

Article

Improved Estimation of Cotton Aboveground Biomass Using a New Developed Multispectral Vegetation Index and Particle Swarm Optimization

Guanyu Wu ^{1,†}, Mingyu Hou ^{1,†}, Yuqiao Wang ¹, Hongchun Sun ¹, Liantao Liu ¹ , Ke Zhang ¹, Lingxiao Zhu ¹, Xiuliang Jin ² , Cundong Li ¹ and Yongjiang Zhang ^{1,*} 

¹ State Key Laboratory of North China Crop Improvement and Regulation, Hebei Key Laboratory of Crop Growth Regulation, College of Agronomy, Hebei Agricultural University, Baoding 071001, China; 15830103006@163.com (G.W.)

² Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China

* Correspondence: zhangyj@hebau.edu.cn

[†] These authors contributed equally to this work.

Abstract

Accurate and rapid estimation of aboveground biomass (AGB) in cotton is crucial for precise agricultural management. However, current AGB estimation methods are limited by data homogeneity and insufficient model accuracy, which fail to comprehensively reflect the cotton growth status. This study introduces a novel approach by coupling cotton canopy Soil and Plant Analyzer Development (SPAD) values with multispectral (MS) data to achieve precise estimation of cotton AGB. Two experimental treatments, involving varied nitrogen fertilizer rates and organic manure applications, were conducted from 2022 to 2023. MS data from UAVs were collected across multiple cotton growth stages, while AGB and canopy SPAD values were synchronously measured. Using the coefficient of variation method, SPAD values were coupled with existing vegetation indices to develop a novel vegetation index termed CGSI_{VI}. Moreover, the applicability of various machine learning algorithms—including Random Forest Regressor (RFR), eXtreme Gradient Boosting (XGBoost), Categorical Boosting (CatBoost), Particle Swarm Optimization-XGBoost (PSO-XGBoost), and Particle Swarm Optimization-CatBoost (PSO-CatBoost)—was evaluated for inverting cotton AGB. The results indicated that, compared to the original vegetation indices, the correlation between the improved vegetation index (CGSI_{VI}) and AGB was enhanced by 13.60% overall, with the CGSI_{CRe} exhibiting the highest correlation with cotton AGB ($R^2 = 0.87$). The overall AGB estimation accuracy across different growth stages, spanning the entire growth period, ranged from 0.768 to 0.949, peaking during the flowering stage. Furthermore, when the CGSI_{VI} was used as an input parameter in comparisons of different machine learning algorithms, the PSO-XGBoost algorithm demonstrated superior estimation accuracy across the entire growth stage and within individual growth stages. This high-throughput crop phenotyping analysis method enables rapid and accurate estimation. It reveals the spatial heterogeneity of cotton growth status, thereby providing a powerful tool for accurately identifying growth differences in the field.

Keywords: unmanned aerial multispectral remote sensing; cotton; aboveground biomass; vegetation index; machine learning algorithm



Academic Editors: Hengbiao Zheng and Xia Yao

Received: 10 November 2025

Revised: 10 December 2025

Accepted: 16 December 2025

Published: 17 December 2025

Citation: Wu, G.; Hou, M.; Wang, Y.; Sun, H.; Liu, L.; Zhang, K.; Zhu, L.; Jin, X.; Li, C.; Zhang, Y. Improved Estimation of Cotton Aboveground Biomass Using a New Developed Multispectral Vegetation Index and Particle Swarm Optimization.

Agriculture **2025**, *15*, 2608.

<https://doi.org/10.3390/agriculture15242608>

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/](https://creativecommons.org/licenses/by/4.0/)

<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Aboveground biomass (AGB), serving as a pivotal indicator for assessing the accumulation of organic matter per unit area in crops, is crucial for monitoring crop growth status, nutrient diagnosis, and yield prediction [1]. In large-scale farmland, traditional measurement methods are time-consuming, labor-intensive, and prone to low precision due to human interference and poor timeliness [2]. In contrast, remote sensing technology, characterized by its low cost and high timeliness, offers the potential for rapidly and accurately acquiring biomass information over broad areas [3].

Unmanned aerial vehicles (UAVs) offer unique advantages in estimating AGB, as they can carry various sensors to capture phenotypic parameter information of crops, significantly enhancing the comprehensiveness and accuracy of the collected data [4]. For instance, multispectral vegetation indices such as GNDVI and NDRE, acquired through multispectral sensors, have improved the estimation accuracy of forest AGB ($R^2 = 0.71$) [5]. Furthermore, by incorporating both multispectral vegetation indices (VIs) and Fourier spectrum textures as input parameters in machine learning algorithms, models for extracting rice growth parameters have been developed, enhancing the precision of estimating rice leaf area index, canopy chlorophyll content, and aboveground biomass [6]. Increasing amounts of research show that integrating various data sources can enhance the estimation accuracy of AGB. The fusion of UAV-derived multisource data and phenological information has been demonstrated to be a reliable predictor of crop biomass [7]. The results indicated that this approach yielded more accurate predictions than using a single indicator [8]. Besides remote sensing data, biochemical indicators like chlorophyll content effectively indicate crop growth. Soil and Plant Analyzer Development (SPAD) values, closely linked to chlorophyll content, serve as a useful parameter for estimating crop phenotypes [9]. While remote sensing data efficiently estimate crop aboveground biomass, few studies have integrated UAV multispectral data with SPAD values for cotton AGB estimation.

In recent years, the swift advancements in computing technology and artificial intelligence have facilitated substantial progress in crop phenotype estimation through machine learning and deep learning models [10]. Nevertheless, the performance of traditional machine learning approaches remains constrained by limitations related to datasets and generalization capabilities [11,12]. In contrast, ensemble learning methods, such as random forest regression (RFR), enhance estimation accuracy by amalgamating predictions from multiple models [13]. Previous research has demonstrated that maize biomass can be predicted using unmanned aerial vehicle hyperspectral data coupled with maize plant height data, and the XGBoost regression yielded superior accuracy in predicting maize dry and fresh weights relative to multiple stepwise regression and random forest regression, with the peak accuracy for dry matter prediction occurring at the V6 growth stage [14]. Separate studies have focused on estimating wheat aboveground biomass, and in these studies, multiple machine learning models were optimized using algorithms like the GOA (Grasshopper Optimization Algorithm) and PSO (Particle Swarm Optimization); the optimization results have revealed that the GOA-XGB model achieves the highest estimation accuracy when multispectral bands and vegetation indices (VIs) are used as combined input parameters [15].

While previous studies have predominantly relied on individual VIs for biomass estimation, often leveraging machine learning techniques, this approach fails to fully capture the synergistic interactions and complementary information among multiple VIs. Moreover, it overlooks the complex and dynamic nature of crop growth processes, which are influenced by multifaceted factors such as phenological stages, environmental conditions, and agronomic practices. A more holistic approach—integrating multiple vegetation indices (VIs) and advanced modeling frameworks—is therefore essential to enhance the

accuracy and robustness of biomass estimation. The overreliance on single VIs limits further improvements in estimation accuracy, whereas ground-based data (e.g., field-measured SPAD values) typically exhibit high accuracy and can complement spectral information. Cotton is an important economic crop, and its canopy structure (such as multiple branches and dense leaf arrangement) is more complex than crops such as wheat and rice. There is still a lack of research on whether coupling spectral indices with SPAD can improve the estimation accuracy of AGB. In this study, ground data SPAD values are combined with multispectral vegetation indices to construct a new type of vegetation index, and by comparing the applicability of different machine learning algorithms in the inversion of cotton AGB. We develop an estimation method of cotton AGB that combines the improvement in vegetation indices and optimization of algorithms, which can provide an alternative and effective method for monitoring crop growth in agricultural production.

