

Prediction of Needle Chlorophyll Model with Different Leaf Ages Based on BP Neural Network and PLSR

yiheng wang

Beihua University

Zhe Wen

Beihua University

Kun Sun

Beihua University

Yanhong Xu

Beihua University

Jinhua Wei (✉ enheng1212@126.com)

Beihua University <https://orcid.org/0000-0002-5565-7697>

Methodology

Keywords: BP neural network, PLSR partial least squares regression, spectral pre-processing, variable selection, chlorophyll

Posted Date: January 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-143311/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Prediction of Needle Chlorophyll Model with Different Leaf Ages Based on BP Neural Network and PLSR

Yiheng Wang, Zhe Wen, Kun Sun, Yanhong Xu, Jinhua Wei*

Forestry College of Beihua University, Jilin 132013, China

* Corresponding author: enheng1212@126.com

Abstract

Background: To explore the changes of chlorophyll content in needles of different leaf ages of *Picea koraiensis* Nakai of different specifications, the study compared the prediction accuracy changes of chlorophyll in needles of *Picea koraiensis* Nakai by two modeling methods: the BP neural network and partial least squares regression (PLSR) methods. The effects of different spectral pre-processing and characteristic band selection methods on the performance accuracy of the model were tested, and the optimal combination model was selected to predict forest growth status and community structure productivity through the physiological and biochemical characteristics of needles at different leaf ages.

Results: 1) the spectral pre-processing method could avoid systematic errors and eliminate background values; 2) the accuracy of the needle chlorophyll fitting model with different leaf ages was much higher than that of mixed needle chlorophyll model, verifying that needle chlorophyll with different leaf ages could better estimate the annual growth and examine the growth status of *Picea koraiensis* Nakai; 3) the accuracy of the BP neural network model was significantly higher than that of the PLSR model, with its R^2 above 0.95, and the validation set's R^2 above 0.86; and 4) the fitting accuracy of different leaf age needle chlorophyll models

of the spectral pre-processing model, variable selection model, PLSR model and BP neural network: triennial needles > annual and biennial needles.

Conclusions: The BP neural network method was more accurate than the PLSR method in predicting pigment content model. In the process of model fitting, it was found that the pigment model fitted by fine classification of needles improves the accuracy of the model, which provides the basis and theoretical support for the establishment of the model by combining remote sensing technology with stoichiometry methods in the future.

Keywords: BP neural network; PLSR partial least squares regression; spectral pre-processing; variable selection; chlorophyll

Background

The data of the eighth national forest resource inventory showed that China's 69.33 million hm² of plantation area was the most in the world, while its afforestation area and accumulation of coniferous pure forest accounted for 73.07% and 74.58% of the world's total plantation area and total accumulation, respectively^[1]. *Picea koraiensis* Nakai is the main conifer species for afforestation and timber in Northeast China, and the cultivation and management of plantations by *Picea koraiensis* Nakai have been included in the National "Thirteenth Five-Year Plan" Key R&D Program^[2]. Coniferous stands are often caused by unreasonable allocation of spatial and non-spatial structures, forcing trees to compete for their spatial niches, resulting in unhealthy stands and unstable structural attributes. However, leaf pigment content, as one of the indicators for evaluating the status of plant and forest health, has been identified as a trait associated with ecosystem structure and its biodiversity relationship^[3]. Its pigment content data can reflect the status of plant growth such as nutrient

stress, photosynthetic capacity and senescence process, which is helpful for the biogeochemistry and nutrient cycle of ecosystems^[4-5]. At present, the traditional method for the determination of plant physiological and biochemical indicators is extremely time consuming and labor-intensive, and even environmentally destructive. An important goal of forestry research is how to quickly monitor the growth status, content of inclusion, health evaluation and other information related to large area stand structure and provide scientific basis and theoretical support for the decision-making of forestry production management. With the evolution of sensor technology, hyperspectral remote sensing technology can easily obtain the reflection information of different targets, and the fine spectral resolution can fully reflect the characteristics and differences of subtle changes in vegetation spectra^[6]. Wen et al.^[7] confirmed that the establishment of a BP neural network model, which is based on band depth analysis, can improve the estimation accuracy of pigment content in rice leaves by spectral techniques. Yuling et al.^[8] showed that the partial least squares regression (PLSR) technique was superior to the stepwise multiple linear regression (SMLR) technique in predicting the model accuracy of soil heavy metals. However, variable selection using the squared multiple correlation (SMC) method can improve the PLSR model accuracy^[9]. Garhwal et al.^[10] used the Variable Importance in Projection (VIP) method for band optimization to reveal the sensitive band of diseased potatoes and improve recognition accuracy.

The use of hyperspectral technology for rapid non-destructive measurement is favored by many scholars, but there are few reports on the application of spectral technology to predict the pigment content of needles at different leaf ages of high-throughput conifers in China,

combined with different spectral pre-processing and characteristic band selection to improve the accuracy of the model. The use of an artificial neural network model to parameter inversion is also in the early stages of research in China^[11]. Therefore, in this study, based on the measured needle chlorophyll content and needle spectral reflectance of annual, biennial and triennial needles of *Picea koraiensis* Nakai, the construction of different specifications of needle chlorophyll models of different leaf ages of *Picea koraiensis* Nakai involves the use of five pre-treatment methods (Savitzky-Golay Raw Spectra (SG-Raw), Savitzky-Golay First Derivative (SG-FD), Savitzky-Golay Second Derivative (SG-SD), Standard Normal Variate (SNV) and Multiplicative Scatter Correction (MSC); two variable selection methods SMC and VIP; and two modeling methods, BP neural network model and PLSR. The main goals of our research are 1) to compare the effects of different spectral pre-processing and variable selection on the performance of the model; 2) to compare the prediction accuracy of BP neural network and PLSR for different leaf age needle chlorophyll and mixed needle chlorophyll models of *Picea koraiensis* Nakai; and 3) to select the optimal combination model, in order to provide data and theoretical support for remote sensing technology in the analysis of physiological and biochemical indicators and evaluation of the health status of needle forests in northeast China in the future.

Results

Chlorophyll Content Statistics and Spectral Preprocessing Analysis

The statistical values of chlorophyll content are shown in Table 1, and the chlorophyll content of different leaf ages of *Picea koraiensis* Nakai with the same ground diameter: triennial needles > biennial needles > annual needles. The chlorophyll content of *Picea koraiensis*

Nakai with the ground diameter of 25 – 30 cm at the same leaf age was greater than that with the ground diameter of 20 – 25 cm, and the dispersion was also less than that of *Picea koraiensis* Nakai with the ground diameter of 20 – 25 cm.

When the Specim hyperspectral camera is used to collect needle reflectance, the surrounding noisy environment, light scattering and diffuse reflection will have a certain impact on the spectral data. Therefore, pre-processing of the spectral data can avoid the errors caused by noise and baseline translation^[12]. Although there was no significant change in the spectral reflectance of *Picea koraiensis* Nakai needles of different sizes at the same leaf age, there was a significant difference in the spectral reflectance of needles of different leaf ages at the same size ($p < 0.05$) (Fig. 2). The spectral reflectance values of annual conifers SG-Raw and MSC in the bands of 513 nm to 616 nm and 720 nm to 988 nm were higher than those in the conifers biennial, triennial conifers and mixed *Picea koraiensis* Nakai conifers. In addition, extreme points were formed near 711 nm and 690 nm in the SG-FD and SG-SD spectra, respectively, which were due to the strong absorption of chlorophyll in different bands. These findings were consistent with results in the existing literature^[13].

Table 1 Statistics of chlorophyll content in needle leaves of *Picea koraiensis* Nakai at different leaf ages with different ground diameters. Note: Lower case letters indicate that the chlorophyll content of *Picea koraiensis* Nakai needles with different leaf ages at different ground diameters was significantly different at the 0.05 level.

Ground diameter	Leaf age	Max	Min	Mean	Standard Deviation	Coefficient of variation
25–30cm	annual	0.96d	0.24d	0.59e	0.16e	0.27b
	biennial	1.20a	0.28c	0.69c	0.18c	0.26c
	triennial	1.20a	0.35a	0.77a	0.19b	0.25d
20–25cm	annual	0.93e	0.16f	0.58f	0.17d	0.29a
	biennial	1.11c	0.23e	0.66d	0.19b	0.29a
	triennial	1.19b	0.32b	0.75b	0.20 a	0.27b

