


## Article

# Comparing Stacking Ensemble Learning and 1D-CNN Models for Predicting Leaf Chlorophyll Content in *Stellera chamaejasme* from Hyperspectral Reflectance Measurements

Xiaoyu Li <sup>1</sup>, Yongmei Liu <sup>1,2,\*</sup>, Huaiyu Wang <sup>1</sup>, Xingzhi Dong <sup>1</sup>, Lei Wang <sup>1,2</sup> and Yongqing Long <sup>1,2</sup> <sup>1</sup> College of Urban and Environmental Sciences, Northwest University, Xi'an 710127, China<sup>2</sup> Shaanxi Key Laboratory of Earth Surface System and Environmental Carrying Capacity, Xi'an 710127, China

\* Correspondence: liuym@nwu.edu.cn

**Abstract:** *Stellera chamaejasme*, a toxic invasive species widespread in degraded alpine grasslands, Qinghai Province, causes a significant threat to the local ecological balance. Accurate monitoring of the leaf chlorophyll content is essential for preventing its expansion over large areas. This study presents an optimal approach by integrating hierarchical dimensionality reduction, stacking ensemble learning, and 1D-CNN models to estimate leaf chlorophyll content in *S. chamaejasme* using hyperspectral reflectance data. Field spectrometry analysis demonstrates that the combination of Pearson correlation, first derivative, and SPA algorithms can efficiently select the most chlorophyll-sensitive wavelengths, red-edge parameters, and spectral indices related to *S. chamaejasme* leaves. The stacking ensemble model outperforms the 1D-CNN model in predicting leaf chlorophyll content of *S. chamaejasme* over the whole growth stage, while the 1D-CNN excels at prediction in each individual growth stage. Comparatively, the 1D-CNN model achieved higher accuracy ( $R^2 > 0.5$ ) in all five growth stages, with optimal performance during the flower bud stage ( $R^2 = 0.787$ , RMSE = 2.476). This study underscores the potential of combining feature spectra selection with machine learning and deep learning models to monitor *S. chamaejasme* growth, offering valuable insights for invasive species control and ecological management.



Academic Editor: Maciej Zaborowicz

Received: 13 December 2024

Revised: 24 January 2025

Accepted: 26 January 2025

Published: 28 January 2025

**Keywords:** leaf chlorophyll; hyperspectral prediction; dimensionality reduction; stacking ensemble learning model; 1D convolutional neural network

**Citation:** Li, X.; Liu, Y.; Wang, H.; Dong, X.; Wang, L.; Long, Y. Comparing Stacking Ensemble Learning and 1D-CNN Models for Predicting Leaf Chlorophyll Content in *Stellera chamaejasme* from Hyperspectral Reflectance Measurements. *Agriculture* **2025**, *15*, 288. <https://doi.org/10.3390/agriculture15030288>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Leaf chlorophyll is crucial for photosynthesis, and converts light energy into chemical energy to support plant growth [1]. The leaf chlorophyll content is an indicator of plant health and is considered a vital biochemical parameter [2–4]. Traditional determination of leaf chlorophyll content mainly adopts the destructive sampling method, a process that is intricate and time-consuming [5,6]. Hyperspectral remote sensing acquires the collection of continuous narrowband spectral data related to the target, and provides swift, accurate, and noninvasive assessment of the leaf chlorophyll content in plants [7].

Current studies mainly utilize various techniques, such as correlation analysis, principal component analysis (PCA), competitive adaptive reweighted sampling (CARS), and the successive projection algorithm (SPA), to select feature wavelengths to calculate chlorophyll-sensitive vegetation indices [8–10]. Furthermore, various regression models, including linear regression, multiple linear regression (MLR), and partial least squares (PLS), are widely used in the estimation of leaf chlorophyll content [11,12]. Nonetheless, these methods exhibit sensitivity to data noise, lack of robustness, and low computational efficiency.

Owing to powerful modeling capability, high processing efficiency, and strong robustness, machine learning and deep learning technologies hold great potential when integrated with hyperspectral remote sensing in the analysis of the physicochemical parameters of plants. Many studies present the application of machine learning and deep learning algorithms in the hyperspectral inversion of chlorophyll content. An et al. [13] and Li et al. [14] employed MLR, support vector machine (SVM), and random forest (RF) models to determine the chlorophyll content in rice and potato leaves; and the results showed that the SVM and RF models markedly promoted estimation accuracy compared to the traditional MLR. Gan et al. [15] and Putra et al. [16] confirmed that the sparse autoencoder (SAE) model outperformed linear regression in estimating the chlorophyll content in longan leaves across different maturity stages. Sonobe et al. [17] found the advantages of the integration of deep belief networks (DBNs) and RF in the estimation of chlorophyll-a and chlorophyll-b contents in tea leaves. In recent years, some studies on the hyperspectral inversion of chlorophyll have focused on natural grasslands. Zhang et al. [18] explored the estimation of grassland chlorophyll content by the combination of fractional-order derivative (FOD), least squares regression, and support vector regression (SVR) models. Ji and Liu [19] applied the backward feature elimination (BFE) method in combination with PLS, RF, and tree-based regression (TBR) models to estimate chlorophyll content in alpine meadows on the Qinghai–Tibet Plateau. These models showed different suitability in terms of the accuracy and certainty of their predictions.

Qinghai Province is one of the five principal pastoral regions in China, characterized by abundant alpine grassland resources. The natural grasslands cover 41.867 million hectares, occupying 60.5% of the Qinghai area. In recent decades, alpine grasslands have suffered degradations due to climate change and human activities, accompanied by the notable increase in toxic weeds. *S. chamaejasme* possesses strong environmental adaptability and population competitiveness, and has become one of the major toxic species in moderately to severely degraded alpine grasslands in Qinghai Province [20]. The rapid expansion of *S. chamaejasme* significantly affects the alpine ecosystem balance and animal husbandry sustainability [21]. Rapid and accurate monitoring of *S. chamaejasme* growth through the hyperspectral prediction of leaf chlorophyll can offer key support for the prevention of *S. chamaejasme* invasion and the management of degraded grasslands.

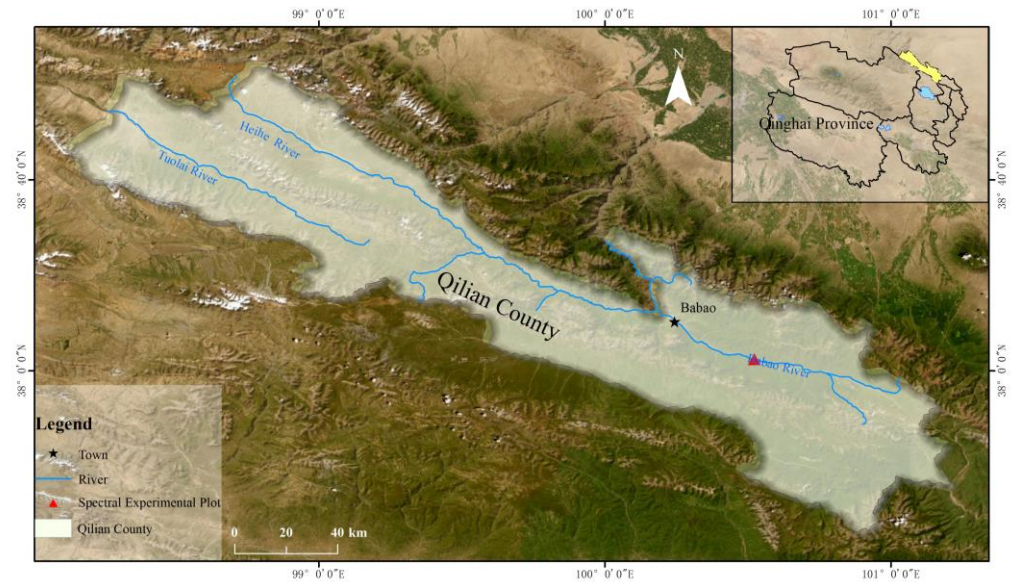
Therefore, the primary objectives of this study are as follows: (1) to select the chlorophyll content-sensitive wavelengths, parameters, and indices related to *S. chamaejasme* leaves, using a hierarchical procedure which combines Pearson correlation analysis, and first derivative and SPA algorithms; (2) to establish the comparative models for predicting the leaf chlorophyll content in *S. chamaejasme* across various growth stages via the stacking ensemble learning model and 1D-CNN model; and (3) to further determine the applicability of machine learning and deep learning algorithms in the hyperspectral estimation of the biochemical properties of toxic invasive species in alpine grasslands.

## 2. Materials and Methods

### 2.1. Experimental Site

*S. chamaejasme* is a perennial herb belonging to the *Thymelaeaceae* family, and the whole plant is toxic. Its height is about 20–50 cm, characterized by a terminal head in florescence, and it is white or red in color. The blooming occurs from late June to late July [22]. The study area is located in Qilian County, Haibei Tibetan Autonomous Prefecture, Qinghai Province, with average elevation of 3070 m. The region has a typical plateau continental climate; the annual average temperature ranges from  $-1.1$  °C to  $0.3$  °C, and the annual average precipitation is about 420 mm. The main vegetation type is alpine meadow, and the soil type is Mat Cry-gelic Cambisols. The experimental site is in Qingyangou, Babao

Town, Qilian County, located between  $100^{\circ}21'38''$  E,  $38^{\circ}9'32''$  N and  $100^{\circ}21'52''$  E,  $38^{\circ}9'40''$  N (Figure 1). *S. chamaejasme* dominates in the area, and is densely distributed in patches. Other dominant species are *Anemone rivularis*, *Thermopsis lanceolata*, *Anaphalis lactea*, *Morina kokonorica*, and so on [23]. The community coverage ranges from 26.0% to 63.0%. The average coverage of *S. chamaejasme* is 15.4%, with a maximum patch coverage of 38.5%.