2. Materials and Methods

2.1. Study Sites and Experimental Design

The experimental site was situated at the Weixian Experimental Station (115°27' E, 36°56' N, Hebei Province) from 2022 to 2023 (Figure 1), characterized by a warm temperate continental semi-arid monsoon climate. The study comprises two experimental treatments. In experiment I, the cotton cultivar Jinongda 36 was used as the experimental material, with six nitrogen levels established: 0 kg/hm² (N1), 75 kg/hm² (N2), 150 kg/hm² (N3), 225 kg/hm² (N4), 300 kg/hm² (N5), and 375 kg/hm² (N6). A randomized block design was employed, with each treatment replicated three times. The ratio of basal nitrogen fertilizer to topdressing (applied at the early flowering stage) was maintained at 6:4. Cotton was planted with a uniform row spacing of 76 cm, and the plot area was standardized at 66 m², with a length of 10 m and a width of 6.6 m. The sowing date was set for April 25th. In experiment II, Shikang 126 was utilized as the test material, and ten distinct treatments were configured. These treatments included CK (control without fertilizer application), CKF (conventional chemical fertilizer application with N:P₂O₅:K₂O ratios of 12 kg/hm²/5 kg/hm²/10 kg/hm²), M1 (moderate organic fertilizer application, specifically chicken manure at 1500 kg/hm²), M2 (high organic fertilizer application, chicken manure at 3750 kg/hm²), and a series of treatments exploring the integration of organic and reduced chemical fertilizers: L1 (moderate organic fertilizer with 10% reduction in chemical fertilizer), L2 (moderate organic fertilizer with 30% reduction in chemical fertilizer), L3 (moderate organic fertilizer with 50% reduction in chemical fertilizer), H1 (high organic fertilizer with 10% reduction in chemical fertilizer), H2 (high organic fertilizer with 30% reduction in chemical fertilizer), and H3 (high organic fertilizer with 50% reduction in chemical fertilizer). A randomized block design was adopted for this experiment as well, with each plot replicated four times (Figure 1d). Each plot measured 10 m in length, 5 m in width, and featured a uniform row spacing of 76 cm. All other management practices followed local high-yield field standards. Both experimental plots in this study featured loam soil. For Plot 1, the soil properties were as follows: organic matter content of 7.7 g/kg, total nitrogen of 0.57 g/kg, available nitrogen of 52.33 mg/kg, available phosphorus of 6.83 mg/kg, and available potassium of 112.12 mg/kg. For Plot 2, the soil properties were organic matter content of 8.39 g/kg, total nitrogen of 0.85 g/kg, available nitrogen of 54.16 mg/kg, available phosphorus of 16.23 mg/kg, and available potassium of 132.45 mg/kg. The irrigation amount was referenced to the standard of local high-yield fields.

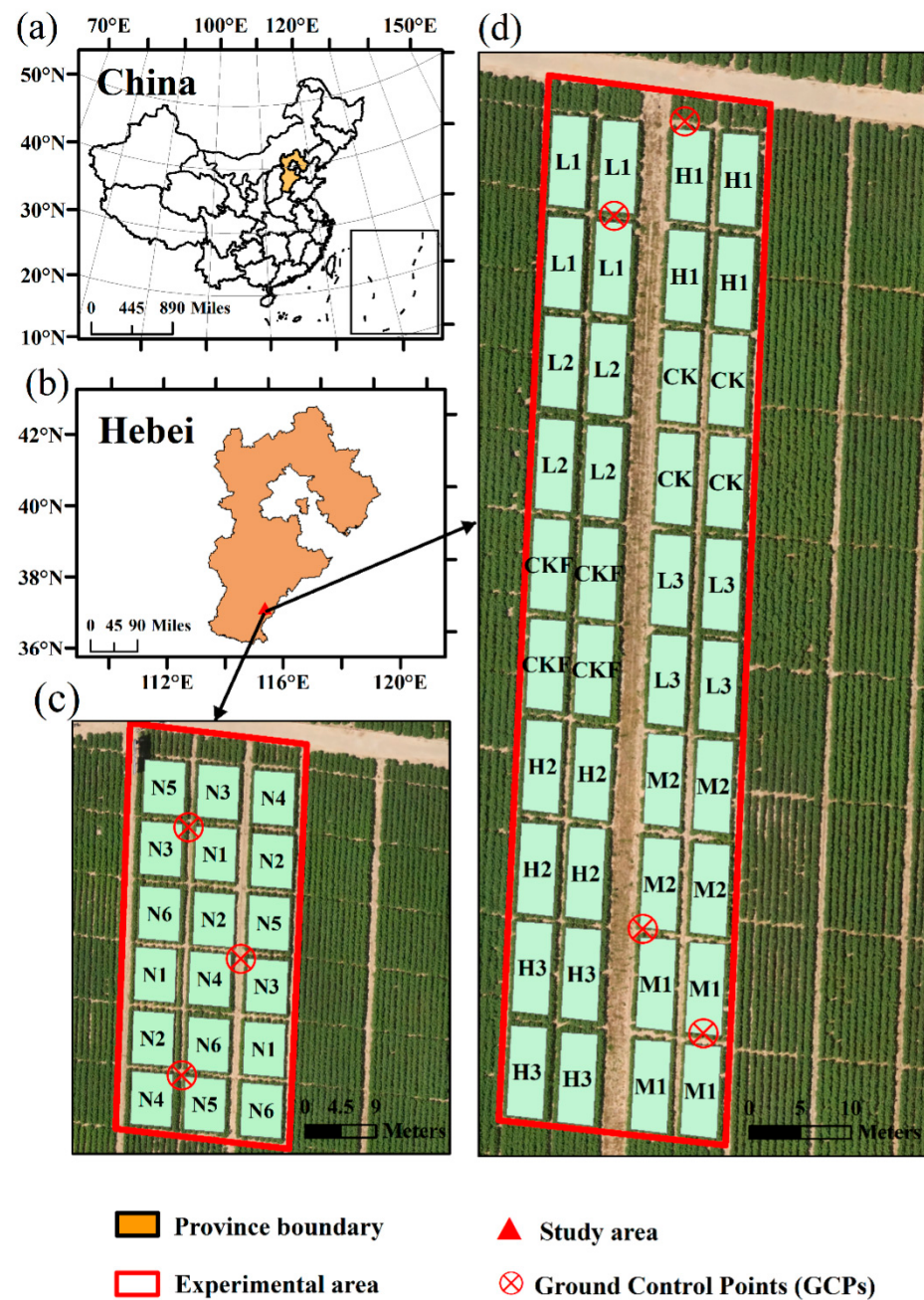


Figure 1. Location of the study area and distribution of experimental sites. (a) Location of Hebei Province in China; (b) location of the experimental station in Hebei Province; (c) distribution of experimental site plot I; (d) distribution of experimental site plot II.

2.2. Data Acquisition

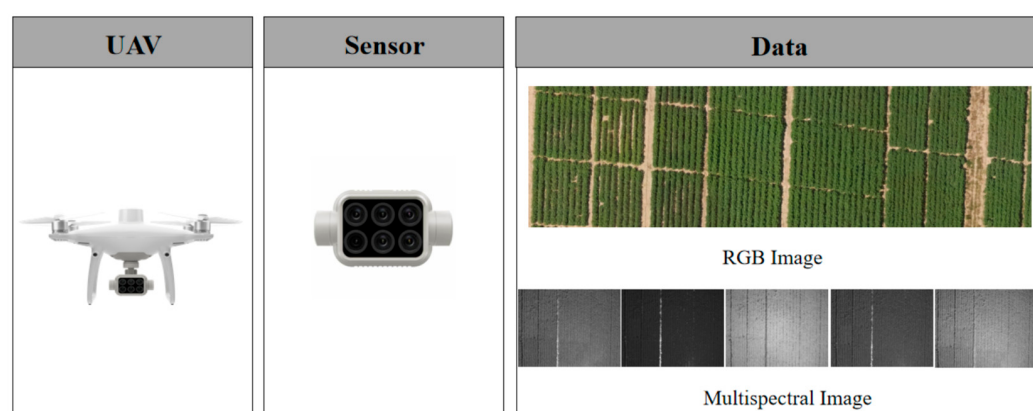
2.2.1. UAV Data Acquisition and Preprocessing

Using a quadcopter DJI Phantom 4 multispectral drone equipped with RGB and multispectral sensors (Table 1), data were collected during clear and cloudless weather, specifically between 11:00 and 14:00. The heading overlap rate was set to 80%, the side overlap rate to 90%, and the flight altitude to 50 m, resulting in a ground resolution of 3.5 cm. The UAV, sensors, and their data acquisition process are illustrated in Figure 2. Before acquiring the multispectral image data, a calibration plate was photographed with the multispectral camera for radiometric calibration. The data were collected at different stages of cotton growth: squaring stage (SS), flowering stage (FS), boll maturation stage (BMS), and boll opening stage (BOS).

Table 1. UAV lens parameters.

Camera Parameter	Parameter Value
Filter	Blue: 450 nm $\pm$ 16 nm; Green: 560 nm $\pm$ 16 nm; Red: 650 nm $\pm$ 16 nm; Red Edge: 730 nm $\pm$ 16 nm; Near-Infrared: 840 nm $\pm$ 26 nm
Color Sensor ISO Range	200–800
Monochrome Sensor Gain	1–8 times
Electronic Global Shutter	1/100–1/20000 s (RGB); 1/100–1/10000 s (Multispectral imaging)
Maximum Photo Resolution	1600 $\times$ 1300 (4: 3.25)

Image stitching was performed using Agisoft PhotoScan software (version 1.4.5, Agisoft LLC, St. Petersburg, Russia), and radiometric calibrations were used to obtain orthophotos. Processes such as band combination and alignment were performed using ENVI (version 5.6, L3Harris Technologies, Inc., Melbourne, FL, USA) and ArcMap (version 10.6, ESRI, Inc., Redlands, CA, USA) software.

**Figure 2.** Experimental equipment and field images.

Detailed parameters of the DJI Elf 4 drone lens are shown in Table 1.

