land

MDPI

Article

Multi-Feature Estimation Approach for Soil Nitrogen Content in Caohai Wetland Based on Diverse Data Sources

Zhuo Dong [1], Yu Zhang [2], Guanglai Zhu [1], Tianjiao Luo [1], Xin Yao [3], Yongxiang Fan [4] and Chaoyong Shen [1,5,6,*]

[1] College of Forestry, Guizhou University, Guiyang 550025, China
[2] College of Resources and Environmental Engineering, Guizhou University, Guiyang 550025, China
[3] The Third Institute of Surveying and Mapping of Guizhou Province, Guiyang 550004, China
[4] Jihua Laboratory, Foshan 528200, China
[5] Guizhou Provincial Key Laboratory of Intelligent Processing and Application of Remote Sensing Big Data, Guiyang 550004, China
[6] The Natural Resources Technical Information Center of Guizhou Province, Guiyang 550004, China
* Correspondence: whimfast@163.com

Abstract

Nitrogen (N) is a key nutrient for sustaining ecosystem productivity and agricultural sustainability; however, achieving high-precision monitoring in wetlands with highly heterogeneous surface types remains challenging. This study focuses on Caohai, a representative karst plateau wetland in China, and integrates Sentinel-2 multispectral and Zhuhai-1 hyperspectral remote sensing data to develop a soil nitrogen inversion model based on spectral indices, texture features, and their integrated combinations. A comparison of four machine learning models (RF, SVM, PLSR, and BPNN) demonstrates that the SVM model, incorporating Zhuhai-1 hyperspectral data with combined spectral and texture features, yields the highest inversion accuracy. Incorporating land-use type as an auxiliary variable further enhanced the stability and generalization capability of the model. The study reveals the spatial enrichment of soil nitrogen content along the wetland margins of Caohai, where remote sensing inversion results show significantly higher nitrogen levels compared to surrounding areas, highlighting the distinctive role of wetland ecosystems in nutrient accumulation. Using Caohai Wetland on the Chinese karst plateau as a case study, this research validates the applicability of integrating spectral and texture features in complex wetland environments and provides a valuable reference for soil nutrient monitoring in similar ecosystems.

Keywords: soil nitrogen; remote sensing; spectral indices; texture features; support vector machine; karst plateau wetland

check for updates

Academic Editor: Ligang Xu

Received: 20 August 2025
Revised: 24 September 2025
Accepted: 27 September 2025
Published: 29 September 2025

Citation: Dong, Z.; Zhang, Y.; Zhu, G.; Luo, T.; Yao, X.; Fan, Y.; Shen, C. Multi-Feature Estimation Approach for Soil Nitrogen Content in Caohai Wetland Based on Diverse Data Sources. *Land* **2025**, *14*, 1967. https://doi.org/10.3390/land14101967

Copyright: © 2025 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Against the backdrop of intensified climate change, ecosystem degradation, and growing pressure on agricultural resources, precise management of soil nutrients has emerged as a critical issue in global ecological and sustainable development research [1]. Soil nitrogen (N) is a fundamental nutrient for ecosystem functioning and agricultural productivity, regulating plant growth, nutrient cycling, and ecosystem stability [2]. Nitrogen imbalance has triggered eutrophication, soil degradation, and declines in ecosystem functioning, which have become global environmental concerns, with particularly pronounced impacts in ecologically sensitive regions such as wetlands [2]. As transitional zones between terrestrial and aquatic ecosystems, wetlands play a crucial role in nitrogen retention, water

purification, and biodiversity conservation. Their high biomass, complex hydrological conditions, and geomorphic features (e.g., low-lying terrain and micro-topographic variations) promote nutrient accumulation but simultaneously pose challenges for soil nutrient monitoring. Therefore, achieving efficient and non-destructive monitoring of soil nitrogen spatial distribution has become a critical task for supporting sustainable land use and global ecosystem management [3]. Traditional laboratory analytical methods (e.g., Kjeldahl digestion) provide high accuracy, but their high sampling costs and limited spatiotemporal coverage constrain their application for large-scale and dynamic monitoring of wetlands.

Remote sensing, with its efficiency, non-invasiveness, and high spatiotemporal resolution, has emerged as an ideal tool for estimating soil nutrients [4]. Multispectral remote sensing indirectly reflects soil nutrient status through vegetation indices (e.g., NDVI, SAVI) and has been widely applied in cropland and grassland ecosystems [5]. Hyperspectral remote sensing, with its continuous bands and high spectral resolution, enables more precise detection of nitrogen-related spectral features (e.g., red-edge and near-infrared bands), thereby significantly improving estimation accuracy [6–8]. In addition, the integration of hyperspectral lidar (HSLidar) technology provides a new perspective for further improving the accuracy of soil nitrogen estimation, particularly in wetland environments with complex terrain and strong heterogeneity, effectively complementing the limitations of traditional remote sensing data [9,10]. However, in complex environments such as wetlands, remote sensing signals are easily affected by vegetation cover, moisture interference, and surface heterogeneity, resulting in unstable estimation accuracy when relying solely on spectral features [11]. Therefore, it is necessary to integrate additional complementary variables and optimization strategies to enhance the reliability and adaptability of remote sensing inversion.

To overcome the limitations of purely spectral variables in complex landscapes, image texture features have been incorporated into remote sensing estimation frameworks due to their ability to represent spatial structural heterogeneity. Using approaches such as the gray-level co-occurrence matrix (GLCM), texture features can capture gray-level variations among pixels and have achieved notable results in applications such as soil classification and vegetation identification [12]. Particularly in plateau wetlands with pronounced micro-topographic variations, texture information enhances the model's ability to capture heterogeneous surface patterns. Moreover, land-use type, as a key indicator of surface functional zoning, strengthens the model's responsiveness to ecological functional heterogeneity, thereby reflecting differences in soil nutrient accumulation mechanisms [13]. However, current research has primarily focused on a single scale or specific extraction methods, with limited systematic evaluation of the synergistic effects of multiple remote sensing features [14].

Although existing remote sensing inversion studies have achieved substantial progress in structurally homogeneous ecosystems with limited disturbances, systematic evaluations of modeling strategies remain insufficient in plateau wetlands dominated by karst landforms, where fragmented terrain and complex remote sensing responses prevail [15]. Located in the heart of the Yunnan–Guizhou Plateau in southwestern China, the Weining Caohai Wetland is one of the largest and most representative karst plateau wetland systems in the country, characterized by plateau geomorphology, land–water interlacing, human disturbances, and ecological transitions [16]. Such an extremely heterogeneous ecosystem provides a critical testbed for evaluating the adaptability, transferability, and robustness of remote sensing approaches. Therefore, this study takes the Weining Caohai Wetland as the research object and proposes a remote sensing-based method for estimating soil nitrogen tailored to karst plateau wetlands. It aims to address the following: In karst plateau wetlands characterized by fragmented terrain and high surface heterogeneity,

such as Weining Caohai, can the multi-feature approach—integrating spectral features, image texture features, and land-use types—improve the accuracy and stability of soil nitrogen content estimation compared to the single spectral feature-based method, and what constitutes the optimal feature combination? Additionally, can the multi-feature fusion strategy tailored for karst plateau wetlands simultaneously enhance the model's adaptability to heterogeneous sub-regions of the wetland (e.g., land–water transition zones, microtopographic gradient zones) to achieve reliable spatial inversion of soil nitrogen across the entire wetland, while also possessing cross-platform applicability?

2. Materials and Data

2.1. Study Area Overview

As a globally significant plateau ecological region, the Yunnan–Guizhou Plateau is of great importance for ecological response studies under global change, owing to its complex geomorphology and diverse land-use patterns. The study area is located in Weining County, Guizhou Province, southwestern China, in the core of the Yunnan–Guizhou Plateau, and represents a typical karst plateau wetland ecosystem. The region has an average elevation of approximately 2200 m with relatively gentle terrain, a subtropical plateau monsoon climate, an average annual precipitation of about 1100 mm, and a mean annual temperature of around 10.5 °C. Land use in the study area is diverse, including grassland, cropland, wetlands, and built-up land, making it highly typical and representative. Soil sampling was conducted in November 2022 under clear weather conditions without precipitation, yielding a total of 40 soil samples. Sampling sites were arranged according to land-use distribution, covering the major soil types and land-use categories within the region, thereby ensuring strong spatial representativeness and ecological heterogeneity. The study area and the spatial distribution of sampling sites are shown in Figure 1.

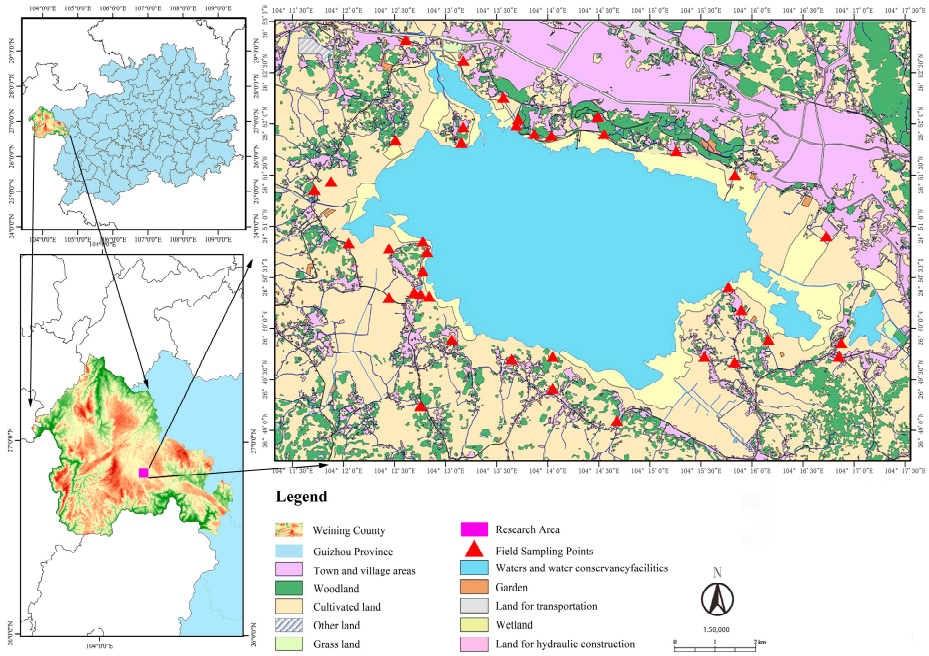


Figure 1. Study area and field sampling sites (coordinate system: WGS 84).

2.2. Soil Sample Collection and Analysis

The 40 soil samples in this study were collected from 40 plots at a depth of 10 cm, with GPS coordinates recorded for each sampling site. All samples were collected in November 2022 under clear weather conditions without precipitation, specifically between 8:00 and 17:00. To improve representativeness, a 3 m × 3 m plot was established at each site, and

one composite sample was obtained by mixing subsamples collected using the five-point method. After collection, all soil samples were air-dried at room temperature (25 °C) to remove impurities and coarse aggregates. The samples were then passed through a 2 mm sieve to remove coarse particles, ensuring consistency and homogeneity. Total soil nitrogen content was subsequently determined using the Kjeldahl digestion method [17,18]. The statistical distribution of nitrogen content for all samples is shown in Table 1.

Table 1. Soil nitrogen content statistical distribution table.

