

Article

Trait-Dependent Effects of Band Selection on Predicting Soybean Biomass, Leaf Area Index, and Canopy Cover from Hyperspectral Reflectance

Etsushi Kumagai ^{1,*}, Takayuki Yabiku ², Yusuke Masuya ^{2,3}, Kensuke Kimura ¹, Erina Fushimi ¹ and Ryosuke Nomiya ⁴

¹ Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization (NARO), Tsukuba, Ibaraki 305-8604, Japan

² Tohoku Agricultural Research Center, National Agriculture and Food Research Organization (NARO), Morioka, Iwate 020-0198, Japan

³ Faculty of Agriculture, Iwate University, Morioka, Iwate 020-8550, Japan

⁴ Kyushu-Okinawa Agricultural Research Center, National Agriculture and Food Research Organization (NARO), Chikugo, Fukuoka 833-0041, Japan

* Correspondence: kumagai.etsushi327@naro.go.jp; Tel.: +81-29-838-8856

Highlights

What are the main findings?

- The effectiveness of band selection depended on the soybean trait. Full-band partial least squares regression (PLS) model performed best for aboveground biomass (AGB), whereas selected-band models maintained comparable accuracy for leaf area index (LAI) and canopy cover (CC).
- Bootstrapped least absolute shrinkage and selection operator (LASSO) (BoLASSO) and the ensemble method selected bands across the visible, red-edge, and near-infrared regions and performed well, particularly for LAI and CC.

What are the implications of the main findings?

- Band selection is not equally effective for all canopy traits and should be tailored to the target trait.
- The selected wavelengths indicate candidate spectral regions for simplified multi-spectral sensing, although further validation is needed for actual broad-band sensor configurations and year-to-year transferability.

Abstract

Predicting canopy traits non-destructively is important for understanding crop growth and improving phenotyping efficiency. Hyperspectral reflectance provides detailed spectral information, but the role of band selection in regression-based trait prediction at the canopy scale remains unclear. In this study, we evaluated the effects of different band-selection algorithms on the prediction accuracy of aboveground biomass (AGB), leaf area index (LAI), and canopy cover (CC) in soybeans using canopy hyperspectral reflectance in the visible to near-infrared (VNIR) range from 501 to 801 nm. The dataset included multiple sites, years, cultivars, and irrigation treatments. We compared a full-band partial least squares regression (PLS) model with three band-selection methods (PLS-Variable Importance in Projection (VIP), Bootstrapped least absolute shrinkage and selection operator (LASSO) (BoLASSO), and an ensemble approach). Model performance was assessed using Kennard–Stone validation and leave-one-year-out cross-validation. The results

Academic Editor: Jaromir Krzyszczyk

Received: 28 May 2026

Revised: 29 June 2026

Accepted: 30 June 2026

Published: 3 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

showed that the effectiveness of band selection depended on the target trait. Full-band PLS performed well for AGB under Kennard–Stone validation, whereas BoLASSO achieved comparable accuracy to PLS for LAI and CC using a reduced number of selected bands. Leave-one-year-out cross-validation showed that year-to-year transferability was more difficult for AGB than for LAI and CC. The selected wavelengths were located mainly in the visible, red-edge, and near-infrared regions. These results indicate that band-selection strategies should be tailored to the target trait and that selected VNIR bands can provide candidate spectral regions for simplified sensing of soybean canopy traits.

Keywords: aboveground biomass; band selection; canopy cover; hyperspectral reflectance; leaf area index; soybean

1. Introduction

Soybeans (*Glycine max* (L.) Merr.) are a globally important crop and a major source of plant-based protein and vegetable oil. They are a staple crop in many countries, including Japan. Advances in breeding and cultivation techniques have led to increased seed yields in many regions [1–3]. From a physiological perspective, yield potential under optimal environmental conditions is determined by the radiation interception efficiency (RIE) of the crop canopy, the radiation use efficiency (RUE) that converts this into dry matter, and the harvest index, which indicates the proportion of dry matter allocated to seeds [4]. Recent increases in yield have been primarily associated with increased production of aboveground dry matter [5]. Leaf Area Index (LAI) is closely related to the light interception of the canopy and influences yield formation through RIE [4]. Therefore, monitoring traits such as aboveground biomass (AGB) and LAI over time provides foundational information for the development of high-yielding cultivars.

To elucidate yield formation mechanisms and establish technologies for stable and high yields, it is essential to quantify canopy-level traits. To do so, AGB, LAI, and RIE need to be estimated simultaneously, non-destructively, and rapidly [4,6]. Conventionally, direct measurement methods based on field surveys have been employed. With these methods, plants within a defined area are harvested. AGB is then determined through drying and weighing, and LAI is calculated based on leaf area measurements. RUE is estimated from the relationship between cumulative intercepted radiation and AGB [4,5]. These approaches are highly reliable and have been widely used as standard methods in eco-physiological research and breeding evaluation [5,7,8].

Direct measurement methods, however, are time-consuming and labor-intensive [4,6]. Time-series measurements of AGB require repeated harvesting, drying, and weighing, and the workload increases with the number of samples. Moreover, destructive sampling prevents repeated measurements on the same plants and may affect final yield evaluation. For LAI, both destructive methods [9] and non-destructive approaches such as plant canopy analyzers (e.g., LAI-2200) or aerial photography [10] have been developed, but limitations remain for large-scale, time-series measurements. In soybeans, rapid canopy development further complicates high-throughput evaluation across multiple sites and genotypes [11,12].

In recent years, high-throughput phenotyping technologies have been developed to acquire physiological and agronomic traits non-destructively and at high frequency [11–14]. Advances in proximal remote sensing, particularly the integration of sensor technologies with data science, have enabled cost-effective phenotyping. Among these, hyperspectral reflectance (HSR) can be used to quantify the vegetation reflectance across a wide

spectral range using hundreds of wavelength bands and to capture the biochemical and structural characteristics of the plant canopy [14–20].

Beyond plot-scale phenotyping, remote sensing has also been increasingly used for large-scale agricultural monitoring, including crop mapping, cropland parcel extraction, and crop planting structure mapping. In these applications, spectral bands, texture features, and temporal features are often combined, and feature selection or feature fusion is used to improve mapping accuracy and model transferability [21,22]. Although these studies focus mainly on classification and mapping, they share a common methodological issue with regression-based trait estimation: identifying informative and non-redundant features from high-dimensional remote sensing data. Therefore, feature-selection findings from mapping studies are also relevant to trait estimation using spectral reflectance.

HSR can be broadly used in two ways: (1) to construct regression models using the full reflectance spectrum as explanatory variables (e.g., partial least squares regression, PLS [23]) and (2) to select spectral regions (features) relevant to the target trait [24]. The former leverages the full spectral signal, whereas the latter includes approaches such as vegetation indices (VIs) based on specific band relationships. These approaches have been applied to estimate various traits, including AGB, LAI, canopy cover (CC), leaf photosynthetic capacity, yield, and nitrogen content [14,17,25,26].

Since HSR data consist of many bands, band selection is not only a means of dimensionality reduction and computational efficiency, but also a key factor influencing predictive performance, generalization, and interpretability [24]. The objectives of band selection can be summarized as: (1) improving the model fit and prediction accuracy [27,28]; (2) enhancing the sensitivity and robustness of vegetation indices for trait prediction [24,29,30]; and (3) improving interpretability by linking spectral features to underlying physiological and molecular processes [31]. Focusing on a subset of bands has been suggested to clarify the relationship between spectral information and traits and improve model applicability across datasets [25,32].

On the other hand, HSR data often exhibit strong correlations among adjacent bands and a limited number of samples relative to the number of variables, which can lead to overfitting and reduced generalization [25,33]. These issues are particularly important in regression tasks [33]. Furthermore, full-band models are sensitive to sensor characteristics and observation conditions, and they may be affected by inter-instrument variations and environmental fluctuations. Practical constraints also exist because hyperspectral sensors are costly and difficult to deploy.

Selecting a limited number of informative bands is considered an effective strategy to improve both the model generalization and practical applicability. However, band-selection results are often dependent on the dataset and analytical procedures, and the selected bands may be unstable [25]. Stepwise regression has been criticized for producing unstable and unreliable band selections [25,31]. Indeed, previous studies have reported a lack of consensus on optimal band sets for chlorophyll estimation [34,35].

Given that no single method has been found to be consistently superior to others [25,35], reliance on a single method is not justified, and frameworks that include inter-method comparisons and stability assessments are required. Ensemble approaches, which combine multiple methods, have been shown to improve prediction accuracy in remote sensing applications [36–39]. In HSR studies, integrating or stacking multiple machine learning models can enhance prediction accuracy [40]. However, whether such multi-method ensembles enable robust band selection remains unclear [25].

Existing band-selection studies have developed primarily around classification tasks [41], and few frameworks simultaneously evaluate stability, generalization performance, and the interpretability of selected bands in regression tasks involving continuous variables at the canopy scale. Previous reviews have pointed out that redundancy among ad-

jacent bands, limited sample size, and differences in study design (e.g., band range and data composition) hinder the evaluation of generalizability [41].

This study aimed to clarify the effects of different band-selection algorithms on the prediction accuracy of AGB, LAI, and CC using canopy hyperspectral reflectance in the visible to near-infrared (VNIR) range from 501 to 801 nm for multiple soybean cultivars across different sites, years, cultivars, and irrigation treatments. We specifically tested the hypothesis that using selected VNIR bands can achieve prediction accuracy comparable to that of full-band VNIR models under varying observation conditions. In addition, we examined the characteristics of the selected bands to explore their implications for efficient band design and practical applications in remote sensing.

2. Materials and Methods

2.1. Study Sites and Experimental Design

Experiments were conducted over five years (2019, 2020, 2021, 2023, and 2024) at three research centers in Japan: the Tohoku Agricultural Research Center (TARC; 39.74°N, 141.13°E) in Morioka, Iwate; the Institute for Agro-Environmental Sciences (NIAES; 36.02°N, 140.11°E) in Tsukuba, Ibaraki; and the Kyushu-Okinawa Agricultural Research Center (KARC; 33.21°N, 130.49°E) in Chikugo, Fukuoka. Table S1 summarizes the year, site, cultivar, number of replicates, water management conditions, and the number of observations for day of year (DOY), AGB, LAI and CC for each experiment.

Meteorological stations were installed at each site to monitor air temperature, solar radiation, and precipitation. During the growing season (June–October), mean air temperature ranged from 19.5 to 26.6 °C, solar radiation from 14.4 to 17.7 MJ m⁻² d⁻¹, and precipitation from 601 to 1088 mm, with variation among sites and years (Table S2).