2.2.2. SPAD and AGB Acquisition

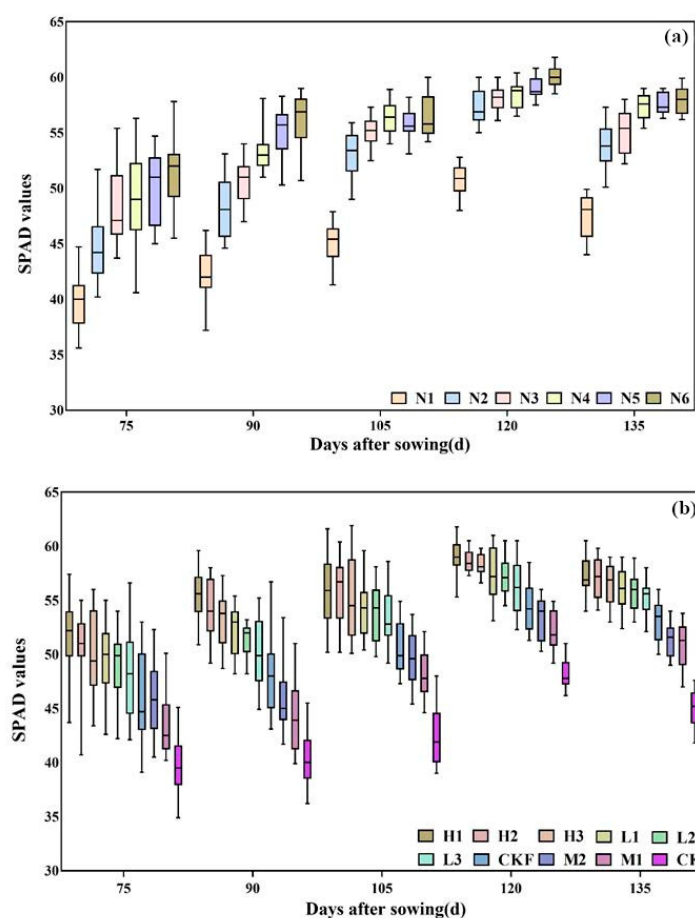
On the day of the UAV flight, cotton SPAD values and AGB were measured synchronously. The upper, middle, and lower parts of the cotton canopy leaves were measured using SPAD-502 (chlorophyll content meter), and the average value was taken as the canopy SPAD value. Three plants were replicated for each treatment. In each plot, five cotton plants of uniform growth were selected, decomposed, and put into kraft paper bags, baked in an oven at 105 °C for 30 min, then dried at 80 °C until constant weight, weighed, and finally recorded. Changes in cotton SPAD values at different stages under different fertilizer treatments can be seen in Figure 3.

2.2.3. Selection of Spectral Indices

In this study, 20 commonly used multispectral and RGB vegetation indices related to crop AGB were selected (Table 2).

Table 2. Vegetation indices extracted from multispectral and RGB images.

Sensor	Vegetation Index	Calculation Formula	References
Multispectral	Normalized Difference Vegetation Index (NDVI)	$NDVI = (NIR - R) / (NIR + R)$	[16]
	Enhanced Vegetation Index (EVI)	$EVI = 2.5 \times (NIR - R) / (NIR + 6 \times R - 7.5 \times B + 1)$	[17]
	Two-Band Enhanced Vegetation Index (EVI2)	$EVI2 = 2.5 \times (NIR - R) / (NIR + 2.4 \times R + 1)$	[18]
	Global Environment Monitoring Index (GEMI)	$GEMI = [2 \times (NIR^2 - R^2) + 1.5 \times NIR + 0.5 \times R] / (NIR + R + 0.5)$	[19]
	Soil-Adjusted Vegetation Index (SAVI)	$SAVI = 1.5 \times (NIR - R) / (NIR + R + 0.5)$	[20]
	Modified Soil-Adjusted Vegetation Index (MSAVI)	$MSAVI = 0.5 \times [2 \times NIR + 1 - \sqrt{(2 \times NIR + 1)^2 - 8 \times (NIR - R)}]$	[21]
	Transformed Soil-Adjusted Vegetation Index (TSAVI)	$TSAVI = 0.33 \times (NIR - 0.33 \times R - 0.5) / [0.5 \times NIR + R - 0.5 \times 0.33 + (1 + 0.33^2)]$	[22]
	Near-Infrared Reflectance times Vegetation (NIRv)	$NIRv = NDVI \times NIR$	[23]
	Green Chlorophyll Vegetation Index (GCVI)	$GCVI = NIR / GREEN - 1$	[24]
	Green Difference Vegetation Index (GDVI)	$GDVI = NIR - GREEN$	[25]
	Green Normalized Difference Vegetation Index (GNDVI)	$GNDVI = (NIR - G) / (NIR + G)$	[26]
	Chlorophyll Index Red Edge (CIre)	$CIre = NIR / REDedge - 1$	[27]
RGB	Red Edge Normalized Difference Vegetation Index (NDVIre)	$NDVIre = (NIR - REDedge) / (NIR + REDedge)$	[28]
	Difference Vegetation Index (DVI)	$DVI = NIR - R$	[29]
	Red to Green Ratio Index (RGRI)	$RGRI = R / G$	[30]
	Green Chromatic Coordinate (GCC)	$GCC = G / (R + G + B)$	[31]
	Green Red Vegetation Index (GRVI)	$GRVI = (GREEN - RED) / (GREEN + RED)$	[32]
	Water Index (WI)	$WI = (G - B) / (R - G)$	[33]
	Visible Atmospherically Resistant Index (VARI)	$VARI = (G - R) / (G + R - B)$	[34]
	Principal component analysis index (IPCA)	$IPCA = 0.994 \times R - B + 0.961 \times G - B + 0.914 \times G - R $	[35]

**Figure 3.** Changes in cotton SPAD values under different fertilizer treatments. (a) Experiment I; (b) experiment II.

2.3. VI-Weight SPAD Variation Coefficient

To quantify the relative importance of SPAD values and VIs in assessing crop growth status, we employed a two-step weighting approach combining variance inflation factor screening and the coefficient of variation (CV) method. The rationale for CV-based weighting is as follows: The relative variability of SPAD values and vegetation indices (VIs) directly reflects their response sensitivity to differences in cotton biomass. On the premise of eliminating multicollinearity ($VIF < 10$), indicators with higher variability better capture spatiotemporal heterogeneity of cotton biomass under different fertilizer treatments and growth stages, and they also make greater informational contributions to biomass prediction [11,36]. Subsequently, CV was calculated for each SPAD indicator and VI to evaluate their variability. CV_i was derived as the ratio of the standard deviation (σ_i) to the mean of each indicator ($\bar{x}_i$), as shown in Equation (1):

$$CV_i = \frac{\sigma_i}{\bar{x}_i} \quad (1)$$

The normalized weight (W_i) for each indicator was then computed by dividing its coefficient of variation by the sum of all coefficients of variation, as expressed in Equation (2):

$$W_i = \frac{CV_i}{\sum_{i=1}^n CV_i} \quad (2)$$

This weighting scheme reflects the relative contribution of each indicator to the overall variability, with higher weights assigned to indicators exhibiting greater dispersion. To integrate SPAD values and VIs into a unified metric, both indicators were normalized to a common scale using max normalization, as described in Equation (3):

$$U_i = \frac{x_i}{\max(x_i)} \quad (3)$$

Finally, a novel vegetation index, termed the CGSI_{VI}, was developed by combining the weighted SPAD values and VIs. This approach not only mitigates the limitations of relying on single indicators but also provides a comprehensive representation of crop growth dynamics, as demonstrated in previous studies [37].

The newly developed vegetation index CGSI_{VI} is ultimately constructed by the weighted fusion of normalized SPAD values and VIs, with the formula presented as follows:

$$CGSI_{VI} = W_{SPAD} \times U_{SPAD} + W_{VI} \times U_{VI} \quad (4)$$

2.4. Setting of Regression Methods

2.4.1. RFR

In the study of AGB estimation based on remote sensing technology, the RFR algorithm has shown good predictive ability in this field [38]. The study implemented the RFR model using the Python 3.7 environment (Google Inc., Mountain View, CA, USA) and achieved effective processing and prediction of random data samples by constructing an integration containing 100 decision trees. Except for specifying the number of trees as 100, all other configurations of the model followed the default parameter settings of the scikit-learn library. The AGB was estimated separately for the two types of independent variables, VI and CGSI_{VI}, and all the metrics of the corresponding type were utilized in each case to assess the performance of the RFR model under different input conditions.

2.4.2. XGBoost

XGBoost is an optimization algorithm based on Gradient Boosting Decision Tree (GBDT). Compared to random forest (RF), XGBoost can correct the residuals more finely by constructing a new tree iteratively and learning a new function in each round to fit the residuals of the previous round, to predict the scores efficiently based on the sample characteristics.

