrates site-specific tracking, including mainstream agronomic practices such as fertilization, irrigation, and tillage. These implementations provide a solid foundation for a comprehensive assessment of the migration mechanism of nitrogen and phosphorus non-point source pollution.

Among these fertilizer treatment modes, the conventional nitrogen fertilizer application rate was 148 kg/hm^2, the phosphorus fertilizer application rate was 67 kg/hm^2, and the potassium fertilizer application rate was 114 kg/hm^2. Organic fertilizer substitution refers to the use of fermented chicken manure to replace 15% of the chemical fertilizer application under conventional fertilization conditions. The rice and corn tillage treatments are conventional tillage. Deep plowing is performed three times before planting, with each tillage reaching a depth of 20–30 cm. Each treatment was set up with 3 replicates, with a total of 60 treatment plots (15 each for rice, corn, vegetables, and citrus). The specifications of the rice, corn, and vegetable plots were 7 m long and 5 m wide. The total area is 35 m^2. The specifications of the citrus plots were 7 m long and 2.5 m wide with a 17.5 m^2 area. The experimental plot is separated by a 40 cm polyethylene (PE) board. The PE board is 20 cm higher than the field soil to prevent runoff infiltration. Two 460 L barrels are set next to each monitoring plot to collect runoff water. Moreover, a small agricultural meteorological station is installed to automatically record the rainwater. The field experiment design and setup in this study are shown in Figure 1d.

2.3. Phosphorus and Nitrogen Measurement of Water Runoff Samples

The total runoff volume for each community/sampling point is calculated by multiplying the average water level (measured in the runoff bucket) by the base area of the runoff bucket. The runoff water samples collected from the two runoff buckets in each community are mixed into one bottle of water sample in equal volumes. A sterile polyethylene stirring rod is used to fully mix the runoff water in the bucket to prevent sediment settling from affecting the sample results. After that, the runoff water samples were collected in a 100 mL polyethylene wide-mouth bottle. All runoff samples were collected overnight following each runoff event. In addition, standard QA/QC procedures were applied during sample collection and analysis, including the use of duplicate samples and blanks. After collecting the water samples, they are immediately transported to the laboratory for storage at 4 °C. The total nitrogen concentration was determined by alkaline potassium persulfate digestion and ultraviolet spectrophotometry (HJ 636-2012 [24]), and the absorbance was measured using an ultraviolet–visible spectrophotometer (UV-2800, Unico, Dayton, NJ, USA). Total phosphorus concentration was determined by continuous injection–ammonium molybdate spectrophotometry (HJ 671-2013 [25]). Samples were digested with potassium persulfate, followed by the addition of ascorbic acid and molybdate solution, and the total phosphorus concentration was determined by a continuous flow analyzer (Skalar, Breda, The Netherlands) after color development. The formula for calculating the amount of nitrogen and phosphorus runoff loss is given by the following equations.

$$F_n = \sum_{i=1}^{n} \left(C_i \times V_i \times 10^{-2} \right) / S$$

$$F_p = \sum_{i=1}^{n} \left(C_i \times V_i \times 10^{-2} \right) / S$$

where F_n/F_p is the amount of nitrogen and phosphorus runoff loss in each plot, kg/hm^2; C_i is the concentration of total nitrogen and total phosphorus measured in the i-th sampling, mg/L; V_i is the total amount of runoff in the runoff bucket of each plot measured in the i-th sampling, L; S is the area of each plot (the area of rice, corn, and vegetable plots is 35 m^2, and the area of citrus plots is 17.5 m^2); and n is the total number of samplings during the crop growth period.

2.4. Correlation Analysis of Meteorological Dataset

To quantify interrelationships among meteorological variables, a correlation analysis was performed on the standardized meteorological dataset (Supplementary Materials). The analysis was implemented in Python 3.10 using pandas for data manipulation and seaborn/matplotlib for visualization. The dataset was standardized (mean = 0, std = 1) prior to analysis. Non-numeric columns (Date and Average Temperature (F)) were excluded to retain only continuous variables. A Pearson correlation matrix was generated using pandas. DataFrame.corr(), capturing linear relationships between all variable pairs. The lower triangle displays correlation coefficients (annotated with a color gradient). The upper triangle uses circle markers, where circle size scales with absolute correlation strength and color indicates direction.

2.5. K-Means Cluster Analysis for Water Runoff Samples with Concentrations of Nitrogen and Phosphorus Fertilizers

This study uses unsupervised machine learning technology to analyze the distribution pattern of nitrogen and phosphorus concentrations in water bodies under different fertilization conditions (NPK, PK, NK, CK, and OF) and their association with meteorological factors. First, for the standardized nutrient concentration data (five features of NPK, PK, NK, CK, and OF), the runoff samples were divided into three characteristic clusters using the K-means clustering algorithm. After clustering, the high-dimensional data were reduced to two-dimensional space by principal component analysis (PCA), and a scatter plot with the first principal component (PCA1) and the second principal component (PCA2) as coordinates was generated. In order to explore the association between clustering results and environmental factors, the average values of eight meteorological characteristics, such as dew point temperature, visibility, average wind speed, maximum sustained wind speed, maximum gust, maximum temperature, minimum temperature, and precipitation in each cluster, were further calculated and visually compared through multi-cluster bar charts. Finally, one-way analysis of variance (ANOVA) was used to quantitatively test the statistically significant differences in each meteorological characteristic among clusters, and $p < 0.05$ was set as the significant threshold. The entire analysis process is implemented based on Python 3.10, and the key dependent libraries include scikit-learn (clustering and dimensionality reduction), SciPy (statistical test), and seaborn/matplotlib (visualization). The clustering goal is to cluster samples X = {x1, x2, . . ., xn} to divide into K clusters so that the within-cluster sum of squares is minimized, as shown in the following equation.

$$min_{C_1,\ldots,C_K} \sum_{k=1}^{K} \sum_{x_i \in C_k} | x_i - \mu_k |^2$$

where C_k is the k-th cluster. μ_k is the centroid of cluster C_k. $| x_i - \mu_k |^2$ is the Euclidean distance squared.

For each meteorological variable V_j, the mean value for different clusters is calculated using the following equation.

$$Vj^{(c)} = 1/n_c \sum i \in C_c V_j^{(i)}$$

where C_c is the sample index set of the c-th cluster. n_c: the number of samples in cluster C_c. $V_j^{(c)}$ is the average value of the c-th cluster on variable V_j. The K-Means cluster code used in this study is available at the following repository: https://github.com/jr198868/Machine_Learning_Nonpoint_Source_Pollution (accessed on 23 November 2025).

2.6. Multiple Linear Regression of Meteorological Conditions for Concentrations of Total Nitrogen and Phosphorus Predicting

This study uses a multivariate linear regression model to quantify the impact of meteorological conditions on water nutrient concentrations (NPK, PK, NK, CK, OF). First, eight meteorological characteristics, including dew point temperature, visibility, average wind speed, maximum sustained wind speed, maximum gust, maximum temperature, minimum temperature, and precipitation, were selected as independent variables (excluding "Average Temperature (F)" to avoid multicollinearity), and the total phosphorus concentration under five fertilization conditions was used as the dependent variable; after reading the CSV format data through the pandas library, sm.add_constant was used to add constant terms to the independent variable matrix, and the align function was used to ensure that the independent variable and the dependent variable samples were accurately matched. Subsequently, each nutrient variable was independently modeled: the data was randomly divided into a training set and a test set in a ratio of 8:2 (test_size = 0.2, random_state = 42), and the ordinary least squares (OLS) method of the statsmodels library was used to fit the regression model (sm.OLS(y_train, X_train).fit()), and evaluation indicators such as the coefficient of determination (R^2), mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) were calculated. The entire analysis process was implemented based on Python 3.10, and the core dependent libraries include statsmodels (v0.14.0), scikit-learn (v1.2.2), and seaborn (v0.12.2). Standard regression diagnostics, including multicollinearity, residual normality, and homoscedasticity, were implemented to ensure the suitability of the linear modeling. For each dependent variable Y_j, the multivariate linear regression model is defined as the following equation.

$$Y_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_p X_p + \varepsilon$$

where Y_j is the response variable corresponding to the j-th fertilization treatment (such as NPK, PK, NK, CK, and OF). X_1, X_2, …, X_p denotes the meteorological factors or other explanatory variables (such as temperature, wind speed, precipitation, etc.). β_0 is the intercept term. β_1, …, β_p is the regression coefficient, which is estimated by the least squares method. ε is the random error term. The multiple linear regression code used in this study is available at the following repository: https://github.com/jr198868/Machine_Learning_Nonpoint_Source_Pollution (accessed on 23 November 2025). The full dataset of multiple linear regression analysis for total nitrogen and phosphorus predicting is in the Supplementary Materials.

2.7. Model Performance Metrics

In this study, three performance indicators, i.e., mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE), are used to evaluate the practicality and effectiveness of the model [26]. MAE is one of the widely used evaluation indicators in machine learning and regression analysis. MSE is another core evaluation indicator, which strengthens the penalty for large errors by calculating the average of the squares of the errors between the predicted values and the actual values. This indicator is relatively insensitive to outliers and is of great value in scenarios where the average error needs to be minimized. MSE is another core evaluation indicator that strengthens the penalty for large errors by calculating the average of the squares of the errors between the predicted values and the actual values. Last but not least, RMSE, as the square root of MSE, can significantly improve the interpretability of the results because its dimension is consistent with that of the target variable. This indicator achieves an effective balance between highlighting

large errors and maintaining interpretability. The formulas for MAE, MSE, and RMSE were indicated in the following equations.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}\left|\mathrm{y}_{\mathrm{test}} - \mathrm{y}_{\mathrm{pred}}\right|$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}\left(\mathrm{y}_{\mathrm{test}} - \mathrm{y}_{\mathrm{pred}}\right)^2$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}\left(\mathrm{y}_{\mathrm{test}} - \mathrm{y}_{\mathrm{pred}}\right)^2}$$

where $\mathrm{y}_{\mathrm{test}}$ represents the actual observed value (actual value) of the *i*-th sample (the true measurement value) in the dataset. $\mathrm{y}_{\mathrm{pred}}$ represents the model-predicted value (predicted value) of the *i*-th sample, which is the estimated value generated by the multivariate linear regression model.

