forests

MDPI

Article

Intelligent Analysis Strategy for the Key Factor of Soil Nitrogen and Phosphorus Loss via Runoff under Simulated Karst Conditions

Yuqi Zhang, Rongchang Zeng, Tianyang Li, Lan Song and Binghui He *

College of Resources and Environment, Southwest University, Chongqing 400715, China; zhangyqi@swu.edu.cn (Y.Z.); 202131051003@mail.bnu.edu.cn (R.Z.); tyli53@swu.edu.cn (T.L.); slan1999@email.seu.edu.cn (L.S.)

* Correspondence: hebinghui@swu.edu.cn; Tel.: +86-23-68251249

Abstract: Given the complex influence of various factors on soil nitrogen (N) and phosphorus (P) loss through runoff in a karst environment, analyzing the importance of different factors to determine the most efficient method for soil nutrient conservation remains a key challenge. Herein, we proposed a novel intelligent analysis strategy based on the Random Forest (RF) regression algorithm to identify the main features and discover the fundamental mechanisms among them under a rock-exposed karst slope with synchronous existence of surface runoff and subsurface leakage. Typically, the results indicated that the rock–soil angle (β) was the main factor influencing soil N and P loss, which was further confirmed based on the RF regression-multifactor analysis. The proposed strategy was used to characterize the relationships of inflow rate, soil bed–ground angle, and rock–soil angle with soil N and P concentrations in soil surface runoff, subsurface runoff, and fissure runoff to study the potential application of soil N and P loss under karst conditions. Our results provide a new approach and promising potential for soil nutrient conservation and related soil and plant research.

Keywords: soil N and P loss; RF regression algorithm; intelligent analysis; feature; karst

Citation: Zhang, Y.; Zeng, R.; Li, T.; Song, L.; He, B. Intelligent Analysis Strategy for the Key Factor of Soil Nitrogen and Phosphorus Loss via Runoff under Simulated Karst Conditions. *Forests* **2023**, *14*, 2109. https://doi.org/10.3390/f14102109

Academic Editor: Choonsig Kim

Received: 3 September 2023
Revised: 18 October 2023
Accepted: 18 October 2023
Published: 21 October 2023

1. Introduction

Soil nitrogen (N) and phosphorus (P) are indispensable nutrient elements to plant recolonization and establishment. The majority of soil's N and P are not used efficiently, but are instead lost through various runoff pathways as a result of water erosion, with annual losses higher than 18.88 kg hm^{-2} for N and 81.05 kg hm^{-2} for P in northwest China. Although great efforts have been made to reduce soil N and P loss through runoff, it is still difficult due to the complex relationship of changes in many environmental factors, such as inflow rate, slope gradients, and geological conditions [1–3]. Recently, it has been reported that, under optimal conditions, the relationship between soil N and P los and rainfall-produced runoff could be fitted as a mathematic function, but the pathways of soil N and P loss in these models vary significantly [4–6]. Higher-level rainfall-produced runoff causes soil N and P loss, mainly through particulate loss, whereas lower-level rainfall runoff causes soluble soil N and P loss. Due to these significant differences, it is not surprising that many parameters could have simultaneous influences on soil N and P loss via runoff, with various factors affecting different pathways (e.g., surface loss, subsurface loss, and leakage) and levels. The relationship between soil N and P loss and these factors would be too complicated to be fitted using specific mathematic functions. To the best of our knowledge, no studies have presented such experiments and fundamental mechanisms of interactions among various features (i.e., inflow rate, slope gradients, and geological conditions) in the complex soil N and P loss system.

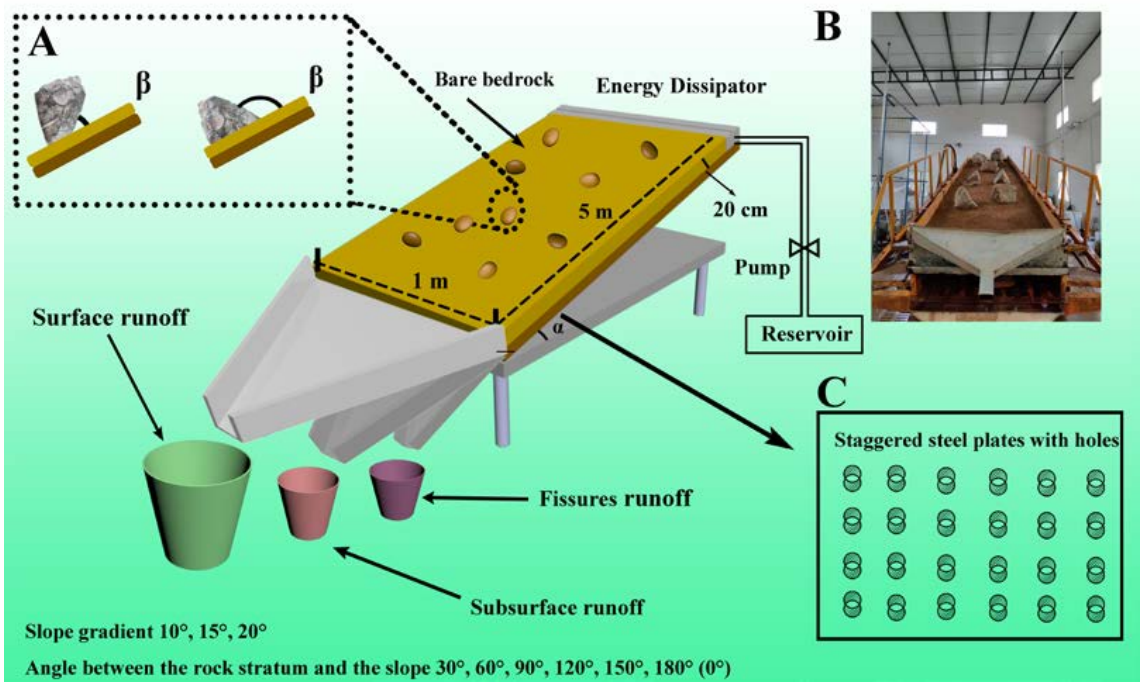
Karst trough valleys are some of the most widespread landforms in the world. The karst regions are crucial water providers, accounting for 10.2% of the Earth's area and

supplying approximately 25% of the world's water [7]. Due to the special geological structure and chemical composition of the bedrock, such as a high concentration level of $CaCO_3$, the karst environment is characterized by significant water erosion and losses of soil nutrients [8]. Specifically, due to the exposed rocks and the synchronous existence of surface runoff and subsurface leakage in karst environments, such as the karst trough valley slope, analyzing the importance of different environmental factors is still a key challenge (e.g., inflow rates, soil bed–ground angles, and rock–soil angles) to find the most efficient method for soil N and P loss prediction. Indeed, characterization of the relationships between these environmental factors and soil N and P loss will be crucial for soil and plant studies in karst trough valleys.