Statistic	Soil Nitrogen Content (Nitrogen)
Sample Size (*n*)	40
Mean	17.55
Standard Deviation (SD)	5.04
Standard Error (SE)	0.80
Minimum (Min)	5.27
Maximum (Max)	28.72
25% Quartile (Q1)	14.22
Median (50%)	17.76
75% Quartile (Q3)	20.59
Skewness	0.20
Kurtosis	−0.14

2.3. Remote Sensing Data and Preprocessing

This study employed Sentinel-2 multispectral imagery and Zhuhai-1 hyperspectral imagery for remote sensing inversion of soil nitrogen content in the Caohai region. Detailed information on these two types of remote sensing data is provided in Table 2.

Table 2. Detailed information on Sentinel-2 and Zhuhai-1.

Satellite	Sentinel-2	Zhuhai-1 HSI
Spatial Resolution	10 m (Multispectral)	10 m (Hyperspectral)
Coverage	290 km	150 km
Revisit Cycle	10 days	6 days
Number of Bands	13 (4 bands used in this study)	13
Wavelength Range	443–945 nm (490–842 nm used)	443–940 nm
Image Acquisition Date	2 November 2022	2 November 2022

Sentinel-2, launched by the European Space Agency (ESA), provides high-resolution multispectral imagery with a short revisit cycle and rich spectral information, and has been widely applied in land resource, ecological, and agricultural monitoring. In this study, four bands at 10 m spatial resolution were selected—blue (B2: 490 nm), green (B3: 560 nm), red (B4: 655 nm), and near-infrared (B8: 842 nm)—for soil spectral feature extraction. Zhuhai-1, an independently developed hyperspectral remote sensing satellite in China, offers high spectral resolution and frequent observation capability, making it suitable for fine-scale object identification and soil nutrient monitoring. In this study, 32 continuous bands covering the 400–1000 nm range with a spatial resolution of 10 m were utilized.

Both types of imagery were preprocessed using ENVI software (5.6), including radiometric calibration, atmospheric correction, and orthorectification. Radiometric calibration was performed with ENVI's Radiometric Calibration tool, atmospheric correction was conducted using the FLAASH module to minimize atmospheric effects, and orthorectification was carried out with ENVI's Orthorectification tool to ensure spatial accuracy. After preprocessing, the imagery was clipped according to the vector boundary of the study area and precisely aligned with field sampling sites using ENVI's georeferencing function, ensuring one-to-one correspondence between spectral data and soil measurements, thereby providing high-quality inputs for subsequent spectral index and texture feature extraction.

3. Methods

Taking the Weining Caohai Wetland as the study area, this research integrates Sentinel-2 multispectral data with Zhuhai-1 hyperspectral data to develop a soil nitrogen estimation method that combines spectral indices, texture features, and land-use types. Optimal features were selected through Pearson correlation analysis to construct multi-feature combination models, and the performance of four machine learning algorithms—random forest (RF), support vector machine (SVM), partial least squares regression (PLSR), and backpropagation neural network (BPNN)—was compared to improve estimation accuracy in complex wetland environments. The innovations of this study include: (1) comparing remote sensing data from different platforms and integrating spectral indices with texture features to enhance the representation of spatial heterogeneity in karst wetlands; (2) incorporating land-use type to optimize model stability and account for the effects of ecological functions and human activities; and (3) validating the applicability of multi-feature integration methods in karst ecosystems using the Caohai Wetland as a case study. This study not only provides an applied demonstration of soil nutrient monitoring in complex ecosystems but also offers a technical framework for remote sensing estimation methods in similar regions. Soil nitrogen inversion models were established separately using Sentinel-2 and Zhuhai-1 data, and the differences between their results were compared. The workflow is illustrated in Figure 2.

3.1. Pearson Correlation Analysis

To quantify the linear relationship between remote sensing features and soil nitrogen (N) content, the Pearson correlation coefficient was applied to pre-screen spectral and texture variables [19]. This coefficient reflects the degree of linear dependence between variables and is commonly used for dimensionality reduction and feature selection in high-dimensional remote sensing data. The Pearson correlation coefficient r is calculated as follows:

$$r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}} \quad (1)$$

In Equation (1), x_i represents the remote sensing feature value of the i-th sample, y_i denotes the measured soil nutrient content, $\overline{x}$ and $\overline{y}$ are their respective means, and n is the total number of samples.

The coefficient ranges from $[-1, 1]$: $r > 0$ indicates a positive correlation, $r < 0$ indicates a negative correlation, and values closer to ± 1 reflect stronger linear relationships. In this study, a two-tailed test ($p < 0.05$) was used to assess statistical significance, and only significantly correlated features were retained for subsequent modeling.

3.2. Construction and Selection of Spectral Indices

To enhance the spectral sensitivity of imagery to soil nitrogen, seven classical vegetation indices—GNDVI [20,21], OSAVI [22,23], NDVI [24,25], SAVI [24,26], DVI [22,23], EVI [22,23], and RVI [23,27]—were constructed using Sentinel-2 multispectral and Zhuhai-1 hyperspectral data. These indices, derived from specific combinations of band reflectance, strengthen the mapping relationship between soil nutrients and spectral features, thereby improving inversion capability. For Sentinel-2, four bands at 10 m resolution were selected—blue (490 nm), green (560 nm), red (665 nm), and near-infrared (842 nm)—primarily based on geometric registration accuracy and the band requirements of commonly used vegetation indices. Compared with the 20 m and 60 m bands, these bands provide clearer spectral responses to nitrogen, and their effectiveness has been validated in previous studies [5,6]. This selection also ensures uniform spatial resolution, reduces mixed-pixel effects, and enhances feature robustness and cross-platform applicability.

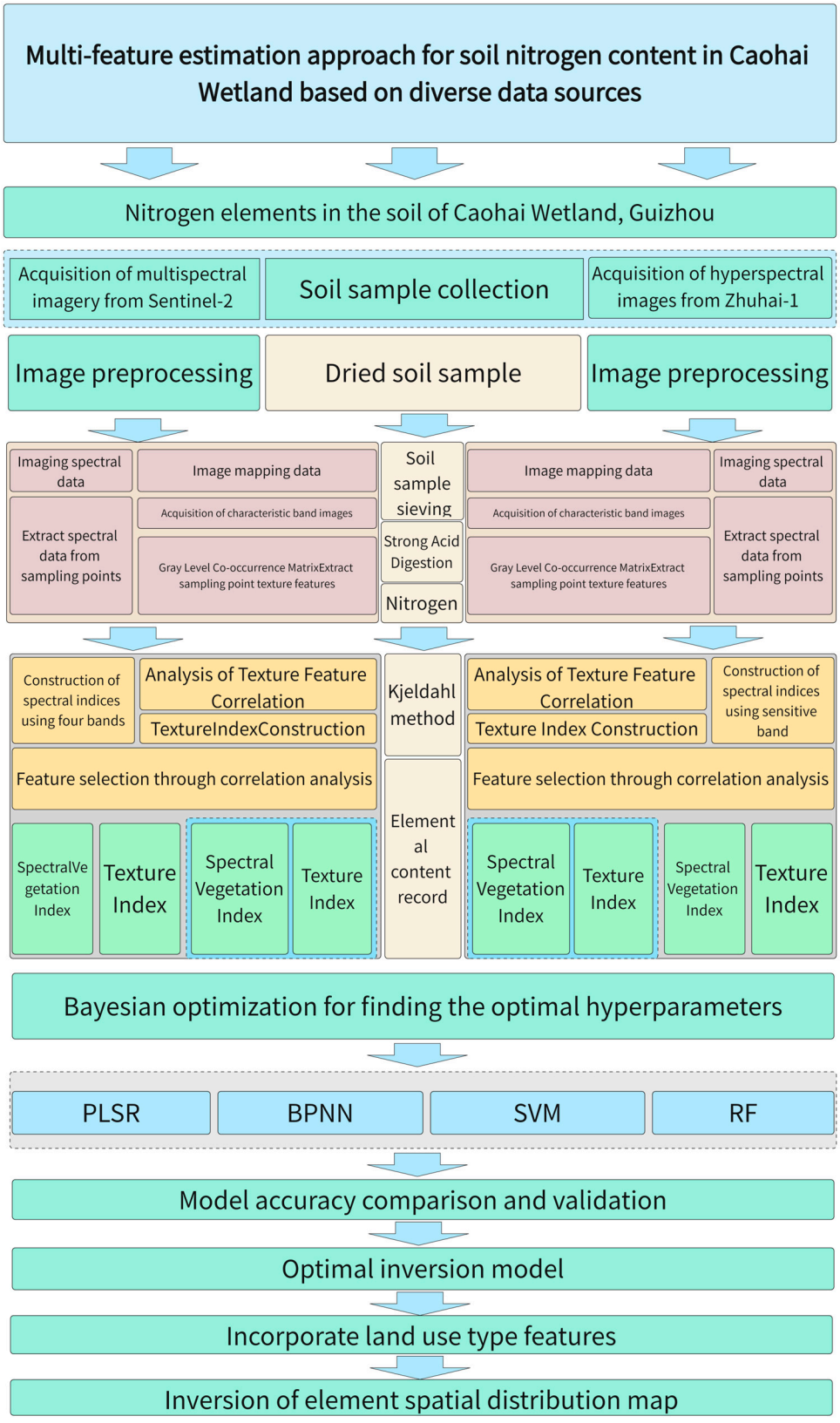


Figure 2. Workflow for multi-source remote sensing estimation of soil nitrogen content.

Considering the differences in band configuration and spectral resolution between the two platforms, band combinations were adjusted accordingly to maintain consistency in index definitions. For Zhuhai-1, the advantage of continuous hyperspectral bands and nitrogen-sensitive red-edge and near-infrared regions was leveraged, and index band com-

binations were optimized based on Pearson correlation analysis. After index construction, a two-tailed Pearson test ($p < 0.05$) was applied to select the indices most significantly correlated with the target variable, which were then used as model inputs.

3.3. Texture Feature Extraction and Index Construction

Texture features, by quantifying the gray-level relationships among image pixels, effectively characterize surface spatial heterogeneity and structural variation. They serve as an important complement to spectral features and hold significant application value in soil nitrogen (N) content inversion [28]. The extraction of texture features and construction of texture indices in this study involved two steps: first, multiple basic texture parameters were derived from the gray-level co-occurrence matrix (GLCM); second, composite texture indices were generated following the principles of vegetation index construction to enhance feature sensitivity to nutrient variation.

3.3.1. Extraction of Basic Texture Features (GLCM Features)

In this study, the gray-level co-occurrence matrix (GLCM) is used to extract texture features from the image. Specifically, a 3 × 3 pixel window is applied to calculate the gray-level relationships between each pixel and its 8 neighboring pixels in the image. The GLCM quantifies the spatial relationship between pixel gray values to characterize texture information. The process of constructing the GLCM involves the following steps:

Determining the relative positions of pixel pairs: For each reference pixel, a 3 × 3 pixel window is used to consider the gray-level relationships between the pixel and the 8 neighboring pixels. These neighboring pixel gray values are calculated in four directions (0°, 45°, 90°, and 135°).

Calculating the GLCM: For each direction, the frequency of occurrence of each gray value pair is calculated to construct the GLCM. The specific formula is:

$$P(i,j) = \sum_{k=1}^{N} \sum_{l=1}^{N} \delta(i,j,k,l) \tag{2}$$

where the term $P(i,j)$ represents the frequency of occurrence of pixel pairs with gray values i and j in the gray-level co-occurrence matrix.

