



Article

Estimating Above-Ground Biomass of Potato Using Random Forest and Optimized Hyperspectral Indices

Haibo Yang ¹, Fei Li ^{1,*}, Wei Wang ² and Kang Yu ³

¹ Inner Mongolia Key Laboratory of Soil Quality and Nutrient Resource, College of Grassland, Resources and Environment, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia 010011, China; hbyang@emails.imau.edu.cn

² ULanqab Institute of Agriculture and Forestry Sciences, ULanqab, Inner Mongolia 012000, China; wlcbbw@163.com

³ Department Life Science Engineering, School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; kang.yu@tum.de

* Correspondence: lifei@imau.edu.cn

Abstract: Spectral indices rarely show consistency in estimating crop traits across growth stages; thus, it is critical to simultaneously evaluate a group of spectral variables and select the most informative spectral indices for retrieving crop traits. The objective of this study was to explore the optimal spectral predictors for above-ground biomass (AGB) by applying Random Forest (RF) on three types of spectral predictors: the full spectrum, published spectral indices (Pub-SIs), and optimized spectral indices (Opt-SIs). Canopy hyperspectral reflectance of potato plants, treated with seven nitrogen (N) rates, was obtained during the tuber formation and tuber bulking from 2015 to 2016. Twelve Pub-SIs were selected, and their spectral bands were optimized using band optimization algorithms. Results showed that the Opt-SIs were the best input variables of RF models. Compared to the best empirical model based on Opt-SIs, the Opt-SIs based RF model improved the prediction of AGB, with R^2 increased by 6%, 10%, and 16% at the tuber formation, tuber bulking, and for across the two growth stages, respectively. The Opt-SIs can significantly reduce the number of input variables. The optimized Blue nitrogen index (Opt-BNI) and Modified red-edge normalized difference vegetation index (Opt-mND705) combined with an RF model showed the best performance in estimating potato AGB at the tuber formation stage ($R^2 = 0.88$). In the tuber bulking stage, only using optimized Nitrogen planar domain index (Opt-NPDI) as the input variable of the RF model produced satisfactory accuracy in training and testing datasets, with the R^2 , RMSE, and RE being 0.92, 208.6 kg/ha, and 10.3%, respectively. The Opt-BNI and Double-peak nitrogen index (Opt-NDDA) coupling with an RF model explained 86% of the variations in potato AGB, with the lowest RMSE (262.9 kg/ha) and RE (14.8%) across two growth stages. This study shows that combining the Opt-SIs and RF can greatly enhance the prediction accuracy for crop AGB while significantly reduces collinearity and redundancies of spectral data.



Citation: Yang, H.; Li, F.; Wang, W.; Yu, K. Estimating Above-Ground Biomass of Potato Using Random Forest and Optimized Hyperspectral Indices. *Remote Sens.* **2021**, *13*, 2339. <https://doi.org/10.3390/rs13122339>

Academic Editor: Brigitte Leblon

Received: 13 May 2021

Accepted: 10 June 2021

Published: 15 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Keywords: potato crops; biomass estimation; machine learning; vegetation indices



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Above-ground biomass (AGB) as an insightful indicator of crop production is essential to guide agricultural management practices [1–3]. Particularly, the AGB is the most important indicator for the calculation of the nitrogen nutrition index, which has been proposed to diagnose the plant nitrogen status for precise nitrogen fertilizer management in crops [4,5]. However, traditional manual measurements of the AGB are time-consuming and labor-intensive. These point-sampling-based methods are also infeasible in regional crop management [6–8]. Therefore, it is imperative to develop effective technologies to rapidly and accurately monitor crops AGB on a regional scale for precision crop nitrogen management [9,10].

Remote sensing is regarded as the most promising technology for timely and non-destructively acquiring crop canopy spatial variation [3,11]. The spectral parameters are widely used to estimate AGB at local, regional, and global scales [12–14]. For decades, spectral indices (SIs) are common parameters for the non-destructive estimation of AGB in crops [15,16]. Much research has reported that the SIs were used to estimate the AGB of maize [6,17,18], wheat [10,19,20], and rice [15,21]. However, the relationship between the SIs and AGB is inconsistent. For example, Venancio [17] found that the enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), and optimized soil-adjusted vegetation index (OSAVI) were the best performing SIs for the estimation of maize AGB estimation, whereas Zhu [18] found that the normalized green red difference index (NGRDI) and simple ratio vegetation index (SRVI) had the highest correlation coefficient with AGB of maize. In contrast, Wang [6] found that the SIs showed poor relationships with maize AGB. The relationships between the selected SIs and crop properties are often inconsistent due to the influence of sites, years, and crop growth stages on the SIs algorithms and corresponding sensitive band combinations [22,23]. Therefore, the selection of suitable SIs algorithms and the sensitive band to develop the optimized SIs is necessary to guarantee the performance of SIs. Considering all possible band combinations according to established index formulations is a popular method to develop optimized SIs. The optimized SIs can improve the robustness and accurate of vegetation properties estimation by identifying optimal band combinations to some degree [24,25]. Many studies have demonstrated that it is an effective method for improving the estimation accuracy of AGB [21,26–28]. However, these SIs based methods keep on being restricted by using a few bands only. We could not ensure whether the spectral indices were the most effective indicators for estimating certain vegetation properties [29,30]. Although the band optimization can improve the robustness of SIs to some degree, the saturation effect is still a problem for quantifying moderate-high AGB in crops [31–34].

Recently, machine learning coupled with remotely sensed data have become a popular approach for vegetation parameters estimation [35,36]. Dayananda [37] used 121 hyperspectral bands from 450 to 998 nm coupling with a random forest algorithm to predict the biomass of maize with lower root mean square error. Similarly, the findings of Yu [38] confirmed that the partial least squares regression (PLSR) model could explain 91% variation of the rice AGB from the full spectrum analysis. However, the full spectrum has strong band continuity and higher dimensionality which reduce the computation efficiency and estimation accuracy of the model, especially for hyperspectral data [12,39,40]. Therefore, identifying the method to reduce the dimension of spectrum input variables is very important for machine learning algorithms. SIs are the mathematical transformation of several specific wavelengths. Much research has demonstrated that the SIs coupling with machine learning algorithms can achieve accurate estimation of crop biomass and other biophysical parameters [15,41]. For instance, Wang [41] predicted the wheat AGB using the support vector regression (SVR), the random forest (RF), and artificial neural network (ANN) based on 15 SIs, e.g., the normalized difference vegetation index (NDVI), optimized soil-adjusted vegetation index (OSAVI) and EVI. The SIs coupling with the SVR, RF, and ANN algorithms could explain 47–61%, 53–79%, and 30–49% of the variations in AGB at different growth stages, respectively. Niu [42] predicted the maize AGB using multivariable linear regression (MLR) based on the normalized green-red difference index (NGRDI), excess green minus excess red (ExGR) with an R^2 of 0.82. The partial least squares regression (PLSR) coupling with the two-band NDVI combinations estimated wheat AGB satisfactorily ($R^2 = 0.89$) [12]. In contrast, Wang [6] showed that the PLSR model coupling with SIs, e.g., NDVI, simple ratio vegetation index (SR), and modified soil-adjusted vegetation index (MSAVI), did not improve the biomass estimation accuracy in maize. Therefore, selecting suitable spectral variables as input parameters is critical for machine learning algorithms.

Machine learning offers a practical method for analyzing spectral information and accurately monitoring vegetation biophysical variables [29,30]. The previous studies have investigated the full spectrum and the SIs as input variables for machine learning to esti-

mate crop AGB. Spectral indices can effectively reduce the dimension and multicollinearity of machine learning input variables. However, to date, limited studies reported the Opt-SIs coupling with machine learning algorithms in the estimation of crop AGB. Among various machine learning methods, the RF algorithm has been regarded as one of the most commonly used prediction approaches for classification and regression due to the advantages of computation efficiency and insensitivity to over-fitting [41]. Therefore, the main objectives of the current research were (1) to evaluate the performances of published and Opt-SIs in estimating potato AGB, (2) to compare the performance of full-spectrum, Pub-SIs, and Opt-SIs coupled with RF models in predicting potato AGB.

2. Materials and Methods

2.1. Study Area and Experimental Design

Experiments were carried out in Wuchuan County, which is located in the middle of Inner Mongolia (extending from 110°31'E, 40°47'N to 111°53'E, 41°23'N), China, during the potato growing seasons of 2015 and 2016. As illustrated in Figure 1, the annual variations in temperature and precipitation are small. The area has a middle temperate arid and semi-arid continental monsoon, with cold winters and cool summers. The annual average precipitation is 398.3 mm and 90–95% of which occurs between April and October. The average temperature is 15.0 °C in the potato growing season. The main crops in this area are potato, sunflower, and oat.

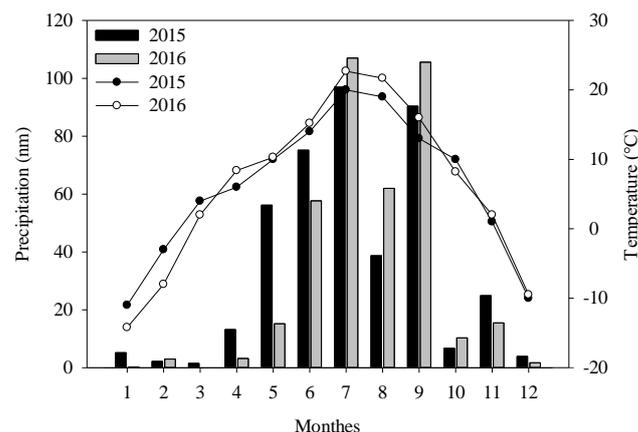


Figure 1. The variations of temperature and precipitation of Wuchuan County in 2015 and 2016.

Two field experiments with different N levels and one cultivar (Kexin 1) were conducted at two villages. Experiment 1 was performed in Dong Liangshan with seven N rates (0, 83, 135, 165, 180, 250, and 424 kg N ha⁻¹) and four replications in 2015. The plot size was 9 m × 9 m. Nitrogen fertilizer applications with urea were split for before sowing and five growth stages, i.e., BBCH code 31, 40, 51, 60, and 70. At Dong Tucheng in 2016, Experiment 2 had seven N treatments, among the one control (0 kg N ha⁻¹), four optimum N treatments using urea, urea with urease inhibitor, urea ammonium nitrate solution, urea ammonium nitrate solution with a urease inhibitor, and two conventional N rates based on urea and urea ammonium nitrate solution. Each treatment had four replications and a plot of 8 m × 9 m. The N fertilizer was used as a fertigation split at five growth stages (BBCH 31, 40, 51, 60, and 70). The design of a completely random block was adopted in all experimental plots.