During the analysis of various characteristics of soil N and P loss, the intrinsic characteristics of complexity, uncertainty, and non-linearity, and especially the nonlinear responses among them, should be considered. The common methods mostly adopt a single regression model including inflow rates, soil bed–ground angles and rock–soil angles and some other related features, which is less practical and thus it is hard to build an accurate predictive computational model [9–12]. In recent years, machine learning algorithms have been tested for their ability to identify main features and the complex relationship among them [13,14], which has gradually become the mainstream alternative resolution for complex systems. Among various machine learning algorithms, the emerging Random Forest (RF) algorithm has been regarded as one of the most precise prediction methods for classification and regression to efficiently model complex interactions among various features due to its capability to discover complex nonlinear relationships among independent or dependent variables without statistical assumptions [15–17].

To date, RF algorithms have been successfully employed in studies of environmental research, such as water quality and its related effects, demonstrating great potential as an efficient resolution to identify complex nonlinear relationships among various factors in complex systems [18,19]. Commonly, RF algorithms use multiple decision trees with the same distribution to set up a forest based on a combination of tree predictors, where each tree is generated using a random vector sampled independently from the input vector and the decision tree is a non-parametric supervised learning method to summarize decision rules from a series of data with features and labels and present these rules to solve classification and the regression problem via the structure of the tree. To the best of our knowledge, few studies have reported the use of the RF regression algorithm for estimating the key factors of soil N and P loss under simultaneous rock-exposed karst conditions with the existence of surface runoff and subsurface leakage, although it may be an efficient and useful model [19].

Herein, we proposed a novel intelligent analysis strategy based on the RF regression algorithm to identify the main features and discover complex nonlinear relationships and fundamental mechanisms among them to generate the key factor of soil N and P loss via runoff under a simulated rock-exposed karst slope with variable inflow rates (v), soil bed–ground angles (α), and rock–soil angles (β). Typically, a laboratory scouring experimental system was used at varying α and β, as shown in Scheme 1, and the collected water was obtained from soil surface runoff, subsurface runoff, and fissure runoff to study the distribution of soil N and P loss via different pathways to evaluate the main features in this system with the RF regression algorithm. Based on the relationships between the main features and soil N and P loss, we expected to propose the main features of various N and P loss pathways through runoff, providing a new avenue for soil and plant research in the karst environment worldwide.



Scheme 1. Schematic diagram of the proposed intelligent analysis strategy based on RF regression algorithm (soil surface runoff, subsurface runoff, and fissure runoff). A: Schematic diagram of the different rock-soil angles; B: The photo of the experimental setup; C: Scheme of the staggered steel plates with holes of the experimental setup.

2. Materials and Methods

2.1. Soil Sample

The test soil in this study was loessial loam collected from farming land in the Jigong Mountain area (106°27′19″ E, 29°47′41″ N), Chongqing City, which is one of the typical karst trough valley areas in southwest China. The bulk soil was first gently crushed and separated using a 10 cm diameter sieve to remove plant roots, rocks, and other debris. The soil particle size distribution and bulk density of the test soil are listed in Table 1 and were determined based on the pipette method according to previous reference [20].

Table 1. Particle size distribution and bulk density of the test soil.

Soil Particle Size Distribution			Bulk Density g cm^{-3}
<0.002 mm	0.002–0.02 mm	>0.02 mm	
28.23%	51.09%	20.68%	1.21

2.2. Experimental Setup

The experimental water scouring simulation was carried out using specially designed experimental equipment that included a water runoff setup, a soil flume, and three water collecting devices (Scheme 1). Specifically, the water runoff setup contained a water supply and simulator with water pipes and a simulated water controlling chamber. The 5.0 m long, 1.0 m wide, and 0.2 m deep soil flume was specially designed with three steel catchment collectors to collect soil surface runoff, subsurface runoff, and fissure runoff in containers. In this study, α was set as 10, 15, and 20 by adjusting the hydraulic jack of the flume, β was set as 30, 60, 90 120, 150, and 180 (0) by changing the angle of the rocks, and v was set as 5, 7.5, and 10 dm^3 min^{-1} by changing the water pressure [21].

Before the start of the test, the crevasses were adjusted to meet the test requirements by adjusting the dislocation of the round holes in the soil tank floor. After fixing the soil, the test fill thickness was 20 cm. After calculating the fill amount according to the lower layer

of 5 cm (bulk weight 1.3 g cm^{-3}) and the upper layer of 15 cm (bulk weight 1.2 g cm^{-3}), the load was filled by weighing layers. In order to reduce the edge effect, when filling the soil tank, Vaseline was applied on both sides of the inner wall, and the height was filled slightly. During the filling process, special wooden boards were used to compact the layers and the soil was raked to ensure that the soil was smooth. After filling the soil, 9 limestone blocks with a diameter greater than 25 cm prepared before were randomly arranged in the soil tank according to the corresponding fixed contact area of rock and soil. The inclinometer was used to adjust the rock–slope angle, their placement was fixed, and the area exposed to the soil surface was measured to ensure that the rock exposure rate is consistent with the natural condition.

2.3. Experimental Processes

In this study, v, α, and β were the three main features used as a model to construct the intelligent analysis strategy and investigate its application characteristics to generate the relationship between them and soil N and P loss. Figure 1 shows the experimental conditions, in which three different conditions were employed for every feature. Before the experimental study, the treated soil sample was placed in the steel flume with a bulk density of 1.21 g cm^{-3} and an increment of 5 cm. Additionally, some rocks with a rock–soil angle were added onto the soil surface to simulate the condition of a karst trough valley area. To achieve uniform water content in the profile, an amount of water was sprayed and a plastic film was applied to the soil surface. After that, the flume was adjusted to the designed soil bed–ground angle before the scouring experiment began. Each scouring experimental run lasted for 18 min, and soil surface runoff, subsurface runoff, and fissure runoff were collected every 1 min for the first 6 min and every 1.5 min for the last 12 min in a plastic bucket. Under varied conditions, soil surface runoff, subsurface runoff, and fissure runoff were used for runoff N and P testing.