The function $\delta(i,j,k,l)$ is an indicator function, which takes the value 1 when the gray values of pixels k and l are i and j, respectively, and 0 otherwise.

Extracting texture features: Using the GLCM, several basic texture features are extracted from the image, including mean (MEA), variance (VAR), homogeneity (HOM), contrast (CON), dissimilarity (DIS), entropy (ENT), second moment (SEC), and correlation (COR). These features effectively describe the spatial structure and gray-level distribution of the image.

Through this process, 32 basic texture features (4 bands × 8 features) were extracted from the four Sentinel-2 multispectral bands and Zhuhai-1 hyperspectral imagery, which serve as input variables for subsequent texture index construction.

3.3.2. Construction and Selection of Texture Indices

To enhance the correlation between texture information and nitrogen content, and to examine linear combinations of texture features, this study followed previous research [29] and applied the principles of vegetation index construction. Based on individual texture features, three texture indices (TI) were developed using the normalized difference spectral index (NDSI), difference spectral index (DSI), and ratio spectral index (RSI) approaches, namely the Normalized Difference Texture Index (NDTI), Ratio Texture Index (RTI), and Difference Texture Index (DTI). These three texture indices were constructed by pairing texture

features in all possible two-by-two combinations, thereby enhancing feature discrimination capability. The detailed combinations are shown in Table 3.

Table 3. Formulas for constructing texture indices.

Index Type	Formula Construction	Source
NDTI	$\frac{(T_1-T_2)}{(T_1+T_2)}$	[29]
RTI	$T_1 - T_2$	[29]
DTI	$\frac{T_1}{T_2}$	[29]

$T1$ and $T2$ represent any two texture features, with $T1 \neq T2$. All texture indices were subjected to Pearson correlation analysis with measured soil nitrogen content, and the single most optimal texture index for nitrogen was selected from each platform as a modeling input.

3.4. Feature Integration and Variable Construction

To systematically evaluate the impact of different remote sensing features on the inversion accuracy of soil nitrogen content, three basic input variable combinations were constructed: spectral indices only (VI), texture indices only (TI), and a combination of spectral and texture indices (VI + TI). Building on this, to further explore the contribution of spatial semantic information to model performance, land-use type (LU) was introduced as an auxiliary variable into the optimal feature combination model, forming the "VI + TI + LU" combination, and its effect on improving modeling accuracy was evaluated. Land-use type was obtained through manual interpretation, covering major categories such as cropland, grassland, water bodies, and built-up land, and was incorporated into the model as a categorical variable. This analysis aimed to investigate the potential regulatory role of regional land-use patterns in shaping the spatial distribution of soil nutrients.

3.5. Regression Models and Inversion

To achieve remote sensing inversion of soil nitrogen content, four commonly used regression models were selected in this study: Partial Least Squares Regression (PLSR), Backpropagation Neural Network (BPNN), Support Vector Machine (SVM), and Random Forest (RF). These models are widely applied in remote sensing inversion and soil parameter estimation, each with distinct modeling mechanisms and applicable domains. By comparatively analyzing modeling performance under different feature combinations, this study evaluated the accuracy and stability of each model in soil nitrogen inversion [30].

3.5.1. Partial Least Squares Regression (PLSR)

Partial Least Squares Regression (PLSR) is a classical linear regression method, particularly suitable for situations where strong multicollinearity exists among input variables. PLSR projects the original variables onto a new set of latent variables, integrating dimensionality reduction with regression, thereby effectively reducing the interference of inter-variable correlations on the modeling results. In soil property inversion tasks using hyperspectral data, PLSR has demonstrated strong performance, particularly in analyzing soil quality where large numbers of variables and complex datasets are involved. Previous studies have shown that PLSR effectively captures spectral variations of surface materials and links them to soil properties, and it has been widely applied in remote sensing estimation of soil nitrogen content [31].

3.5.2. Backpropagation Neural Network (BPNN)

The Backpropagation Neural Network (BPNN) is an artificial neural network based on a multilayer feedforward structure, which learns complex nonlinear relationships between inputs and outputs by adjusting connection weights. BPNN has shown excellent performance in various remote sensing modeling tasks, particularly in estimating soil

heavy metal content and retrieving water parameters, demonstrating strong adaptability. With its outstanding nonlinear modeling capability, BPNN can effectively address complex problems such as soil nitrogen estimation, thereby significantly improving prediction accuracy. Previous studies have demonstrated that the application of BPNN in soil nutrient modeling can effectively improve inversion accuracy and has been widely employed in remote sensing estimation of soil nitrogen content [32].

3.5.3. Support Vector Machine Regression (SVM)

Support Vector Machine Regression (SVM) is a nonlinear regression method based on the principle of structural risk minimization. By mapping input features into a high-dimensional space through kernel functions, SVM overcomes the limitations of linear regression methods in handling high-dimensional problems. Owing to its excellent generalization ability, SVM is particularly well suited for modeling small-sample and high-dimensional datasets. SVM has been widely applied in remote sensing inversion studies of soil nitrogen and phosphorus, effectively handling complex nonlinear relationships in soils. Studies have shown that SVM provides more accurate predictions of soil properties, demonstrating superior modeling capability, particularly with high-dimensional datasets [33].

3.5.4. Random Forest (RF)

Random Forest (RF) is a non-parametric model based on an ensemble learning strategy, which makes predictions by aggregating the outputs of multiple decision trees. RF has demonstrated excellent stability and robustness in soil parameter estimation and environmental remote sensing inversion. Its core concept is to randomly select features and samples to construct multiple decision trees, and then derive the final prediction through voting or averaging. RF is particularly well suited for high-dimensional inputs and situations involving complex interactions among variables. In soil nitrogen estimation, RF effectively addresses the high dimensionality of the feature space and automatically manages multicollinearity among input variables. Studies have shown that RF provides relatively accurate estimates in the remote sensing inversion of various soil properties, with particularly strong performance in predicting soil fertility and nutrient content [34].

3.6. Validation

In this study, the coefficient of determination (R^2), root mean square error (*RMSE*), and residual predictive deviation (*RPD*) were employed to evaluate model accuracy. To enhance the robustness and generalization of the model results, model training and testing were conducted using a 10-fold cross-validation strategy. The specific formulas are as follows [35]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

$$RPD = \frac{SD}{RMSE} \tag{5}$$

In Equations (1)–(3), y_i denotes the observed value, $\hat{y}_i$ represents the predicted value, $\overline{y}$ is the mean of the observed values, SD is the standard deviation of the observed values, and n is the total number of samples.

4. Results and Analysis

4.1. Correlation Analysis Between Soil Nitrogen and Remote Sensing Features

4.1.1. Correlation Analysis Between Soil Nitrogen and Multispectral Vegetation Indices

Based on the constructed spectral indices, Pearson correlation analysis was used to calculate the linear correlation coefficients between seven Sentinel-2 vegetation indices and soil nitrogen content. The indices used and their calculation formulas are detailed in Table 4.

Table 4. Construction formulas of Sentinel-2 spectral vegetation indices.

Index	Formula	Source
GNDVI	$\frac{R_{NIR} - R_{green}}{R_{NIR} + R_{green}}$	[20,21]
OSAVI	$\frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red} + 0.16}$	[22,23]
NDVI	$\frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}}$	[24,25]
SAVI	$1.5 \times \frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red} + 0.5}$	[24,26]
DVI	$R_{NIR} - R_{red}$	[22,23]
EVI	$2.5 \times \frac{R_{NIR} - R_{red}}{R_{NIR} + 6 \times R_{red} - 7.5 \times R_{blue} + 1}$	[22,23]
RVI	$\frac{R_{NIR}}{R_{red}}$	[23,27]

First, the vegetation indices in Table 4 are calculated using the reflectance of different bands, aiming to assess vegetation growth status and environmental impacts. In the statistical analysis shown in Figure 3, EVI (enhanced vegetation index) performs the best. EVI combines data from the red, blue, and near-infrared bands, effectively reducing atmospheric interference and soil background influences. As a result, it shows strong correlation and significance in soil nitrogen estimation. The Pearson correlation coefficient for EVI exceeds 0.4, indicating a strong positive correlation with soil nitrogen compared to other indices. Moreover, as shown in Figure 3, its statistical significance (log10(p) value) is also high, reaching 0.01, which indicates that this index is statistically significant and suitable for remote sensing monitoring in the current ecological environment.

Following closely, GNDVI (Green–Red Vegetation Index) is another prominent index. GNDVI uses the ratio of the green band to the near-infrared band to reflect the photosynthetic capacity and moisture status of vegetation. Its Pearson correlation coefficient is also close to 0.4, indicating a positive correlation with soil nitrogen, and it demonstrates high statistical significance. This suggests that GNDVI is effective in reflecting the spatial variation of soil nitrogen content and is a powerful tool for soil nitrogen monitoring.

SAVI (Soil Adjusted Vegetation Index) performs relatively weakly in this study. Although SAVI helps reduce soil background interference, its correlation with soil nitrogen is lower, with a Pearson correlation coefficient of 0.32. The log10(p) value exceeds 1.3, and its p-value reaches 0.05 but does not meet the 0.01 threshold, indicating limited statistical significance. As a result, its application in soil nitrogen estimation is less effective.

Next, OSAVI (Optimized Soil Adjusted Vegetation Index) and NDVI (Normalized Difference Vegetation Index) show relatively poor performance. OSAVI adjusts for soil background to improve estimation accuracy but has a lower Pearson correlation coefficient of 0.31 in this study, with low statistical significance, and its p-value does not even reach 0.05. NDVI, a classic vegetation index, despite being widely used, performs fairly poorly in this study. It has a Pearson correlation coefficient of 0.27, similar to DVI (Difference

Vegetation Index), and its log10(p) value does not exceed 1.1, indicating weak correlation with soil nitrogen, thus limiting its effectiveness in soil nitrogen estimation.

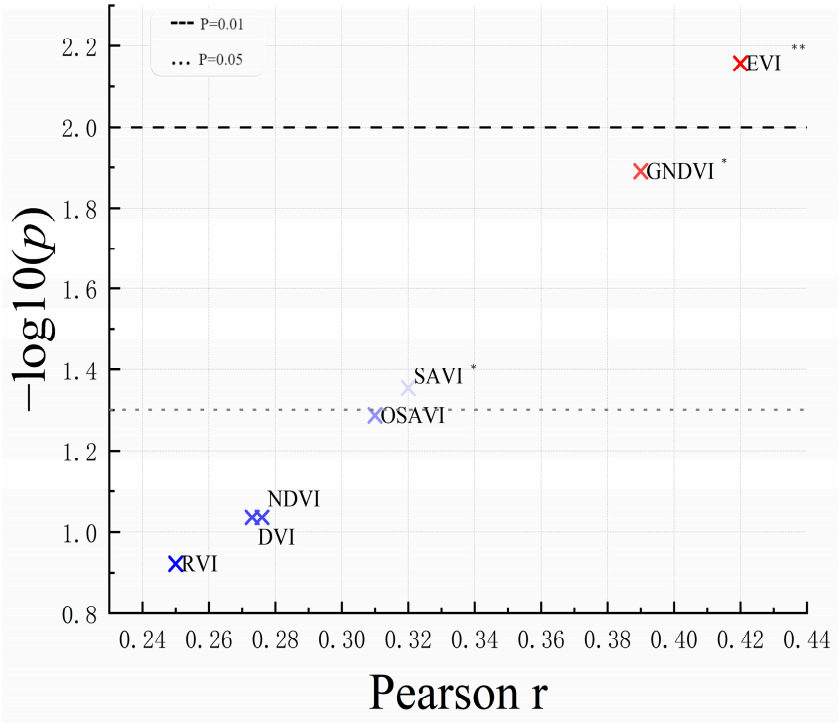


Figure 3. Pearson correlation between Sentinel-2 vegetation indices and nitrogen content (* $p < 0.05$ ** $p < 0.01$).