3. Result and Discussion

3.1. Total Concentration of Nitrogen and Phosphorus Through Surface Water Runoff

According to Figure 2, total nitrogen concentration data were recorded for four crop types under five fertilization treatments during the five-year runoff water sample monitoring. The results indicated that different fertilization treatments had a significant effect on total nitrogen concentration, and the performance was also different among different crops. The NPK group not only showed relatively high total nitrogen concentrations with average values of 4.86–8.37 mg/L in all crops but also showed the highest fluctuation. The peak values even reached more than 15 mg/L in 2020 in corn plots. The total nitrogen concentration of the NK group was also high (4.52–7.83 mg/L), similar to or slightly lower than that of the NPK group. For example, the five-year average concentration of the NK treatment was 7.83 mg/L in cornfields, indicating that nitrogen application itself was sufficient to significantly increase the runoff nitrogen level. The total nitrogen concentration of the NK group also fluctuated greatly, exceeding the third fluctuating fertilizer group (OF) and tending to an intermediate level in other years. It should be noted that a stable nitrogen concentration can still be detected in the PK group and the CK group. The total nitrogen concentration of the PK group (no nitrogen fertilizer) and the CK group (no fertilizer) was relatively low and fluctuated little. The concentration ranged from 1.5 to 3.5 mg/L in most years. Especially in the rice, vegetable, and cornfields, the concentration difference between the two treatment groups was not significant, which suggested that the background soil nitrogen was the dominant source. The potential releasing source of nitrogen could be from decomposition and release of soil organic matter, such as ammonium nitrogen and nitrate nitrogen. It could also result from previous fertilizer residues and the slow release of nitrogen stored in the soil.

Regarding the total phosphorus (Figure 3), the NPK and PK groups generally had the highest total phosphorus concentration among all crops, especially in citrus plots. The average values of total phosphorus concentration were 0.64 mg/L and 0.66 mg/L for the NPK and PK groups, respectively. It is almost two times the total phosphorus concentration in vegetables and cornfields. The concentration of the NK group (no phosphorus fertilizer) was significantly lower than that of the NPK and PK groups. There was still an occasional increase in citrus and vegetable plots, with the annual average value mostly falling in the range of 0.1–0.3 mg/L, indicating that considerable phosphorus concentrations can still be detected under non-phosphorus fertilization conditions. The OF group had a wide overall fluctuation range, with concentrations approaching or even exceeding NPK levels in some years, which was particularly evident in vegetable and cornfields. The NK and CK groups showed low concentrations and relatively stable fluctuations, with annual average

concentrations mostly maintained at 0.062–0.184 mg/L, especially in rice and cornfields with the smallest fluctuations. The full dataset of the total concentration of nitrogen and phosphorus is in the Supplementary Materials.

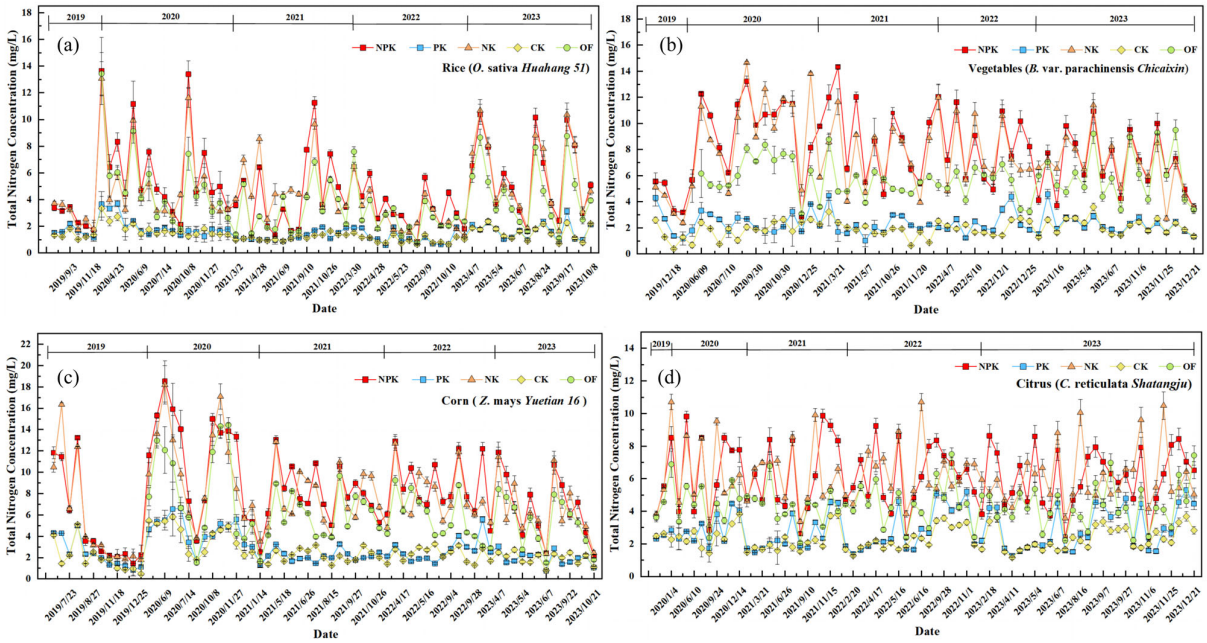


Figure 2. Dynamic changes in total nitrogen concentration in runoff water of four crops (**a**) rice, (**b**) vegetables, (**c**) corn, and (**d**) citrus under different water and fertilizer management modes from 2019 to 2023.

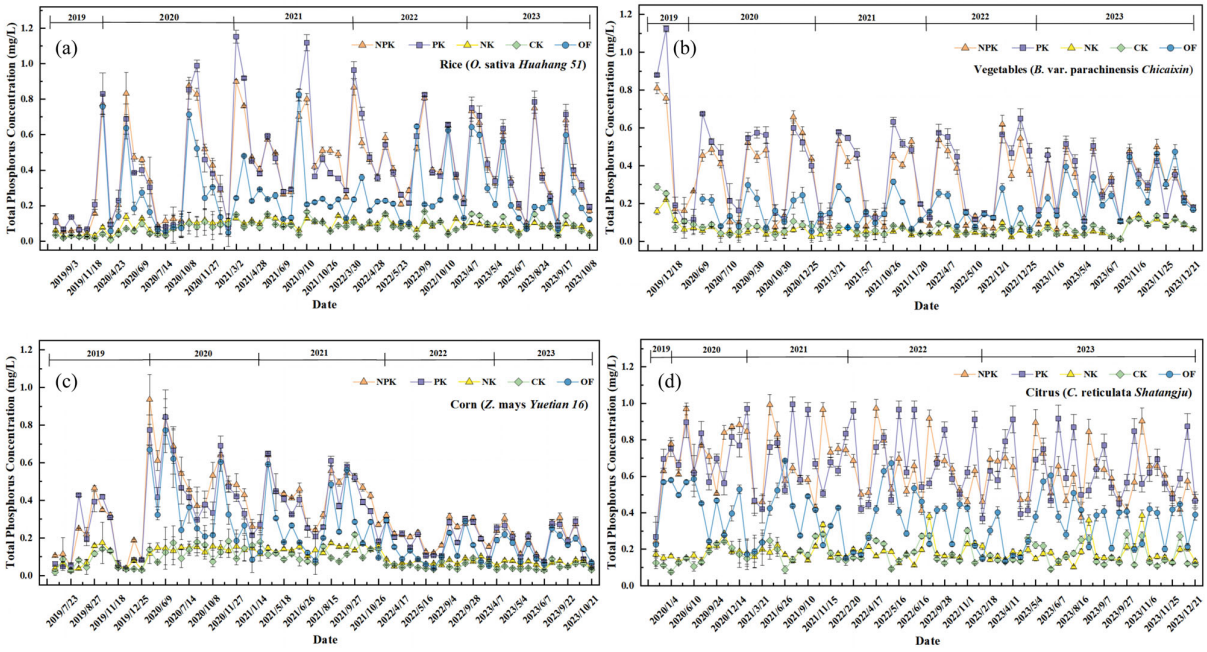


Figure 3. Dynamic changes in total phosphorus concentration in runoff water of (**a**) rice, (**b**) vegetables, (**c**) corn, and (**d**) citrus under different water and fertilizer management modes from 2019 to 2023.