At TARC, the soil in the experimental plots was Andosol. In each year, 3 g m⁻² of nitrogen, 12.5 g m⁻² of phosphorus (as P₂O₅), and 5 g m⁻² of potassium (as K₂O) were applied prior to sowing and incorporated into the soil by rotary tillage. The number of cultivars tested was nine in 2019 (Exp. ID 1) and 2020 (Exp. ID 2), six in 2021 (Exp. ID 3), and two in 2023 (Exp. ID 4) and 2024 (Exp. ID 8). Sowing dates were 5 June 2019; 4 June 2020; 8 June 2021; 30 May 2023; and 18 June 2024. All sowing was performed by hand. Seeds were treated with a mixture of insecticide and fungicide (CruiserMaxx, Syngenta Co., Tokyo, Japan) prior to sowing. Immediately after sowing, the prescribed amount of herbicide (Ecotop, Maruwa Biochemical, Tokyo, Japan) was applied, and insecticides and fungicides were sprayed as needed. The planting pattern was 75 cm between rows and 15 cm between plants in all years. The experiment was conducted with three replicates in 2019 and 2020, and two replicates in 2021, 2023, and 2024. The experimental plots consisted of six rows, each 4.8 m long, in 2019 and 2020; nine rows, each 3.0 m long, in 2021; and seven rows, each 3.6 m long, in 2023 and 2024. In 2019, 2020, and 2021, the crops were grown under rain-fed conditions. In 2023, irrigation was applied using irrigation tubes, and three levels of soil moisture—high, medium, and low—were established based on the distance from the tubes. In 2024, irrigation was also performed using irrigation tubes, and two levels of soil moisture—high and low—were established.

At NIAES, the soil in the experimental plots was either Andosol (Exp. IDs 5, 6, and 10) or Typic Endoaquept (Exp. ID 9). In each year, 1.8 g m⁻² of nitrogen, 14 g m⁻² of phosphorus (as P₂O₅), and 6 g m⁻² of potassium (as K₂O) were applied before sowing and incorporated by rotary tillage. The number of test cultivars was two for Exp. ID 5 and four for Exp. ID 6 in 2023, and two for Exp. ID 9 and four for Exp. ID 10 in 2024. Sowing dates were 13 June 2023, and 11 June 2024, and all sowing was done by hand. Seeds were treated with a combined insecticide and fungicide (CruiserMaxx) prior to sowing. Immediately after sowing, the herbicide (Ecotop) was applied at the recommended rate, and insecticides and fungicides were sprayed as needed. The planting pattern was 70 cm between

rows and 15 cm between plants in all years. The experiment was conducted with two replicates for Exp. ID 5 and three replicates for Exp. ID 6 in 2023, and two replicates for Exp. ID 9 and three replicates for Exp. ID 10 in 2024. The experimental plots consisted of six rows (7.2 m in length) for Exp. IDs 5 and 9, and six rows (2.5 m) for Exp. IDs 6 and 10. In Exp. ID 5 in 2023, treatments included a rain-fed control, an early-flooding treatment (July 11–13), a late-flooding treatment (July 22–28), and a rain-shielded treatment (July 24–August 6) with rain-fed conditions outside the treatment periods. In Exp. ID 9 (2024), five treatments were established: rain-fed control, early flooding (July 10–16), late flooding (July 26–August 1), combined early and late flooding, and the rain shelter treatment (July 25–August 14).

At KARC, the soil in the experimental plots was Typic Endoaquept. In 2023, no fertilizer was applied. In 2024, 2.4 g m⁻² of nitrogen, 0.9 g m⁻² of phosphorus (as P₂O₅), and 0.2 g m⁻² of potassium (as K₂O) were applied before sowing and incorporated into the soil by rotary tillage. Two cultivars were tested in both 2023 and 2024. Sowing dates were 15 July 2023 and 17 July 2024. Sowing was performed by hand. Seeds were treated with a mixed insecticide and fungicide (CruiserMaxx) prior to sowing. Immediately after sowing, herbicide (Creaturn, Kumiai Chemical Industry Co., Ltd., Tokyo, Japan) was applied at the recommended rate, and insecticides and fungicides were sprayed as needed. The planting pattern was 70 cm between rows and 20 cm between plants. The experiment was conducted with two replicates in both years. The experimental plots consisted of four rows (4.2 m in length) in 2023 and seven rows (1.0 m) in 2024. In 2023, rain-fed (control) and tube-irrigated plots were established. In 2024, a lysimeter equipped with a rain shelter was used, with four treatments: a control with a constant groundwater level of 35 cm, a dry treatment with the water table set at 90 cm (August 19–September 2), an early-stage waterlogged treatment (0 cm; August 2–9), and a late-stage waterlogged treatment (0 cm; August 19–26).

2.2. Ground Truth Data for AGB, LAI, and CC

At TARC, destructive surveys to measure AGB and LAI were conducted four times in 2019 (Exp. ID 1) and 2020 (Exp. ID 2), twice in 2021 (Exp. ID 3), three times in 2023 (Exp. ID 4), and twice in 2024 (Exp. ID 8) (Table S1). In Exp. IDs 1–3, 0.68 m² was harvested from three rows in each plot. In Exp. IDs 4 and 8, 0.45 m² was harvested from one row in each plot.

At NIAES, destructive surveys were conducted once for Exp. ID 5 and six times for Exp. ID 6 in 2023, and once for Exp. ID 9 and nine times for Exp. ID 10 in 2024. In Exp. IDs 5 and 9, 0.42 m² was harvested from one row in each plot, whereas in Exp. IDs 6 and 10, 0.84 m² was harvested from four rows in each plot.

At KARC, destructive surveys were conducted once in each of 2023 (Exp. ID 7) and 2024 (Exp. ID 11). In Exp. ID 7, 0.56 m² was harvested from one row in each plot, and in Exp. ID 11, 0.84 m² was harvested from two rows in each plot.

As described in Section 2.3, because the sensor footprint was approximately 0.35 m², all sampling areas sufficiently covered the measurement area.

Leaves were separated from the collected plant material, and leaf area was measured. At TARC, an automatic leaf area meter (AAM-8, Hayashidenko Co., Ltd., Tokyo, Japan) was used, whereas at NIAES and KARC, a leaf area meter (LI-3100C, LI-COR, Lincoln, NE, USA) was used. LAI was calculated by dividing leaf area by sampling area. After drying the plant samples at 80 °C for three days, total dry weight was measured and converted to AGB based on the sampling area.

Fractional canopy cover (CC) was estimated using digital image analysis at TARC and NIAES. Images were captured using a camera (GACS1; Kimura Ouyou-Kougei Co., Ltd., Osaka, Japan) mounted at a height of 1.5 m above ground level and pointing directly

downward (nadir). CC was calculated according to the method of Kumagai and Takahashi [42]. Imaging was conducted at time points corresponding to the destructive surveys. CC data were not available for all experiments; CC was not measured in the 2021 TARC experiment or in the 2023 and 2024 KARC experiments (Table S1). As a result, the number of CC observations was 390, whereas the numbers of AGB and LAI samples were 472 and 462, respectively. Nevertheless, the CC dataset included observations from four years, two sites, multiple cultivars and water management treatments, and sampling dates from early to late canopy development.

2.3. Measurement of Canopy HSR

In this study, to obtain canopy-scale HSR at high frequency in experimental plots where multiple cultivars were grown, we developed a portable measurement system. Drawing on the method described by Kimm et al. [43], we adopted a single-fiber configuration to reduce weight and simplify the system. The configuration of the measurement system and the observation geometry are shown in Figure S1.

The Flame-S spectrometer (Ocean Optics, Orlando, FL, USA) was used for the measurements. This instrument measures wavelengths in the 501–801 nm range and has a spectral resolution of 0.92 nm (FWHM). A fiber with a core diameter of 600 μm (Ocean Optics) was used. Incident light was measured using a standard reflector (Spectralon, Labsphere Inc., North Sutton, NH, USA), and upward radiation from the canopy was measured at a field of view of 25°. Reflectance was calculated as the ratio of the two measurements.

The fiber tip was fixed approximately 1.5 m above the ground, and a level was used to confirm a zenith angle of 0° (nadir) during measurements. Three to five measurements were taken per plot, and the average value was used as the representative value. The measurement time per plot was less than one minute. Standard reflector measurements were performed before each plot measurement. Dark current was acquired by blocking the light path using an INLINE-TTL-S Electronic Shutter (Ocean Optics), and corrections were made with particular attention to fluctuations in solar irradiance and temperature. When solar irradiance changed, the integration time was adjusted using the standard reflector to avoid saturation of the DN values. Measurements were conducted between 10:00 and 15:00 on clear days, before and after the destructive surveys for AGB and LAI. In most cases, spectral measurements and destructive surveys were conducted on the same day; when weather conditions were not suitable for HSR measurements, the measurement date was shifted by up to one day before or after the destructive survey. This adjustment was made to avoid measurements under cloudy or unstable light conditions. The acquired spectra were recorded using OceanView software 1.6.7 (Ocean Optics) and smoothed using a Savitzky–Golay filter (third-order polynomial, 9-point moving window) [44]. The Savitzky–Golay filter was applied to reduce high-frequency noise while preserving the overall shape of the reflectance spectra. Because the spectra were converted to 1-nm wavelength intervals, the 9-point window corresponded to an approximately 9-nm smoothing range. This setting was used as a relatively conservative smoothing procedure to reduce measurement noise without strongly altering broad spectral features in the VNIR region. The resulting reflectance data from 501 to 801 nm were used for subsequent analyses.

2.4. Model Development and Validation

Machine learning regression models were applied to predict AGB, LAI, and CC using visible-near-infrared (VNIR) spectral reflectance data (301 bands, 501–801 nm) at the canopy scale. A total of 472, 462, and 390 samples were collected for AGB, LAI, and CC, respectively (Table S1). All retained extreme values were checked against the original records and were confirmed to be genuine measurements, with no negative or physically implausible values remaining in the dataset.

The dataset was divided into training and validation sets in a 2:1 ratio using the Kennard–Stone algorithm [45], which selects representative samples to evenly cover the data space based on distances in multivariate space. The Kennard–Stone split was performed using the spectral reflectance data only. All model fitting, hyperparameter optimization, band selection, and variable reduction were performed using the training dataset, and the validation dataset was used only for model evaluation.