Finally, RVI (Ratio Vegetation Index) performs the worst. Its correlation with soil nitrogen is very low, with a Pearson correlation coefficient of 0.25 and extremely low statistical significance, making it nearly ineffective in providing reliable soil nitrogen predictions. Overall, none of the vegetation indices had a correlation coefficient exceeding 0.5 with nitrogen content, indicating that their explanatory power for nitrogen variation remains limited. In this study, EVI was selected as the optimal vegetation index for nitrogen inversion using the Sentinel-2 platform.

4.1.2. Correlation Analysis Between Soil Nitrogen and Hyperspectral Vegetation Indices

Based on the Pearson correlation coefficients between nitrogen content and Zhuhai-1 reflectance shown in Figure 4, this study found varying degrees of correlation between soil nitrogen content and multiple hyperspectral bands. In particular, bands around the 730 nm region exhibited stronger correlations with soil nitrogen content. Moreover, correlation peaks were observed in the 500–600 nm and 700–800 nm ranges, which is consistent with findings from previous studies [36]. These bands exhibit strong responses in capturing soil nitrogen characteristics and provide robust support for hyperspectral remote sensing inversion of soil nitrogen content, indicating that hyperspectral data enable more precise identification of nitrogen-related spectral features. Based on the correlation analysis between 32 bands and nitrogen content, and in consideration of spectroscopic mechanisms, this study optimized the band combinations for vegetation indices. Specifically, bands with higher correlations were selected and aligned with the spectroscopic characteristics of nitrogen to ensure a more accurate reflection of soil nitrogen variations. The optimized band selections are presented in Table 5, and Pearson correlation analyses between the indices and nitrogen content are illustrated in Figure 5.

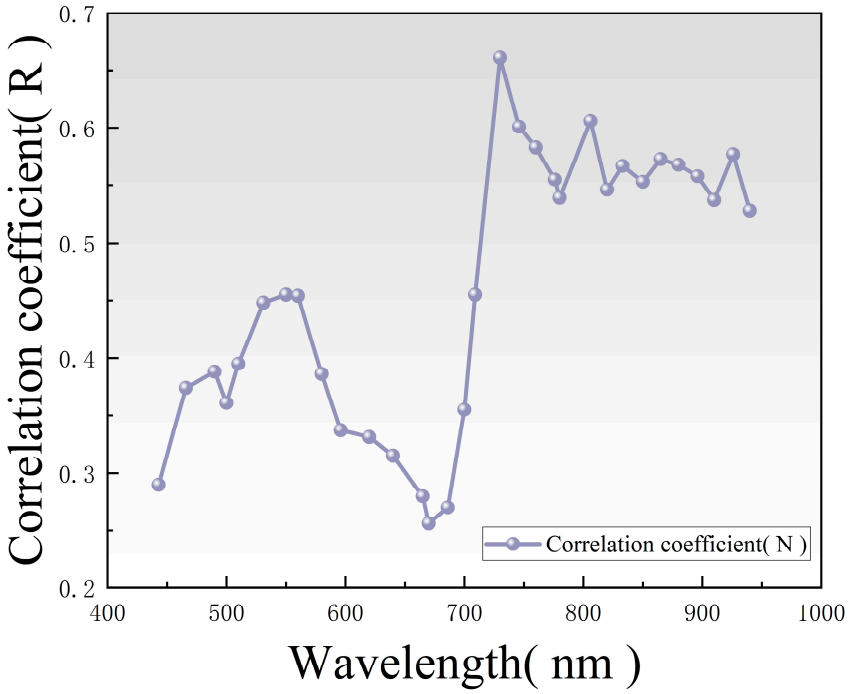


Figure 4. Correlation between Zhuhai-1 reflectance and nitrogen content.

Table 5. Formulas of spectral vegetation indices for nitrogen inversion using Zhuhai-1 data.

Index	Definition (N)	Source
GNDVI	$\frac{R_{730} - R_{550}}{R_{730} + R_{550}}$	[20,21]
OSAVI	$\frac{R_{730} - R_{670}}{R_{730} + R_{670} + 0.16}$	[22,23]
NDVI	$\frac{R_{730} - R_{670}}{R_{730} + R_{670}}$	[24,25]
SAVI	$1.5 \times \frac{R_{730} - R_{670}}{R_{730} + R_{670} + 0.5}$	[24,26]
DVI	$R_{730} - R_{670}$	[22,23]
EVI	$2.5 \times \frac{R_{730} - R_{670}}{R_{730} + 6 \times R_{670} - 7.5 \times R_{490} + 1}$	[22,23]
RVI	$\frac{R_{730}}{R_{670}}$	[23,27]

The optimization of the band combinations for vegetation indices in Table 5 is based on spectral mechanisms, the correlation analysis of soil nitrogen content, and the rationality of the formula structure. Through the correlation analysis of different bands with soil nitrogen content, we optimized the selection of vegetation index bands by combining the spectral characteristics of the bands and their responses to nitrogen content variations.

For ratio-based indices (such as GNDVI, NDVI, and RVI), we selected bands with higher correlations to nitrogen content changes. GNDVI uses the 730 nm and 550 nm bands, where the 550 nm band is in the green light region and shows a higher correlation with soil nitrogen content, while the 730 nm band is in the near-infrared region and has the highest correlation with nitrogen. This band effectively reflects the plant growth state. By combining these two bands, GNDVI can better capture soil nitrogen content changes, particularly in areas with active vegetation growth.

NDVI uses the 730 nm and 670 nm bands. The 730 nm band, located in the near-infrared region, responds most strongly to vegetation growth and nitrogen content changes,

while the 670 nm band, located in the red light region, has the lowest correlation with soil nitrogen content. Although the reflectance of the 670 nm band has a weaker correlation with nitrogen content, it still enhances sensitivity to nitrogen content changes through its difference in reflectance with the 730 nm band, especially in reflecting the relationship between vegetation health and soil nitrogen content.

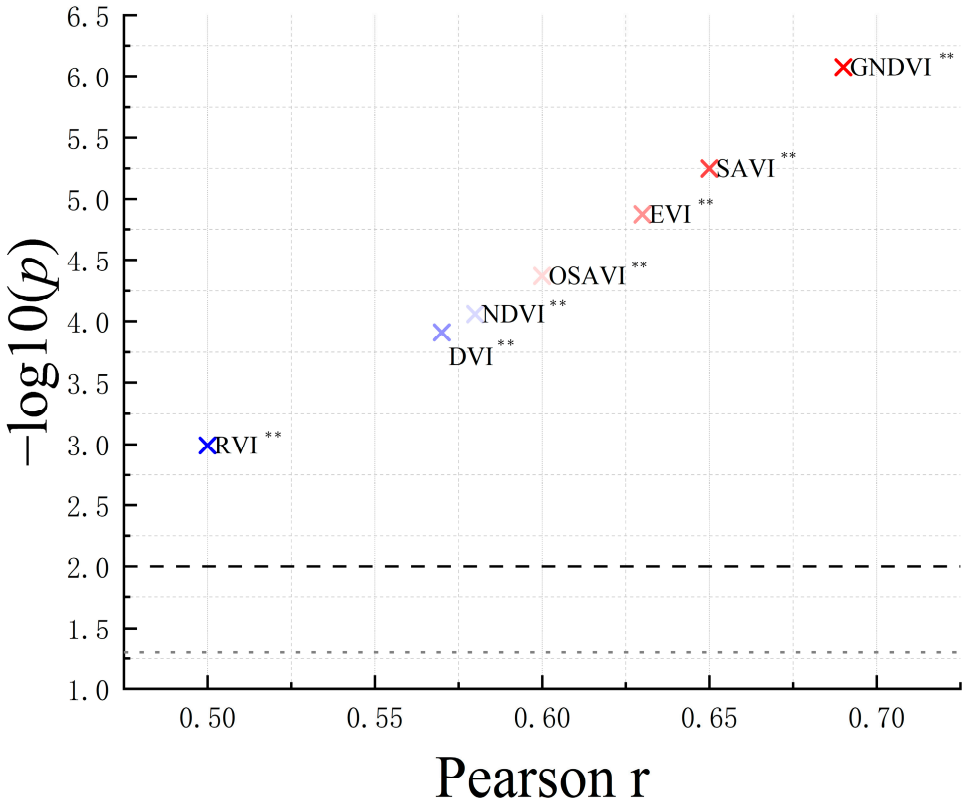


Figure 5. Pearson correlation between Zhuhai-1 vegetation indices and nitrogen content (** $p < 0.01$).

RVI selects the 730 nm and 670 nm bands, where the reflectance ratio between these two bands effectively reflects soil nitrogen content, especially in areas with exposed soil or sparse vegetation. The combination of the 730 nm and 670 nm bands, through the ratio calculation, directly reflects changes in soil and vegetation, thereby improving nitrogen content estimation accuracy.

For difference-based indices, DVI uses the reflectance difference between the 730 nm and 670 nm bands. By calculating this stark difference between these two bands, we can directly reflect soil nitrogen content changes, especially in low vegetation cover or arid regions. The difference-based index is particularly effective at capturing nitrogen content changes in such areas.

Regulation-based indices (such as SAVI and OSAVI) introduce constants to adjust for soil background or atmospheric interference, enhancing the response to vegetation and reducing the impact of external factors. SAVI uses the 730 nm and 670 nm bands and adjusts with a constant of 0.5. The reflectance difference between these two bands is highly sensitive to changes in vegetation growth and nitrogen content, effectively reducing soil background interference. OSAVI also uses the 730 nm and 670 nm bands and introduces a constant of 0.16 to further adjust for soil background effects, particularly in regions with complex soil backgrounds, allowing for more accurate reflection of nitrogen content changes.

Finally, the enhanced vegetation index (EVI) combines the 730 nm, 670 nm, and 490 nm bands. The 490 nm band, located in the blue light region, has the highest correlation with soil nitrogen content in the blue light range, so it was chosen to enhance sensitivity to

nitrogen content changes. The selection of the 730 nm and 670 nm bands is also to enhance the index's response to nitrogen. Although the 670 nm band has a lower correlation with nitrogen content, the 730 nm band is the one that responds most strongly to changes in soil nitrogen content. The inclusion of the 490 nm band helps reduce atmospheric interference, and through the weighted combination of multiple bands, further improves the accuracy of soil nitrogen content estimation.

Through this optimization of band combinations, this study ensures that each vegetation index can maximize the accuracy of soil nitrogen content estimation. The selection of bands is not only based on the high correlation of the bands but also considers the formula structure of each index to make them more suitable for soil nitrogen content detection. By optimizing different types of vegetation indices, we ensure that we can more accurately reflect changes in soil nitrogen content and reduce the influence of external interference factors.