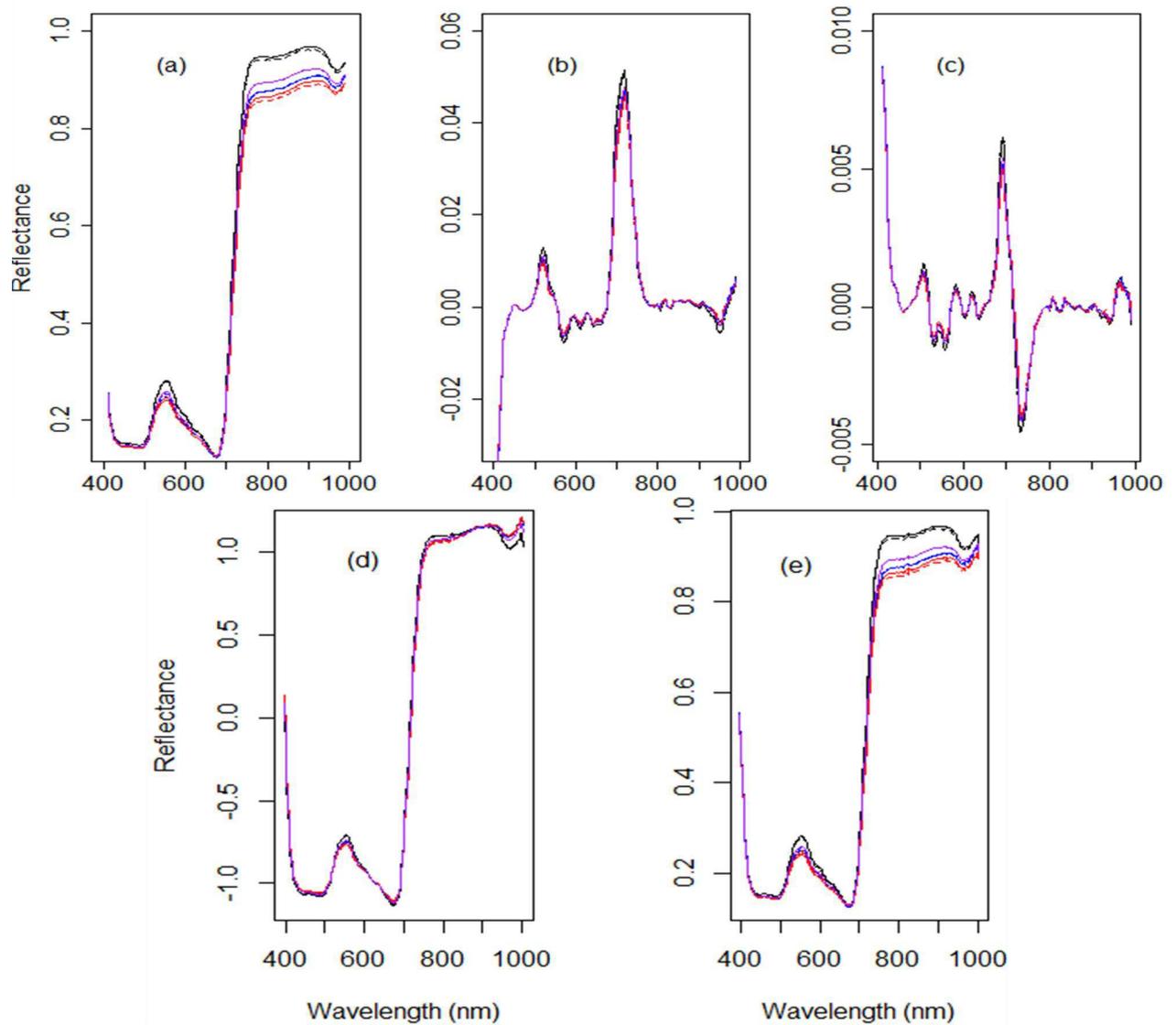


Figure 2. Spectral preprocessing curve; black: annual needles, blue: biennial needles, red: triennial needles, purple: mixed *Picea koraiensis* Nakai; solid line: *Picea koraiensis* Nakai with a diameter of 25–30 cm, dashed line: *Picea koraiensis* Nakai with a diameter of 20–25 cm; (a): SG-Raw, (b): SG-FD, (c): SG-SD, (d): SNV, (e): MSC.

PLSR Model Development

Response of Different Leaf Age Needle Pigment Models to Different Spectral Pretreatments

Five spectral pre-processing methods were applied to the PLSR model of conifer chlorophyll at different leaf ages (Fig. 3). The fitting accuracy of the *Picea koraiensis* Nakai model with a ground diameter of 25 – 30 cm was generally higher than that of *Picea koraiensis* Nakai with a ground diameter of 20 – 25 cm, and the fitting accuracy of triennial needles was higher than

that of annual and biennial needles. This may be related to the fact that the greater the leaf age and the stronger the conversion ability, the higher the pigment content^[14]. However, the fitting accuracy of mixed *Picea koraiensis* Nakai needle model is much lower than that of different specifications and different leaf age needles (Fig. 3). Therefore, the construction of the pigment content model for fine classification of needles can improve the robustness of the model.

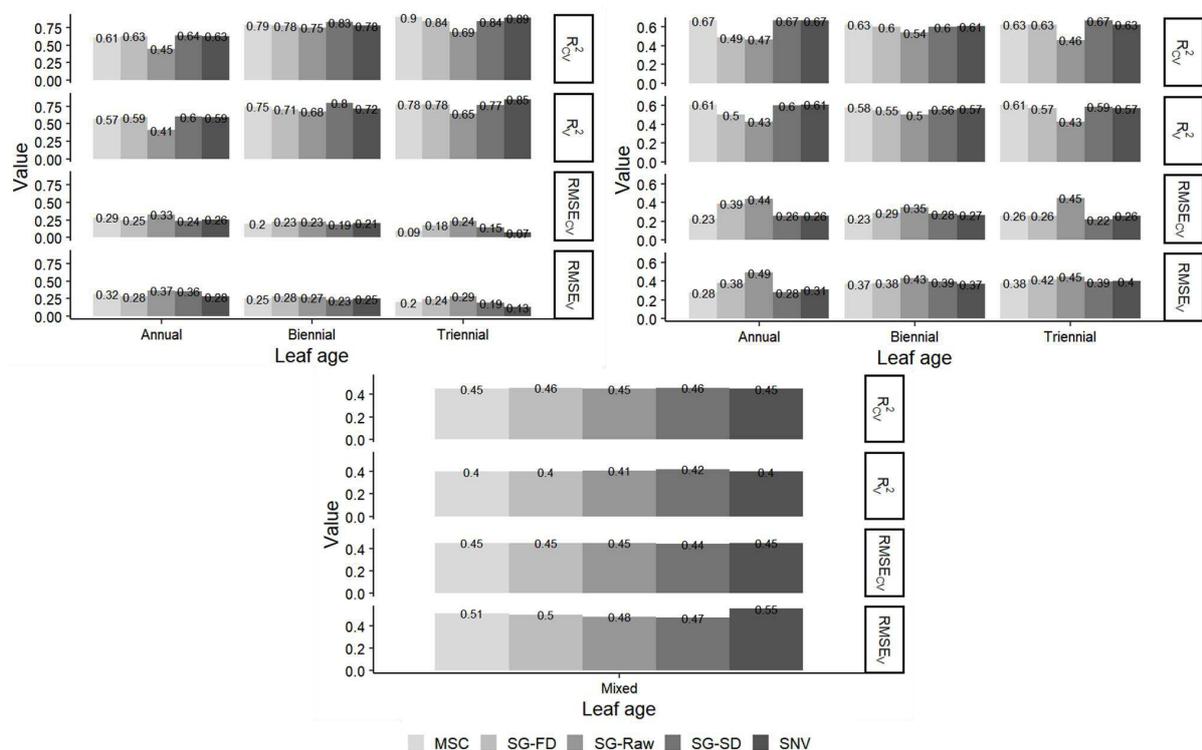


Figure 3. Fitting and evaluation of chlorophyll PLSR model for spectral pretreatment of needles of *Picea koraiensis* Nakai; upper left: 25 – 30 cm ground diameter *Picea koraiensis* Nakai, upper right: 20 – 25 cm ground diameter *Picea koraiensis* Nakai, lower: mixed *Picea koraiensis* Nakai; Annual: annual conifers, Biennial: biennial conifers, Triennial: triennial conifers, Mixed: mixed *Picea koraiensis* Nakai conifers.

Influence of Feature Band Extraction on Model Performance

To obtain a more robust yet simplified PLSR model with stronger predictive ability, the SMC and VIP variable selection methods were used to extract the characteristic bands and eliminate useless or nonlinear variables. After variable selection treatment, some models showed an increasing accuracy trend (Fig. 4), and the number of model components decreased, which

was consistent with the results of Yanjie L^[15]. Compared with the optimal models of annual, biennial, triennial needles and mixed needles that had been pretreated with SMC, and VIP treatments (Fig. 5), the models obtained after VIP treatment of annual and triennial needles of *Picea koraiensis* Nakai with a ground diameter of 25 - 30 cm were optimal, and the combinations treated by SMC to obtain the optimal models included: annual needles with a ground diameter of 25 - 30 cm + VIP + SG-Raw; triennial needles with a ground diameter of 25 - 30 cm + VIP + SG-FD, with R² of 0.9307 and 0.9344; the biennial needles of *Picea koraiensis* Nakai with a ground diameter of 25 - 30 cm; and annual, biennial and triennial needles with a ground diameter of 20 - 25 cm. The combinations also included: biennial needles with a ground diameter of 25 - 30 cm + SG-Raw; biennial needles with a ground diameter of 20 - 25 cm + annual needles + SG-SD; biennial needles with a ground diameter of 20 - 25 cm + MSC; and biennial needles with a ground diameter of 20 - 25 cm + triennial needles + SG-SD, with R² of 0.8847, 0.8040, 0.8612, and 0.8700, respectively. The mixed *Picea koraiensis* Nakai conifer chlorophyll model had the lowest fitting accuracy, while *Picea koraiensis* Nakai with 25–30 cm ground diameter had the highest accuracy.

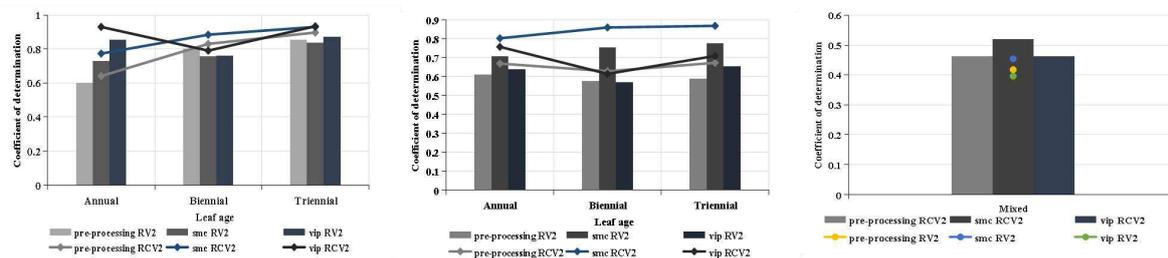


Figure 5. Comparison of spectral preprocessing and variables selection optimal model R² of different leaf ages and mixed *Picea koraiensis* Nakai in PLSR model; left: 25–30 cm ground diameter *Picea koraiensis* Nakai, middle: 20–25 cm ground diameter *Picea koraiensis* Nakai, right: mixed *Picea koraiensis* Nakai.