**Figure 1.** Location of the study area.

## 2.2. Field Data Collection

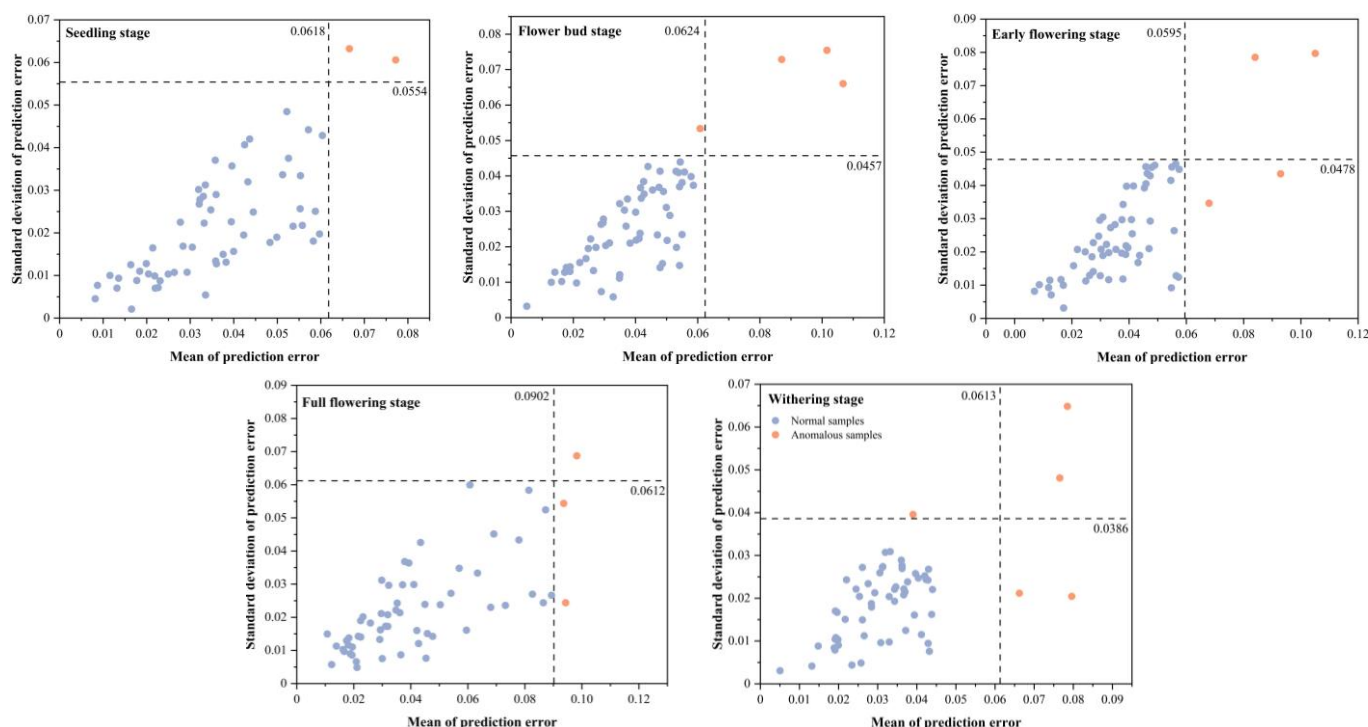
Field spectral reflectance and SPAD (Soil and Plant Analyzer Development) values of *S. chamaejasme* leaves were collected in the middle of July 2020 and 2021. *S. chamaejasme* growth was categorized into five distinct stages: the seedling stage, flower bud stage, early flowering stage, full flowering stage, and withering stage. Plant leaves of different stages were randomly selected in the experimental site. When sampling, the small leaves from one side of one plant were used for spectral measurements and the leaves from the other side were used for SPAD measurements. A total of 307 samples were collected across the five growth stages, with sample sizes of 60, 63, 62, 61, and 61.

The reflectance spectra of *S. chamaejasme* leaves were measured using a leaf clip loaded in an ASD Field Spec4 Hi-RES spectroradiometer. The instrument covers a wavelength measurement range of 350–2500 nm, with a spectral resolution of 3 nm at 700 nm and 8 nm at 1400/2100 nm, a spectral sampling interval of 1.4 nm from 350–1000 nm and 1.1 nm from 1001–2500 nm, a wavelength accuracy of  $\pm 0.1$  nm, and a field of view of  $25^{\circ}$ . Standard white plate calibration was conducted prior to the spectral measuring. Five to ten leaves were arranged in the leaf chamber flatly, ensuring no gaps between them. The measuring was conducted 10 times for each plant sample, and the average value was calculated as its reflectance. Konica Minolta chlorophyll meter SPAD-502 was used to measure SPAD value of *S. chamaejasme* leaves. The SPAD value functions as an indicator of the chlorophyll concentration as discussed by Yadava [24] and Ruiz-Espinoza et al. [25].

## 2.3. Data Preprocessing

In this study, the input spectra for the modeling of leaf chlorophyll content in *S. chamaejasme* was set to 350–1000 nm. Firstly, the Savitzky-Golay algorithm [26] with a smoothing window of  $3 \times 3$  was used for denoising via ViewSpecPro (version 5.6) software. The Monte Carlo method [27] can effectively identify and remove both spectral outliers and SPAD outliers and reasonably determine the number of samples in both the modeling and

prediction sets. Thus, the Monte Carlo method was then applied to examine the abnormal samples of *S. chamaejasme* leaf spectra and SPAD values, with the threshold of 2.5 times the mean and standard deviation of the prediction error for the sample set. The number of abnormal samples detected in the seedling stage, flower bud stage, early flowering stage, full flowering stage, and withering were 2, 4, 4, 3, and 5, respectively. The remaining samples of 58, 59, 58, 58, and 56 for the various stages were adopted for modeling and prediction (Figure 2). The Sample Set Partitioning Based on Joint X-Y Distance (SPXY) algorithm [28] can integrate both the spectral reflectance and SPAD value of each sample in determining sample spacing, thereby enhancing the predictive capability. Finally, the SPXY algorithm was employed to partition *S. chamaejasme* leaf samples for the various stages into a modeling set and a validation set, maintaining a 7:3 ratio (Table 1). Data preprocessing was performed via ViewSpecPro and MATLAB R2019b.



**Figure 2.** Detection of abnormal *S. chamaejasme* leaf samples for various growth periods via the Monte Carlo method. The numbers on the left of the vertical dashed lines represent 2.5 times the mean prediction error of SPAD values. The numbers under the horizontal dashed lines represent 2.5 times the mean standard deviation of SPAD values.

**Table 1.** *S. chamaejasme* leaf samples collected at different growth stages.

Growth Stage	Sample Set	Number	SPAD Value	Mean	Standard Deviation
Seedling stage	I	41	20.4~41.4	34.25	6.588
	II	17	22.9~42.2	34.65	5.665
Flower bud stage	I	42	20.2~38.1	34.46	6.039
	II	17	19.3~43	33.39	6.979
Early flowering stage	I	41	26.6~42.5	36.48	5.578
	II	17	27.8~45.3	37.72	4.565
Full flowering stage	I	41	28.5~42.5	37.85	5.028
	II	17	30.6~43	36.41	4.153



Table 1. Cont.

Growth Stage	Sample Set	Number	SPAD Value	Mean	Standard Deviation
Withering stage	I	40	35.4~45.5	40.15	4.634
	II	16	31.6~46	40.7	4.65
Whole growth stage	I	203	19.3~47.6	36.21	6.658
	II	86	22.9~42.7	37.33	4.942

I: Modeling set; II: Validation set.

### 2.4. Methodology

Figure 3 shows the workflow used to predict leaf chlorophyll content in *S. chamaejasme* leaves. First, hyperspectral data preprocessing was performed, followed by feature spectra selection using hierarchical dimensionality reduction. Next, hyperspectral prediction models for leaf chlorophyll content of *S. chamaejasme* were established using the stacking ensemble learning model and 1D-CNN model. Finally, the performance of the two models was evaluated and the applicability in various growth stages of *S. chamaejasme* was compared.

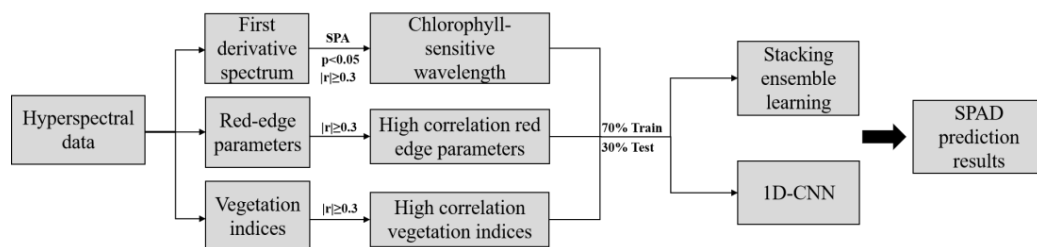


Figure 3. Workflow of hyperspectral prediction of the leaf chlorophyll content in *S. chamaejasme*.

#### 2.4.1. Hierarchical Dimensionality Reduction

The first derivative (FD) of *S. chamaejasme* leaf reflectance was calculated to enhance the spectral characteristics of the biochemical properties of leaves. The Pearson correlation coefficients ( $r$ ) between the FD values and the SPAD values were then determined for various growth stages. Based on the above steps, the leaf chlorophyll content-sensitive wavelengths with  $|r| \geq 0.3$  were detected ( $p < 0.05$ ). The SPA algorithm [29] can remove redundant information and reduce spectral dimensionality. Therefore, the SPA algorithm was further employed to refine and determine the set of leaf chlorophyll-sensitive wavelengths. The SPA algorithm was executed via MATLAB R2019b.

