forests

MDPI

Article

Mechanical Property Prediction of *Larix gmelinii* Wood Based on Vis-Near-Infrared Spectroscopy

Chunxu Li, Yaoxiang Li *, Yanzheng Zhao, Zheyu Zhang and Zichun Wang

College of Engineering and Technology, Northeast Forestry University, Harbin 150040, China
* Correspondence: yaoxiangli@nefu.edu.cn; Tel.: +86-4518-219-1115

Abstract: *Larix gmelinii* is the major tree species in Northeast China. The wood properties of different *Larix gmelinii* are quite different and under strong genetic controls, so it can be better improved through oriented breeding. In order to detect the longitudinal compressive strength (LCS), modulus of rupture (MOR) and modulus of elasticity (MOE) in real-time, fast and non-destructively, a prediction model of wood mechanical properties with high precision and stability is constructed based on visible-near-infrared spectroscopy (Vis-NIRS) technology. The featured wavelengths were selected with the algorithms of competitive adaptive reweighted sampling (CARS), successive projection algorithm (SPA), uninformative variable elimination (UVE), synergy interval partial least squares (SiPLS) and their combinations. The prediction models were then developed based on the partial least square regression (PLSR). The predictive ability of models was evaluated with coefficient of determination (R^2) and root mean square error (RMSE). It indicated that CARS performed the best among the four methods examined in terms of wavelength-variable selection. The combined featured wavelength selecting method of SiPLS-CARS showed better performance than the single wavelength selection method. The optimal models of LCS, MOR and MOE are the SiPLS-CARS-PLSR model, with the R^2 of the calibration set and the validation set are both greater than 0.99, and RMSE the smallest. The NIR optimal models for wood mechanical properties predictions has high predictive accuracy and good robustness.

Keywords: near-infrared spectroscopy; wood mechanical property; wavelength selection; competitive adaptive reweighted sampling; synergy interval partial least square

check for updates

Citation: Li, C.; Li, Y.; Zhao, Y.; Zhang, Z.; Wang, Z. Mechanical Property Prediction of *Larix gmelinii* Wood Based on Vis-Near-Infrared Spectroscopy. *Forests* **2022**, *13*, 1995. https://doi.org/10.3390/f13121995

Academic Editor: Byung-Dae Park

Received: 31 October 2022
Accepted: 23 November 2022
Published: 25 November 2022

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

Copyright: © 2022 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

As a deciduous tree belonging to the genus Larix in the Pinaceae family, *Larix gmelinii* is the main tree species in Northeast China [1]. Due to its good mechanical properties, clear texture and strong decay resistance, *Larix gmelinii* has been widely used in civil construction, railway sleepers, textile machinery, mining columns and piles and other fields [2,3]. Although the wood properties of different *Larix gmelinii* are quite different and under strong genetic controls [4], the overall wood properties of timber forests can be well improved through oriented breeding [5]. The mechanical properties of wood are not only measures of the ability to resist external force, but also the main material property indexes of structural timber.

Longitudinal compressive strength (LCS), defined as the maximum capacity of the wood to bear the compressive load parallel to grain, is mainly used to calculate and select the allowable acting stress on the compression members of structural and building materials. The modulus of rupture (MOR), which refers to the maximum capacity of wood to withstand gradually added bending loads, is the measure of the ability of wood to bear lateral loads and is mainly applied to beams of various cabinets in furniture, trusses of buildings, floors and bridges and other easy-to-bend members. The modulus of elasticity (MOE) is the ratio of stress to strain within the proportional limit when the wood is bent under force, which represents the rigidity or elasticity of wood, and the ability to resist

bending or deformation [6–9]. The MOE of wood is used to calculate the deformation of beams or trusses under bending loads and to calculate safe loads. Evaluating the physical and mechanical properties of the wood of *Larix gmelinii* can provide technical supports for the selection of *Larix gmelinii* plantation and efficient use of the wood [8].

Traditional wood mechanics testing is destructive and not suitable for rapid testing in online production due to lengthy testing times, harsh conditions, poor stability and reproducibility. As a green, fast and non-destructive testing method, visible-near-infrared spectroscopy (Vis-NIRS) has been widely used in agriculture, petrochemicals, life science and other fields [10,11]. Due to its advantages of a simple operational process, accurate analysis, strong model generalization ability, and real-time online monitoring, Vis-NIRS can meet the needs of rapid and accurate determination of large numbers of samples in an area [12]. The analytical information carried in the near-infrared spectral region is mainly the frequency multiplication and combination characteristic information from the vibrations of hydrogen-containing groups in the molecule [13,14]. Therefore, Vis-NIRS testing can obtain the wood mechanical properties without damaging the wood and its original shape, structure and state. The prediction of wood mechanical properties by Vis-NIRS has become one of the commonly used and promising technical means for wood quality assessment and elite variety breeding of *Larix gmelinii.*

Vis-NIRS has the shortcomings of serious band overlap and weak effective signal [15]. Researchers have conducted a lot of research on the near-infrared prediction modeling by combining traditional mechanical property detection methods with multivariate analysis methods. The results demonstrated that the existence of abnormal samples can reduce the correlation between the spectrum and the true value and the prediction accuracy of the model. Therefore, it is necessary to discriminate and remove abnormal samples [16,17]. The methods for dividing the sample set and preprocessing spectral data have great influence on the modeling results. Yuan et al. [18–20] compared the influences of four training set sample-selection methods on the model, that is, random sampling, Kennard-Stone, duplex, and sample-set partitioning based on joint x–y distance (SPXY). The results showed that the model established by the training set selected by the SPXY method is better than the other three methods. Horvath and Fujimoto [21,22] et al. built prediction models of wood compressive strength, flexural strength and elastic modulus with good accuracy based on NIR spectroscopy technology. Leblon B [23] et al. made a comprehensive summary of the near-infrared band of wood, as well as some practical considerations important for the application of NIR to wood. Thumm [24] et al. used NIR spectroscopy to predict the MOE of *Pinus radiata* jointless wood, and the prediction effect in the NIR longwave region was slightly better than in the full-spectrum region. Guo et al. [25] applied four wavelength-variable selection algorithms, that is, synergy interval (SI), successive projection algorithm (SPA), genetic algorithm (GA) and competitive adaptive reweighted sampling (CARS), to reduce the dimension of apple-soluble solids spectral data. The results showed that all methods could simplify the partial least square regression (PLSR) model to a certain extent, while the correlation coefficients were all above 0.9. Among the four algorithms, CARS has the best optimization performance.

This study aims to model and predict the mechanical properties of larch wood based on Vis-NIRS. In this experiment, choosing the *Larix gmelinii* in Xinlin Forest Farm as the research object, the Vis-NIRS was used to model and predict the LCS, MOR and MOE in the transverse, radial and tangential sections of *Larix gmelinii*. Different wavelength-variable extraction algorithms were used to select the characteristic wavelength variables of the full-band spectrum. Taking the optimal characteristic wavelength obtained after characteristic variable selecting as the input variable, a prediction model for mechanical properties of *Larix gmelinii* was developed. The prediction model can provide technical supports for the real-time and rapid determination of the mechanical properties of *Larix gmelinii* wood and a theoretical basis for the elite variety breeding and forest management of *Larix gmelinii* [26].

2. Material and Methods

2.1. Wood Sample Collection

The sampling site is located at Xinlin Forest Farm, Greater Khingan Mountains, Northeast China. The altitude is relatively high and the average altitude is about 600 m, the terrain is gentle and the slope is mostly below 6°. The main forest type in the sampling site is natural secondary forest, and the dominant species are *Larix gmelinii* and *Betula platyphylla*. In this study, the *Larix gmelinii* wood samples were collected from five typical sample plots (30 m × 30 m) with similar geographical conditions in the experimental forest farm. In each plot, one *Larix gmelinii* average stem was selected as the standard wood. According to the "Method of Sample Tree Collection for Physical and Mechanical Tests of Wood" (GB/T 1936-2009), 81 test pieces of 30 mm × 20 mm × 20 mm longitudinal compressive strength were made, and 99 samples of bending strength and bending elastic modulus of 300 mm × 20 mm × 20 mm were made. Before the test, the samples were placed in a constant temperature and humidity chamber with the moisture content adjusted to 12%.

The mechanical measuring instrument used in this study is the GMT-6305 electromechanical universal testing machine of Zhuhai Sansi Measuring Instrument Company. The compressive strength of wood samples was measured with reference to "Method of testing in compressive strength parallel to grain of wood" (GB/T 1936-2009). The flexural strength and flexural modulus of elasticity were measured with reference to the "Method for determination of the Modulus of Elasticity of Wood" (GB/T 1936-2009). Using IBM SPSS 26.0 for statistical analysis, the wood mechanical properties of *Larix gmelinii* are shown in Figure 1.

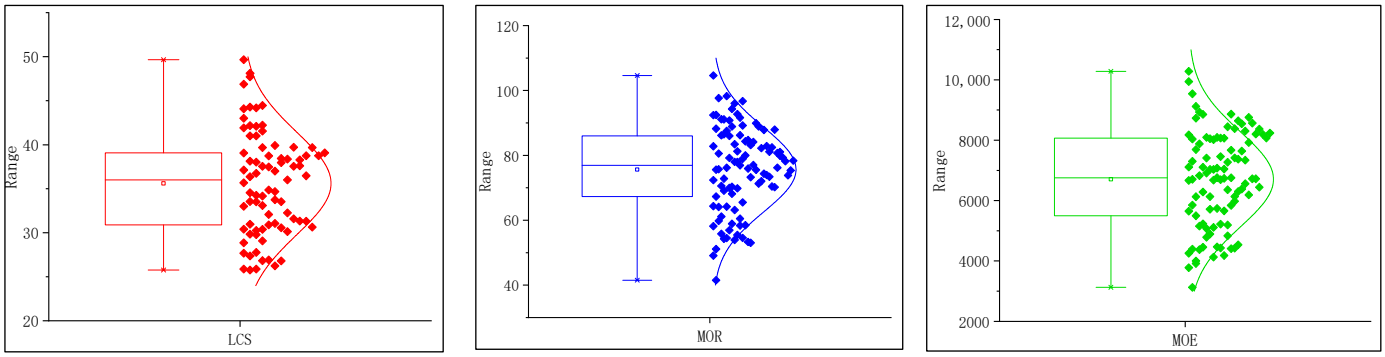

Figure 1. Ranges of mechanical properties of larch wood.