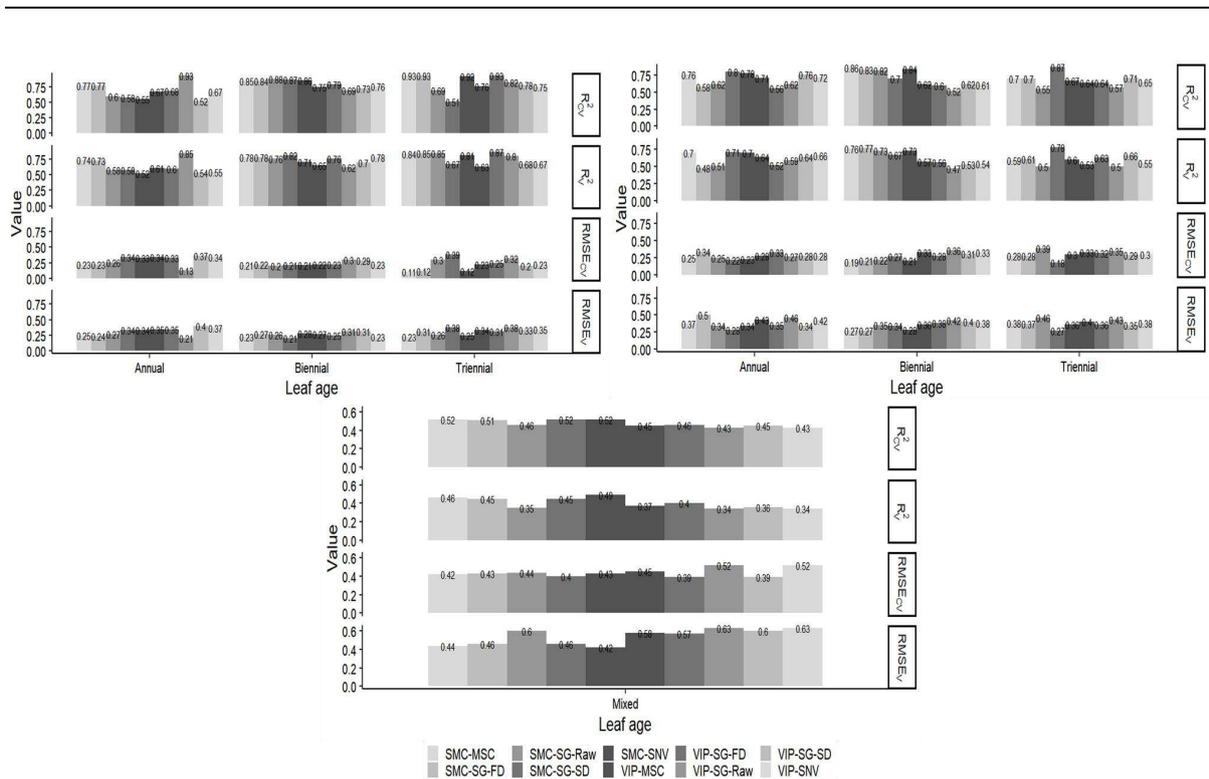


Figure 4. Fitting and evaluation of chlorophyll PLSR model spectra pretreatment and variable selection of *Picea koraiensis* Nakai; upper left: 25 – 30 cm ground diameter *Picea koraiensis* Nakai, upper right: 20 – 25 cm ground diameter *Picea koraiensis* Nakai, lower: mixed *Picea koraiensis* Nakai.

Construction of BP Neural Network Model

The BP neural network is trained by forward signal transmission and back error propagation mode, and the optimal weight value and threshold value are calculated to fit the model. In addition to the low fitting accuracy of the MSC-processed spectral model, the fitting accuracy of different specifications of the *Picea koraiensis* Nakai needle chlorophyll content model is greatly improved, which may be related to the fact that MSC weakens the signal of characteristic variables while correcting the spectral differences. This results in a large deviation of the weights of each neuron in the hidden and output layers of the neural network, leading to low training model accuracy, while the extremely strong error of SG-FD spectral value signal is small. Except for the annual needles of *Picea koraiensis* Nakai with a ground diameter of 25 – 30 cm, which had the best fitting accuracy with the SG-SD spectral model (Fig. 6), the other models had the highest model fitting accuracy on the basis of SG-FD

spectral pre-processing, with R^2 of 0.9543, 0.9858, 0.9874, 0.9667, 0.9661 and 0.9699. The accuracy of the fitted models of SG-FD and SG-SD was the highest in mixed *Picea koraiensis* Nakai needles, with R^2 of 0.9192 and 0.8898. The fitting accuracy of the chlorophyll model was higher in the triennial needles than in the annual and biennial ones, regardless of the size of *Picea koraiensis* Nakai. As can be seen in Fig. 4, the RMSE value of the BP neural network model is much lower than that of PLSR model.

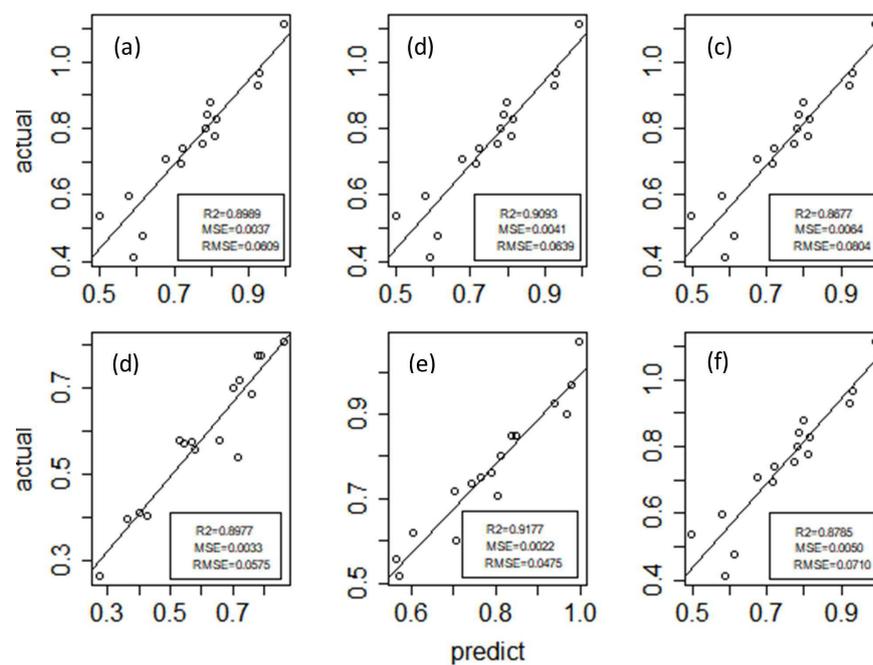


Figure 7. Scatter plot of the fit between the measured and predicted values of the BP neural network chlorophyll model. (a), (b), and (c) are annual, biennial and triennial needles of 25–30 cm ground diameter *Picea koraiensis* Nakai, respectively, and (d), (e), and (f) are annual, biennial and triennial needles of 20–25 cm ground diameter *Picea koraiensis* Nakai..

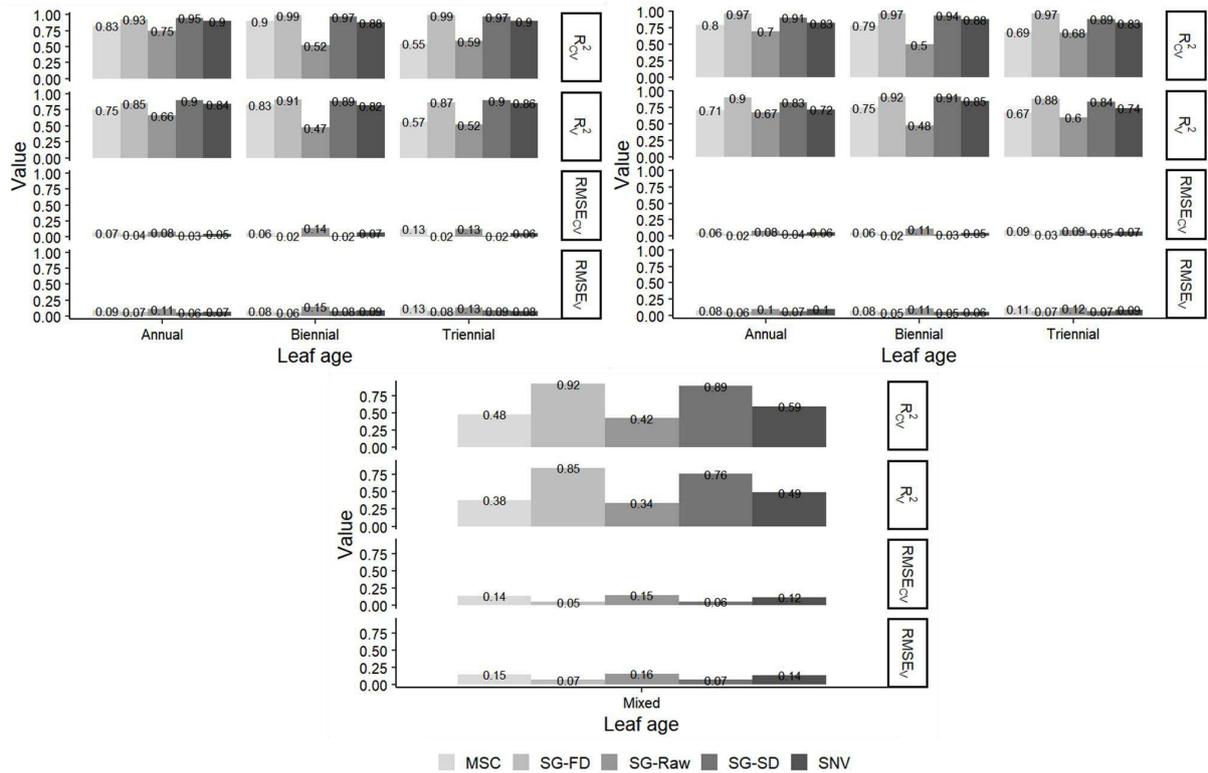


Figure 6. Fitting and evaluation of BP neural network chlorophyll model of *Picea koraiensis* Nakai; upper left: 25 – 30 cm ground diameter *Picea koraiensis* Nakai, upper right: 20 – 25 cm ground diameter *Picea koraiensis* Nakai, lower: mixed *Picea koraiensis* Nakai.