#### 2.4.2. Red-Edge Parameter and Spectral Index Calculation

The red-edge parameter is closely related to the leaf chlorophyll content, which changes as the growth stage progresses, and in turn leads to a more significant red shift [30]. The spectral index is created by the linear or nonlinear combination of specific spectral bands, which serves as the indicator of vegetation growth status [31]. Qiao et al. [32] explored the combination of red-edge parameters with spectral indices in order to improve the monitoring of vegetation health and chlorophyll content. In order to accurately simulate the chlorophyll content of *S. chamaejasme*, this study referenced Cui and Zhou [33] and Tong and He [34] and selected three red edge parameters and 12 spectral indices (Table 2). The red-edge parameters and spectral indices were calculated from the spectra of *S. chamaejasme* leaves at various growth stages. Then, Pearson correlation analysis was performed; the red-edge parameters and vegetation indices, which were significantly correlated with SPAD values ( $p < 0.05$ ), were identified as the leaf chlorophyll content-sensitive parameters/indices. These parameters and indices were calculated via Python 3.9 software.

As a result, the selected wavelengths, red-edge parameters, and spectral indices were used as feature spectra parameters for the following hyperspectral prediction.

**Table 2.** Red-edge parameters and vegetation indices.

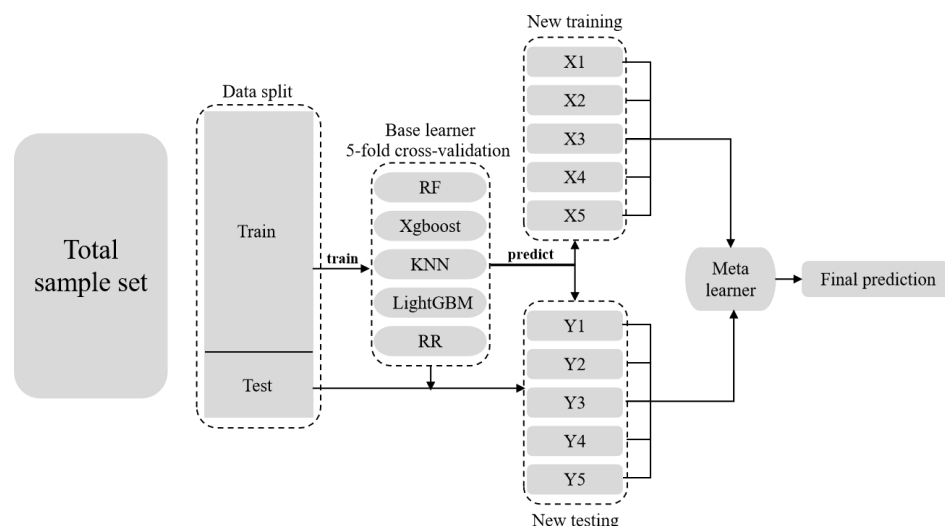
Red-Edge Parameter	Definition
Red-edge position ( $\lambda_r$ ) [35]	The wavelength corresponding to the maximum value of the first derivative spectrum within 680–760 nm.
Red-edge area (Sr) [36]	The area enclosed by the first derivative spectrum within 680–760 nm.
Red-edge amplitude ( $D\lambda_r$ ) [37]	The maximum value of the first derivative spectrum within 680–760 nm.
Spectral index	Calculation formula
Normalized Difference Vegetation Index (NDVI705) [38]	$(R_{750} - R_{705}) / (R_{750} + R_{705})$
Normalized Difference Vegetation Index (NDVI) [39]	$(R_{NIR} - R_{RED}) / (R_{NIR} + R_{RED})$
Normalized Chlorophyll Index (NCPI) [40]	$(R_{680} - R_{430}) / (R_{680} + R_{430})$
Lobate Vegetation Index (LCI) [41]	$(R_{850} - R_{710}) / (R_{850} - R_{680})$
Photochemical Reflectance Index (PRI) [42]	$(R_{531} - R_{570}) / (R_{531} + R_{570})$
Modified Chlorophyll Absorption Reflectance Index (MCARI) [43]	$[(R_{702} - R_{671}) - 0.2(R_{702} - R_{549})] (R_{702} / R_{671})$
Green Normalized Difference Vegetation Index (GNDVI) [44]	$(R_{750} - R_{550}) / (R_{750} + R_{550})$
Modified Normalized Difference Vegetation Index (MNDVI) [45]	$(R_{750} - R_{705}) / (R_{750} + R_{705} - 2 \times R_{445})$
Soil Adjusted Vegetation Index (SAVI) [46]	$1.5(R_{800} - R_{670}) / (R_{800} - R_{670} + 0.5)$
Enhanced Vegetation Index (EVI) [47]	$[2.5(R_{800} - R_{700})] / (R_{800} + 6R_{700} - 7.5R_{436} + 1)$
Difference Vegetation Index (DVI) [48]	$R_{NIR} - R_{RED}$
Ratio Vegetation Index (RVI) [49]	$R_{NIR} / R_{RED}$

#### 2.4.3. Stacking Ensemble Learning

The stacking ensemble learning model [50] integrates multiple base models through a meta model. The learning structure consists of two levels: a primary learner and a secondary learner. The establishment of secondary learner depends on the primary learner's outputs during the training processing. The procedure effectively promotes the outcome of the base learners via a meta model, making it appropriate for addressing more intricate issues. In this study, five algorithms were selected as base models.

- The random forest (RF) method exhibits high learning efficiency and robust generalization capabilities, making it appropriate for high-dimensional datasets.
- Extreme gradient boosting (XGBoost) accommodates custom loss functions, thereby facilitating a reduction in training errors.
- K-nearest neighbour (KNN) classifies a target point on the basis of the categories of the k-nearest sample data and operates without prior knowledge.
- The Light Gradient Boosting Machine (LightGBM) achieves high-precision predictions from a small set of samples through the implementation of the GOSS and EFB techniques.
- Ridge regression (RR) addresses the issue of multiple collinearities by modifying the regularization coefficient to mitigate overfitting.

Linear regression was chosen as the meta model for developing a model to estimate the leaf chlorophyll content in *S. chamaejasme* leaves. Using the feature spectra parameters as input, the five base models were trained via fivefold cross-validation. Then, based on the derived new training set and testing set, the meta model was established by linear regression (Figure 4). A grid search method for optimizing the parameters of base models was also involved. The method systematically explores all potential values of each parameter to determine the optimal parameter combination for the base modes (Table 3).



**Figure 4.** Framework of stacking ensemble model for predicting leaf chlorophyll content of *S. chamaejasme*.

**Table 3.** Optimal parameter combinations for predicting leaf chlorophyll content of *S. chamaejasme*.

Model	Parameter	Value					
		Seedling Stage	Flower Bud Stage	Early Flowering Stage	Full Flowering Stage	Withering Stage	Whole Growth Stage
RF	n_estimators	104	105	105	107	104	109
	Max_features	5	5	5	5	5	5
Xgboost	n_estimators	100	100	100	100	100	100
	Max_depth	3	4	3	3	3	5
KNN	K-neighbours	5	5	3	4	3	5
LightGBM	Max_depth	8	8	10	10	10	10
	Learning_rate	0.3	0.3	0.3	0.2	0.1	0.1
RR	Alpha	0.1	0.1	0.1	0.1	0.1	0.1

#### 2.4.4. One-Dimensional Convolutional Neural Network

The convolutional neural network (CNN) is superior in local connection and parameter sharing, which reduces the number of optimal parameters and improves model training efficiency. The application of this method in the spectral analysis of vegetation biochemical parameters has significant advantages. Therefore, this study explored the prediction of the leaf chlorophyll content in *S. chamaejasme* via a one-dimensional convolutional neural network (1D-CNN). Table 4 presents the model structure and parameter settings. The model comprised two convolutional layers, each utilizing a set of filters with a specified size of 5. The initial layer employed a standard convolution operation with a dilation factor of 1, facilitating the fundamental features extraction from *S. chamaejasme* leaf spectra. The second layer, performed with a dilation factor of 2, aimed to obtain broader contextual information and identify intricate spectral features. To improve the model’s nonlinear mapping ability, an ReLU activation function was incorporated with the two convolution layers. Then, a fully connected layer was employed as the output of the model to map the features extracted by the convolutional layers to the final prediction. The Adam optimizer was selected for model training, with a learning rate of 0.001 over 100 iterations. Also,

5-fold cross-validation was implemented on the training dataset to simulate the prediction model via 1D-CNN.

**Table 4.** 1D-CNN model parameter settings.

Network Layer	Model Parameters
Input layer	Feature spectra parameter of <i>S. chamaejasme</i> leaves
Average pooling layer	Pool size =10
Convolutional layer C1	filters = 16, filter size = 5, dilation = 1, ReLu activation function
Convolutional layer C2	filters = 32, filter size = 5, dilation = 2, ReLu activation function
Fully connected layer	Linear activation function
Output layer	Output prediction result

In this study, several strategies were implemented to enhance the model's generalization ability. Early stopping was initially used to track the loss variation on the validation set, thereby preventing model overfitting. Data augmentation was then employed to increase the diversity of the training data, thus improving the model's adaptability to unfamiliar data. Ultimately, dropout layers were introduced during training to mitigate the model's dependence on the training data and enhancing its robustness.

#### 2.4.5. Accuracy Evaluation

The performance of the model was evaluated via the coefficient of determination ( $R^2$ ) and the root mean square error (RMSE). The closer  $R^2$  is to 1, the greater the degree of agreement between the model prediction and the true value. The lower the RMSE value is, the more robust the model. The model accuracy evaluation was completed on the PyCharm platform via Python 3.9.

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

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - f(x_i))^2}, \quad (2)$$

where  $\hat{y}_i$  is the predicted value;  $y_i$  is the actual value; and  $\bar{y}$  is the mean value.