Figure 5 shows the Pearson correlation coefficients between different vegetation indices and soil nitrogen content. GNDVI exhibits the highest correlation with nitrogen content (r = 0.69), indicating that this index is the most effective in capturing nitrogen variations. SAVI (r = 0.65) and EVI (r = 0.63) follow closely, suggesting that they also have strong responsiveness to nitrogen variation. OSAVI (r = 0.58), NDVI (r = 0.57), and DVI (r = 0.57) show lower correlations; although their performance in nitrogen inversion is weaker than that of GNDVI and NDVI, they still retain a certain explanatory capacity. RVI (r = 0.50) demonstrates the weakest correlation, indicating a relatively limited sensitivity to changes in soil nitrogen content.

Overall, GNDVI was identified as the optimal vegetation index for nitrogen inversion using the Zhuhai-1 remote sensing platform.

4.1.3. Correlation Analysis Between Soil Nitrogen and Texture Indices

Eight texture features were extracted from four Sentinel-2 multispectral bands using the gray-level co-occurrence matrix, and their correlations with nitrogen content were then established, as shown in Figure 6. For convenience, the notation XY is used to represent the texture feature Y extracted from parameter X, such as B2mea indicating the mean (mea) texture feature derived from band B2.

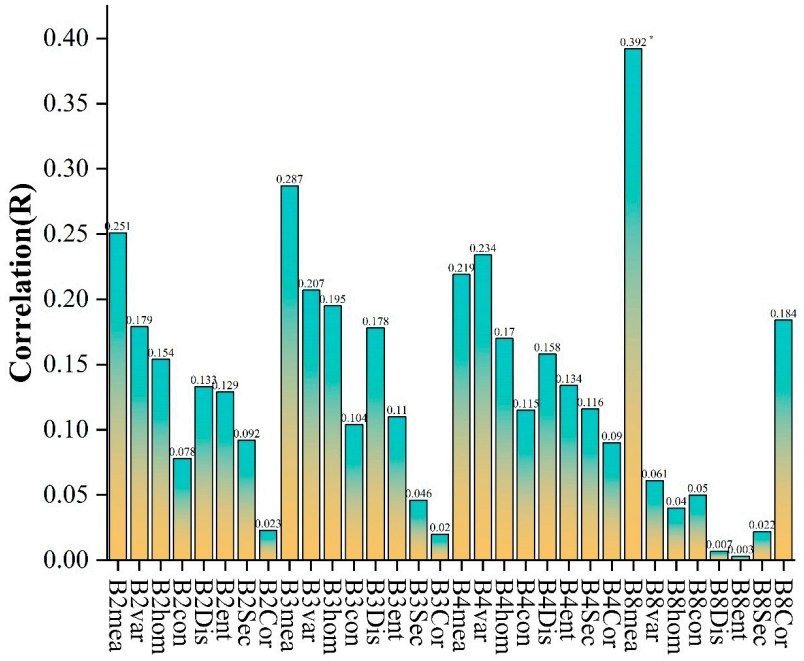


Figure 6. Pearson correlation between Sentinel-2 texture features and nitrogen content (* $p < 0.05$).

As shown in Figure 6, the sensitivity of the eight texture features extracted from different bands to nitrogen content varies. The correlation between B8mea and nitrogen content passed the significance test at $p < 0.05$, with a correlation coefficient of 0.39, which is the highest among the features. The remaining texture features show no significant correlations with nitrogen content, as their correlation coefficients did not pass the significance test.

As shown in Figure 4, the 730 nm, 806 nm, 746 nm, and 760 nm bands exhibit the highest correlations with nitrogen; therefore, eight texture features were extracted from these bands, and Pearson correlation analyses with nitrogen content were performed. The results are presented in Figure 7. Figure 7 illustrates the correlations between texture features extracted from four hyperspectral nitrogen-sensitive bands of Zhuhai-1 and nitrogen content. The results indicate that 730dis, 806hom, 746hom, and 746sec passed the significance test at $p < 0.05$, while 730mea, 730hom, 730sec, 806mea, 746mea, and 760mea passed at $p < 0.01$. Among them, 730mea shows the highest correlation with nitrogen, with a coefficient of 0.49.

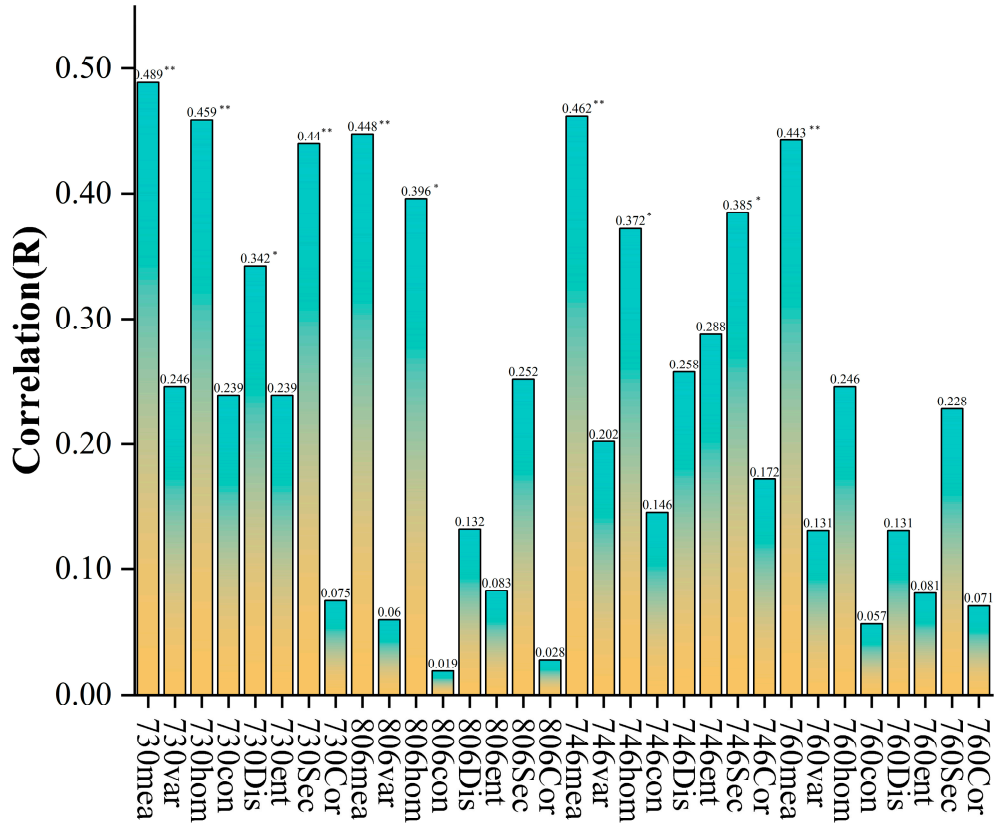


Figure 7. Pearson correlation between Zhuhai-1 texture features and nitrogen content (* $p < 0.05$ ** $p < 0.01$).

Based on Figures 6 and 7, individual texture features show varying degrees of correlation with soil nitrogen, but their overall sensitivity is low. To address the weak correlations of single texture features with soil nitrogen, this study combined them in pairs to construct texture indices. The resulting correlations are presented in Table 5.

As shown in Table 6, the three optimal combined texture indices selected from the Sentinel-2 platform for nitrogen all passed the significance test at $p < 0.01$, indicating highly significant correlations. Among them, DTI (B4var, B8var) shows the strongest correlation with nitrogen content, with a coefficient of 0.65. Overall, DTI (B4var, B8var) was identified as the optimal texture index for nitrogen inversion using the Sentinel-2 platform.

Table 6. Correlation between Sentinel-2 texture indices and nitrogen. (** $p < 0.01$).

Element	Texture Index	Selected Texture T1	T2	Correlation Coefficient	Significance
Nitrogen	NDTI	B2Cor	B4Sec	0.565729	**
	RTI	B2Hom	B8Hom	0.565164	**
	DTI	B4Var	B8Var	0.648657	**

As shown in Table 6, three optimal composite texture indices for nitrogen estimation using the Zhuhai-1 platform all passed the $p < 0.01$ significance test. These indices exhibited highly significant correlations. Among them, RTI (730Hom, 760Sec) showed the highest correlation with nitrogen content, reaching 0.78. This study selected RTI (730Hom, 760Sec) as the optimal texture index for nitrogen inversion on the Zhuhai-1 platform.

The results in Tables 6 and 7 demonstrate that texture indices derived from the combination of multiple texture features exhibit stronger correlations with soil nitrogen. This indicates that incorporating texture indices can improve the inversion accuracy of soil nitrogen.

Table 7. Correlation of Zhuhai-1 texture indices with nitrogen content. (** $p < 0.01$).

Element	Texture Index	Selected Texture T1	T2	Correlation Coefficient	Significance
Nitrogen	NDTI	730Hom	760Sec	0.7822	**
	RTI	730Hom	760Sec	0.78396	**
	DTI	730Dis	760Dis	0.760255	**

4.2. Inversion Accuracy Metrics

Based on the optimal features identified in Section 4.1, three feature combinations were constructed: spectral indices alone (VI), texture indices alone (TI), and a combined set of spectral and texture indices (VI + TI). Using Sentinel-2 and Zhuhai-1 data, the performance of four machine learning models—Random Forest (RF), Support Vector Machine (SVM), Partial Least Squares Regression (PLSR), and Back Propagation Neural Network (BPNN)—was compared for soil nitrogen estimation. Model performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and ratio of performance to deviation (RPD). Ten-fold cross-validation was employed to ensure the robustness of the results.

Data source differences

Figure 8a–d compares the modeling performance of different data sources in nitrogen inversion. Using the combined feature set (VI + TI) as an example, Zhuhai-1 hyperspectral data consistently outperformed Sentinel-2 multispectral data across all models. In the SVM model, the R^2 of hyperspectral data reached 0.66, representing a 106% improvement over the multispectral value of 0.32. RMSE decreased from 4.38 to 2.89 (−34.0%), while RPD increased from 1.27 to 1.79 (+40.9%). Similar trends were observed for RF, BPNN, and PLSR models, with R^2 improvements of 113%, 100%, and 96.7%, respectively. RMSE reductions were around 25%, and RPD values all exceeded 1.6. Apart from the combined features, most single-feature combinations (VI or TI) also exhibited the same trend, although the magnitude of improvement was smaller. Overall, the results demonstrate that under the same feature inputs, hyperspectral data provide higher modeling accuracy than multispectral data, with superior performance in goodness-of-fit, error control, and prediction stability.

