





Article

A Leaf Chlorophyll Content Estimation Method for *Populus deltoides* (*Populus deltoides* Marshall) Using Ensembled Feature Selection Framework and Unmanned Aerial Vehicle Hyperspectral Data

Zhulin Chen ^{1,2}, Xuefeng Wang ^{1,2,*} , Shijiao Qiao ³, Hao Liu ^{1,4}, Mengmeng Shi ^{1,2} , Haiying Jiang ⁵  and Huimin Zou ⁶ 

- ¹ Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China; chenzl@ifrit.ac.cn (Z.C.); liuhao@ifrit.ac.cn (H.L.); mengms_182@163.com (M.S.); chenxj@ifrit.ac.cn (X.C.)
 - ² State Forestry and Grassland Administration, Key Laboratory of Forest Management and Growth Modelling, Beijing 100091, China
 - ³ State Key Laboratory of Remote Sensing Science, Innovation Research Center of Satellite Application, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; 202321051206@mail.bnu.edu.cn
 - ⁴ Key Laboratory of Forestry Remote Sensing and Information System, National Forestry and Grassland Administration, Beijing 100091, China
 - ⁵ College of Computer Science and Technology, Harbin Engineering University, Harbin 150006, China; jianghaiying@hrbeu.edu.cn
 - ⁶ National Ocean Technology Center, Tianjin 300112, China; hmzou1349@163.com
- * Correspondence: xuefeng@ifrit.ac.cn

Abstract: Leaf chlorophyll content (LCC) is a key indicator in representing the photosynthetic capacity of *Populus deltoides* (*Populus deltoides* Marshall). Unmanned aerial vehicle (UAV) hyperspectral imagery provides an effective approach for LCC estimation, but the issue of band redundancy significantly impacts model accuracy and computational efficiency. Commonly used single feature selection algorithms not only fail to balance computational efficiency with optimal set search but also struggle to combine different regression algorithms under dynamic set conditions. This study proposes an ensemble feature selection framework to enhance LCC estimation accuracy using UAV hyperspectral data. Firstly, the embedded algorithm was improved by introducing the SHapley Additive exPlanations (SHAP) algorithm into the ranking system. A dynamic ranking strategy was then employed to remove bands in steps of 10, with LCC models developed at each step to identify the initial band subset based on estimation accuracy. Finally, the wrapper algorithm was applied using the initial band subset to search for the optimal band subset and develop the corresponding model. Three regression algorithms including gradient boosting regression trees (GBRT), support vector regression (SVR), and gaussian process regression (GPR) were combined with this framework for LCC estimation. The results indicated that the GBRT-Optimal model developed using 28 bands achieved the best performance with R^2 of 0.848, RMSE of 1.454 $\mu\text{g}/\text{cm}^2$ and MAE of 1.121 $\mu\text{g}/\text{cm}^2$. Compared with a model performance that used all bands as inputs, this optimal model reduced the RMSE value by 24.37%. In addition to estimating biophysical and biochemical parameters, this method is also applicable to other hyperspectral imaging tasks.

Keywords: leaf chlorophyll content estimation; hyperspectral data; unmanned aerial vehicle; feature selection; *Populus deltoides* Marshall



Citation: Chen, Z.; Wang, X.; Qiao, S.; Liu, H.; Shi, M.; Chen, X.; Jiang, H.; Zou, H. A Leaf Chlorophyll Content Estimation Method for *Populus deltoides* (*Populus deltoides* Marshall) Using Ensembled Feature Selection Framework and Unmanned Aerial Vehicle Hyperspectral Data. *Forests* **2024**, *15*, 1971. <https://doi.org/10.3390/f15111971>

Academic Editor: Henning Buddenbaum

Received: 27 September 2024
Revised: 4 November 2024
Accepted: 6 November 2024
Published: 8 November 2024



Copyright: © 2024 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

Populus deltoides (*Populus deltoides* Marshall) is an important fast-growing timber species extensively cultivated across the temperate plain regions of China [1]. Renowned for its rapid growth, early maturation, high productivity, and easy regeneration, *Populus deltoides* has extensive applications in both ecological protection forestry and industrial timber

plantations [2]. There are abundant germplasm resources of *Populus deltoides* with various phenotypic traits. Understanding these traits not only provides a solid foundation for biological breeding but also offers valuable insights for developing improved varieties [3]. Therefore, investigating these phenotypic traits is essential for selecting superior genotypes that can adapt to various environmental conditions and management practices.

Leaf chlorophyll content (LCC) is a critical phenotypic trait that reflects photosynthetic capacity and the nitrogen status of plants. Monitoring LCC levels and their changes is essential for identifying *Populus deltoides* genotypes with strong resistance to drought and saline-alkali conditions [4]. Traditional LCC determination methods rely on laboratory-based chemical analyses, which are destructive, time-consuming, and unsuitable for accurately estimating large areas. In contrast, remote sensing technology offers a rapid and cost-effective approach to LCC estimation [5]. Previous studies have demonstrated that the green (G), red (R), and red-edge (RE) spectral regions exhibit significant absorption and reflection characteristics associated with LCC [6–8]. These distinctive spectral responses provide a theoretical foundation for LCC inversion using remote sensing data. Currently, data from satellites such as Landsat-8 and Sentinel-2 have been successfully used in LCC estimation across a wide range of crops [9,10]. However, compared to multispectral data, hyperspectral data offers hundreds of narrow bands that enable more precise detection of subtle variations in canopy reflectance [11]. Despite these advantages, the limited number of hyperspectral satellites in orbit and their low spatial resolution restrict their capability for fine-scale LCC monitoring. In recent years, unmanned aerial vehicle (UAV) remote sensing has rapidly advanced due to its high spatial resolution (at the centimeter level), low cost, and flexibility to meet specific flight requirements [12]. Supported by the UAV systems, hyperspectral data have been easily acquired and applied for precision agriculture and forestry LCC estimation [13,14]. In addition to using only optical sensors, recent studies have also employed LiDAR for LCC estimation. Some researchers have explored the potential of integrating optical sensors and LiDAR to improve the accuracy of LCC estimation [15,16]. In these studies, LiDAR data were used for leaf area index extraction to mitigate the influence of canopy structure when developing LCC estimation models. Other studies have attempted to integrate optical and LiDAR data to generate the three-dimensional distribution of LCC [17,18]. These methods provide a more comprehensive understanding of biochemical pigment distribution within canopies. Moreover, some studies have demonstrated that LiDAR-derived intensity images can serve as substitutes for infrared bands [19,20]. Therefore, LiDAR data can be integrated with RGB imagery to enhance spectral information, providing a more cost-effective approach for LCC estimation.

Although hyperspectral remote sensing has advantages in LCC estimation, its high dimensionality complicates model construction by increasing computational complexity and the risk of overfitting [21]. Therefore, the original hyperspectral data are often processed using dimensionality reduction algorithms. To ensure the interpretability of the estimation model, algorithms that perform mapping and transformation within the same feature space are preferred. These algorithms are referred to as feature selection methods. From a strategic perspective, feature selection algorithms can be categorized into three types: filter, wrapper, and embedded, each with its advantages and disadvantages [22]. Filter algorithms independently evaluate and rank features based on their statistical significance or relevance to the target variable. However, they have a limited capability to account for interactions and dependencies among features during the selection process. In contrast, wrapper algorithms consider feature interactions with a predefined learning algorithm and are capable of searching for an optimal subset of features. Nonetheless, these algorithms are more time-consuming due to the large search space, making them unsuitable for high-dimensional data. Embedded algorithms combine the strengths of both filter and wrapper methods by efficiently evaluating feature relevance during model training. Among these three types, embedded methods are commonly used for processing hyperspectral data. For instance, Zhang et al. [23] employed the least absolute shrinkage and selection operator (LASSO) algorithm to select important hyperspectral bands for apple tree LCC estimation.

Zhao et al. [24] used the random forest (RF) algorithm to select sensitive wavelengths for in situ LCC estimation. However, although embedded algorithms are preferred in relative studies, the feature ranking provided is often constrained by the predefined regression algorithm, which limits its adaptability. For example, previous study has demonstrated that features selected using RF algorithm achieve lower accuracy when applied with other regression algorithms such as support vector regression (SVR) [25]. More importantly, the importance of features changes as the number of features in the model varies, but embedded algorithms fail to account for the impact of dynamic feature sets on ranking. Additionally, a single-type feature selection algorithm often fails to balance computational efficiency with the search for an optimal subset. Therefore, it is crucial to develop a multitype integrated feature selection algorithm framework to enhance applicability across different scenarios.

Besides choosing appropriate features, another way to improve model performance is to find a suitable regression algorithm. According to previous studies, LCC estimation can be conducted using physical models, parametric models, and non-parametric models [26,27]. Physical models, such as the radiative transfer model, are grounded in physical laws and are less affected by sample characteristics or environmental background. However, they are typically highly complex and require a large number of input parameters for the accurate simulation of surface conditions. The number of these input parameters often exceeds the available remote sensing observations, making parameter inversion particularly challenging [28]. Parametric algorithms are used to estimate the parameters of a statistical model or probability distribution based on observed data. They assume that the data come from a known form of distribution and employ statistical techniques to determine the parameters. Due to the specific distributional assumption of data, parametric models have limited flexibility and generalization ability. In contrast, non-parametric algorithms rely solely on observed data to define the model structure. They are more flexible and better suited to handling complex data distributions. Machine learning (ML) is a non-parametric method that is often used as the predefined algorithm in the wrapper and embedded feature selection algorithm. By capturing complex nonlinear relationships between canopy reflectance and LCC, ML models often achieve more accurate estimations than traditional linear regression algorithms [29,30]. However, the performance of different ML algorithms varies depending on the data characteristics and specific tasks. Therefore, to improve the accuracy of LCC estimation models for *Populus deltoides*, it is still necessary to evaluate different ML algorithms.

This study aims to develop an accurate LCC estimation algorithm for *Populus deltoides* based on UAV hyperspectral data. The objectives of this study are outlined as follows: (1) improving embedded feature selection algorithm to be applicable to different regression algorithms and adaptable to dynamic feature set conditions; (2) developing an ensembled feature selection framework by integrating the advantages of various algorithms; (3) evaluating different ML regression algorithms in the LCC estimation model for *Populus deltoides* based on the ensembled feature selection framework.

2. Materials and Methods

2.1. Study Area

Figure 1 shows the geographic location of the study area for data collection. It is located within the National Superior *Populus* Seed Base in Shishou county, Hubei province, China. This county lies in the southern part of Hubei, along the middle reaches of the Yangtze River. The region is predominantly flat, with some low hills in the southwest. Shishou exhibits a subtropical monsoon climate with distinct seasonal variations. The average annual temperature ranges from 16 °C to 18 °C, with hot, rainy summers and mild winters. The county receives abundant rainfall, averaging between 1200 mm and 1400 mm annually, mostly concentrated during the summer. Rich in natural resources, especially water and wetland ecosystems, the area plays a crucial role in supporting agriculture and biodiversity. In Figure 1, the planting area of *Populus deltoides* is approximately 83,600 m². It contains over 3000 *Populus deltoides* trees with different genotypes.