### 3. Results

#### 3.1. Hyperspectral Response of *S. chamaejasme* Leaves

Figure 5 shows that the maximum, minimum, and average SPAD value of *S. chamaejasme* leaves consistently increased with the progression of the growth cycle. From the seedling stage to the flower bud stage, the minimum and average SPAD values of *S. chamaejasme* leaves remained constant, whereas the maximum value decreased. From the early flowering stage via full flowering stage to withering stage, all three statistical values increased to different degrees, peaking at the withering stage. Notably, the minimum value showed the most significant increase.

Figure 6 shows that the reflectance spectra of *S. chamaejasme* leaves at various growth stages are consistently similar to green plants. In the visible light spectrum (350–689 nm), the leaf chlorophyll strongly absorbs blue and red light, resulting in the formation of a blue valley at 395–405 nm and a red valley at 670–675 nm, alongside a green peak at 550 nm owing to the partial reflection of green light. The spectral reflectance in the red edge band (690–749 nm) increases significantly, exceeding 0.45, and a highly reflective platform appears in the near-infrared band (750–1000 nm) with a range of 0.45–0.6. The reflectance spectra of *S. chamaejasme* leaves across various growth stages are similar within the visible



light band; however, the reflectances corresponding to the flower bud and early flowering stages are markedly lower than those corresponding to the other growth stages.

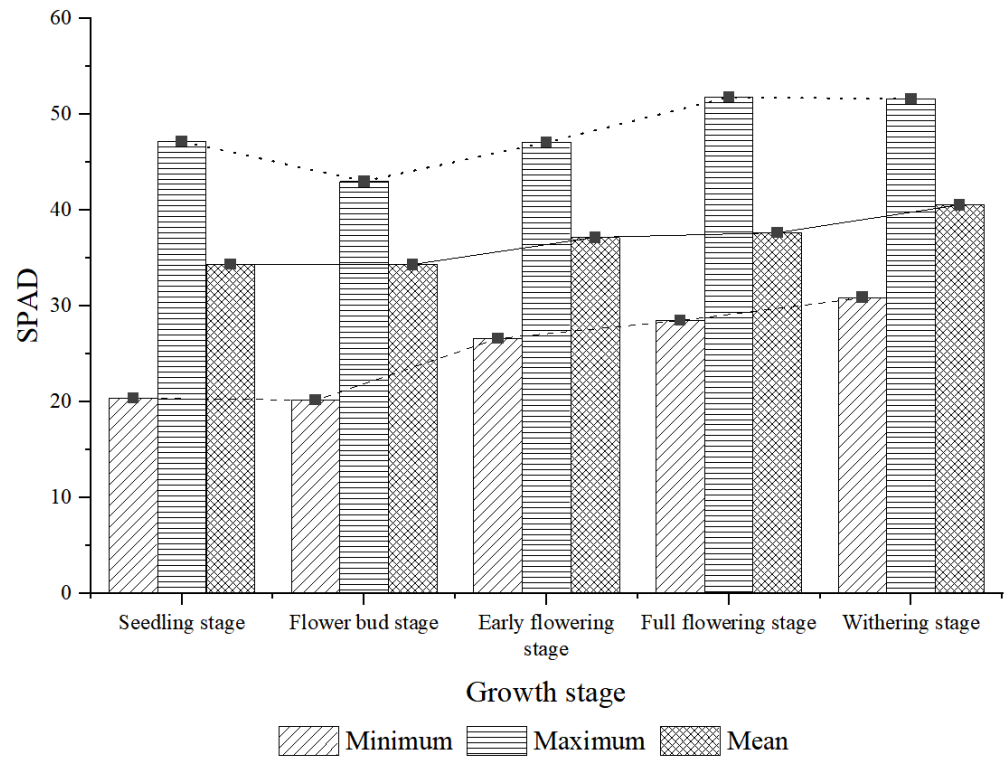


Figure 5. Statistical parameters of the SPAD value of *S. chamaejasme* leaves at different growth stages.

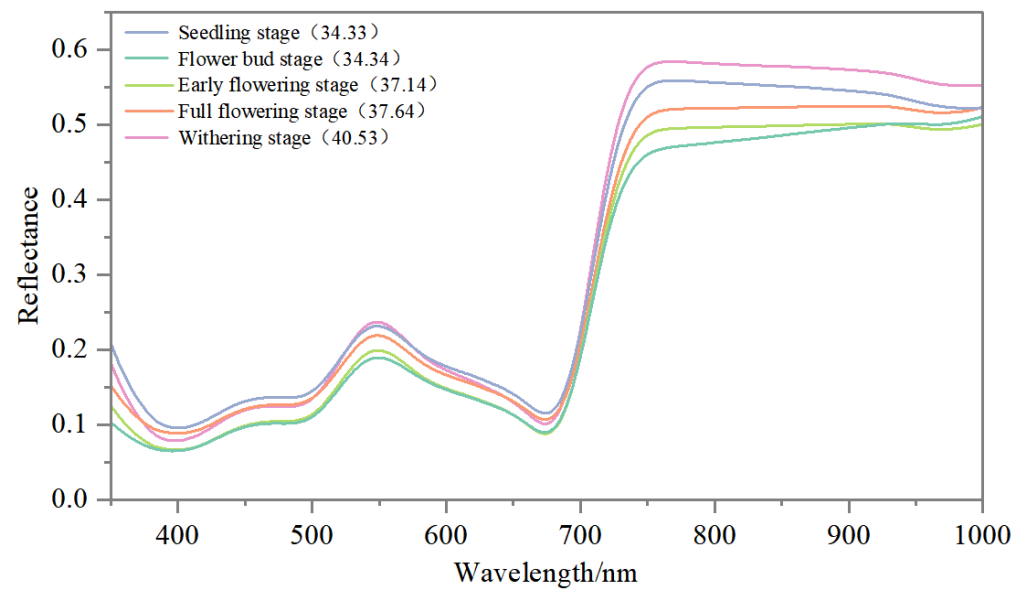


Figure 6. Reflectance spectra of *S. chamaejasme* leaves in the range of 350–1000 nm. Numbers in the brackets represent the average SPAD values of *S. chamaejasme* leaves at different growth stages.

In the near-infrared band, the reflectance of leaves exhibited notable variations across the various growth stages, which consistently increased throughout the growth cycle, beginning at the flower bud stage and peaking at the withering stage. As the growth period progressed from the flower bud stage, both the SPAD value and the spectral reflectance gradually increased. The result showed that the leaf chlorophyll content increases in

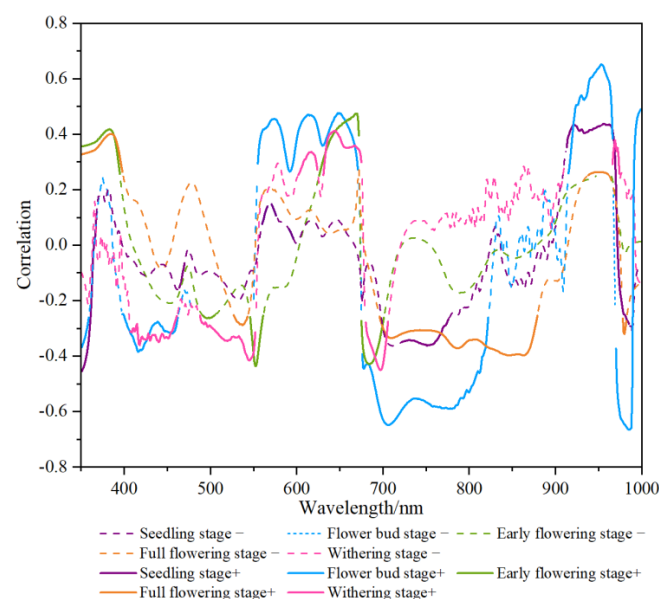
conjunction with an increase in the reflectance of *S. chamaejasme* leaves during the growth period, except the seedling stage.

### 3.2. Extraction of Leaf Chlorophyll-Sensitive Feature Spectra

Figure 7 shows the significant differences in the correlation between the first derivative values of the spectra of *S. chamaejasme* leaves at various growth stages and the SPAD values within the range of 350–1000 nm. The distributions of the positive and negative correlation coefficients are relatively balanced, with values ranging from  $-0.7$  to  $0.7$ . In this study, the wavelengths with the  $|r| \geq 0.3$  are emphasized. Table 5 indicates that the wavelengths sensitive to leaf chlorophyll content are primarily within the ranges of 351–400 nm, 555–675 nm, 680–880 nm, and 910–970 nm. The correlation between the first-order derivative spectrum and the SPAD value is most pronounced during the flower bud stage than during the other growth stages. The highest number of sensitive wavelengths (447) appears in this stage, and the strongest positive and negative correlations are at 350 nm and 985 nm, with correlation coefficient values of 0.652 and  $-0.665$ , respectively. There are 225 and 186 sensitive wavelengths for the full flowering and withering stages, demonstrating a significant correlation with the SPAD values. The correlations are relatively weak during the seedling and early flowering stages, with 140 and 120 sensitive wavelengths identified, respectively.

Leaf chlorophyll-sensitive wavelengths with  $|r| \geq 0.3$  at each growth stage were identified as input values, and SPAD values served as response values. SPA was further performed to identify the most sensitive wavelengths at various growth stages, yielding 13, 16, 15, 19, 13, and 28, respectively. These wavelengths were selected as feature wavelengths for hyperspectral prediction of *S. chamaejasme* leaf chlorophyll content (Table 6).

Table 7 indicates that with the exception of SAVI and  $Sr$ , the other red-edge parameters and spectral indices exhibit significant correlations with the leaf SPAD values ( $p < 0.05$ ). The correlation displays positive or negative, varying with different red-edge parameters and spectral indices. The proportion of high correlation ( $r \geq 0.7$ ) and medium correlation ( $0.3 \leq r < 0.7$ ) is 90%, whereas the proportion of weak correlation ( $r < 0.3$ ) is merely 10%. Consequently, the red-edge parameters and spectral indices of high and medium correlations were selected as the feature parameters and indices for the leaf chlorophyll estimation across various growth periods.