2.2. Data Collection

During the experimental periods, ground-based hyperspectral reflectance was measured at key growth stages of tuber formation and tuber bulking using a Field Spec Pro FR spectroradiometer (tec5, Oberursel, Germany). This instrument records reflectance between 300 and 1150 nm with a bandwidth resolution of 3.3 nm. As described by Li [22,27,32], the

measuring head of this device consists of two optics: the upper one is used to quantify the incoming light as a reference, and the lower one records the reflectance from the vegetation and ground with a 12° FOV. The bi-directional spectrometer was calibrated once with a Spectralon white reference panel. In the critical growing season of tuber formation and tuber bulking, which corresponded to the BBCH codes 40–49 and 60–69, potato canopy reflectance was measured by holding the sensor at nadir, from 0.5–0.8 m above the canopy, and walking a distance of 4–6 m with a constant speed along the potato ridges for each plot at a clear and windless day condition between 10:00 a.m. and 2:00 p.m. Subsequently, randomly selected two 1 m (1.8 m²) consecutive rows of potato in the spectrometer-scanned area of each plot, all the above-ground plants were sampled and weighed the fresh weight. After chopping and mixing samples, 400–600 g sub-samples were oven-dried at 70 °C. to constant weight, after which the dry weight was determined. The AGB can be calculated as the following equation:

$$\text{AGB (kg/ha)} = \frac{\text{FW}}{1 + \text{M}} / 1.8 \text{ m}^2 \times 10000 \text{ m}^2 \quad (1)$$

FW: fresh weight of all the above-ground plants.

M: moisture content of potato plants.

In the current studies, the sample of tuber formation and tuber bulking stages is 79 and 78, respectively. For each stage, the pooled data from 2015 and 2016 were randomly divided into a training dataset (70% of the pooled data) and an independent testing dataset (30% of the pooled data). For the training dataset, the number of samples was 55 at the tuber formation, 54 at the tuber bulking stage. For the test dataset, there were 24 samples for each growth stage. The pooled datasets (All) were derived from the combination of the data of the above two stages, including 109 samples in the training dataset and 48 samples in the testing dataset. The training dataset was used to establish models to investigate the performance of RF algorithm coupling with different spectral variables. The test dataset was used to test the accuracy and robustness of each prediction model.

2.3. Spectral Indices

Many SIs have been developed to estimating crop biophysical parameters. Among them, the two-band spectral indices simple ratio vegetation index [43], normalized difference vegetation index (NDVI) [44], and difference vegetation index [45] are the most classic spectral indices algorithms. Subsequently, the three-band spectral indices were proposed by adding new wavebands and changing the formula formats. For example, Sims and Gamon [46] modified the normalized difference (ND) formula formats by using blue reflectance (R₄₅₅) as a measure and proposed the new three-band spectral indices ND705. In the current study, we chose 12 Pub-SIs (Table 1) and used the “lambda-by-lambda” band-optimization algorithm which is widely used for the optimization of spectral indices [47–49] to determine the best band combination for different spectral indices formula formats. The Opt-SIs were used as predictors in the RF models for potato AGB estimation. More details regarding the process of selecting spectral index sensitive bands are described in our previous work [27,50,51].

2.4. Random Forest Regression

The RF regression model is an ensemble learning algorithm that combines a large number of decision trees (*ntree*). When each tree was built, two-third of the training samples were used to training the model, and one-third of training samples, called out-of-bag (OOB) samples were left out [3,52]. The prediction result was determined by averaging over all the trees [3,41,53].

RF algorithm has been regarded as one of the most accurate prediction approaches for classification and regression [41]. The framework parameters (e.g., *n_estimators*, *oob_score* and *criterion*) and decision tree parameters (e.g., *max_depth*, *max_features*, and *max_leaf_nodes*) are an important part of the RF algorithm. Among them, the number of decision trees

(*ntree*) and input variables per node (*mtry*) are the most frequently concerned parameters to improve operating efficiency and estimation accuracy. However, most of the studies have by far focused on the *ntree* and *mtry*, and have not recognized the importance of input variables to achieve their best performance. Therefore, three types of spectral input variables: full-spectrum, Pub-SIs, and Opt-SIs were applied in the current study to investigate the influence of spectral input variables on the estimation ability of random forest algorithm.

To evaluate the influences of different input variables coupling with the RF model to predict potato AGB, we set three decision tree levels: high (*ntree* = 5000), medium (*ntree* = 2000), and low (*ntree* = 500), and other parameters were set to default values according to the *scikit-learn* software package (<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>). Subsequently, the number of the *mtry* and *ntree* were further optimized according to the ‘importance scores’ and root square error (RMSE) for simplifying the prediction model and enhancing the operational efficiency and accuracy.

2.5. Variable Importance Score

The RF model can assess the variable importance score according to the out-of-bag (OOB) error estimates in the model [21]. Importance score as an evaluation indicator was widely used to select the input variables to simplify the RF model [52,54]. The importance score can be calculated as follow:

$$\text{Importance score (X)} = \sum_{i=1}^n \frac{\text{errOOB2} - \text{errOOB1}}{n} \quad (2)$$

where *errOOB1* represents the error of out-of-bag for variable X with one decision tree, *errOOB2* represents the error of adding noise to variable X with one decision tree, and *n* represents the number of decision trees.

2.6. Model Accuracy

The performances of the different models of RF and spectral indices were evaluated by comparing the coefficients of determination (R^2) [55] relative error (RE, %) [56] and root square error (RMSE) [57] in predictions. The higher the R^2 and the lower the RMSE and RE, the better the precision and accuracy of the models. The R^2 , RMSE, and RE were calculated as following equations:

$$R^2 = \frac{\sum (y_i - \bar{y})^2}{\sum (y_i - \hat{y}_i)^2} \quad (3)$$

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

$$RE(\%) = \frac{RMSE}{\bar{y}} * 100 \quad (5)$$

where \hat{y}_i , y_i , and \bar{y} are the measured, predicted, and mean values of AGB, respectively, and *n* is the number of samples.

Table 1. Spectral indices used in this study.

Types	Spectral Indices	Abbreviations	Formulas	Algorithms	References
Two-band spectral indices	Ratio vegetation index	RVI	R_{800}/R_{670}	$R_{\lambda 1}/R_{\lambda 2}$	[43]
	Normalized difference vegetation index	NDVI	$(R_{800} - R_{680})/(R_{800} + R_{680})$	$(R_{\lambda 1} - R_{\lambda 2})/(R_{\lambda 1} + R_{\lambda 2})$	[44]
	Different vegetation index	DVI	$R_{800} - R_{680}$	$R_{\lambda 1} - R_{\lambda 2}$	[45]
	Modified soil-adjusted vegetation index	MSAVI	$0.5 \times (2 \times R_{810} + 1 - ((2 \times R_{810} + 1) \times 2 - 8 \times (R_{810} - R_{670})) \times 0.5)$	$0.5 \times (2 \times R_{\lambda 1} + 1 - ((2 \times R_{\lambda 1} + 1) \times 2 - 8 \times (R_{\lambda 1} - R_{\lambda 2})) \times 0.5)$	[58]
	The renormalized difference vegetation index	RDVI	$(R_{800} - R_{670})/\text{sqrt}(R_{800} + R_{670})$	$(R_{\lambda 1} - R_{\lambda 2})/\text{sqrt}(R_{\lambda 1} + R_{\lambda 2})$	[59]
	Optimal vegetation index	VI _{opt}	$(1 + 0.45) \times ((R_{800}) \times 2 + 1)/(R_{670} + 0.45)$	$(1 + 0.45) \times ((R_{\lambda 2}) \times 2 + 1)/(R_{\lambda 1} + 0.45)$	[60]
Three-band spectral indices	Canopy chlorophyll content index	CCCI	$(NDRE - NDRE_{MIN})/(NDRE_{MAX} - NDRE_{MIN})$	$(NDRE - NDRE_{MIN})/(NDRE_{MAX} - NDRE_{MIN})$	[61]
	Modified red-edge normalized difference vegetation index	mND705	$(R_{750} - R_{705})/(R_{750} + R_{705} - 2 \times R_{445})$	$(R_{\lambda 1} - R_{\lambda 2})/(R_{\lambda 1} + R_{\lambda 2} - 2 \times R_{\lambda 3})$	[46]
	Blue nitrogen index	BNI	$R_{434}/(R_{496} + R_{401})$	$R_{\lambda 1}/(R_{\lambda 2} + R_{\lambda 3})$	[62]
	Nitrogen planar domain index	NPDI	$(CI_{\text{green edge}} - CI_{\text{green edge MIN}})/$ $(CI_{\text{green edge MAX}} - CI_{\text{green edge MIN}})$	$(CI_{\text{green edge}} - CI_{\text{green edge MIN}})/$ $(CI_{\text{green edge MAX}} - CI_{\text{green edge MIN}})$	[51]
	Double-peak nitrogen index	NDDA	$(R_{755} + R_{680} - 2 \times R_{705})/(R_{755} - R_{680})$	$(R_{\lambda 1} + R_{\lambda 2} - 2 \times R_{\lambda 3})/(R_{\lambda 1} - R_{\lambda 2})$	[63]
Modified red-edge ratio	mRER	$(R_{759} - 1.8 \times R_{419})/(R_{742} - 1.8 \times R_{419})$	$(R_{\lambda 1} - 1.8 \times R_{\lambda 2})/(R_{\lambda 3} - 1.8 \times R_{\lambda 2})$	[64]	

R: the abbreviation of reflectance. λ : the wavebands of spectral indices.

3. Results

3.1. Variations in Potato AGB

As illustrated in Figure 2, the potato AGB increased from tuber formation to tuber bulking stage. The AGB in the training dataset ranged from 492.8 kg/ha to 3881.0 kg/ha with a CV value of 48.1%, while it varied from 497.2 kg/ha to 3353.5 kg/ha with a CV value of 44.9% for the testing dataset during the tuber formation and tuber bulking stage. The training and testing datasets exhibited a similar statistical distribution of AGB, avoiding potentially biased estimations in model calibration and validation.