2.2. Visible-Near-Infrared Spectra Collection

The visible-near-infrared spectra were collected using a LabSpec Pro FR/A114260 (Analytical Spectral Devices, Inc., Boulder, CO, USA). The wavelengths ranged from 350 to 2500 nm, with a minimum interval range of 2 nm, and a total of 2151 wavelength points were collected. The sampling time was 10 times/s. The sampling points for NIRS were the center points of wood transverse, radial and tangential section, as shown in Figure 2. In order to eliminate the noise signal and improve the accuracy of spectral acquisition, the number of scans was increased to 30 and the average value of the repeated measurement spectra was taken as the original spectra, and the spectral signal-to-noise ratio was subsequently improved. Then raw spectra were imported into the spectral data processing software ViewSpecPro (Caliper Corporation, Newton, MA, USA) to realize the conversion of diffuse reflectance into absorbance values according to the Kubelka-Munk theory.

2.3. Modeling and Optimization

In this study, MATLAB R2016a (MathWorks, Natick, MA, USA) was used for the Near-Infrared model development and optimization.

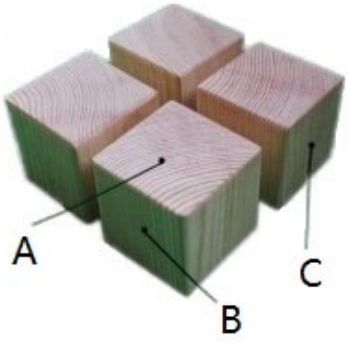

Figure 2. Schematic of different scanning section. A—Transverse Section, B—Radial Section, C—Tangential Section.

The Monte Carlo cross-validation (MCCV) is a method of identifying singular samples by using the statistical laws of singular samples in Monte Carlo cross-validation, which can be used to solve complex statistical models and matrix high-dimensional problems. Compared to traditional methods, it can identify the singular samples better. The Monte Carlo sampling method takes 80% of the samples as the calibration set to establish the partial least squares regression (PLSR) model, and the remaining part as the prediction set. By repeating the cycle 1000 times, a set of predicted residuals (PRESS) for each sample was obtained. If a sample has a large PRESS in the PLSR model and the frequency of occurrence is high, it is obviously deviated and judged as an abnormal sample [27–30].

Wood samples were partitioned into a calibration set and validation set based on joint x–y distances (SPXY).The SPXY method is derived from the classical sample classification algorithm, the Kennard-Stone method, which first selects the two samples with the farthest Euclidean distances into the calibration set, then selects the samples with the largest and smallest distances into the calibration set by calculating the Euclidean distances from each remaining sample to each known sample in the calibration set, and so on, until the number of samples in the calibration set reaches the specified number [31]. SPXY is similar to the selection process of the Kennard-Stone method, with the difference that SPXY integrates the spectral response matrix and concentration column vectors of the samples to achieve effective coverage of the multidimensional vector space, increase the variability and representativeness of the samples, and improve the stability of the established model [32]. The distance calculation equation of SPXY dividing the sample set is as follows:

$$d_x(p,q) = \sqrt{\sum_{j=1}^{N}\left[x_p(j) - x_q(j)\right]^2} \quad (1)$$

$$d_y(p,q) = \sqrt{\left(y_p - y_q\right)^2} = \left|y_p - y_q\right| \quad (2)$$

$$d_{xy}(p,q) = \frac{d_x(p,q)}{\max\limits_{p,q\in[1,N]} d_x(p,q)} + \frac{d_y(p,q)}{\max\limits_{p,q\in[1,N]} d_y(p,q)} \quad (3)$$

where $p, q \in [1, N]$. X_p and X_q denote two different samples and N is the number of wavelength points of the samples.

2.3.1. Spectral Preprocessing and Band Optimization

Vis-NIRS has the characteristics of high dimension and redundancy. In order to find the weak interference information that can characterize the samples and integrate multivariate information to eliminate the strong interference information, wavelength-variable selection becomes a key issue in the Vis-NIRS application to wood mechanical property analysis. Wavelength-variable selection can be divided into two types, that is, point screening and band screening [33]. This study explores the effects of three wavelength

point screening methods, CARS, SPA and UVE, and the band screening method SiPLS on model quality improvement, and it also explores the screening effect of four combined strategies: UVE-CARS, SiPLS-UVE, SiPLS-CARS and SiPLS-SPA.

CARS based on the principle of 'survival of the fittest' in evolutionary theory had the characteristics of fast computation and a high screening efficiency [34]. This study set the number of Monte Carlo samples to 100 and selects the subset with the smallest RMSE by 10-fold transverse-validation to find the set of optimal solution variables [35].

The SPA was capable of extracting the few characteristic wavelengths in the full waveband that contain the least redundant information and the least covariance, thus eliminating a larger number of irrelevant variables in the original spectral matrix, and was widely used in the field of high-precision analysis of Vis-NIRS [36]. For the original eigenspectral matrix $X_{m\times n}$, where m was the number of samples and n the number of wavelengths, $x_n(0)$ and K were the initial iteration vectors and the number of wavelengths to be extracted, respectively. SPA was a forward cyclic variable selection method that minimizes vector space covariance [37].

According to the relationship between the spectral matrix and the concentration matrix, the UVE method was used to select characteristic wavelengths based on the regression coefficients of the PLSR model [38]. Based on the PLSR model, the randomly generated matrix was combined with the noise matrix with the same number of dependent variables to form a new matrix. By excluding the C_i, which is in the interval of [1, n], with the value of C_i less than C_{max}, the final extracted variables were obtained. The specific formula is as follows:

$$C_i = \frac{mean(b_i)}{std(b_i)},\ (i = 1, 2, \cdots 2n) \tag{4}$$

SiPLS was an extension and improvement of the interval partial least squares (IPLS) method. With IPLS method, the full-band data was split into several equal-width subintervals. Then a PLSR model was built for each subinterval separately, and the optimal modeling subinterval was selected by taking RMSE as the model indicator [39,40]. However, IPLS was sensitive to the interval width, and the selected subinterval may not be exactly the information interval. SiPLS overcomes the drawback of IPLS single-interval modeling by calculating all possible PLSR combination models for two or more subintervals in the same interval division based on the idea of permutation and combination, and the optimal joint subinterval was used for PLSR. In this study, the number of intervals was 20, and the number of combined subintervals was 2, 3 or 4 [41–43].

2.3.2. Modeling Methods and Model Evaluation

PLSR combines the advantages of principal component analysis (PCA), classical correlation analysis (CCA) and multiple linear regression (MLR) to achieve high-dimensional data structure simplification, regression modeling, and analysis of multiple correlations of independent variables [44]. A PLSR model was established by using Vis-NIRS combined with the multivariate selection method to quantitatively analyze the mechanical properties of *Larix gmelinii* in Xinlin. Coefficient of determination (R^2) and root mean square error (RMSE) were selected for the model evaluation. The closer R^2 is to 1 in the range of 0 to 1, the better the stability of the model and the higher the fitting degree [45]. RMSE was used to detect the deviation between the estimated value and the true value. The smaller the RMSE, the better the prediction ability of the model [46]. The computations for these criteria are as follows [46]:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{5}$$

$$RMSE = \sqrt{\sum_{i=1}^{m} \frac{(\hat{y}_i - y_i)^2}{m}} \tag{6}$$

3. Results

Figure 3 shows the NIR spectra absorbance of three sections of *Larix gmelinii* wood mechanical test pieces. Although the absorbance varying trends of three sections over the full-band spectrum were consistent, the value of absorbance among the three sections were obviously different. The transverse section had the strongest spectra absorption capacity, while the radial and tangential sections had relatively weak and similar spectral absorbance. However, there was no significant difference between the spectra of different mechanical test pieces. The strongest spectra absorption capacity of the wood transverse section of *Larix gmelinii* was related to the arrangement of the cells in the wood and the growth characteristics of the wood. Therefore, different sections of wood had an influence on the construction of NIR spectra. The NIR spectra of *Larix gmelinii* wood samples had strong absorption peaks at 1431 nm, 1857 nm and 2056 nm. Wood is a complex naturally polymeric organic material, and the hydrogen-containing groups in the organic material led to strong absorption peaks in this wavelength region [12], which indicates a close relationship between the mechanical properties of *Larix gmelinii* wood and the NIR spectra. Therefore, the NIR spectra can be used to construct models to predict the mechanical properties of wood.

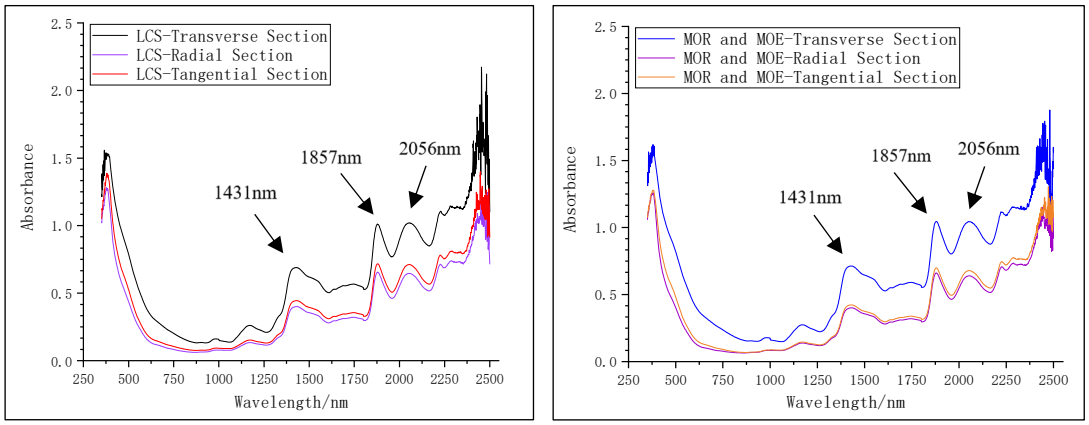

Figure 3. Three sections of spectra of mechanical testing pieces.