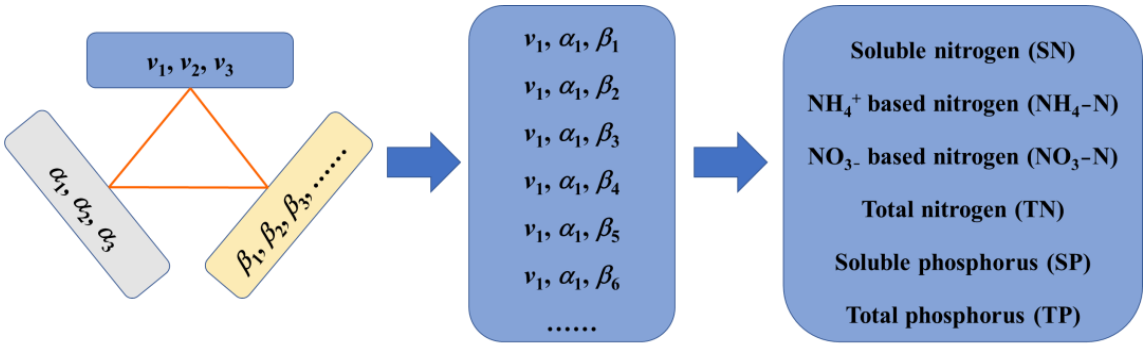


Figure 1. Experimental conditions in this study with different v, α, and β.

According to previous references with some modifications, soluble nitrogen (SN) concentration in different runoff waters was measured using the alkaline potassium persiflage digestion UV spectrophotometric method with separation, ammonium (NH_4^+)-based nitrogen (NH_4-N) concentration was measured based on Nessler's colorimetric method, NO_3^--based nitrogen (NO_3-N) concentration was measured using the UV spectrophotometric method, total nitrogen (TN) concentration was measured using the alkaline potassium persiflage digestion UV spectrophotometric method, soluble phosphorus (SP) concentration was measured based on the ammonium molybdate spectrophotometric method with separation, and total phosphorus (TP) concentration was measured based on the ammonium molybdate spectrophotometric method [22–25]. In these analytical measurements, absorption spectra were measured using a UV-vis spectrophotometer, type UV-2450 (Shimadzu, Kyoto, Japan). A pH-3C digital pH-meter (Shanghai LeiCi Device Works, Shanghai, China) was employed in all of the pH tests in the experiment. Ultrapure water with a resistivity of 18.2 MΩ·cm was used throughout this study.

2.4. Data Set

The data set mainly contained three model features including v, α and β. More than three factors for each feature were employed as input characteristics of the experimental data, and SN, TN, NH_4-N, NO_3-N, SP, and TP concentrations were employed as the output characteristics to study the relationship between model features and soil N and P loss with different pathways and mechanisms. Furthermore, RF algorithms depend on various conditions and characteristics performances were used to generate the relationships with more than 5000 results, and the employed program code is supported in the Supporting Information.

2.5. Data Analysis with Random Forest Regression

RF is a machine learning algorithm based on a classification tree developed by Breimanz on the basis of the Bagging algorithm, which can handle classification and regression problems [26–29]. When the dependent variable is a continuous variable, the RF regression model can be used to explain the influence of several independent variables on the dependent variable. The basic principle of the RF algorithm is to use the bootstrap repeated sampling method to randomly extract some observations of the dependent variable from the data set, randomly select a specified number of variables from the independent variables to determine the nodes of the classification tree, and randomly construct hundreds of classification trees. After selecting the one with the highest degree of repetition, regression trees are used to form a combined model, and the predicted value of the model is formed by calculating the average of multiple regression trees. In this RF system, multiple linear regression requires the data to satisfy many assumptions, while the RF algorithm does not need to pre-set the function form, is insensitive to multicollinearity, and can more accurately fit up to thousands of independent variables. The interaction between variables can be considered in the calculation process to complicate the nonlinear relationship with dependent variables. When the RF algorithm deals with multi-level categorical variables, it can avoid the problem of a large increase in estimated parameters and over-fitting. The RF algorithm has no effect on outliers and has strong anti-interference ability. It can still show a robust prediction effect even when there are many discrete and missing values, and it can evaluate the effect of each independent variable on the dependent variable importance.

2.6. Statistical Analysis

In this study, all experimental results were subjected to analysis of variance using the Statistical Studies for Social Sciences software package SPSS v. 17.0 (IBM, New York, NY, USA). Principal component analysis (PCA) was employed to evaluate the relationships of various chemicals and to calculate soil nutrition with N and P.

3. Results

3.1. Characteristics Analysis based on the Normal Single Factor Analysis

In this study, the relationship between the N concentration and v, α, and β in the soil surface runoff, subsurface runoff and fissure runoff, respectively, were characterized based on the normal single factor analysis. As shown in Figure 2, there was a correlation between the SN concentration and v, α, and β in the soil surface runoff, subsurface runoff, and fissure runoff, respectively. In the soil surface runoff, the SN concentration increased slightly when the inflow rate increased from 5 to 10 dm^3 min^{-1}. Although there was an obvious linear relationship between v and SN based on the average value of SN with the same v and different a and b, the correlation coefficient and slope value of the linear function was about 0, indicating that there was no significant relationship between v and SN based on the functional correlation, which could be confirmed by the Pearson correlation coefficient between v and SN of 0.024 (v-SN) based on the statistical analysis, which was significantly lower than 0.05. Similar results were obtained for the soil bed–ground angle. The SN concentration increased when the rock–soil angle was in the range of 30 to 90 and then decreased when the rock–soil angle was higher than 90. It is important to note that the

relative standard deviation for all of these results was clearly higher than that of the used method, and the relative standard deviations of the v and α were significantly higher than that of β, which may have resulted from the effect of different features. Accurately, the v and α had minimal influence on the SN concentration, resulting in a lower relative standard deviation in the relationship between the SN concentration and β. The β had a large influence on the SN concentration and thus resulted in a higher relative standard deviation in the relationship between the SN concentration and the v and α, indicating that different features had different influences on the SN concentration and β may be the main feature of the SN concentration in the soil surface runoff rather than v and α. An accurate relationship among them could not be calculated based on the normal single factor analysis, although it has been widely employed in related studies. Similar results were obtained from the results of the relationship between the TN, NH_4-N, NO_3-N, SP, and TP concentrations and v, α, and β in the soil subsurface runoff and fissure runoff, respectively, which are shown in Supplementary Figures S1–S5.