Optimal Combination Models

Comparing Figs. 3, 4 and 6, the fitting accuracy of the needle chlorophyll models with different leaf ages by the BP neural network was significantly higher than the fitting accuracy of the PLSR model. The optimal combination models of different leaf ages of *Picea koraiensis* Nakai were annual needles of *Picea koraiensis* Nakai with ground diameter of 25 – 30 cm: BP + SG-SD; biennial, triennial needles of *Picea koraiensis* Nakai with ground diameter of 25 – 30 cm; and annual, biennial and triennial needles of *Picea koraiensis* Nakai with ground diameter of 20 – 35 cm: BP + SG-FD. Fig. 7 shows the scatterplot of measured vs. predicted chlorophyll content of the optimal model. The validation set R^2 was above 0.8677, indicating the robustness of the model.

Discussion

Compared with conventional chemical methods, spectroscopy is faster, more chemical-free, easier to use and non-destructive. Traditional spectral analysis generally uses full spectral modeling^[16]. However, the full spectrum contains a good deal of useless information or interference variables, which not only increase the model's complexity but also reduces its prediction performance. After pre-processing of the spectrum in this test, the model accuracy of SG-FD, SG-SD, SNV and MSC was greater than that of SG-Raw, which was generally consistent with the results of combining SNV and FD studied by Jie L et al.^[17] to eliminate the interference of factors such as background noise and baseline drift and to highlight the effective information of the spectrum to improve the model's discrimination accuracy. Based on spectral pre-processing, the selection of SMC and VIP variables greatly improves the prediction accuracy of the *Picea koraiensis* Nakai pigment model and reduces the RMSE value. This is consistent with the results of the Xiangzhong^[18] trial. Although some scholars have cited neural networks in spectral analysis modeling and discriminant analysis to improve the prediction accuracy and fine classification of plant physiological and biochemical index models^{[19], [20], [21]}, there are few comparative studies on the accuracy of BP neural network models and PLSR models. In this experiment, it is verified that the accuracy of the BP neural network model is higher than that of the PLSR model, which is generally consistent with the results of Jingyang et al.^[22] who used the nonlinear mapping relationship of the BP neural network model to treat variables better than the linear relationship of the multiple linear regression model.

It was found that the fit of the needle pigment content model with different specifications and

different leaf ages was much higher than that of the mixed *Picea koraiensis* Nakai model, which may be related to the dispersion of pigment content, indicating that the model of fine classification and fitting of needles is helpful to improve the robustness of the model. Whether it is the spectral pre-processing model, the variable selection model or the PLSR or BP neural network model, the fitting accuracy of the chlorophyll model of triennial conifers is always higher than that of the annual and biennial needles. Yingzi et al.^[23] showed that within a certain range, the greater the leaf age, the higher the pigment content. Analysis of chlorophyll content at different leaf ages in *Picea koraiensis* Nakai revealed that the dispersion of its chlorophyll content decreased with increasing leaf age ($p < 0.05$), which may directly affect the fitting accuracy of the model. Therefore, the high fitting accuracy of the triennial needle model may be closely related to the level and dispersion of pigment content. In summary, the model of fine classification fitting of needles helps to improve the robustness of the model.

Conclusion

Both the PLSR and BP neural network models can rapidly predict the chlorophyll content of the *Picea koraiensis* Nakai needle leaf, but the BP neural network model has higher accuracy and a lower RMSE.

(1) Spectral pre-processing can avoid systematic errors and eliminate background values, but it has little effect on improving model accuracy. After the PLSR model was processed by SMC and VIP variable selection, the model accuracy was greatly improved and the RMSE value was reduced.

(2) Compared with the PLSR model, the BP neural network model had higher accuracy, with R^2 above 0.95, and the validation set R^2 above 0.86.

(3) The fit of the models with different sizes and different leaf ages for needle chlorophyll content was much higher than that of the mixed *Picea koraiensis* Nakai model, indicating that the pigment model with fine classification and fitting of needles improves the robustness of the model. The fitting accuracy of the spectral pre-processing model, variable selection model, PLSR model and BP neural network model of triennial needles is always higher than for annual and biennial needles.

Pigment content is an evaluation index for studying plant metabolism, stress resistance and ornamental characteristics. In this experiment, a highly robust chlorophyll model was established by the BP neural network model to provide a statistical method basis and theoretical support for the use of a combination of remote sensing technology and stoichiometry methods to predict plant physiological and biochemical indicators in the future.

Methods

Materials

In this study, *Picea koraiensis* Nakai from Mopanshan Practice Base and the Botanical Garden of Beihua University was used as the experimental material. Two sizes of ground diameter (20–25 cm and 25–30 cm) plants were selected for collection at both sites, and annual, biennial and triennial healthy undamaged needles were collected with high branch scissors from four orientations: upper, middle and lower canopies as well as East, South, West and North. Each plant was sampled in triplicate.

Methods

Determination of Conifer Pigment Content

We took 0.3 g needle leaves of *Picea koraiensis* Nakai, cut them into pieces and put them into

a mortar, added a small amount of quartz sand and 95% ethanol, ground them into homogenate, transferred them to a test tube and added 95% ethanol to constant volume of 10 ml. We then sealed the tubes and placed them in the dark to soak for 24 h. Then we performed filtration, using an Agilent Cary60 UV-Visible spectrophotometer to measure the absorbance of chlorophyll extract at 649 nm and 665 nm.

$$\text{Chlorophyll a} = (13.95D_{665} - 6.88D_{649}) V/M/1000$$

$$\text{Chlorophyll b} = (24.96D_{649} - 7.32D_{665}) V/M/1000$$

$$\text{Chlorophyll} = \text{Chlorophyll a} + \text{Chlorophyll b}$$

Spectrum Acquisition and Processing Method

Needle reflectance was measured using a Specim handheld intelligent AISA hyperspectral camera with a wavelength range of 400 – 1000 nm, spectral resolution of 7 nm, image resolution of 512 × 512 pixels and 204 spectral bands. We adjusted the height of spectral camera frame to 130 cm and the light height to 180 cm, selected “simultaneous” mode in the WR method, set up a new white board, corrected the integration time, and collected the spectral data.

The Kennard-Stone method was used to model 80% of the samples as the modeling dataset and 20% as the validation set. Five methods of spectral preprocessing were used: SG-Raw, SG-FD, SG-SD, SNV and MSC.

The formula for calculating the first derivative of spectrum is $R'_{\lambda_n} = \frac{R_{\lambda_{n+1}} - R_{\lambda_n}}{\Delta\lambda}$. (1)

In Equation (1), R'_{λ_n} is the first derivative of the spectrum in the band from n to $n + 1$, and $R_{\lambda_{n+1}}$ and R_{λ_n} are the original spectral reflectance values at $n + 1$, n , respectively. $\Delta\lambda$ denotes the interval from wavelength λ_{n-1} to λ_n .

The formula for calculating the second derivative of spectrum is $R''_{\lambda_n} = \frac{R'_{\lambda_{n+1}} - R'_{\lambda_n}}{\Delta\lambda}$. (2)

In Equation (2), R''_{λ_n} is the second derivative of the spectrum in the band from n to $n + 1$, and $R'_{\lambda_{n+1}}$ and R'_{λ_n} are the original spectral reflectance values at $n + 1$, n , respectively. $\Delta\lambda$ denotes the interval from wavelength λ_{n-1} to λ_n . Differential processing can eliminate the influence of system error and background noise on spectral values.

The formula for calculating SNV is $SNV_i = \frac{x_i - \bar{x}_i}{s_i}$. (3)

In Equation (3), x_i is the spectral reflectance of the i th observation, \bar{x}_i is the mean, and s_i is the standard deviation. SNV standardizes each spectrum and eliminates the effect of dimensional effects and the size of the variable's own variation, that is, the size of the value^[24].

The formula for calculating MSC is $MSC_{x_i} = \frac{a_i - (m_i x_r + a_i)}{m_i}$. (4)

In Equation (4), x_i is the spectral reflectance of the i th observed value, x_r is the average of all spectral data as the ideal spectral value, and a_i and m_i are constant terms. MSC eliminates scattering levels resulting in spectral differences and corrects the baseline translation and offset phenomena of spectral data by ideal spectra.

Model Construction and Validation

Partial Least Squares Regression

PLSR combines the measured individual chemical components with reflection spectra for model calibration^[25], and finds the best function of the data for matching by minimizing the sum of squares of the errors^[26]. PLSR, which is widely used in modeling statistical analysis.