3.2. Total Losses of Nitrogen and Phosphorus Through Surface Water Runoff

According to Figure 4 and Table 2, observations for five consecutive years showed that nitrogen and phosphorus losses in surface runoff showed significant crop differences and fertilization responses. In general, the loss of nitrogen is higher than that of phosphorus, especially in the corn planting system. Among all crops, the total nitrogen loss of corn

plots reached the highest value when NPK and NK treatments were applied (18.75 and 17.68 kg/hm^2, respectively), which may be related to the large nitrogen requirement of corn, loose soil, and rainfall runoff intensity. In contrast, the total nitrogen loss of vegetables and citrus under CK treatments was relatively low (2.10 and 2.42 kg/hm^2, respectively), indicating that under nitrogen deficiency or no fertilization conditions, the migration of nitrogen to surface water bodies is limited. The organic fertilizer (OF) treatment showed moderate levels of total nitrogen and phosphorus loss in most crop types, such as the average total nitrogen loss of rice was 8.22 kg/hm^2, and total phosphorus was 0.62 kg/hm^2, indicating that organic fertilizers may improve soil fertility but also pose certain risks to water bodies. In terms of total phosphorus loss, the performance of different crops was more dispersed. For example, the average total phosphorus loss of vegetable CK and NK treatments was the lowest (both below 0.13 mg/hm^2), while the PK and NPK treatments of rice were significantly higher, reaching 0.80 and 0.75 kg/hm^2, respectively, indicating that the phosphorus loss in the rice system is more serious under the conditions of sufficient phosphorus supply or nitrogen application. In addition, the total phosphorus loss of vegetables was the lowest when NK treatment was applied, indicating that phosphorus migration was significantly inhibited under the condition of no phosphorus application. In summary, different crop types and fertilization combinations will significantly affect the migration behavior of nitrogen and phosphorus in surface runoff. From the perspective of ecological security, in high rainfall areas or farmlands with large terrain slopes, the NK or CK mode should be preferred to reduce total phosphorus loss, and total nitrogen loss should be regulated by combining an organic fertilizer strategy to achieve a balance between yield and environmental risks. Based on this, it is recommended to implement differentiated crop regulation, in which the corn system gives priority to the deep application of phosphorus fertilizers. For these four crops, it is a priority to apply nitrogen fertilizers deeply under PK and CK treatments regarding minimum total nitrogen loss. It is a priority to apply phosphorus fertilizers deeply under NK and CK treatments regarding minimum total phosphorus loss.

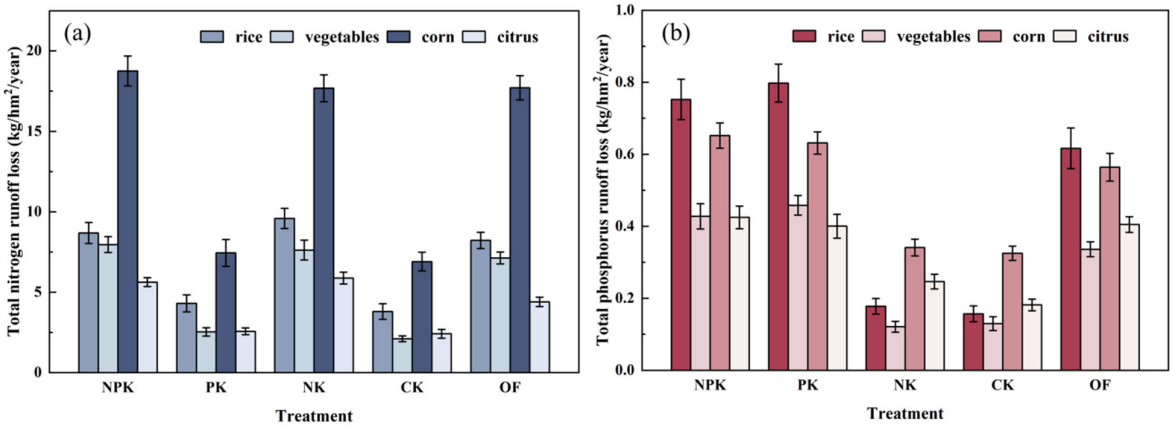


Figure 4. Average annual runoff losses of (**a**) total nitrogen and (**b**) total phosphorus from 2019 to 2023 for four crops under different water and fertilizer management modes.

Table 2. Total nitrogen and phosphorus runoff losses of four crops under different water–fertilizer management patterns.

Crop Type	Treatment	Total Nitrogen Runoff Loss (kg/hm^2)					Total Phosphorus Runoff Loss (kg/hm^2)				
		2019	2020	2021	2022	2023	2019	2020	2021	2022	2023
Rice	NPK [a]	4.22 ± 0.29	10.01 ± 0.95	8.21 ± 0.95	9.66 ± 0.48	11.29 ± 0.56	0.13 ± 0.01	0.89 ± 0.09	0.87 ± 0.09	0.96 ± 0.05	0.90 ± 0.05
	PK [b]	2.48 ± 0.19	5.70 ± 1.03	5.41 ± 1.03	4.10 ± 0.21	3.85 ± 0.19	0.17 ± 0.03	1.06 ± 0.07	0.88 ± 0.07	0.94 ± 0.05	0.93 ± 0.05
	NK [c]	4.38 ± 0.46	11.54 ± 0.77	10.04 ± 0.77	10.18 ± 0.51	11.81 ± 0.59	0.06 ± 0.01	0.22 ± 0.04	0.22 ± 0.04	0.23 ± 0.01	0.16 ± 0.01
	CK [d]	2.10 ± 0.12	4.56 ± 0.96	4.66 ± 0.96	4.08 ± 0.20	3.60 ± 0.18	0.04 ± 0.01	0.18 ± 0.04	0.17 ± 0.04	0.19 ± 0.01	0.20 ± 0.01
	OF [e]	6.19 ± 0.00	8.75 ± 0.56	8.56 ± 0.57	8.35 ± 0.42	9.23 ± 0.46	0.43 ± 0.00	0.71 ± 0.08	0.57 ± 0.08	0.73 ± 0.04	0.64 ± 0.03
Vegetables	NPK [a]	6.08 ± 0.33	8.73 ± 0.64	8.12 ± 0.64	9.03 ± 0.45	7.84 ± 0.39	0.58 ± 0.04	0.40 ± 0.05	0.37 ± 0.05	0.47 ± 0.02	0.31 ± 0.02
	PK [b]	2.99 ± 0.49	2.54 ± 0.29	2.26 ± 0.29	2.59 ± 0.13	2.31 ± 0.12	0.64 ± 0.06	0.50 ± 0.02	0.45 ± 0.02	0.39 ± 0.02	0.31 ± 0.02
	NK [c]	5.35 ± 0.73	8.99 ± 0.79	8.09 ± 0.79	7.37 ± 0.37	8.28 ± 0.41	0.17 ± 0.03	0.14 ± 0.02	0.13 ± 0.02	0.13 ± 0.01	0.03 ± 0.00
	CK [d]	1.64 ± 0.18	2.44 ± 0.09	2.17 ± 0.40	1.98 ± 0.10	2.28 ± 0.11	0.21 ± 0.03	0.14 ± 0.03	0.12 ± 0.03	0.13 ± 0.01	0.05 ± 0.01
	OF [e]	7.49 ± 0.00	8.63 ± 0.79	7.42 ± 0.10	6.01 ± 0.30	6.10 ± 0.31	0.25 ± 0.00	0.44 ± 0.03	0.38 ± 0.03	0.37 ± 0.02	0.24 ± 0.01
Corn	NPK [a]	23.19 ± 1.44	22.26 ± 0.84	17.81 ± 0.84	16.83 ± 0.84	13.67 ± 0.68	0.82 ± 0.05	0.69 ± 0.03	0.64 ± 0.03	0.69 ± 0.03	0.41 ± 0.02
	PK [b]	10.52 ± 0.66	8.73 ± 1.47	7.07 ± 1.47	7.21 ± 0.36	3.67 ± 0.18	0.80 ± 0.05	0.73 ± 0.03	0.59 ± 0.03	0.66 ± 0.03	0.37 ± 0.02
	NK [c]	24.22 ± 0.83	19.77 ± 0.96	15.81 ± 0.96	16.45 ± 0.82	12.14 ± 0.61	0.34 ± 0.05	0.43 ± 0.02	0.41 ± 0.02	0.39 ± 0.02	0.14 ± 0.01
	CK [d]	8.77 ± 0.48	7.82 ± 0.93	6.72 ± 0.93	7.34 ± 0.37	3.87 ± 0.19	0.29 ± 0.03	0.44 ± 0.02	0.38 ± 0.02	0.41 ± 0.02	0.10 ± 0.00
	OF [e]	27.72 ± 0.00	19.88 ± 0.90	16.70 ± 0.70	13.42 ± 0.67	10.81 ± 0.54	0.73 ± 0.00	0.70 ± 0.06	0.59 ± 0.06	0.47 ± 0.02	0.33 ± 0.02
Citrus	NPK [a]	4.84 ± 0.21	5.72 ± 0.27	5.32 ± 0.27	6.36 ± 0.32	5.89 ± 0.29	0.46 ± 0.07	0.34 ± 0.02	0.39 ± 0.02	0.43 ± 0.02	0.50 ± 0.03
	PK [b]	2.55 ± 0.04	2.71 ± 0.37	2.42 ± 0.37	2.55 ± 0.13	2.61 ± 0.13	0.51 ± 0.07	0.40 ± 0.03	0.36 ± 0.03	0.18 ± 0.01	0.55 ± 0.03
	NK [c]	4.8 ± 0.10	6.27 ± 0.55	5.64 ± 0.55	6.13 ± 0.31	6.58 ± 0.33	0.17 ± 0.02	0.19 ± 0.02	0.17 ± 0.02	0.42 ± 0.02	0.29 ± 0.01
	CK [d]	2.68 ± 0.09	2.15 ± 0.51	2.21 ± 0.51	2.48 ± 0.12	2.58 ± 0.13	0.13 ± 0.02	0.18 ± 0.02	0.16 ± 0.02	0.17 ± 0.01	0.27 ± 0.01
	OF [e]	4.31 ± 0.11	4.52 ± 0.44	3.89 ± 0.44	5.09 ± 0.25	4.21 ± 0.21	0.42 ± 0.02	0.41 ± 0.02	0.35 ± 0.02	0.39 ± 0.02	0.46 ± 0.02

[a] NPK refers to the conventional application of nitrogen, phosphorus, and potassium fertilizers. [b] PK refers to the conventional application of phosphorus and potassium fertilizers. [c] NK refers to the conventional application of nitrogen and potassium fertilizers. [d] CK refers to the no-fertilizer treatment. [e] OF refers to the treatment with organic fertilizers replacing conventional fertilizers.