3.1. Screening of Singular Samples

The PRESS distribution of the MCCV of the wood samples is shown in Figure 4. As shown in Figure 4A, four singular samples of LCS in the transverse section of test pieces were identified, which were samples with labels 21, 34, 49 and 70. Three singular samples labelled with 69, 70 and 75 in the radial section and four labelled with 51, 69, 70 and 75 in the tangential section were identified in Figure 4B,C, respectively. The number of singular samples of MOR in transverse section is eight, that is, samples with labels 17, 31, 54, 57, 79, 80, 85 and 98, as shown in Figure 4D. Nine singular samples labelled with 4, 5, 9, 43, 50, 58, 65, 79, and 87 in the radial section and thirteen labelled with 17, 22, 23, 50, 57, 61, 62, 64, 67, 74, 87, 97, and 98 in the tangential section were identified in Figure 4E,F, respectively. There were three singular samples of modulus elasticity labelled with 54, 66 and 69 in the transverse section, six labelled with 3, 7, 50, 66, 78 and 79 in the radial section, and eight labelled with 42, 44, 50, 54, 59, 60, 66 and 69 in the tangential section, as shown in Figure 4G–I, respectively.

In order to verify the identifying results of singular samples, PLSR modeling was carried out on the selected samples, with the results outlined in Table 1. It can be seen that the R^2 of PLSR models with different mechanical properties in different sections were all improved. The R^2 of LCS in three sections were increased by 0.0716, 0098 and 0.0613, respectively. The R^2 of MOR were increased by 0.088, 0.1453 and 0.1047, and MOE increased by 0.1872, 0.1378 and 0.0698, respectively. The results show that the existence of singular samples had a significant influence on the quality of the PLSR model. There

are many reasons for the appearance of singular samples, such as instrumental errors and experimental operation errors. Therefore, the identification of singular samples has always been one of the hot spots in multivariate correction research [47].

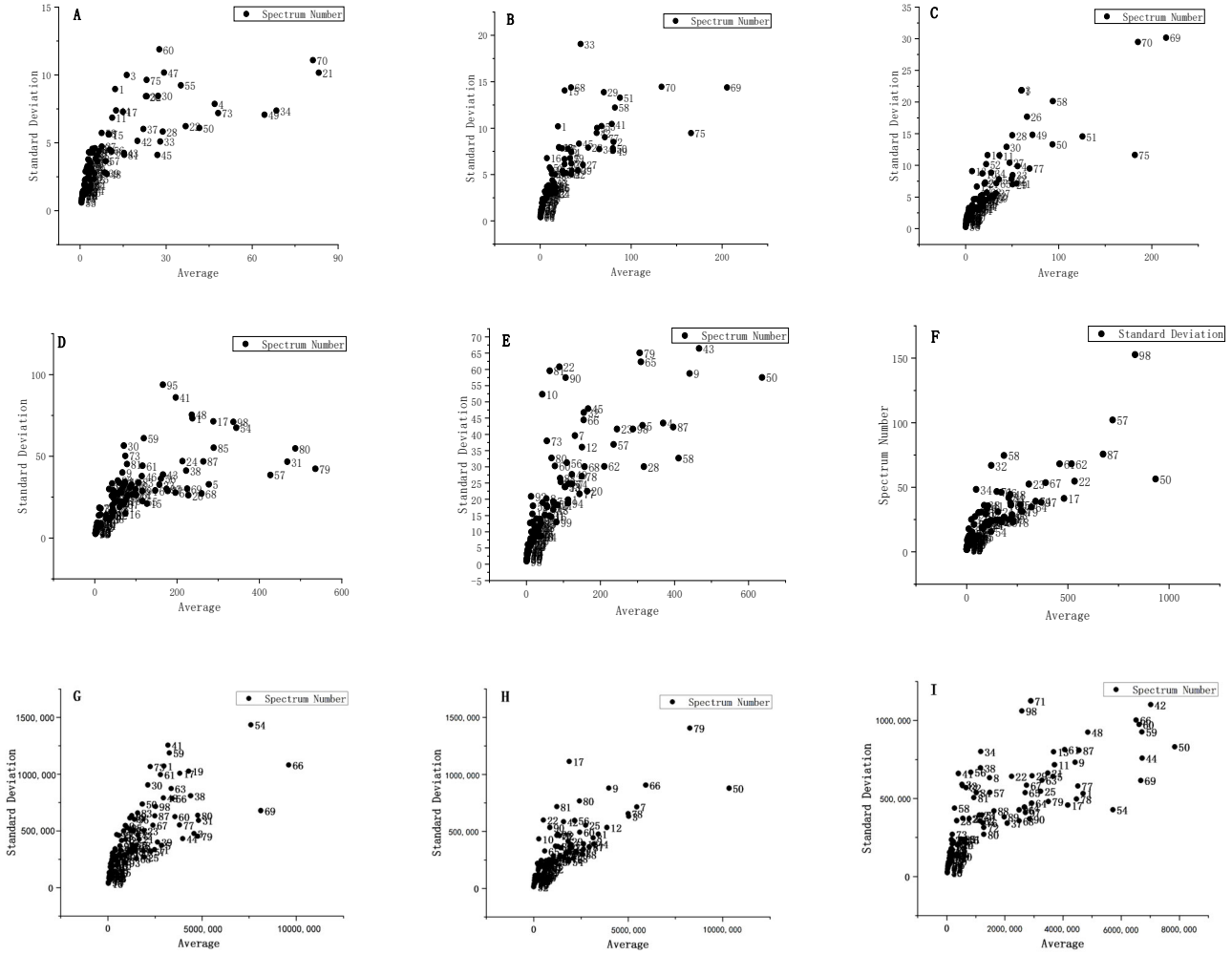

Figure 4. PRESS distributions of MCCV of different sections of *Larix gmelinii*. (**A**) LCS of transverse section, (**B**) LCS of radial section, (**C**) LCS of tangential section, (**D**) MOR of transverse section, (**E**) MOR of radial section, (**F**) MOR of tangential section, (**G**) MOE of transverse section, (**H**) MOE of radial section, (**I**) MOE of tangential section.

Table 1. MCCV screening of wood mechanical properties.

Mechanical Property	Section	Before Screening			After Screening		
		Sample Number	RMSE	R^2	Sample Number	RMSE	R^2
LCS	Trans Section	81	3.8035	0.5690	77	3.3119	0.6406
	Rad Section	81	5.5474	0.0831	78	5.0671	0.0929
	Tang Section	81	5.5111	0.0950	77	4.7498	0.1563
MOR	Trans Section	99	10.4792	0.3535	91	9.1556	0.4415
	Rad Section	99	9.9123	0.4215	90	8.0021	0.5668
	Tang Section	99	11.9027	0.1659	86	9.2123	0.2706
MOE	Trans Section	99	1293.5190	0.3260	96	1066.2353	0.5132
	Rad Section	99	1188.6372	0.4308	93	971.3656	0.5686
	Tang Section	99	1357.9858	0.2571	91	1177.7148	0.3269

3.2. Partition of Sample Sets

In this study, the sample-set dividing was conducted by SPXY. With the dividing ratio of 2:1, the statistics of mechanical properties of wood corresponding to different spectral sample sets are shown in Table 2.

Table 2. Statistics of wood mechanical property.

Mechanical Property	Section	Sample Grouping	Number of Samples	Minimum (Mpa)	Maximum (Mpa)	Average (Mpa)	Standard Deviation (%)
LCS	Trans Section	cal	51	25.76	49.65	36.25	5.84
		val	26	26.8	44.19	34.31	4.82
	Rad Section	cal	52	25.76	46.87	35.38	5.44
		val	26	25.87	44.19	34.57	5.22
	Tang Section	cal	51	25.87	44.27	35.35	4.89
		val	26	25.76	44.46	34.18	5.80
MOR	Trans Section	cal	60	54.24	104.61	78.94	11.74
		val	31	53.01	92.7	71.61	11.62
	Rad Section	cal	60	41.5	95.99	74.89	12.87
		val	30	53.01	97.65	74.01	11.00
	Tang Section	cal	57	53.92	98.25	76.51	11.21
		val	29	53.31	92.39	75.55	10.27
MOE	Trans Section	cal	64	3124.91	9942.79	6865.20	1558.43
		val	32	4128.55	8642.89	6145.76	1393.05
	Rad Section	cal	62	3124.91	9536.82	6527.45	1477.71
		val	31	4255.43	8931.31	7077.10	1460.16
	Tang Section	cal	60	3777.54	9536.82	6762.04	1483.91
		val	31	4181.45	8544.13	6563.14	1375.92

3.3. Wavelength Selection

In order to improve the model accuracy of wood mechanical properties NIR spectroscopy and reduce the difficulty of modeling, the study explored the effect of eight methods of feature variable selection: UVE, CARS, SPA, SiPLS, UVE-CARS, SiPLS-UVE, SiPLS-CARS and SiPLS-SPA.

The effect of UVE screening is shown in Figure 5, where the left side of the vertical dividing line shows the original wavelength variables, the right side of the red area shows the randomly introduced noise variables, and the upper and lower solid lines parallel to the horizontal axis are the threshold lines for assessing the stability. From the principle of UVE, it can be seen that the wavelength points between the threshold lines do not contribute to the modeling and are uninformative variables. Moreover, the parts marked with green "*" beyond the threshold are the feature variables containing modeling information. As shown in Table 3, the optimal UVE-PLS models of the three sections with different mechanical properties were the radial section models. The accuracy of the UVE-PLS models of LCS and MOE was less than that of the optimal full-spectrum model, and only the radial section model of MOR was better than the full-spectrum PLSR model; 129 and 660 wavelength variables were screened, with Rc^2 = 0.6895 and Rv^2 = 0.8123, RMSEc = 7.1141 and RMSEv = 4.6872.