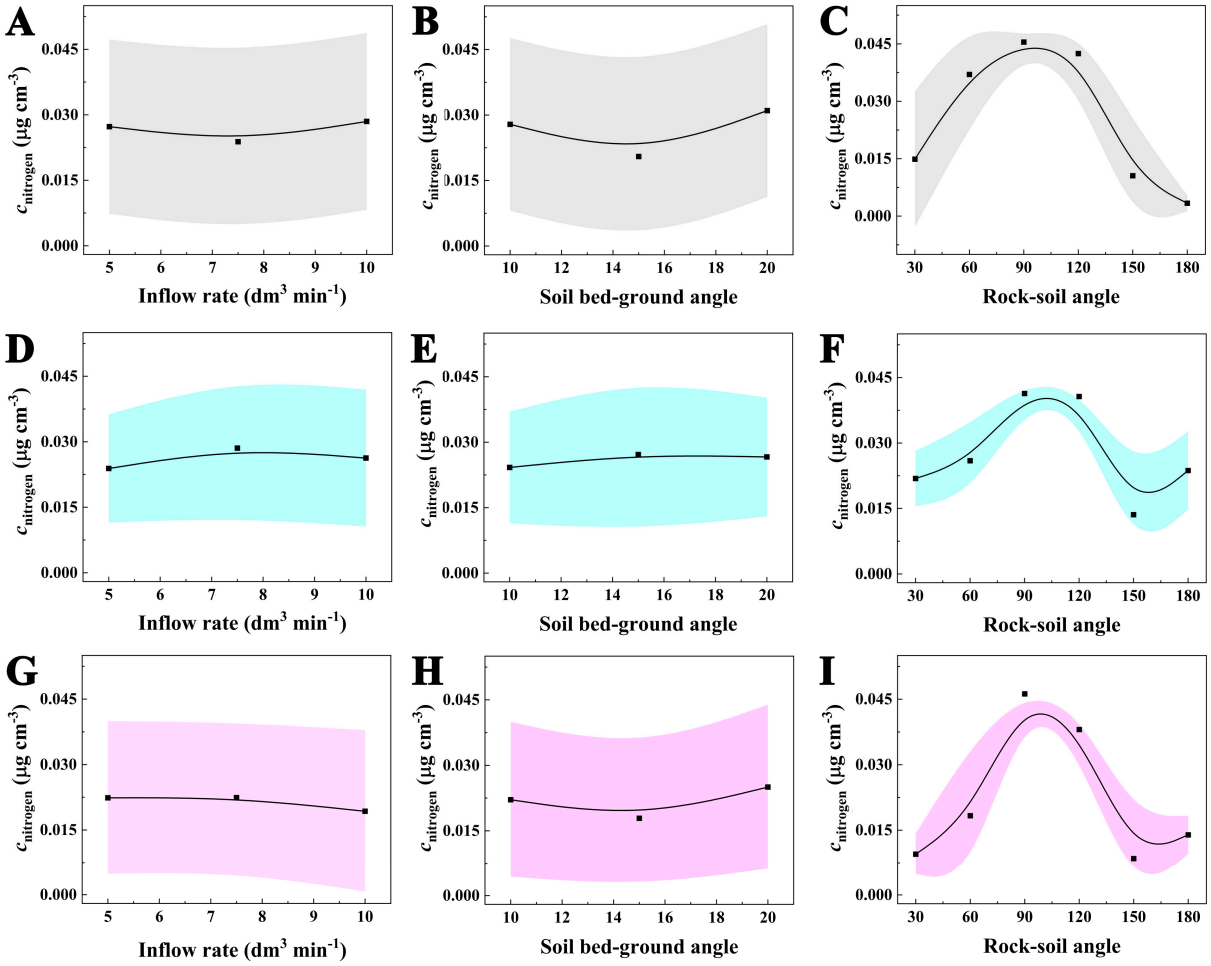


Figure 2. Characteristic relationships between the SN concentration and v, α, and β in soil surface runoff (**A**–**C**), subsurface runoff (**D**–**F**), and fissure runoff (**G**–**I**), respectively. The dot means the average nitrogen concentration and the graph means the standard deviation.

3.2. Characteristics of the Features Based on the Random Forest Regression

To clearly characterize the effect of different features with v, α, and β as a model on the soil N and P loss with the SN, TN, NH_4-N, NO_3-N, SP, and TP concentrations in soil surface runoff, subsurface runoff, and fissure runoff, respectively, the RF regression was used to investigate the effect of these different features with a special value named as "Importance", which was in range of 0 to 1, indicating the possible effect and relationship value between

each variable, such as the soil N and P loss with the SN, TN, NH_4-N, NO_3-N, SP, and TP, and dependent variables such as the v, α, and β. In this system, the relationship between the soil N and P loss and some potential features could be generated as the function f (v, α, β). Due to the different effects of these features, the function was nonlinear, and each single function had a different weight, named f_1 (v), f_2 (α), and f_3 (β), and thus the function f (v, α, β) could be generated as $a f_1$ (v) + $b f_2$ (α) + $c f_3$ (β), in which a, b, and c were the weight of every single function. As shown in Figure 3, the obviously different importance was obtained, indicating the different effects of these features on soil N and P loss. For the SN concentration in the soil surface runoff, an importance of approximately 0.75 was obtained for β, which was significantly higher than that of approximately 0.1 for v and α. Similar results were obtained for the SN concentration in interflow and soil percolating water. All the results indicated that β would be the main feature for the SN concentration in the soil surface runoff, subsurface runoff, and fissure runoff, as indicated in the previous reference that the lower-level rainfall runoff may generate higher soluble soil N and P loss in special conditions. Further studies were performed for the characterization of the v, α, and β of the TN, NH_4-N, NO_3-N, and TP concentrations in the soil surface runoff, subsurface runoff, and fissure runoff, respectively, and similar results were obtained (Figure S6). More importantly, these results generate the Importance factor to estimate the effect of every feature on the soil N and P loss. Additionally, for the results of soil surface runoff, subsurface runoff, and fissure runoff, there was a little difference, indicating that there would be a distribution for the soil N and P loss, which may be important for soil N and P loss, soil fertility, and some other related research.