**Figure 7.** Correlation between SPAD values and first derivative spectra of *S. chamaejasme* leaves in 350–1000 nm. +: significant; -: not significant ( $p < 0.05$ ).

**Table 5.** Leaf chlorophyll-sensitive wavelength of *S. chamaejasme* selected by correlation analysis ( $|r| \geq 0.30$ ).

Growth Stage	Seedling Stage	Flower Bud Stage	Early Flowering Stage	Full Flowering Stage	Withering Stage
Wavelength range	351~368, 381~385, 464~470, 478~479, 540~543, 560~566, 571~576, 591~596, 692~696, 722~776, 778~784, 797~804, 820~829, 834~841, 864, 876~877, 881~882, 890~892, 896~907, 909~911, 913~952, 954~985	351~359, 398~462, 479~549, 555~671, 676~822, 915~966, 970~989, 991~1000	351~395, 491~505, 546~561, 626~672, 675~701, 948~964	351~398, 527~543, 672, 696~879, 938~961, 978~982	407~461, 488~552, 574~583, 602~624, 631~675, 680~706, 862~865, 868~870, 909~912, 966~974, 976~977
Wavelength Number	140	447	120	225	186
Max_r	-0.459	-0.665	0.475	0.401	-0.451
Wavelength_Max_r	350	985	669	384	697

**Table 6.** Leaf chlorophyll-sensitive wavelength of *S. chamaejasme* selected by SPA (sorted by importance).

Growth Stage	Wavelength/nm
Seedling stage	751, 756, 995, 677, 479, 721, 774, 561, 639, 350, 358, 934, 951
Flower bud stage	432, 577, 618, 652, 704, 778, 954, 983, 985, 699, 643, 720, 387, 470, 924, 362
Early flowering stage	379, 554, 671, 686, 652, 627, 669, 412, 428, 395, 674, 518, 961, 465, 375
Full flowering stage	388, 539, 650, 785, 866, 979, 772, 384, 387, 744, 400, 854, 428, 449, 540, 578, 374, 519, 720
Withering stage	426, 545, 694, 774, 813, 697, 962, 574, 979, 382, 750, 605, 548
Whole growth period	916, 946, 880, 968, 820, 768, 586, 569, 695, 678, 665, 621, 982, 737, 386, 389, 373, 383, 400, 711, 551, 522, 646, 380, 394, 408, 426, 473

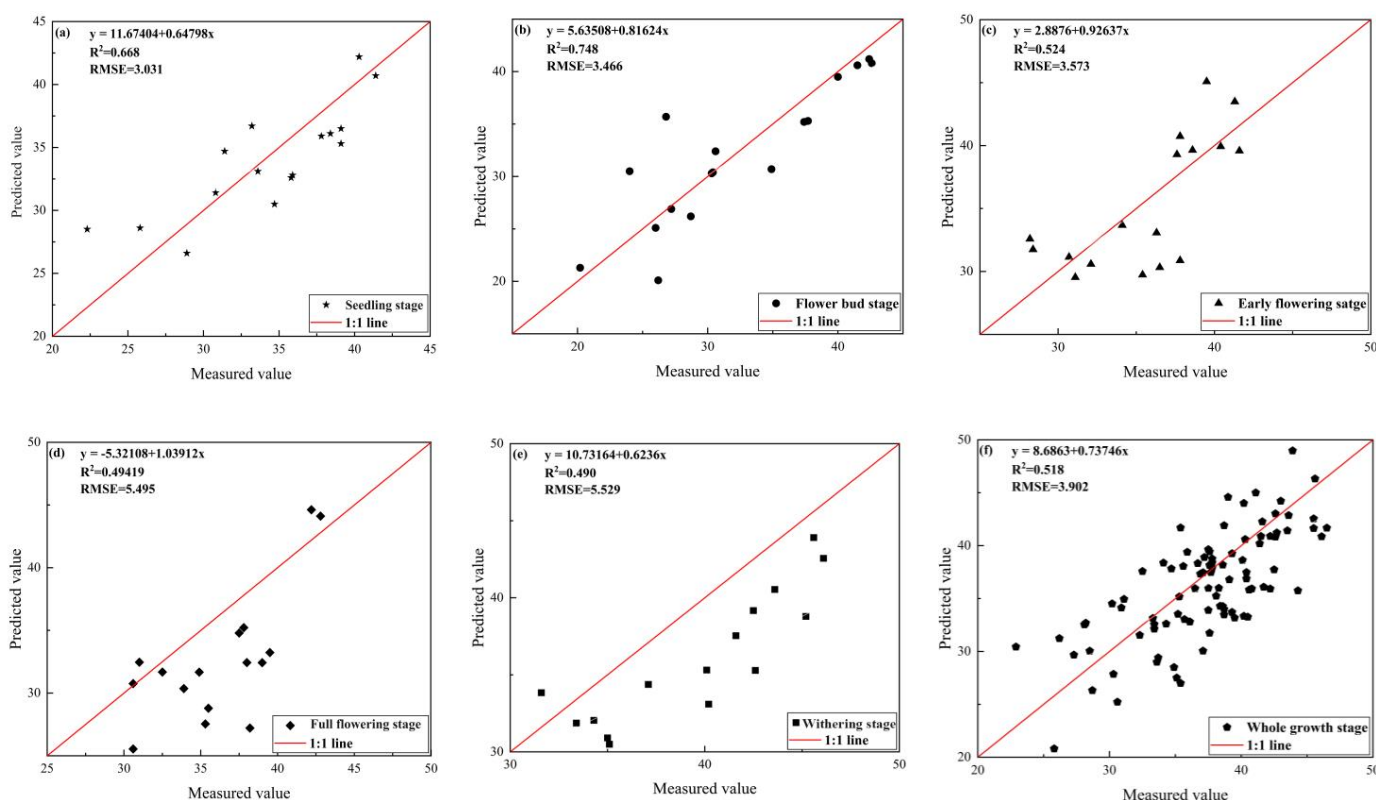
**Table 7.** Pearson correlation between SPAD values and red-edge parameters and spectral indices.

Spectral Index	Correlation Coefficient (r)					
	Seedling Stage	Flower Bud Stage	Early Flowering Stage	Full Flowering Stage	Withering Stage	Whole Growth Stage
NDVI	-0.382 *	-0.332 *	-0.442 *	-0.210 *	-0.441 *	-0.423 *
NDVI <sub>705</sub>	-0.427 *	-0.245 *	-0.563 *	-0.331 *	-0.562 *	-0.545 *
NCPI	-0.562 *	-0.367 *	-0.604 *	-0.452 *	-0.603 *	-0.666 *
LCI	0.368 *	0.489 *	0.187 *	0.573 *	0.345 *	0.444 *
PRI	0.588 *	0.512 *	0.209 *	0.644 *	0.467 *	0.567 *
SAVI	0.302	0.634	0.330	0.232	0.589	0.600
MCARI	-0.412 *	-0.156 *	-0.451 *	-0.353 *	-0.621 *	-0.468 *
GNDVI	0.651 *	0.278 *	0.572 *	0.474 *	0.368 *	0.588 *
MNDVI	0.653 *	0.390 *	0.613 *	0.595 *	0.480 *	0.620 *
EVI	0.611 *	0.412 *	0.198 *	0.616 *	0.511 *	0.489 *
DVI	0.568 *	0.534 *	0.220 *	0.301 *	0.632 *	0.510 *
RVI	-0.391 *	-0.656 *	-0.341 *	-0.422 *	-0.399 *	-0.633 *
$\lambda_r$	0.609 *	0.378 *	0.462 *	0.543 *	0.410 *	0.499 *
$S_r$	0.333	0.399	0.583	0.664	0.531	0.555
$D\lambda_r$	-0.554 *	-0.321 *	-0.624 *	-0.320 *	-0.652 *	-0.645 *

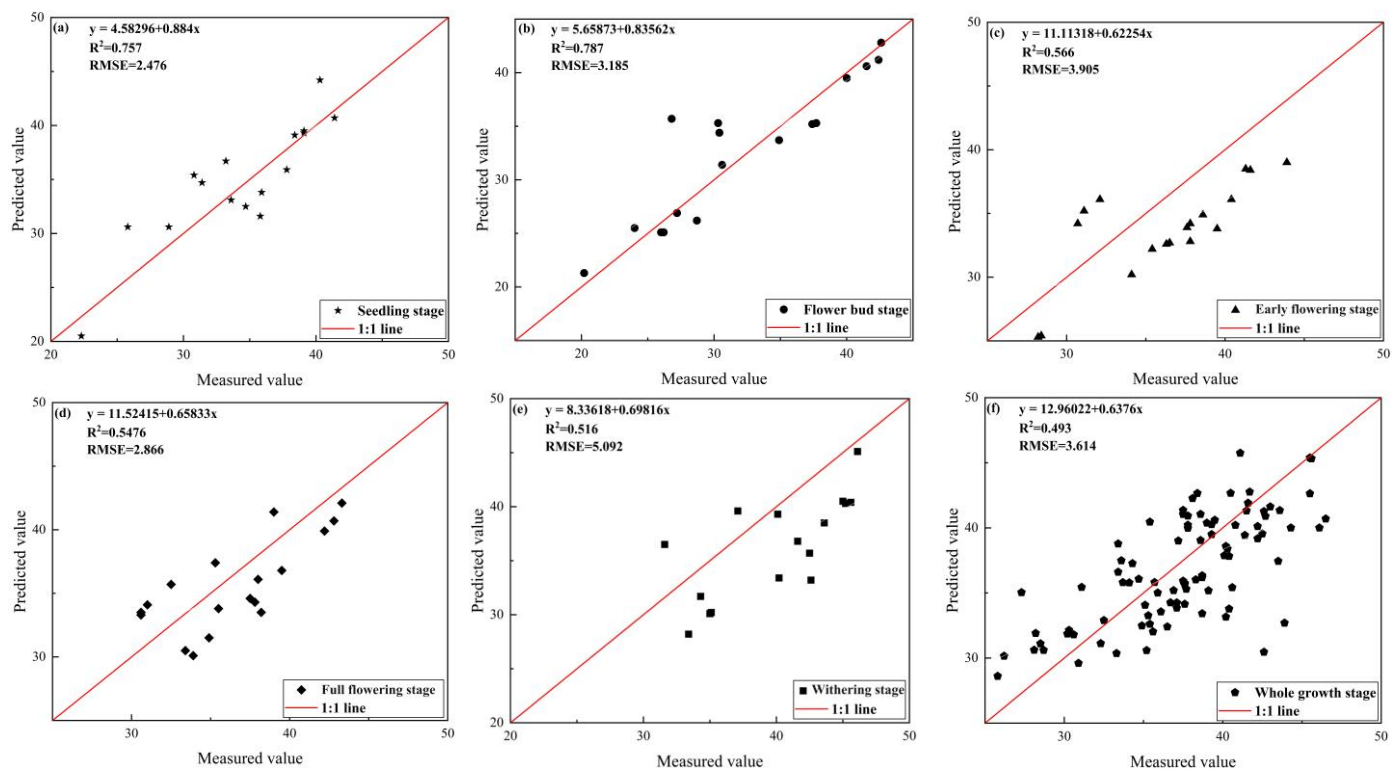
\*  $p < 0.05$ .