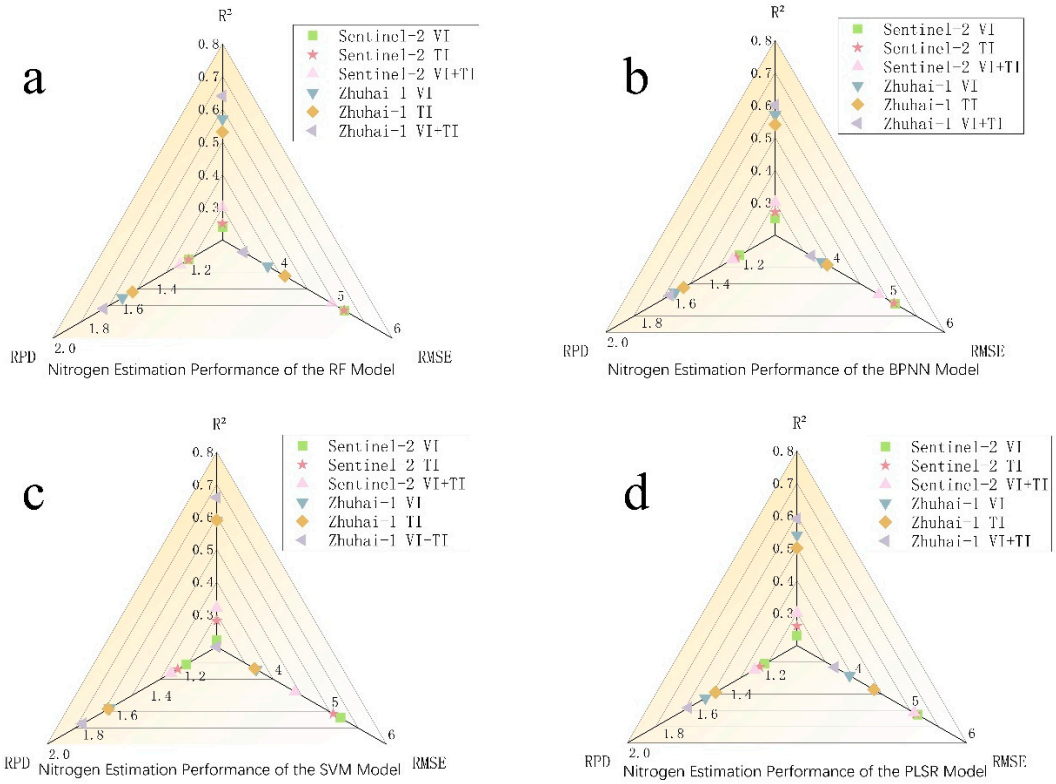


Figure 8. (**a**) RF model; (**b**) BPNN model; (**c**) SVM model; (**d**) PLSR model. Comparison of four machine learning models (RF, SVM, PLSR, and BPNN) for nitrogen inversion.

Differences in feature combinations

Across the four regression models for nitrogen estimation, the combination of spectral indices and texture features (VI + TI) significantly outperformed the use of spectral indices alone (VI), achieving higher goodness-of-fit and more stable predictive performance. Across the four regression models for nitrogen estimation, the combination of spectral indices and texture features (VI + TI) significantly outperformed the use of spectral indices alone (VI), achieving higher goodness-of-fit and more stable predictive performance. In the RF model, feature fusion also improved performance: hyperspectral R^2 increased from 0.57 to 0.64, with RMSE decreasing from 3.80 to 3.37. Under multispectral conditions, R^2 improved to 0.30 and RMSE dropped to 4.93. Accuracy improvements for BPNN and PLSR were relatively limited, particularly with hyperspectral data, where R^2 increased only slightly to 0.60 and 0.59, and RMSE reductions were within 0.2. However, under multispectral data, feature fusion still yielded R^2 improvements exceeding 20%. Overall, the results indicate that hyperspectral data achieve higher modeling accuracy when driven by feature fusion, while the performance improvements for multispectral data are even more pronounced, particularly in the SVM and RF models. Overall, the results indicate that hyperspectral data achieve higher modeling accuracy when driven by feature fusion, while the performance improvements for multispectral data are even more pronounced, particularly in the SVM and RF models. By combining the spectral sensitivity to nitrogen status with the texture response to surface heterogeneity, feature fusion demonstrates stronger overall modeling capacity in nitrogen inversion.

Differences in model performance

Different models exhibited distinct performance across data sources and feature combinations. The SVM model showed the strongest overall fitting ability, particularly under the combined feature set (VI + TI), with R^2 improvements of 3–46%, RPD increases of 2–10%,

and RMSE reductions of 6–22%, making it the model with the most notable performance gains. As illustrated in Figure 8c, its indicator points were consistently closer to the R^2 and RPD axes and further from the RMSE apex, reflecting strong accuracy and stability. The RF model (Figure 8a) ranked second, with more concentrated indicator distributions, particularly in the hyperspectral fusion group, demonstrating stable fitting performance. In contrast, the BPNN model (Figure 8b) showed greater variability across different feature combinations, with more scattered indicator points. The PLSR model (Figure 8d) tended to cluster in low-R^2 and low-RPD regions, with some combinations yielding R^2 values of only 0.25–0.28 and RPD values near 1.2, approaching the threshold of practical utility. Overall, the differences in data source, feature combination, and model architecture significantly affected the accuracy of nitrogen inversion. Among all combinations, the Zhuhai-1 hyperspectral fusion group (VI + TI) within the SVM model (Figure 8c) performed best, with a test-set R^2 of 0.66, RMSE of 2.89, and RPD of 1.79. In summary, the Zhuhai-1 hyperspectral + VI + TI + SVM combination achieved the highest fitting accuracy and optimal error control in soil nitrogen inversion, making it the best-performing modeling approach in this study.

4.3. Effect of Land-Use Type on Optimizing the Best-Performing Soil Nitrogen Inversion Model

The karst plateau wetlands are characterized by fragmented terrain and complex land-use distributions, resulting in pronounced spatial heterogeneity within the region. Influenced by topography, vegetation, and human disturbances, different land-use types exhibit marked differences in soil physicochemical properties, moisture conditions, and nutrient dynamics. These differences not only affect the processes of nitrogen migration and accumulation on the surface but also lead to structural inconsistencies in the relationships between remote sensing features and soil nitrogen content across land-use types. Under unified modeling conditions, mixing samples from different land-use types can introduce structural errors into the model and reduce fitting accuracy. To address this issue, this study incorporated land-use stratification during model construction, aiming to reduce the interference caused by spatial heterogeneity, enhance model stability and adaptability, and further evaluate the effectiveness of this strategy in nitrogen inversion within karst plateau wetlands.

Figure 9 compares the fitting performance of the SVM model constructed with Zhuhai-1 spectral and texture features for nitrogen prediction. Panel (a) shows the scatter plot before incorporating land-use optimization, while panel (b) presents the results after optimization. In Figure 9a, the model achieved an R^2 of 0.66, an RMSE of 2.89, and an RPD of 1.79. Although both training and testing samples were generally distributed around the 1:1 line, considerable dispersion was observed. In the high-nitrogen range, most points fell below the 1:1 line, indicating an underestimation of measured values, whereas in the low-nitrogen range, some points were located above the line, reflecting slight overestimation. The fitted line lay below the 1:1 line in the medium-to-high nitrogen range, with the gap widening as nitrogen content increased, suggesting limited model accuracy in this interval.

After incorporating land-use type as an additional feature (Figure 9b), the model's performance improved, with R^2 increasing to 0.75, RMSE decreasing to 2.28, and RPD rising to 2.11. The scatter points of both the training and testing sets became more concentrated around the 1:1 line, with reduced vertical deviation. The underestimation observed in the high-nitrogen range was notably alleviated, while the distribution of low-nitrogen samples aligned more closely with the 1:1 line, reducing the number of overestimated points. The fitted line exhibited a slope closer to 1 than in panel (a), and its deviation from the 1:1 line narrowed across different nitrogen levels, indicating enhanced consistency of predictions. Furthermore, the distribution patterns of the training and testing sets became more aligned,

reflecting improved model stability under varying sample conditions. Compared with panel (a), the model in panel (b) demonstrated gains in both fitting accuracy and stability, providing a more reliable basis for subsequent spatial analysis of nitrogen distribution.

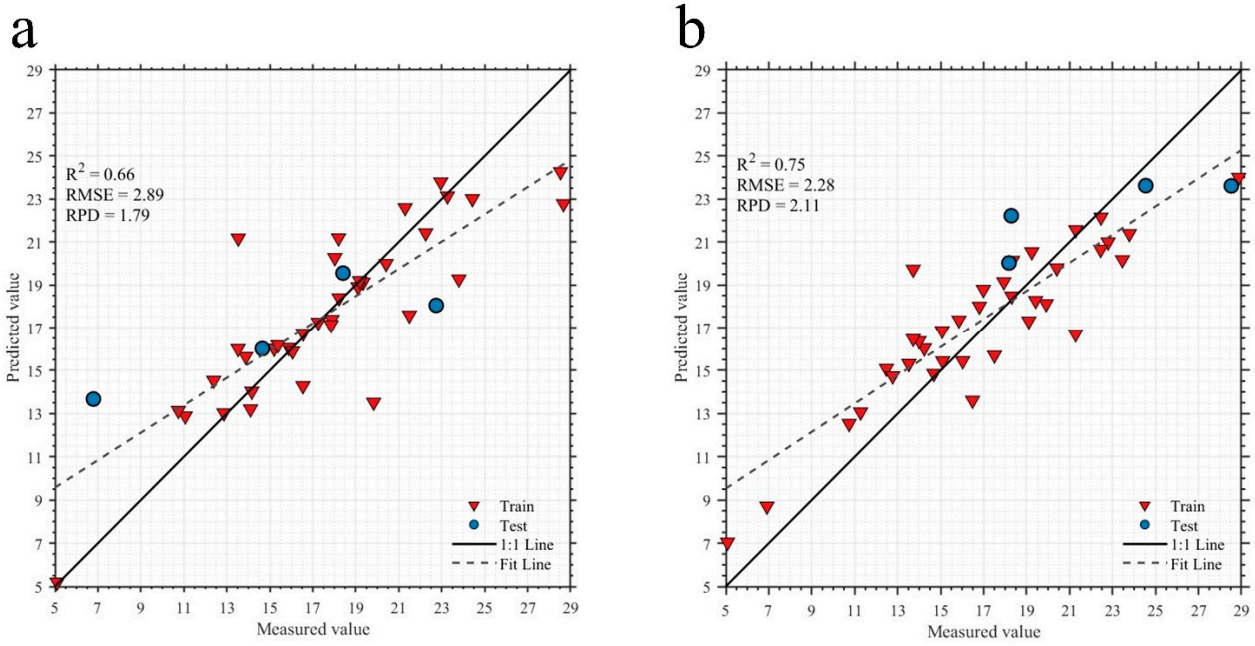


Figure 9. Scatter plots of nitrogen inversion using Zhuhai-1 {VI(Vegetation Index) + TI(Texture Index)} SVM model. (**a**) Scatter plot of nitrogen inversion before incorporating land-use type; (**b**) scatter plot of nitrogen inversion after incorporating land-use type.

4.4. Nitrogen Content Inversion in the Study Area

Figure 10 illustrates the spatial distribution of soil nitrogen content in the study area. A pronounced high-value belt is observed along the edge of Caohai Lake, dominated by red and orange colors, corresponding to nitrogen contents mostly above 20 g/kg, which are markedly higher than those in the adjacent outer areas shown in green and yellow (primarily below 15 g/kg). Outside the lake, medium- and low-value zones are interspersed with medium- and high-value areas, forming a mosaic pattern of green patches and red–orange patches in the spatial distribution.

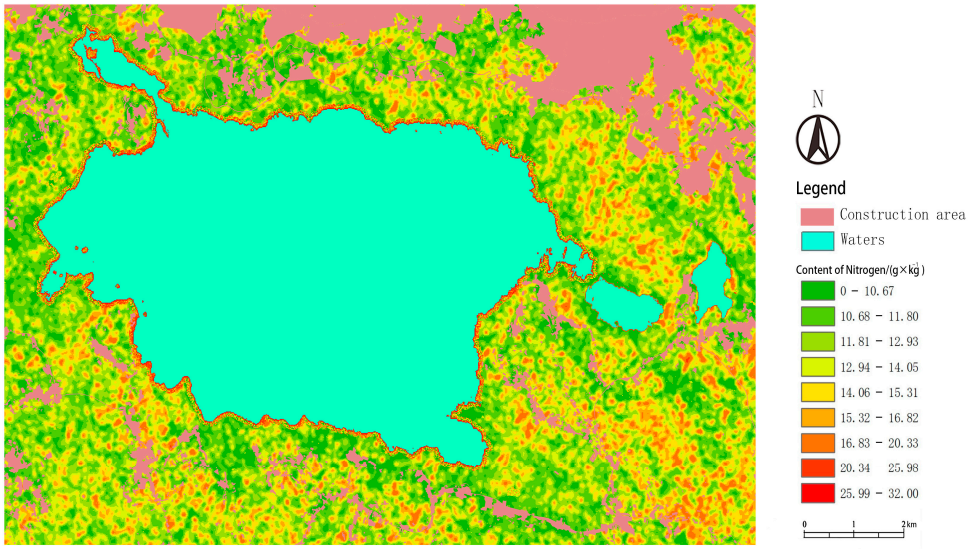


Figure 10. Spatial inversion of soil nitrogen content in study area.