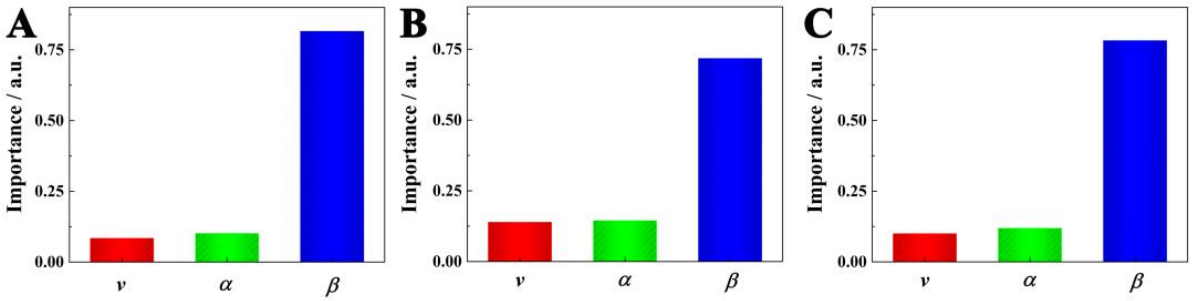


Figure 3. Characterizations of the Importance factor to estimate the effect of every feature, including the v, α, and β, on the soluble nitrogen concentration in the soil surface runoff (**A**), subsurface runoff (**B**), and fissure runoff (**C**), respectively.

3.3. Characteristics Analysis Based on the RF Regression Multifactor Analysis

To verify the practicality of the proposed intelligent analysis strategy based on the RF regression algorithm to identify the main features, a complex nonlinear multifactor analysis was structured with v, α, and β on the TN, NH_4-N, NO_3-N, and TP concentrations in the soil surface runoff, subsurface runoff, and fissure runoff, respectively. As shown in Figure 4A for the SN concentration in the soil surface runoff, there was no significant difference when v and α changed. Clearly, the SN concentration was strongly related to the rock–soil angle (β) due to the lack of an obvious difference with different v and α, and similar shape to that shown in Figure 4B,C. The change in SN concentration in interflow and soil percolating water was similar, although the influence of β was minimal, indicating that β has the greatest effect on the SN concentration in different waters flows. Further studies were conducted based on the complex nonlinear multifactor analysis for the characterization of v, α, and β on TN, NH_4-N, NO_3-N, and TP concentrations in the soil surface runoff, subsurface runoff, and fissure runoff, respectively, and similar results were obtained (Figures S7–S11). All these results confirmed the Importance factor analysis based on the RF regression and generated the main factors under different conditions, with the Importance factor being slightly different under some special conditions, such as the results in NH_4-N characterizations.

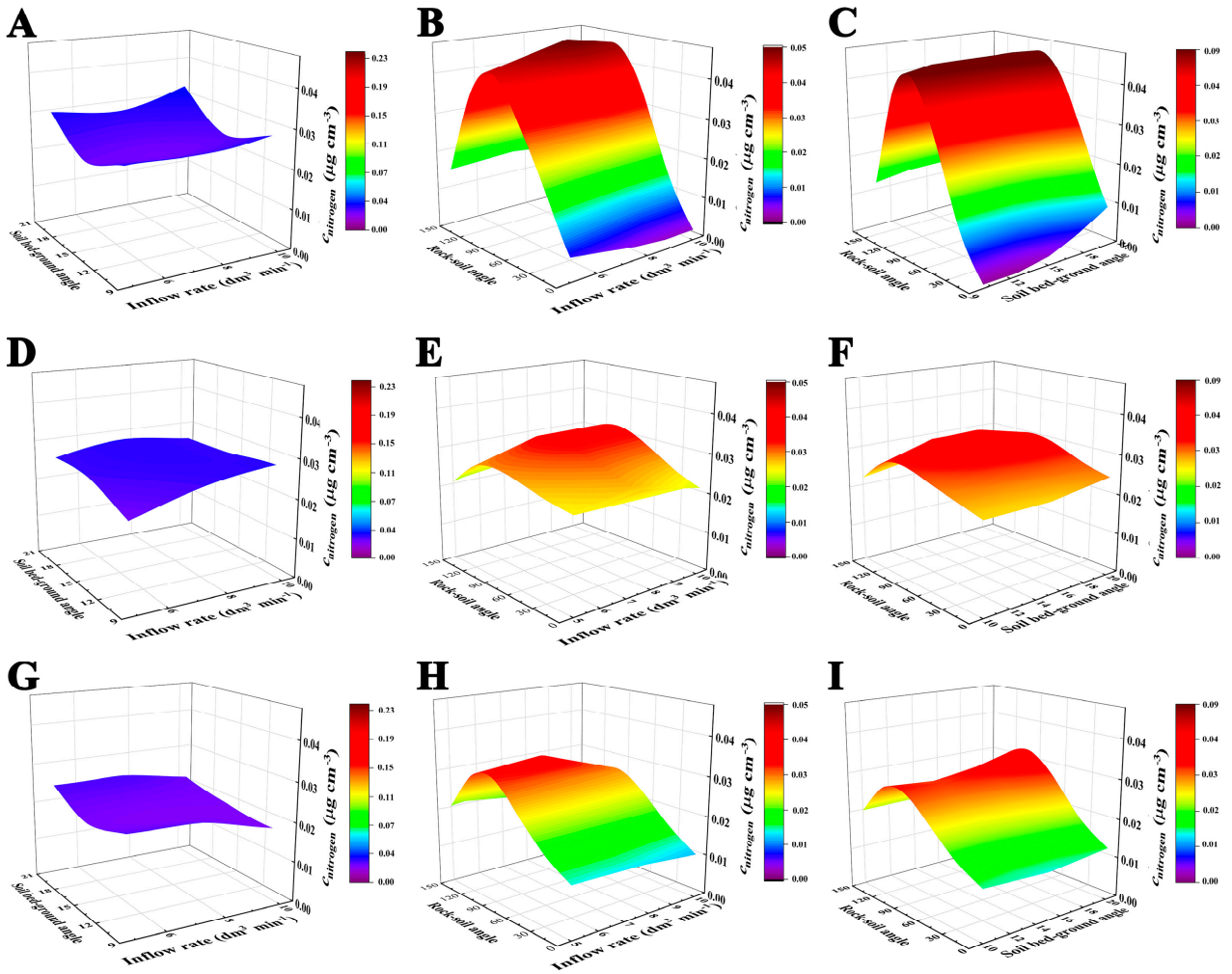


Figure 4. Characterizations of the relationship between the SN concentration in the soil surface runoff (**A**–**C**), subsurface runoff (**D**–**F**), and fissure runoff (**G**–**I**) and v, α, and β, respectively.