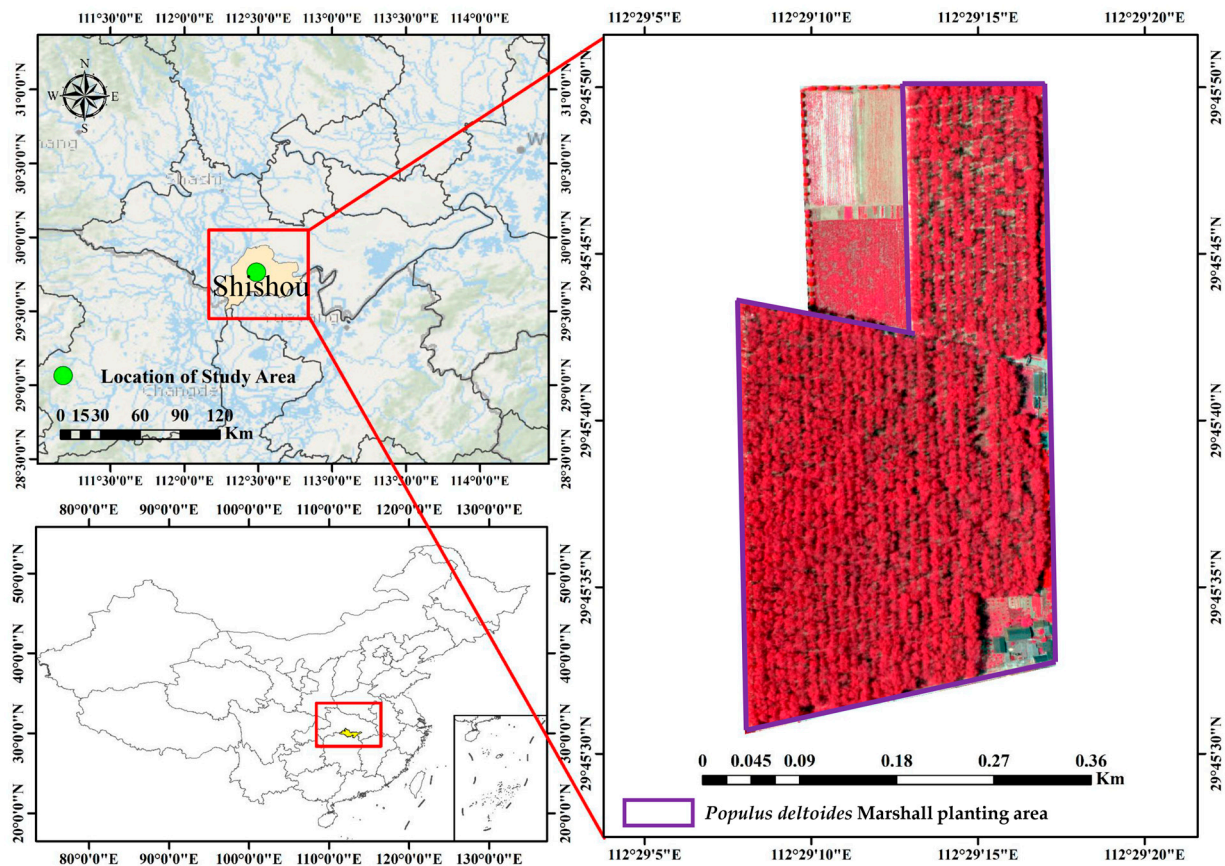


Figure 1. The location of study area.

2.2. Data

2.2.1. UAV Hyperspectral Data Acquisition and Preprocessing

As shown in Figure 2, hyperspectral imagery was acquired using a 300TC hyperspectral camera (Beijing IRIS Remote Sensing Technology Limited, Inc., Beijing, China) loaded on a DJI M350 four-rotator UAV (Shenzhen Dajiang Innovation Technology Co., Ltd., Shenzhen, China). This camera has 308 spectral channels ranging from 393 to 1007 nm, with a spectral resolution of 2 nm. The UAV flight mission was conducted on May 14th between 10:00 and 15:00 under clear sky and low wind conditions. The spatial resolution of the hyperspectral imagery is 20 cm per pixel. Before the flight, radiation calibration was performed using a target cloth covered with Lambertian coating. The radiometric, geometric, and orthorectification processes were all conducted using the Mega Cube software (version 11.0.13), which developed independently by IRIS Corporation. After preprocessing, the hyperspectral reflectance data were acquired. Based on previous studies and the result of reflectance checks, bands above 891 nm were excluded due to spectral noise and distortions [31]. The remaining bands were then smoothed using the Savitzky–Golay filter [32]. To reduce the impact of localization errors on modeling, the average reflectance within a 3×3 pixels window centered on the leaf sampling point was used as the final reflectance in this study.



Figure 2. The DJI 350M UAV with 300TC hyperspectral camera.

2.2.2. LCC Measurement

From 16 May to 22 May 2024, leaf samples were collected from 240 *Populus deltoides* trees. The samples were obtained from the south-facing side of each tree at a height of 15 to 17 m. The locations of samples were recorded using the real time kinematic GPS. The collected leaves were immediately transported to the laboratory and stored in a refrigerator at 4 °C to preserve freshness for subsequent analysis. Each day after leaf collection, LCC was determined using the spectrophotometric method. For each sample, twelve circular leaf discs (0.5 cm in diameter) were excised using a hole punch and placed in test tubes containing ethanol (Figure 3). After a 24 h incubation in darkness, the absorbance of the extraction solution was measured at 665 nm for chlorophyll-a and 649 nm for chlorophyll-b using a spectrophotometer. Chlorophyll concentrations were then calculated using Equations (1) and (2) based on the absorbance values [33]. Finally, LCC was converted from a per-unit weight basis (mg/g) to a per-unit area basis ($\mu\text{g}/\text{cm}^2$) using Equation (3).

$$\text{Chl}_a \left(\frac{\text{mg}}{\text{g}} \right) = \frac{[(13.95A_{665} - 6.88A_{649}) \times V_T(\text{ml})]}{[\text{FW}(\text{g}) \times 1000]} \quad (1)$$

$$\text{Chl}_b \left(\frac{\text{mg}}{\text{g}} \right) = \frac{[(24.96A_{649} - 7.32A_{665}) \times V_T(\text{ml})]}{[\text{FW}(\text{g}) \times 1000]} \quad (2)$$

$$\text{Chl}_{a+b} \left(\frac{\mu\text{g}}{\text{cm}^2} \right) = \frac{[(\text{Chl}_a(\text{mg}/\text{g}) + \text{Chl}_b(\text{mg}/\text{g})) \times \text{FW}(\text{g}) \times 1000]}{\text{LA}(\text{cm}^2)} \quad (3)$$

where Chl_a and Chl_b represent the concentrations of chlorophyll *a* and chlorophyll *b*, respectively. A_{665} and A_{649} correspond to the absorbance of the extracted pigments at wavelengths of 665 nm and 649 nm, respectively. Chl_{a+b} denotes the combined concentration of Chl_a and Chl_b , commonly referred to as the LCC. V_T , FW , and LA refer to the volume of the extraction solution, fresh weight of the sample, and leaf area, respectively.

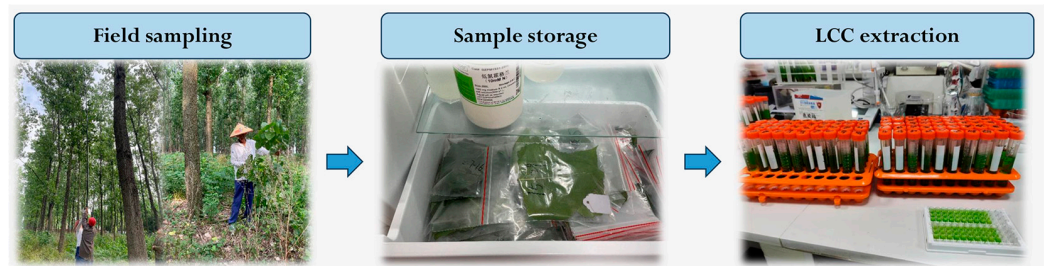


Figure 3. *Populus deltoides* leaves collection, storage and LCC extraction.

2.3. Ensembled Feature Selection Framework

2.3.1. Embedded Algorithm Improvements

Before developing the ensembled feature selection framework, the embedded algorithm was improved. To allow integration with various regression algorithms rather than relying on a predefined regression algorithm, the SHapley Additive exPlanations (SHAP) algorithm was used for feature ranking. The core principle of SHAP is based on the Shapley value from cooperative game theory, which ensures the fair distribution of a total gain among the features based on their contributions [34]. SHAP values are calculated by considering all possible combinations of features and measuring the marginal contribution of each feature to every possible coalition [35]. This ensures a consistent and fair allocation of contributions among features. The SHAP value ϕ_i for a feature i is computed using the formula:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \tag{4}$$

where N is the set of all features; S is a subset of N that does not include feature i ; $f(S)$ is the model prediction based on the subset of features S ; $|S|$ is the number of features in subset S ; $|N|$ is the total number of features.

During the feature selection process, removing certain features alters the composition of the feature set. This change subsequently affects the importance of the remaining features. As a result, the feature ranking obtained from the original feature set may not be reliable for subsequent rounds of feature selection. To address these dynamic changes, the LCC estimation model is reconstructed after each band-removal step in this study, where the 10 lowest-ranked bands are removed. The newly updated feature ranking is then used for the next round of dimensionality reduction. The flowchart of this improved embedded algorithm is shown in the ensembled framework (Figure 4).