### 3.3. Establishment of Models for Predicting Leaf Chlorophyll Content

The selected feature wavelength, parameters, and indices served as the input, and the SPAD values as the output; the models for predicting leaf chlorophyll content were developed via the stacking ensemble learning model and the 1D-CNN model. Figures 8 and 9 indicate that the prediction accuracy for the seedling and flower bud stages is optimal with  $R^2$  greater than 0.65 and RMSE less than 3.5. In contrast, the prediction accuracy for the early flowering, full flowering, withering stages, and the entire growth stage is relatively lower, as indicated by  $R^2$  values less than 0.6 and RMSE values exceeding 3.5, with the exception of 1D-CNN in the full-blooming stage. Compared with the stacking ensemble learning model, the 1D-CNN model enhances the prediction accuracy of SPAD values of *S. chamaejasme* leaves across various growth stages (Table 8). The  $R^2$  of the modeling set increases by 0.027 to 0.087, whereas the RMSE decreases by 0.213 to 2.429. The  $R^2$  value for the validation set improves by 0.026 to 0.089, whereas the RMSE decreases by 0.281 to 2.629. The prediction accuracy for the flower bud stage based on the 1D-CNN model is the highest, with a validation  $R^2$  of 0.787 and an RMSE of 2.476. Conversely, the prediction accuracy for the withering stage based on the stacking ensemble learning model is the lowest, with a validation  $R^2$  of 0.490 and an RMSE of 5.529. Compared with the 1D-CNN model, the stacking ensemble learning model achieves better prediction accuracy for the SPAD values in the whole growth period, with  $R^2 = 0.518$  and RMSE = 3.902 on the validation set.



**Figure 8.** Prediction accuracy of leaf chlorophyll content in *S. chamaejasme* leaves based on stacking ensemble learning models. (a) Seedling stage; (b) Flower bud stage; (c) Early flowering stage; (d) Full flowering stage; (e) Withering stage; (f) Whole growth stage.



**Figure 9.** Prediction accuracy of leaf chlorophyll content in *S. chamaejasme* based on 1D-CNN model. (a) Seedling stage; (b) Flower bud stage; (c) Early flowering stage; (d) Full flowering stage; (e) Withering stage; (f) Whole growth stage.

**Table 8.** Accuracy evaluation of the prediction models for the leaf chlorophyll content of *S. chamaejasme*.

Growth Stage	Stacking Ensemble Learning				1D-CNN			
	Model Accuracy		Validation Accuracy		Model Accuracy		Validation Accuracy	
	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE
Seedling stage	0.677	2.905	0.668	3.031	0.764	2.334	0.757	2.476
Flower bud stage	0.755	3.278	0.748	3.466	0.792	3.065	0.787	3.185
Early flowering stage	0.529	3.353	0.524	3.573	0.582	3.875	0.566	3.905
Full flowering stage	0.497	5.238	0.494	5.495	0.574	2.809	0.548	2.866
Withering stage	0.494	5.312	0.490	5.529	0.521	5.078	0.516	5.092
Whole growth stage	0.533	3.028	0.518	3.902	0.507	3.386	0.493	3.614

Generally, the 1D-CNN model is more effective for predicting the chlorophyll content in *S. chamaejasme* leaves across various growth periods, whereas the stacking ensemble learning model is preferable for the whole growth stage. It also can be observed that the prediction accuracy of the *S. chamaejasme* leaf chlorophyll content consistently decreases as the growth stage progresses, demonstrating optimal prediction results for the flower bud stage and suboptimal results for the withering stage.

### 4. Discussion

#### 4.1. The Application of Feature Spectra Selection in Leaf Chlorophyll Content Prediction

Hyperspectral data possess numerous bands, leading to significant redundancy and noise. This complicates the extraction of meaningful biochemical information and limits the generalizability of predictive models. Dimensionality reduction effectively mitigates overfitting multicollinearity, which is essential for extract plant properties [51]. First-order



derivative transformation, correlation analysis, and the SPA algorithm are commonly employed for the spectral dimension reduction. First-order derivative transformation enhances the spectral response of plant characteristics in comparison with the original spectrum [52–54]. Correlation analysis identifies significant wavelengths that are sensitive to plant biochemical parameters [55]. SPA effectively reduces the dimensionality of spectral data and enhances modeling accuracy. However, it is difficult to select the most reasonable feature wavelengths just using a single algorithm. Grassland spectra are influenced by various environmental factors such as soil type, vegetation structure, and atmospheric conditions. Compared to extensively cultivated crops, the leaf chlorophyll content prediction for grassland species is more challenging, due to unobvious variations in chlorophyll contamination throughout their growth periods. So, the application of hierarchical dimensionality reduction strategies is particularly essential. Zhang et al. [56] used hyperspectral data combined with first-order derivative spectra and PCA to estimate the chlorophyll content in the Hulunbuir grassland of Inner Mongolia, and the performance of the prediction model was significantly improved. In our study, the spectral wavelength number was reduced from 651 to a minimum of 13 through the three levels of dimensionality reduction, obtaining a reduction efficiency of 98%. As a result, the optimal feature bands for the chlorophyll content prediction of *S. chamaejasme* were successfully identified (Table 6). Our result is consistent with the previous findings, which achieved overall reduction efficiencies of approximately 97.9% and 97.6% [57,58]. The studies demonstrate the distinct benefits of employing multiple dimensionality reduction techniques for the extraction of optimal wavelengths in vegetation monitoring. Incorporating red-edge parameters and spectral indices can enhance the model's sensitivity to chlorophyll concentration in plants. The combination of vegetation indices with dimensionality reduction techniques improves the accuracy of chlorophyll content estimation models, particularly for grass species with significant variations in canopy structure [59]. Also, the integration of vegetation indices with machine learning models markedly enhances the robustness and applicability of chlorophyll estimation models in northern Australian grasslands [60]. In this study, Pearson correlation analysis was employed to identify the red-edge parameters and spectral indices sensitive to SPAD values of *S. chamaejasme* leaves, thereby fortifying the spectral responses related to the leaf chlorophyll content in *S. chamaejasme*.

#### 4.2. The Accuracy of Prediction Models via Machine Learning and Deep Learning

The stacking ensemble learning model exhibits distinct advantages and potential in crops growth monitoring. Yang et al. [61] reported that the model outperformed a single machine learning algorithm in predicting potato leaf chlorophyll content, achieving accuracies of  $R^2 = 0.839$  and  $RMSE = 0.261$ . Similarly, Chen et al. [62] reported that the model excelled in simulating the physiological parameters of maize under drip irrigation, with an  $R^2$  of 0.9 and an  $RMSE$  of 0.23, which was an 11% promotion in accuracy compared to those of a single model. These studies provide new ways for monitoring the growth of grassland invasive species. Our work reveals that, in addition to the full flowering and withering stages, the stacking ensemble learning model obtains relatively good prediction accuracy for predicting the chlorophyll content in *S. chamaejasme* leaves in other growth stages and the whole growth period, achieving a validation  $R^2 > 0.5$  and  $RMSE < 3.5$ . The prediction accuracy for the flower bud stage is the highest ( $R^2 = 0.748$ ,  $RMSE = 3.466$ ). It is noted that the estimation of chlorophyll content in *S. chamaejasme* leaves is inferior to crops leaves, which is mainly caused by the obvious ecological and physical variations between natural plants and cultivated crops. The concurrence of *S. chamaejasme* plants at various growth stages results in a complex spectral response. This diminishes the spectral

difference of chlorophyll content in the leaves, thereby impacting the model's effectiveness and generalizability.

Recent studies have extensively examined the use of deep convolutional neural network models in vegetation spectral analysis. Padarian et al. [63] and Furbank et al. [64] reported that the 1D-CNN model outperforms machine learning models in the analysis of soil spectra and physiological traits of wheat leaves, with an  $R^2$  improvement exceeding 15%. In this study, the 1D-CNN model exhibits generally better prediction accuracy for the chlorophyll content in *S. chamaejasme* leaves in the various growth stages, compared with the stacking ensemble learning model. The optimal prediction was achieved during the flower bud stage ( $R^2 = 0.787$ , RMSE = 3.185).