2.4.3. CatBoost

This study applied the Yandex open source CatBoost algorithm, which has a built-in category feature processing mechanism with a gradient boosting framework [39]. The parameter configurations follow the default values of CatBoost, except for the dataset partitioning, to evaluate the algorithm's performance with standard configurations.

2.4.4. PSO-CatBoost

The PSO-CatBoost algorithm combines PSO with CatBoost for optimizing model parameters [40]. The range of hyperparameters was set to cover the number of iterations (400~1100), tree depth (1~10), learning rate (0.01~0.5), and regularization term (1~10). The PSO configuration was set to 80 particles with an upper limit of 200 iterations. The stopping criterion for PSO optimization adopts a dual-threshold mechanism: either reaching the maximum number of iterations (200) or the variation in the optimal fitness value of the population being less than 1×10^{-5} for 10 consecutive iterations (the fitness function aims to maximize the R^2 of the validation set). The iteration stops when either condition is satisfied.

2.4.5. PSO-XGBoost

The PSO-XGBoost algorithm is applied to integrate the PSO and XGBoost techniques in order to optimize the model parameters. The key hyperparameters are set as number of trees (1000~1500), maximum depth of trees (1~10), learning rate (0.1~0.5), and regularization factor (1~10). The stopping criterion is consistent with that of PSO-CatBoost.

The dataset was split into the training set and the test set at a ratio of 7:3. During the splitting process, spatial and temporal correlations were strictly controlled. The `sklearn.train_test_split` function was employed, with `random_state = 42` set to fix the random sampling results.

2.5. Correlation Analysis and Accuracy Assessment

Accuracy Assessment

Performance metrics such as coefficient of determination (R^2) and relative root mean square error (rRMSE) are calculated on the test set. The two metrics are calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (5)$$

$$rRMSE = \frac{\sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}}{\bar{y}} \times 100\% \quad (6)$$

In the above equation, x_i represents the estimated value of cotton AGB, y_i represents the measured value of cotton AGB, $\bar{y}_i$ is the mean of the sample estimates, and n represents the total sample size.

3. Results

3.1. Descriptive Statistics of Cotton AGB

Cotton AGB increased with the increase in fertilizer rate during 2022–2023, with standard deviations ranging from 0.290 to 0.616 t/hm² for single stage (116 samples were taken at each growth stage, reaching a maximum value at the boll opening stage, Table 3). For the whole growth stage (464 samples in total), the standard deviation of AGB was 2.665 t/hm², with significant variability, and the larger variation was favorable for the establishment and validation of the inversion model.

Table 3. Statistics for cotton AGB of training and testing datasets.

Growth Stage	Dataset	Number of Samples	Maximum Value (t/hm ²)	Minimum Value (t/hm ²)	Average Value (t/hm ²)	Standard Deviation (t/hm ²)
Entire growth stage	Training	324	10.007	0.633	4.585	2.754
	Testing	140	9.605	0.765	4.516	2.665
Squaring stage	Training	81	2.480	0.633	1.694	0.519
	Testing	35	2.417	0.666	1.679	0.477
Flowering stage	Training	81	3.623	2.522	3.041	0.290
	Testing	35	3.621	2.592	3.051	0.316
Boll maturation stage	Training	81	5.245	4.187	4.726	0.294
	Testing	35	5.192	4.113	4.750	0.329
Boll opening stage	Training	81	10.007	6.477	8.809	0.616
	Testing	35	9.646	7.205	8.866	0.594

3.2. Correlation Analysis of VI and CGSI_{VI} with AGB

VIs were calculated using raw spectra, and 40 indices were derived by the CV method. The degree of correlation between these indices and cotton AGB was quantified using Pearson correlation analysis (Figure 4).

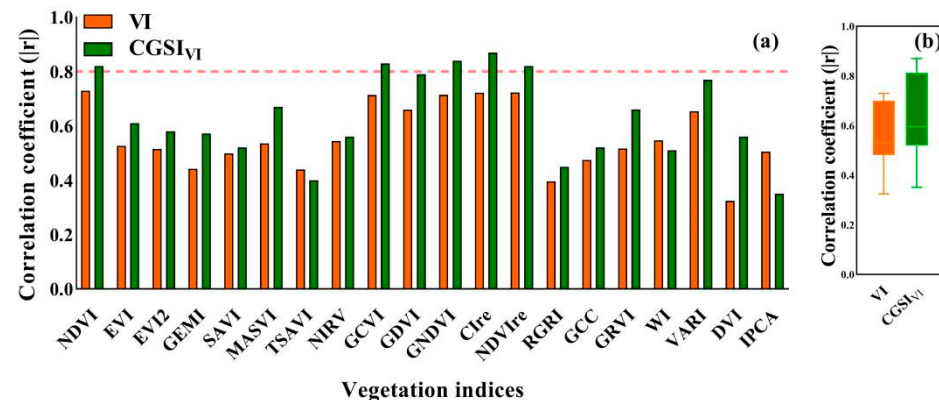


Figure 4. Correlation analysis of indicators and biomass of cotton. (a) Multispectral vegetation index and biomass correlation; (b) correlation coefficient box plot.

As shown in Figure 4a, the absolute correlation coefficients for all VIs remained below 0.80, the threshold for strong correlation, indicating a relatively weak correlation between these VIs and cotton AGB. Among them, the absolute values of correlation coefficients for five VIs, namely GCVI, NDVI, GNDVI, CIre, and NDVIre, were relatively higher (0.70–0.72), but they did not reach a significant level. Among the new VIs derived through the CV method, five indices had absolute values of correlation coefficients exceeding the threshold of 0.80, including the CGSI_{NDVI}, CGSI_{GCVI}, CGSI_{GNDVI}, CGSI_{CIre}, and CGSI_{NDVIre}. Notably, the CGSI_{CIre} exhibited the highest correlation coefficient with cotton AGB, reaching 0.87. As shown in Figure 4b, the CGSI_{VI} generally had a higher correlation with cotton AGB, with an overall improvement of 13.6% compared to VIs calculated from the original spectra.

Through correlation analysis, the index with the highest correlation coefficient with cotton AGB was selected, and the relationship between CIre and CGSI_{CIre} with cotton

AGB across different growth stages was compared (Figure 5). As shown in the figure, the variation range of the vegetation index CI_{re} is relatively small, approximately around 0.10, and there is no obvious pattern of change. In contrast, the trend of the $CGSI_{CI_{re}}$ across different growth stages exhibits a similar increasing trend to that of cotton AGB.

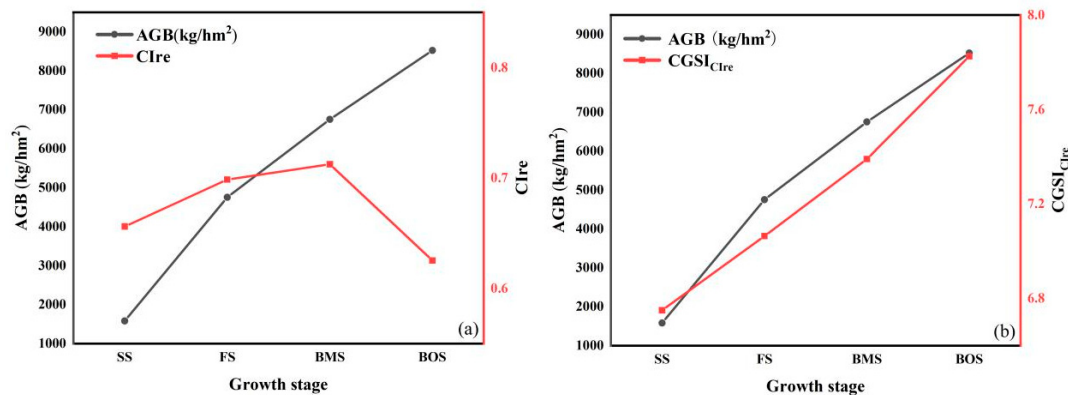


Figure 5. Indicators and trend of biomass of cotton during the different growth stages. (a) Changes in vegetation index CI_{re} and AGB at different stages; (b) changes in $CGSI_{CI_{re}}$ and AGB at different stages.

3.3. AGB Estimation with Different Machine Learning Algorithms at the Whole Stage

In this study, RFR, XGBoost, CatBoost, PSO-XGBoost, and PSO-CatBoost were utilized to estimate the AGB of cotton throughout its entire growth cycle. The input parameters included SPAD values, VI , and the $CGSI_{VI}$. As demonstrated in Table 4, the incorporation of the $CGSI_{VI}$ as an input parameter significantly improved the estimation accuracy across all models, yielding R^2 values ranging from 0.94 to 0.97 and rRMSE between 11.97% and 14.73%. In contrast, the use of SPAD values alone as the input parameter resulted in the lowest estimation accuracy, with R^2 values ranging from 0.43 to 0.65 and rRMSE values between 38.40% and 45.49%. These findings underscore the superior performance of the $CGSI_{VI}$ in capturing the complex dynamics of cotton AGB compared to traditional SPAD-based approaches.