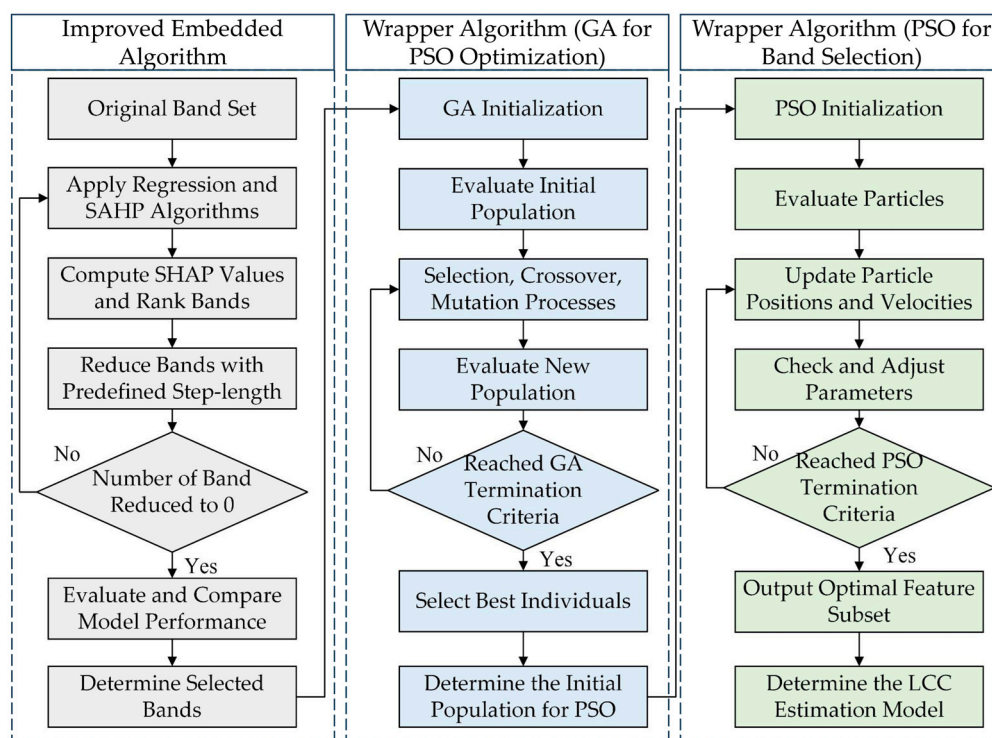


Figure 4. Framework of the ensembled feature selection method.

2.3.2. Wrapper Algorithm

Although the improved embedded feature selection methods are more robust and efficient, they still have limitations in identifying the optimal feature subset. Therefore, this study proposes an ensembled framework that uses the improved embedded method

to determine the initial band subset and then applies the wrapper algorithm for optimal subset search. Particle swarm optimization (PSO) is a typical wrapper feature selection algorithm used in related studies [36]. It is a global optimization algorithm that efficiently searches the feature space to identify the optimal subset by integrating the regression model. In PSO, each particle represents a potential solution (a subset of bands) encoded as a binary vector [37]. The algorithm begins by initializing a swarm of particles with random positions and velocities. Each particle's fitness is evaluated by training a regression model on the selected bands and calculating an evaluation index. The particles then update their positions and velocities based on their own best positions and the global best position found by the swarm. The PSO algorithm demonstrates strong local search capabilities and faster convergence rates. However, it is susceptible to becoming trapped in local optima. To address this limitation, a genetic algorithm (GA) is introduced to perform a global search and identify individuals with the highest fitness levels [38]. These selected individuals are then used as the initial population for the PSO algorithm. This hybrid approach enhances the efficiency of PSO by improving its convergence behavior and effectively avoiding entrapment in local optima [39]. Therefore, the framework of this ensembled feature selection algorithm is shown in Figure 4.

2.4. Regression Algorithm

2.4.1. GBRT

GBRT is a machine learning algorithm based on boosting principles. It improves model performance by combining the predictions of multiple weak learners into a single, robust model. The fundamental concept behind GBRT is to iteratively construct an ensemble of decision trees, where each subsequent tree is trained to correct the errors of the preceding ensemble [40]. This approach uses gradient descent to optimize a predefined loss function while incorporating regularization techniques to prevent overfitting, thereby achieving both accuracy and robustness. The performance of GBRT is highly dependent on several hyperparameters, including the number of trees, tree depth, and learning rate. In this study, these hyperparameters were optimized through grid search combined with cross-validation to enhance the estimation accuracy.

2.4.2. SVR

SVR is another ML algorithm based on the principles of Support Vector Machines. This algorithm aims to find a hyperplane in a high-dimensional space that best fits the data within a specified margin, denoted as ϵ [41]. This margin defines the tolerance for prediction errors. The fundamental concept of SVR is to minimize a regularized loss function, which balances model complexity with prediction accuracy. Specifically, SVR solves the optimization problem by minimizing the norm of the coefficient vector while constraining the prediction errors to fall within the ϵ margin. To handle both linear and nonlinear relationships, SVR employs kernel functions that map the input data into a higher-dimensional space. Common kernel functions include linear, polynomial, and radial basis function (RBF) kernels. In this study, the RBF kernel was chosen due to its effectiveness in capturing complex, nonlinear patterns. The hyperparameters, including the regularization parameter (C) and ϵ , were optimized using grid search combined with cross-validation.

2.4.3. GPR

Gaussian Process Regression (GPR) is a non-parametric Bayesian approach used for regression analysis. It provides a flexible and probabilistic framework for modeling complex relationships between target and features. In this algorithm, the underlying function is treated as a realization of a Gaussian process, which is a collection of random variables where any finite subset follows a joint Gaussian distribution [42]. The Gaussian process is characterized by its mean function and covariance function. The choice of covariance function, such as the Radial Basis Function (RBF) kernel, determines the smoothness and

flexibility of the model. During the regression process, GPR updates the prior Gaussian process using observed data to derive a posterior distribution. This posterior distribution provides predictions in terms of both the mean and the associated uncertainty (variance) at new data points. In this study, the RBF kernel is utilized as the covariance function. Additionally, other hyperparameters, including the length scale and signal variance, are optimized using maximum likelihood estimation to improve model performance.

2.5. Comparative Experiments

To demonstrate the effectiveness of the proposed feature selection framework, three comparative experiments were designed. The first two experiments were based on a single feature selection algorithm: filter and embedded. They were used to demonstrate the superiority of the ensembled feature selection algorithm. The third experiment was based on an ensembled framework and aimed to demonstrate the effectiveness of the improved embedded algorithm. As in the proposed method, the GBRT, SVR, and GPR algorithms were used in each comparative experiment.

The methods for each comparative experiment are as follows:

- (1) The first experiment was conducted using a simpler filter method based on the absolute value of the Pearson correlation coefficient (APCC). The bands with APCC values greater than 0.6, which are considered linearly correlated, were used to construct the LCC estimation model.
- (2) The second experiment was conducted using the ensemble algorithm. A fixed GBRT ranking and the stepwise band reduction process were used to identify the bands for LCC estimation.
- (3) The third experiment used the results of the second experiment and employed the same wrapper algorithm used in the framework to identify the optimal bands for LCC estimation.

To differentiate between the various models, the forms “X-NV model”, “C-X-NV model”, “X-Optimal model”, and “C-X-Optimal model” were introduced, where “X” denotes the regression algorithm used in the model, “N” indicates the number of bands used in the model, and “C” signifies that the model was generated from the comparative experiment. Models without “C” were produced using the proposed method. The “X-NV model” and “C-X-NV model” were derived from the initial band subset, while the “X-Optimal model” and “C-X-Optimal model” were based on the optimal band subset.

2.6. Model Performance Evaluation

In this study, 2/3 of samples were randomly selected as the training dataset, while the remaining samples were used for model performance evaluation. The coefficients of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE) were used as evaluation indicators. R^2 measures the ability of a model to explain the variability in the data. RMSE and MAE quantify the differences between estimated and measured LCC values. RMSE emphasizes the sensitivity to larger errors while MAE offers a balanced consideration of all errors. To evaluate the changes in model performance following dimensionality reduction, this paper employs Δ RMSE as a quantitative measure. The calculation formulas for all the relevant indicators are provided below.

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

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y'_i - y_i)^2} \quad (6)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \quad (7)$$

$$\Delta\text{RMSE} = \frac{\text{RMSE}_{\text{basic model}} - \text{RMSE}_{\text{model}}}{\text{RMSE}_{\text{basic model}}} \times 100\% \quad (8)$$

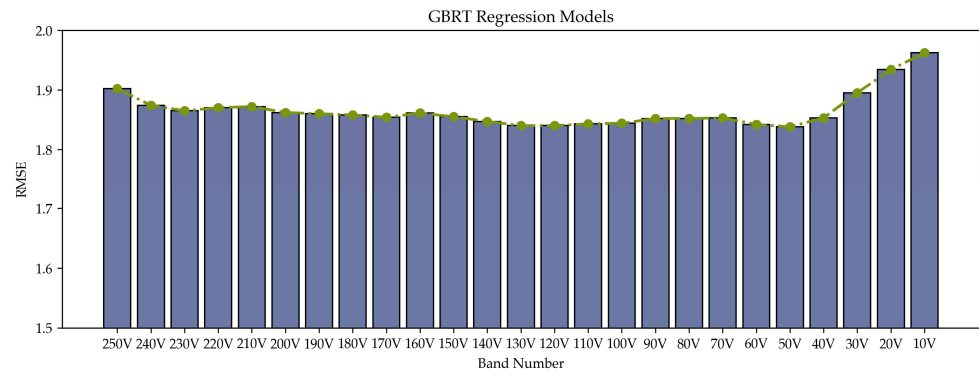
where y_i, \bar{y} and y'_i are measured LCC, mean value of measured LCC and the estimated LCC. The $\text{RMSE}_{\text{model}}$ refers to the RMSE value of the model after feature selection process, while the $\text{RMSE}_{\text{basic model}}$ refers to the RMSE value of the model that used all bands as inputs.

3. Results

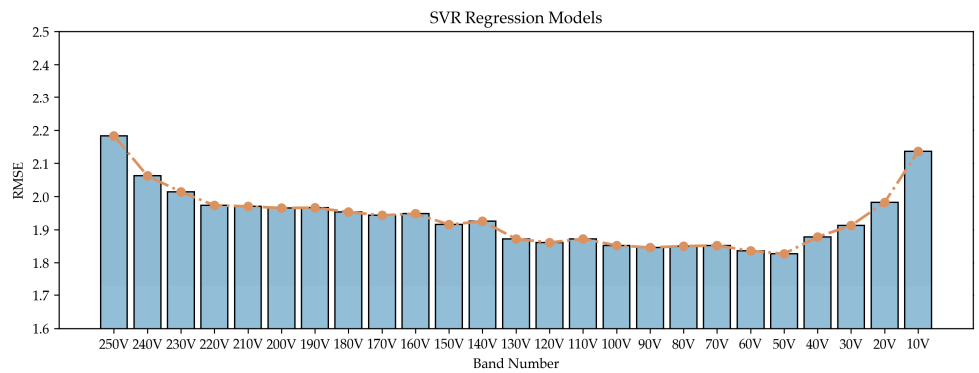
3.1. Initial Band Subset Developed Using the Improved Embedded Feature Selection Algorithm

Figure 5 illustrates the changes in RMSE of the LCC estimation model across different regression algorithms during the stepwise dimensionality reduction process guided by the improved embedded algorithm. As the number of bands decreased, all three regression models initially exhibited a reduction in RMSE, followed by an upward trend. The lowest RMSE for all models occurs when the band count is reduced to 50 V. While the overall trend is similar across the models, they differ in the magnitude of RMSE changes and the rate at which these changes occur. When using all bands as inputs, the GBRT model achieved the best model performance with an RMSE of $1.902 \mu\text{g}/\text{cm}^2$. While the RMSE values of the SVR and the GPR model are $2.183 \mu\text{g}/\text{cm}^2$ and $2.142 \mu\text{g}/\text{cm}^2$, respectively. When the number of bands is reduced from 250 V to 50 V, the RMSE of the GBRT model decreases by only $0.064 \mu\text{g}/\text{cm}^2$. While the RMSE values of the SVR and GPR models decrease by $0.356 \mu\text{g}/\text{cm}^2$ and $0.253 \mu\text{g}/\text{cm}^2$, respectively. Although both the SVR and GPR models achieved significant improvements in LCC estimation accuracy, the SVR model exhibited more stable RMSE variations during the stepwise dimensionality reduction process. In contrast, the RMSE of the GPR model increased several times as the number of bands was reduced from 250 V to 130 V. For example, the RMSE of the GPR model increased by $0.072 \mu\text{g}/\text{cm}^2$ when the band count decreased from 230 V to 220 V. Finally, as the number of bands was further reduced from 50 V to 10 V, a decline in accuracy was observed across all three models. Specifically, the RMSE increased by $0.124 \mu\text{g}/\text{cm}^2$ for the GBRT model, while the SVR and GPR models showed increases of $0.309 \mu\text{g}/\text{cm}^2$ and $0.567 \mu\text{g}/\text{cm}^2$, respectively.