3.3. Correlation Analysis of Meteorological Condition Dataset

Figure 5 shows a heat map of correlations among eight meteorological variables, where the correlation coefficients are double-coded by the size and color of the circular markers. The red color represents positive correlations, and the blue represents negative correlations. The color depth was proportional to the strength of the correlation. Overall, there was a strong linear relationship between air temperature, wind speed, and dew point temperature, while visibility and precipitation are relatively independent. Among these variables, the strongest positive correlation was between dew point temperature and minimum temperature (r = 0.93). In addition, the correlation between maximum temperature and minimum temperature was also high (r = 0.87), indicating that the temperature difference between day and night was relatively stable. In contrast, the correlation of other pair variables was generally weak, which may be affected by complex terrain or particle distribution. There was a strong linear correlation between some variables, such as the minimum temperature and dew point temperature, suggesting that the model could face the challenge of multicollinearity. The potential redundancy between variables can be preliminarily identified by observing the correlation coefficient.

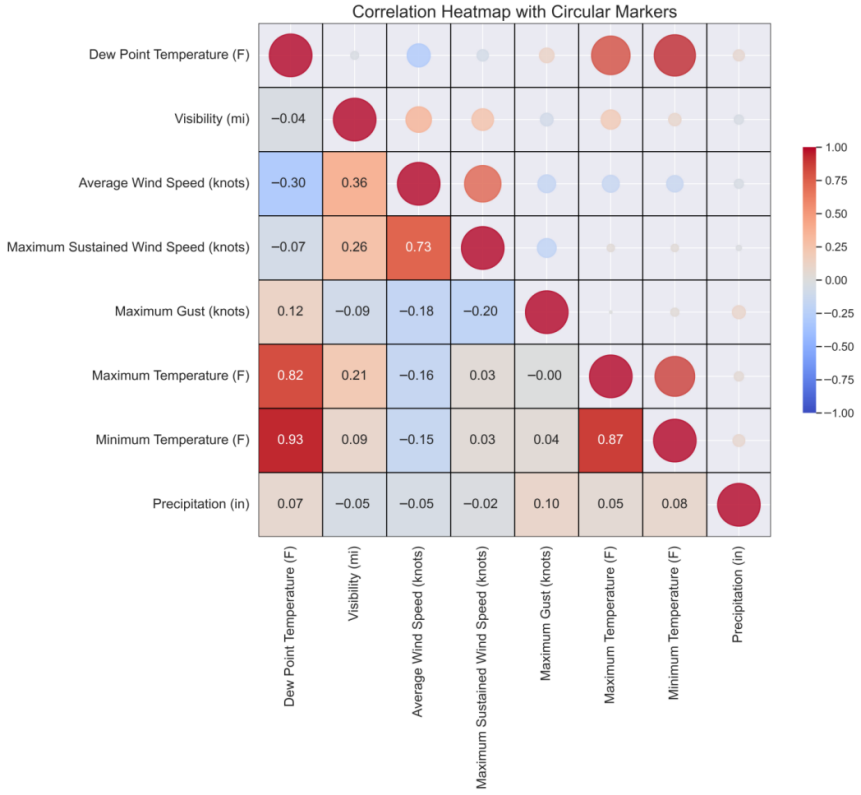


Figure 5. Heat map of meteorological variables correlation, which shows the correlation coefficients between dew point temperature (°F), visibility (mi), average wind speed (kn), maximum sustained wind speed (kn), maximum gust (kn), maximum temperature (°F), minimum temperature (°F), and precipitation (in). The correlation is presented by the color depth and the size of the circular mark: red indicates positive correlation, blue indicates negative correlation, and the larger the circle, the stronger the correlation.

3.4. Cluster Analysis for Water Runoff Samples with Concentrations of Nitrogen and Phosphorus Fertilizers

Figure 6 shows that the citrus field had a relatively clear three-cluster distribution pattern along the first and second principal components in the PCA scatter plot of the total nitrogen data. For rice and cornfields, most of the data points were located in the

low-value range of PC1, representing samples with low overall nitrogen concentration. For the vegetable field, the blue and green clusters were not completely separable in the first two principal component spaces, which indicates that the projection values of the green and blue samples have a high degree of similarity. These two clusters were not completely different in all nutrient indicators. As a result, they could be close in some dimensions (such as PK/OF ratio or CK content), resulting in mixing in low-dimensional visualization. This three-sample group divided by concentration gradient suggested that total nitrogen may be dominated by different basin processes (such as land use or seasonal degradation) and lays the foundation for subsequent comparative analysis of meteorological conditions of different clusters.

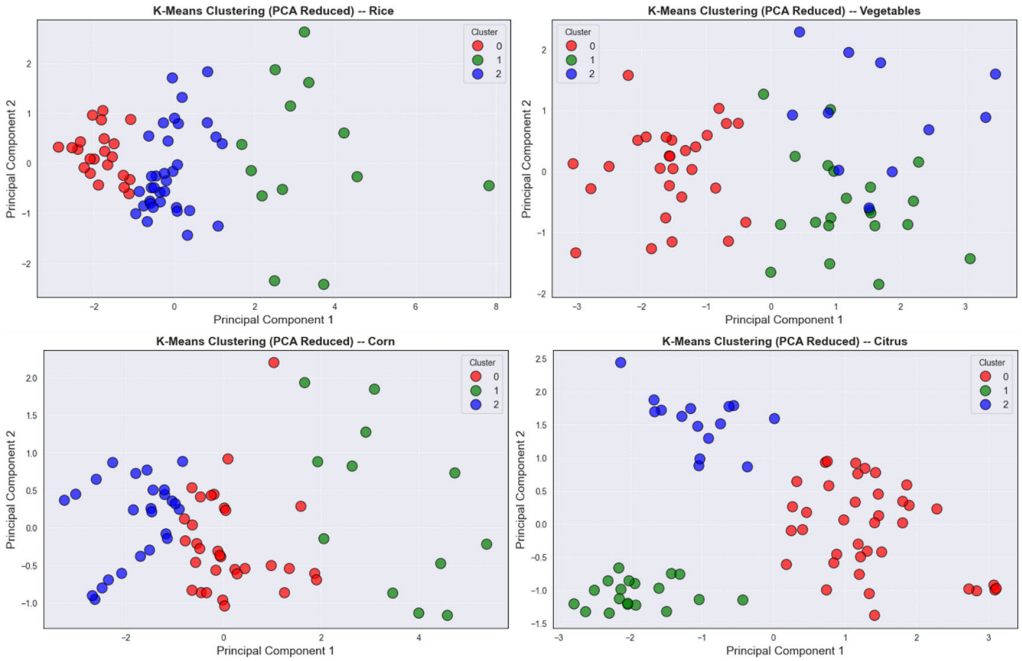


Figure 6. K-Means clustering results (total nitrogen) using principal component analysis (PCA) for rice, corn, vegetables, and citrus crops.

Regarding the total phosphorus data, the corn and citrus fields had a clear three-cluster distribution pattern along the first and second principal components in the PCA scatter plot (Figure 7). Most of the data points were located in the low-value range of principal components representing samples with low overall nitrogen concentration. The green cluster (Cluster 1) of rice fields was distributed in the high PC1 and PC2 regions corresponding to the relatively high nitrogen load, while the blue cluster (Cluster 2) was concentrated in the low-value area, indicating a low nitrogen content. Regarding the PCA scatter plot of the vegetable field, there was an overlap between the red cluster (Cluster 0) and the green cluster (Cluster 1), indicating that the changes in phosphorus components (such as OF and CK) between samples are more continuous and difficult to separate. The fuzziness of cluster boundaries reflects the heterogeneity of phosphorus migration and loss mechanisms, which could be affected by incidental events such as rainfall runoff and particle binding.

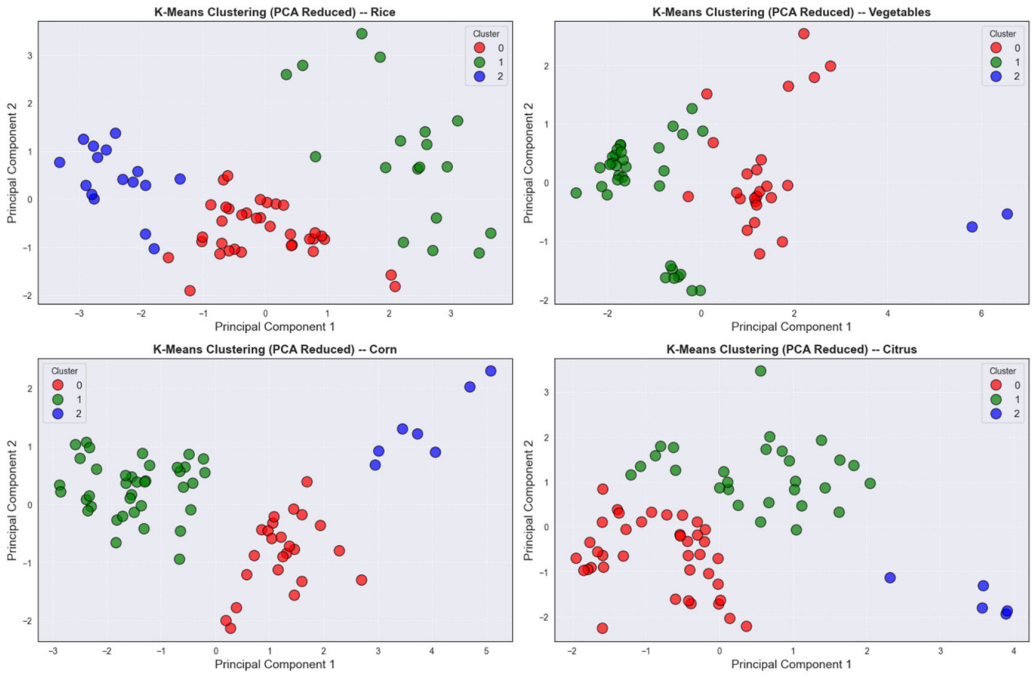


Figure 7. K-Means clustering results (total phosphorus) using principal component analysis (PCA) for rice, corn, vegetables, and citrus crops.