Table 4. Accuracy of cotton AGB estimation using different data combinations and algorithms.

Data	RFR		XGBoost		CatBoost		PSO-XGBoost		PSO-CatBoost	
	R^2	rRMSE (%)	R^2	rRMSE (%)	R^2	rRMSE (%)	R^2	rRMSE (%)	R^2	rRMSE (%)
SPAD	0.43	44.61	0.51	45.49	0.49	41.83	0.63	39.79	0.65	38.40
VI	0.94	16.60	0.95	14.91	0.94	16.16	0.96	12.35	0.94	15.93
$CGSI_{VI}$	0.94	14.73	0.95	13.14	0.94	14.56	0.97	11.97	0.95	13.97

Figures 6 and 7 present scatter plots of cotton AGB estimations using VI s and the $CGSI_{VI}$ as input parameters for XGBoost, PSO-XGBoost, CatBoost, and PSO-CatBoost, respectively. As shown in the figures, AGB values above 8000 kg/hm² exhibited significant scatter. This scattering may be mainly attributed to spectral saturation—when AGB is high, cotton canopy spectral reflectance plateaus, reducing VI s' ability to distinguish biomass differences. Alternative factors include insufficient model learning of high-biomass samples, limited samples with AGB > 9000 kg/hm² (only eight in the test set), and slight field measurement errors in high-biomass plant processing. By comparing the scatter plots of Figures 6 and 7, it is found that using the $CGSI_{VI}$ as input parameters in machine learning algorithms improves the estimation accuracy of cotton AGB. Among them, the

PSO-XGBoost model exhibits the best performance in AGB estimation, with an R^2 value of 0.966 and an rRMSE of 11.97%.

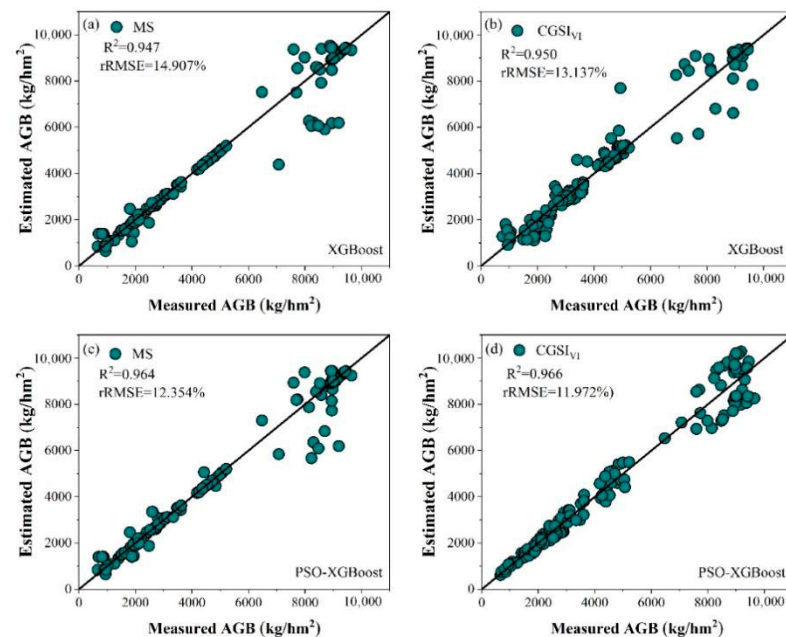


Figure 6. Scatter plots for estimating AGB using XGBoost and PSO-XGBoost algorithms: (a) VI + XGBoost, (b) CGSI_{VI} + XGBoost, (c) VI + PSO-XGBoost, and (d) CGSI_{VI} + PSO-XGBoost.

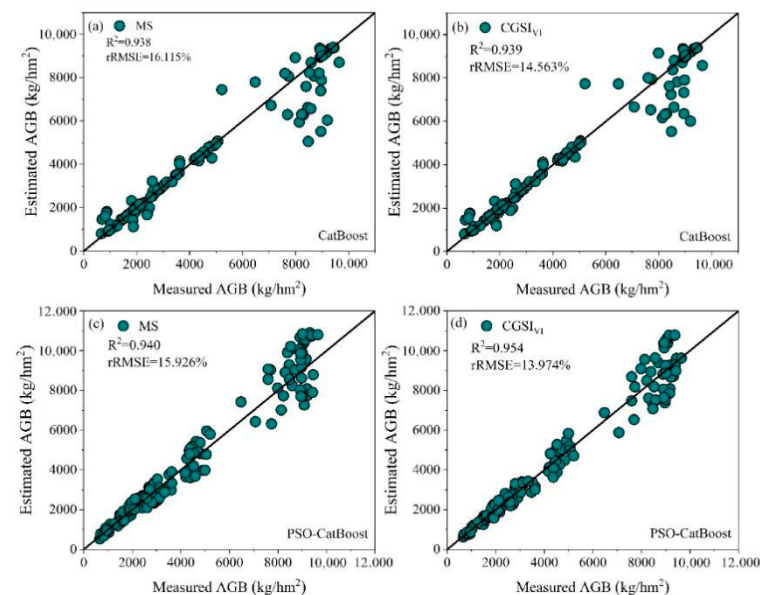


Figure 7. Scatterplot of AGB estimated by CatBoost and PSO-CatBoost algorithms. (a) MS+ CatBoost, (b) CGSI_{VI}+ CatBoost, (c) MS+PSO-CatBoost, and (d) CGSI_{VI}+PSO-CatBoost.

The comparison between R^2 and rRMSE indicates that the R^2 values for RFR, XGBoost, and CatBoost fluctuate within the ranges of 0.43–0.94, 0.51–0.95, and 0.49–0.94, respectively (Figure 8). Meanwhile, their corresponding rRMSE values fall within the ranges of 3.61–44.61%, 3.26–45.49%, and 3.42–41.83%, respectively. Among these three algorithms, both XGBoost and CatBoost show better prediction results. After applying PSO optimization to these models, the PSO-XGBoost algorithm achieves an R^2 ranging from 0.63 to 0.97 and an rRMSE between 2.80% and 39.80%. Compared to RFR, CatBoost, and XGBoost,

the PSO-XGBoost algorithm exhibits significantly improved performance, indicating its enhanced accuracy in estimating AGB.

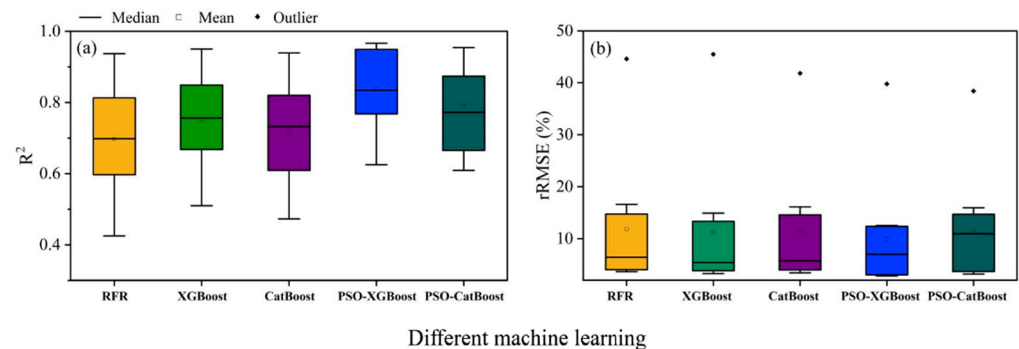


Figure 8. AGB estimation accuracy of different algorithms. (a) R^2 ; (b) rRMSE.

3.4. Estimates of AGB at Different Growth Stages

As can be seen from Table 5, the accuracy of cotton AGB estimation varied at different growth stages. The R^2 values of AGB estimation at the squaring stage ranged from 0.61 to 0.85, and the rRMSE ranged from 3.53% to 14.70%, and the R^2 was improved by 10.32%, on average, using the CGSI_{VI} as the input parameter compared with VI. As cotton entered the flowering stage, the accuracy of the AGB estimation reached the maximum, the R^2 values increased to 0.79~0.95, while the rRMSE was 2.80%~14.70%. The R^2 increased by an average of 4.49% with the CGSI_{VI} as an input parameter compared to VI as an input parameter into the five models. The R^2 values at the flowering stage were 0.47~0.87, showing that a decreasing trend was observed, while the rRMSE values ranged from 3.03% to 4.76%. As can be seen from Figure 9, the estimation accuracy of AGB showed a tendency of increasing and then decreasing with the advancement of the cotton reproductive process, and it reached its highest at the flowering stage.

Table 5. Comparison of AGB estimation accuracy and algorithms for different cotton stages.