The PLSR pairs m independent variables ($x_1, x_2, x_3, \dots, x_m$) and q dependent variables ($y_1, y_2, y_3, \dots, y_q$), and then normalizes them. Then the linear combination of the first components t_1

and u_1 is extracted from it to establish the initial variable equation about T1:

$$y_0 = t_1 \beta_1' + f_1. \quad (5)$$

In Equation (5), β_1' is the parameter vector at time T1 of the independent variable, and f_1 is an $n \times q$ residual matrix. If the first component combination extracted cannot meet the model accuracy requirements, the above steps are continued to extract the components from the independent variables ($t_2, t_3, t_4, \dots, t_k$).

$$t_k = w_{k1}x_1 + w_{k2}x_2 + \dots + w_{kn}x_n \quad (6)$$

$$y_i = t_1\beta_{1i} + t_2\beta_{2i} + \dots + t_k\beta_{ki} \quad (i = 1, 2, 3, \dots, q) \quad (7)$$

Bring Equation (6) into Equation (7) to obtain the final PLSR equation

$$y_j = a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jm}x_m. \quad (8)$$

To avoid the problems of low accuracy and poor robustness of the PLSR fitting model, the PLSR model was subjected to variable selection to screen feature bands for improving model performance. SMC can reduce the impact of unrelated variables on the model and highlight related variables to improve model accuracy^[27]. VIP is used for band optimization^[28].

BP Neural Network Algorithm

The BP neural network algorithm can perform the complex and nonlinear mapping function, predict the nonlinear function approximation and other problems, and has the ability of self-learning promotion and generalization^[29]. It uses the square of network error as the objective function, uses the gradient descent method, adjusts the network weight value and threshold through back propagation, and calculates the minimum value of the objective function. The BP neural network, shown in Fig. 1, includes the forward transmission of signal and the reverse propagation of error. The forward transmission includes three layers. Each

layer of neuron affects only the next layer of state. If the output of the trained neural network cannot obtain the expected output value, the reverse propagation process of error is performed. The weight and threshold are adjusted layer by layer until the expected error range is reached.

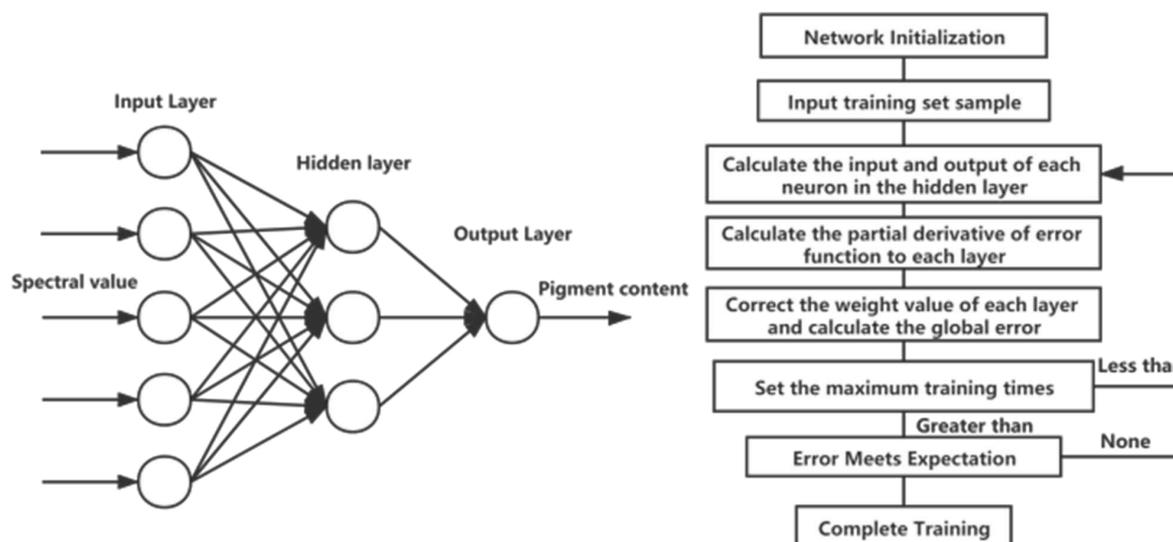


Figure 1. Topology diagram of BP neural network model (left) and flowchart of learning algorithm (right)

Data Analysis

R was used for data collation, model establishment and drawing analysis. The prospectr package^[30] was used for spectral processing and KS sample^[31] division, and the spectral curve was calculated with Savitzky-Golay filtering^[32] with a window of 11; the pls package^[33] was used for leave-one-out cross-validation and PLSR model fitting; the plsvartr package^[34] was used to select SMC and VIP variables for the PLSR model and to extract important variables affecting the model; the grid, MASS and neuralnet package^[35] were used for BP neural network model fitting; and the ggplot2 package was used for chart drawing.

Abbreviations

PLSR: partial least squares regression; SG-Raw: Savitzky-Golay Raw Spectra; SG-FD: Savitzky-Golay First Derivative; SG-SD: Savitzky-Golay Second Derivative; SNV: Standard

Normal Variate; MSC: Multiplicative Scatter Correction; VIP: Variable Importance in Projection; R^2_{CV} : coefficient of determination for cross-validation; R^2_V : coefficient of determination for validation; $RMSE_{CV}$: cross-validation; $RMSE_V$: root mean square error of validation.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

The authors gratefully acknowledge the funding from the Research and Demonstration of Efficient Production Technology for High-quality Seedlings of *Picea koraiensis* Nakai (2017YFD0600606-08).

Authors' contributions

YW designed the study, participated in the experiment and wrote the manuscript. JW and YX supervised the experiments and reviewed the manuscript. ZW and KS supported the data

collection and field experiment. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Tianrong L. The 8th Forest Resource Inventory in China[J]. Paper and Paper-making,2014, 33(04):84.
2. Lei Z, Hui W, Shuli W. Study on nutrient contents and ecological stoichiometric characteristics in needles of *Picea koraiensis* in different stands[J]. Journal of Plant Resources and Environment,2020, 29(03):19-25+33.
3. Miraglio T, Adeline K, Huesca M, et al. Monitoring LAI, Chlorophylls, and Carotenoids Content of a Woodland Savanna Using Hyperspectral Imagery and 3D Radiative Transfer Modeling[J]. Remote Sensing,2019, 12(1):28.
4. Ma S, Baldocchi DD, Mambelli S, et al. Are temporal variations of leaf traits responsible for seasonal and inter-annual variability in ecosystem CO₂ exchange?[J]. Functional Ecology,2011, 25(1):258-270.
5. Shipley B, Lechowicz MJ, Wright I, et al. FUNDAMENTAL TRADE-OFFS GENERATING THE WORLDWIDE LEAF ECONOMICS SPECTRUM[J]. Ecology,2006, 87(3):535-547.
6. Qingxi T, Bing Z, Lifu Z. Current progress of hyperspectral remote sensing in China[J]. Journal of Remote Sensing,2016, 20(05):689-707.

-
7. Wen Z, Jin M, Mengke Y, et al. Hyperspectral estimation of rice pigment content based on band depth analysis and BP neural network[J]. Chinese Journal of Eco-Agriculture,2017, 25(08):1224-1235.
 8. Yuling Z, Nannan Y, Haixia Z, et al. Study on the Statistical Estimation Model of Soil Heavy Metals in Handan City Based on Hyper-spectral[J]. Ecology and Environmental Sciences,2020, 29(04):819-826.
 9. Yanjie L, M AC. Effects of variable selection and processing of NIR and ATR-IR spectra on the prediction of extractive content in Eucalyptus bosistoana heartwood[J]. Spectrochimica acta Part A, Molecular and biomolecular spectroscopy,2019, 213:111-117.
 10. Garhwal AS, Pullanagari RR, Li M, et al. Hyperspectral imaging for identification of Zebra Chip disease in potatoes[J]. Biosystems Engineering,2020, 197:306-317.
 11. Xinmei W, Jinsong Z, Ping M, et al. Estimation of Nitrogen Content in Walnut Canopy Based on GS-2 Hyperspectral Image[J]. Transactions of the Chinese Society for Agricultural Machinery,2020:1-16.
 12. Divya Y, Gopinathan P. Soil water content measurement using hyper-spectral remote sensing techniques – A case study from north-western part of Tamil Nadu, India[J]. Remote Sensing Applications: Society and Environment,2019, 14:1-7.
 13. Fedotov YV, Belov ML, Kravtsov DA, et al. Selection of emission detection ranges for the laser method of plant stress revealing at a fluorescence excitation wavelength of 355 nm[J]. IOP Conference Series: Materials Science and Engineering,2018, 450(6):062005.

-
14. Nuo L, Pei W, Edou PCA, et al. Leaf morphology and photosynthetic characteristics at different leaf ages in Chinese fir[J]. *Subtropical Agriculture Research*,2018, 14(03):167-171.
 15. Yanjie L, Xin D, Yang S, et al. An efficient method to reduce grain angle influence on NIR spectra for predicting extractives content from heartwood stem cores of *Toona sinensis*[J]. *Plant methods*,2020, 16:77.
 16. Marena M. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials[J]. *Chemical Society reviews*,2014, 43(24):8200.
 17. Jie L, Shangke L, Liwen J, et al. A Nondestructive Method Identifying Varieties of Green Tea Based on Near Infrared Spectroscopy and Chemometrics[J]. *Journal of Instrumental Analysis*,2020, 39(11):1344-1350.
 18. Xiangzhong S, Guo T, Luda Z, et al. Research Advance of Variable Selection Algorithms in Near Infrared Spectroscopy Analysis[J]. *Spectroscopy and Spectral Analysis*,2017, 37(04):1048-1052.
 19. Tongyu X, Weinan Y, Yuncheng Z, et al. Estimation of Dicotyledon Leaf Chlorophyll Value Based on Hyperspectrum and BP Neural Network[J]. *Journal of Shenyang Agricultural University*,2018, 49(06):749-755.
 20. Huang F, Zhang J, Shan Q, et al. The research of the plant leaves identification method based on 3-layers BP neutral network[J]. *Cluster Computing*,2019, 22(5):11143-11152.
 21. Xuelian W, Yuzhi S, Fanfan K, et al. Estimation of Chlorophyll-a Concentration in Taihu Lake by Using Back Propagation (BP) Neural Network Forecast Model[J]. *Chinese Journal of Agrometeorology*,2016, 37(04):408-414.