SPA performed characteristic wavelength extraction on the near-infrared spectrum of wood mechanics, and the results are shown in Figure 6, where the red quadrilateral is marked with the location of the characteristic wavelength. The characteristic wavelength extracted by SPA from the NIR spectrum of wood mechanical properties has a high coincidence degree, and the variables are mainly concentrated in 2000–2500.The SPA-PLS results are shown in Table 3. The SPA had poor screening results for wood mechanics NIR spectra. The optimal SPA-PLS model of LCS was the transverse section model, but its model accuracy was less than the full-band spectral model. Eleven characteristic wavelengths were screened. The R^2 of calibration set and validation set decreased by 0.3643 and 0.0704, and RMSE was increased by 1.5956 and 0.3171. Among the SPA-PLS models

with different sections of MOR, only the radial section model had higher accuracy than the corresponding full-band spectral model. SPA obtained 10 characteristic wavelengths; the R^2 of calibration set and validation set were 0.4787 and 0.7635, and RMSE were 9.2175 and 5.2613, respectively. The accuracy of the SPA-PLS model in different sections of MOE was lower than the full-band spectral model, and the radial section SPA-PLS model was the best. SPA selects 10 characteristic wavelengths of radial section to construct PLSR model, with Rc^2 = 0.5598 and Rv^2 = 0.8394, RMSEc = 972.4311 and RMSEv = 575.7233.

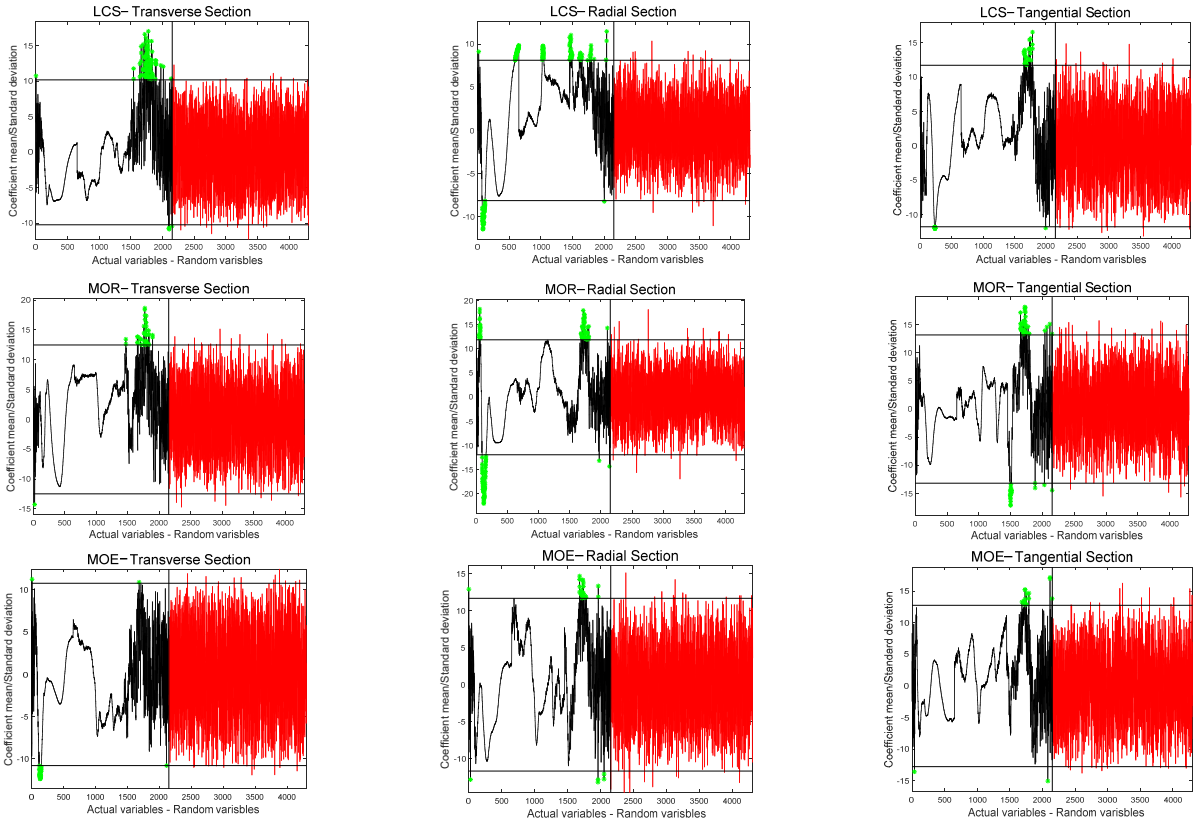

Figure 5. Band screening results of wood mechanical properties Vis-NIRS based on UVE.

CARS optimized the band of the near-infrared spectrum of wood mechanical properties, and the results are shown in Figure 7. The optimal sampling times of wood mechanics were all less than 70 times. When the sampling times increase from 0 to 70, RMSECV fluctuates with the increase of sampling times, but the overall trend shows a downward trend. CARS selected characteristic variables to build the PLSR model, and the results are shown in Table 3. The performance of the CARS-PLS model in different sections of LCS was better than that of the full-band model. The transverse section CARS-PLS model was the best; the optimal sampling times were 55 and 48, and the characteristic variables of calibration set and validation set were 48 and 78, R^2 was 0.9149 and 0.9805, and RMSE was 1.6866 and 0.6599, respectively. The accuracy of the transverse section CARS-PLS model of MOR was lower than that of the full-band spectral model, and the performance of radial and tangential sectional CARS-PLS models was improved. The radial-section CARS-PLS model was the best, and the optimal sampling times were 39 and 51. The calibration set and the validation set selected 148 and 63 characteristic wavelengths, accounting for 6.8805% and 2.9289% of the global variables; Rc^2 = 0.9832 and Rv^2 = 0.9670, RMSEc = 1.6556 and RMSEv = 1.9667. The radial section CARS-PLS model of MOE had the best performance, high model accuracy and good stability. The optimal sampling times of the calibration set and the validation set were 52 and 63, and the selected characteristic wavelengths account for 2.7429% and 1.353482% of the global variables, Rc^2 = 0.9680 and Rv^2 = 0.9545, RMSEc = 262.2065 and RMSEv = 306.3651.

Table 3. PLSR modeling results of different variable selection methods.

	Wavelength Selection	Trans Section				Rad Section				Tang Section			
		Calibration Models		Validation Models		Calibration Models		Validation Models		Calibration Models		Validation Models	
		RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
LCS	–	3.0187	0.7274	2.3166	0.7596	5.0737	0.1119	4.8146	0.1140	4.2422	0.2315	4.0989	0.4807
	UVE	4.3095	0.4444	2.8919	0.6253	3.9129	0.4718	2.1699	0.8200	4.4751	0.1448	3.2692	0.6696
	SPA	4.6143	0.3631	2.6337	0.6892	5.1146	0.0975	3.7034	0.4758	4.1448	0.2664	3.3530	0.6525
	CARS	1.6866	0.9149	0.6599	0.9805	2.3512	0.8093	3.4580	0.5430	2.6530	0.6994	1.5944	0.9214
	SiPLS	4.2458	0.4607	3.0068	0.5949	0.4484	0.9931	1.8195	0.8735	1.4138	0.9146	3.6286	0.5930
	UVE-CARS	3.6219	0.6076	3.3401	0.5002	4.8225	0.1977	3.1118	0.6299	4.1233	0.2740	3.1086	0.7013
	SiPLS-UVE	3.8154	0.5645	3.3771	0.4890	3.9434	0.4635	1.5985	0.9023	3.4262	0.4987	3.8762	0.5355
	SiPLS-SPA	3.2541	0.6832	3.0000	0.5968	2.7740	0.7345	3.2727	0.5906	3.7434	0.4016	2.7159	0.7720
	SiPLS-CARS	0.7017	0.9853	0.6771	0.9795	0.2242	0.9983	0.2010	0.9985	0.4762	0.9903	0.6005	0.9889
MOR	–	7.2867	0.6086	7.1487	0.6087	8.1888	0.5885	5.7562	0.7170	9.1872	0.3161	8.6206	0.2703
	UVE	10.4640	0.1928	7.8751	0.5251	7.1141	0.6895	4.6872	0.8123	7.2857	0.5699	6.9493	0.5258
	SPA	10.9926	0.1092	9.3522	0.3302	9.2175	0.4787	5.2613	0.7635	9.1262	0.3251	8.4597	0.2973
	CARS	8.1744	0.5074	1.3085	0.9869	1.6556	0.9832	1.9667	0.9670	5.0289	0.7951	4.6230	0.7901
	SiPLS	6.7776	0.6614	5.7152	0.7499	6.2629	0.7593	5.2525	0.7643	5.5616	0.7494	7.7502	0.4102
	UVE-CARS	8.7328	0.4378	7.8751	0.5251	5.0520	0.8434	1.7270	0.9745	7.9975	0.4817	6.9493	0.5258
	SiPLS-UVE	6.6482	0.6742	5.9283	0.7309	7.6109	0.6446	5.4087	0.7501	7.6510	0.5257	6.3259	0.6071
	SiPLS-SPA	7.7680	0.5552	8.5603	0.4389	7.6952	0.6367	5.0751	0.7800	8.1958	0.4557	7.4592	0.4537
	SiPLS-CARS	0.8714	0.9944	1.1315	0.9902	0.4873	0.9985	0.7267	0.9955	0.8082	0.9947	1.2694	0.9842
MOE	–	782.2642	0.7440	567.9018	0.8284	922.2126	0.6041	803.4178	0.6872	1060.6739	0.4804	945.1786	0.5124
	UVE	1173.5775	0.4239	1191.7883	0.2445	928.9296	0.5983	534.5773	0.8615	1112.7769	0.4281	861.5252	0.5949
	SPA	1374.5134	0.2098	917.1758	0.5525	972.4311	0.5598	575.7233	0.8394	1156.8858	0.3819	891.5219	0.5662
	CARS	760.5009	0.7581	474.2760	0.8803	262.2065	0.9680	306.3651	0.9545	717.4283	0.7623	176.4556	0.9830
	SiPLS	402.3331	0.9323	759.8211	0.6929	770.1811	0.7239	712.5736	0.7539	990.3318	0.5471	950.6278	0.5067
	UVE-CARS	1173.5775	0.4239	1191.7883	0.2445	928.9296	0.5983	534.5773	0.8615	1112.7769	0.4281	671.6445	0.7538
	SiPLS-UVE	951.5708	0.6213	668.1904	0.7625	864.6130	0.6520	800.3360	0.6896	1351.3077	0.1567	937.5018	0.5203
	SiPLS-SPA	1105.6073	0.4887	1154.8332	0.2906	730.6229	0.7515	778.6521	0.7061	1135.6491	0.4044	1052.0384	0.3959
	SiPLS-CARS	112.7340	0.9947	100.4788	0.9946	129.9646	0.9921	130.0499	0.9918	146.6291	0.9901	165.7741	0.9850