Taking practical applications into account, the number of selected bands used as explanatory variables in the final model was limited to 10. This was intended to enable implementation using 10-band multispectral cameras (e.g., the MicaSense dual camera system) mounted on commercially available UAVs. This constraint was used as a practical upper limit because increasing the number of bands generally increases sensor complexity, data volume, and operational cost in UAV-based monitoring. In this study, we compared PLS using all bands; variable importance in projection (VIP), which extracts bands with high contribution from the PLS model (PLS-VIP); bootstrapped least absolute shrinkage and selection operator (LASSO) (BoLASSO), which mitigates the selection instability of LASSO using bootstrapping; and an ensemble method that integrates band importance based on multiple regression models. In addition, conventional vegetation index models were evaluated as benchmark models using the normalized difference vegetation index (NDVI; $[R800 - R670]/[R800 + R670]$), green normalized difference vegetation index (GNDVI; $[R750 - R550]/[R750 + R550]$), normalized difference red-edge index (NDRE; $[R750 - R705]/[R750 + R705]$), and photochemical reflectance index (PRI; $[R531 - R570]/[R531 + R570]$) [46–49].

In PLS-VIP, we first standardized the explanatory and response variables and determined the optimal number of latent variables via cross-validation to construct a PLS model. Next, we calculated VIP scores for each band to evaluate their contribution to the model. The VIP score is an indicator calculated by integrating the score, weight, and relationship with the response variable for each latent variable; higher values indicate greater contributions to the prediction [23]. Subsequently, the top 10 candidate bands were selected based on the VIP scores, and unnecessary variables were removed using stepwise backward selection based on the Akaike information criterion (AIC) to determine the final combination of explanatory variables.

The workflows for BoLASSO and the ensemble method are shown in Figure 1. BoLASSO was applied with some modifications to the method proposed by Bach [50]. First, the explanatory variables and response variable in the training data were standardized, and the LASSO regularization parameter was determined using 5-fold cross-validation. Subsequently, 1000 bootstrap samples were generated, and LASSO regression was applied to each sample. For each band, the number of times the regression coefficient was non-zero was counted as the selection frequency, and the top 10 candidate bands with the highest selection frequencies were extracted. Using these 10 bands as candidate variables, unnecessary variables were removed using an AIC-based variable reduction to construct the final multiple regression model.

For the ensemble method, we referred to the approach by Feilhauer et al. [24]. Band importance was calculated using four methods: PLS, Ridge regression, LASSO, and linear support vector regression (LSVR). For each method, the absolute value of the standardized regression coefficient was used as the importance score. The importance scores from each method were weighted by the coefficient of determination (R^2) from cross-validation and summed across methods to calculate the ensemble importance (IE). The top 10 candidate bands were selected based on IE, and similar to BoLASSO, a final multiple regression model was constructed using an AIC-based variable reduction.

To evaluate the sensitivity of model performance to the number of candidate bands, additional analyses were conducted for BoLASSO and the ensemble method by varying

the number of candidate bands from 3 to 15. Model performance was compared between models using the top-k candidate bands directly and those using AIC-based variable reduction after candidate-band selection.

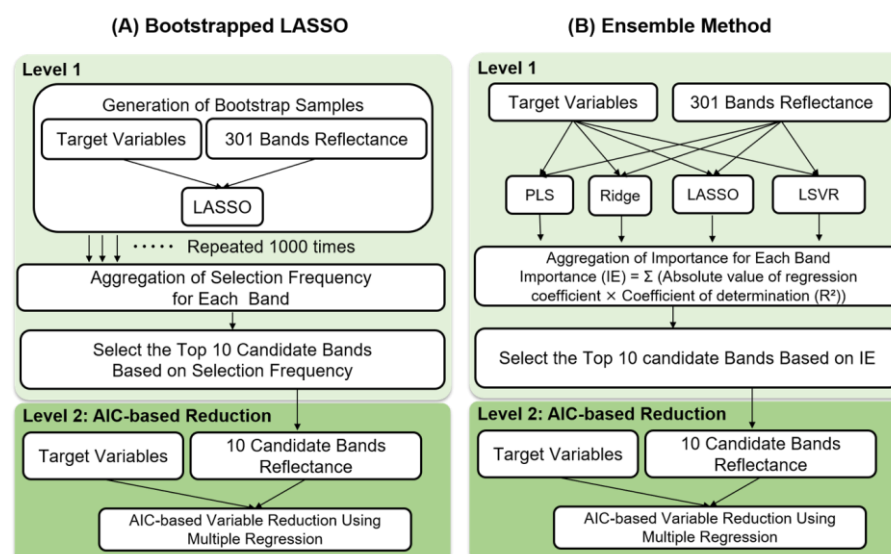


Figure 1. Workflow of band-selection methods used in this study. **(A)** Bootstrapped LASSO (BoLASSO): bootstrap resampling and LASSO fitting were repeated 1000 times to evaluate the selection frequency of each band, and the top 10 candidate bands were selected based on their selection frequency. The dotted line in the BoLASSO workflow indicates this repeated bootstrap–LASSO procedure. **(B)** Ensemble learning approach: band importance was calculated from multiple regression models (PLS, Ridge, LASSO, and LSVR) and aggregated as an importance index (IE). The top 10 candidate bands were selected based on IE. In both approaches, AIC-based variable reduction using multiple regression was then applied to the 10 candidate bands to construct the final selected-band model.

The prediction accuracy of the models was evaluated using R^2 and root mean square error (RMSE) for both the training and validation data. Furthermore, the normalized RMSE (nRMSE, %) was calculated by dividing the RMSE by the mean of the observed values and multiplying by 100. For the Kennard–Stone validation results, uncertainty in model performance was evaluated using bootstrap resampling of the validation dataset. Observed and predicted values were resampled as paired observations 2000 times, and R^2 , RMSE, and nRMSE were recalculated for each bootstrap sample. The 95% confidence intervals were obtained from the 2.5th and 97.5th percentiles of the bootstrap distributions.

As an additional test of year-to-year transferability, leave-one-year-out cross-validation was conducted. In this analysis, data from one year were used as the validation dataset, and data from the remaining years were used as the training dataset. Model fitting, hyperparameter optimization, band selection, and AIC-based variable reduction were repeated independently within each training dataset. Overall R^2 and nRMSE were calculated using all pooled predictions from the hold-out years, and fold-level R^2 and nRMSE values were used to evaluate variation among hold-out years.

The hyperparameters for each regression method were optimized using cross-validation. For PLS, the number of latent variables was selected by 5-fold cross-validation from the feasible range determined by the training data size and matrix rank. For BoLASSO, the LASSO regularization parameter was determined automatically using 5-fold cross-validation. For the ensemble method, the Ridge regularization parameter was searched over 2^{-5} to 2^9 , the LASSO regularization parameter was searched from 0.01 to

0.70 at intervals of 0.01, and the LSVR parameters C and ε were searched over 2^{-5} to 2^4 and 2^{-10} to 2^{-1} , respectively.

Data analysis and visualization were performed in Python 3.11.13 using Jupyter Notebook 4.1.6 and the scikit-learn 1.7.1, pandas 2.3.1, NumPy 2.3.2, statsmodels 0.14.5, SciPy 1.16.1, and Matplotlib 3.10.5.

3. Results

3.1. Data Characteristics

Figure 2 shows the summary of canopy spectral reflectance in the visible to near-infrared region (501–801 nm). Reflectance remained low in the visible region (500–700 nm), with strong absorption observed particularly in the red region. In contrast, reflectance increased sharply beyond approximately 700 nm, confirming the characteristic red-edge feature. The mean reflectance varied systematically with wavelength, and the standard deviation (SD) tended to increase in the near-infrared region. The ranges of maximum and minimum values also expanded in the near-infrared region, indicating that data variability was wavelength-dependent.

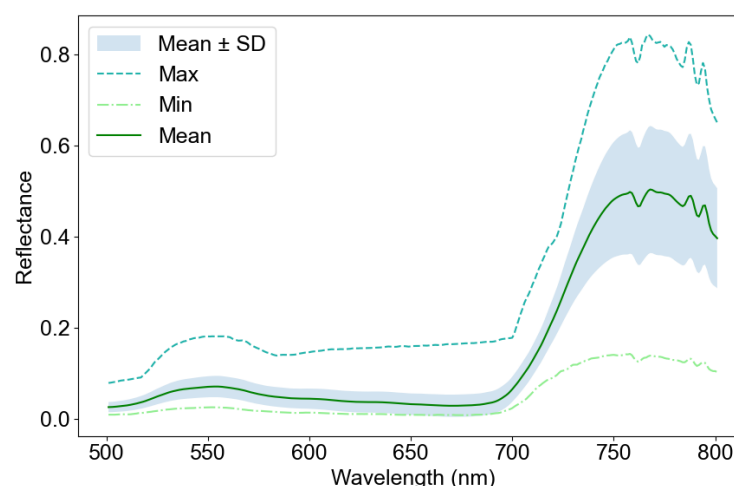


Figure 2. Canopy spectral reflectance across all observations used in this study in the visible to near-infrared range (501–801 nm). The solid line, shaded area, and dashed lines indicate the mean, mean $\pm$ standard deviation, and maximum and minimum values, respectively.

Table 1 shows the descriptive statistics for AGB, LAI, and CC. The dataset covered wide ranges of AGB, LAI, and CC, reflecting variation in growth stages and canopy structures across multiple years, sites, cultivars, and treatments. AGB and LAI showed relatively large variability, with CV values of 74% and 58%, respectively, whereas CC showed lower but still substantial variability, with a CV of 29%.

Table 1. Descriptive statistics for soybean traits (aboveground biomass, AGB; leaf area index, LAI; and canopy cover, CC) obtained across multiple years, sites, growth stages, and cultivars.

Target Variables	Mean	Max	Min	SD	CV
AGB (g m^{-2})	385	1239	8.34	287	74
LAI ($\text{m}^2 \text{m}^{-2}$)	3.72	10.21	0.12	2.17	58
CC (%)	80	100	5	24	29

3.2. Prediction Accuracy Under Kennard–Stone Validation

We compared the prediction accuracy of AGB, LAI and CC between PLS models using all bands and three types of band-selection models (PLS-VIP, BoLASSO, and the ensemble method) using the Kennard–Stone validation dataset (Tables 2–4; Figures 3–5). For reference, the corresponding relationships between observed and predicted values in the training dataset are shown in Figures S2–S4.