According to Table 3, the corresponding average value of the meteorological feature showed relatively consistent seasonally driven characteristics. Cluster 0 usually corresponded to the summer high-load meteorological feature with high temperature, high humidity, and low precipitation. Taking the rice field as an example, the dew point temperature (0.492) and the maximum temperature (0.510) were the highest among the three clusters, while the precipitation (−0.106) was the lowest. Cluster 1 was characterized by high visibility (rice: 0.466), low wind speed (rice average wind speed −0.091), and extremely low precipitation (rice: −0.137), which is the background level sample of the clear and dry period. Cluster 2 generally had strong precipitation and moderate temperature and humidity (rice precipitation 0.367, dew point temperature 0.137), corresponding to the "rush peak" during heavy rain or monsoon. For vegetables and corn, Cluster 2 had relatively prominent wind speed and precipitation, Cluster 0 had higher temperature and humidity, and Cluster 1 had the mildest conditions. Citrus had the highest wind speed (0.614) and higher precipitation (0.331) in Cluster 1. It is suggested that local storm processes also have a significant impact on total nitrogen. In general, high-concentration clusters of total nitrogen were often closely related to high temperature and humidity (summer growing season) or heavy precipitation events (runoff water peak flushing).

Table 3. Mean values of weather variables for each cluster identified via K-Means, along with statistical comparison (ANOVA *p*-values).

Cluster-Based Summary of Meteorological Conditions for Total Nitrogen									
Crops	**Cluster**	**Dew Point Temperature (°F)**	**Visibility (mi)**	**Average Wind Speed (Knots)**	**Maximum Sustained Wind Speed (Knots)**	**Maximum Gust (Knots)**	**Maximum Temperature (°F)**	**Minimum Temperature (°F)**	**Precipitation (in)**
Rice	0	0.492	0.189	0.483	0.326	0.006	0.510	0.551	−0.106
	1	0.112	0.466	−0.091	−0.207	0.325	0.026	0.065	−0.137
	2	0.137	0.243	0.334	0.400	−0.259	0.292	0.239	0.367
Vegetables	0	−0.269	−0.105	0.094	−0.153	0.092	−0.263	−0.334	0.167
	1	−0.211	0.121	0.382	0.261	−0.335	−0.199	−0.103	−0.130
	2	−0.756	−0.294	0.320	0.425	0.553	−0.653	−0.797	−0.136
Corn	0	0.472	0.114	0.094	0.049	0.096	0.456	0.457	0.407
	1	0.246	0.294	0.249	0.542	−0.737	0.436	0.484	−0.132
	2	0.215	−0.133	0.000	0.072	0.294	0.220	0.161	−0.121
Citrus	0	0.044	0.207	0.286	0.288	0.094	−0.057	−0.049	0.097
	1	−0.284	0.228	0.614	0.175	0.361	−0.102	−0.261	0.331
	2	0.320	−0.028	0.193	−0.004	0.298	0.176	0.265	−0.126
Cluster-Based Summary of Meteorological Conditions for Total Phosphorus									
Crops	**Cluster**	**Dew Point Temperature (°F)**	**Visibility (mi)**	**Average Wind Speed (Knots)**	**Maximum Sustained Wind Speed (Knots)**	**Maximum Gust (Knots)**	**Maximum Temperature (°F)**	**Minimum Temperature (°F)**	**Precipitation (in)**
Rice	0	0.291	0.295	0.322	0.294	0.009	0.282	0.295	0.110
	1	0.014	0.488	0.307	0.129	−0.017	0.292	0.141	0.323
	2	0.405	−0.039	0.253	0.341	−0.279	0.397	0.513	−0.081
Vegetables	0	−0.390	0.103	0.237	0.132	0.177	−0.332	−0.434	0.209
	1	−0.238	−0.181	0.194	0.059	−0.083	−0.275	−0.213	−0.127
	2	−1.435	0.018	0.934	0.354	0.101	−0.718	−1.304	−0.140
Corn	0	0.219	0.256	0.106	0.038	−0.576	0.456	0.351	−0.134
	1	0.386	−0.057	0.141	0.236	0.520	0.315	0.312	0.327
	2	0.500	−0.015	−0.272	0.002	−0.624	0.345	0.598	−0.112
Citrus	0	0.013	−0.012	0.526	0.234	0.070	−0.186	−0.062	−0.110
	1	−0.009	0.324	0.106	0.142	0.383	0.161	−0.062	0.466
	2	0.159	0.635	0.387	0.216	0.293	0.259	0.261	−0.131

In contrast, the differences in meteorological characteristics among the three clusters for total phosphorus were more subtle. In the rice field, Cluster 2 still had a high dew point (0.405) and moderate precipitation (−0.080), but Cluster 1 had the second-highest precipitation (0.323) and visibility (0.488). It indicated that the meteorological feature of the rice field in this study is usually observed after a short period of heavy rainfall. The dew point (−1.435) and minimum temperature (−1.304) of Cluster 2 of vegetables are relatively low, and the wind speed (0.934) was the highest as compared with that of the other two clusters, suggesting that this cluster represents a dry, cold, and high-wind weather event. Clusters 1 and 2 of corn and citrus also had peaks in dew point or wind speed, respectively. However, the overall precipitation swing was smaller than that of the total nitrogen data. These facts indicated that although the high load of total phosphorus was also driven by meteorological conditions (such as heavy rain), it is dominantly affected by extreme short-term events (dry cold storms) rather than seasonal high humidity or background drought phases.

3.5. Results of Multiple Linear Regression of Meteorological Conditions for Concentrations of Total Nitrogen and Phosphorus Predicting

According to Table 4 and Figure 8, the multiple linear regression model for total nitrogen prediction showed a relatively good linear fitting ability in the rice and corn fields, with high determination coefficients (R^2) for the PK (0.274, 0.415) and CK (0.309, 0.337) groups (Table 4). The MAE of PK and CK groups for rice were 0.493 and 0.395, respectively. The RMSE was maintained in the range of 0.446–0.547 (Table 4). For the cornfield, the MAE of the PK and CK groups reached 1.009 and 1.052, and the RMSE was controlled below 1.3. This indicated that in the context of rice and corn planting, the linear regression model established with meteorological factors as independent variables can capture the trend of total nitrogen concentration changes. Combined with the clustering results in Table 3, samples of the summer high-load meteorological feature were generally distributed in Cluster 0. Taking rice as an example, the dew point (0.492) and maximum temperature (0.510) were the highest among all clusters, while its precipitation was the lowest (−0.106), thereby forming a relatively high-temperature and dry environment. This environment is conducive to the accumulation and redispersion of nitrogen in the soil surface of the experimental site, which is the main driving factor for the outstanding model fitting effect. In addition, there were some interference factors such as high precipitation (rice: 0.367 in Cluster 2; corn: 0.407 in Cluster 0) and speed disturbance (rice: average wind speed = 0.483 in cluster 0; corn: maximum sustained wind speed = 0.542 in Cluster 1) in samples of rice and corn. These interference factors increased the fitting error of some dependent variables in the model, especially for the conventional fertilizer treatment group (NPK) and organic fertilizer group (OF), such as the R^2 of rice-NPK being 0.193 and the R^2 of rice-OF being 0.215 (Table 4). Similar cases in corn samples were also observed. The R^2 of the NPK and OF groups were only 0.189 and 0.168, respectively. The R^2 of other treatment groups was as high as 0.415 for PK and 0.337 for CK. This suggests that under heavy rainfall or extreme weather events, the migration behavior of nitrogen may go beyond the linear mechanism. It involves complex processes such as leakage, runoff transport, and microbial activity, thereby reducing the accuracy of the model.

Table 4. Performance metrics of multiple linear regression models for predicting total nitrogen and total phosphorus concentrations in citrus, corn, rice, and vegetable crops under different fertilizer treatments.