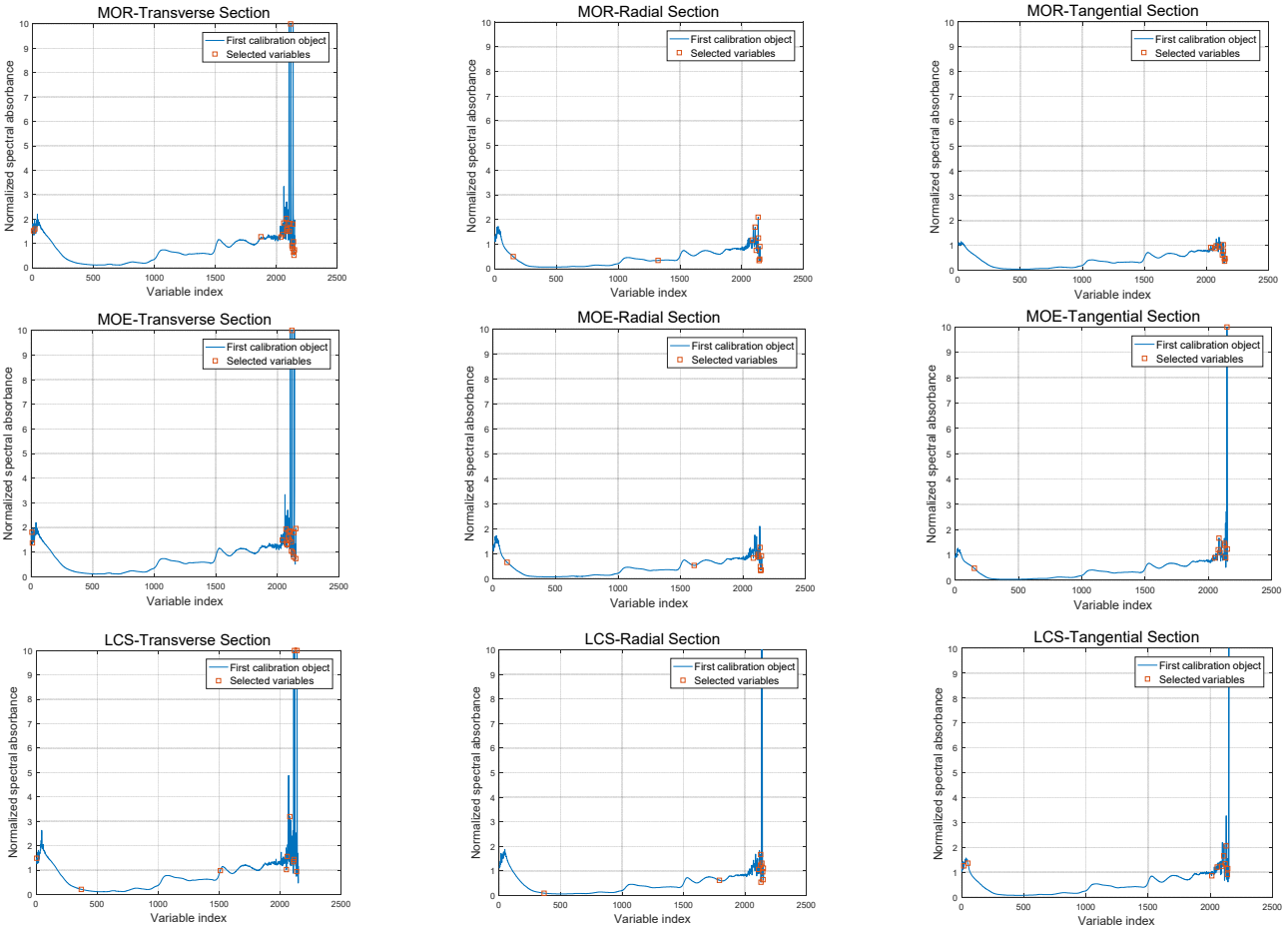

Figure 6. Wavelength extraction results of wood mechanical properties VIS-NIRS based on SPA.

The interval combination number of SiPLS were 2, 3 and 4, and the optimal band of the NIR spectrum in different sections of wood mechanical properties is shown in Figure 8. The PLSR model was built for the optimal near-infrared band of SiPLS, and the results are shown in Table 4. The number of combination intervals had a significant impact on the performance of the SiPLS-PLS model. The radial section NIR spectrum of LCS had the minimum RMSE when the interval number was 4, and the model performance was the best. The PLSR model was constructed with 431 wavelengths, Rc^2 = 0.9931 and Rv^2 = 0.8735, RMSEc = 0.4484 and RMSEv = 1.8195. The optimal SiPLS-PLS model of MOR was a radial section model with a combined interval of 2, the wavelength band was 675–890 nm, R^2 was 0.7593 and 0.7643, and RMSE was 6.2629 and 5.2525, respectively. When the combined interval number of MOE was 4, the SiPLS-PLS model of transverse section was optimal, accounting for 20.0372% of the global variables, and the calibration model Rc^2 = 0.9323, RMSEc = 402.3331; the validation model Rv^2 = 0.6929, RMSEv = 759.8211.

To compensate for the weakness of UVE in eliminating invalid variables, CARS was used to further refine the UVE. The accuracy of the UVE-CARS-PLS model of LCS was lower than that of the optimal all-band PLSR model, and the transverse sectional UVE-CARS-PLS was relatively better. The selected variables of the calibration set and the validation set were 6 and 88, R^2 were 0.6076 and 0.5002, and RMSE were 3.6219 and 3.3401, respectively. The radial section UVE-CARS-PLS model of MOR had the highest accuracy. The calibration set and validation set obtained 18 and 15 wavelengths, which were 13.9535% and 2.2727% of UVE screening results, respectively. R^2 increased to 0.8434 and 0.9745, and RMSE decreased to 5.0520 and 1.7270, respectively. UVE-CARS showed a good result in the radial section of the MOE NIR spectrum screening. The PLSR model was constructed

with 27 wavelengths, accounting for 1.2552% of the global variable. The calibration model Rc^2 = 0.5983, RMSEc = 928.9296; the validation model Rv^2 = 0.8615, RMSEv = 534.5773.

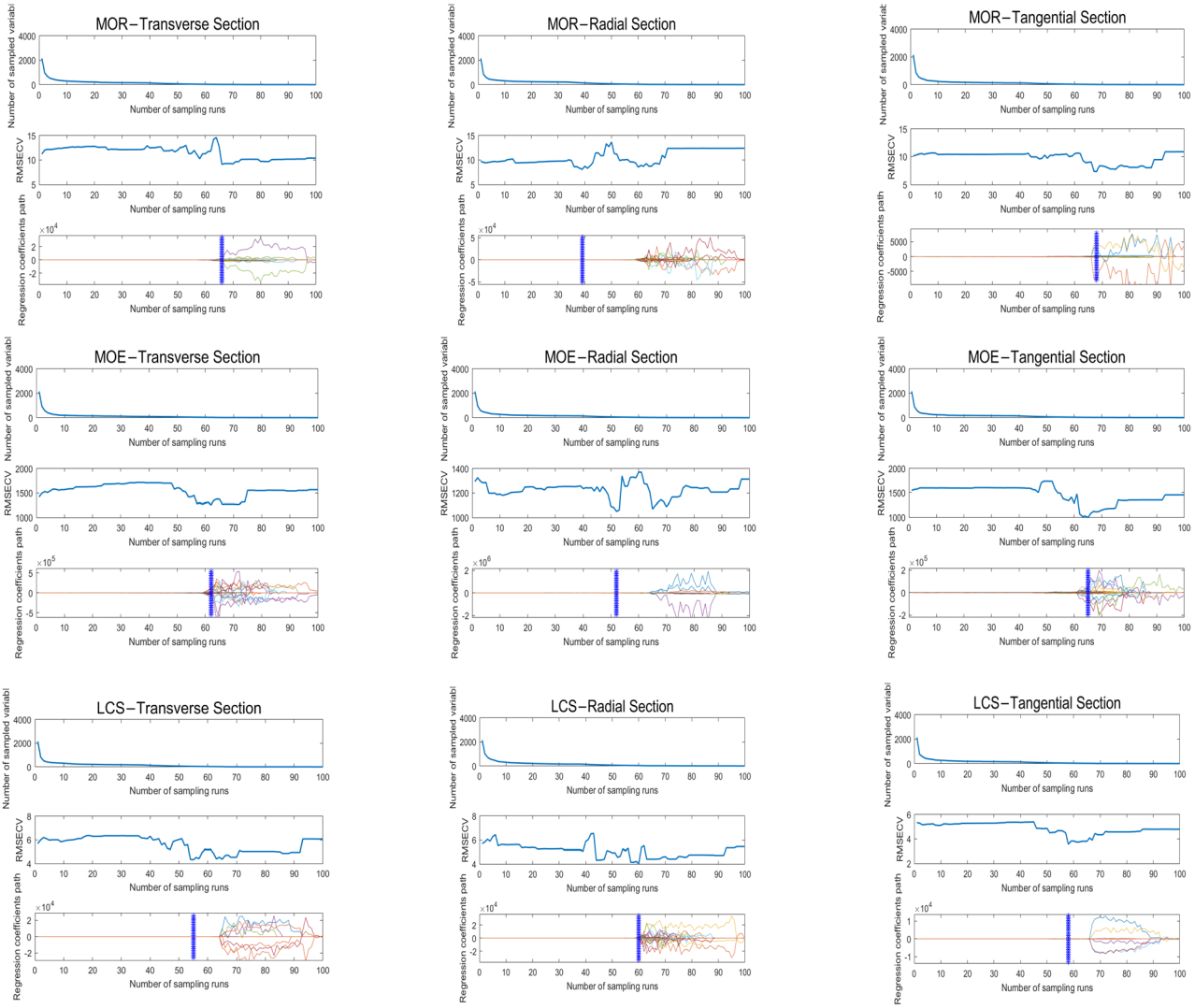

Figure 7. Band screening results of wood mechanical properties Vis-NIRS based on CARS.