For AGB, the full-band PLS model showed the lowest nRMSE among the four methods, with $R^2 = 0.750$ and $\text{nRMSE} = 31.8\%$ (Table 2, Figure 3). Among the selected-band models, the ensemble method showed performance close to that of the full-band PLS model ($R^2 = 0.719$, $\text{nRMSE} = 33.7\%$). BoLASSO showed lower prediction accuracy ($R^2 = 0.582$, $\text{nRMSE} = 41.1\%$), whereas PLS-VIP showed the lowest accuracy ($R^2 = 0.384$, $\text{nRMSE} = 49.9\%$). These results indicate that reducing the number of bands tended to decrease the accuracy of AGB prediction, although the ensemble method retained relatively high prediction accuracy using seven selected bands.

For LAI, the full-band PLS model, BoLASSO, and the ensemble method showed similar validation accuracy (Table 3, Figure 4). The full-band PLS model showed $R^2 = 0.771$ and $\text{nRMSE} = 25.1\%$, while BoLASSO showed $R^2 = 0.763$ and $\text{nRMSE} = 25.5\%$. The ensemble method also showed comparable accuracy ($R^2 = 0.754$, $\text{nRMSE} = 25.9\%$). In contrast, PLS-VIP showed lower accuracy ($R^2 = 0.620$, $\text{nRMSE} = 32.3\%$). These results indicate that LAI could be predicted with similar accuracy using selected-band models, particularly BoLASSO and the ensemble method, compared with the full-band PLS model.

For CC, BoLASSO and the ensemble method achieved accuracy comparable to the full-band PLS model (Table 4, Figure 5). The full-band PLS model showed $R^2 = 0.883$ and $\text{nRMSE} = 8.6\%$, while BoLASSO and the ensemble method showed $R^2 = 0.887$ and 0.886 and $\text{nRMSE} = 8.4\%$ and 8.5% , respectively. In contrast, PLS-VIP showed lower accuracy ($R^2 = 0.673$, $\text{nRMSE} = 14.4\%$). These results indicate that CC prediction was less affected by band reduction than AGB prediction.

In summary, the effect of band selection differed among the target traits. For AGB, the full-band PLS model showed the lowest nRMSE, whereas BoLASSO and the ensemble method maintained accuracy comparable to full-band PLS for LAI and CC. In contrast, PLS-VIP tended to show lower accuracy for all three traits. Additional error analysis for high-value subsets showed that prediction errors increased particularly under high-AGB conditions, whereas errors for high LAI and high CC remained relatively low for PLS, BoLASSO, and the ensemble method (Table S3).

Table 2. Performance of regression models for predicting aboveground biomass (AGB) using canopy spectral reflectance data in the Kennard–Stone validation dataset and the corresponding selected wavelength bands.

Methods	PLS	PLS-VIP	BoLASSO	Ensemble
R^2	0.750 [0.685–0.802]	0.384 [0.282–0.468]	0.582 [0.492–0.656]	0.719 [0.650–0.776]
nRMSE (%)	31.8 [28.4–35.3]	49.9 [44.9–55.2]	41.1 [36.4–45.8]	33.7 [29.6–37.9]
Number of x variables	301	7	7	7
Selected wavelength bands (nm)	All bands	677, 679, 680, 683, 762, 793, 797	509, 520, 525, 674, 716, 720, 793	520, 535, 595, 692, 716, 720, 794

Values in square brackets indicate bootstrap 95% confidence intervals based on 2000 paired resampling iterations of observed and predicted values. R^2 , coefficient of determination; nRMSE, normalized root mean square error. The number of x variables indicates the final number of explanatory variables retained in each model. For PLS, all 301 wavelength bands were used. For the selected-band models, candidate bands were selected using each method and further refined by AIC-based variable reduction using multiple regression. All band selection, AIC-based variable reduction

tion, and model fitting procedures were performed using only the training dataset, and the validation dataset was used only for model evaluation.

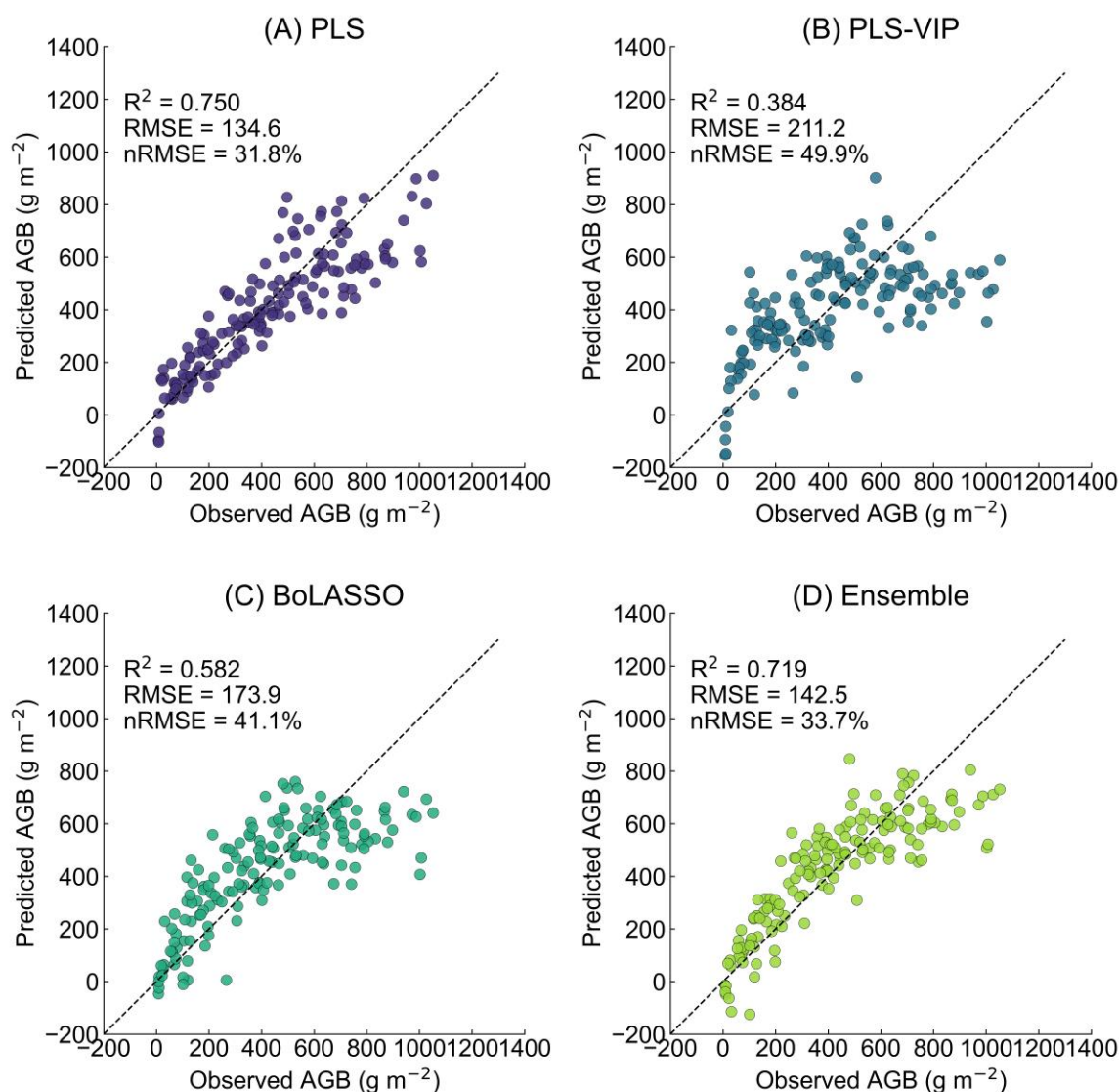


Figure 3. Relationship between observed and predicted aboveground biomass (AGB) in the Kennard–Stone validation dataset for four regression methods: (A) PLS, (B) PLS-VIP, (C) BoLASSO, and (D) Ensemble. The dashed line indicates the 1:1 relationship, and R^2 , RMSE, and nRMSE (%) are shown in each panel.

Table 3. Performance of regression models for predicting leaf area index (LAI) using canopy spectral reflectance data in the Kennard–Stone validation dataset and the corresponding selected wavelength bands.

Methods	PLS	PLS-VIP	BoLASSO	Ensemble
R^2	0.771 [0.694–0.836]	0.620 [0.517–0.708]	0.763 [0.685–0.829]	0.754 [0.676–0.821]
nRMSE (%)	25.1 [21.1–29.8]	32.3 [27.9–37.1]	25.5 [21.4–30.0]	25.9 [21.9–30.6]
Number of x variables	301	6	6	10
Selected wavelength bands (nm)	All bands	785, 790, 793, 796, 797, 798	501, 505, 595, 704, 729, 793	504, 694, 728, 729, 785, 786, 787, 792, 793, 794

See Table 2 for definitions of R^2 , nRMSE, confidence intervals, the number of x variables, and details of variable selection.

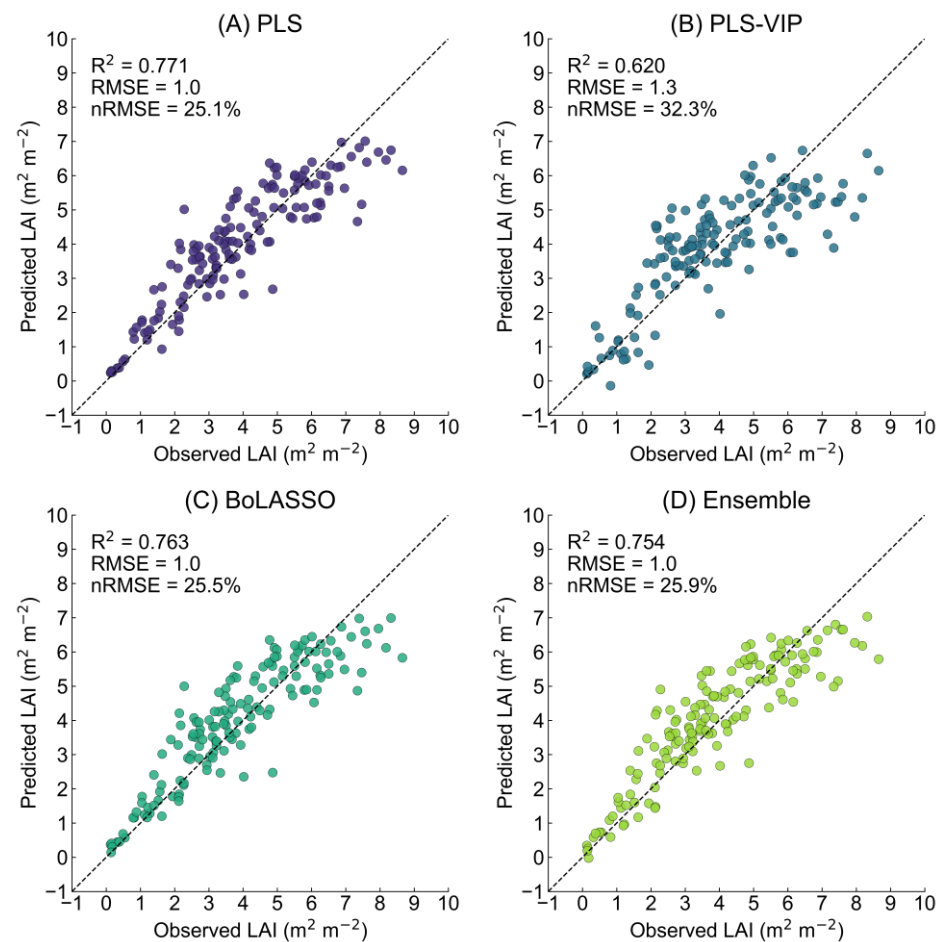


Figure 4. Relationship between observed and predicted leaf area index (LAI) in the Kennard–Stone validation dataset for four regression methods: (A) PLS, (B) PLS-VIP, (C) BoLASSO, and (D) Ensemble. The dashed line indicates the 1:1 relationship, and R^2 , RMSE, and nRMSE (%) are shown in each panel.