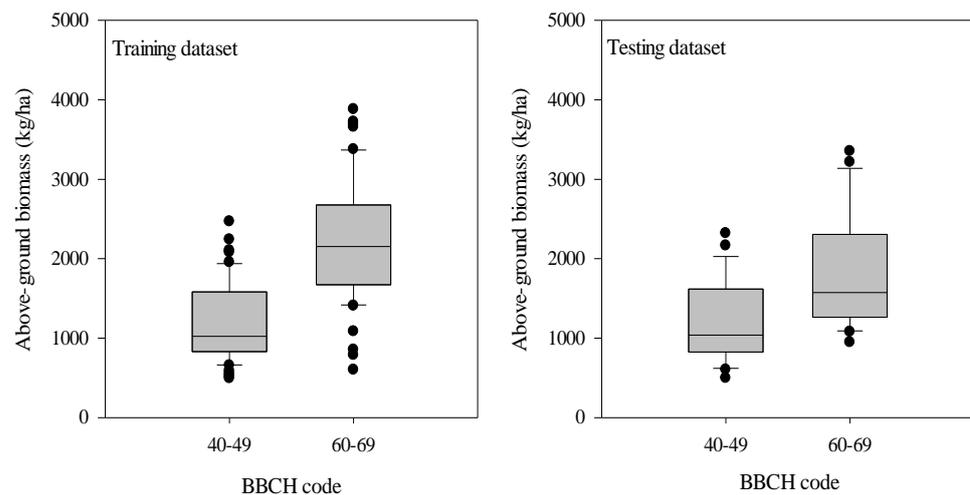


Figure 2. Variation of the above-ground biomass of potato at tuber formation and tuber bulking stages. The distribution is characterized by box-and-whisker plots, where the boxes show the 25th and 75th percentiles and the whiskers the 10th and the 90th percentiles. The median is represented by the line in the box and is provided as a number above the box plot.

3.2. Relationships between Spectral Indices and Potato AGB

Table 2 shows the relationships between Pub-SIs and potato AGB based on the training dataset. The results indicated that the Pub-SIs could explain 6–62% of the variations in potato AGB for different growth stages. The growth stages significantly affect the performances of Pub-SIs. The best performing Pub-SIs were mND705, CCCI, and RVI for the tuber formation stage, tuber bulking stage, and the combination of two stages, respectively.

Table 2. The band (nm) combination of published spectral indices (Pub-SIs) and corresponding relationships between Pub-SIs and potato AGB at different growth stages.

Pub-SIs	Band Combinations			R ²		
	R _{λ1}	R _{λ2}	R _{λ3}	Tuber Formation	Tuber Bulking	All
RVI	800	670		0.36	0.43	0.62
NDVI	800	680		0.28	0.34	0.52
DVI	800	680		0.06	0.48	0.26
MSAVI	810	670		0.27	0.32	0.50
RDVI	800	670		0.13	0.48	0.39
VLopt	800	670		0.28	0.49	0.61
CCCI	800	720	670	0.28	0.58	0.17
mND705	750	705	445	0.52	0.53	0.47
BNI	434	496	401	0.31	0.53	0.31
NPDI	806	738	560	0.15	0.25	0.23
NDDA	755	680	705	0.50	0.52	0.22
mRER	759	419	742	0.46	0.56	0.44

Pub: the abbreviations of published. R: the abbreviations of reflectance.

The Opt-SIs significantly improved the performances compared to the Pub-SIs. The Opt-SIs explained 42–80% of the variations in potato AGB (Table 3). The predicting ability of three-band Opt-SIs outperformed the two-band Opt-SIs. The optimized NPDI, BNI, and mRER were the best performing Opt-SIs for the tuber formation stage, tuber bulking stage, and the combination of two stages, respectively. Compared with the Pub-SIs, the Opt-SIs had a great variation in the sensitive band combinations due to the influence of the growth stages (Tables 2 and 3). The sensitive bands were mainly located at ultraviolet (350–450 nm), blue (450–500 nm), and near-infrared (NIR, 800–1100 nm) at all three datasets.

Table 3. The band (nm) combination of Opt-SIs and corresponding relationships between Opt-SIs and potato AGB at different growth stages.

Opt-SIs	Tuber Formation				Tuber Bulking				All			
	R _{λ1}	R _{λ2}	R _{λ3}	R ²	R _{λ1}	R _{λ2}	R _{λ3}	R ²	R _{λ1}	R _{λ2}	R _{λ3}	R ²
RVI	608	472		0.59	1096	1094		0.74	820	600		0.66
NDVI	608	472		0.59	1096	1094		0.75	1020	936		0.71
DVI	980	946		0.62	1096	1094		0.75	1014	1008		0.67
MSAVI	944	728		0.63	1096	1094		0.75	1008	936		0.72
RDVI	974	936		0.68	1096	1094		0.74	998	934		0.73
VLOpt	562	324		0.42	772	304		0.56	770	658		0.61
CCCI	402	404	486	0.72	1096	1094	308	0.74	822	986	812	0.71
mND705	492	386	412	0.72	1094	308	1096	0.76	998	934	1148	0.74
BNI	412	404	418	0.74	1094	416	1096	0.78	1020	908	1034	0.75
NPDI	362	444	458	0.76	1084	1096	1094	0.75	946	558	1020	0.72
NDDA	492	406	394	0.73	1096	308	1094	0.75	812	822	986	0.71
mRER	730	760	1140	0.73	498	1096	1094	0.80	932	1146	986	0.73

All: the combination of two growth stages.

Figure 3 shows the accuracy and precision of the AGB estimation based on the best performing Opt-SIs. The results showed that the Opt-SIs had the higher R² (0.39–0.81), lower RMSE (231.8–386.1 kg/ha), and RE% (14.1–25.4%) in different testing datasets compared to Pub-SIs.

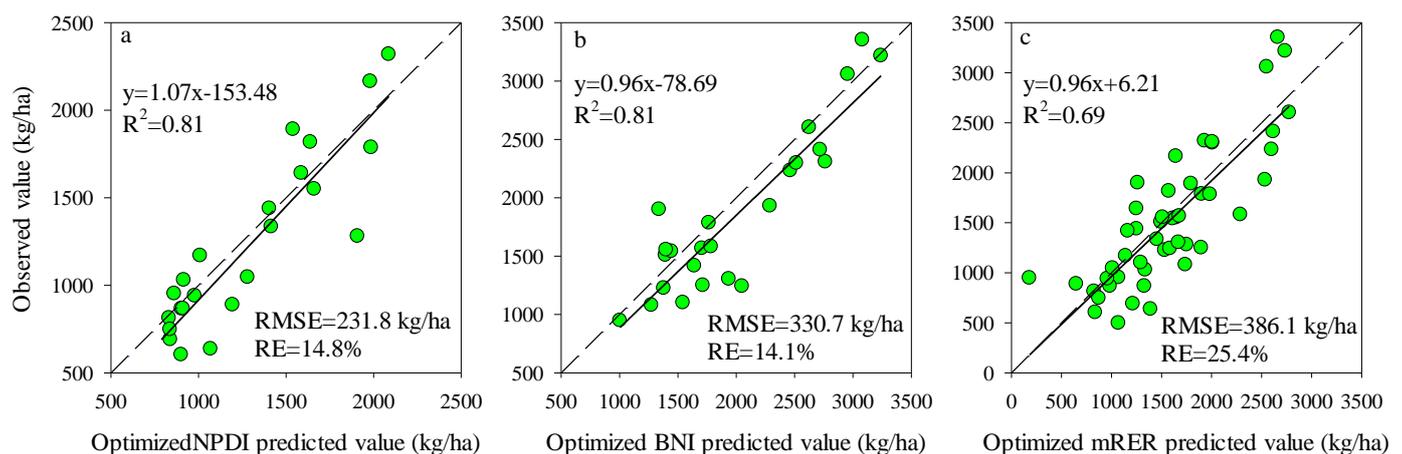


Figure 3. The relationships between the observed value and predicted value for the testing datasets using the best Opt-SIs at (a) tuber formation stage, (b) tuber bulking stage, and (c) the combination of growth stages.

3.3. Estimation of Potato AGB Using RF Model

Comparison analysis in the performance of RF model-based AGB estimation was conducted for the training dataset using different spectrum input variables. The relationships between observed and predicted AGB at different growth stages are presented in Figure 4. The input variables had a significant influence on the predicting ability of RF

models. Compared with the published and Opt-SIs, the full spectrum coupling with the RF model had the highest RMSE and RE% in the estimation of AGB. RF model coupled with the Opt-SIs had the best performances in different stages and decision trees. Using the Opt-SIs as the input variables of the RF model significantly improved the prediction of potato AGB in different training datasets. The number of regression trees, i.e., 500, 2000, and 5000, did not affect the performance of the models, indicating that the initial number of decision trees is adequate to explain the variation of AGB in the training dataset.

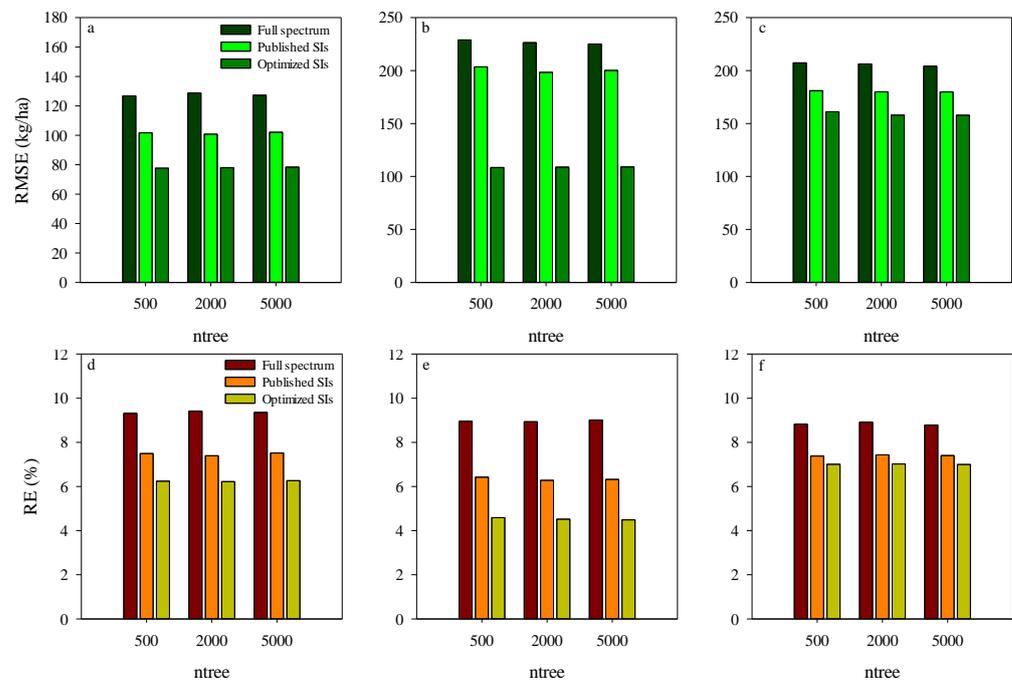


Figure 4. The RMSE and RE of training dataset using RF model based on different input variations at tuber formation stage (a,d), tuber bulking stage (b,e), and the combined stage of tuber formation and tuber bulking (c,f).