The 1D-CNN model excels at capturing the nonlinear and complex differences in hyperspectral data. The algorithm outperformed in predicting the chlorophyll content in *S. chamaejasme* leaves at each individual growth stage. The 1D-CNN model faces challenges in the case of a large number of wavelengths across the whole growth stage. This reduces the network's generalization ability and consequently lowers its predictive performance [65]. The stacking ensemble learning model is well-suited for large datasets and capable of handling diverse data [66]. Thus, the model demonstrates superior prediction accuracy for the chlorophyll content in *S. chamaejasme* leaves during the whole growth period. The result reveals the different performances of machine learning and deep learning algorithms in the hyperspectral prediction of plant biochemical properties. Due to the similarity and stability in the characteristics of leaf chlorophyll of *S. chamaejasme*, the proposed procedure provides a good foundation to enable timely monitoring and effective management of this species over large scales. Also, our research demonstrates the broad potential of machine learning and deep learning in the hyperspectral inversion of key traits of other toxic invasive species in alpine grasslands.

## 5. Conclusions

This paper proposes a procedure that integrates hierarchical dimensionality reduction and machine learning/deep learning algorithms to establish prediction models for the chlorophyll content in *S. chamaejasme* leaves across various growth stages. In comparison, the 1D-CNN model achieves superior prediction accuracy across the various growth stages, whereas the stacking ensemble learning model yields the most effective prediction results during the whole growth period. This work offers a reference for developing rapid, efficient, and non-destructive methods to predict vegetation biochemical parameters via hyperspectral data. Future research will be carried out to optimize input parameters, enhance the learning ability of the chlorophyll prediction model, and improve the applicability of the proposed prediction model.

**Author Contributions:** Conceptualization and methodology, Y.L. (Yongmei Liu) and X.L.; software and formal analysis, X.L.; investigation, Y.L. (Yongmei Liu), H.W., X.D.; resources, Y.L. (Yongmei Liu); writing—original draft preparation, X.L. and Y.L. (Yongmei Liu); writing—review and editing, Y.L. (Yongmei Liu) and X.L.; visualization, X.L.; supervision, L.W. and Y.L. (Yongqing Long); funding acquisition, Y.L. (Yongmei Liu). All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the National Natural Science Foundation of China (Grant No. 41871335).

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available on request from the corresponding author.

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

## References

- Croft, H.; Chen, J.M.; Wang, R.; Mo, G.; Luo, S.; Luo, X.; He, L.; Gonsamo, A.; Arabian, J.; Zhang, Y.; et al. The Global Distribution of Leaf Chlorophyll Content. *Remote Sens. Environ.* **2020**, *236*, 111479. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Croft, H.; Chen, J.M.; Luo, X.; Bartlett, P.; Chen, B.; Staebler, R.M. Leaf Chlorophyll Content as a Proxy for Leaf Photosynthetic Capacity. *Glob. Change Biol.* **2017**, *23*, 3513–3524. [\[CrossRef\]](#)
- Xu, M.; Liu, R.; Chen, J.M.; Liu, Y.; Shang, R.; Ju, W.; Wu, C.; Huang, W. Retrieving Leaf Chlorophyll Content Using a Matrix-Based Vegetation Index Combination Approach. *Remote Sens. Environ.* **2019**, *224*, 60–73. [\[CrossRef\]](#)
- Porra, R.J.; Thompson, W.A.; Kriedemann, P.E. Determination of Accurate Extinction Coefficients and Simultaneous Equations for Assaying Chlorophylls a and b Extracted with Four Different Solvents: Verification of the Concentration of Chlorophyll Standards by Atomic Absorption Spectroscopy. *Biochim. Biophys. Acta-Bioenerg.* **1989**, *975*, 384–394. [\[CrossRef\]](#)
- Steele, M.R.; Gitelson, A.A.; Rundquist, D.C. A Comparison of Two Techniques for Nondestructive Measurement of Chlorophyll Content in Grapevine Leaves. *Agron. J.* **2008**, *100*, 779–782. [\[CrossRef\]](#)
- Zhang, K.; Li, W.; Li, H.; Luo, Y.; Li, Z.; Wang, X.; Chen, X. A Leaf-Patchable Reflectance Meter for In Situ Continuous Monitoring of Chlorophyll Content. *Adv. Sci.* **2023**, *10*, 2305552. [\[CrossRef\]](#)
- Croft, H.; Chen, J.M.; Zhang, Y. The Applicability of Empirical Vegetation Indices for Determining Leaf Chlorophyll Content over Different Leaf and Canopy Structures. *Ecol. Complex.* **2014**, *17*, 119–130. [\[CrossRef\]](#)
- Ma, W.; Wang, X. Progress on Grassland Chlorophyll Content Estimation by Hyperspectral Analysis. *Prog. Geogr.* **2016**, *35*, 25–34. [\[CrossRef\]](#)
- Moharana, S.; Dutta, S. Spatial Variability of Chlorophyll and Nitrogen Content of Rice from Hyperspectral Imagery. *ISPRS J. Photogramm. Remote Sens.* **2016**, *122*, 17–29. [\[CrossRef\]](#)
- Zhao, X.; Sun, X.; Wang, F.; Xie, X.; Guo, X. A Summary of the Researches on Hyperspectral Remote Sensing Monitoring of Rice. *Acta Agric. Univ. Jiangxiensis* **2019**, *41*, 1–12. [\[CrossRef\]](#)
- Ali, A.M.; Darvishzadeh, R.; Skidmore, A.; Gara, T.W.; O'Connor, B.; Roeoesli, C.; Heurich, M.; Paganini, M. Comparing Methods for Mapping Canopy Chlorophyll Content in a Mixed Mountain Forest Using Sentinel-2 Data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *87*, 102037. [\[CrossRef\]](#)
- An, G.; Xing, M.; He, B.; Liao, C.; Huang, X.; Shang, J.; Kang, H. Using Machine Learning for Estimating Rice Chlorophyll Content from In Situ Hyperspectral Data. *Remote Sens.* **2020**, *12*, 3104. [\[CrossRef\]](#)
- Li, C.; Liu, Y.; Qin, T.; Wang, Y. Estimation of Chlorophyll Content in Potato Leaves Based on Machine Learning. *Spectrosc. Spect. Anal.* **2024**, *44*, 1117–1127. [\[CrossRef\]](#)
- Gan, H.; Yue, X.; Hong, T.; Ling, K.; Wang, L.; Cen, Z. A Hyperspectral Inversion Model for Predicting Chlorophyll Content of Longan Leaves Based on Deep Learning. *J. South China Agric. Univ.* **2018**, *39*, 102–110. [\[CrossRef\]](#)
- Putra, B.T.W.; Wirayuda, H.C.; Syahputra, W.N.H.; Prastowo, E. Evaluating In-Situ Maize Chlorophyll Content Using an External Optical Sensing System Coupled with Conventional Statistics and Deep Neural Networks. *Measurement* **2021**, *189*, 110482. [\[CrossRef\]](#)
- Sonobe, R.; Hirono, Y.; Oi, A. Quantifying Chlorophyll-*a* and *b* Content in Tea Leaves Using Hyperspectral Reflectance and Deep Learning. *Remote Sens. Lett.* **2020**, *11*, 933–942. [\[CrossRef\]](#)
- Zhang, A.; Yin, S.; Wang, J.; He, N.; Chai, S.; Pang, H. Grassland Chlorophyll Content Estimation from Drone Hyperspectral Images Combined with Fractional-Order Derivative. *Remote Sens.* **2023**, *15*, 5623. [\[CrossRef\]](#)
- Ji, T.; Liu, X. Establishing a Hyperspectral Model for the Chlorophyll and Crude Protein Content in Alpine Meadows Using a Backward Feature Elimination Method. *Agriculture* **2024**, *14*, 757. [\[CrossRef\]](#)
- Bao, G.; Wang, Y.; Song, M.; Wang, H.; Yin, Y. Effects of *Stellera Chamaejasme* Patches on the Surrounding Grassland Community and on Soil Physicochemical Properties in Degraded Grasslands Susceptible to *S. Chamaejasme* Invasion. *Acta Prataculturae Sin.* **2019**, *28*, 51–61. [\[CrossRef\]](#)
- Guo, L.; Wang, K. Research Progress on Biology and Ecology of *Stellera chamaejasme* L. *Acta Agrestia Sin.* **2018**, *26*, 525–532. [\[CrossRef\]](#)
- Shi, Z. *Important Poisonous Plants in Grassland of China*; China Agriculture Press: Beijing, China, 1997; ISBN 7-109-04658-3.
- Liu, Y.; Dong, X.; Long, Y.; Zhu, Z.; Wang, L. Classification of *Stellera Chamaejasme* Communities and Their Relationships with Environmental Factors in Degraded Alpine Meadow in the Central Qilian Mountains, Qinghai Province. *Acta Prataculturae Sin.* **2022**, *31*, 1–11. [\[CrossRef\]](#)
- Yadava, U.L. A Rapid and Nondestructive Method to Determine Chlorophyll in Intact Leaves. *HortScience* **1986**, *21*, 1449–1450. [\[CrossRef\]](#)