Table 4. Performance of regression models for predicting canopy cover (CC) using canopy spectral reflectance data in the Kennard–Stone validation dataset and the corresponding selected wavelength bands.

Methods	PLS	PLS-VIP	BoLASSO	Ensemble
R^2	0.883 [0.806–0.924]	0.673 [0.477–0.757]	0.887 [0.812–0.925]	0.886 [0.808–0.925]
nRMSE (%)	8.6 [7.2–10.1]	14.4 [12.3–16.6]	8.4 [7.1–9.9]	8.5 [7.2–9.9]
Number of x variables	301	4	6	9
Selected wavelength bands (nm)	All bands	797, 798, 800, 801	549, 599, 683, 693, 720, 762	536, 599, 683, 693, 719, 720, 721, 793, 797

See Table 2 for definitions of R^2 , nRMSE, confidence intervals, the number of x variables, and details of variable selection.

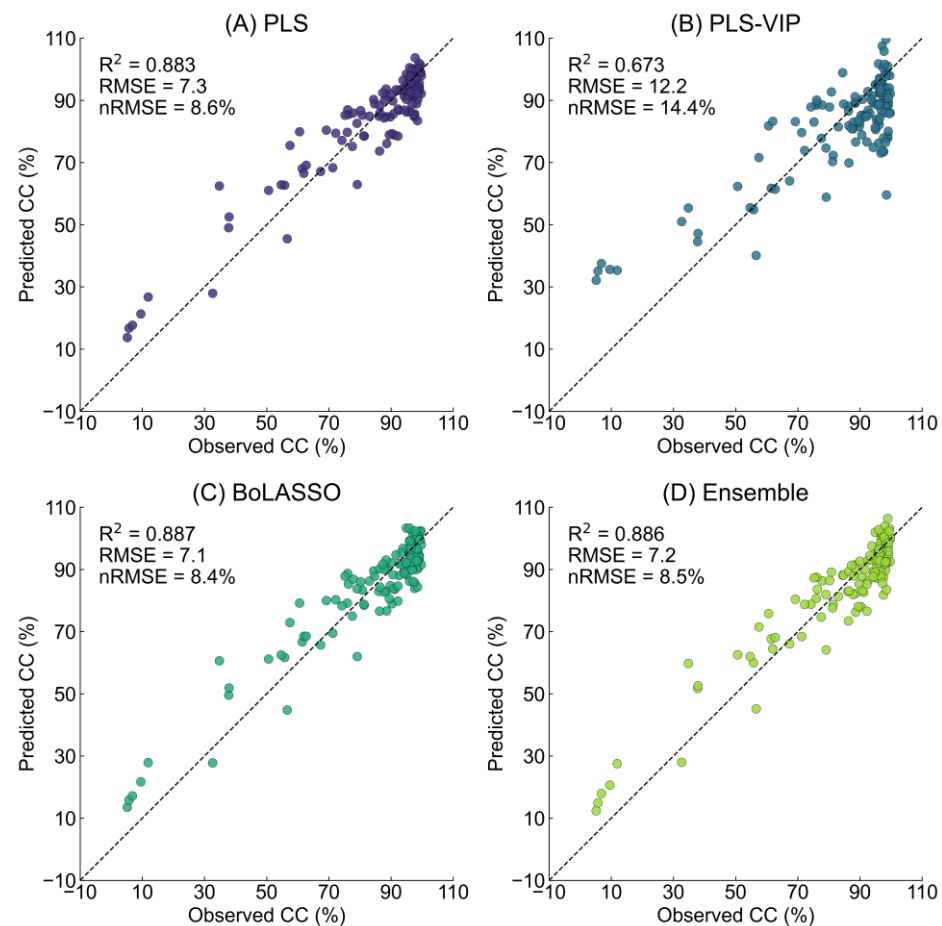


Figure 5. Relationship between observed and predicted canopy cover (CC) in the Kennard–Stone validation dataset for four regression methods: (A) PLS, (B) PLS-VIP, (C) BoLASSO, and (D) Ensemble. The dashed line indicates the 1:1 relationship, and R^2 , RMSE, and nRMSE (%) are shown in each panel.

3.3. Comparison with Conventional Vegetation Indices

We compared the prediction accuracy of conventional vegetation index models with that of the regression models using NDVI, GNDVI, NDRE, and PRI as benchmark indices (Table S4). Among the four indices, GNDVI and NDRE generally showed higher accuracy than NDVI and PRI. For AGB and LAI, however, the accuracy of the conventional vegetation index models was lower than that of the full-band PLS model and the selected-band models. In contrast, for CC, GNDVI and NDRE showed relatively high accuracy, with nRMSE values of 9.4% and 10.1%, respectively. These results indicate that conventional vegetation indices can provide useful information, particularly for CC, but multivariate models using spectral reflectance provided higher or more stable prediction accuracy for AGB and LAI.

3.4. Wavelength Distribution and Importance Scores

The selected wavelength bands differed among the band-selection methods and target traits (Tables 2–4; Figures S5–S7). For AGB, PLS-VIP selected bands mainly around 677–683 nm and 762–797 nm, whereas BoLASSO and the ensemble method selected bands from wider spectral regions, including the visible, red-edge, and near-infrared regions (Table 2). A similar tendency was observed for LAI and CC. For LAI, PLS-VIP selected closely spaced bands in the near-infrared region around 785–798 nm, while BoLASSO and the ensemble method selected bands from the visible, red-edge, and near-infrared regions (Table 3). For CC, PLS-VIP selected only near-infrared bands around 797–801 nm, whereas

BoLASSO and the ensemble method selected bands across the visible to near-infrared regions (Table 4). Although AIC-based variable reduction was applied after candidate-band selection, all candidate bands were retained in some cases, such as the LAI ensemble model.

The wavelength-wise importance scores also showed differences among the methods (Figures S5–S7). In PLS-VIP, high VIP scores and final selected bands tended to be concentrated in adjacent or closely spaced wavelengths. In contrast, BoLASSO showed several distinct peaks in selection frequency, and the ensemble method also showed peaks across multiple spectral regions. These results indicate that PLS-VIP tended to select redundant neighboring wavelengths, whereas BoLASSO and the ensemble method selected bands distributed across broader spectral regions.

3.5. Sensitivity to the Number of Candidate Bands

The sensitivity of prediction accuracy to the number of candidate bands was evaluated for BoLASSO and the ensemble method (Figure 6). Model performance was compared using the top-k candidate bands directly and after AIC-based variable reduction. The trends in nRMSE were similar before and after AIC-based reduction, although AIC-based reduction slightly improved or maintained prediction accuracy in some cases.

For AGB, validation nRMSE decreased as the number of candidate bands increased, particularly for the ensemble method. For LAI, BoLASSO reached stable accuracy with five or more candidate bands, whereas the ensemble method improved gradually with increasing band number. For CC, nRMSE decreased from 3 to 7 candidate bands and then remained constant. These results indicate that the effect of the number of candidate bands depended on the target trait. AGB prediction appeared to benefit from a larger number of candidate bands, whereas LAI and CC could be estimated with relatively few selected bands.

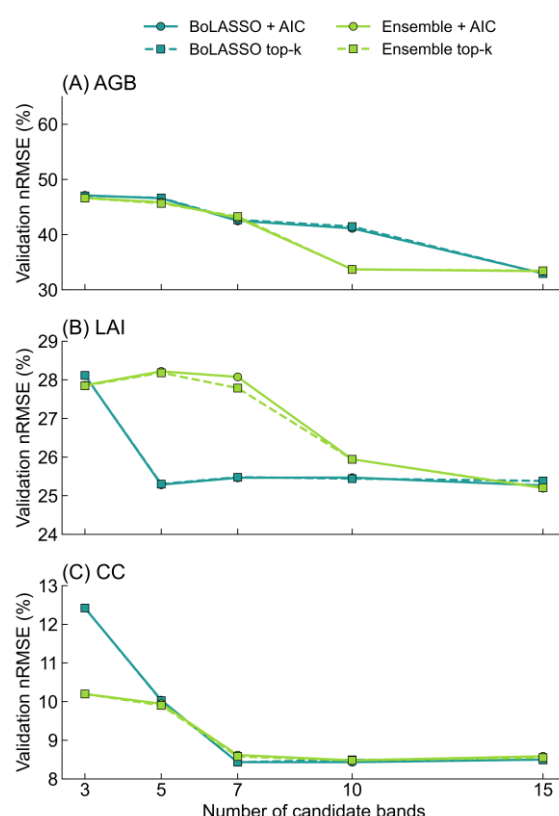


Figure 6. Sensitivity of prediction accuracy to the number of candidate wavelength bands for (A) aboveground biomass (AGB), (B) leaf area index (LAI), and (C) canopy cover (CC). Validation

nRMSE (%) is shown for BoLASSO and the ensemble method using either the top-k candidate bands directly or after AIC-based variable reduction using multiple regression. The number of candidate bands varied from 3 to 15.