3.4. Nitrogen and Phosphorus Nutrition Value Characterizations

To fully evaluate soil nutrition using N and P, a new N and P index known as the NP value was calculated using PCA to overcome the limitation of a single feature, which was calculated based on the complete evaluation of SN, TN, NH_4-N, NO_3-N, SP, and TP concentrations in soil surface runoff, subsurface runoff, and fissure runoff, two main components with effect values of 1.792 and 1.405, when the cumulative percentage was 53.30%. Thus, a complete principal component matrix could be obtained, as illustrated in Figure 5. As can be seen, NO_3-N concentration had a major impact on the NP value. Furthermore, a modified NP value could be calculated using the PCA equation, especially as NP value = 0.25 TN + 0.39 TP + 0.35 SN + 0.28 SP − 0.20 NH_4 + 0.87 NO_3. Thus, the relationship between the NP value and the related effects could be evaluated, including v, α, and β in the soil surface runoff, subsurface runoff, and fissure runoff, respectively. As shown in Figure 6, v, α, and β had an obvious effect on the NP value, with the rock–soil angle having the largest effect of approximately 90, providing a new perspective on N and P loss.

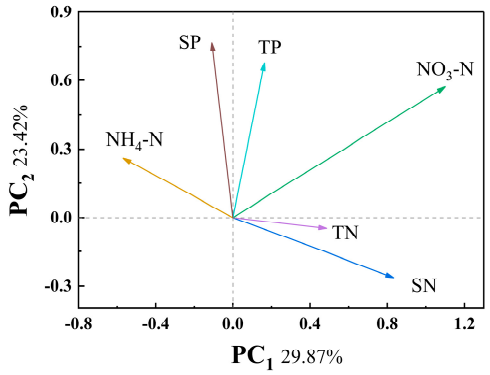


Figure 5. Principal component analysis of SN, TN, NH_4-N, NO_3-N, SP, and TP concentrations in the soil surface runoff, subsurface runoff, and fissure runoff. Variations among the data points could be largely explained by PC_1 and PC_2 with a cumulative percentage of 53.30%.

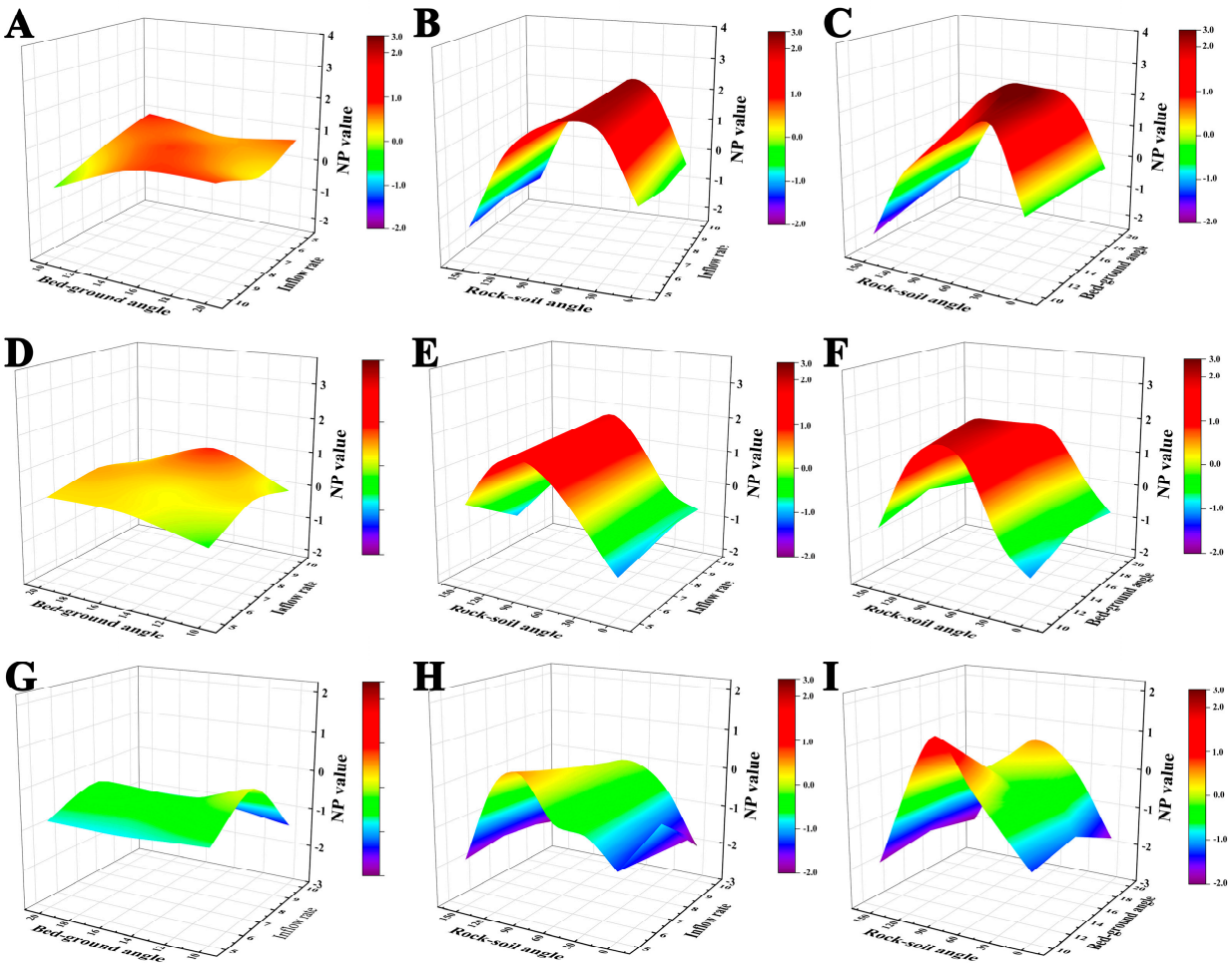


Figure 6. Characterizations of the relationship between the NP value in the soil surface runoff (**A**–**C**), subsurface runoff (**D**–**F**), and fissure runoff (**G**–**I**) and v, α, and β, respectively.