To further evaluate the performance of RF model coupling with different input variables, the estimation accuracy of RF models was validated with the testing datasets at each stage. The results showed that the RF models coupled with Opt-SIs had the best predicting performance in the estimation of potato AGB at all growth stages, with the highest R^2 (0.85–0.91), the lowest RMSE (192.2–273.4 kg/ha), and RE% (11.7–13.4%) (Figure 5a–c). Compared to the Opt-SIs, the full spectrum and Pub-SIs coupling with RF models showed the least performance at different growth stages, especially the Pub-SIs due to the influence of growth stages.

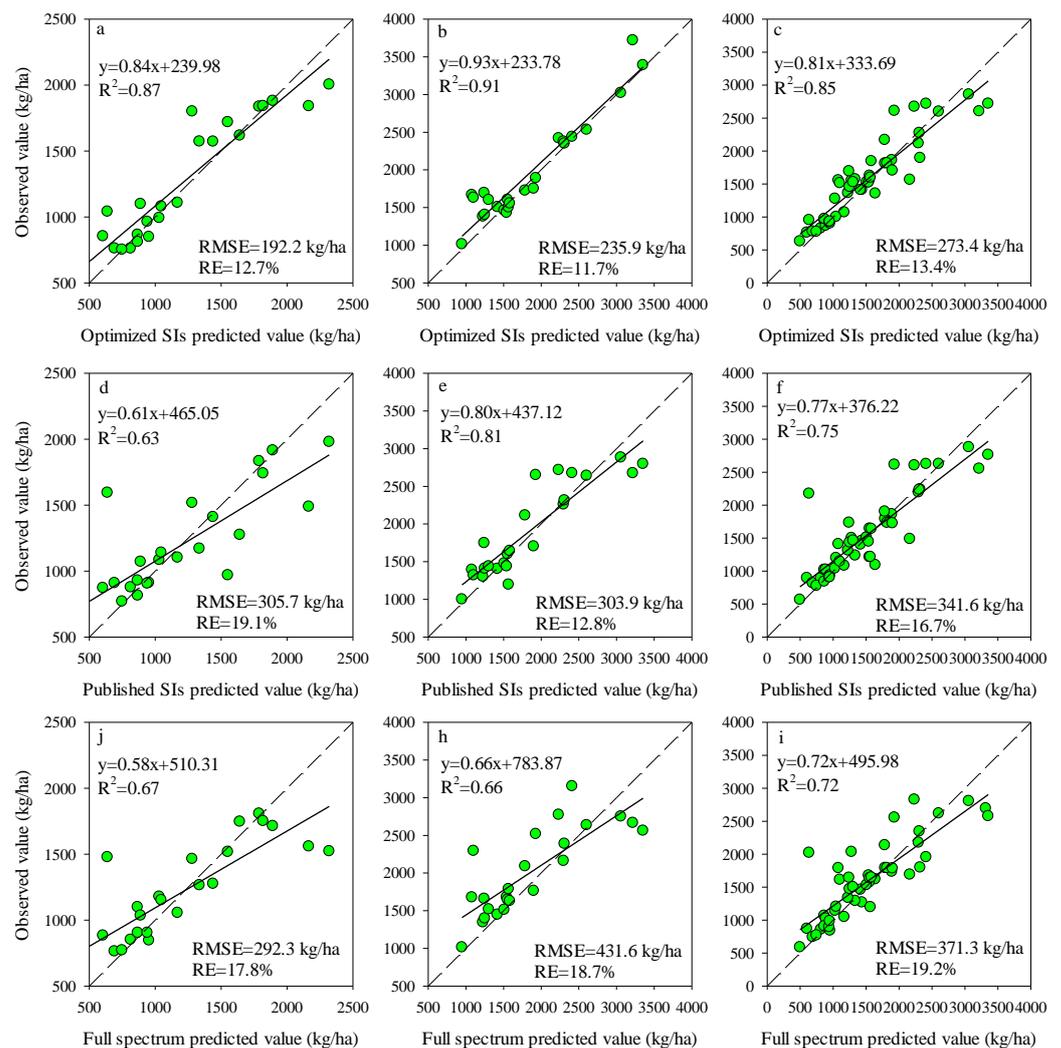


Figure 5. The relationships between the predicted value and observed value for different spectral variables of RF model at tuber formation (a,d,j), tuber bulking (b,e,h), and combined stage (c,f,i).

3.4. The Optimization of RF Model

To obtain the most efficient model for estimating potato AGB, the *ntree* and *mtry* were further optimized using the training dataset for the models of Opt-SIs and full-spectrum coupling with the RF algorithm. For each stage, the *mtry* values from 1 to 12 with an interval of length 1 were tested with the *ntree* values of 150, 300, and 500. The *ntree* and *mtry* values with the lowest RMSE were regarded as the best selection. According to Figure 6, the values of *ntree* and *mtry* were 350 and 5 at the tuber formation (RMSE = 74.9 kg/ha), 150 and 8 at tuber bulking (RMSE = 103.2 kg/ha), and 150 and 3 at the combination of two growth stages (RMSE = 150.8 kg/ha) were the best performing parameters for the models of Opt-SIs coupling with RF algorithm. Compared with the Opt-SIs, the training results of the RF model were poor, with the higher RMSE of 176.4–188.6 kg/ha, 211.0–230.1 kg/ha, and 122.0–134.0 kg/ha at tuber formation, tuber bulking and the combination of growth stage, respectively. The best optimized RF models were tested using the testing datasets and explained 85–91% variation of potato AGB with lower RMSE (185.4–273.5 kg/ha) and RE% (11.1–13.7%) (Figure 7).

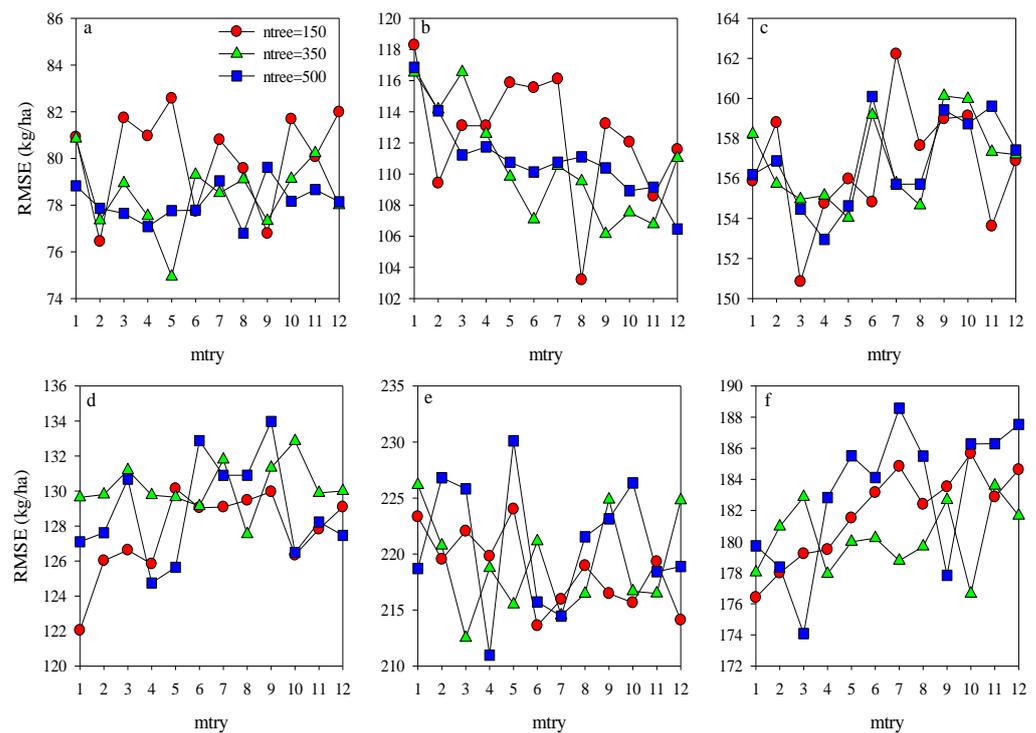


Figure 6. Optimization of random forest parameters (ntree and mtry) using RMSE at different growth stages for training dataset (a–c represent the Opt-SIs coupling with RF algorithm at (a) tuber formation, (b) tuber bulking, and (c) the combined stage; (d–f) represent the full spectrum coupling with RF algorithm at (d) tuber formation, (e) tuber bulking, and (f) the combined stage.).

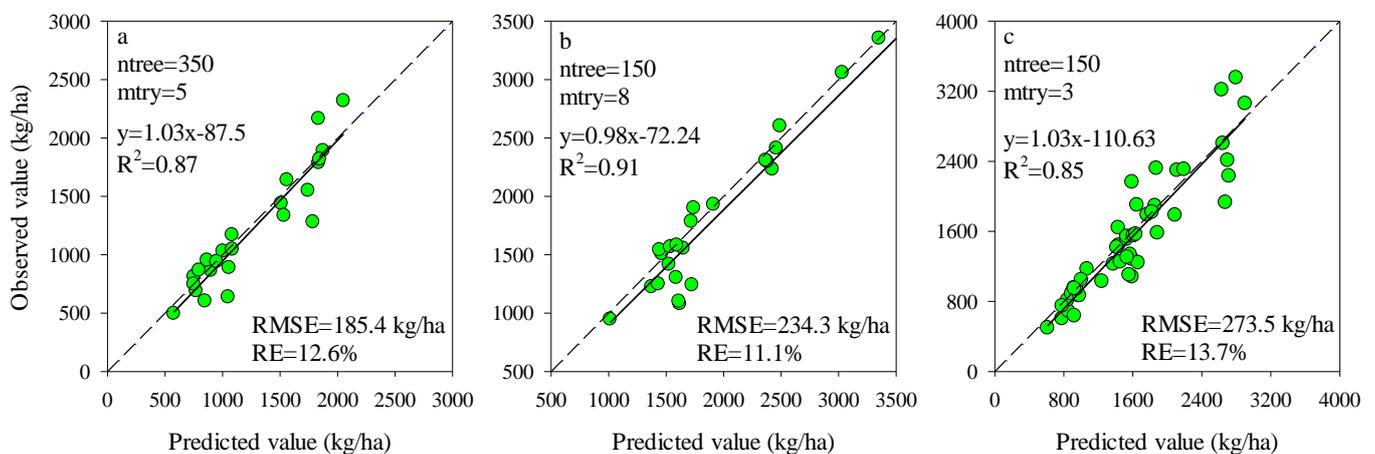


Figure 7. Predictive performance of RF models based on the best modeling in Figure 4 at (a) tuber formation, (b) tuber bulking, and (c) the combined growth stage using testing dataset.