-
22. Jingyang Z, Guangyou P. Comparison and Application of Multiple Regression and BP Neural Network Prediction Model[J]. Journal of Kunming University of Science and Technology(Natural Science),2013, 38(06):61-67.
 23. Yingzi W, Wei H, Chengzhen W, et al. Variation of Chloroplast Pigments of Various Age Leaves along Elevation Gradients of Dominant Species in *Castanopsis carlessi* Forests in Lingshishan National Forest Park[J]. Scientia Silvae Sinicae,2010, 46(11):43-51.
 24. Yuqiang L, Tianhong P, Haoran L, et al. NIR Spectral Feature Selection Using Lasso Method and Its Application in the Classification Analysis[J]. Spectroscopy and Spectral Analysis,2019, 39(12):3809-3815.
 25. Yanjie L, Yang S, Jingmin J, et al. Spectroscopic determination of leaf chlorophyll content and color for genetic selection on *Sassafras tzumu*[J]. Plant methods,2019, 15:73.
 26. Sicheng Y, Zaixi S, Yang C. Identification of Rice Varieties Based on Hyperspectral Image[J]. Spectroscopy and Spectral Analysis,2019, 39(10):3273-3280.
 27. Tran TN, Afanador NL, Buydens LMC, et al. Interpretation of variable importance in Partial Least Squares with Significance Multivariate Correlation (sMC)[J]. Chemometrics and Intelligent Laboratory Systems,2014, 138:153-160.
 28. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics[J]. Chemometrics and Intelligent Laboratory Systems,2001, 58(2):109-130.
 29. Keiner LE, Yan X-H. A Neural Network Model for Estimating Sea Surface Chlorophyll and Sediments from Thematic Mapper Imagery[J]. Remote Sensing of Environment,1998, 66(2):153-165.

-
30. Mullen KM, Stokkum IHMv. An Introduction to the `Special Volume Spectroscopy and Chemometrics in R`[J]. *Journal of Statistical Software*,2007, 18(1).
 31. Kennard R W, Stone L A. Computer Aided Design of Experiments[J]. *Technometrics*,1969, 11(1):137-148.
 32. J S, Y T, J D. Smoothing and differentiation of data by simplified least square procedure[J]. *Analytical chemistry*,1972, 44(11):1906-1909.
 33. Rogers MS, Chung T, Fergusson S. Oscillometric Measurement of Mean Arterial Pressure: The Effect of Position on Prediction of Pregnancy Induced Hypertension[J]. *M S Rogers;T Chung;Sally Fergusson*,2009, 15(4):221-225.
 34. Mehmood T, Liland KH, Snipen L, et al. A review of variable selection methods in Partial Least Squares Regression[J]. *Chemometrics and Intelligent Laboratory Systems*,2012, 118:62-69.
 35. Anastasiadis AD, Magoulas GD, Vrahatis MN. New globally convergent training scheme based on the resilient propagation algorithm[J]. *Neurocomputing*,2004, 64:253-270.

Figures

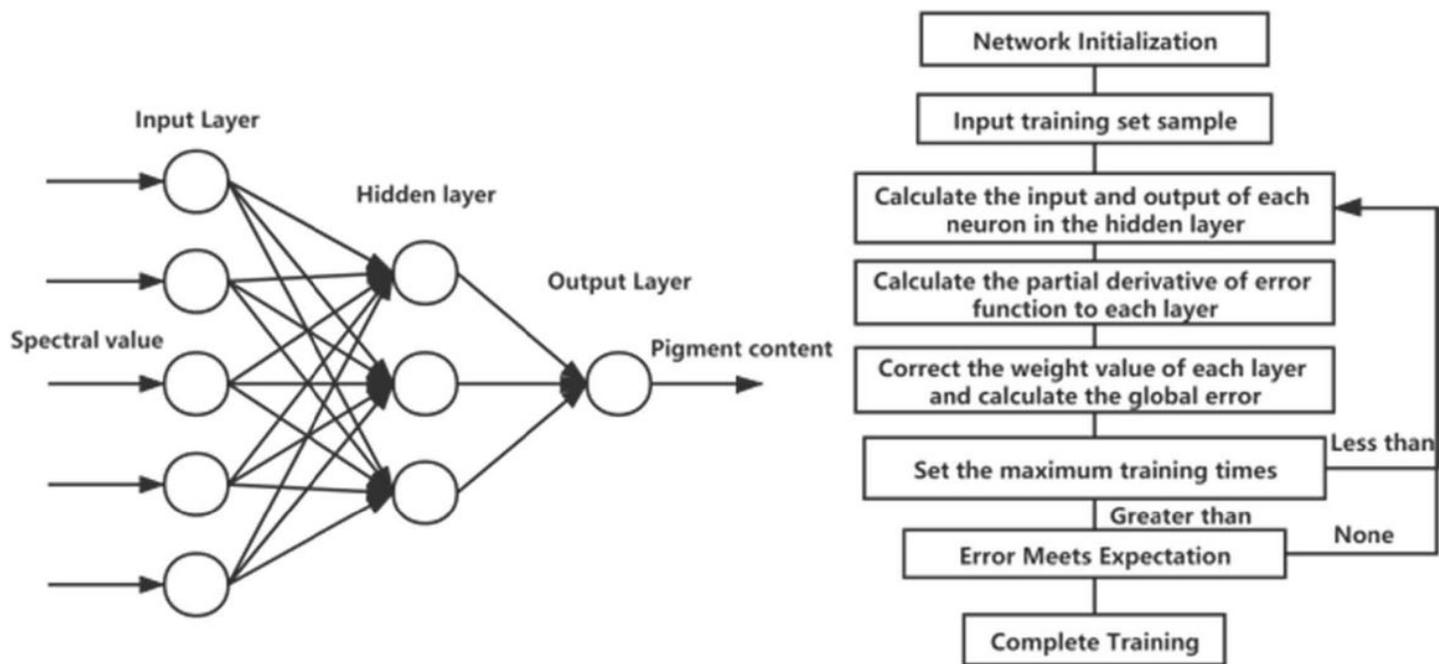


Figure 1

Topology diagram of BP neural network model (left) and flowchart of learning algorithm (right)

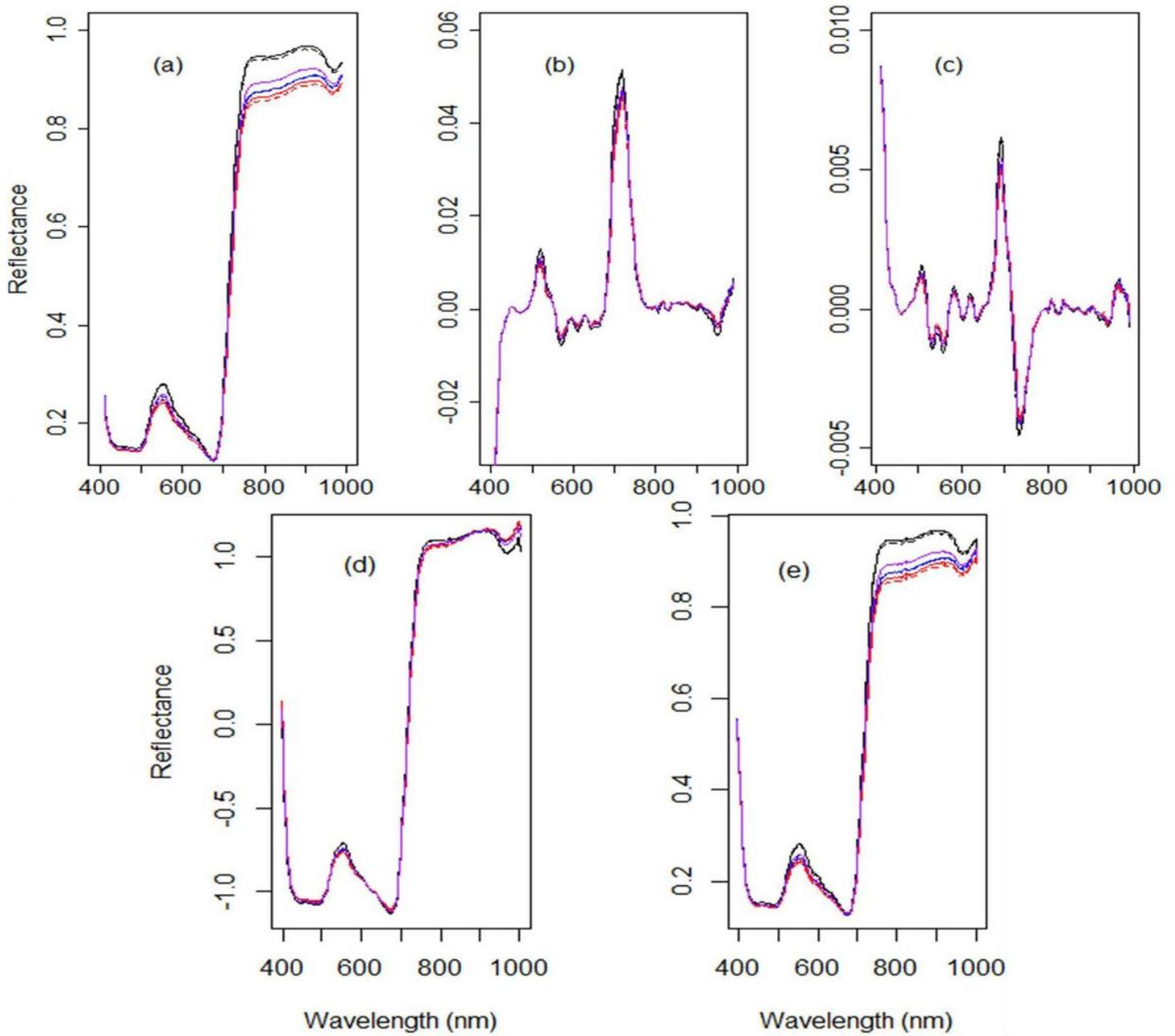


Figure 2

Spectral preprocessing curve; black: annual needles, blue: biennial needles, red: triennial needles, purple: mixed *Picea koraiensis* Nakai; solid line: *Picea koraiensis* Nakai with a diameter of 25–30 cm, dashed line: *Picea koraiensis* Nakai with a diameter of 20–25 cm; (a): SG-Raw, (b): SG-FD, (c): SG-SD, (d): SNV, (e): MSC.