	Citrus				**Corn**				**Rice**				**Vegetables**			
Total nitrogen	R^2	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE
NPK	0.179	1.902	5.126	2.264	0.189	3.361	16.304	4.038	0.193	2.436	9.777	3.127	0.121	2.305	7.473	2.734
PK	0.154	0.982	1.233	1.11	0.415	1.009	1.553	**1.246**	0.274	0.493	0.299	0.547	0.16	0.631	0.73	0.854
NK	0.151	1.472	3.393	1.842	0.202	2.884	12.651	3.557	0.151	1.999	7.034	2.652	0.104	2.067	6.367	2.523
CK	**0.203**	0.632	0.577	**0.76**	0.337	1.052	1.6	1.265	**0.309**	0.395	0.199	**0.446**	0.132	0.544	0.531	**0.729**
OF	0.125	0.783	0.901	0.949	0.168	2.244	10.617	3.258	0.215	1.945	5.497	2.345	**0.137**	1.409	3.65	1.91
Total phosphorus	R^2	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2	MAE	MSE	RMSE
NPK	0.145	0.181	0.046	0.214	0.272	0.147	0.037	0.192	0.093	0.276	0.101	0.318	0.183	0.143	0.029	0.17
PK	0.03	0.024	0.001	**0.032**	0.243	0.141	0.029	0.17	**0.154**	0.238	0.085	0.292	0.23	0.138	0.027	0.164
NK	**0.255**	0.206	0.064	0.253	**0.328**	0.029	0.002	**0.045**	0.145	0.03	0.001	**0.032**	**0.271**	0.019	0.001	**0.032**
CK	0.216	0.033	0.002	0.045	0.301	0.032	0.002	0.045	0.119	0.034	0.002	0.045	0.199	0.064	0.031	0.176
OF	0.134	0.115	0.019	0.138	0.141	0.072	0.008	0.089	0.048	0.175	0.052	0.228	0.248	0.067	0.009	0.095

The highest R^2 values and the lowest RMSE values are highlighted in bold to emphasize the best-performing models.

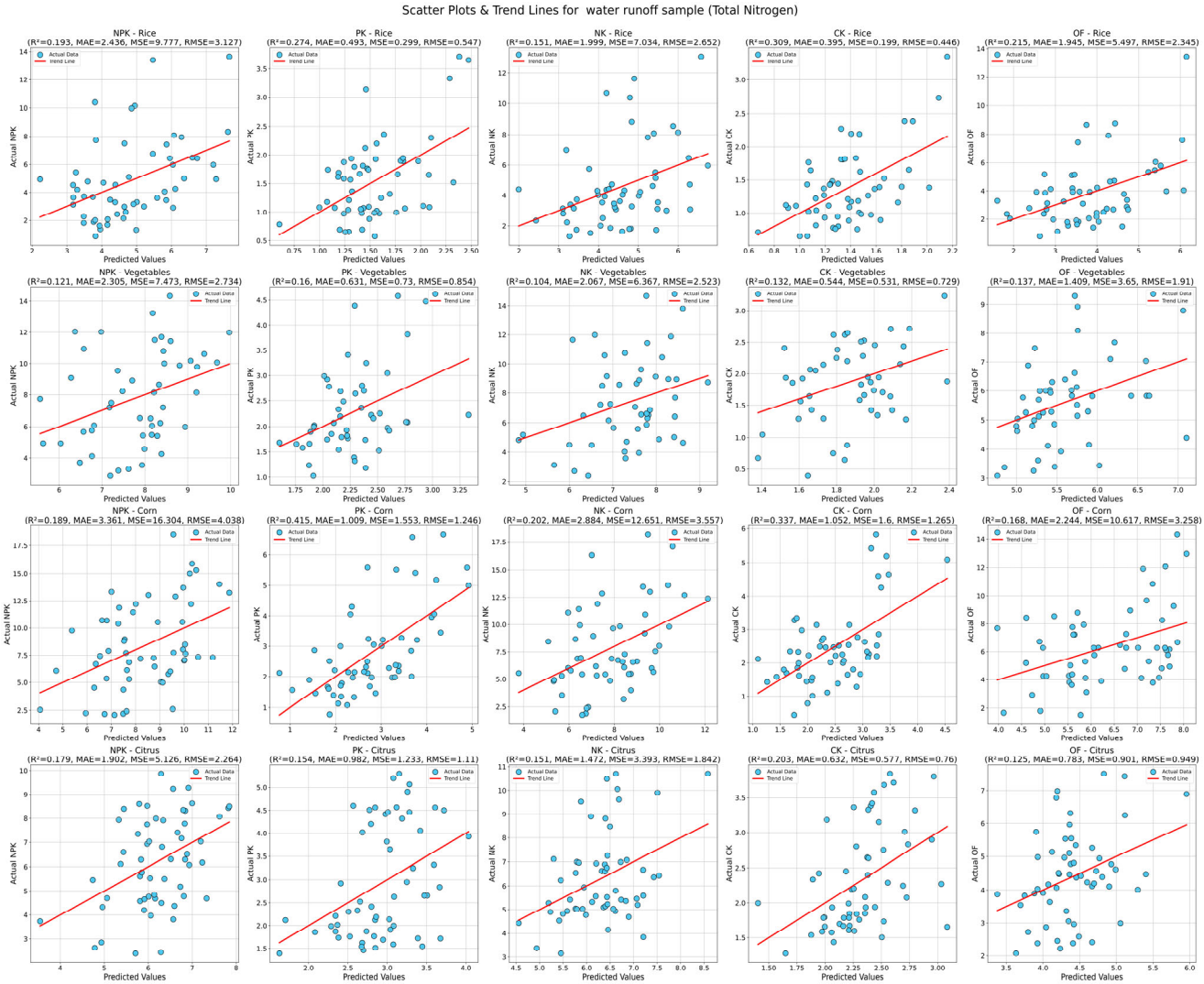


Figure 8. Multiple linear regression of total nitrogen prediction in water runoff, which shows the correspondence between the actual and predicted values of total nitrogen under different fertilization combinations (NPK, PK, NK, NP, and OF) and crop types (rice, vegetables, corn, and citrus). Relevant statistical indicators are marked in the figure, including Pearson correlation coefficient (R), mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE).

Compared with total nitrogen, the multiple linear regression model for total phosphorus showed less obvious dependence on meteorological factors (Figure 9). Only the results from the cornfield showed relatively high goodness of fit. The R^2 values of the NPK, PK, NK, and CK groups are 0.272, 0.243, 0.328, and 0.301. The R^2 values of all fertilizer treatment groups for the other three crops remained at relatively low levels. For example, the R^2 values of the vegetable field ranged from 0.183 to 0.271, with MAE from 0.019 to 0.143. The citrus field had R^2 values ranging from 0.03 to 0.255, and the model error was generally under control. However, in the rice field, the model accuracy dropped significantly, especially for the NPK and OF groups, where R^2 drops to 0.093 and 0.048, respectively, and RMSE rose to 0.045–0.318, showing nonlinear interference characteristics. Cluster 2 of vegetables showed dry and cold weather conditions accompanied by strong disturbances in which the dew point is −1.435, the minimum temperature is −1.304, and the average wind speed is as high as 0.934. Cluster 1 of citrus showed a similar weather condition, with the dew point of −0.009 and the minimum temperature of −0.062, with 0.106 average wind speed. This type of climate background could be triggered by cold gusts or seasonal high-pressure intrusion,

which has impacts on the surface phosphorus migration and particle binding mechanism. Therefore, it is difficult for the model to accurately reconstruct the concentration changes in a linear regression model.