The importance scores of predictors for RF models can be used to screen the input variables of prediction models. The importance scores of Opt-SIs in the best performing RF model were further summarized and analyzed and found that the importance of the Opt-SIs differed among models (Figure 8). The three-band-based Opt-SIs had relatively higher importance scores compared to two-band-based Opt-SIs. The number of model input parameters was optimized according to the importance scores. The results showed that the optimized BNI (412, 404, and 418 nm) and mND705 (492, 386, and 412 nm) coupling with the RF model achieved good prediction results in the training dataset and the testing dataset at the tuber formation stage (Figure 9). For the tuber bulking stage, the best prediction accuracy for potato AGB was achieved using only one predictor: optimized

NPDI, with sensitive bands located at 1084 nm, 1096 nm, and 1094 nm. The RF model based on optimized NPDI can explain 92% variation of potato AGB when the number of trees was 150, the RMSE and RE% were 208.6 kg/ha and 10.3%, respectively (Figure 10). Compared with a single growth stage, the optimized BNI (1020, 908, and 1034 nm) coupling with RF had better performance in the training dataset, while the RMSE, RE%, and R^2 were relatively high in the testing dataset at the combination of two growth stages (Figure 11a,d). However, using two or three Opt-SIs, e.g., optimized BNI (1020, 908 and 1034 nm), NDDA (812, 822 and 986 nm), and mND705 (998, 934, and 1148 nm) enables efficient and accurate prediction of potato AGB in both of the training and testing datasets (Figure 11b,c,e,f). Across the two growth stages, the optimized RF model could explain 86–87% variation of potato AGB.

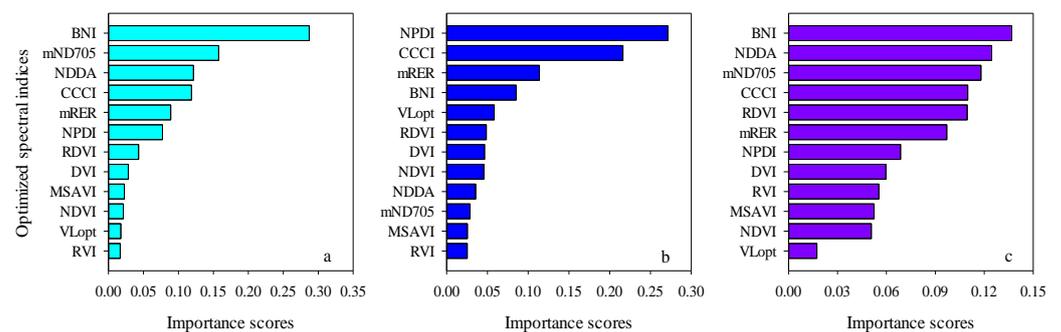


Figure 8. Importance scores both Opt-SIs in the best RF modeling for predicting potato AGB at (a) tuber formation, (b) tuber bulking, and (c) the combined growth stage.

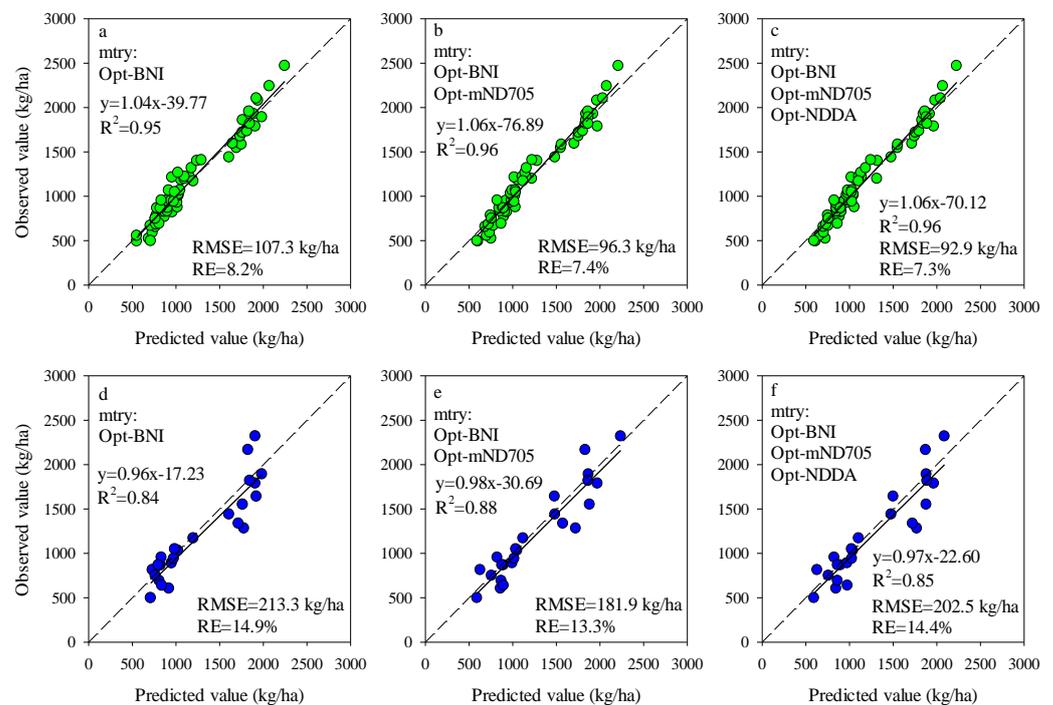


Figure 9. Optimizing the number of Opt-SIs predictors according to important scores at the tuber formation stage for training dataset (a–c) and testing dataset (d–f).

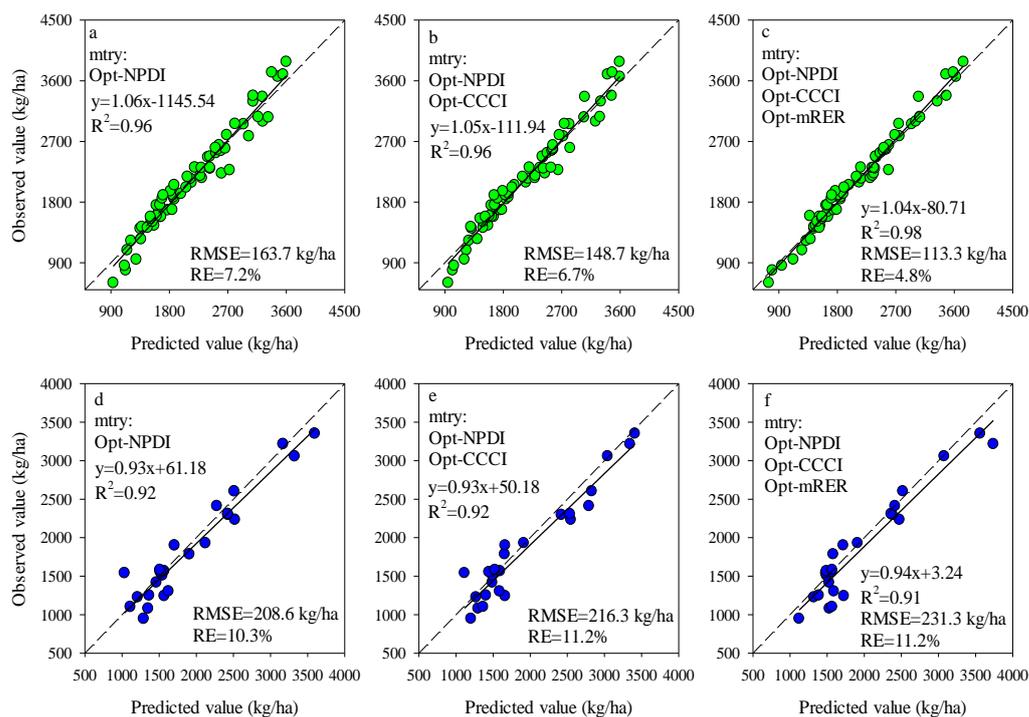


Figure 10. Optimizing the number of Opt-SIs predictors according to important scores at the tuber formation stage for training dataset (a–c) and testing dataset (d–f).

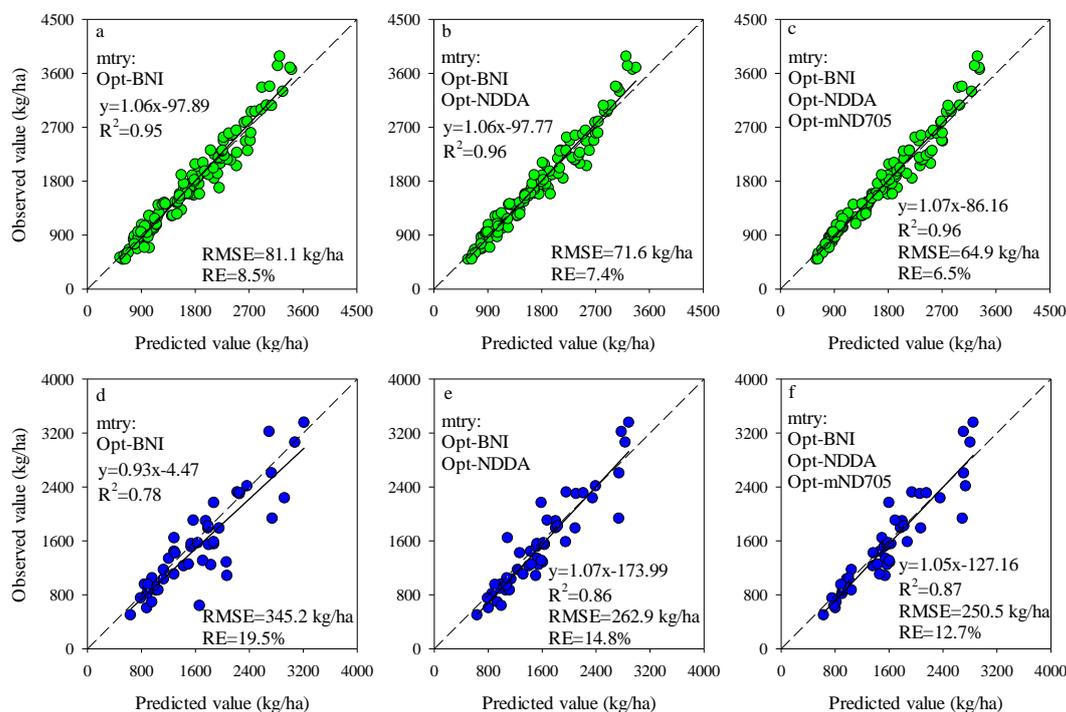


Figure 11. Optimizing the number of Opt-SIs predictors according to important scores at the combination of growth stage for training dataset (a–c) and testing dataset (d–f).