Growth Stage	Data	RFR		XGBoost		CatBoost		PSO-XGBoost		PSO-CatBoost	
		R^2	rRMSE (%)	R^2	rRMSE (%)	R^2	rRMSE (%)	R^2	rRMSE (%)	R^2	rRMSE (%)
Squaring stage	VI	0.61	6.41	0.77	5.15	0.66	5.73	0.83	12.56	0.77	14.70
	CGSI _{VI}	0.70	5.91	0.75	5.37	0.77	5.20	0.89	3.53	0.85	4.19
Flowering stage	VI	0.79	3.84	0.81	3.81	0.76	3.85	0.91	2.98	0.84	3.20
	CGSI _{VI}	0.81	3.61	0.85	3.26	0.82	3.42	0.95	2.80	0.87	3.17
Boll maturation stage	VI	0.51	4.76	0.58	4.09	0.47	4.96	0.71	3.39	0.61	3.94
	CGSI _{VI}	0.60	4.00	0.67	3.63	0.61	3.94	0.77	3.03	0.67	3.64
Boll opening stage	VI	0.66	12.75	0.67	13.31	0.66	13.64	0.80	9.08	0.77	12.25
	CGSI _{VI}	0.70	13.32	0.76	11.96	0.73	13.05	0.82	6.98	0.77	10.92

Using the CGSI_{VI} as an input parameter effectively improved the accuracy of the AGB estimation at different growth stages (Figure 10), and this improvement was consistent with the trend observed throughout the growth stages (Table 5). It is shown that the fusion technique of multispectral sensor data and SPAD values improves the accuracy of AGB estimation both in the whole growth stage and single stage.

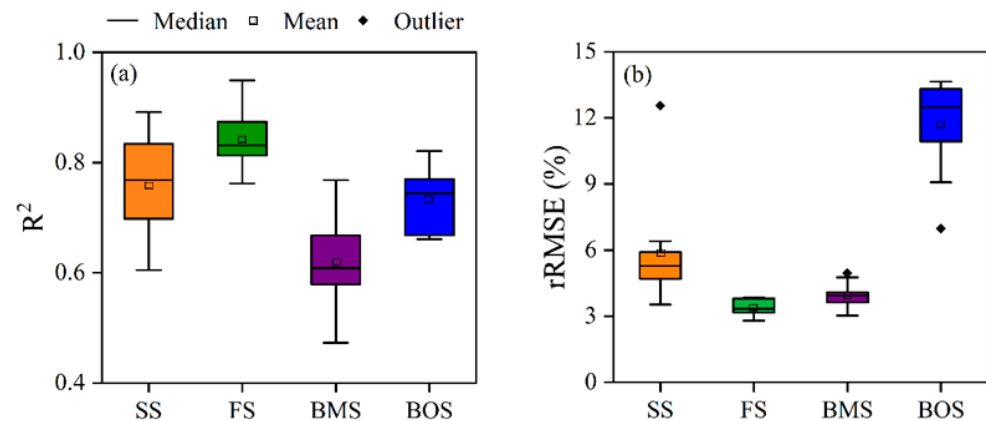


Figure 9. Statistics of AGB estimation accuracy at different growth stages. (a) R^2 ; (b) rRMSE.

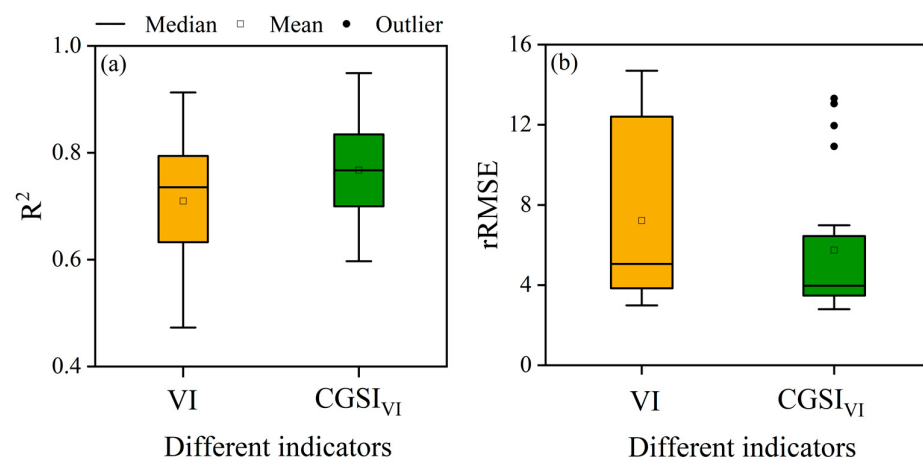


Figure 10. Statistics of AGB estimation accuracy for different vegetation indices. (a) R^2 ; (b) rRMSE.

3.5. Spatial and Temporal Distribution of Cotton AGB

Figures 11 and 12 show the spatial distribution of AGB based on RGB-VIs in each growth cycle of cotton. AGB showed a significant growth trend as cotton phenological development progressed. As can be seen from Figure 11, at six different nitrogen fertilizer treatments (N1 to N6), it was found that cotton plants from N5 and N6 treatments exhibited higher AGB levels and grew better than the other treatment groups. In contrast, plants in the N1 treatment showed significantly lower AGB levels.

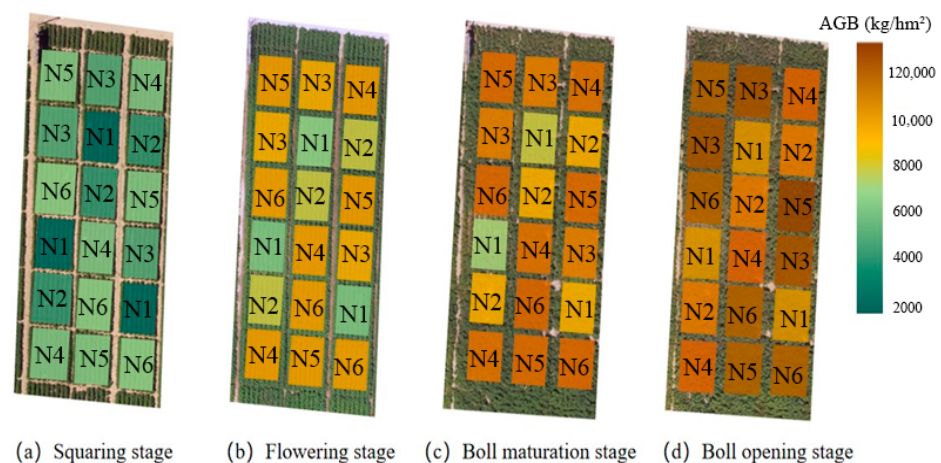


Figure 11. PSO-XGBoost-based CGSI_{VI} estimation of AGB spatiotemporal distribution under different nitrogen treatments.

As can be seen in Figure 12, the AGB of cotton also showed a significant increasing trend with the increase in fertilizer application in the 10 organic fertilizer treatment trials. Among them, the H1 treatment exhibited the highest AGB level due to its higher organic fertilizer input, which indicated that reasonable organic fertilizer input had a significant promotion effect on cotton growth.

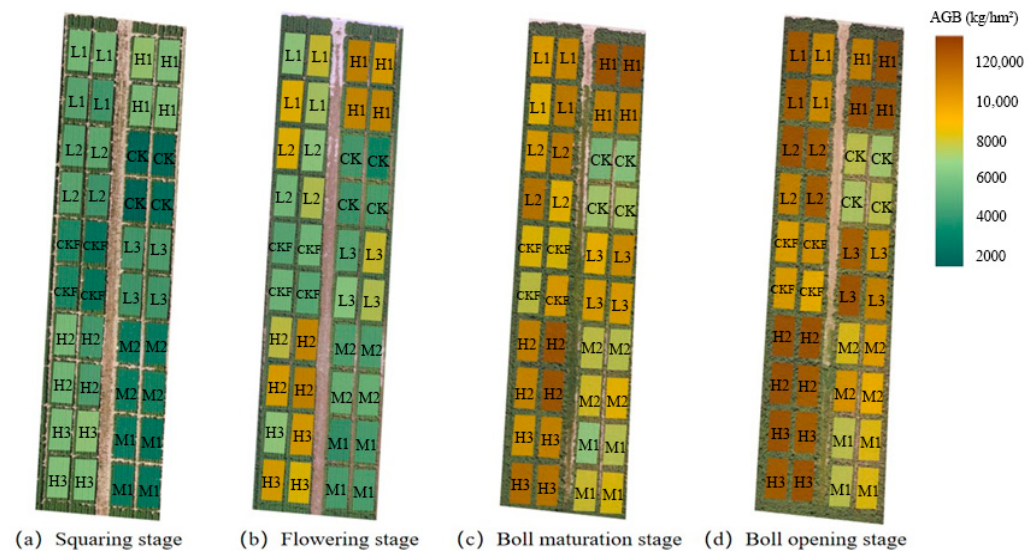


Figure 12. PSO-XGBoost-based CGSI_{VI} estimation of spatial and temporal distribution of AGB under different organic fertilizer treatments.