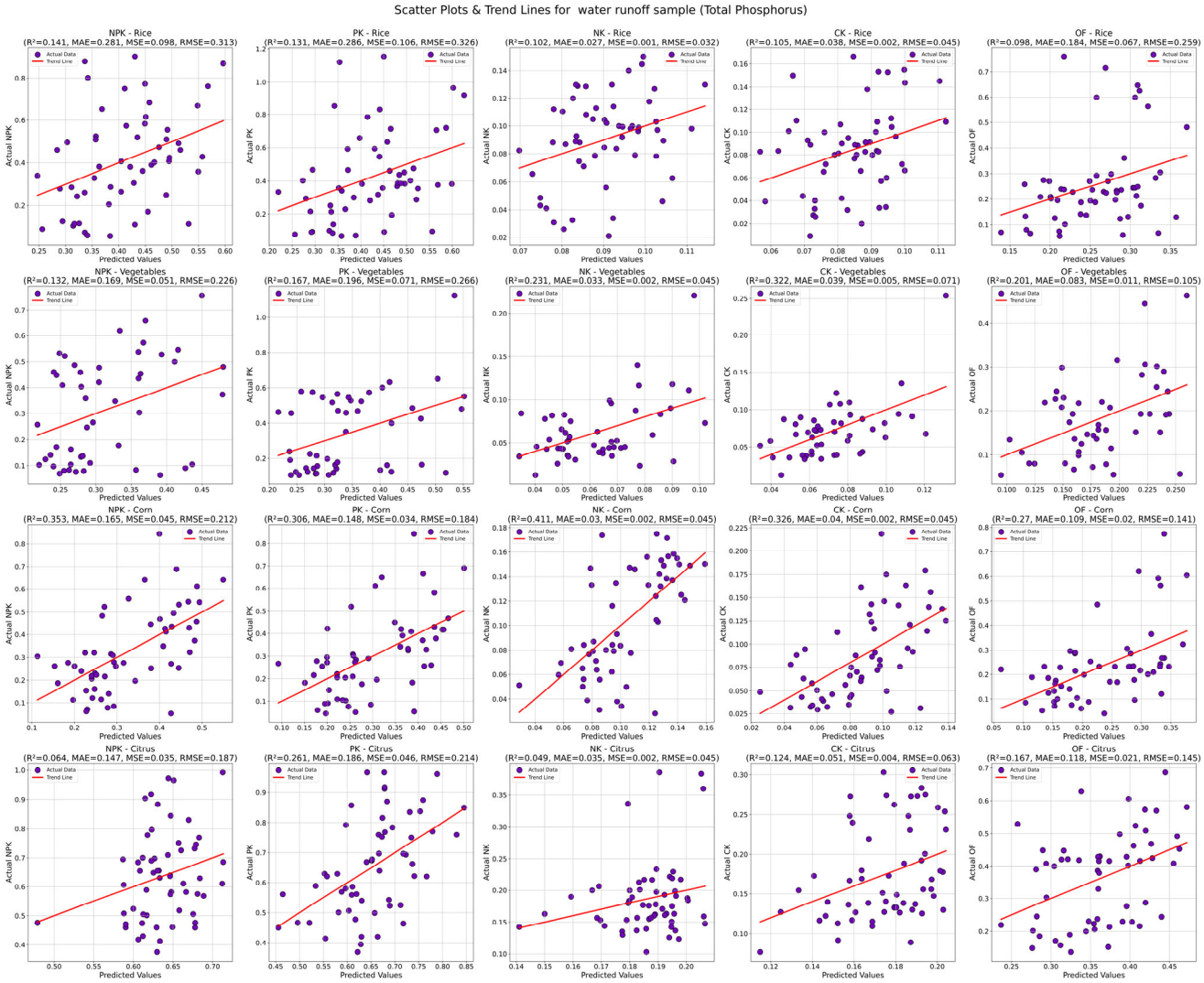


Figure 9. Multiple linear regression of total phosphorus prediction in water runoff, which shows the correspondence between the actual and predicted values of total phosphorus under different fertilization combinations (NPK, PK, NK, NP, and OF) and crop types (rice, vegetables, corn, and citrus). Relevant statistical indicators are marked in the figure, including Pearson correlation coefficient (R), mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE).

4. Discussion

4.1. Causes for Fluctuation of Nitrogen and Phosphorus Concentrations Under Different Fertilization Treatments

Figures 2 and 3 show that the total nitrogen and phosphorus concentrations under different fertilization treatments fluctuated significantly during the five-year monitoring period, especially in the NPK, NK, and OF groups, where the concentrations varied widely and had significant interannual differences. This fluctuation was mainly affected by two factors: (1) fertilization significantly increased the effective concentration of nutrients in the soil, resulting in an enhanced ability of nutrients to migrate in runoff; (2) rainfall intensity and frequency had a regulatory effect on the dilution and scouring of nutrients, especially in years of heavy rainfall, when the total nitrogen concentration was affected by the dilution effect and showed a downward trend. It should be noted that even in

the NK group without phosphate fertilizer, the total nitrogen concentration is still high (for example, the average corn-NK group reached 7.83 mg/L), indicating that nitrogen application has a dominant effect on runoff nitrogen concentration, and this effect does not depend on whether phosphorus is applied simultaneously. The organic fertilizer treatment group (OF) has a larger fluctuation range, which may be related to the mineralization speed, component diversity, and time lag of organic fertilizer release, resulting in dynamic instability of nitrogen release. The total nitrogen concentrations of PK (no nitrogen) and CK (no fertilizer) groups were relatively stable in each year, with average values mostly in the range of 1.5–3.5 mg/L. This indicates that the background nitrogen concentration under the non-fertilization treatment may come from mineral nitrogen (such as ammonium nitrogen and nitrate nitrogen) generated by the decomposition of organic matter in the soil or from historically accumulated fertilizer residues [27].

4.2. Correlation Between Crop Types and Nitrogen/Phosphorus Loss

The characteristics of different crops also have a regulatory effect on total nitrogen and phosphorus migration. This study involves four types of crops: rice, vegetables, corn, and citrus, with different root depths, canopy structures, planting densities, and tillage management methods, which affect soil moisture distribution and nutrient transformation pathways [28]. Taking rice as an example, the fields are flooded all year round, resulting in an oxygen-deficient environment, which is conducive to the denitrification and ammonification of nitrogen by the microorganisms, thereby enhancing the release rate of nitrogen. Cornfields usually rely on surface irrigation systems and are more prone to surface scouring during rainfall events, resulting in rapid loss of nitrogen and phosphorus [29]. Vegetable fields may trigger particulate phosphorus binding and surface migration processes in the short and medium term due to high planting density and frequent tillage. Citrus fields have high vegetation coverage and less soil disturbance, which may inhibit runoff erosion [30]. It should be noted that the cold and dry weather can potentially reduce the activity of soil microorganisms, thereby disturbing the phosphorus adsorption and desorption process and leading to regression model prediction deviation. Therefore, different crop systems not only affect the initial concentration of nutrients but also jointly regulate the migration path of nutrients in water bodies through microclimate and root activity. Nutrient management should consider crop-specific practices rather than only increased fertilizer inputs on the basis of the differences among crop systems. For rice and corn systems, improved timing and placement of fertilizers under high-rainfall conditions can help reduce nitrogen and phosphorus losses. Regarding vegetable systems, minimizing surface disturbance can improve fertilizing performance. Last but not least, maintaining canopy cover of citrus and soil stability can reduce nutrient runoff.

4.3. Characteristics and Limitations of Multiple Linear Regression Models Driven by Meteorology Data

The multivariate linear regression model in this study uses eight types of meteorological variables as independent variables, which can achieve a good prediction of nitrogen and phosphorus concentrations in some crops. In rice and corn, the CK and PK treatment groups showed significant fitting advantages, such as the R^2 of rice-CK reaching 0.309. MAE was 0.395, and the R^2 of corn-PK was as high as 0.415. It indicates that meteorological conditions have a strong driving force on nutrient concentration under the background of fertilizer deficiency or single-factor fertilization. However, for compound fertilization groups such as NPK and OF, the model fitting ability decreased, such as the R^2 of rice-NPK dropping to 0.193 and the R^2 of corn-OF being only 0.168. The hypothesis is that the linear effect of meteorological variables may be disturbed by nonlinear processes such as soil microbial activity and organic matter mineralization in complex fertilization scenarios [31].

In addition, the overall performance of the model was limited in the modeling of total phosphorus. A relatively good fit was achieved in corn (such as NPK group R^2 = 0.272) and in vegetables and citrus. R^2 was mostly below 0.25, indicating that the phosphorus migration mechanism is more complex and more sensitive to surface disturbance and particulate matter.

5. Conclusions

The results show that fertilization treatments had a significant effect on the total concentration and losses of nitrogen and phosphorus. The total nitrogen concentration in runoff water of the four crops in NPK and NK groups was generally higher (4.52–8.37 mg/L), and the amount of nitrogen loss in the cornfield was the largest (NPK up to 18.75 kg/hm^2), which was closely related to the large nitrogen requirement of corn, loose soil, and erosion intensity of rainfall runoff. The NPK and PK groups generally had the highest total phosphorus concentration among all crops, especially in citrus (0.64–0.66 mg/L), which could be related to the soil phosphorus accumulation and root distribution characteristics of citrus. Meteorological factors also had a relatively high impact on the losses of nitrogen and phosphorus. High temperature, high humidity, or heavy precipitation events can enrich nitrogen in surface runoff. While meteorological variables were ineffective predictors for phosphorus runoff ($R^2 < 0.3$). Phosphorus loss was more affected by short-term extreme weather, such as dry cold storms. This matches the results of K-Means cluster analysis, in which the high nitrogen concentration sample clusters mostly correspond to summer meteorological characteristics, while phosphorus was more closely related to sudden weather events. Under the conditions of fertilizer deficiency or single-factor fertilization (CK and PK), the multiple linear regression model has a good predictive effect on nitrogen concentration in rice and corn (R^2 up to 0.415). However, the model's fitting accuracy significantly decreases in the complex fertilization scenarios (NPK and OF) due to the interference of nonlinear processes such as soil microbial activities and organic matter mineralization. It can be deduced that different crop types could regulate nitrogen and phosphorus migration through root depth and tillage management methods. Rice fields promote the release of nitrogen in the anoxic environment formed by long-term flooding. Cornfields rely on surface irrigation, and it is prone to severe surface erosion during rainfall, thereby aggravating the loss of nitrogen and phosphorus. Due to the high planting density and frequent tillage in vegetable fields, the combination and migration of particulate phosphorus may be triggered. Higher vegetation coverage of citrus can inhibit runoff erosion to a certain extent. This study also has some limitations. For example, the field experiments were conducted at a single site, and the related analysis only focused on surface runoff processes. In addition, linear regression models could not fully capture nonlinear nutrient dynamics under extreme weather conditions. It would further enhance the generalizability of the findings if future studies used multi-site observations and nonlinear modeling.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/su18020590/s1.