4. Discussion

In this study, we have developed a novel intelligent analysis strategy based on the RF regression algorithm to identify the main features and discover complex nonlinear relationships and fundamental mechanisms among them for the key factor of soil N and P loss under simultaneous rock-exposed karst conditions with the existence of surface runoff and subsurface leakage. Similarly, Wright's group put a lot of effort into investigating the

predictors of soil microbial biomass C, N, and P to incorporate the broad-scale soil microbial nutritional properties and the associated processes into biogeochemical models [1]. The storages of soil C, N, and P in the top soil surface have been successfully estimated. Although the different chemical forms of soil C, N, and P have been strongly identified, the response of these different chemical forms to specific environmental factors is still unclear, such as the inflow rate, soil bed–ground angle, and rock–soil angle in this study. Xu and coauthors described the various impacts on river water quality based on the case of Dongjiang Lake Basin, generating the different environmental factors that could influence water indicators based on the different paths and mechanisms, including C, N, and P with different chemical forms [30]. However, the relationships still needed to be studied further. Kotha and coauthors used machine learning algorithms to calculate the multiplex relationships among various factors, especially based on the RF and some other algorithm models, performing successful multiplex system research [30]. The applications of machine learning algorithms, such as the RF model, under different experimental conditions confirmed the great potential of incorporating machine learning algorithms into the biogeochemical model, which reveal the cycling mechanisms of representative soil nutrients in detail, such as the study of N and P elements here. These studies provided new insights and approaches to ecological informatics, ecological analysis, soil nutrient conservation, and related plant studies.

5. Conclusions

Herein, with the characterization of the relationship between v, α, and β and SN, TN, NH_4-N, NO_3-N, SP, and TP concentrations in soil surface runoff, subsurface runoff, and fissure runoff, respectively, as a model, the proposed RF regression algorithm was successful in generating the main factors of different features within a complex nonlinear relationship. Furthermore, the relationship between the main feature and soil N and P loss was identified using the nonlinear multifactor analysis. As expected, this study proposed these relationships, each of which had significant importance for soil N and P loss, providing a new avenue and promising potential for soil conservation and the related soil and plant studies, indicating that v, α, and β presented combined effects on soil N and P loss, and should be considered as the key parameters when developing a universal soil nutrient loss model in karst trough valleys. Importantly, these results will not only contribute to the understanding of the mechanisms of nutrient exports via surface runoff and subsurface leakage in the karst region worldwide, but also provide new insights into the main factors in different features and new avenues for soil conservation and the related soil and plant studies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/f14102109/s1, Figures S1–S5: Characteristic relationships between the SP concentration, NH_4-N, NO_3-N, TN, and TP and v, α, and β in soil surface runoff, subsurface runoff, and fissure runoff, respectively; Figure S6: Characterizations of the Importance factor to estimate the effect of every feature, including the v, α, and β, on the TN, NH_4-N, NO_3-N, SP, and TP concentration in the soil surface runoff, subsurface runoff, and fissure runoff, respectively; Figures S7–S10: Characterizations of the relationship between the SP, NH_4-N, NO_3-N, and TN concentration in the soil surface runoff (A–C), subsurface runoff (D–F), and fissure runoff (G–I) and the v, α, and β, respectively. Figure S11: Characterizations of the relationship between the TP concentration in the soil surface runoff (A–C), subsurface runoff (D–F), and fissures runoff (G–I) and the v, α, and β, respectively.

Author Contributions: Conceptualization, Y.Z., T.L. and B.H.; methodology, Y.Z.; software, Y.Z.; validation, Y.Z. and R.Z.; formal analysis, Y.Z. and R.Z.; investigation, Y.Z., L.S. and R.Z.; resources, Y.Z. and R.Z.; data curation, Y.Z. and R.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, T.L. and B.H.; visualization, T.L. and B.H.; supervision, T.L. and B.H.; project administration, T.L. and B.H.; funding acquisition, B.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financially supported by the National Key Research and Development Program of China (2022YFF1302903), the Innovation Research 2035 Pilot Plan of Southwest University (SWU-XDZD22003), and the Fundamental Research Funds for the Central Universities (SWU-KT22060).