SiPLS was used in combination with UVE, SPA and CARS to screen wavelength points in the optimal wavelength range obtained by SiPLS. The number of variables selected by SiPLS-UVE decreased compared with UVE. The SiPLS-UVE-PLS of LCS and MOE were smaller than the optimal full-band spectral model, and the transverse section and radial section of MOR were larger than the original spectral model. The SiPLS-UVE-PLS model of MOR had a high stability, with 22 and 8 wavelengths selected, accounting for 5.0926% and 1.8519% of SiPLS bands; R^2 increased to 0.6742 and 0.7309, and RMSE decreased to 6.6482 and 5.9283. SiPLS-SPA method had the worst effect among the three combinations of SiPLS. Among the three sections of the SiPLS-SPA-PLS model of MOR, only the radial section model had a higher accuracy than the full-band spectral model, but its model stability was less than the SiPLS-PLS model. Ten wavelengths were selected to construct SiPLS-SPA-PLS, and R^2 of calibration set and validation set increased by 0.0281 and 0.1713, RMSEc increased by 0.4085 and RMSEv decreased by 2.0736 compared with the optimal full-band spectral model. The model complexity was significantly reduced by first collecting valid information intervals by SiPLS and then filtering out redundant variables by CARS. The SiPLS-CARS model of wood mechanical properties NIR spectrum had the best performance, and the R^2 of SiPLS-CARS-PLS was all greater than 0.9700. The optimal model of LCS

was the radial section SiPLS-CARS-PLS model. Forty and 53 wavelength variables were selected, accounting for 9.2807% and 12.2970% of the SiPLS band; R^2 was 0.9983 and 0.9985, and RMSE was 0.2242 and 0.2010, respectively. The optimal model of MOR was radial section SiPLS-CARS-PLS. Compared with the optimal full-band spectral model, R^2 increased by 0.3899 and 0.4868, and RMSE decreased by 6.7994 and 6.4220. The transverse sectional SiPLS-CARS-PLS model of MOE had the best performance. Wavelength variables 57 and 16 were selected to construct the model. The calibration set Rc^2 = 0.9947 and RMSEC = 112.7340, and the validation set Rv^2 = 0.9946 and RMSEV = 100.4788.

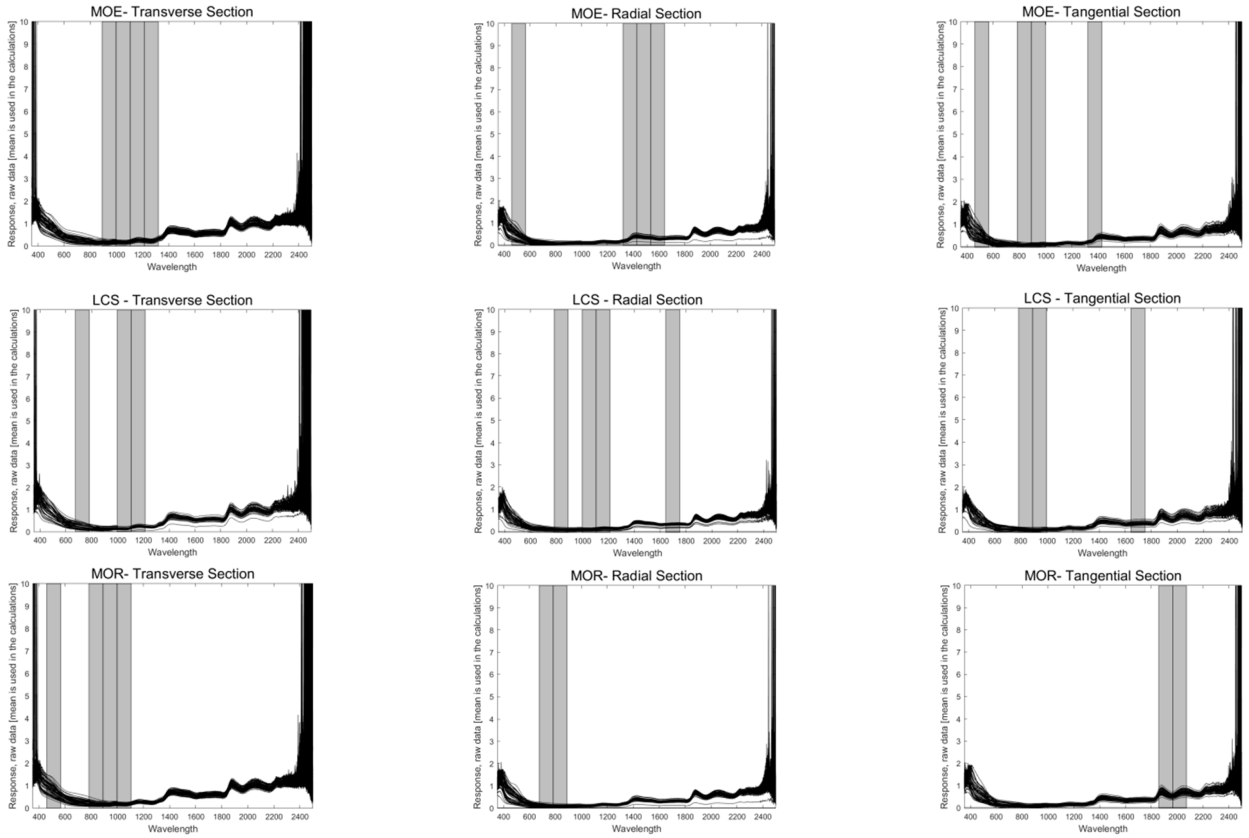

Figure 8. Band screening results of wood mechanical properties Vis-NIRS based on SiPLS.

The wavelength variables of NIRS spectra of different sections of wood mechanical properties were selected, and the results of constructing the PLSR model are shown in Table 3. The transverse sectional CARS model of the single wavelength-variable selection method of LCS was better. SiPLS-CARS-PLS was the optimal model in the combination of wavelength-variable selection methods for different sections of LCS, the R^2 of the optimal model was both above 0.9700, and the RMSE was less than 1.0000. The CARS model of radial section of MOR NIR spectrum had the highest model accuracy among the single wavelength-variable selection methods. In the combination of different wavelength-variable selection methods, the optimal model of MOR was SiPLS-CARS-PLSR, R^2 was greater than 0.9800, RMSE was smaller than 2.0000, and the SiPLS-CARS-PLSR model with diameter section had the highest accuracy. The optimal model for single-wavelength selection of MOE was the CARS-PLSR model of radial section. The optimal method of MOE wavelength-variable selection method combined strategy was SiPLS-CARS, and the transverse section SiPLS-CARS-PLSR model had the best performance, with R^2 of the calibration set and the validation set both greater than 0.9900, while RMSE were less than 120.

Table 4. PLSR modeling results based on SiPLS.

Mechanical Property	Section	Number of Interval Combinations	Calibration Models		Validation Models	
			RMSE	R^2	RMSE	R^2
LCS	Trans Section	2	4.4797	0.3997	3.1738	0.5487
		3	4.2458	0.4607	3.0068	0.5949
		4	4.4279	0.4135	1.8014	0.8546
	Rad Section	2	2.1382	0.8423	2.5097	0.7593
		3	2.5294	0.7793	1.7309	0.8855
		4	0.4484	0.9931	1.8195	0.8735
	Tang Section	2	2.2984	0.7744	3.2551	0.6725
		3	1.4138	0.9146	3.6286	0.5930
		4	3.2874	0.5385	3.7984	0.5540
MOR	Trans Section	2	7.4541	0.5904	5.8208	0.7405
		3	6.9913	0.6397	0.3208	0.9992
		4	6.7776	0.6614	5.7152	0.7499
	Rad Section	2	6.2629	0.7593	5.2525	0.7643
		3	8.4734	0.5594	4.9739	0.7887
		4	7.3755	0.6662	4.9026	0.7947
	Tang Section	2	5.5616	0.7494	7.7502	0.4102
		3	7.0872	0.5930	6.2916	0.6113
		4	7.8786	0.4970	6.0092	0.6454
MOE	Trans Section	2	619.4787	0.8395	1106.1957	0.3491
		3	435.2256	0.9208	754.0843	0.6975
		4	402.3331	0.9323	759.8211	0.6929
	Rad Section	2	903.7286	0.6198	803.3156	0.6443
		3	770.2608	0.7238	781.9038	0.7037
		4	770.1811	0.7239	712.5736	0.7539
	Tang Section	2	1034.6431	0.5056	1177.3204	0.2434
		3	1076.639	0.4647	970.9668	0.4854
		4	990.3318	0.5471	950.6278	0.5067

The optimal models for LCS, MOR and MOE in different sections were all SiPLS-CARS-PLSR, and the R^2 of calibration set and validation set were all greater than 0.9900. The PLSR modeling results are shown in Figure 9. The trend line of calibration model and target were nearly coincident, indicating that the SiPLS-CARS-PLSR model can effectively predict the wood mechanical properties of *Larix gmelinii*. In the meantime, the high fit of the trend line of validation model and the target indicates that the model was robust.