3.6. Year-to-Year Transferability

Leave-one-year-out cross-validation was conducted to evaluate year-to-year transferability of the models (Table 5). Prediction accuracy was lower than that obtained with the Kennard–Stone validation, particularly for AGB. For AGB, the full-band PLS model showed the highest overall accuracy ($R^2 = 0.538$, nRMSE = 50.5%), followed by BoLASSO ($R^2 = 0.495$, nRMSE = 52.8%) and the ensemble method ($R^2 = 0.442$, nRMSE = 55.5%). In contrast, PLS-VIP showed low transferability for AGB, with $R^2 = 0.032$ and nRMSE = 73.2%.

For LAI, BoLASSO showed the highest overall accuracy ($R^2 = 0.769$, nRMSE = 28.1%), followed by the full-band PLS model ($R^2 = 0.742$, nRMSE = 29.6%) and the ensemble method ($R^2 = 0.721$, nRMSE = 30.8%). PLS-VIP showed lower accuracy than the other methods ($R^2 = 0.548$, nRMSE = 39.2%). For CC, BoLASSO and the full-band PLS model showed similar overall accuracy, with $R^2 = 0.831$ and 0.825 and nRMSE = 12.1% and 12.3%, respectively. The ensemble method showed lower accuracy ($R^2 = 0.733$, nRMSE = 15.2%), and PLS-VIP again showed the lowest accuracy ($R^2 = 0.407$, nRMSE = 22.6%).

The fold-level results showed that prediction accuracy varied among hold-out years (Table S5). For AGB, fold nRMSE ranged from 36.4% to 68.9% for PLS, from 35.7% to 62.3% for BoLASSO, and from 37.1% to 71.3% for the ensemble method. For LAI and CC, the ranges of fold nRMSE were smaller than those for AGB. These results indicate that year-to-year transferability differed among traits, with AGB showing the largest reduction in prediction accuracy and the largest variation among hold-out years.

Table 5. Leave-one-year-out cross-validation performance of the four regression methods for predicting aboveground biomass (AGB), leaf area index (LAI), and canopy cover (CC).

Trait	Method	Overall R^2	Overall nRMSE (%)
AGB	PLS	0.538	50.5
AGB	PLS-VIP	0.032	73.2
AGB	BoLASSO	0.495	52.8
AGB	Ensemble	0.442	55.5
LAI	PLS	0.742	29.6
LAI	PLS-VIP	0.548	39.2
LAI	BoLASSO	0.769	28.1
LAI	Ensemble	0.721	30.8
CC	PLS	0.825	12.3
CC	PLS-VIP	0.407	22.6
CC	BoLASSO	0.831	12.1
CC	Ensemble	0.733	15.2

For AGB and LAI, leave-one-year-out cross-validation was conducted using five yearly folds, whereas for CC it was conducted using four yearly folds because CC was not measured in the 2021 TARC experiment or in the KARC experiments. In each fold, data from one year were used as the validation dataset, and data from the remaining years were used for model development. Overall R^2 and nRMSE were calculated using all pooled predictions from the hold-out years. Fold-level results and sample sizes are provided in Table S5.

4. Discussion

4.1. Trait-Dependent Effects of Band Selection

This study showed that the effect of band selection depended on the target trait. HSR has been widely used to estimate crop canopy traits such as biomass, LAI-related variables, and nitrogen status, but prediction performance often varies depending on the target trait and spectral modeling approach [15–20]. In this study, the full-band PLS model showed the lowest nRMSE for AGB, whereas band selection tended to reduce prediction accuracy (Table 2). In contrast, for LAI and CC, BoLASSO and the ensemble method maintained accuracy comparable to that of the full-band PLS model using a reduced number of selected bands (Tables 3 and 4). These results are consistent with previous studies showing that useful spectral information and model performance can vary among soybean crop variables and other canopy traits [17,19,20].

This trait-dependent response likely reflects the complexity of the canopy information represented by each trait. LAI and CC are closely related to canopy leaf area, canopy closure, and light interception, which strongly affect reflectance in the visible, red-edge, and near-infrared regions through chlorophyll absorption, red-edge shifts, and multiple scattering within the canopy [14–20]. Therefore, a limited number of bands in these regions could capture much of the spectral information needed to estimate LAI and CC. This suggests that LAI and CC are mainly controlled by optical processes that strongly affect VNIR reflectance, such as canopy closure, leaf area expansion, and multiple scattering. In contrast, AGB includes biomass components that are not always optically exposed at the canopy surface, such as stems and pods, and therefore may require spectral information from broader or multiple wavelength regions. AGB is also influenced by organ allocation and growth stage. These components may be less directly represented by surface reflectance in the 501–801 nm VNIR range, particularly when only a small number of selected bands are used. The additional error analysis under high-value conditions supported this interpretation: all models tended to underestimate AGB in the high-AGB subset, whereas errors for high LAI and high CC remained relatively low for PLS, BoLASSO, and the ensemble method (Table S3).

The spectral range used in this study should also be considered when interpreting the AGB results. The measurements were limited to the VNIR range from 501 to 801 nm and did not include the shortwave infrared region, which is sensitive to water content, dry matter, cellulose, and lignin [31]. Thus, the present results should be interpreted as trait-dependent effects of band selection within the VNIR region. The advantage of the full-band PLS model for AGB may also partly reflect this spectral limitation. Band selection should therefore be regarded as a trait-specific modeling step that needs to balance prediction accuracy, model simplicity, and biological interpretability [24,25,27,41].

4.2. Redundancy in PLS-VIP Band Selection

To further interpret the trait-dependent differences described above, we examined the selected wavelengths and their distribution. BoLASSO and the ensemble method selected wavelengths across the visible, red-edge, and near-infrared regions, although the exact wavelengths differed among traits (Tables 2–4). These regions are associated with pigment absorption, red-edge shifts, leaf internal structure, and multiple scattering within the canopy [14–20,24,29,30]. Thus, the selected wavelengths were physiologically reasonable, but the way in which each method selected bands differed.

Differences in band-selection algorithms also affected prediction accuracy. In particular, PLS-VIP tended to show lower accuracy than BoLASSO and the ensemble method for all three traits (Tables 2–4). PLS-VIP evaluates band importance based on contributions to latent variables in a PLS model [23]. This approach is useful for identifying important

spectral regions, but it does not necessarily remove redundancy among adjacent wavelengths. In hyperspectral data, neighboring bands are highly correlated, and importance can be distributed across continuous spectral regions rather than isolated wavelengths [25,33].

In the present study, PLS-VIP tended to select closely spaced wavelengths, especially in the red-edge or near-infrared regions (Tables 2–4; Figures S5–S7). Such concentration of selected bands may have reduced the amount of independent spectral information retained in the final model. This may partly explain the lower prediction accuracy of PLS-VIP. Similar instability in variable selection under multicollinearity has also been reported in hyperspectral modeling [51].

4.3. Practical Implications for Multispectral Sensor Design

The results have implications for selecting spectral regions for simplified multispectral sensing of soybean canopy traits. BoLASSO and the ensemble method selected bands from broader spectral regions than PLS-VIP and generally showed higher prediction accuracy (Tables 2–4). BoLASSO achieved accuracy comparable to that of the full-band PLS model for LAI and CC and maintained a relatively high level of accuracy for AGB. This likely reflects the ability of bootstrap-based variable selection to identify stable bands while reducing redundancy among adjacent bands. Bootstrapping is widely used to improve the stability of variable selection and can help reduce overfitting and selection bias in high-dimensional data [50].

The ensemble method also outperformed PLS-VIP, although its performance differed among traits. This may reflect the fact that integrating multiple methods reduces the influence of relying on a single model structure. Ensemble strategies have been used in remote sensing to improve prediction accuracy and robustness [24,40]. In this study, the ensemble method was particularly effective for AGB, for which information from multiple spectral regions appeared to be useful.

The number of candidate bands was limited to 10 in this study, considering the possible use of multispectral sensors mounted on UAVs. This restriction was not intended to determine the optimal number of bands, but to examine whether prediction accuracy could be maintained with a band number close to that of practical sensor configurations. The sensitivity analysis showed that the effect of the number of candidate bands differed among traits (Figure 6). For AGB, prediction accuracy improved as the number of candidate bands increased. In contrast, LAI and CC reached relatively stable accuracy with fewer bands, particularly for BoLASSO. These results indicate that AGB prediction may require information from more spectral regions, whereas LAI and CC can be estimated with fewer selected bands.

The comparison with conventional vegetation indices also supports this interpretation (Table S4). Conventional indices such as GNDVI and NDRE showed relatively high accuracy for CC, but their accuracy was lower for AGB and LAI than that of multivariate models using selected bands. This suggests that single-index approaches can be useful for simple canopy traits, whereas multivariate models are more suitable when the target trait is influenced by several canopy characteristics.

The selected bands identified in this study should not be interpreted as exact specifications for commercial multispectral sensors. The present analysis was based on 1-nm hyperspectral reflectance data, whereas most UAV multispectral sensors use broader bands with fixed center wavelengths and sensor-specific spectral response functions. Therefore, the selected wavelengths should be regarded as candidate spectral regions rather than exact sensor bands. Further evaluation using actual multispectral sensor data, or simulated broad-band reflectance based on sensor response functions, will be needed before applying these bands directly to UAV-based monitoring.

4.4. Transferability

Although the Kennard–Stone algorithm provided representative training and validation datasets, this validation scheme does not fully evaluate model transferability across independent years, sites, cultivars, or management conditions. In the Kennard–Stone split, samples from the same years, sites, cultivars, and treatments can be included in both the training and validation datasets. Therefore, the results from the Kennard–Stone validation should be interpreted as model performance within the observed distribution of the pooled dataset.

To further examine year-to-year transferability, leave-one-year-out cross-validation was added. This validation showed lower prediction accuracy than the Kennard–Stone validation, particularly for AGB, indicating that year-to-year transferability remains a challenge for biomass prediction. For AGB, the best overall nRMSE under leave-one-year-out cross-validation was 50.5% for the full-band PLS model, whereas the selected-band models showed nRMSE values of 52.8–73.2% (Table 5). In contrast, LAI and CC showed better year-to-year transferability, with BoLASSO achieving the lowest nRMSE for both traits (28.1% for LAI and 12.1% for CC). These results suggest that selected-band models can be reasonably transferable for LAI and CC, but that AGB prediction is more sensitive to year-to-year variation. The fold-level results in Table S5 further support this interpretation, showing larger year-to-year variation in prediction accuracy for AGB than for LAI and CC.