4. Discussion

The multispectral vegetation index, derived from spectral reflectance across multiple bands through specific algorithms, serves as a critical parameter for monitoring crop growth. However, its accuracy in biomass estimation is often compromised by external environmental factors such as crop growth status and soil conditions [41]. To address this limitation, integrating additional data sources, such as SPAD values, has been proposed to mitigate sensor bias and enhance estimation reliability. SPAD values, which reflect chlorophyll content and photosynthetic activity, provide a robust proxy for crop growth [42]. In this study, the newly developed CGSI_{VI}, which incorporates SPAD values, demonstrated a significantly stronger correlation with cotton AGB across all growth stages compared to traditional multispectral vegetation indices. The coefficient of determination (R^2) between the CGSI_{VI} and aboveground biomass (AGB) was 0.87, representing a 13.6% improvement compared to the highest R^2 value (0.72) of traditional vegetation indices (VIs). Taking the PSO-XGBoost model as an example, when the CGSI_{VI} was used as the input parameter, the relative root mean square error (rRMSE) for AGB estimation across the entire growth period was 11.97%, which was a 3.08% reduction compared to the rRMSE (12.35%) obtained with traditional VIs as inputs. The most significant improvement in estimation accuracy was observed at the flowering stage, where the R^2 value increased from 0.91 to 0.95. The correlation coefficient between the CGSI_{VI} and AGB exceeded 0.8, aligning with the findings from previous studies [37]. However, while prior research has primarily focused on the integration of spectral indices and environmental factors, this study highlights the unique contribution of SPAD values in enhancing biomass estimation accuracy. Unlike traditional vegetation indices, the CGSI_{VI} captures the dynamic interplay between chlorophyll content and biomass accumulation, offering a more comprehensive representation of crop growth dynamics. This distinction is particularly evident in the sensitivity of the CGSI_{VI} to biomass variations across different growth stages and environmental conditions, a feature less pronounced in conventional indices [43]. Furthermore, the close relationship between

SPAD values and photosynthetic efficiency enables the CGSI_{VI} to better explain biomass variability, a finding consistent with earlier studies [44]. This study further verified that the overall AGB estimation accuracy across different growth stages ranged from 0.768 to 0.949. This variation is closely linked to cotton's phenological characteristics: the flowering stage achieved the highest accuracy ($R^2 = 0.95$) due to the stable canopy structure and comprehensive synergistic information captured by the CGSI_{VI} (canopy structure from VIs and leaf physiology from SPAD). In contrast, lower accuracy at the squaring stage ($R^2 \approx 0.768$) is attributed to soil background interference, while the boll opening stage is affected by leaf senescence and uneven boll maturation, leading to moderate accuracy ($R^2 \approx 0.82$). These results align with prior crop biomass estimation studies, where inversion accuracy is highly dependent on crop growth stage-specific canopy traits [37,44]. Despite these advantages, the reliance on SPAD values introduces certain limitations, such as the need for high-quality, complete datasets and increased computational complexity due to the coefficient of variation method. These challenges highlight the importance of future research to optimize data collection and processing techniques, ensuring broader applicability and improved accuracy in diverse agricultural contexts.

Despite the excellent performance of the proposed CGSI_{VI} in cotton aboveground biomass (AGB) estimation, certain limitations in its practical application need to be clarified. The core limitation lies in its reliance on field-based manual SPAD measurements, which are time-consuming and labor-intensive, which conflicts with the goal of high-efficiency monitoring. However, the integration of SPAD data is indispensable for improving the accuracy and reliability of AGB estimation, and its advantages fully justify this trade-off: First, it compensates for the lack of physiological information in traditional remote sensing by supplementing chlorophyll-related physiological data to multispectral vegetation indices (VIs), facilitating the capture of subtle differences in cotton growth under complex fertilizer treatments; second, it enhances estimation performance under complex growth conditions, increasing the correlation between the CGSI_{VI} and AGB by 13.6% compared with traditional VIs (with a maximum $R^2 = 0.87$). Particularly in the middle and late growth stages, where canopy structure is complex and spectral saturation is prone to occur, it can distinguish biomass differences that cannot be identified by single VIs alone. Third, it improves the environmental adaptability of the model, reducing dependence on spectral reflectance and mitigating the impacts of external interferences.

5. Conclusions

Based on different fertilizer treatment experiments in the field, the ability of UAV multispectral data to improve the estimation of aboveground biomass in cotton was explored. An improved vegetation index, the CGSI_{VI}, for estimating the aboveground biomass of cotton was constructed. By analyzing the correlation between the multispectral vegetation index and aboveground biomass of cotton, it was found that the CGSI_{VI} model performed better than the original vegetation index, with the highest correlation between the CGSI_{Clre} and AGB. The overall estimation accuracy of AGB across different growth stages throughout the entire growth stage ranged from 0.768 to 0.949, peaking during the flowering stage. When the CGSI_{VI} was used as an input parameter in the comparison of different machine learning algorithms, the PSO-optimized XGBoost algorithm exhibited higher estimation accuracy both across the entire growth stage and within individual growth stages.

Author Contributions: G.W., M.H. and Y.Z. initiated and designed the study. G.W., M.H., Y.W., H.S., L.L. and K.Z. performed the experiments and collected the data. G.W., L.Z., C.L. and X.J. analyzed the data. G.W., M.H. and Y.Z. wrote and revised the manuscript. All authors contributed to the article and approved the submitted version. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Key Research and Development Program of Hebei Province (22326403D) and the Natural Science Foundation of Hebei Province (C2022204158).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Liu, B.; Asseng, S.; Wang, A.; Wang, S.; Tang, L.; Cao, W.; Zhu, Y.; Liu, L. Modelling the Effects of Post-Heading Heat Stress on Biomass Growth of Winter Wheat. *Agric. For. Meteorol.* **2017**, *247*, 476–490. [\[CrossRef\]](#)
2. Kharel, T.P.; Tyler, H.L.; Mubvumba, P.; Huang, Y.; Bhandari, A.B.; Fletcher, R.S.; Anapalli, S.; Joshi, D.R.; Mengistu, A.; Birru, G.; et al. Machine Learning on Multi-spectral Imagery to Estimate Nutrient Yield of Mixed-species Cover Crops. *Agric. Environ. Lett.* **2025**, *10*, e70009. [\[CrossRef\]](#)
3. Anees, S.A.; Mehmood, K.; Khan, W.R.; Sajjad, M.; Alahmadi, T.A.; Alharbi, S.A.; Luo, M. Integration of Machine Learning and Remote Sensing for above Ground Biomass Estimation through Landsat-9 and Field Data in Temperate Forests of the Himalayan Region. *Ecol. Inform.* **2024**, *82*, 102732. [\[CrossRef\]](#)
4. Zhang, P.; Lu, B.; Ge, J.; Wang, X.; Yang, Y.; Shang, J.; La, Z.; Zang, H.; Zeng, Z. Using UAV-Based Multispectral and RGB Imagery to Monitor above-Ground Biomass of Oat-Based Diversified Cropping. *Eur. J. Agron.* **2025**, *162*, 127422. [\[CrossRef\]](#)
5. David, R.M.; Rosser, N.J.; Donoghue, D.N.M. Improving above Ground Biomass Estimates of Southern Africa Dryland Forests by Combining Sentinel-1 SAR and Sentinel-2 Multispectral Imagery. *Remote Sens. Environ.* **2022**, *282*, 113232. [\[CrossRef\]](#)
6. Houborg, R.; McCabe, M. High-Resolution NDVI from Planet's Constellation of Earth Observing Nano-Satellites: A New Data Source for Precision Agriculture. *Remote Sens.* **2016**, *8*, 768. [\[CrossRef\]](#)
7. Shen, S.; Li, W.; Zou, J.; Harrison, M.T.; Liu, S.; Eyshi Rezaei, E.; Liu, K.; Kamali, B.; Wang, Z.; Zhang, D.; et al. Fusing UAV Multiple Data and Phenology to Predict Crop Biomass. *Inf. Process. Agric.* **2025**, S2214317325000460. [\[CrossRef\]](#)
8. Biswal, S.; Pathak, N.; Chatterjee, C.; Mailapalli, D.R. Estimation of Aboveground Biomass from Spectral and Textural Characteristics of Paddy Crop Using UAV-Multispectral Images and Machine Learning Techniques. *Geocarto Int.* **2024**, *39*, 2364725. [\[CrossRef\]](#)
9. Uddling, J.; Gelang-Alfredsson, J.; Piikki, K.; Pleijel, H. Evaluating the Relationship Between Leaf Chlorophyll Concentration and SPAD-502 Chlorophyll Meter Readings. *Photosynth. Res.* **2007**, *91*, 37–46. [\[CrossRef\]](#)
10. Yu, D.; Zha, Y.; Sun, Z.; Li, J.; Jin, X.; Zhu, W.; Bian, J.; Ma, L.; Zeng, Y.; Su, Z. Deep Convolutional Neural Networks for Estimating Maize Above-Ground Biomass Using Multi-Source UAV Images: A Comparison with Traditional Machine Learning Algorithms. *Precis. Agric.* **2023**, *24*, 92–113. [\[CrossRef\]](#)
11. Zhai, W.; Li, C.; Cheng, Q.; Mao, B.; Li, Z.; Li, Y.; Ding, F.; Qin, S.; Fei, S.; Chen, Z. Enhancing Wheat Above-Ground Biomass Estimation Using UAV RGB Images and Machine Learning: Multi-Feature Combinations, Flight Height, and Algorithm Implications. *Remote Sens.* **2023**, *15*, 3653. [\[CrossRef\]](#)
12. Allu, A.R.; Mesapam, S. Crop Health Assessment from Predicted AGB and NPK Derived from UAV Spectral Indices and Machine Learning Techniques. *Agronomy* **2025**, *15*, 2059. [\[CrossRef\]](#)
13. Ramesh, V.; Kumaresan, P. Stacked Ensemble Model for Accurate Crop Yield Prediction Using Machine Learning Techniques. *Environ. Res. Commun.* **2025**, *7*, 035006. [\[CrossRef\]](#)
14. Zhang, Y.; Xia, C.; Zhang, X.; Cheng, X.; Feng, G.; Wang, Y.; Gao, Q. Estimating the Maize Biomass by Crop Height and Narrowband Vegetation Indices Derived from UAV-Based Hyperspectral Images. *Ecol. Indic.* **2021**, *129*, 107985. [\[CrossRef\]](#)
15. Han, Y.; Tang, R.; Liao, Z.; Zhai, B.; Fan, J. A Novel Hybrid GOA-XGB Model for Estimating Wheat Aboveground Biomass Using UAV-Based Multispectral Vegetation Indices. *Remote Sens.* **2022**, *14*, 3506. [\[CrossRef\]](#)
16. Groten, S.M.E. NDVI—Crop Monitoring and Early Yield Assessment of Burkina Faso. *Int. J. Remote Sens.* **1993**, *14*, 1495–1515. [\[CrossRef\]](#)
17. Wiegand, C.L.; Richardson, A.J.; Escobar, D.E.; Gerbermann, A.H. Vegetation Indices in Crop Assessments. *Remote Sens. Environ.* **1991**, *35*, 105–119. [\[CrossRef\]](#)