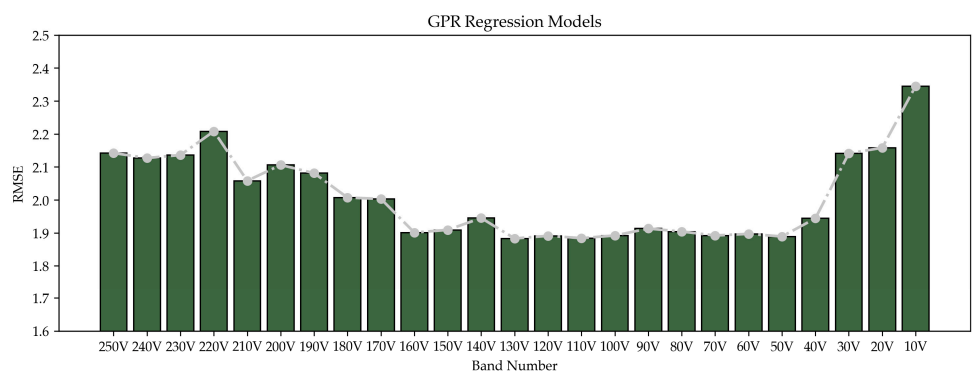
Figure 6 shows the LCC estimation results of the GBRT, SVR and GPR models using 50 bands selected by the improved embedded algorithm. The SVR-50V model achieved the best performance with an R^2 of 0.760, an RMSE of $1.827 \mu\text{g}/\text{cm}^2$ and an MAE of $1.531 \mu\text{g}/\text{cm}^2$. The GBRT-50V model followed closely, with an R^2 of 0.757, an RMSE of $1.838 \mu\text{g}/\text{cm}^2$ and an MAE of $1.544 \mu\text{g}/\text{cm}^2$. In contrast, the GPR-50V model achieved the lowest accuracy with an R^2 of 0.744, an RMSE of $1.889 \mu\text{g}/\text{cm}^2$ and an MAE of $1.569 \mu\text{g}/\text{cm}^2$. The GPR-50V model demonstrated a more pronounced underestimation of high LCC values and overestimation of low LCC values. To provide a more comprehensive assessment of model performance across different LCC ranges, residual boxplots were generated in this study (Figure 7). In the range of 20 to $25 \mu\text{g}/\text{cm}^2$, all samples are overestimated by using the GBRT-50V and GPR-50V models. However, the GBRT-50V model achieved a lower degree of overestimation compared to the GPR-50V model. In the range of 25– $35 \mu\text{g}/\text{cm}^2$, all three models exhibited relatively large residuals, which were the primary source of reduced estimation accuracy. Compared to the other two models, the GPR-50V model still showed significant overestimation within the 25 to $30 \mu\text{g}/\text{cm}^2$ range. In the range of 35 to $40 \mu\text{g}/\text{cm}^2$, both the GBRT-50V and GPR-50V models displayed notable underestimation. Although the performance of the GBRT-50V model is similar to that of the SVR-50V model, their residual distributions vary across LCC ranges. Specifically, the SVR-50V model demonstrates greater residual variability in the 20 to $25 \mu\text{g}/\text{cm}^2$ and 35 to $40 \mu\text{g}/\text{cm}^2$ ranges. This result indicates that some low-value and high-value samples were estimated more accurately by the SVR-50V model.



(a)



(b)



(c)

Figure 5. (a) RMSE value of the GBRT model during the stepwise dimensionality reduction process; (b) RMSE value of the SVR model during the stepwise dimensionality reduction process; (c) RMSE value of the GPR model during the stepwise dimensionality reduction process.

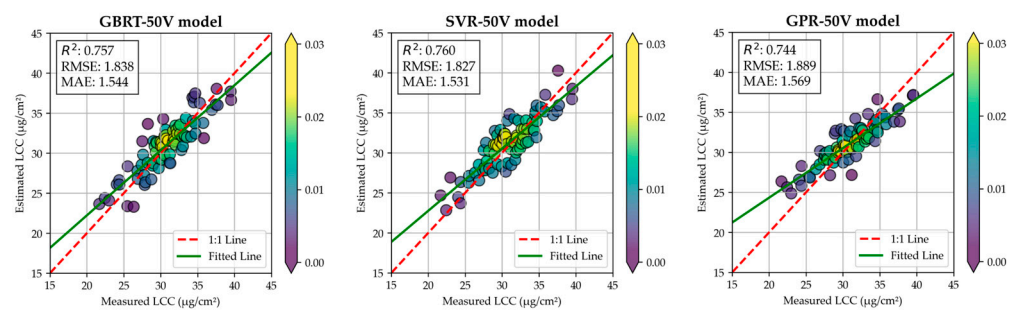


Figure 6. Performances of the GBRT-50V model, SVR-50V model and GRP-50V model.

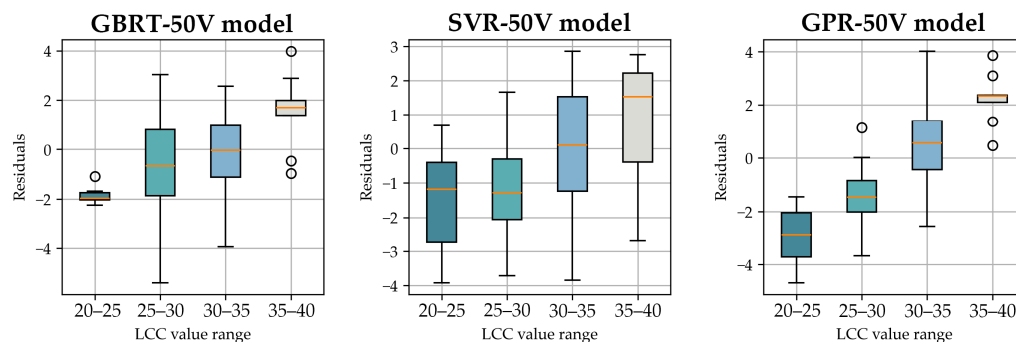


Figure 7. Residual distribution in different LCC range of the GBRT-50V model, SVR-50V model and GPR-50V model.

3.2. Optimal Models Developed Using the Ensembled Feature Selection Framework

Figure 8 illustrates the LCC estimation performance of three optimal models. The GBRT-Optimal model achieved the highest estimation accuracy with R^2 of 0.848, RMSE of $1.454 \mu\text{g}/\text{cm}^2$ and MAE of $1.121 \mu\text{g}/\text{cm}^2$. The SVR-Optimal model demonstrated lower accuracy with R^2 of 0.816, RMSE of $1.600 \mu\text{g}/\text{cm}^2$ and MAE of $1.356 \mu\text{g}/\text{cm}^2$. Compared to the results shown in Figure 6, the GBRT model showed significant accuracy improvements than the SVR model following the wrapper feature selection process using the PSO algorithm. The GPR-Optimal model also showed significant improvements with R^2 increased by 0.057, RMSE reduced by $0.227 \mu\text{g}/\text{cm}^2$ and MAE reduced by $0.231 \mu\text{g}/\text{cm}^2$. However, despite this enhancement, the GPR-Optimal model is still less accurate than both the GBRT-Optimal and SVR-Optimal models. This is because the GPR-Optimal model exhibits larger estimation errors for certain specific samples, resulting in a decrease in R^2 and an increase in RMSE. Nevertheless, since MAE is robust to outliers, the MAE of the GPR-Optimal model is lower than that of the SVR-Optimal model. In comparison to the models of initial band subsets (GBRT-50V, SVR-50V, GPR-50V), the results of optimal models exhibited closer alignment with the 1:1 line (Figure 9). In particular, the GBRT-Optimal model demonstrated improved estimation accuracy across both low and high LCC value ranges. Compared to using the initial band subsets as inputs (Figure 6), the PSO feature selection process improved the LCC estimation accuracy of all models (Table 1). The RMSE reductions for the GBRT, SVR, and GPR models were 20.89%, 12.42%, and 12.02%, respectively. Overall, the ensemble feature selection framework resulted in RMSE reductions of 24.37%, 31.91%, and 25.41% for the GBRT, SVR, and GPR models, respectively.

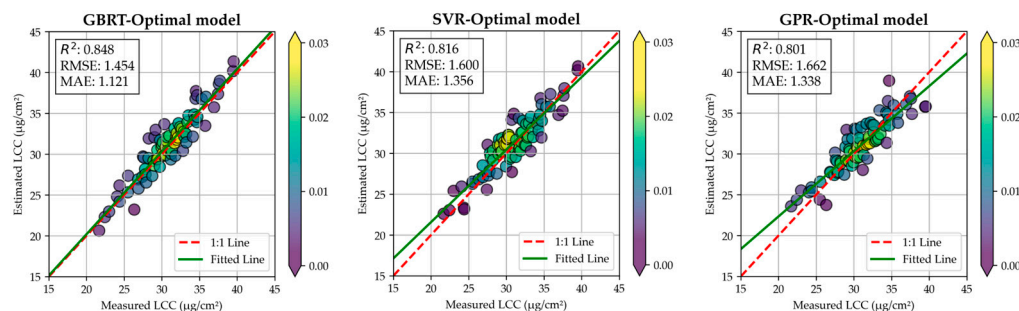


Figure 8. Performances of the GBRT-Optimal model, SVR-Optimal model and GPR-Optimal model.

Table 1. Model performance improvements of three LCC estimation models.