This additional validation is important because spectral relationships can vary among years due to differences in weather conditions, growth stage distribution, canopy development, and soil background conditions. Even when a model performs well in a representative validation dataset, its performance may decline when applied to an independent year. Therefore, both representative validation, such as the Kennard–Stone split, and external-type validation, such as leave-one-year-out cross-validation, are useful for evaluating spectral prediction models.

4.5. Limitations

This study has several limitations. First, the spectral range was limited to 501–801 nm. As discussed above, this range includes information related to pigments, red-edge position, and canopy structure, but it does not include the shortwave infrared region, which is related to water content and dry matter composition. The lack of shortwave infrared information may have limited the ability to estimate AGB, particularly under high-biomass conditions. Second, this study was based on ground-based, canopy-scale reflectance data. Direct application to UAV or satellite observations requires consideration of observation angle, spatial resolution, atmospheric correction, background soil effects, and differences in sensor bandwidth [11–13]. Third, although the dataset included multiple sites, years, cultivars, and irrigation treatments, further validation is needed for different climate zones, cropping systems, and cultivars.

Overall, this study showed that the effectiveness of band selection depends on both the target trait and the band-selection method. For LAI and CC, selected-band models, especially BoLASSO and the ensemble method, maintained accuracy comparable to that of the full-band PLS model. In contrast, for AGB, full-band PLS remained advantageous, indicating that reducing the number of bands involves a trade-off between model simplicity and prediction accuracy. These results suggest that band selection should be applied according to the target trait and the intended use of the model.

The results also suggest that AGB and CC can be estimated from the same HSR dataset. Because CC is related to canopy light interception and AGB represents dry matter accumulation, their combined estimation may provide useful information for evaluating canopy productivity and RUE-related traits. Conventional RUE estimation requires re-

peated destructive sampling and radiation measurements [4–7,42]. Therefore, simultaneous estimation of AGB and CC by HSR could support a more rapid and non-destructive assessment of canopy growth under field conditions.

5. Conclusions

This study evaluated the impact of different band-selection algorithms on the accuracy of predicting soybean canopy traits using canopy hyperspectral reflectance in the VNIR range from 501 to 801 nm. The results showed that the effectiveness of band selection depends on the target trait. For AGB, the full-band PLS model provided the highest accuracy under the Kennard–Stone validation, and leave-one-year-out cross-validation indicated that year-to-year transferability remains a challenge for biomass prediction. In contrast, for LAI and CC, selected-band models, particularly BoLASSO, achieved accuracy comparable to that of full-band PLS while using fewer bands.

The selected wavelengths were mainly located in the visible, red-edge, and near-infrared regions, reflecting spectral information related to pigments and canopy structure. These results indicate that trait-specific band selection can contribute to simpler spectral models for estimating soybean canopy traits, especially LAI and CC. However, the selected wavelengths should be interpreted as candidate spectral regions within the VNIR range, and further validation using broader spectral ranges and actual multispectral sensor configurations is needed for practical UAV-based applications.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs18132179/s1>: Table S1. Summary of field experiments conducted at three sites (TARC, NIAES, and KARC) from 2019 to 2024, including experimental design, cultivars, water management treatments, sampling dates (day of year, DOY), and the number of samples for aboveground biomass (AGB), leaf area index (LAI), and canopy cover (CC); Table S2. Monthly mean air temperature (°C), solar radiation (MJ m⁻² d⁻¹), and precipitation (mm) during the soybean growing season (June–October) at each experimental site; Table S3. Prediction errors of the four regression models under high-value conditions for aboveground biomass (AGB), leaf area index (LAI), and canopy cover (CC) in the Kennard–Stone validation dataset; Table S4. Performance of conventional vegetation index models for predicting aboveground biomass (AGB), leaf area index (LAI), and canopy cover (CC) in the Kennard–Stone validation dataset; Table S5. Fold-level results of leave-one-year-out cross-validation for aboveground biomass (AGB), leaf area index (LAI), and canopy cover (CC); Figure S1. Field-portable spectral reflectance measurement system used in this study; Figure S2. Relationship between observed and predicted aboveground biomass (AGB) in the Kennard–Stone training dataset for four regression methods: (A) PLS, (B) PLS-VIP, (C) BoLASSO, and (D) Ensemble; Figure S3. Relationship between observed and predicted leaf area index (LAI) in the Kennard–Stone training dataset for four regression methods: (A) PLS, (B) PLS-VIP, (C) BoLASSO, and (D) Ensemble; Figure S4. Relationship between observed and predicted canopy cover (CC) in the Kennard–Stone training dataset for four regression methods: (A) PLS, (B) PLS-VIP, (C) BoLASSO, and (D) Ensemble; Figure S5. Wavelength-wise importance scores for the selected-band methods used to predict aboveground biomass (AGB); Figure S6. Wavelength-wise importance scores for the selected-band methods used to predict leaf area index (LAI); Figure S7. Wavelength-wise importance scores for the selected-band methods used to predict canopy cover (CC).

Author Contributions: Conceptualization, E.K.; methodology, E.K.; formal analysis, E.K.; investigation, E.K., T.Y., Y.M., K.K., E.F. and R.N.; resources, E.K., T.Y., Y.M., K.K., E.F. and R.N.; data curation, E.K.; writing—original draft preparation, E.K.; writing—review and editing, T.Y., Y.M., K.K., E.F. and R.N.; visualization, E.K.; supervision, E.K.; project administration, E.K.; funding acquisition, E.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Cross-ministerial Strategic Innovation Promotion Program (SIP), “Building a sustainable food chain that provides abundant and nutritious food (tentative)” (funding agency: Bio-oriented Technology Research Advancement Institution) and by the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Numbers JP18H02190 and JP23K26892).

Data Availability Statement: The datasets presented in this article are not readily available because they are part of an ongoing research project and are subject to intellectual property considerations and contractual restrictions related to data sharing. During the project period, the data are shared only among authorized project members. Requests to access the datasets should be directed to the corresponding author. Access may be granted after consultation with the project authorities and relevant collaborators, provided that such access does not conflict with ongoing analyses, planned publications, intellectual property management, or contractual obligations. After completion of the project and related intellectual property and publication procedures, part of the dataset may be considered for public release following appropriate data processing and with proper acknowledgment of the data source.

Acknowledgments: We thank Toshihiro Hasegawa, Eri Ogiso-Tanaka, Kaori Hirata, Tetsuya Yamada, Kazumi Suenaga, Ayumi Hatakeyama, Toyoshi Omizu, Hirokazu Tomiyama, Takahiro Mikuni, Akihito Okubo, Kanae Saito, Yuki Muta, Rina Suzuki, Hiroyuki Kato, Yoshihiro Nakao, Tomoe Abe, Kohei Nishiya, Hiroko Yamaura, Miyuki Nakajima, Mari Namikawa, Kafumi Segawa, Yuko Suzuki, Erika Sasaki, Koya Yoshida, Hisashi Tamura, Eisaku Kumagai, Fumihiko Saito, Hisaya Tanaka, Hiroshi Kawamukai, Yuto Omori, Yasuaki Kamata, Koki Yamato, Haruko Shimoda, Yasutaka Kawaguchi, Masao Minamoto, Isao Aoki, Kazumi Tominaga, and Kanako Tanaka (NARO) for their technical assistance. We also thank the members of the SIP project for their valuable comments and constructive suggestions on the methodological aspects of this study. During the preparation of this manuscript, the authors used ChatGPT GPT-5.5 through Microsoft Copilot to support language editing, grammar checking, and readability improvement. The authors reviewed and edited all AI-assisted outputs and take full responsibility for the content of the manuscript.

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

References

1. Jin, J.; Liu, X.; Wang, G.; Mi, L.; Shen, Z.; Chen, X.; Herbert, S.J. Agronomic and physiological contributions to the yield improvement of soybean cultivars released from 1950 to 2006 in northeast China. *Field Crop. Res.* **2010**, *115*, 116–123. <https://doi.org/10.1016/j.fcr.2009.10.016>.
2. Rincker, K.; Nelson, R.; Specht, J.; Sleper, D.; Cary, T.; Cianzio, S.R.; Casteel, S.; Conley, S.; Chen, P.; Davis, V.; et al. Genetic improvement of U.S. soybean in maturity groups ii, iii, and iv. *Crop Sci.* **2014**, *54*, 1419–1432. <https://doi.org/10.2135/cropsci2013.10.0665>.
3. Todeschini, M.H.; Milioli, A.S.; Rosa, A.C.; Dallacorte, L.V.; Panho, M.C.; Marchese, J.A.; Benin, G. Soybean genetic progress in south brazil: Physiological, phenological and agronomic traits. *Euphytica* **2019**, *215*, 124. <https://doi.org/10.1007/s10681-019-2439-9>.
4. Monteith, J.L. Climate and efficiency of crop production in britain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1977**, *281*, 277–294. <https://doi.org/10.1098/rstb.1977.0140>.
5. Koester, R.P.; Skoneczka, J.A.; Cary, T.R.; Diers, B.W.; Ainsworth, E.A. Historical gains in soybean (*Glycine max* Merr.) seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. *J. Exp. Bot.* **2014**, *65*, 3311–3321. <https://doi.org/10.1093/jxb/eru187>.
6. Sinclair, T.R.; Muchow, R.C. Occam’s Razor, Radiation-use efficiency, and vapor pressure deficit. *Field Crop. Res.* **1999**, *62*, 239–243. [https://doi.org/10.1016/S0378-4290\(99\)00011-8](https://doi.org/10.1016/S0378-4290(99)00011-8).
7. Kumagai, E.; Yabiku, T.; Hasegawa, T. A strong negative trade-off between seed number and 100-seed weight stalls genetic yield gains in northern Japanese soybean cultivars in comparison with Midwestern US cultivars. *Field Crop. Res.* **2022**, *283*, 108539. <https://doi.org/10.1016/j.fcr.2022.108539>.