25. Ruiz-Espinoza, F.H.; Murillo-Amador, B.; García-Hernández, J.L.; Fenech-Larios, L.; Rueda-Puente, E.O.; Troyo-Diéguez, E.; Kaya, C.; Beltrán-Morales, A. Field Evaluation of the Relationship between Chlorophyll Content in Basil Leaves and a Portable Chlorophyll Meter (Spad-502) Readings. *J. Plant Nutr.* **2010**, *33*, 423–438. [[CrossRef](#)]
26. Luo, J.; Ying, K.; Bai, J. Savitzky–Golay Smoothing and Differentiation Filter for Even Number Data. *Signal Process.* **2005**, *85*, 1429–1434. [[CrossRef](#)]
27. Wisnowski, J.; Montgomery, D.; Simpson, J. A Comparative Analysis of Multiple Outlier Detection Procedures in the Linear Regression Model. *Comput. Stat. Data Anal.* **2001**, *36*, 351–382. [[CrossRef](#)]
28. Wang, S.; Han, P.; Cui, G.; Wang, D.; Liu, S.; Zhao, Y. The NIR Detection Research of Soluble Solid Content in Watermelon Based on SPXY Algorithm. *Spectrosc. Spect. Anal.* **2019**, *39*, 738–742.
29. Soares, S.F.C.; Gomes, A.A.; Araujo, M.C.U.; Filho, A.R.G.; Galvão, R.K.H. The Successive Projections Algorithm. *Trends Anal. Chem.* **2013**, *42*, 84–98. [[CrossRef](#)]
30. Horler, D.N.H.; Dockray, M.; Barber, J.; Barringer, A.R. Red Edge Measurements for Remotely Sensing Plant Chlorophyll Content. *Adv. Space Res.* **1983**, *3*, 273–277. [[CrossRef](#)]
31. Sun, W.; Du, Q. Hyperspectral Band Selection: A Review. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 118–139. [[CrossRef](#)]
32. Qiao, L.; Tang, W.; Gao, D.; Zhao, R.; An, L.; Li, M.; Sun, H.; Song, D. UAV-Based Chlorophyll Content Estimation by Evaluating Vegetation Index Responses under Different Crop Coverages. *Comput. Electron. Agric.* **2022**, *196*, 106775. [[CrossRef](#)]
33. Cui, S.; Zhou, K. A Comparison of the Predictive Potential of Various Vegetation Indices for Leaf Chlorophyll Content. *Earth Sci. Inf.* **2017**, *10*, 169–181. [[CrossRef](#)]
34. Tong, A.; He, Y. Estimating and Mapping Chlorophyll Content for a Heterogeneous Grassland: Comparing Prediction Power of a Suite of Vegetation Indices across Scales between Years. *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 146–167. [[CrossRef](#)]
35. Clevers, J.G.P.W.; Jong, H.; Epema, G.F.; Van Der Meer, F.; Bakker, W.; Skidmore, A.; Addink, E.A. MERIS and the Red-Edge Position. *Int. J. Appl. Earth Obs. Geoinf.* **2001**, *3*, 313–320. [[CrossRef](#)]
36. Guo, B.; Zhu, Y.; Feng, W.; He, L.; Wu, Y.; Zhou, Y.; Ren, X.; Ma, Y. Remotely Estimating Aerial N Uptake in Winter Wheat Using Red-Edge Area Index from Multi-Angular Hyperspectral Data. *Front. Plant Sci.* **2018**, *9*, 675. [[CrossRef](#)]
37. Filella, I.; Penuelas, J. The Red Edge Position and Shape as Indicators of Plant Chlorophyll Content, Biomass and Hydric Status. *Int. J. Remote Sens.* **1994**, *15*, 1459–1470. [[CrossRef](#)]
38. Evangelides, C.; Nobajas, A. Red-Edge Normalised Difference Vegetation Index (NDVI<sub>705</sub>) from Sentinel-2 Imagery to Assess Post-Fire Regeneration. *Remote Sens. Appl.* **2020**, *17*, 100283. [[CrossRef](#)]
39. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with ERTS. *NASA Spec. Publ.* **1974**, *1*, 309–317.
40. Yoder, B.J.; Waring, R.H. The Normalized Difference Vegetation Index of Small Douglas-Fir Canopies with Varying Chlorophyll Concentrations. *Remote Sens. Environ.* **1994**, *49*, 81–91. [[CrossRef](#)]
41. Zhang, N.; Xiong, H.; Jin, Y. Hyperspectral Estimation Models for Chlorophyll Content Based on the Measured Spectra of Safflower. *Hubei Agric. Sci.* **2016**, *55*, 5651–5658. [[CrossRef](#)]
42. Gamon, J.A.; Peñuelas, J.; Field, C.B. A Narrow-Waveband Spectral Index That Tracks Diurnal Changes in Photosynthetic Efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [[CrossRef](#)]
43. Daughtry, C.S.; Walthall, C.; Kim, M.; De Colstoun, E.B.; McMurtrey Iii, J. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
44. Daughtry, C.S.T.; Gallo, K.P.; Goward, S.N.; Prince, S.D.; Kustas, W.P. Spectral Estimates of Absorbed Radiation and Phytomass Production in Corn and Soybean Canopies. *Remote Sens. Environ.* **1992**, *39*, 141–152. [[CrossRef](#)]
45. Jurgens, C. The Modified Normalized Difference Vegetation Index (mNDVI) a New Index to Determine Frost Damages in Agriculture Based on Landsat TM Data. *Int. J. Remote Sens.* **1997**, *18*, 3583–3594. [[CrossRef](#)]
46. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
47. Huete, A.R.; Liu, H.Q.; Batchily, K.; Leeuwen, W. van A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [[CrossRef](#)]
48. Roujean, J.-L.; Breon, F.-M. Estimating PAR Absorbed by Vegetation from Bidirectional Reflectance Measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [[CrossRef](#)]
49. Birth, G.S.; McVey, G.R. Measuring the Color of Growing Turf with a Reflectance Spectrophotometer. *Agron. J.* **1968**, *60*, 640–643. [[CrossRef](#)]
50. Ganaie, M.A.; Hu, M.; Tanveer, M.; Suganthan, P.N.; Malik, A.K. Ensemble Deep Learning: A Review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105151. [[CrossRef](#)]
51. Mojaradi, B.; Abrishami-Moghaddam, H.; Valadan Zoej, M.J.; Duin, R.P.W. Dimensionality Reduction of Hyperspectral Data via Spectral Feature Extraction. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2091–2105. [[CrossRef](#)]
52. Yao, X.; Tian, Y.; Liu, X.; Cao, W.; Zhu, Y. Comparative Study on Monitoring Canopy Leaf Nitrogen Status on Red Edge Position with Different Algorithms in Wheat. *Sci. Agric. Sin.* **2010**, *43*, 2661–2667.

53. Wang, Y.; Li, F.; Wang, W.; Chen, X.; Chang, Q. Hyper-Spectral Remote Sensing Stimulation of Shoot Biomass of Winter Wheat Based on SPA and Transformation Spectra. *J. Trop. Crop.* **2020**, *40*, 1389–1398.
54. Yang, Y.; Nan, R.; Mi, T.; Song, Y.; Shi, F.; Liu, X.; Wang, Y.; Sun, F.; Xi, Y.; Zhang, C. Rapid and Nondestructive Evaluation of Wheat Chlorophyll under Drought Stress Using Hyperspectral Imaging. *Int. J. Mol. Sci.* **2023**, *24*, 5825. [[CrossRef](#)] [[PubMed](#)]
55. Cao, Y.; Xu, H.; Song, J.; Yang, Y.; Hu, X.; Wiyao, K.T.; Zhai, Z. Applying Spectral Fractal Dimension Index to Predict the SPAD Value of Rice Leaves under Bacterial Blight Disease Stress. *Plant Methods* **2022**, *18*, 67. [[CrossRef](#)]
56. Zhang, A.; Li, M.; Shi, J.; Pang, H. Hyperspectral Inversion Method for Natural Grassland Canopy SPAD Value Based on Scaling Up of Green Coverage Rate. *Spectrosc. Spect. Anal.* **2024**, *44*, 3513–3523.
57. Adam, E.; Mutanga, O. Spectral Discrimination of Papyrus Vegetation (*Cyperus Papyrus* L.) in Swamp Wetlands Using Field Spectrometry. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 612–620. [[CrossRef](#)]
58. Fernandes, M.R.; Aguiar, F.C.; Silva, J.M.N.; Ferreira, M.T.; Pereira, J.M.C. Spectral Discrimination of Giant Reed (*Arundo Donax* L.): A Seasonal Study in Riparian Areas. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 80–90. [[CrossRef](#)]
59. Ji, T.; Wang, B.; Yang, J.; Li, Q.; Liu, Z.; Guan, W.; He, G.; Pan, D.; Liu, X. Construction of Chlorophyll Hyperspectral Inverse Model of Alpine Grassland Community in Eastern Qilian Mountains. *Grassl. Turf.* **2021**, *41*, 25–33. [[CrossRef](#)]
60. Amiri, R.; Beringer, J.; Isaac, P. Narrowband Spectral Indices for the Estimation of Chlorophyll along a Precipitation Gradient. In Proceedings of the 2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lisbon, Portugal, 6–9 June 2011; pp. 1–4.
61. Yang, H.; Hu, Y.; Zheng, Z.; Qiao, Y.; Zhang, K.; Guo, T.; Chen, J. Estimation of Potato Chlorophyll Content from UAV Multispectral Images with Stacking Ensemble Algorithm. *Agronomy* **2022**, *12*, 2318. [[CrossRef](#)]
62. Chen, Z.; Zhu, Z.; Sun, S.; Wang, Q.; Su, T.; Fu, Y. Estimation of Daily Evapotranspiration and Crop Coefficient of Maize under Mulched Drip Irrigation by Stacking Ensemble Learning Model. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 95–104. [[CrossRef](#)]
63. Padarian, J.; Minasny, B.; McBratney, A.B. Using Deep Learning to Predict Soil Properties from Regional Spectral Data. *Geoderma Reg.* **2019**, *16*, e00198. [[CrossRef](#)]
64. Furbank, R.T.; Silva-Perez, V.; Evans, J.R.; Condon, A.G.; Estavillo, G.M.; He, W.; Newman, S.; Poiré, R.; Hall, A.; He, Z. Wheat Physiology Predictor: Predicting Physiological Traits in Wheat from Hyperspectral Reflectance Measurements Using Deep Learning. *Plant Methods* **2021**, *17*, 108. [[CrossRef](#)] [[PubMed](#)]
65. Shi, X.; Wu, Y.; Tang, L.; Huang, X. Application of Neural Network Model with Partial Least-Square Regression in Prediction of Peak Velocity of Blasting Vibration. *Shock. Vib.* **2013**, *32*, 45–49. [[CrossRef](#)]
66. Dietterich, T.G. Ensemble Methods in Machine Learning. In *Multiple Classifier Systems*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 1–15.

**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.