Regression Algorithm	Δ RMSE Achieved Using the Initial Band Subset	Δ RMSE Achieved Using the Optimal Band Subset
GBRT	20.89%	24.37%
SVR	12.42%	31.91%
GPR	12.02%	25.41%

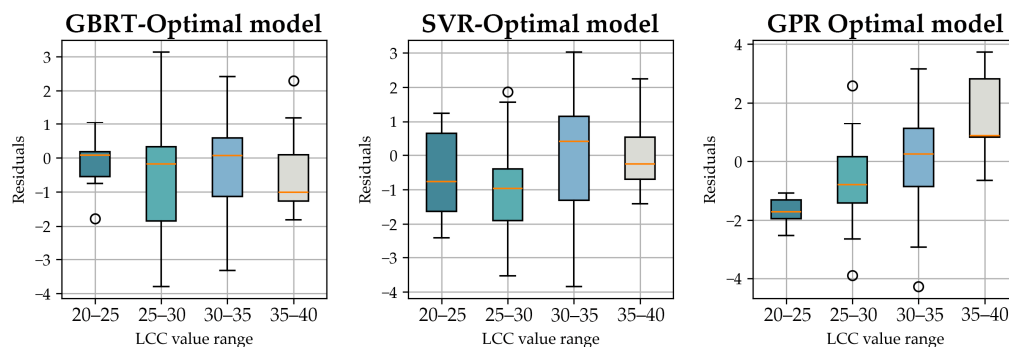


Figure 9. Residual distribution in different LCC range of the GBRT-Optimal model, SVR-Optimal model, and GPR-Optimal model.

Figure 9 shows the boxplot of residuals for each optimal model across different LCC value ranges. Compared to the results shown in Figure 7, the distribution of residuals for each model has changed. For instance, the median of residuals for samples within the LCC range of 20 to 25 $\mu\text{g}/\text{cm}^2$ is close to zero in the GBRT-Optimal model. This indicated that the overestimation has been mitigated. Moreover, within the LCC range of 35 to 40 $\mu\text{g}/\text{cm}^2$, the median of residuals is negative for the GBRT-Optimal model, which indicates that the underestimation phenomenon has also been mitigated. For the SVR-Optimal model, the residuals have been significantly reduced in the range of 35 to 40 $\mu\text{g}/\text{cm}^2$. In the GPR-Optimal model, the overestimation has been mitigated in the range of 20 to 25 $\mu\text{g}/\text{cm}^2$.

Figure 10 and Table 2 present the bands included in the three initial band subsets and the three optimal band subsets. Although each initial subset comprises 50 bands, the distribution of these selected bands varies significantly. As shown in Figure 10a, the bands used in the GBRT model are primarily concentrated in the B bands (420 nm to 460 nm) and the RE bands (680 nm to 760 nm). In contrast, the bands used in the SVR and GPR models are more widely distributed across the 400 nm to 900 nm spectrum, with a greater emphasis on G and R bands. Compared to the initial subsets, some bands with relatively high correlation were removed in the optimal subsets. For instance, the initial subset of the GBRT model includes several continuous bands between 660 nm and 760 nm, whereas the optimal subset shows a reduced number of continuous bands. The number of bands used in the GBRT-Optimal, SVR-Optimal, and GPR-Optimal models is 28, 36, and 29, respectively. This reduction demonstrates that the ensemble feature selection framework effectively decreased the number of bands to 12%–15% of the original set, resulting in improved LCC estimation accuracy, as evidenced by a 20%–30% reduction in RMSE (Table 1).

Table 2. Bands used in three optimal LCC estimation models.

Model	Wavelength of Bands Used in the Model
GBRT-Optimal	393 nm, 395 nm, 397 nm, 431 nm, 437 nm, 439 nm, 441 nm, 449 nm, 461 nm, 519 nm, 555 nm, 669 nm, 679 nm, 681 nm, 683 nm, 691 nm, 695 nm, 697 nm, 701 nm, 703 nm, 715 nm, 723 nm, 731 nm, 741 nm, 747 nm, 751 nm, 887 nm, 891 nm
SVR-Optimal	393 nm, 395 nm, 397 nm, 417 nm, 431 nm, 439 nm, 441 nm, 457 nm, 459 nm, 487 nm, 523 nm, 525 nm, 537 nm, 539 nm, 555 nm, 557 nm, 579 nm, 611 nm, 641 nm, 679 nm, 691 nm, 695 nm, 709 nm, 721 nm, 729 nm, 739 nm, 741 nm, 751 nm, 761 nm, 763 nm, 801 nm, 855 nm, 861 nm, 873 nm, 879 nm, 885 nm
GPR-Optimal	395 nm, 405 nm, 425 nm, 431 nm, 471 nm, 473 nm, 479 nm, 509 nm, 513 nm, 519 nm, 521 nm, 575 nm, 603 nm, 633 nm, 675 nm, 689 nm, 697 nm, 699 nm, 707 nm, 711 nm, 725 nm, 737 nm, 805 nm, 843 nm, 859 nm, 875 nm, 879 nm, 883 nm, 889 nm

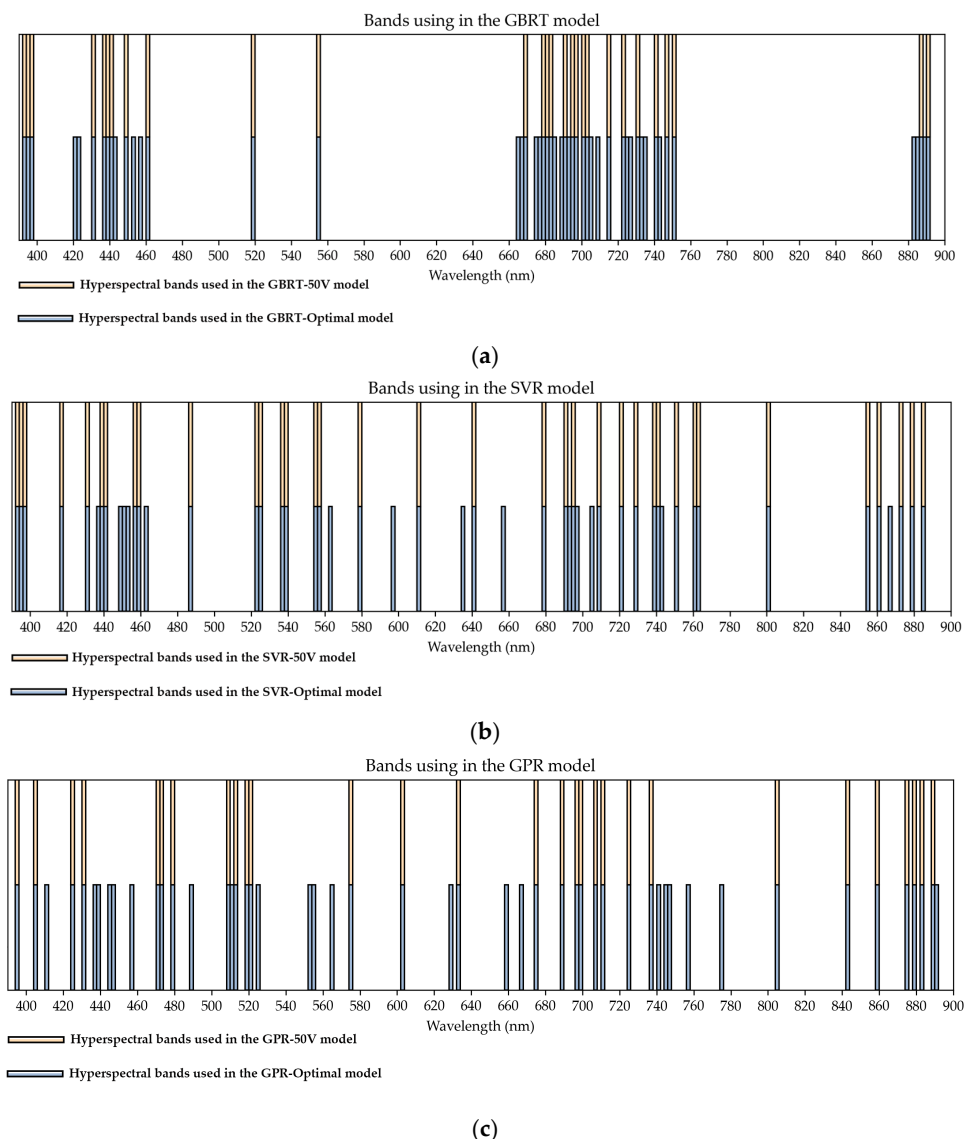


Figure 10. (a) Hyperspectral bands used in the GBRT-50V and GBRT-Optimal models; (b) Hyperspectral bands used in the SVR-50V and SVR-Optimal models; (c) Hyperspectral bands used in the GPR-50V and GPR-Optimal models.

3.3. Model Performances of Comparative Experiments

Figure 11 shows the selected 49 bands that achieved APCC values higher than 0.6. The bands are primarily concentrated within the ranges of 420 to 460 nm and 620 to 680 nm. These bands exhibit significant correlation with each other. As shown in Figure 12, the 49 bands were used as an input for developing the LCC estimation model. The C-SVR-49V model achieved the best performance, with R^2 of 0.664, RMSE of $2.164 \mu\text{g}/\text{cm}^2$ and MAE of $1.697 \mu\text{g}/\text{cm}^2$. In contrast, the C-GPR-49V model achieved the lowest accuracy, with R^2 of 0.602, RMSE of $2.354 \mu\text{g}/\text{cm}^2$ and MAE of $1.891 \mu\text{g}/\text{cm}^2$. Compared with the results of the proposed ensemble framework, the RMSE values of the C-GBRT-49V, C-SVR-49V, and C-GPR-49V models increased by $0.741 \mu\text{g}/\text{cm}^2$, $0.564 \mu\text{g}/\text{cm}^2$, and $0.692 \mu\text{g}/\text{cm}^2$, respectively.

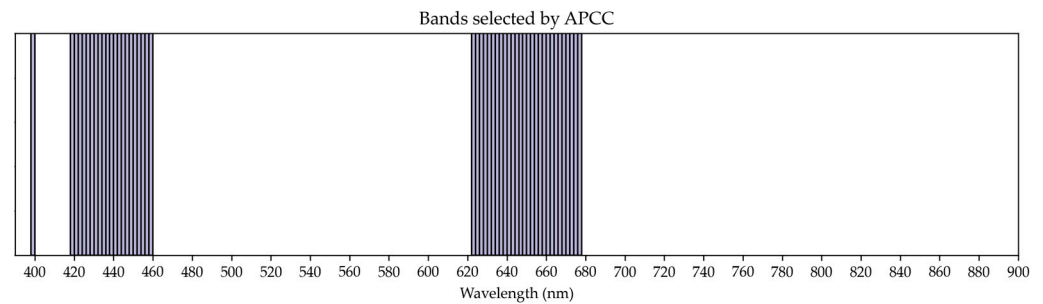


Figure 11. Bands selected according to APCC value.