8. Murcia, M.A.L.; Moreira, F.F.; Hearst, A.; Cherkauer, K.; Rainey, K.M. Physiological breeding for yield improvement in soybean: Solar radiation interception-conversion, and harvest index. *Theor. Appl. Genet.* **2022**, *135*, 1477–1491. <https://doi.org/10.1007/s00122-022-04048-5>.
9. Watson, D.J. Comparative physiological studies on the growth of field crops: I. Variation in net assimilation rate and leaf area between species and varieties, and within and between years. *Ann. Bot.* **1947**, *11*, 41–76. <https://doi.org/10.1093/oxfordjournals.aob.a083148>.
10. Welles, J.M.; Norman, J.M. Instrument for indirect measurement of canopy architecture. *Agron. J.* **1991**, *83*, 818–825. <https://doi.org/10.2134/agronj1991.00021962008300050009x>.
11. Luis Araus, J.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. <https://doi.org/10.1016/j.tplants.2013.09.008>.
12. Borra-Serrano, I.; De Swaef, T.; Quataert, P.; Aper, J.; Saleem, A.; Saeys, W.; Somers, B.; Roldán-Ruiz, I.; Lootens, P. Closing the phenotyping gap: High resolution UAV time series for soybean growth analysis provides objective data from field trials. *Remote Sens.* **2020**, *12*, 1644. <https://doi.org/10.3390/rs12101644>.
13. White, J.W.; Andrade-Sanchez, P.; Gore, M.A.; Bronson, K.F.; Coffelt, T.A.; Conley, M.M.; Feldmann, K.A.; French, A.N.; Heun, J.T.; Hunsaker, D.J.; et al. Field-based phenomics for plant genetics research. *Field Crop. Res.* **2012**, *133*, 101–112. <https://doi.org/10.1016/j.fcr.2012.04.003>.
14. Fu, P.; Montes, C.M.; Siebers, M.H.; Gomez-Casanovas, N.; McGrath, J.M.; Ainsworth, E.A.; Bernacchi, C.J. Advances in field-based high-throughput photosynthetic phenotyping. *J. Exp. Bot.* **2022**, *73*, 3157–3172. <https://doi.org/10.1093/jxb/erac077>.
15. Flynn, K.C.; Baath, G.; Lee, T.O.; Gowda, P.; Northup, B. Hyperspectral reflectance and machine learning to monitor legume biomass and nitrogen accumulation. *Comput. Electron. Agric.* **2023**, *211*, 107991. <https://doi.org/10.1016/j.compag.2023.107991>.
16. Vollmann, J.; Rischbeck, P.; Pachner, M.; Đorđević, V.; Manschadi, A.M. High-throughput screening of soybean di-nitrogen fixation and seed nitrogen content using spectral sensing. *Comput. Electron. Agric.* **2022**, *199*, 107169. <https://doi.org/10.1016/j.compag.2022.107169>.
17. Chiozza, M.V.; Parmley, K.A.; Higgins, R.H.; Singh, A.K.; Miguez, F.E. Comparative prediction accuracy of hyperspectral bands for different soybean crop variables: From leaf area to seed composition. *Field Crop. Res.* **2021**, *271*, 108260. <https://doi.org/10.1016/j.fcr.2021.108260>.
18. Kawamura, K.; Ikeura, H.; Phongchanmaixay, S.; Khanthavong, P. Canopy hyperspectral sensing of paddy fields at the booting stage and pls regression can assess grain yield. *Remote Sens.* **2018**, *10*, 1249. <https://doi.org/10.3390/rs10081249>.
19. Gnyp, M.L.; Miao, Y.; Yuan, F.; Ustin, S.L.; Yu, K.; Yao, Y.; Huang, S.; Bareth, G. Hyperspectral canopy sensing of paddy rice aboveground biomass at different growth stages. *Field Crop. Res.* **2014**, *155*, 42–55. <https://doi.org/10.1016/j.fcr.2013.09.023>.
20. Hansen, P.M.; Schjoerring, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542–553. [https://doi.org/10.1016/S0169-7439\(03\)00131-7](https://doi.org/10.1016/S0169-7439(03)00131-7).
21. Xie, J.; Li, Y.; Wu, H.; Wu, Z.; Zhao, R.; Lin, A.; Adami, M.; Li, G.; Zhang, J. Integrating random forest with boundary enhancement for mapping crop planting structure at the parcel level from remote sensing images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 9934–9953. <https://doi.org/10.1109/JSTARS.2025.3554922>.
22. Wu, H.; Xie, J.; Deng, W.; Lin, A.; Mohamed Shariff, A.R.; Akmalov, S.; Wu, W.; Li, Z.; Yu, Q.; Wang, Q.; et al. CT-HiffNet: A contour-texture hierarchical feature fusion network for cropland field parcel extraction from high-resolution remote sensing images. *Comput. Electron. Agric.* **2025**, *239*, 111010. <https://doi.org/10.1016/j.compag.2025.111010>.
23. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1).
24. Feilhauer, H.; Asner, G.P.; Martin, R.E. Multi-method ensemble selection of spectral wavelengths related to leaf biochemistry. *Remote Sens. Environ.* **2015**, *164*, 57–65. <https://doi.org/10.1016/j.rse.2015.03.033>.
25. Inoue, Y.; Guerif, M.; Baret, F.; Skidmore, A.; Gitelson, A.; Schlerf, M.; Darvishzadeh, R.; Oliso, A. Simple and robust methods for remote sensing of canopy chlorophyll content: A comparative analysis of hyperspectral data for different types of vegetation. *Plant Cell Environ.* **2016**, *39*, 2609–2623. <https://doi.org/10.1111/pce.12815>.
26. Kumagai, E.; Burroughs, C.H.; Pederson, T.L.; Montes, C.M.; Peng, B.; Kimm, H.; Guan, K.; Ainsworth, E.A.; Bernacchi, C.J. Predicting biochemical acclimation of leaf photosynthesis in soybean under in-field canopy warming using hyperspectral reflectance. *Plant Cell Environ.* **2022**, *45*, 80–94. <https://doi.org/10.1111/pce.14204>.
27. Andersen, C.M.; Bro, R. Variable selection in regression—A tutorial. *J. Chemom.* **2010**, *24*, 728–737. <https://doi.org/10.1002/cem.1360>.

28. Genuer, R.; Poggi, J.-M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. <https://doi.org/10.1016/j.patrec.2010.03.014>.
29. Blackburn, G.A. Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *Int. J. Remote Sens.* **1998**, *19*, 657–675. <https://doi.org/10.1080/014311698215919>.
30. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X).
31. Curran, P.J. Remote sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271–278. [https://doi.org/10.1016/0034-4257\(89\)90069-2](https://doi.org/10.1016/0034-4257(89)90069-2).
32. Kokaly, R.F.; Clark, R.N. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sens. Environ.* **1999**, *67*, 267–287. [https://doi.org/10.1016/S0034-4257\(98\)00084-4](https://doi.org/10.1016/S0034-4257(98)00084-4).
33. Fu, Y.; Yang, G.; Pu, R.; Li, Z.; Li, H.; Xu, X.; Song, X.; Yang, X.; Zhao, C. An overview of crop nitrogen status assessment using hyperspectral remote sensing: Current status and perspectives. *Eur. J. Agron.* **2021**, *124*, 126241. <https://doi.org/10.1016/j.eja.2021.126241>.
34. Blackburn, G.A. Hyperspectral remote sensing of plant pigments. *J. Exp. Bot.* **2006**, *58*, 855–867. <https://doi.org/10.1093/jxb/erl123>.
35. Ustin, S.L.; Gitelson, A.A.; Jacquemoud, S.; Schaepman, M.; Asner, G.P.; Gamon, J.A.; Zarco-Tejada, P. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sens. Environ.* **2009**, *113*, S67–S77. <https://doi.org/10.1016/j.rse.2008.10.019>.
36. Bates, J.M.; Granger, C.W.J. The combination of forecasts. *J. Oper. Res. Soc.* **1969**, *20*, 451–468. <https://doi.org/10.1057/jors.1969.103>.
37. Waske, B.; Linden, S.V.D. Classifying multilevel imagery from SAR and optical sensors by decision fusion. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1457–1466. <https://doi.org/10.1109/TGRS.2008.916089>.
38. Engler, R.; Waser, L.T.; Zimmermann, N.E.; Schaub, M.; Berdos, S.; Ginzler, C.; Psomas, A. Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. *For. Ecol. Manag.* **2013**, *310*, 64–73. <https://doi.org/10.1016/j.foreco.2013.07.059>.
39. Du, P.; Xia, J.; Chanussot, J.; He, X. Hyperspectral remote sensing image classification based on the integration of support vector machine and random forest. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 174–177.
40. Fu, P.; Meacham-Hensold, K.; Guan, K.Y.; Bernacchi, C.J. Hyperspectral leaf reflectance as proxy for photosynthetic capacities: An ensemble approach based on multiple machine learning algorithms. *Front. Plant Sci.* **2019**, *10*, 13. <https://doi.org/10.3389/fpls.2019.00730>.
41. Rogers, M.; Blanc-Talon, J.; Urschler, M.; Delmas, P. Wavelength and texture feature selection for hyperspectral imaging: A systematic literature review. *J. Food Meas. Charact.* **2023**, *17*, 6039–6064. <https://doi.org/10.1007/s11694-023-02044-x>.
42. Kumagai, E.; Takahashi, T. Soybean (*Glycine max* (L.) Merr.) yield reduction due to late sowing as a function of radiation interception and use in a cool region of northern Japan. *Agronomy* **2020**, *10*, 14. <https://doi.org/10.3390/agronomy10010066>.
43. Kimm, H.; Guan, K.; Burroughs, C.H.; Peng, B.; Ainsworth, E.A.; Bernacchi, C.J.; Moore, C.E.; Kumagai, E.; Yang, X.; Berry, J.A.; et al. Quantifying high-temperature stress on soybean canopy photosynthesis: The unique role of sun-induced chlorophyll fluorescence. *Glob. Change Biol.* **2021**, *27*, 2403–2415. <https://doi.org/10.1111/gcb.15603>.
44. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. <https://doi.org/10.1021/ac60214a047>.
45. Kennard, R.W.; Stone, L.A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148. <https://doi.org/10.1080/00401706.1969.10490666>.
46. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with ERTS. In *NASA Special Publication*; Freden, S.C., Mercanti, E.P., Becker, M.A., Eds.; NASA: Washington, DC, USA, 1974; Volume 351, p. 309.
47. 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. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7).
48. Barnes, E.M.; Clarke, T.R.; Richards, S.E. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In *Proceedings of the Fifth International Conference on Precision Agriculture*; Robert, P.C., Rust, R.H., Larson, W.E., Eds.; American Society of Agronomy: Madison, WI, USA, 2000.
49. Gamon, J.P.; Field, C.B. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44.

50. Bach, F. Bolasso: model consistent Lasso estimation through the bootstrap. *arXiv* **2008**, arXiv:0804.1302. <https://doi.org/10.48550/arXiv.0804.1302>.
51. Chong, I.-G.; Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112. <https://doi.org/10.1016/j.chemolab.2004.12.011>.

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.