4. Discussion

4.1. The Performances of RF Models Coupling with Different Spectrum Variables

In this study, we investigated the RF model for the estimation of AGB in potato plants. Based on the canopy hyperspectral reflectance data, the full spectrum bands, published and Opt-SIs were compared for their performances of being used as predictors of RF models.

The types and numbers of input variables significantly influenced the performance of RF models in the estimation of potato AGB. The input variables using full-spectrum bands for the RF model had relatively high RMSE and RE for training and testing datasets in different stages. Compared with the full spectrum, the Pub-SIs coupling with the RF could improve by 3–15% of the prediction ability in potato AGB at tuber bulking and the combination of two growth stages. Similar to the findings of Wang [41] and Niu [42], the machine learning models (RF, ANN, SVR, and MLR) coupling with spectral indices had better performances in predicting AGB at different growth stages. Nevertheless, the performances of different machine learning models can be different. For example, compared with ANN and SVR models, SIs combined with the RF algorithm had the best performances in predicting wheat AGB ($R^2 = 0.53–0.79$) [41]. Those results showed that SIs coupling with machine learning is a promising method to predict the AGB. In contrast, the results of Wang [6] suggested that the PLSR model coupling with Pub-SIs failed to improve the biomass estimation accuracy of maize AGB. One reason for the results may be that these Pub-SIs had high multicollinearity [6]. Another major cause might be that the original SIs used probably was insensitive to the estimation of maize AGB. In the current study, selecting 12 typical SIs and optimized the band combinations with different formula formats. Compared with the full spectrum bands and Pub-SIs, the Opt-SIs can significantly enhance the feature information. The Opt-SIs as input variables significantly increased the performance of RF algorithms in the estimation of potato AGB, suggesting that the Opt-SIs could potentially reduce the disturbance of growth stages and sites.

4.2. The Comparison of Sensitive Bands

Extracting sensitive bands from numerous spectral reflectance wavebands is very important for enhancing the prediction accuracy of SIs. At the tuber formation stages, the ultraviolet, violet and blue bands from 350–500 nm and NIR (900–1100 nm) are sensitive regions to potato AGB. However, the NIR (800–1100 nm) regions are important to AGB estimation in potato at tuber bulking and the combination of two growth stages. Similar to the findings of Fu [34], the NIR was the best sensitive region for winter wheat AGB. These areas are sensitive to dry matter and vegetation water content [65]. In contrast, many studies have indicated that the wavelengths in the red edge area contain useful information in the estimation of vegetation AGB [27,66–69]. For example, Kross [69] and Kanke [67] found that the red-edge SIs had better performance estimating crops AGB. Most of these results were reported in specific growth stages in wheat, corn, and rice, which suggest that the crop cultivars and growth stages have a great influence on the selection of sensitive bands. Therefore, the wavebands of SIs for the RF algorithms should be further optimized under different conditions to enhance the estimation accuracy.

Regarding the RF model based on full-spectrum bands, the sensitive bands were mainly located at the NIR radiation (1050–1150 nm) regions at the tuber formation stage. In the tuber bulking, however, the visible bands including ultraviolet (340–360 nm), blue (400–430 nm), and red edge (720–780 nm) were the most sensitive spectral regions to predict potato AGB. When two growth stages were combined, the spectral ranges from the red to red-edge (650–700 nm) and NIR (1080–1100 nm) were sensitive regions for the RF model to predict potato AGB (Figure 12). However, the findings of Dayananda [37] showed that the most important bands contributing to the RF method for biomass estimation were in the wavelength ranges of 546–910 nm (lablab), 750–794 nm (maize), and 686–814 nm (finger millet). These results confirm that the growth stages and crop cultivars have a great influence on the selection of sensitive bands.

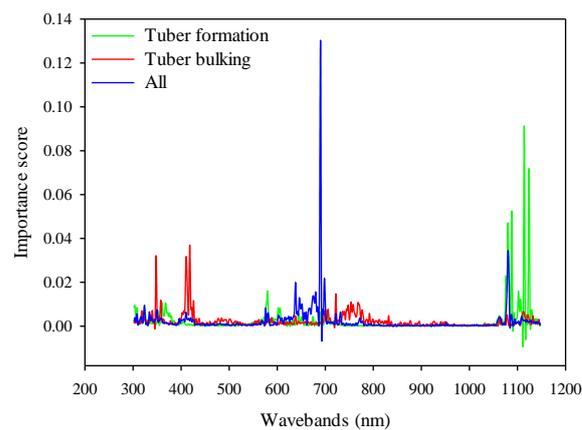


Figure 12. Comparison of the importance score for RF model based on full spectrum at different growth stages.

Compared with the full spectrum, the sensitive bands for the Opt-SIs as input variables were significantly different for each growth stage. One reason for the results is that the SIs combined several bands and specific formula formats to enhance spectral features sensitive to vegetation biochemical properties [70]. The wavebands of SIs contain some reference bands that are insensitive to AGB, and those bands are helpful to increase the signal-to-noise ratio [71]. Although the sensitive bands were different between the uses of Opt-SIs and full-spectrum, the sensitive spectral regions are consistent with some satellite bands, e.g., RapidEye and Sentinel-2A. It has been shown that the spectral indices and spectral bands from satellite image coupling with the artificial neural network algorithm could enhance the performance of the estimation model [72]. Therefore, those broad-band spectral reflectances or corresponding SIs coupling with machine learning algorithms are expected to improve the estimation of AGB on a large scale.

4.3. The Evaluation of Opt-SIs and RF Model

Existing studies have demonstrated that the growth stages have a significant influence on SIs performance in estimating the crops AGB [73,74]. In this study, the saturation effect appeared when the potato AGB was more than 2500 kg/ha (Figure 13). Similar to the current study, the saturation effect was found in wheat [33,34], maize [75], and rice AGB estimation based on SIs [15,76]. Therefore, many studies indicated that a single growth stage could improve the performance for using spectral indices to estimate the vegetable properties [49,77]. Compared to the best performing empirical models based on the optimized BNI, the Opt-SIs coupling with the RF algorithm can overcome the influences of saturation effect and significantly increase the estimation accuracy of potato AGB.

Following the variable selection and calibration, the most important step for RF is model optimization according to the predictor importance scores. It is particularly critical for reducing the probability of high-dimensional problems [78]. In the current study, the high prediction accuracy of potato AGB can be obtained by only using 1–3 Opt-SIs as the predictors of RF compared to the use of full-spectrum (424 bands), which was consistent across different growth stages. Similarly, Fu [34] found that the PLSR models based on optimized normalized difference vegetation index and optimized soil adjusted vegetation index and band parameters produced lower estimation errors (RMSE). Therefore, Opt-SIs provide an efficient way to reducing the high-dimensional problems of spectral analysis while enhancing the performance of machine learning.

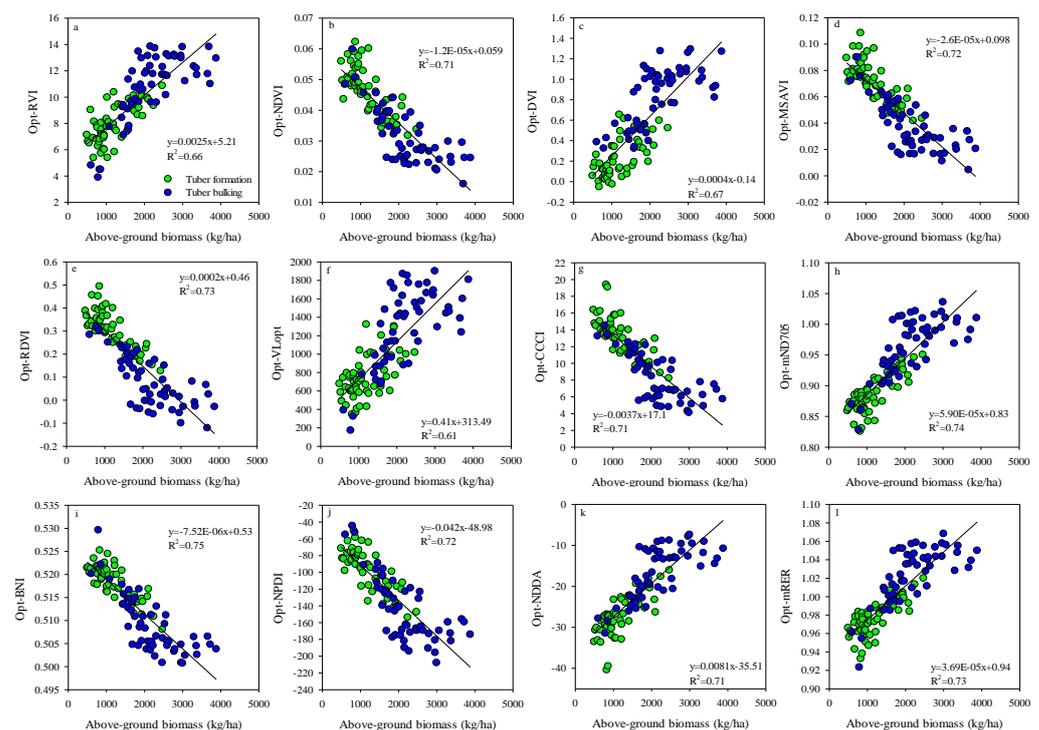


Figure 13. Relationships between potato above-ground biomass and (a) Opt-RVI, (b) Opt-NDVI, (c) Opt-DVI, (d) Opt-MSAVI, (e) Opt-RDVI, (f) Opt-VLopt, (g) Opt-CCCI, (h) Opt-mND705, (i) Opt-BNI, (j) Opt-NPDI (k) Opt-NDDA and (l) Opt-mRER at the combined growth stages.