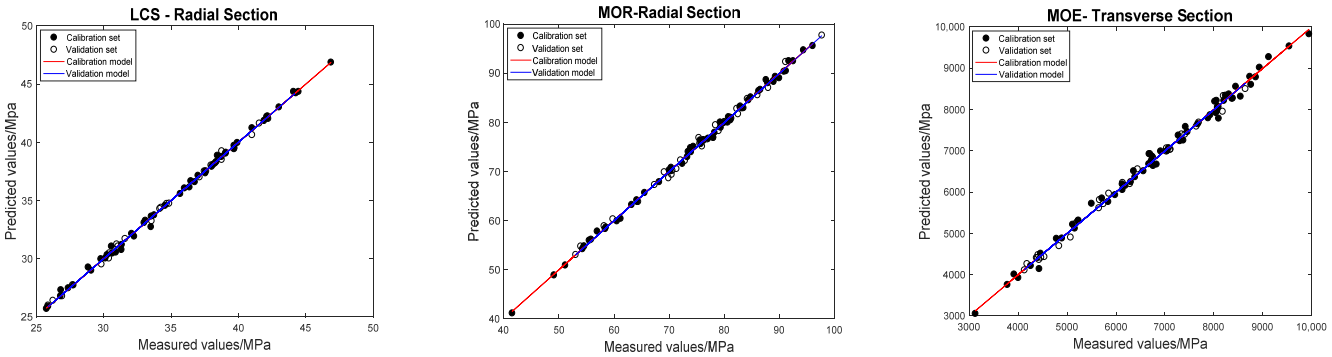

Figure 9. The SiPLS-CARS-PLSR predicting model for *Larix gmelinii* wood mechanical properties.

4. Conclusions and Discussion

The advantages of Vis-NIRS are non-destructiveness and fast detection, which can make up for the shortcomings of traditional destructive mechanical detection. Vis-NIRS

wood mechanical testing can provide an important reference for the scientific development of oriented breeding of *Larix gmelinii*, improvement of wood properties, and processing and utilization of solid wood [26]. In this study, choosing the *Larix gmelinii* in Xinlin Forest Farm as the research object, the modeling and prediction of LCS, MOR and MOE of *Larix gmelinii* in transverse, radial and tangential section were determined by using Vis-NIRS. The differences in the structural characteristics and wood compositions make the spectra absorbance of the three sections different. In order to remove the interference of singular values on the model performance, the MCCV method is used to screen out singular samples. Compared to the prediction results of the PLSR model without screening singular samples, the model predicting ability after eliminating the singular samples is significantly improved [27,48]. In order to ensure that the representativeness of the calibration set samples is comprehensive and that all samples are uniformly distributed in each set, the SPXY method is used to divide the sample sets.

The performance of nine different models, PLSR, CARS-PLSR, SPA-PLSR, UVE-PLSR, SiPLS-PLSR, UVE-CARS-PLSR, SiPLS-UVE-PLSR, SiPLS-CARS-PLSR and SiPLS-SPA-PLSR, was compared to select variables for modeling and to determine the best modeling method. The optimal single wavelength-variable selection method for different mechanical properties of wood was CARS, and the R^2 of calibration set and validation set was more than 0.9100, indicating that CARS had a strong ability to screen wavelength variables. The optimal models for different mechanical properties of wood were SiPLS-CARS-PLSR, the optimal model of LCS and MOR was the radial section model, the optimal model of MOE was the transverse sectional model. Rc^2 and Rv^2 of the optimal model were greater than 0.9900, RMSEC were 0.2242, 0.4873 and 112.7340, and RMSEV were 0.2010, 0.7267 and 100.4788, respectively. The results show that the performance of the Vis-NIRS model was significantly different with different sections of *Larix gmelinii*. The optimal models of LCS, MOR and MOE were the SiPLS-CARS-PLSR model, with R_c^2 of calibration set and R_v^2 of validation set both being greater than 0.9900, and RMSE the smallest. Due to the good prediction accuracy and robustness of the model, the mechanical properties of *Larix gmelinii* can be effectively predicted. SiPLS-CARS was the optimal characteristic wavelength-variable selection method, which could significantly improve the accuracy and stability of the model. SiPLS-CARS can effectively predict the mechanical properties of larch, provide technical support for the determination of wood mechanical properties, and provide a theoretical basis for the breeding of *Larix gmelinii* varieties and forest management [29].

Among the four single classical wavelength selection algorithms, UVE and SPA showed poor results, and SiPLS and CARS showed good results. Comparing CARS and SiPLS, CARS calculated quickly and obtained fewer characteristic variables, while SiPLS took longer to calculate, and the screened interval contained irrelevant variables. Therefore, the optimal single-wavelength-variable selection method is CARS, which is consistent with the results of soil optimization [10]. Different wavelength selection methods have their limitations. In order to make up for the defects of the classical algorithm, four combined strategies were used to select the characteristic variables of the near-infrared spectrum of wood mechanical properties. When the two variable selection methods are coupled, due to the different wavelength selection principles, the optimization effects are not the same [48]. Comprehensive comparison showed that SiPLS and CARS were better, and SiPLS-CARS-PLS was optimal. Combining characteristic variable selecting method for Vis-NIRS modeling is better than the single wavelength-variable selection method. This is because not all characteristic variable selection methods can find strong information variables and eliminate weak and unrelated information variables, due to the differences in principles of wavelength selection. The combination of some methods can achieve the function complementarity [49,50].

Wood mechanical properties are important indicators of wood properties, but wood density and wood chemical properties are also significant in wood utilization. In order to better and comprehensively study wood properties, we will study other wood indicators

and explore the relationship between these indicators in the future. The NIRS model cannot accurately detect the experimental target due to changes in external conditions or instrument replacement, but the cost of model reconstruction is high. Therefore, we will study the model transfer that can improve the applicability and accuracy of the NIRS model in the future.

Author Contributions: Conceptualization, Y.L. and C.L.; methodology, C.L. and Y.L.; software, C.L.; investigation and sample collection, C.L., Y.Z., Z.Z. and Z.W.; writing—original draft preparation, C.L.; writing—review and editing, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Applied Technology Research and Development Plan of Heilongjiang Province (GA19C006) and The Key Research and development plan of Heilongjiang Province (GA21C030).

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

References

1. Li, J. A Brief Discussion of Preparing the Standard Volume Table of the Artificial Larch of the Daxing'anling Forestry Region in Inner Mongolia. *Inn. Mong. For. Investig. Des.* **2014**, *37*, 59–61,107. [CrossRef]
2. Mu, C.; Sun, X.; Ni, Z.; Yang, M.; Zhang, N. Comprehensive Evaluation of the Effects Planting in Swamp-Forest in Daxing'an Mountains. *Sci. Silvae Sin.* **2007**, *43*, 51–58. [CrossRef]
3. Xin, X.; Li, Y. Economy Analysis of Larix Gmelini. *For. Sci. Technol. Inf.* **2009**, *41*, 6–7. [CrossRef]
4. Zhao, X. Study on the Strength Properties of Larch Dimension Lumber. Ph.D. Thesis, Chinese Academy of Forestry, Beijing, China, 2010.
5. Sorensson, C.; Bian, R.; Wellauer, B.; Alley, R. Structural lumber conversion potential of clonal radiata. *Proc. Wood Qual.* **2004**. [CrossRef]
6. Huang, T.; Fu, Y.; Li, N. Physical and Mechanical Properties of *Shorea chinensis* Wood. *J. Northwest For. Univ.* **2013**, *28*, 160–163. [CrossRef]
7. Huang, T.; Wang, J.; Luan, J.; Song, L.; Lei, F. Evaluation of Wood Physical and Mechanical Properties in *Eucalyptus urophylla* × *E. grandis* Clone Plantations at Different Ages. *Chin. J. Trop. Agric.* **2021**, *41*, 95–100. [CrossRef]
8. Sun, X.; Chu, X.; Zhang, S.; Liu, J. Timber Evaluation on Physical and Mechanical Properties of Species and Hybrids of Larix. *Sci. Silvae Sin.* **2012**, *48*, 153–159. [CrossRef]
9. Li, J. *Wood Science Research*; Science Press: Beijing, China, 2009; pp. 163–164.
10. Li, C.; Zhao, J.; Li, Y.; Meng, Y.; Zhang, Z. Modeling and Prediction of Soil Organic Matter Content Based on Visible-Near-Infrared Spectroscopy. *Forests* **2021**, *12*, 1809. [CrossRef]
11. Meng, Y.; Zhang, Y.; Li, C.; Zhao, J.; Wang, Z.; Wang, C.; Li, Y. Prediction of the Carbon Content of Six Tree Species from Visible-Near-Infrared Spectroscopy. *Forests* **2021**, *12*, 1233. [CrossRef]
12. Chu, X.; Chen, P.; Li, J.; Liu, D.; Xu, Y. Progresses and Perspectives of Near Infrared Spectrocopy Analytical Technology. *J. Instrum. Anal.* **2020**, *39*, 1181–1188. [CrossRef]
13. Li, Y.; Via, B.K.; Cheng, Q.; Zhao, J.; Li, Y. New Pretreatment Methods for Visible–Near-Infrared Calibration Modeling of Air-Dry Density of *Ulmus pumila* Wood. *For. Prod. J.* **2019**, *69*, 188–194. [CrossRef]
14. Li, Y.; Via, B.K.; Young, T.; Li, Y. Visible-near infrared spectroscopy and chemometric methods for wood density prediction and origin/species identification. *Forests* **2019**, *10*, 1078. [CrossRef]
15. Yu, H.; Zhao, R.; Fu, F.; Fei, B.; Jiang, Z. Prediction Mechanical Properties of Chinese Fir Wood by Near Infrared Spectroscopy. *J. Northwest For. Univ.* **2007**, *22*, 149–154. [CrossRef]
16. Fernández-Cabanás, V.; Garrido-Varo, A.; Pérez-Marín, D.; Dardenne, P. Evaluation of pretreatment strategies for near-infrared spectroscopy calibration development of unground and ground compound feedingstuffs. *Appl. Spectrosc.* **2006**, *60*, 17–23. [CrossRef] [PubMed]
17. Silalahi, D.D.; Midi, H.; Arasan, J.; Mustafa, M.S.; Caliman, J.-P. Robust generalized multiplicative scatter correction algorithm on pretreatment of near infrared spectral data. *Vib. Spectrosc.* **2018**, *97*, 55–65. [CrossRef]
18. Yuan, Y.; Wang, W.; Chu, X.; Xi, M. Selection of Characteristic Wavelengths Using SPA and Qualitative Discrimination of Mildew Degree of Corn Kernels Based on SVM. *Spectrosc. Spectr. Anal.* **2016**, *36*, 226–230. [CrossRef]
19. Tian, H.; Zhang, L.; Li, M.; Wang, Y.; Sheng, D.; Liu, J.; Wang, C. Weighted SPXY method for calibration set selection for composition analysis based on near-infrared spectroscopy. *Infrared Phys. Technol.* **2018**, *95*, 88–92. [CrossRef]
20. Zhan, X.-R.; Zhu, X.-R.; Shi, X.-Y.; Zhang, Z.-Y.; Qiao, Y.-J. Determination of hesperidin in tangerine leaf by near-infrared spectroscopy with SPXY algorithm for sample subset partitioning and Monte Carlo cross validation. *Spectrosc. Spectr. Anal.* **2009**, *29*, 964–968. [CrossRef]