Author Contributions: R.J.: Conceptualization, data curation, formal analysis, investigation, methodology, software, validation, writing—original draft, and writing—review and editing; Y.X.: conceptualization, data curation, formal analysis, investigation, methodology, validation, writing—original draft, and writing—review and editing; Z.H.: data curation, formal analysis, investigation, methodology, validation, writing—original draft, and writing—review and editing; X.Y.: data curation, formal analysis, investigation, methodology, validation, and writing—review and editing; X.L.: formal analysis, investigation, project administration, validation, visualization, and writing—review and editing; W.D.: data curation, formal analysis, investigation, methodology, and visualization; F.Z.: conceptualization, data curation, methodology, investigation, and validation; T.C.: formal analysis, data

curation, methodology, and investigation; X.W.: formal analysis, data curation, investigation, and validation; X.H.: data curation, investigation, validation, and visualization; Z.Z.: data curation, funding acquisition, project administration, resources, and supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (Grant No. 2023YFD1901300, Center for Science and Technology Development, Ministry of Agriculture and Rural Affairs, China) and Guangdong Provincial Key R&D Program (Grant No. 2023B0202030001, Department of Science and Technology of Guangdong Province, China).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: Author Wenbin Duan is employed by Anyang Huanshui Park Management Station. Author Feifan Zeng is employed by Quliang Electronics Co., Ltd. Author Tianyi Chen is employed by Maoming Energy Conservation Center. These employments are unrelated to the present research. The employers did not provide funding for this study and had no role in the study design, data collection, data analysis, decision to publish, or preparation of the manuscript. All authors declare that they have no additional commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Xie, Z.J.; Ye, C.; Li, C.H.; Shi, X.G.; Shao, Y.; Qi, W. The global progress on the non-point source pollution research from 2012 to 2021: A bibliometric analysis. *Environ. Sci. Eur.* **2022**, *34*, 17. [CrossRef]
2. Fleming, P.M.; Stephenson, K.; Collick, A.S.; Easton, Z.M. Targeting for nonpoint source pollution reduction: A synthesis of lessons learned, remaining challenges, and emerging opportunities. *J. Environ. Manag.* **2022**, *308*, 10. [CrossRef]
3. Wu, Y.P.; Chen, J. Investigating the effects of point source and nonpoint source pollution on the water quality of the East River (Dongjiang) in South China. *Ecol. Indic.* **2013**, *32*, 294–304. [CrossRef]
4. Delkash, M.; Al-Faraj, F.A.M.; Scholz, M. Impacts of Anthropogenic Land Use Changes on Nutrient Concentrations in Surface Waterbodies: A Review. *Clean-Soil Air Water* **2018**, *46*, 10. [CrossRef]
5. Pretty, J.; Sutherland, W.J.; Ashby, J.; Auburn, J.; Baulcombe, D.; Bell, M.; Bentley, J.; Bickersteth, S.; Brown, K.; Burke, J.; et al. The top 100 questions of importance to the future of global agriculture. *Int. J. Agric. Sustain.* **2010**, *8*, 219–236. [CrossRef]
6. You, H.Y.; Li, J.W.; Xu, F.Y. Off-farm employment and nonpoint source pollution from chemical fertilizers in China: Mediating role of farmland transfer. *Environ. Dev. Sustain.* 2024. *online ahead of print*. [CrossRef]
7. Sun, B.; Zhang, L.X.; Yang, L.Z.; Zhang, F.S.; Norse, D.; Zhu, Z.L. Agricultural Non-Point Source Pollution in China: Causes and Mitigation Measures. *Ambio* **2012**, *41*, 370–379. [CrossRef]
8. Ju, X.T.; Gu, B.J.; Wu, Y.Y.; Galloway, J.N. Reducing China's fertilizer use by increasing farm size. *Glob. Environ. Chang.-Hum. Policy Dimens.* **2016**, *41*, 26–32. [CrossRef]
9. Wang, C.; Shen, Y.; Fang, X.T.; Xiao, S.Q.; Liu, G.Y.; Wang, L.G.; Gu, B.J.; Zhou, F.; Chen, D.L.; Tian, H.Q.; et al. Reducing soil nitrogen losses from fertilizer use in global maize and wheat production. *Nat. Geosci.* **2024**, *17*, 22. [CrossRef]
10. Bechmann, M.; Stålnacke, P. Agricultural nitrogen and phosphorus pollution in surface waters. In *Oxford Research Encyclopedia of Environmental Science*; Oxford University Press: Oxford, UK, 2019.
11. Wurtsbaugh, W.A.; Paerl, H.W.; Dodds, W.K.J.W.I.R.W. Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. *Wiley Interdiscip. Rev. Water* **2019**, *6*, e1373. [CrossRef]
12. Wang, X.B.; Wang, X.L.; Sheng, H.J.; Wang, X.Z.; Zhao, H.T.; Feng, K. Excessive Nitrogen Fertilizer Application Causes Rapid Degradation of Greenhouse Soil in China. *Pol. J. Environ. Stud.* **2022**, *31*, 1527–1534. [CrossRef]
13. Srivastav, A.L.; Patel, N.; Rani, L.; Kumar, P.; Dutt, I.; Maddodi, B.S.; Chaudhary, V.K. Sustainable options for fertilizer management in agriculture to prevent water contamination: A review. *Environ. Dev. Sustain.* **2024**, *26*, 8303–8327. [CrossRef]
14. Hussain, F.; Ahmed, S.; Muhammad Zaigham Abbas Naqvi, S.; Awais, M.; Zhang, Y.; Zhang, H.; Raghavan, V.; Zang, Y.; Zhao, G.; Hu, J.J.A. Agricultural Non-Point Source Pollution: Comprehensive Analysis of Sources and Assessment Methods. *Agriculture* **2025**, *15*, 531. [CrossRef]
15. Agrawal, N.; Govil, H.; Kumar, T. Agricultural land suitability classification and crop suggestion using machine learning and spatial multicriteria decision analysis in semi-arid ecosystem. *Environ. Dev. Sustain.* **2025**, *27*, 13689–13726. [CrossRef]

16. Huan, J.; Fan, Y.X.; Xu, X.G.; Zhou, L.W.; Zhang, H.; Zhang, C.; Hu, Q.C.; Cai, W.X.; Ju, H.R.; Gu, S.L. Deep learning model based on coupled SWAT and interpretable methods for water quality prediction under the influence of non-point source pollution. *Comput. Electron. Agric.* **2025**, *231*, 14. [CrossRef]
17. Yin, M.W.; Wu, Z.J.; Zhang, Q.; Su, Y.Y.; Hong, Q.; Jia, Q.Q.; Wang, X.; Wang, K.; Cheng, J.R. Combining SWAT with Machine Learning to Identify Primary Controlling Factors and Their Impacts on Non-Point Source Pollution. *Water* **2024**, *16*, 3026. [CrossRef]
18. Zeng, F.F.; Zuo, Z.; Mo, J.C.; Chen, C.Y.; Yang, X.J.; Wang, J.J.; Wang, Y.; Zhao, Z.Q.; Chen, T.Y.; Li, Y.T.; et al. Runoff Losses in Nitrogen and Phosphorus from Paddy and Maize Cropping Systems: A Field Study in Dongjiang Basin, South China. *Front. Plant Sci.* **2021**, *12*, 15. [CrossRef]
19. Ho, K.C.; Chow, Y.L.; Yau, J.T.S. Chemical and microbiological qualities of the East River (Dongjiang) water, with particular reference to drinking water supply in Hong Kong. *Chemosphere* **2003**, *52*, 1441–1450. [CrossRef]
20. Ding, J.; Jiang, Y.; Liu, Q.; Hou, Z.; Liao, J.; Fu, L.; Peng, Q. Influences of the land use pattern on water quality in low-order streams of the Dongjiang River basin, China: A multi-scale analysis. *Sci. Total Environ.* **2016**, *551*, 205–216. [CrossRef]
21. Kleinman, P.J.A.; Srinivasan, M.S.; Dell, C.J.; Schmidt, J.P.; Sharpley, A.N.; Bryant, R.B. Role of rainfall intensity and hydrology in nutrient transport via surface runoff. *J. Environ. Qual.* **2006**, *35*, 1248–1259. [CrossRef]
22. Kulkarni, S. Climate change, soil erosion risks, and nutritional security. In *Climate Change and Resilient Food Systems: Issues, Challenges, and Way Forward*; Springer: Singapore, 2021; pp. 219–244. [CrossRef]
23. Costa, D.; Sutter, C.; Shepherd, A.; Jarvie, H.; Wilson, H.; Elliott, J.; Liu, J.; Macrae, M. Impact of climate change on catchment nutrient dynamics: Insights from around the world. *Environ. Rev.* **2022**, *31*, 4–25. [CrossRef]
24. *HJ 636-2012*; Water Quality—Determination of Total Nitrogen—Alkaline Potassium Persulfate Digestion UV Spectrophotometric Method. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 2012.
25. *HJ 671-2013*; Water Quality—Determination of Total Phosphorus—Continuous Injection–Ammonium Molybdate Spectrophotometric Method. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 2013.
26. Hodson, T.O. Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev.* **2022**, *15*, 5481–5487. [CrossRef]
27. Wang, X.; Wang, B.; Gu, W.; Li, J.J.A. Effects of carbon-based fertilizer on soil physical and chemical properties, soil enzyme activity and soil microorganism of maize in northeast China. *Agronomy* **2023**, *13*, 877. [CrossRef]
28. Shah, F.; Wu, W. Soil and Crop Management Strategies to Ensure Higher Crop Productivity within Sustainable Environments. *Sustainability* **2019**, *11*, 1485. [CrossRef]
29. Sherman, J.F.; Young, E.O.; Jokela, W.E.; Casler, M.D.; Coblentz, W.K.; Cavadini, J.J.S.S. Influence of soil and manure management practices on surface runoff phosphorus and nitrogen loss in a corn silage production system: A paired watershed approach. *Soil Syst.* **2020**, *5*, 1. [CrossRef]
30. Liu, R.; Zhang, Y.T.; Wang, Z.C.; Zhang, X.L.; Xu, W.J.; Zhang, J.W.; Zhang, Y.Q.; Hu, B.; Shi, X.J.; Rennenberg, H. Groundcover improves nutrition and growth of citrus trees and reduces water runoff, soil erosion and nutrient loss on sloping farmland. *Front. Plant Sci.* **2024**, *15*, 16. [CrossRef]
31. Heinen, M.; Assinck, F.; Groenendijk, P.; Schoumans, O. Soil dynamic models: Predicting the behavior of fertilizers in the soil. In *Biorefinery of Inorganics: Recovering Mineral Nutrients from Biomass and Organic Waste*; Wiley: New York, NY, USA, 2020; pp. 405–435.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.