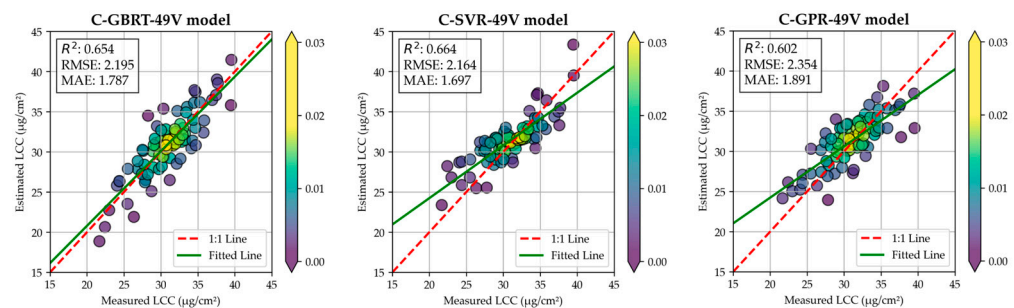


Figure 12. Performances of the C-GBRT-49V model, C-SVR-49V model, and C-GPR-49V model.

As shown in Figure 13, this study conducted the stepwise dimensionality reduction process based on the fixed GBRT ranking and different regression algorithm (second comparative performance). When using the GBRT regression algorithm, the RMSE increased as the number of bands decreased from 250 V to 170 V. As the number of bands further decreased from 170 V to 50 V, the RMSE stabilized within the range of $1.930 \mu\text{g}/\text{cm}^2$ to $1.960 \mu\text{g}/\text{cm}^2$. A significant increase in RMSE was observed when the number of bands dropped below 40 V. The lowest RMSE of $1.900 \mu\text{g}/\text{cm}^2$ was achieved at 190 V. For the SVR models, the RMSE increased as the number of bands decreased from 250 V to 70 V, followed by irregular fluctuations. In contrast, when using the GPR algorithm, the RMSE exhibited inconsistent fluctuations as the number of bands decreased from 250 V to 180 V. The best SVR and GPR models were achieved at 240 V ($\text{RMSE} = 2.163 \mu\text{g}/\text{cm}^2$) and 70 V ($\text{RMSE} = 1.981 \mu\text{g}/\text{cm}^2$), respectively. Compared to the results obtained using the SHAP ranking (Figure 5), the single feature selection process based on the fixed GBRT ranking not only resulted in larger errors in LCC estimation but also failed to maintain high estimation accuracy at lower dimensionalities. When developing the SVR and GPR models with the fixed GBRT ranking, the RMSE ranged from $1.900 \mu\text{g}/\text{cm}^2$ to $2.700 \mu\text{g}/\text{cm}^2$, which is significantly higher than the results achieved with the improved embedded algorithm, where the RMSE ranged from $1.800 \mu\text{g}/\text{cm}^2$ to $2.200 \mu\text{g}/\text{cm}^2$. These findings suggest that GBRT ranking is not well-suited for developing models when combined with other regression algorithms.

The third comparative experiment is based on the band subsets that achieved the highest accuracy for each model, as shown in Figure 13. To reduce computational complexity, the initial band subset dimensions for GBRT, SVR, and GPR were set at 50 V, 60 V, and 70 V, respectively. Compared to using all bands as inputs, employing the initial band subset increased the RMSE of the GBRT and SVR models by 1.47% and 10.31%, respectively (Table 3). Only the GPR model demonstrated an improvement in accuracy, with its RMSE decreasing by 7.52%. This study utilized the same wrapper algorithm to determine the optimal band subset for each model. Figure 14 illustrates the performance of the three optimal LCC estimation models. The C-GPR-Optimal model achieved the highest accuracy with R^2 of 0.767, RMSE of $1.800 \mu\text{g}/\text{cm}^2$ and MAE of $1.411 \mu\text{g}/\text{cm}^2$. In contrast, the C-SVR-Optimal model had the lowest accuracy with R^2 of 0.740, RMSE of $1.902 \mu\text{g}/\text{cm}^2$ and MAE of $1.577 \mu\text{g}/\text{cm}^2$. Although the accuracy of the C-GBRT-Optimal model was

lower than that of the C-GPR-Optimal model, its fitted line was closer to the 1:1 line. Table 3 presents the changes in RMSE across different models. Compared to using 250 V as input, the RMSEs of the C-GBRT-Optimal, C-SVR-Optimal, and C-GPR-Optimal models decreased by 1.31%, 12.87%, and 15.97%, respectively. Furthermore, the optimal band sets for GBRT, SVR, and GPR were determined at 39 V, 41 V, and 58 V, respectively. These results suggest that the fixed GBRT ranking is inadequate for providing a reasonable ranking for the dynamic model, resulting in only a modest improvement in the accuracy of the LCC estimation model.

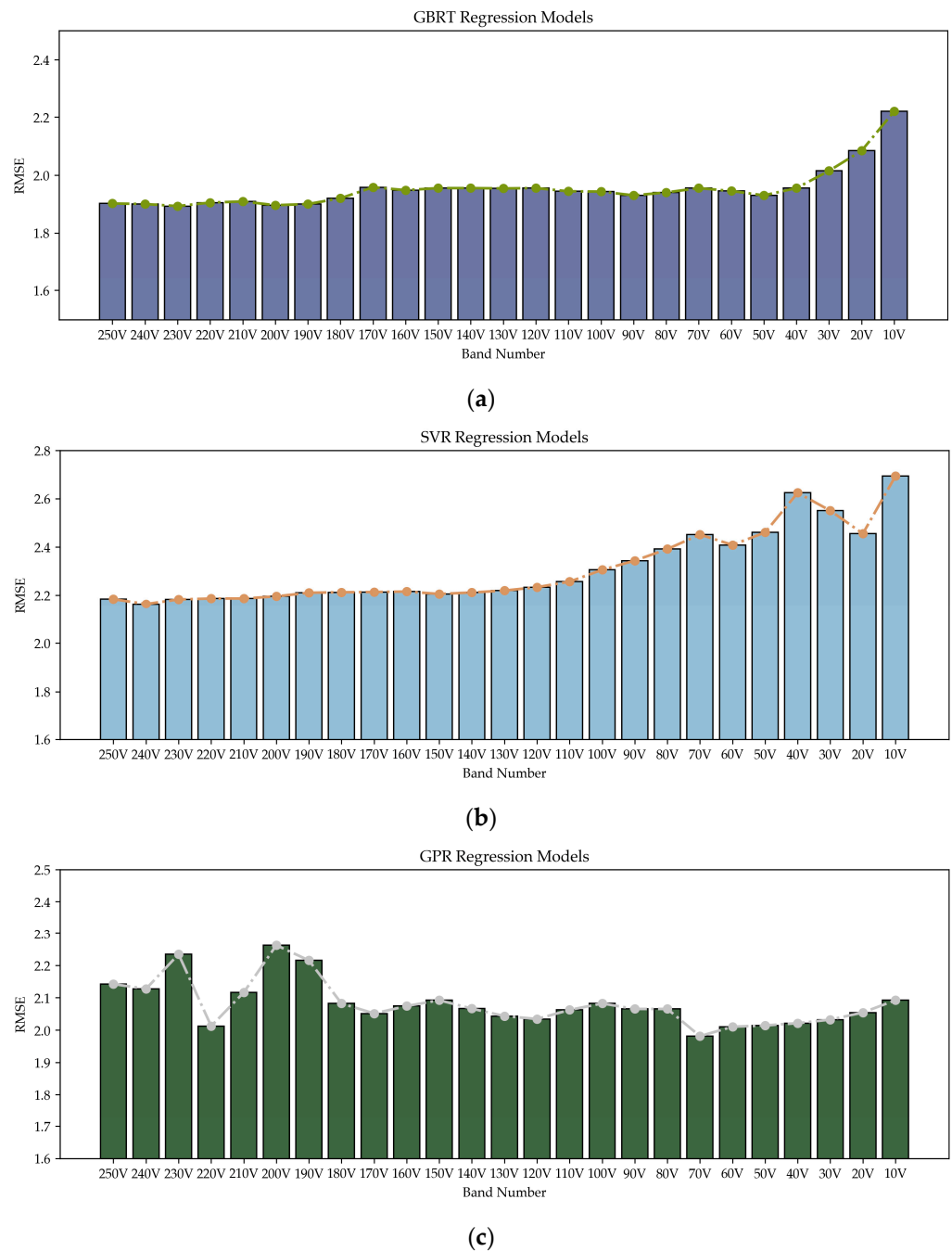
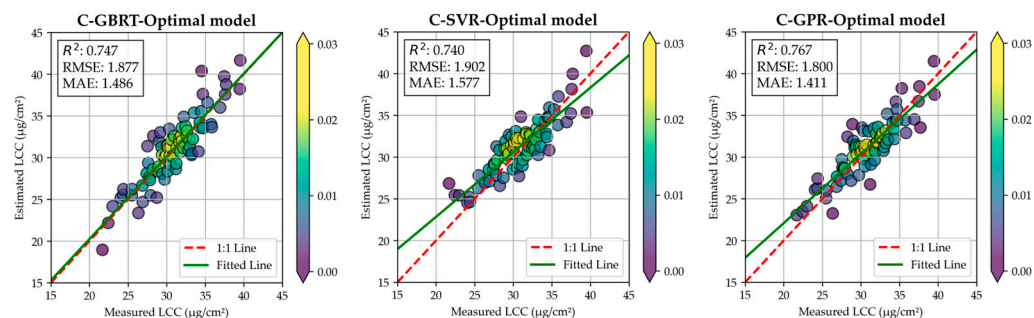


Figure 13. (a) RMSE value of GBRT model during the stepwise dimensionality reduction process based on GBRT ranking; (b) RMSE value of SVR model during the stepwise dimensionality reduction process based on GBRT ranking; (c) RMSE value of GPR model during the stepwise dimensionality reduction process based on GBRT ranking.

Table 3. Model performance improvements of three LCC estimation model using GBRT ranking.

Regression Algorithm	Δ RMSE Achieved Using the Initial Band Subset	Δ RMSE Achieved Using the Optimal Band Subset
GBRT	−1.47%	1.31%
SVR	−10.31%	12.87%
GPR	7.52%	15.97%

**Figure 14.** Performances of the C-GBRT-Optimal model, C-SVR-Optimal model and C-GPR-Optimal model.

4. Discussion

UAV-based hyperspectral imagery provides fine-scale spectral and spatial resolution for estimating LCC in *Populus deltoides* at the individual tree level. However, the strong correlation between hyperspectral bands significantly affects the accuracy and computational efficiency of LCC estimation models. Therefore, identifying the sensitive bands for LCC estimation in *Populus deltoides* is crucial. This study proposed an ensemble feature selection framework that identified 28 sensitive bands. Using these bands, the GBRT estimation model achieved accuracy with R^2 of 0.848 and, RMSE of $1.454 \mu\text{g}/\text{cm}^2$ and MAE of $1.121 \mu\text{g}/\text{cm}^2$. Compared to traditional embedded algorithms, this framework significantly improved model accuracy while reducing model complexity. In addition to biophysical and biochemical parameter estimation, this method is applicable to various hyperspectral imaging tasks, including classification and object detection.