5. Conclusions

To determine the most informative spectral predictors to be used in the RF model for the estimation of potato AGB, the full spectrum, Pub-SIs and Opt-SIs were used to train RF models using a dataset across growth stages. The Opt-SIs coupling with the RF model can significantly increase the potato AGB estimation accuracy for a single grow stage and across growth stages while significantly reduce the number of input variables. RF model based on the optimized BNI and NDDA could explain 86% of the variations in potato AGB. The Opt-SIs are promising predictors for training RF models for achieving robust and accurate estimation of crop biomass.

Author Contributions: Experiments were designed by F.L.; H.Y., and W.W. undertook the above-ground biomass extractions in the field; H.Y. compiled the data and performed the machine learning analysis; H.Y. wrote the initial draft of the manuscript and F.L. and K.Y. edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Programs for Key Science and Technology Development of Inner Mongolia in 2019 and 2020 (2019GG248 and 2020GG0038), the National Natural Science Foundation of China (41361079), and Ph.D. research startup foundation of Inner Mongolia Agricultural University (BJ08-6).

Data Availability Statement: Not applicable.

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

References

- Maimaitijiang, M.; Ghulam, A.; Sidike, P.; Hartling, S.; Maimaitiyiming, M.; Peterson, K.; Shavers, E.; Fishman, J.; Peterson, J.; Kadam, S.; et al. Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS J. Photogramm. Remote Sens.* **2017**, *134*, 43–58. [[CrossRef](#)]
- Avolio, M.L.; Hoffman, A.M.; Smith, M.D. Linking gene regulation, physiology, and plant biomass allocation in *Andropogon gerardii* in response to drought. *Plant Ecol.* **2018**, *219*, 1–15. [[CrossRef](#)]

3. Li, B.; Xu, X.M.; Zhang, L.; Han, J.W.; Bian, C.S.; Li, G.C.; Liu, J.G.; Jin, L.P. Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 161–172. [[CrossRef](#)]
4. Zhao, B. Determining of a critical dilution curve for plant nitrogen concentration in winter barley. *Field Crop. Res.* **2014**, *160*, 64–72. [[CrossRef](#)]
5. Du, L.J.; Li, Q.; Li, L.; Wu, Y.W.; Zhou, F.; Liu, B.X.; Zhao, B.; Li, X.L.; Liu, Q.L.; Kong, F.L.; et al. Construction of a critical nitrogen dilution curve for maize in Southwest China. *Sci. Rep.* **2020**, *10*, 1–10. [[CrossRef](#)]
6. Wang, C.; Nie, S.; Xi, X.H.; Luo, S.Z.; Sun, X.F. Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sens.* **2017**, *9*, 11. [[CrossRef](#)]
7. Walter, J.; Edwards, J.; McDonald, G.; Kuchel, H. Photogrammetry for the estimation of wheat biomass and harvest index. *Field Crop. Res.* **2018**, *216*, 165–174. [[CrossRef](#)]
8. Jin, X.L.; Zarco-Tejada, P.J.; Schmidhalter, U.; Reynolds, M.P.; Hawkesford, M.J.; Varshney, R.K.; Yang, T.; Nie, C.W.; Li, Z.H.; Ming, B.; et al. High-throughput estimation of crop traits: A review of ground and aerial phenotyping platforms. *IEEE Geosci. Remote. Sens. Mag.* **2020**, 1–33. [[CrossRef](#)]
9. Bendig, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote Sens.* **2014**, *6*, 10395–10412. [[CrossRef](#)]
10. Fu, Y.Y.; Yang, G.J.; Song, X.Y.; Li, Z.H.; Xu, X.G.; Feng, H.K.; Zhao, C.J. Improved Estimation of Winter Wheat Aboveground Biomass Using Multiscale Textures Extracted from UAV-Based Digital Images and Hyperspectral Feature Analysis. *Remote Sens.* **2021**, *13*, 581. [[CrossRef](#)]
11. Yang, S.X.; Feng, Q.S.; Liang, T.G.; Liu, B.K.; Zhang, W.J.; Xie, H.J. Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region. *Remote Sens. Environ.* **2018**, *204*, 448–455. [[CrossRef](#)]
12. 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. [[CrossRef](#)]
13. Ene, L.T.; Gobakken, T.; Andersen, H.E.; Næsset, E.; Cook, B.D.; Morton, D.C.; Morton, H.E.; Babcock, C.; Nelson, R. Large-area hybrid estimation of aboveground biomass in interior Alaska using airborne laser scanning data. *Remote Sens. Environ.* **2018**, *204*, 741–755. [[CrossRef](#)]
14. Zhang, Y.Z.; Liang, S.L.; Yang, L. A review of regional and global gridded forest biomass datasets. *Remote Sens.* **2019**, *11*, 2744. [[CrossRef](#)]
15. Zheng, H.B.; Cheng, T.; Zhou, M.; Li, D.; Yao, X.; Tian, Y.C.; Cao, W.X.; Zhu, Y. Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. *Precis. Agric.* **2019**, *20*, 611–629. [[CrossRef](#)]
16. Li, C.; Zhou, L.; Xu, W. Estimating Aboveground Biomass Using Sentinel-2 MSI Data and Ensemble Algorithms for Grassland in the Shengjin Lake Wetland, China. *Remote Sens.* **2021**, *13*, 1595. [[CrossRef](#)]
17. Venancio, L.P.; Mantovani, E.C.; do Amaral, C.H.; Neale, C.M.U.; Gonçalves, I.Z.; Filgueiras, R.; Eugenio, F.C. Potential of using spectral vegetation indices for corn green biomass estimation based on their relationship with the photosynthetic vegetation sub-pixel fraction. *Agric. Water Manag.* **2020**, 236. [[CrossRef](#)]
18. Zhu, Y.H.; Zhao, C.J.; Yang, H.; Yang, G.J.; Han, L.; Li, Z.H.; Feng, H.K.; Xu, B.; Wu, J.T.; Lei, L. Estimation of maize above-ground biomass based on stem-leaf separation strategy integrated with LiDAR and optical remote sensing data. *PeerJ* **2019**, *7*. [[CrossRef](#)]
19. Yue, J.B.; Yang, G.J.; Li, C.C.; Li, Z.H.; Wang, Y.J.; Feng, H.K.; Xu, B. Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sens.* **2017**, *9*, 708. [[CrossRef](#)]
20. Meng, J.H.; Du, X.; Wu, B.F. Generation of high spatial and temporal resolution NDVI and its application in crop biomass estimation. *Int. J. Digit. Earth.* **2013**, *6*, 203–218. [[CrossRef](#)]
21. Cen, H.Y.; Wan, L.; Zhu, J.P.; Li, Y.J.; Li, X.R.; Zhu, Y.M.; Weng, H.Y.; Wu, W.K.; Yin, W.X.; Xu, C.; et al. Dynamic monitoring of biomass of rice under different nitrogen treatments using a lightweight UAV with dual image-frame snapshot cameras. *Plant Methods* **2019**, *15*, 1–16. [[CrossRef](#)] [[PubMed](#)]
22. Li, F.; Mistele, B.; Hu, Y.C.; Chen, X.P.; Schmidhalter, U. Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression. *Eur. J. Agron.* **2014**, *52*, 198–209. [[CrossRef](#)]
23. Stroppiana, D.; Boschetti, M.; Brivio, P.A.; Bocchi, S. Plant nitrogen concentration in paddy rice from field canopy hyperspectral radiometry. *Field Crop. Res.* **2009**, *111*, 119–129. [[CrossRef](#)]
24. Mariotto, I.; Thenkabail, P.S.; Huete, A.; Slonecker, E.T.; Platonov, A. Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HypSPiRI mission. *Remote Sens. Environ.* **2013**, *139*, 291–305. [[CrossRef](#)]
25. Rivera, J.P.; Verrelst, J.; Delegido, J.; Veroustraete, F.; Moreno, J. On the semi-automatic retrieval of biophysical parameters based on spectral index optimization. *Remote Sens.* **2014**, *6*, 4927–4951. [[CrossRef](#)]
26. Gnyp, M.L.; Miao, Y.X.; Yuan, F.; Ustin, S.L.; Yu, K.; Yao, Y.K.; Huang, S.Y.; Bareth, G. Hyperspectral canopy sensing of paddy rice aboveground biomass at different growth stages. *Field Crop. Res.* **2014**, *155*, 42–55. [[CrossRef](#)]
27. Li, F.; Mistele, B.; Hu, Y.C.; Chen, X.P.; Schmidhalter, U. Optimising three-band spectral indices to assess aerial N concentration, N uptake and aboveground biomass of winter wheat remotely in China and Germany. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 112–123. [[CrossRef](#)]
28. Schirrmann, M.; Giebel, A.; Gleiniger, F.; Pflanz, M.; Lentschke, J.; Dammer, K.H. Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. *Remote Sens.* **2016**, *8*, 706. [[CrossRef](#)]