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: We thank Xianchun Zou and Hongbin Hu (College of Computer and Information Science, Southwest University) for computer technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, Z.; Zhao, M.; Yan, Z.; Yang, Y.; Niklas, K.J.; Huang, H.; Mipam, T.D.; He, X.; Hu, H.; Wright, S.J. Global patterns and predictors of soil microbial biomass carbon, nitrogen, and phosphorus in terrestrial ecosystems. *Catena* **2022**, *211*, 106037. [CrossRef]
2. Mu, H.; Fu, S.; Liu, B.; Yu, B.; Wang, A. Influence of soil and water conservation measures on soil fertility in the Beijing mountain area. *Environ. Monit. Assess.* **2018**, *190*, 504–515. [CrossRef] [PubMed]
3. Tittonell, P.; Scopel, E.; Andrieu, N.; Posthumus, H.; Mapfumo, P.; Corbeels, M.; van Halsema, G.E.; Lahmar, R.; Lugandu, S.; Rakotoarisoa, J.; et al. Agroecology-based aggradation-conservation agriculture (ABACO): Targeting innovations to combat soil degradation and food insecurity in semi-arid Africa. *Field Crop. Res.* **2012**, *132*, 168–174. [CrossRef]
4. Bochet, E.; Garcia-Fayos, P. Factors controlling vegetation establishment and water erosion on motorway slopes in Valencia, Spain. *Restor. Ecol.* **2004**, *12*, 166e174. [CrossRef]
5. Wildemeersch, J.C.; Garba, M.; Sabiou, M.; Sleutel, S.; Cornelis, W. The effect of water and soil conservation (WSC) on the soil chemical, biological, and physical quality of a Plinthosol in Niger. *Land Degrad. Dev.* **2015**, *26*, 773–783. [CrossRef]
6. Zhao, X.; Wu, P.; Gao, X.; Persaud, N. Soil quality indicators in relation to land use and topography in a small catchment on the loess plateau of China. *Land Degrad. Dev.* **2015**, *26*, 54–61. [CrossRef]
7. Lucon, T.N.; Costa, A.T.; Galvão, P.; Leite, M.G.P. Cadmium behavior in a karst environment hydrological cycle. *Environ. Sci. Pollut. Res. Int.* **2020**, *27*, 8965–8979. [CrossRef]
8. Chen, Z.; Auler, A.S.; Bakalowicz, M.; Drew, D.; Griger, F.; Hartmann, J.; Veni, G. The world karst aquifer mapping project: Concept, mapping procedure and map of Europe. *Hydrogeol. J.* **2017**, *25*, 771–785. [CrossRef]
9. Guo, T.; Wang, Q.; Li, D. Sediment and solute transport on soil slope under simultaneous influence of rainfall impact and scouring flow. *Hydrol. Process.* **2010**, *24*, 1446–1454. [CrossRef]
10. Rouhipour, H.; Ghadiri, H.; Rose, C.W. Investigation of the interaction between flow-driven and rainfall-driven erosion processes. *Aust. J. Soil Res.* **2006**, *44*, 503–514. [CrossRef]
11. Tormo, J.; Bochet, E.; García-Fayos, P. Roadfill revegetation in semiarid mediterranean environments. Part II: Topsoiling, species selection, and hydroseeding. *Restor. Ecol.* **2007**, *15*, 97–102. [CrossRef]
12. Zheng, F.L.; Huang, C.H.; Norton, L.D. Vertical hydraulic gradient and run-on water and sediment effects on erosion processes and sediment regimes. *Soil Sci. Soc. Am.* **2000**, *64*, 4–11. [CrossRef]
13. Ching, J.; Phoon, K.K. Constructing Site-Specific Multivariate Probability Distribution Model Using Bayesian Machine Learning. *J. Eng. Mech.* **2019**, *145*, 04018126. [CrossRef]
14. Zhou, Y.; Li, S.; Zhou, C.; Luo, H. Intelligent Approach Based on Random Forest for Safety Risk Prediction of Deep Foundation Pit in Subway Stations. *J. Comput. Civil. Eng.* **2019**, *33*, 05018004. [CrossRef]
15. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
16. Rodriguez-Galiano, V.; Mendes, M.P.; Garciasoldado, M.J.; Chicaolmo, M.; Ribeiro, L. Predictive Modeling of Groundwater Nitrate Pollution Using Random Forest and Multisource Variables Related to Intrinsic and Specific Vulnerability: A Case Study in an Agricultural Setting (Southern Spain). *Sci. Total Environ.* **2014**, *476*, 189–206. [CrossRef] [PubMed]
17. Rodriguez-Galiano, V.; Sanchez-Castillo MChica-Olmo, M.; Chica-Rivas, M. Machine Learning Predictive Models for Mineral Prospectivity: An Evaluation of Neural Networks, Random Forest, Regression Trees and Support Vector Machines. *Ore Geol. Rev.* **2015**, *71*, 804–818. [CrossRef]
18. Wang, F.; Wang, Y.X.; Zhang, K.; Hu, M.; Weng, Q.; Zhang, H.C. Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation. *Environ. Res.* **2021**, *202*, 111660. [CrossRef]
19. Singh, B.; Sihag, P.; Singh, K. Modelling of impact of water quality on infiltration rate of soil by random forest regression. *Model. Earth Syst. Environ.* **2017**, *3*, 999–1004. [CrossRef]
20. Li, T.Y.; He, B.H.; Chen, Z.P.; Zhang, Y.; Liang, C. Effects of gravel on con-centrated flow hydraulics and erosion in simulated landslide deposits. *Catena* **2017**, *156*, 197–204. [CrossRef]
21. Zeng, R.C.; Zhang, Y.Q.; He, B.H.; Li, T.Y.; Zeng, C. Hydraulics characteristics of concentration flow relative to angle between rock and slope in the karst trough valley area. *Acta Pedol. Sin.* **2021**, *58*, 1179–1189.
22. *GB/T 7479-1987*; Water Quality—Determination of Ammonium—Nessler's Reagent Colorimetric Method. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 1987.

23. *GB/T 7480-1987*; Water Quality—Determination of Nitrate—Spectrophotometric Method with Phenol Disulfonic Acid. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 1987.
24. *GB/T 11894-1989*; Water Quality—Determination of Total Nitrogen—Alkaline Potassium Persulfate Digestion-UV Spectro Photo Metric Method. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 1989.
25. *GB/T 11893-1989*; Water Quality—Determination of Total Phosphorus—Ammonium Molybdate Spectrophotometric Method. Ministry of Environmental Protection of the People's Republic of China: Beijing, China, 1989.
26. Jin, X.L.; Diao, W.Y.; Xiao, C.H.; Wang, F.Y.; Chen, B.; Wang, K.R.; Li, S.K. Estimation of wheat agronomic parameters using new spectral indices. *PLoS ONE* **2013**, *8*, e72736. [CrossRef] [PubMed]
27. Puissant, A.; Rougier, S.; Stumpf, A. Object-oriented mapping of urban trees using random forest classifiers. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 235–245. [CrossRef]
28. Nitze, I.; Barrett, B.; Cawkwell, F. Temporal optimization of image acquisition for land cover classification with random forest and MODIS time-series. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 136–146.
29. Mo, W.B.; Yang, N.; Zhao, Y.L.; Xu, Z.G. Impacts of land use patterns on river water quality: The case of Dongjiang Lake Basin, China. *Ecol. Inform.* **2003**, *75*, 102083. [CrossRef]
30. Prasad, P.; Loveson, V.J.; Chandra, P.; Kotha, M. Evaluation and comparison of the earth observing sensors in land cover/land use studies using machine learning algorithms. *Ecol. Inform.* **2022**, *68*, 101522. [CrossRef]