The ensemble feature selection framework proposed in this study introduces three critical improvements. First, the SHAP algorithm was integrated into the regression process. Previous studies have indicated that the feature rankings provided by certain embedded algorithms reflect only the contribution of each feature within the specific model used [25,43]. Our findings are consistent with these results. For example, the initial band subsets for the GBRT, SVR, and GPR models were different (Figure 10). Additionally, applying the fixed GBRT ranking to the SVR and GPR models resulted in a significant loss of accuracy (Figure 13). This discrepancy may be due to the specificity of the algorithms and their differing objective functions. The GBRT algorithm constructs models using tree-based methods that capture complex interactions and nonlinear relationships, while the SVR algorithm aims to find a hyperplane that minimizes error [44]. The GPR algorithm uses a probabilistic approach with kernel functions [45]. Therefore, the GBRT ranking may not correspond to those prioritized by SVR or GPR. This mismatch affects the transferability of feature rankings between models. In this study, the SHAP algorithm offers a unified framework for interpreting model outputs by attributing feature contributions through cooperative game theory. By allowing integration with various regression algorithms, this method can generate more robust and reliable feature rankings. Second, dynamic ranking was introduced into the improved embedded feature selection algorithm. During feature selection, the number of features influences interactions, modifies redundancy, and alters the overall model structure. Consequently, the contribution of each feature shifts as the subset changes. In machine learning models, feature changes can affect the importance of other features, especially when they are correlated or provide overlapping

information [46]. Additionally, the internal parameters of the models are impacted by the number of features, leading to shifts in feature importance. This dynamic variation is particularly evident in models that capture complex, nonlinear relationships, such as tree-based models and neural networks [47]. In our study, using the fixed GBRT ranking during the stepwise dimensionality reduction process led to a continuous increase in RMSE (Figure 13a), whereas dynamic ranking significantly improved the accuracy of the LCC estimation model (Figure 5). Lastly, the integration of embedded and wrapper algorithms provides complementary benefits that enhanced the computational efficiency of optimal feature subset search. By initially using an embedded algorithm to identify a suitable feature subset, the number of features is reduced, which leads to decreased computational complexity. This strategy also reduces the risk of overfitting by focusing on representative features, improving model accuracy. Furthermore, it enables flexible algorithm selection tailored to specific datasets. Overall, this combined approach leverages the strengths of both methods, resulting in a more efficient and robust feature selection process.

In this study, the GBRT estimation model achieved the best performance using 28 narrow bands. As shown in Table 2, the optimal band subset includes 12 RE bands ranging from 680 nm to 760 nm. This finding is consistent with previous studies, which have demonstrated that RE band reflectance is highly correlated with LCC. For example, Sun et al. [48] and Chen et al. [49] analyzed the sensitivity of Sentinel-2 bands to LCC using a radiative transfer model. The results showed that RE2 band with wavelength of 704 nm and bandwidth of 15 nm exhibited the highest sensitivity. Similarly, Yue et al. [50] employed UAV-based hyperspectral imagery for soybean LCC estimation and found that reflectance between 700 nm and 800 nm was highly correlated with LCC. Furthermore, previous studies have suggested that using RE bands can reduce the underestimation of high values in LCC estimation models [51,52]. The findings of this study align with this conclusion. For instance, both the initial and optimal band subsets of the GBRT model contain a larger number of RE bands. As a result, the GBRT-50V and GBRT-Optimal models exhibit lower levels of underestimation and overestimation compared to the SVR and GPR models. This is because RE bands are less prone to saturation in areas with high vegetation coverage or chlorophyll content [53,54]. In plants with high chlorophyll content, red light absorption often saturates, while RE bands lie in the transition zone between the R and NIR regions. In this region, reflectance is highly sensitive to variations in leaf area index and LCC due to multiple scattering between leaf layers and chlorophyll absorption [55,56]. Consequently, RE bands are particularly sensitive to variations in LCC.

5. Conclusions

This study developed an ensemble feature selection framework to identify the optimal UAV-based hyperspectral band subset for estimating LCC in *Populus deltoides*. By using 28 narrow bands and a GBRT model, the proposed framework achieved an accurate LCC estimation with R^2 of 0.848, RMSE of $1.454 \mu\text{g}/\text{cm}^2$ and MAE of $1.121 \mu\text{g}/\text{cm}^2$. Compared with using all bands as inputs, this method reduced the RMSE by 24.37%. The main conclusions are as follows:

- (1) The improved embedded feature selection algorithm produced more robust and reliable band rankings. By introducing SHAP values and a dynamic ranking strategy, redundant and low-correlation bands were identified with greater precision across different regression algorithms.
- (2) The ensemble feature selection framework improved the efficiency of optimal subset selection in hyperspectral data. Although the improved embedded algorithm had a limited impact on model accuracy, it provided a strong band subset for the wrapper algorithm. The combination of both feature selection approaches reduced the risk of converging to a local optimum.
- (3) RE bands are critical for developing accurate LCC estimation model for *Populus deltoides*. By incorporating several RE bands in the 680–760 nm range, the underestimation problem at high LCC levels was mitigated.

Author Contributions: Conceptualization, Z.C.; methodology, Z.C. and X.W.; software, S.Q.; validation, Z.C. and H.L.; formal analysis, Z.C. and X.W.; investigation, Z.C., H.J., X.C. and M.S.; resources, Z.C.; writing—original draft preparation, Z.C.; writing—review and editing, Z.C., H.J. and H.Z.; funding acquisition, Z.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Scientific and Technological Innovation 2030—Major Projects, grant number 2023ZD0405605.

Data Availability Statement: Data will be made available on request.

Acknowledgments: The authors would like to express their sincere gratitude to the reviewers for their valuable feedback and constructive comments. Special thanks to Bo Yuan from the Chinese Research Academy of Environmental Sciences for revising this work and providing insightful suggestions.

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

References

- Chen, C.; Chu, Y.; Ding, C.; Su, X.; Huang, Q. Genetic diversity and population structure of black cottonwood (*Populus deltoides*) revealed using simple sequence repeat markers. *BMC Genet.* **2020**, *21*, 2. [[CrossRef](#)] [[PubMed](#)]
- Maier, C.A.; Burley, J.; Cook, R.; Ghezehei, S.B.; Hazel, D.W.; Nichols, E.G. Tree water use, water use efficiency, and carbon isotope discrimination in relation to growth potential in *Populus deltoides* and hybrids under field conditions. *Forests* **2019**, *10*, 993. [[CrossRef](#)]
- Fan, J.; Zhang, Y.; Wen, W.; Gu, S.; Lu, X.; Guo, X. The future of Internet of Things in agriculture: Plant high-throughput phenotypic platform. *J. Clean. Prod.* **2021**, *280*, 123651. [[CrossRef](#)]
- Shi, C.; Yang, F.; Liu, Z.; Li, Y.; Di, X.; Wang, J.; Lin, J. Uniform water potential induced by salt, alkali, and drought stresses has different impacts on the seedling of *Hordeum jubatum*: From growth, photosynthesis, and chlorophyll fluorescence. *Front. Plant Sci.* **2021**, *12*, 733236. [[CrossRef](#)]
- Patane, P.; Vibhute, A. Chlorophyll and nitrogen estimation techniques: A Review. *Int. J. Eng. Res.* **2014**, *2*, 33–41.
- Daughtry, C.S.; Walthall, C.L.; Kim, M.S.; De Colstoun, E.B.; McMurtrey Iii, J.E. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
- 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](#)]
- Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [[CrossRef](#)] [[PubMed](#)]
- Zhou, X.; Zhang, J.; Chen, D.; Huang, Y.; Kong, W.; Yuan, L.; Ye, H.; Huang, W. Assessment of leaf chlorophyll content models for winter wheat using Landsat-8 multispectral remote sensing data. *Remote Sens.* **2020**, *12*, 2574. [[CrossRef](#)]
- Li, Y.; Ma, Q.; Chen, J.M.; Croft, H.; Luo, X.; Zheng, T.; Zheng, T.; Rogers, C.; Liu, J. Fine-scale leaf chlorophyll distribution across a deciduous forest through two-step model inversion from Sentinel-2 data. *Remote Sens. Environ.* **2021**, *264*, 112618. [[CrossRef](#)]
- Zhang, H.; Ge, Y.; Xie, X.; Atefi, A.; Wijewardane, N.K.; Thapa, S. High throughput analysis of leaf chlorophyll content in sorghum using RGB, hyperspectral, and fluorescence imaging and sensor fusion. *Plant Methods* **2022**, *18*, 60. [[CrossRef](#)] [[PubMed](#)]
- Tao, H.; Feng, H.; Xu, L.; Miao, M.; Long, H.; Yue, J.; Li, Z.; Yang, G.; Yang, X.; Fan, L. Estimation of crop growth parameters using UAV-based hyperspectral remote sensing data. *Sensors* **2020**, *20*, 1296. [[CrossRef](#)] [[PubMed](#)]
- Zhu, W.; Sun, Z.; Yang, T.; Li, J.; Peng, J.; Zhu, K.; Li, S.; Gong, H.; Lyu, Y.; Li, B.; et al. Estimating leaf chlorophyll content of crops via optimal unmanned aerial vehicle hyperspectral data at multi-scales. *Comput. Electron. Agric.* **2020**, *178*, 105786. [[CrossRef](#)]
- Zhang, Y.; Hui, J.; Qin, Q.; Sun, Y.; Zhang, T.; Sun, H.; Li, M. Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data. *Remote Sens. Environ.* **2021**, *267*, 112724. [[CrossRef](#)]
- Xu, L.; Shi, S.; Gong, W.; Shi, Z.; Qu, F.; Tang, X.; Chen, B.; Sun, J. Improving leaf chlorophyll content estimation through constrained PROSAIL model from airborne hyperspectral and LiDAR data. *Int. J. Appl. Earth Obs.* **2022**, *115*, 103128. [[CrossRef](#)]
- Yan, P.; Feng, Y.; Han, Q.; Hu, Z.; Huang, X.; Su, K.; Kang, S. Enhanced cotton chlorophyll content estimation with UAV multispectral and LiDAR constrained SCOPE model. *Int. J. Appl. Earth Obs.* **2024**, *132*, 104052. [[CrossRef](#)]
- Shen, X.; Cao, L.; Coops, N.C.; Fan, H.; Wu, X.; Liu, H.; Wang, G.; Cao, F. Quantifying vertical profiles of biochemical traits for forest plantation species using advanced remote sensing approaches. *Remote Sens. Environ.* **2020**, *250*, 112041. [[CrossRef](#)]
- Hosoi, F.; Umeyama, S.; Kuo, K. Estimating 3D chlorophyll content distribution of trees using an image fusion method between 2D camera and 3D portable scanning lidar. *Remote Sens.* **2019**, *11*, 2134. [[CrossRef](#)]
- Buckley, S.J.; Kurz, T.H.; Howell, J.A.; Schneider, D. Terrestrial lidar and hyperspectral data fusion products for geological outcrop analysis. *Comput. Geosci.* **2013**, *54*, 249–258. [[CrossRef](#)]
- Incekara, A.H.; Seker, D.Z.; Bayram, B. Qualifying the LIDAR-derived intensity image as an infrared band in NDWI-based shoreline extraction. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 5053–5062. [[CrossRef](#)]
- Sonobe, R.; Sugimoto, Y.; Kondo, R.; Seki, H.; Sugiyama, E.; Kiriwa, Y.; Suzuki, K. Hyperspectral wavelength selection for estimating chlorophyll content of muskmelon leaves. *Eur. J. Remote Sens.* **2021**, *54*, 513–524. [[CrossRef](#)]