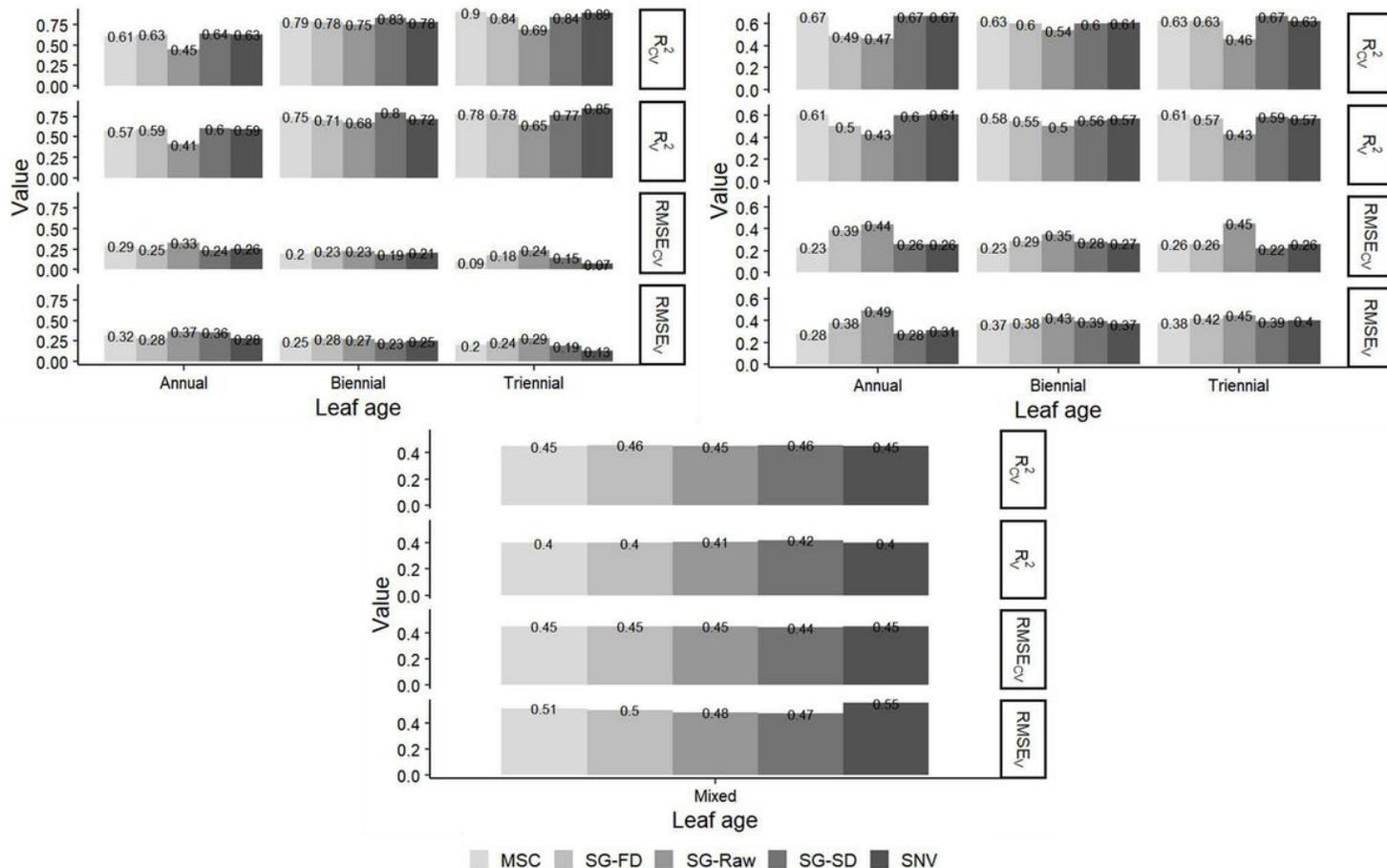


Figure 3

Fitting and evaluation of chlorophyll PLSR model for spectral pretreatment of needles of *Picea koraiensis* Nakai; upper left: 25 – 30 cm ground diameter *Picea koraiensis* Nakai, upper right: 20 – 25 cm ground diameter *Picea koraiensis* Nakai, lower: mixed *Picea koraiensis* Nakai; Annual: annual conifers, Biennial: biennial conifers, Triennial: triennial conifers, Mixed: mixed *Picea koraiensis* Nakai conifers.

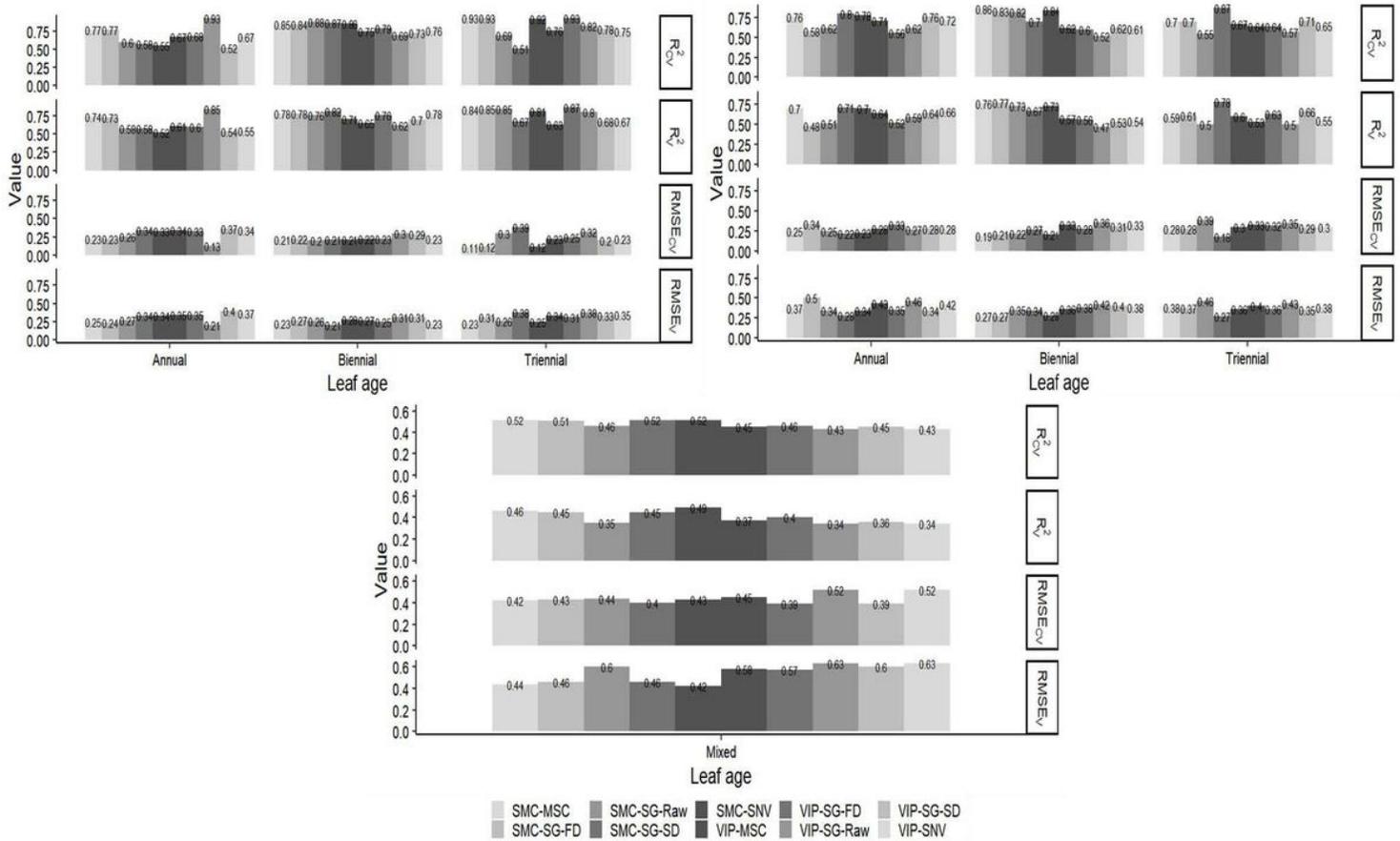


Figure 4

Fitting and evaluation of chlorophyll PLSR model spectra pretreatment and variable selection of *Picea koraiensis* Nakai; upper left: 25 – 30 cm ground diameter *Picea koraiensis* Nakai, upper right: 20 – 25 cm ground diameter *Picea koraiensis* Nakai, lower: mixed *Picea koraiensis* Nakai.