21. Horvath, L.; Peszlen, I.; Peralta, P.; Kelley, S. Use of transmittance near-infrared spectroscopy to predict the mechanical properties of 1- and 2-year-old transgenic aspen. *Wood Sci. Technol.* **2014**, *45*, 303–314. [CrossRef]
22. Fujimoto, T.; Yamamoto, H.; Tsuchikawa, S. Estimation of wood stiffness and strength properties of hybrid larch by near-infrared spectroscopy. *Appl. Spectrosc.* **2007**, *61*, 882–888. [CrossRef]
23. Leblon, B.; Adedipe, O.; Hans, G.; Haddadi, A.; Tsuchikawa, S.; Burger, J.; Stirling, R.; Pirouz, Z.; Groves, K.; Nader, J.; et al. A review of near-infrared spectroscopy for monitoring moisture content and density of solid wood. *For. Chron.* **2013**, *89*, 595–606. [CrossRef]
24. Thumm, A.; Meder, R. Stiffnes prediction of radiata pine clearwood test pieces using infrared spectroscopy. *J. Near Infrared Spectrosc.* **2001**, *9*, 117–122. [CrossRef]
25. Guo, Z.; Wang, M.; Agyekum, A.A.; Wu, J.; Chen, Q.; Zuo, M.; El-Seedi, H.R.; Tao, F.; Shi, J.; Ouyang, Q. Quantitative detection of apple watercore and soluble solids content by near infrared transmittance spectroscopy. *J. Food Eng.* **2020**, *279*, 109955. [CrossRef]
26. Zhang, P.; Shang, X.; Wu, Z. Difference in Wood Physical and Mechanical Properties of Five Kinds of Eucalyptus. *J. Northwest For. Univ.* **2020**, *35*, 184–190. [CrossRef]
27. Li, S.; Shan, Y.; Fan, W.; Yin, Y.; Zhou, Z.; Li, G. Analysis of pH and Acidity of Honey by Near Infrared Spectroscopy Based on MCCV Outlier Dectection and CARS Variable Selection. *Food Sci.* **2011**, *32*, 182–185.
28. Liu, Y.; Ying, Y.; Jiang, H. Rapid determination of maturity in apple using outlier detection and calibration model optimization. *Trans. ASABE* **2006**, *49*, 91–95. [CrossRef]
29. Liu, C.; Hu, Y.; Wu, S.; Sun, X.; Dou, S.; Miao, Y.; Dou, Y. Outlier Sample Eliminating Methods for Building Calibration Model of Near Infrared Spectroscopy Analysis. *J. Food Sci. Technol.* **2014**, *32*, 74–79. [CrossRef]
30. Xu, F.; Fu, D.; Wang, Q.; Xiao, Z.; Wang, B. Nondestructive detection of sugar content and acidity in Red Globe table grapes using visible near infrared spectroscopy based on Monte-Carlo cross validation-competitive adaptive reweighted sampling-random forest (MCCV-CARS-RF). *Shipin Kexue Food Sci.* **2018**, *39*, 149–154. [CrossRef]
31. Yang, Z.; Xiao, H.; Zhang, L.; Feng, D.; Zhang, F.; Jiang, M.; Sui, Q.; Jia, L. Fast determination of oxide content in cement raw meal using NIR spectroscopy with the SPXY algorithm. *Anal. Methods* **2019**, *11*, 3936–3942. [CrossRef]
32. Chen, Y.; Qi, T.; Huang, Y.; Wan, Y.; Zhao, R.; Qi, L.; Zhang, C.; Fei, T. Optimization method of calibration dataset for VIS_NIR spectral inversion model of soil organic matter content. *Trans. Chin. Soc. Agric. Eng.* **2017**, *33*, 107–114. [CrossRef]
33. Mei, C.; Chen, Y.; Yin, L.; Jiang, H.; Chen, X.; Ding, Y.; Liu, G. Wavelength Selection by SiPLS-LASSO for NIR Spectroscopy and Its Application. *Spectrosc. Spectr. Anal.* **2018**, *38*, 436–440. [CrossRef]
34. Gan, L.; Sun, T.; Liu, J.; Liu, M. Double Pulse LIBS Combined with Variable Screening to Detect Procymidone Content. *Spectrosc. Spectr. Anal.* **2019**, *39*, 584–588. [CrossRef]
35. Huo, Y.; Zhang, C.; Li, Y.; Zhi, W.; Zhang, J.; Liu, J. Nondestructive detection for kiwifruit based on the hyperspectral technology and machine learning. *J. Chin. Agric. Mech.* **2019**, *40*, 71–77. [CrossRef]
36. Zhang, Y.; Ren, D.; Han, Y.; Li, J. Air target reference spectrum selection based on characteristic wavelengths extracted by successive projections algorithm. *Infrared Laser Eng.* **2021**, *50*, 232–242. [CrossRef]
37. Huang, P.; Li, Y.; Yu, Q.; Wang, K.; Yin, H.; Hou, D.; Zhang, G. Classification of Organic Contaminants in Water Distribution Systems Developed by SPA and Multi-Classification SVM Using UV-Vis Spectroscopy. *Spectrosc. Spectr. Anal.* **2020**, *40*, 2267–2272. [CrossRef]
38. Sun, T.; Wu, Y.; Liu, X.; Mo, X.; Liu, M. Detection of Chromium Content in Soybean Oil by Laser Induced Breakdown Spectroscopy and UVE Method. *Spectrosc. Spectr. Anal.* **2016**, *36*, 3341–3345. [CrossRef]
39. Miao, X.; Miao, Y.; Gong, H.; Tao, S.; Chen, Y.; Chen, Z. Determination of Moisture Content in Rice by Near Infrared Spectroscopy with Different Partial Least Squares. *J. Anal. Sci.* **2019**, *35*, 643–649. [CrossRef]
40. Yang, H.; Zhu, M. Study of Rapid Detection of Soil Organic Matter Based on Characteristic Wavelength Selection of Visible-near Infrared Spectra. *Infrared* **2015**, *36*, 42–48. [CrossRef]
41. Yang, Z.; Xiao, H.; Zhang, L.; Feng, D.; Zhang, F.; Jiang, M.; Sui, Q.; Jia, L. Fast determination of oxides content in cement raw meal using NIR spectroscopy combined with synergy interval partial least square and different preprocessing methods. *Measurement* **2020**, *149*, 106990. [CrossRef]
42. Chen, Y.; Di, Y.; Tang, X.; Cui, X.; Gao, X.; Cao, J.; Li, S. Combination Weight COD Concentration Prediction Model Based on BiPLS and SiPLS. *Spectrosc. Spectr. Anal.* **2019**, *39*, 2176–2181. [CrossRef]
43. Cheng, B.; Chen, D.; Wu, X. Near Infrared Spectral Wavelength Selection Based on Moving Window-Iterative Genetic Algorithm Method. *Chin. J. Anal. Chem.* **2006**, *34*, 123–130. [CrossRef]
44. Shen, L.; Gao, M.; Yan, J.; Yao, Y. Estimation model of soil organic matter based on SVR and PLSR. *China Agric. Inform.* **2019**, *31*, 58–71. [CrossRef]
45. Di Bucchianico, A. Coefficient of determination (R^2). *Encycl. Stat. Qual. Reliab.* **2008**, *1*. [CrossRef]
46. Hong, Y.; Yu, L.; Zhu, Y.; Li, S.; Guo, L.; Liu, J.; Nie, S.; Zhou, Y. Using Orthogonal Signal Correction Algorithm Removing the Effects of Soil Moisture on Hyperspectral Reflectance to Estimate Soil Organic Matter. *Sci. Agric. Sin.* **2017**, *50*, 3766–3777. [CrossRef]
47. Liu, Z.; Cai, W.; Shao, X. Monte Carlo Cross Validation for Identification of Singular Samples in Near Infrared Spectroscopy. *Sci. Sin.* **2008**, *38*, 316–323. [CrossRef]

48. Jia, M.; Li, W.; Wang, K.; Zhou, C.; Cheng, T.; Tian, Y.; Zhu, Y.; Cao, W.; Yao, X. A newly developed method to extract the optimal hyperspectral feature for monitoring leaf biomass in wheat. *Comput. Electron. Agric.* **2019**, *165*, 104942. [CrossRef]
49. Zhang, Y.; Su, Y.; Li, C.; Men, H. Analysis of MOR and MOE prediction model of *Quercus mongolica* wood by near infrared spectroscopy. *J. Beijing For. Univ.* **2016**, *38*, 99–105. [CrossRef]
50. Hu, M.; Wen, Z. A New Wavelength Selection Algorithm Based on the Fusion of Multiple Models. *Spectrosc. Spectr. Anal.* **2010**, *30*, 2088–2092. [CrossRef]