22. Agrawal, P.; Abutarboush, H.F.; Ganesh, T.; Mohamed, A.W. Metaheuristic algorithms on feature selection: A survey of one decade of research (2009–2019). *IEEE Access* **2021**, *9*, 26766–26791. [[CrossRef](#)]
23. Zhang, Y.; Chang, Q.; Chen, Y.; Liu, Y.; Jiang, D.; Zhang, Z. Hyperspectral estimation of chlorophyll content in apple tree leaf based on feature band selection and the CatBoost model. *Agronomy* **2023**, *13*, 2075. [[CrossRef](#)]
24. Zhao, R.; Tang, W.; Liu, M.; Wang, N.; Sun, H.; Li, M.; Ma, Y. Spatial-spectral feature extraction for in-field chlorophyll content estimation using hyperspectral imaging. *Biosyst. Eng.* **2024**, *246*, 263–276. [[CrossRef](#)]
25. Chen, Z.; Shi, T.; Zhang, X.; Jia, K.; Jiang, H.; Yuan, B. A hybrid leaf area index estimation method of dioscorea polystachya turczaninow using Sentinel-2 vegetation indices. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4415713. [[CrossRef](#)]
26. Chakhvashvili, E.; Siegmann, B.; Muller, O.; Verrelst, J.; Bendig, J.; Kraska, T.; Rascher, U. Retrieval of crop variables from proximal multispectral UAV image data using PROSAIL in maize canopy. *Remote Sens.* **2022**, *14*, 1247. [[CrossRef](#)]
27. Impollonia, G.; Croci, M.; Blandinières, H.; Marcone, A.; Amaducci, S. Comparison of PROSAIL model inversion methods for estimating leaf chlorophyll content and LAI using UAV imagery for hemp phenotyping. *Remote Sens.* **2022**, *14*, 5801. [[CrossRef](#)]
28. Wang, B.; Jia, K.; Liang, S.; Xie, X.; Wei, X.; Zhao, X.; Yao, Y.; Zhang, X. Assessment of Sentinel-2 MSI spectral band reflectances for estimating fractional vegetation cover. *Remote Sens.* **2018**, *10*, 1927. [[CrossRef](#)]
29. Ta, N.; Chang, Q.; Zhang, Y. Estimation of apple tree leaf chlorophyll content based on machine learning methods. *Remote Sens.* **2021**, *13*, 3902. [[CrossRef](#)]
30. Li, D.; Hu, Q.; Ruan, S.; Liu, J.; Zhang, J.; Hu, C.; Liu, Y.; Dian, Y.; Zhou, J. Utilizing Hyperspectral Reflectance and Machine Learning Algorithms for Non-Destructive Estimation of Chlorophyll Content in Citrus Leaves. *Remote Sens.* **2023**, *15*, 4934. [[CrossRef](#)]
31. Ge, X.; Wang, J.; Ding, J.; Cao, X.; Zhang, Z.; Liu, J.; Li, X. Combining UAV-based hyperspectral imagery and machine learning algorithms for soil moisture content monitoring. *Peer J.* **2019**, *7*, e6926. [[CrossRef](#)] [[PubMed](#)]
32. Sun, J.; Yang, W.; Zhang, M.; Feng, M.; Xiao, L.; Ding, G. Estimation of water content in corn leaves using hyperspectral data based on fractional order Savitzky-Golay derivation coupled with wavelength selection. *Comput. Electron. Agric.* **2021**, *182*, 105989. [[CrossRef](#)]
33. Lichtenthaler, H.K.; Wellburn, A.R. Determinations of total carotenoids and chlorophyll a and b of leaf extracts in different solvents. *Biochem. Soc. Trans.* **1983**, *603*, 591–592. [[CrossRef](#)]
34. Antwarg, L.; Miller, R.M.; Shapira, B.; Rokach, L. Explaining anomalies detected by autoencoders using Shapley Additive Explanations. *Expert Syst. Appl.* **2021**, *186*, 115736. [[CrossRef](#)]
35. Nohara, Y.; Matsumoto, K.; Soejima, H.; Nakashima, N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Comput. Methods Programs Biomed.* **2022**, *214*, 106584. [[CrossRef](#)] [[PubMed](#)]
36. Amoozegar, M.; Minaei-Bidgoli, B. Optimizing multi-objective PSO based feature selection method using a feature elitism mechanism. *Expert Syst. Appl.* **2018**, *113*, 499–514. [[CrossRef](#)]
37. Mistry, K.; Zhang, L.; Neoh, S.C.; Lim, C.P.; Fielding, B. A micro-GA embedded PSO feature selection approach to intelligent facial emotion recognition. *IEEE Trans. Cybern.* **2016**, *47*, 1496–1509. [[CrossRef](#)] [[PubMed](#)]
38. Nazir, M.; Majid-Mirza, A.; Ali-Khan, S. PSO-GA based optimized feature selection using facial and clothing information for gender classification. *J. Appl. Res. Technol.* **2014**, *12*, 145–152. [[CrossRef](#)]
39. Ghosh, M.; Guha, R.; Alam, I.; Lohariwal, P.; Jalan, D.; Sarkar, R. Binary genetic swarm optimization: A combination of GA and PSO for feature selection. *Int. J. Intell. Syst.* **2019**, *29*, 1598–1610. [[CrossRef](#)]
40. Li, L.; Dai, S.; Cao, Z.; Hong, J.; Jiang, S.; Yang, K. Using improved gradient-boosted decision tree algorithm based on Kalman filter (GBDT-KF) in time series prediction. *J. Supercomput.* **2020**, *76*, 6887–6900. [[CrossRef](#)]
41. Rodríguez-Pérez, R.; Bajorath, J. Evolution of support vector machine and regression modeling in chemoinformatics and drug discovery. *J. Comput.-Aided Mol. Des.* **2022**, *36*, 355–362. [[CrossRef](#)] [[PubMed](#)]
42. Yadav, A.; Bareth, R.; Kochar, M.; Pazoki, M.; Sehiemy, R.A.E. Gaussian process regression-based load forecasting model. *IET Gener. Transm. Distrib.* **2024**, *18*, 899–910. [[CrossRef](#)]
43. Liu, Y.; Liu, Z.; Luo, X.; Zhao, H. Diagnosis of Parkinson’s disease based on SHAP value feature selection. *Biocybern. Biomed. Eng.* **2022**, *42*, 856–869. [[CrossRef](#)]
44. Huang, H.; Jia, R.; Shi, X.; Liang, J.; Dang, J. Feature selection and hyper parameters optimization for short-term wind power forecast. *Appl. Intell.* **2021**, *51*, 6752–6770. [[CrossRef](#)]
45. Behkamal, B.; Entezami, A.; De Michele, C.; Arslan, A.N. Elimination of thermal effects from limited structural displacements based on remote sensing by machine learning techniques. *Remote Sens.* **2023**, *15*, 3095. [[CrossRef](#)]
46. Aggarwal, M.; Khullar, V.; Goyal, N.; Singh, A.; Tolba, A.; Thompson, E.B.; Kumar, S. Pre-trained deep neural network-based features selection supported machine learning for rice leaf disease classification. *Agriculture* **2023**, *13*, 936. [[CrossRef](#)]
47. Chandrashekar, G.; Sahin, F. A survey on feature selection methods. *Comput. Electr. Eng.* **2014**, *40*, 16–28. [[CrossRef](#)]
48. Sun, Y.; Qin, Q.; Ren, H.; Zhang, T.; Chen, S. Red-Edge band vegetation indices for leaf area index estimation from Sentinel-2/MSI imagery. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 826–840. [[CrossRef](#)]
49. Chen, Z.; Jia, K.; Wei, X.; Liu, Y.; Zhan, Y.; Xia, M.; Yao, Y.; Zhang, X. Improving leaf area index estimation accuracy of wheat by involving leaf chlorophyll content information. *Comput. Electron. Agric.* **2022**, *196*, 106902. [[CrossRef](#)]
50. Yue, J.; Yang, H.; Feng, H.; Han, S.; Zhou, C.; Fu, Y.; Gao, W.; Ma, X.; Qiao, H.; Yang, G. Hyperspectral-to-image transform and CNN transfer learning enhancing soybean LCC estimation. *Comput. Electron. Agric.* **2023**, *211*, 108011. [[CrossRef](#)]

51. Gao, S.; Yan, K.; Liu, J.; Pu, J.; Zou, D.; Qi, J.; Mu, X.; Yan, G. Assessment of remote-sensed vegetation indices for estimating forest chlorophyll concentration. *Ecol. Indic.* **2020**, *162*, 112001. [[CrossRef](#)]
52. Gu, C.; Li, J.; Liu, Q.; Zhang, H.; Liu, L.; Mumtaz, F.; Dong, Y.; Zhao, J.; Wang, X.; Liu, C. Retrieving decametric-resolution leaf chlorophyll content from GF-6 WFV by assessing the applicability of red-edge vegetation indices. *Comput. Electron. Agric.* **2023**, *215*, 108455. [[CrossRef](#)]
53. Sun, Y.; Wang, B.; Zhang, Z. Improving leaf area index estimation with chlorophyll insensitive multispectral red-edge vegetation indices. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 3568–3582. [[CrossRef](#)]
54. Li, W.; Weiss, M.; Jay, S.; Wei, S.; Zhao, N.; Comar, A.; Lopez-Lozano, R.; Solan, B.; Yu, Q.; Wu, Q.; et al. Daily monitoring of Effective Green Area Index and Vegetation Chlorophyll Content from continuous acquisitions of a multi-band spectrometer over winter wheat. *Remote Sens. Environ.* **2024**, *300*, 113883. [[CrossRef](#)]
55. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [[CrossRef](#)]
56. Heckmann, D.; Schlüter, U.; Weber, A.P. Machine learning techniques for predicting crop photosynthetic capacity from leaf reflectance spectra. *Mol. Plant.* **2017**, *10*, 878–890. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.