5. Discussion

Wetlands play a critical role in the global nutrient cycle. Their unique hydrological regimes, high biological productivity, and substantial organic matter accumulation make them favorable sinks for nitrogen retention [37,38]. Plant residue decomposition and root exudates contribute to increasing soil organic matter, thereby enhancing its capacity for nutrient adsorption and retention. At the same time, the reductive environment of wetland soils effectively suppresses nitrogen denitrification, reducing the risk of nutrient migration and loss [39,40]. Remote sensing inversion results revealed that soil nitrogen content along the edges of the Caohai wetland was markedly higher than in adjacent non-wetland areas, exhibiting a pronounced spatial clustering pattern. This distribution is likely driven by the combined effects of hydrological processes, microtopographic structures, and vegetation distribution: dense and vigorous root systems promote a dynamic cycle of nutrient uptake and release, while low-lying terrain facilitates surface nutrient retention, thereby enhancing soil nutrient-holding capacity. Such spatial characteristics highlight the unique role of wetland edge ecosystems in nutrient accumulation, provide explanatory support for regional differences in remote sensing models, and underscore the need to account for the ecological regulatory functions of wetland types when designing inversion strategies.

Differences in spectral resolution across remote sensing data sources are among the key factors influencing the accuracy of soil nitrogen inversion. Multispectral remote sensing is widely applied due to its easy accessibility and broad utility in surface information extraction. However, its limited number of bands and fixed central wavelengths restrict its ability to capture the subtle spectral responses of complex soils. In contrast, hyperspectral remote sensing, with its higher resolution and continuous spectral coverage, can more precisely resolve spectral variations of surface materials, thereby demonstrating greater sensitivity and discrimination capacity in soil nutrient monitoring [41]. This study compared nitrogen inversion performance under combined spectral and texture features using Sentinel-2 multispectral and Zhuhai-1 hyperspectral data. The results showed that, for the SVM model, the R^2 values reached 0.32 and 0.66, with RMSE values of 4.38 and 2.89 and RPD values of 1.27 and 1.79, respectively. These metrics were all markedly superior to those of the multispectral model, confirming the advantage of hyperspectral data in terms of both modeling adaptability and accuracy [42]. It is worth emphasizing that, given the inherent heterogeneity of remote sensing data sources, the inclusion of Sentinel-2 data in this study was not intended merely for accuracy comparison, but to evaluate the robustness of the constructed feature variables under reduced spectral dimensionality and band-limited conditions. The feature construction approach, based on correlation screening and reconstruction of sensitive bands, is independent of fixed band combinations and designed to capture the spectral mechanisms of nitrogen responses. If this strategy performs effectively within the simplified band system of Sentinel-2, it would demonstrate adaptability to different spectral structures and transferability across platforms, thereby reducing the risk of overfitting to environmental features under hyperspectral conditions. This framework strengthens model robustness and provides methodological guidance for cross-platform applications in soil nitrogen inversion [43].

The dimensionality and representational capacity of input features in remote sensing inversion models directly determine estimation accuracy. Traditional approaches often rely on a single spectral index or texture feature, which, although providing some predictive capability, are constrained under complex surface conditions by the limitations of single-variable inputs. To enhance model robustness and adaptability, strategies that integrate multiple remote sensing features have increasingly become a research focus, and they have been widely applied, particularly within multi-source remote sensing frameworks [44]. This study systematically compared the modeling performance of spectral indices, texture

indices, and their combined inputs under Sentinel-2 and Zhuhai-1 remote sensing data. The results showed that, under the same modeling approach, models integrating spectral and texture features consistently outperformed those using a single feature type. For example, in nitrogen inversion with the SVM model based on Zhuhai-1 hyperspectral data, R^2 increased from 0.59 with spectral input alone to 0.66 with fused inputs, while RPD rose from 1.63 to 1.79, indicating that feature fusion significantly enhanced model stability and predictive accuracy. This improvement can be attributed to the complementary information provided by spectral and texture features: spectral indices capture indirect responses related to element absorption or vegetation status, while texture indices reflect surface structure and spatial heterogeneity. In regions with high background noise or complex surface conditions, reliance on spectral or texture features alone often leads to overfitting or underfitting. In contrast, feature fusion expands input dimensionality and strengthens the model's ability to learn the coupled spectral–structural mechanisms, thereby improving overall inversion performance [44]. In summary, under the same remote sensing platform, multi-feature combinations significantly improved soil nitrogen inversion performance, confirming the effectiveness and transferability of the feature-fusion modeling strategy. This approach is particularly suitable for typical wetland ecosystems where variable response signals are weak and background disturbances are strong.

It is worth noting that with the continuous advancement of remote sensing technologies, future research could consider incorporating more advanced tools, such as hyperspectral lidar (HSLidar) technology. As a novel remote sensing tool, hyperspectral lidar can collect data day and night and capture hyperspectral data from a three-dimensional perspective, offering unique technical advantages. By acquiring 3D spatial and 3D spectral information, HSLidar can effectively overcome the limitations of traditional remote sensing technologies in complex terrains and environments, especially for soil nitrogen inversion in small regions such as wetlands. The complexity and high heterogeneity of wetland ecosystems often result in challenges such as vegetation cover, moisture interference, and surface heterogeneity in traditional remote sensing data. However, the three-dimensional information provided by hyperspectral lidar can significantly improve inversion accuracy and enhance model robustness. Therefore, integrating hyperspectral lidar technology for soil nitrogen content inversion in the future could provide more precise technical support for wetland management and ecological conservation [9].

Land-use type plays an important regulatory role in the distribution and accumulation of soil nitrogen, as different land-use practices influence soil physicochemical properties, vegetation cover, and hydrological processes, thereby driving nutrient dynamics [44]. Studies have shown that wetlands, due to sufficient water availability, dense vegetation, and organic matter accumulation, tend to retain more nitrogen, whereas agricultural land, influenced by fertilization and tillage disturbances, exhibits greater spatial variability [45]. Building on the integration of spectral and texture features, this study further introduced land-use type as an auxiliary variable to optimize the inversion performance of the SVM model. Using Zhuhai-1 hyperspectral nitrogen inversion as an example, R^2 increased from 0.66 to 0.75, RMSE decreased from 2.89 to 2.28, and RPD rose from 1.79 to 2.11, indicating substantial improvements in model fitting accuracy and stability. This enhancement reflects the important role of land-use type in characterizing surface nutrient accumulation mechanisms and helps capture soil variability arising from different management practices and natural conditions. Although land use does not directly determine nutrient concentrations, the spatial structural information it provides enhances the model's adaptability to complex surface environments, making it particularly suitable for highly heterogeneous wetland ecosystems [46].

6. Conclusions

This study developed soil nitrogen estimation models based on multi-source remote sensing data and compared the inversion performance of multispectral and hyperspectral imagery. The main conclusions are as follows:

Soil nitrogen content along the wetland–water boundaries of Caohai exhibited significant enrichment, higher than in surrounding areas, highlighting the unique role of karst wetlands in nutrient accumulation.

The SVM model combining hyperspectral data with spectral and texture features achieved higher nitrogen estimation accuracy, with R^2 of 0.66 and RMSE of 2.89, outperforming multispectral data.

Incorporating land-use type further improved model accuracy, with R^2 increasing to 0.75, RMSE decreasing to 2.28, and RPD rising to 2.11, demonstrating the complementary value of surface functional information.

Author Contributions: All authors contributed to the conception and design of the study. Data curation, formal analysis, and original draft preparation were performed by Z.D. Conceptualization, funding acquisition, and writing—review and editing were carried out by C.S. Investigation was conducted collaboratively by Y.Z., G.Z., T.L. and X.Y. Y.F. put forward revision suggestions on the revised manuscript, including those related to graphic processing and manuscript writing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Department of Science and Technology Program of Guizhou Province of China (Qiankehe Foundation-ZK (2023) General 196) and the Guizhou Provincial Key Technology R&D Program—Guizhou Provincial Key Laboratory of Intelligent Processing and Application of Remote Sensing Big Data (No. Qiankehe platform [2025]014).

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: All authors sincerely thank the editors and anonymous reviewers for their valuable feedback on this study.

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

References

1. Johnston, A.E.; Poulton, P.R.; Coleman, K. *Soil Organic Matter: Its Importance in Sustainable Agriculture and Carbon Dioxide Fluxes*; Academic Press: San Diego, CA, USA, 2008; Chapter 1.
2. Sulieman, S.; Mühling, K.H. Utilization of soil organic phosphorus as a strategic approach for sustainable agriculture. *J. Plant Nutr. Soil Sci.* **2021**, *184*, 311–319. [CrossRef]
3. Sahabiev, I.; Smirnova, E.; Giniyatullin, K. Spatial prediction of Agrochemical properties on the scale of a single field using machine learning methods based on remote sensing data. *Agronomy* **2021**, *11*, 2266. [CrossRef]
4. Zhai, M. Inversion of organic matter content in wetland soil based on Landsat 8 remote sensing image. *J. Vis. Commun. Image Represent.* **2019**, *64*, 102645. [CrossRef]
5. Li, W.; Dou, Z.; Cui, L.; Wang, R.; Zhao, Z.; Cui, S.; Lei, Y.; Li, J.; Zhao, X.; Zhai, X. Suitability of hyperspectral data for monitoring nitrogen and phosphorus content in constructed wetlands. *Remote Sens. Lett.* **2020**, *11*, 495–504. [CrossRef]
6. Piccoli, F.; Barbato, M.P.; Peracchi, M.; Napoletano, P. Estimation of Soil Characteristics from Multispectral Sentinel-3 Imagery and DEM Derivatives Using Machine Learning. *Sensors* **2023**, *23*, 7876. [CrossRef] [PubMed]
7. Sun, M.; Li, Q.; Jiang, X.; Ye, T.; Li, X.; Niu, B. Estimation of Soil Salt Content and Organic Matter on Arable Land in the Yellow River Delta by Combining UAV Hyperspectral and Landsat-8 Multispectral Imagery. *Sensors* **2022**, *22*, 3990. [CrossRef]
8. Wan, S.; Hou, J.; Zhao, J.; Clarke, N.; Kempenaar, C.; Chen, X. Predicting Soil Organic Matter, Available Nitrogen, Available Phosphorus and Available Potassium in a Black Soil Using a Nearby Hyperspectral Sensor System. *Sensors* **2024**, *24*, 2784. [CrossRef] [PubMed]
9. Du, L.; Jin, Z.; Chen, B.; Chen, B.; Gao, W.; Yang, J.; Shi, S.; Song, S.; Wang, M.; Gong, W.; et al. Application of Hyperspectral LiDAR on 3-D Chlorophyll-Nitrogen Mapping of Rohdea Japonica in Laboratory. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 9667–9679. [CrossRef]