18. Potopová, V.; Trnka, M.; Hamouz, P.; Soukup, J.; Castraveţ, T. Statistical Modelling of Drought-Related Yield Losses Using Soil Moisture-Vegetation Remote Sensing and Multiscalar Indices in the South-Eastern Europe. *Agric. Water Manag.* **2020**, *236*, 106168. [\[CrossRef\]](#)
19. Carpenter, G.D. GEMI: Fostering Environmental Excellence by Business World-Wide. *Environ. Conserv.* **1991**, *18*, 181–182. [\[CrossRef\]](#)
20. Bausch, W.C. Soil Background Effects on Reflectance-Based Crop Coefficients for Corn. *Remote Sens. Environ.* **1993**, *46*, 213–222. [\[CrossRef\]](#)
21. Begue, A.; Myneni, R. Operational NOAA Vegetation Indices-Absorbed PAR Relationships for Sahelian Vegetation Canopies. In Proceedings of the 1995 International Geoscience and Remote Sensing Symposium, IGARSS '95. Quantitative Remote Sensing for Science and Applications, Firenze, Italy, 10–14 July 1995; Volume 2, pp. 1296–1298.
22. Baret, F.; Guyot, G.; Major, D. TSAVI: A Vegetation Index Which Minimizes Soil Brightness Effects on LAI and APAR Estimation. In Proceedings of the 12th Canadian Symposium on Remote Sensing Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 10–14 July 1989; pp. 1355–1358.
23. Badgley, G.; Anderegg, L.D.L.; Berry, J.A.; Field, C.B. Terrestrial Gross Primary Production: Using NIR_V to Scale from Site to Globe. *Glob. Change Biol.* **2019**, *25*, 3731–3740. [\[CrossRef\]](#)
24. Ismail, M. *Assessment of Genetic Improvement in Doubled Haploid Lines of Bread Wheat under Heat-Stressed Environment*; University of Gezira: Wad Medani, Sudan, 2007.
25. Wu, W. The Generalized Difference Vegetation Index (GDVI) for Dryland Characterization. *Remote Sens.* **2014**, *6*, 1211–1233. [\[CrossRef\]](#)
26. Daughtry, C. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [\[CrossRef\]](#)
27. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [\[CrossRef\]](#)
28. Leblon, B.; Guerif, M.; Baret, F. The Use of Remotely Sensed Data in Estimation of PAR Use Efficiency and Biomass Production of Flooded Rice. *Remote Sens. Environ.* **1991**, *38*, 147–158. [\[CrossRef\]](#)
29. Adam, E.; Mutanga, O.; Rugege, D. Multispectral and Hyperspectral Remote Sensing for Identification and Mapping of Wetland Vegetation: A Review. *Wetl. Ecol Manag.* **2010**, *18*, 281–296. [\[CrossRef\]](#)
30. Villa, P.; Stroppiana, D.; Fontanelli, G.; Azar, R.; Brivio, P. In-Season Mapping of Crop Type with Optical and X-Band SAR Data: A Classification Tree Approach Using Synoptic Seasonal Features. *Remote Sens.* **2015**, *7*, 12859–12886. [\[CrossRef\]](#)
31. Gillespie, A.R.; Kahle, A.B.; Walker, R.E. Color Enhancement of Highly Correlated Images. II. Channel Ratio and “Chromaticity” Transformation Techniques. *Remote Sens. Environ.* **1987**, *22*, 343–365. [\[CrossRef\]](#)
32. Chen, A.; Orlov-Levin, V.; Meron, M. Applying High-Resolution Visible-Channel Aerial Imaging of Crop Canopy to Precision Irrigation Management. *Agric. Water Manag.* **2019**, *216*, 196–205. [\[CrossRef\]](#)
33. Torres-Sánchez, J.; Peña, J.M.; De Castro, A.I.; López-Granados, F. Multi-Temporal Mapping of the Vegetation Fraction in Early-Season Wheat Fields Using Images from UAV. *Comput. Electron. Agric.* **2014**, *103*, 104–113. [\[CrossRef\]](#)
34. Grenzdörffer, G.J. Crop Height Determination with UAS Point Clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *XL-1*, 135–140. [\[CrossRef\]](#)
35. Fallahi, H.; Ebrahimnejad, S.; Sabouri, H.; Lakzadeh, I.; Ghasemi, M.; Kohkan, S. Evaluation of Genotype × Interaction Effects on Grain Yield of Barely Genotypes Using Additive Main Effects and Multiplicative Interactions (AMMI). *J. Crop Breed.* **2018**, *9*, 144–151. [\[CrossRef\]](#)
36. O'Brien, R.M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Qual. Quant.* **2007**, *41*, 673–690. [\[CrossRef\]](#)
37. Zhai, W.; Li, C.; Fei, S.; Liu, Y.; Ding, F.; Cheng, Q.; Chen, Z. CatBoost Algorithm for Estimating Maize Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Multi-Source Sensor Data and SPAD Values. *Comput. Electron. Agric.* **2023**, *214*, 108306. [\[CrossRef\]](#)
38. Su, H.; Shen, W.; Wang, J.; Ali, A.; Li, M. Machine Learning and Geostatistical Approaches for Estimating Aboveground Biomass in Chinese Subtropical Forests. *For. Ecosyst.* **2020**, *7*, 64. [\[CrossRef\]](#)
39. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased Boosting with Categorical Features. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 3–8 December 2018.
40. Kennedy, J.; Eberhart, R. Particle Swarm Optimization. In Proceedings of the ICNN'95—International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
41. Fei, S.; Hassan, M.A.; Xiao, Y.; Su, X.; Chen, Z.; Cheng, Q.; Duan, F.; Chen, R.; Ma, Y. UAV-Based Multi-Sensor Data Fusion and Machine Learning Algorithm for Yield Prediction in Wheat. *Precis. Agric.* **2023**, *24*, 187–212. [\[CrossRef\]](#)
42. Islam, M.R.; Haque, K.; Akter, N.; Karim, M.A. Leaf Chlorophyll Dynamics in Wheat Based on SPAD Meter Reading and Its Relationship with Grain Yield. *Sci. Agric.* **2014**, *8*, 13–18.

43. Liu, Y.; Xiao, J.; Ju, W.; Zhu, G.; Wu, X.; Fan, W.; Li, D.; Zhou, Y. Satellite-Derived LAI Products Exhibit Large Discrepancies and Can Lead to Substantial Uncertainty in Simulated Carbon and Water Fluxes. *Remote Sens. Environ.* **2018**, *206*, 174–188. [[CrossRef](#)]
44. Yoshida, H.; Horie, T. A Process Model for Explaining Genotypic and Environmental Variation in Growth and Yield of Rice Based on Measured Plant N Accumulation. *Field Crops Res.* **2009**, *113*, 227–237. [[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.