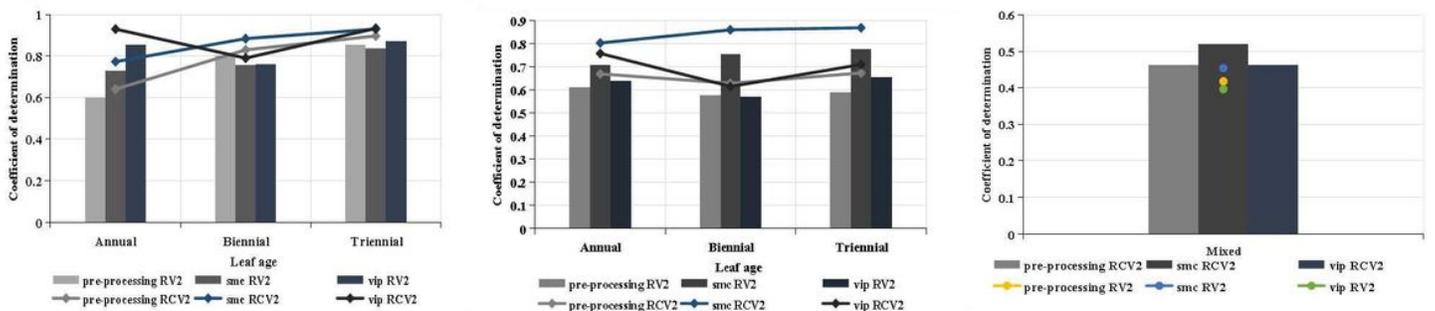


Figure 5

Comparison of spectral preprocessing and variables selection optimal model R^2 of different leaf ages and mixed *Picea koraiensis* Nakai in PLSR model; left: 25–30 cm ground diameter *Picea koraiensis* Nakai, middle: 20–25 cm ground diameter *Picea koraiensis* Nakai, right: mixed *Picea koraiensis* Nakai.

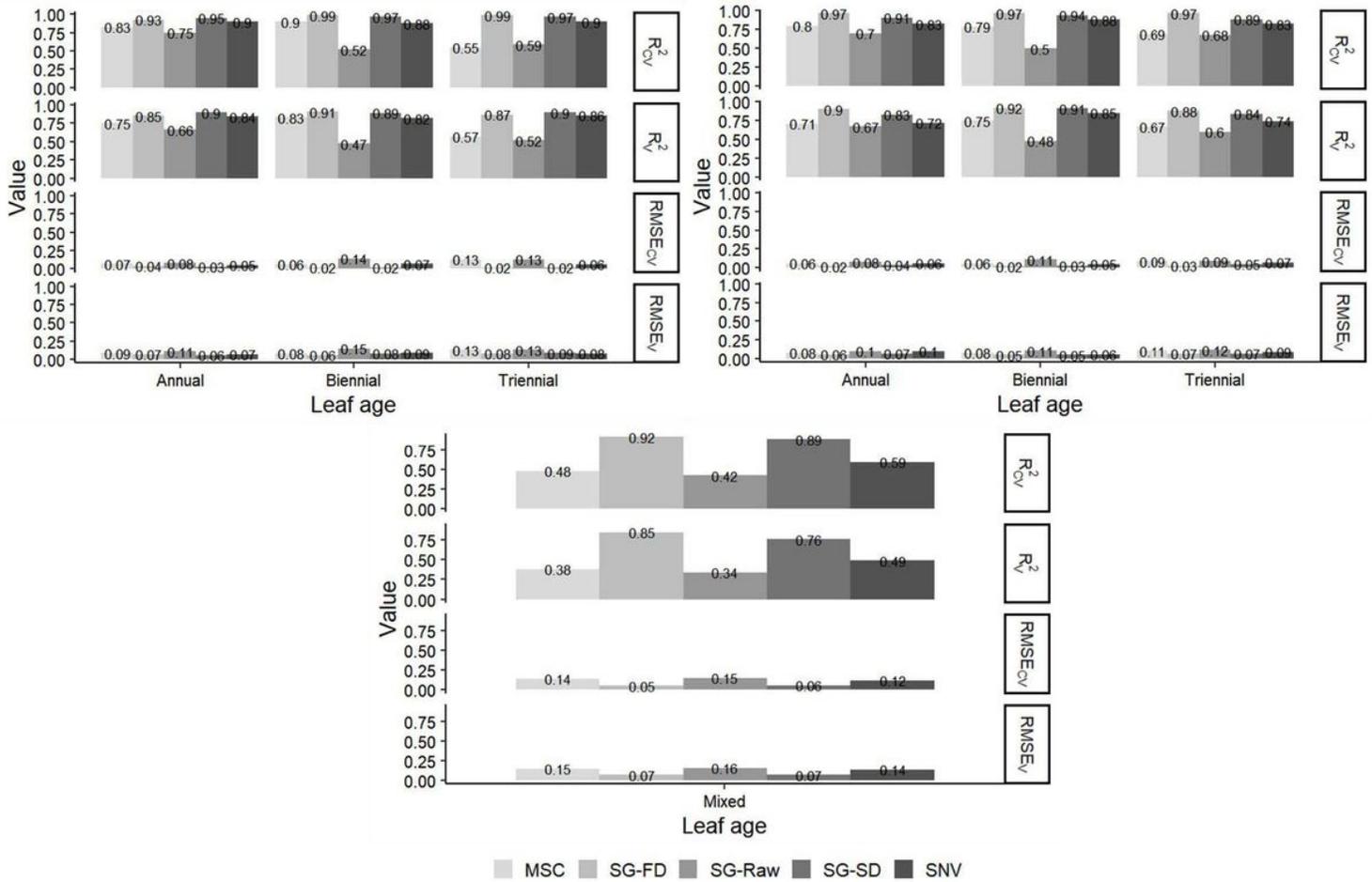


Figure 6

Fitting and evaluation of BP neural network chlorophyll model of *Picea koraiensis* Nakai; upper left: 25 – 30 cm ground diameter *Picea koraiensis* Nakai, upper right: 20 – 25 cm ground diameter *Picea koraiensis* Nakai, lower: mixed *Picea koraiensis* Nakai.

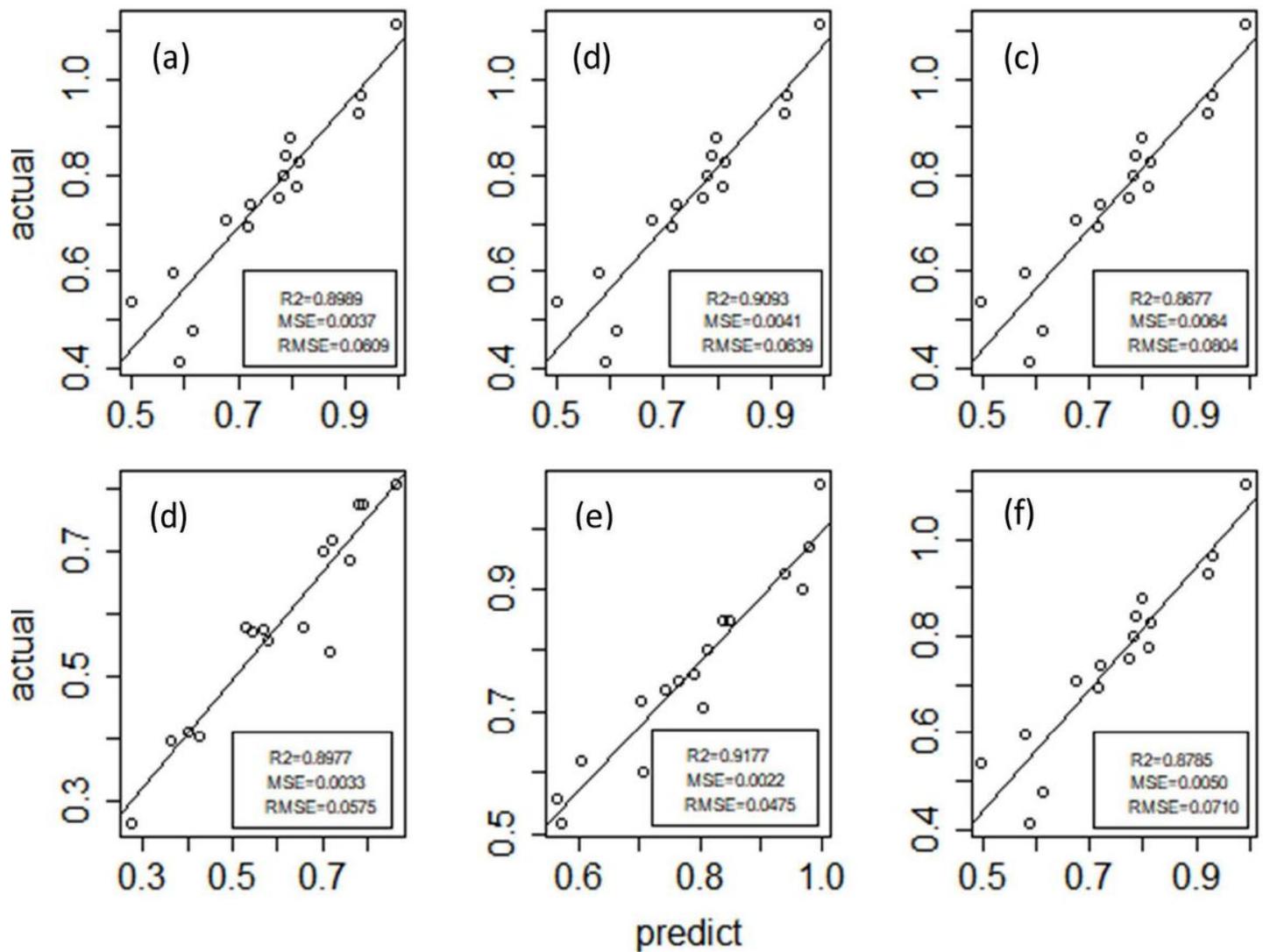


Figure 7

Scatter plot of the fit between the measured and predicted values of the BP neural network chlorophyll model. (a), (b), and (c) are annual, biennial and triennial needles of 25–30 cm ground diameter *Picea koraiensis* Nakai, respectively, and (d), (e), and (f) are annual, biennial and triennial needles of 20–25 cm ground diameter *Picea koraiensis* Nakai..