29. Verrelst, J.; Malenovsky, Z.; Van der Tol, C.; Camps-Valls, G.; Gastellu-Etchegorry, J.P.; Lewis, P.; North, P.; Moreno, J. Quantifying vegetation biophysical variables from imaging spectroscopy data: A review on retrieval methods. *Surv. Geophys.* **2019**, *40*, 589–629. [[CrossRef](#)]
30. Verrelst, J.; Camps-Valls, G.; Muñoz-Marí, J.; Rivera, J.P.; Veroustraete, F.; Clevers, J.G.P.W.; Moreno, J. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—A review. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 273–290. [[CrossRef](#)]
31. Gao, X.; Huete, A.R.; Ni, W.; Miura, T. Optical-biophysical relationships of vegetation spectra without background contamination. *Remote Sens. Environ.* **2000**, *74*, 609–620. [[CrossRef](#)]
32. Li, F.; Miao, Y.X.; Chen, X.P.; Zhang, H.L.; Jia, L.L.; Bareth, G. Estimating winter wheat biomass and nitrogen status using an active crop sensor. *Intell. Autom. Soft Comput.* **2010**, *16*, 1221–1230.
33. Erdle, K.; Mistele, B.; Schmidhalter, U. Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. *Field Crop. Res.* **2011**, *124*, 74–84. [[CrossRef](#)]
34. Fu, Y.Y.; Yang, G.J.; Wang, J.H.; Song, X.Y.; Feng, H.K. Winter wheat biomass estimation based on spectral indices, band depth analysis and partial least squares regression using hyperspectral measurements. *Comput. Electron. Agric.* **2014**, *100*, 51–59. [[CrossRef](#)]
35. Wang, J.J.; Chen, Y.Y.; Chen, F.Y.; Shi, T.Z.; Wu, G.F. Wavelet-based coupling of leaf and canopy reflectance spectra to improve the estimation accuracy of foliar nitrogen concentration. *Agric. For. Meteorol.* **2018**, *248*, 306–315. [[CrossRef](#)]
36. Han, L.; Yang, G.J.; Dai, H.Y.; Xu, B.; Yang, H.; Feng, H.K.; Li, Z.H.; Yang, X.D. Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods* **2019**, *15*, 1–19. [[CrossRef](#)]
37. Dayananda, S.; Astor, T.; Wijesingha, J.; Thimappa, S.C.; Chowdappa, H.D.; Mudalagiriappa; Nidamanuri, R.R.; Nautiyal, S.; Wachendorf, M. Multi-temporal monsoon crop biomass estimation using hyperspectral imaging. *Remote Sens.* **2019**, *11*, 1771. [[CrossRef](#)]
38. Yu, K.; Gnyp, M.L.; Gao, J.; Miao, Y.; Chen, X.; Bareth, G. Using Partial Least Squares (PLS) to Estimate Canopy Nitrogen and Biomass of Paddy Rice in China’s Sanjiang Plain. In *Proceedings of the Workshop on UAV-Based Remote Sensing Methods for Monitoring Vegetation, Cologne, Germany, 9–10 June 2013*; Bendig, J., Bareth, G., Eds.; Kölner Geographische Arbeiten: Cologne, Germany, 2014; Volume 94, pp. 99–103.
39. Wilkes, P.; Disney, M.; Vicari, M.B.; Calders, K.; Burt, A. Estimating urban above ground biomass with multi-scale LiDAR. *Carbon Balanc. Manag.* **2018**, *13*, 1–20. [[CrossRef](#)] [[PubMed](#)]
40. Yue, J.B.; Feng, H.K.; Yang, G.J.; Li, Z.H. A comparison of regression techniques for estimation of above-ground winter wheat biomass using near-surface spectroscopy. *Remote Sens.* **2018**, *10*, 66. [[CrossRef](#)]
41. Wang, L.A.; Zhou, X.D.; Zhu, X.K.; Dong, Z.D.; Guo, W.S. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop. J.* **2016**, *4*, 212–219. [[CrossRef](#)]
42. Niu, Y.X.; Zhang, L.Y.; Zhang, H.H.; Han, W.T.; Peng, X.S. Estimating above-ground biomass of maize using features derived from UAV-based RGB imagery. *Remote Sens.* **2019**, *11*, 1261. [[CrossRef](#)]
43. Jordan, C.F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
44. Rouse, J.W., Jr.; Haas, R.H.; Deering, D.W.; Schell, J.A.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. [Great Plains Corridor]*; NASA: Washington, DC, USA, 1974.
45. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
46. 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. [[CrossRef](#)]
47. Thenkabail, P.S.; Smith, R.B.; De Pauw, E. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sens. Environ.* **2000**, *71*, 158–182. [[CrossRef](#)]
48. Le Maire, G.; François, C.; Soudani, K.; Berveiller, D.; Pontailier, J.Y.; Bréda, N.; Genet, H.; Davi, H.; Dufrêne, E. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sens. Environ.* **2008**, *112*, 3846–3864. [[CrossRef](#)]
49. Yu, K.; Li, F.; Gnyp, M.L.; Miao, Y.X.; Bareth, G.; Chen, X. Remotely detecting canopy nitrogen concentration and uptake of paddy rice in the Northeast China Plain. *ISPRS J. Photogramm. Remote Sens.* **2013**, *78*, 102–115. [[CrossRef](#)]
50. Hasituya; Li, F.; Elsayed, S.; Hu, Y.C.; Schmidhalter, U. Passive reflectance sensing using optimized two-and three-band spectral indices for quantifying the total nitrogen yield of maize. *Comput. Electron. Agric.* **2020**, *105*, 403. [[CrossRef](#)]
51. Li, F.; Mistele, B.; Hu, Y.; Yue, X.; Yue, S.C.; Miao, Y.X.; Schmidhalter, U. Remotely estimating aerial N status of phenologically differing winter wheat cultivars grown in contrasting climatic and geographic zones in China and Germany. *Field Crop. Res.* **2012**, *138*, 21–32. [[CrossRef](#)]
52. Chi, D.; Degerickx, J.; Yu, K.; Somers, B. Urban Tree Health Classification Across Tree Species by Combining Airborne Laser Scanning and Imaging Spectroscopy. *Remote Sens.* **2020**, *12*, 2435. [[CrossRef](#)]
53. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
54. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* **2006**, *9*, 181–199. [[CrossRef](#)]

55. *Introductory Statistics*, Textbook Equity ed.; OpenStax College, Rice University: Houston, TX, USA, 2013; Volume 1, ISBN 978-1-304-89164-8.
56. Li, F.; Miao, Y.; Hennig, S.D.; Gnyp, M.L.; Chen, X.P.; Jia, L.; Bareth, G. Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages. *Precis. Agric.* **2010**, *11*, 335–357. [[CrossRef](#)]
57. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geosci. Model. Dev.* **2014**, *7*, 1247–1250. [[CrossRef](#)]
58. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [[CrossRef](#)]
59. Roujean, J.L.; Breon, F.M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [[CrossRef](#)]
60. Reyniers, M.; Walvoort, D.J.; De Baardemaaker, J. A linear model to predict with a multi-spectral radiometer the amount of nitrogen in winter wheat. *Int. J. Remote Sens.* **2006**, *27*, 4159–4179. [[CrossRef](#)]
61. Fitzgerald, G.; Rodriguez, D.; O’Leary, G. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index—The canopy chlorophyll content index (CCCI). *Field Crop. Res.* **2010**, *116*, 318–324. [[CrossRef](#)]
62. Tian, Y.C.; Yao, X.; Yang, J.; Cao, W.X.; Hannaway, D.B.; Zhu, Y. Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground-and space-based hyperspectral reflectance. *Field Crop. Res.* **2011**, *120*, 299–310. [[CrossRef](#)]
63. Feng, W.; Guo, B.B.; Wang, Z.J.; He, L.; Song, X.; Wang, Y.H.; Guo, T.C. Measuring leaf nitrogen concentration in winter wheat using double-peak spectral reflection remote sensing data. *Field Crop. Res.* **2014**, *159*, 43–52. [[CrossRef](#)]
64. Feng, W.; Guo, B.B.; Zhang, H.Y.; He, L.; Zhang, Y.S.; Wang, Y.H.; Guo, T.C. Remote estimation of above ground nitrogen uptake during vegetative growth in winter wheat using hyperspectral red-edge ratio data. *Field Crop. Res.* **2015**, *180*, 197–206. [[CrossRef](#)]
65. Bowyer, P.; Danson, F.M. Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level. *Remote Sens. Environ.* **2004**, *92*, 297–308. [[CrossRef](#)]
66. Ren, H.; Zhou, G.S.; Zhang, X.S. Estimation of green aboveground biomass of desert steppe in Inner Mongolia based on red-edge reflectance curve area method. *Biosyst. Eng.* **2011**, *109*, 385–395. [[CrossRef](#)]
67. Manjunath, K.R.; Ray, S.S.; Panigrahy, S. Discrimination of spectrally-close crops using ground-based hyperspectral data. *J. Indian Soc. Remote Sens.* **2011**, *39*, 599–602. [[CrossRef](#)]
68. Kanke, Y.; Tubana, B.; Dalen, M.; Harrell, D. Evaluation of red and red-edge reflectance-based vegetation indices for rice biomass and grain yield prediction models in paddy fields. *Precis. Agric.* **2016**, *17*, 507–530. [[CrossRef](#)]
69. Kross, A.; McNairn, H.; Lapen, D.; Sunohara, M.; Champagne, C. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 235–248. [[CrossRef](#)]
70. Glenn, E.P.; Huete, A.R.; Nagler, P.L.; Nelson, S.G. Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors* **2008**, *8*, 2136–2160. [[CrossRef](#)]
71. Jay, S.; Gorretta, N.; Morel, J.; Maupas, F.; Bendoula, R.; Rabatel, G.; Baret, F.; Dutartre, D.; Comar, A. Estimating leaf chlorophyll content in sugar beet canopies using millimeter-to centimeter-scale reflectance imagery. *Remote Sens. Environ.* **2017**, *198*, 173–186. [[CrossRef](#)]
72. Guo, Y.M.; Ni, J.; Liu, L.B.; Wu, Y.Y.; Guo, C.Z.; Xu, X.; Zhong, Q.L. Estimating aboveground biomass using Pléiades satellite image in a karst watershed of Guizhou Province, Southwestern China. *J. Mt. Sci.* **2018**, *15*, 1020–1034. [[CrossRef](#)]
73. Zhang, J.H.; Wang, K.; Bailey, J.S.; Wang, R.C. Predicting nitrogen status of rice using multispectral data at canopy scale. *Pedosphere* **2006**, *16*, 108–117. [[CrossRef](#)]
74. Poley, L.G.; McDermid, G.J. A Systematic Review of the Factors Influencing the Estimation of Vegetation Aboveground Biomass Using Unmanned Aerial Systems. *Remote Sens.* **2020**, *12*, 1052. [[CrossRef](#)]
75. Li, W.; Niu, Z.; Huang, N.; Wang, C.; Gao, S.; Wu, C.Y. Airborne LiDAR technique for estimating biomass components of maize: A case study in Zhangye City, Northwest China. *Ecol. Indic.* **2015**, *57*, 486–496. [[CrossRef](#)]
76. Jiang, Q.; Fang, S.H.; Peng, Y.; Gong, Y.; Zhu, R.S.; Wu, X.T.; Duan, B.; Ma, Y.; Liu, J. UAV-based biomass estimation for rice-combining spectral, TIN-based structural and meteorological features. *Remote Sens.* **2019**, *11*, 890. [[CrossRef](#)]
77. Luo, S.; He, Y.B.; Li, Q.; Jiao, W.H.; Zhu, Y.Q.; Zhao, X.H. Nondestructive estimation of potato yield using relative variables derived from multi-period LAI and hyperspectral data based on weighted growth stage. *Plant Methods* **2020**, *16*, 1–14. [[CrossRef](#)] [[PubMed](#)]
78. Strobel, J.; Hawkins, C. An exploration of design phenomena in second life. In Proceedings of the E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, Vancouver, BC, Canada, 26–30 October 2009; pp. 3702–3709.