10. Bai, J.; Niu, Z.; Wang, L. A theoretical demonstration on the independence of distance and incidence angle effects for small-footprint hyperspectral LiDAR: Basic physical concepts. *Remote Sens. Environ.* **2024**, *315*, 114452. [CrossRef]
11. Zhang, Y.; Xiao, J.; Yan, K.; Lu, X.; Li, W.; Tian, H.; Wang, L.; Deng, J.; Lan, Y. Advances and Developments in Monitoring and Inversion of the Biochemical Information of Crop Nutrients Based on Hyperspectral Technology. *Agronomy* **2023**, *13*, 2163. [CrossRef]
12. Zhao, W.; Ma, F.; Yu, H.; Li, Z. Inversion Model of Salt Content in Alfalfa-Covered Soil Based on a Combination of UAV Spectral and Texture Information. *Agriculture* **2023**, *13*, 1530. [CrossRef]
13. Hu, P.; Zhang, W.; Kuzyakov, Y.; Xiao, L.; Xiao, D.; Xu, L.; Chen, H.; Zhao, J.; Wang, K. Linking bacterial life strategies with soil organic matter accrual by karst vegetation restoration. *Soil Biol. Biochem.* **2022**, *177*, 108925. [CrossRef]
14. Li, Z.; Zhang, Y.; Lu, J.; Wang, Y.; Zhao, C.; Wang, W.; Wang, J.; Zhang, H.; Huo, Z. Better inversion of rice nitrogen nutrition index at early panicle initiation stage using spectral features, texture features, and wavelet features based on UAV multispectral imagery. *Eur. J. Agron.* **2025**, *168*, 127654. [CrossRef]
15. Guo, C.; Guo, X. Estimating leaf chlorophyll and nitrogen content of wetland emergent plants using hyperspectral data in the visible domain. *Spectrosc. Lett.* **2015**, *49*, 180–187. [CrossRef]
16. Zhou, C.W.; Yang, R.; Yu, L.F.; Zhang, Y.; Yan, L.B. Hydrological and ecological effects of climate change in caohai watershed based on swat model. *Appl. Ecol. Environ. Res.* **2019**, *17*, 161–172. [CrossRef]
17. Viscarra Rossel, R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [CrossRef]
18. Ministry of Environmental Protection of the People's Republic of China. *Soil Quality—Determination of Total Nitrogen—Modified Kjeldahl Method: HJ 717—2014*; China Environmental Science Press: Beijing, China, 2015.
19. Song, Y.-Q.; Zhao, X.; Su, H.-Y.; Li, B.; Hu, Y.-M.; Cui, X.-S. Predicting Spatial Variations in Soil Nutrients with Hyperspectral Remote Sensing at Regional Scale. *Sensors* **2018**, *18*, 3086. [CrossRef] [PubMed]
20. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
21. Li, W.; Wang, K.; Han, G.; Wang, H.; Tan, N.; Yan, Z. Integrated diagnosis and time-series sensitivity evaluation of nutrient deficiencies in medicinal plant (*Ligusticum chuanxiong* Hort.) based on UAV multispectral sensors. *Front. Plant Sci.* **2023**, *13*, 1092610. [CrossRef]
22. Mashaba-Munghemezulu, Z.; Chirima, G.J.; Munghemezulu, C. Modeling the Spatial Distribution of Soil Nitrogen Content at Smallholder Maize Farms Using Machine Learning Regression and Sentinel-2 Data. *Sustainability* **2021**, *13*, 11591. [CrossRef]
23. Van Le, C.; Pham, L.T. Estimating leaf phosphorus concentration in rice by combining vegetation indices, texture features, and water indices from UAV multispectral imagery. *Agrosystems Geosci. Environ.* **2025**, *8*, e70160. [CrossRef]
24. Xu, Y.; Smith, S.E.; Grunwald, S.; Abd-Elrahman, A.; Wani, S.P.; Nair, V.D. Estimating soil total nitrogen in smallholder farm settings using remote sensing spectral indices and regression kriging. *CATENA* **2018**, *163*, 111–122. [CrossRef]
25. Mazur, P.; Gozdowski, D.; Wójcik-Gront, E. Soil Electrical Conductivity and Satellite-Derived Vegetation Indices for Evaluation of Phosphorus, Potassium and Magnesium Content, pH, and Delineation of Within-Field Management Zones. *Agriculture* **2022**, *12*, 883. [CrossRef]
26. Gracia-Romero, A.; Kefauver, S.C.; Vergara-Díaz, O.; Zaman-Allah, M.A.; Prasanna, B.M.; Cairns, J.E.; Araus, J.L. Comparative Performance of Ground vs. Aerially Assessed RGB and Multispectral Indices for Early-Growth Evaluation of Maize Performance under Phosphorus Fertilization. *Front. Plant Sci.* **2017**, *8*, 2004. [CrossRef]
27. Xu, D.; Chen, S.; Zhou, Y.; Ji, W.; Shi, Z. Spatial Estimation of Soil Organic Matter and Total Nitrogen by Fusing Field Vis–NIR Spectroscopy and Multispectral Remote Sensing Data. *Remote Sens.* **2025**, *17*, 729. [CrossRef]
28. He, Y.; Yin, H.; Chen, Y.; Xiang, R.; Zhang, Z.; Chen, H. Soil Salinity Estimation Based on Sentinel-1/2 Texture Features and Machine Learning. *IEEE Sensors J.* **2024**, *24*, 15302–15310. [CrossRef]
29. Sun, X.; Yang, Z.; Su, P.; Wei, K.; Wang, Z.; Yang, C.; Wang, C.; Qin, M.; Xiao, L.; Yang, W.; et al. Non-destructive monitoring of maize LAI by fusing UAV spectral and textural features. *Front. Plant Sci.* **2023**, *14*, 1158837. [CrossRef] [PubMed]
30. Yin, Q.; Zhang, Y.; Li, W.; Wang, J.; Wang, W.; Ahmad, I.; Zhou, G.; Huo, Z. Better Inversion of Wheat Canopy SPAD Values before Heading Stage Using Spectral and Texture Indices Based on UAV Multispectral Imagery. *Remote Sens.* **2023**, *15*, 4935. [CrossRef]
31. Tzanakakis, V.A.; Mauromoustakos, A.; Angelakis, A.N. Prediction of Biomass Production and Nutrient Uptake in Land Application Using Partial Least Squares Regression Analysis. *Water* **2015**, *7*, 1–11. [CrossRef]
32. Zhao, W.; Ma, H.; Zhou, C.; Zhou, C.; Li, Z. Soil Salinity Inversion Model Based on BPNN Optimization Algorithm for UAV Multispectral Remote Sensing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 6038–6047. [CrossRef]
33. Zhu, Q.; Wang, Y.; Luo, Y. Improvement of multi-layer soil moisture prediction using support vector machines and ensemble Kalman filter coupled with remote sensing soil moisture datasets over an agriculture dominant basin in China. *Hydrol. Process.* **2021**, *35*, e14154. [CrossRef]

34. Zhao, Z.-D.; Zhao, M.-S.; Lu, H.-L.; Wang, S.-H.; Lu, Y.-Y. Digital Mapping of Soil pH Based on Machine Learning Combined with Feature Selection Methods in East China. *Sustainability* **2023**, *15*, 12874. [CrossRef]
35. Wang, F.; Yang, M.; Ma, L.; Zhang, T.; Qin, W.; Li, W.; Zhang, Y.; Sun, Z.; Wang, Z.; Li, F.; et al. Estimation of Above-Ground Biomass of Winter Wheat Based on Consumer-Grade Multi-Spectral UAV. *Remote Sens.* **2022**, *14*, 1251. [CrossRef]
36. Shao, Y.; Bao, Y.; He, Y. Soil Nitrogen Content Influence on Canopy Reflectance Spectra. *Trans. ASABE* **2011**, *54*, 2343–2349. [CrossRef]
37. Hes, E.M.A.; Yatoi, R.; Laisser, S.L.; Feyissa, A.K.; Irvine, K.; Kipkemboi, J.; van Dam, A.A. The effect of seasonal flooding and livelihood activities on retention of nitrogen and phosphorus in Cyperus papyrus wetlands, the role of aboveground biomass. *Hydrobiologia* **2021**, *848*, 4135–4152. [CrossRef]
38. Tercero, M.d.C.; Álvarez-Rogel, J.; Conesa, H.M.; Párraga-Aguado, I.; González-Alcaraz, M.N. Phosphorus retention and fractionation in an eutrophic wetland: A one-year mesocosms experiment under fluctuating flooding conditions. *J. Environ. Manag.* **2017**, *190*, 197–207. [CrossRef]
39. Martínez-Espinosa, C.; Sauvage, S.; Al Bitar, A.; Green, P.A.; Vörösmarty, C.J.; Sánchez-Pérez, J.M. Denitrification in wetlands: A review towards a quantification at global scale. *Sci. Total Environ.* **2021**, *754*, 142398. [CrossRef]
40. Mastný, J.; Bárta, J.; Kaštovská, E.; Picek, T. Decomposition of peatland DOC affected by root exudates is driven by specific r and K strategic bacterial taxa. *Sci. Rep.* **2021**, *11*, 18677. [CrossRef]
41. Jarocińska, A.; Kopeć, D.; Kycko, M.; Piórkowski, H.; Błońska, A. Hyperspectral vs. Multispectral data: Comparison of the spectral differentiation capabilities of Natura 2000 non-forest habitats. *ISPRS J. Photogramm. Remote. Sens.* **2022**, *184*, 148–164. [CrossRef]
42. Liu, M.; Zhang, S.; Gao, J.; Wang, B.; Fang, K.; Liu, L.; Lv, S.; Zhang, Q. Soil Water-Soluble Ion Inversion via Hyperspectral Data Reconstruction and Multi-Scale Attention Mechanism: A Remote Sensing Case Study of Farmland Saline–Alkali Lands. *Agronomy* **2025**, *15*, 1779. [CrossRef]
43. Anand, A.; Malhi, R.K.M.; Srivastava, P.K.; Singh, P.; Mudaliar, A.N.; Petropoulos, G.P.; Kiran, G.S. Optimal band characterization in reformation of hyperspectral indices for species diversity estimation. *Phys. Chem. Earth Parts A/B/C* **2022**, *126*, 103040. [CrossRef]
44. Yang, K.; Gong, Y.; Fang, S.; Duan, B.; Yuan, N.; Peng, Y.; Wu, X.; Zhu, R. Combining Spectral and Texture Features of UAV Images for the Remote Estimation of Rice LAI throughout the Entire Growing Season. *Remote Sens.* **2021**, *13*, 3001. [CrossRef]
45. Alan, L.; Environmental Consequences of Water Withdrawals and Drainage of Wetlands. UF/IFAS Extension Publication Series/Report Number: SL 302; University of Florida, Institute of Food and Agricultural Sciences (UF/IFAS). Available online: https://edis.ifas.ufl.edu (accessed on 24 June 2025).
46. Long, C.; Zhang, Q.; Chen, Q.; Cheng, X. Divergent Controls on Leaf and Root Litter Decay Linking to Soil C, N, and P Pools Under a Subtropical Land-use Change. *Ecosystems* **2023**, *26*, 1209–1223. [CrossRef]

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.