Near-infrared Spectroscopy Detection Method for Compressive Strength of Fraxinus mandschurica

Hao Liang, Linyin Xing, Jian Wen, Chao Gao, and Jianhui Lin

Abstract—This study used near-infrared (NIR) spectroscopy as a non-destructive test to predict the compressive strength (i.e., modulus of rupture (MOR) and the modulus of elasticity (MOE)) of Fraxinus mandschurica parallel to the wood grain. Tests were conducted with 120 small and clear wood samples to obtain the diffuse NIR reflectance spectra of the radial and tangent surfaces of the wood samples. Standard normal variable transformation (SNV) combined with Savitzky-Golay (SG) convolution smoothing algorithm was used to filter the raw NIR spectra. Uninformative variables elimination (UVE) and a genetic algorithm (GA) were utilized to identify specific wavelengths in the spectra that directly correlated to compression strength. Finally, a partial least squares (PLS) regression model was developed with the identified wavelengths to determine the MOR and MOE of the samples. The results showed the correlation coefficients of the prediction models for MOR and MOE were 0.88 and 0.89, respectively. The root mean square errors of prediction for MOR and MOE models were 0.37 and 0.49, respectively. Based on these results, it is feasible to accurately estimate the compressive strength of Fraxinus mandschurica (parallel to the grain) using NIR spectroscopy.

Keywords—Near-infrared spectroscopy; Compression strength (parallel to wood grain), Uninformative variables elimination, Genetic algorithm.

I. INTRODUCTION

Fraxinus mandschurica, a commonly used structural material, requires a high degree of structural performance and reliability, particularly with its compressive strength parallel to the wood grain. Traditional testing of wood compressive strength is to conduct destructive tests on small and flawless wood samples using a universal testing machine in accordance to standardized laboratory protocols. This method is accurate and reliable, but the sample preparation is cumbersome and the amount of testing time required is high. It cannot meet the actual needs of the forestry industry and the wood processing industry. Therefore, a non-destructive laboratory test method that can measure the mechanical properties of wood would have important application value and practical benefits to engineers.

Near-infrared (NIR) spectroscopy is a fast, non-destructive, and indirect analysis technology that has been widely used with some success in the areas of agriculture, food, medicine, paper testing, petroleum processing, and winemaking, in addition to other fields. In recent years, the application of NIR spectroscopy in the wood sciences has become increasingly extensive. This tool has been used to estimate the lignin and cellulose content in trees, and to analyze the mechanical, physical and chemical properties of wood [1][2][3]. Thumm and Meder used NIR to assess the stiffness of dry radiata pine clearwood and demonstrated that when the load is applied to the radial face then NIR spectra obtained from the radial face are preferred to that obtained from the tangential face, due to spectral information being obtained from both latewood and earlywood. Tong and Zhang estimated the mechanical properties of thermally-modified softwood (southern pine) using NIR; the authors observed a close relationship between the NIR spectral peaks and the mechanical properties of the wood. Eom et al. measured the surface moisture content of yellow poplar in real-time using a NIR technique. Moreover, NIR has been used by several investigators to detect wood surface defects [4][5], as well as to classify the species and origins of the wood specimens [6][7].

In this study, a fast and non-destructive testing method for measuring the compressive strength of Fraxinus mandschurica (parallel to the wood grain) was developed using near-infrared spectroscopy. First, a standard normal variable transformation combined with the Savitzky-Golay convolution smoothing algorithm was used to filter the collected NIR absorption spectra. Then, uninformative variables elimination (UVE) and genetic algorithm (GA) analyses were utilized on the recorded spectra to identify specific NIR wavelengths that correlated with the compressive strength of Fraxinus mandschurica. Finally, a calibration and a prediction model for the compressive strength (i.e., modulus of rupture (MOR) and the modulus of elasticity (MOE)) were developed using a partial least squares (PLS) regression algorithm. Through these analyses and evaluations of the models, it was deduced that there is a close correlation...
between the NIR spectra and the compressive MOR and MOE values for *Fraxinus mandshurica*. The main purpose of this study is to develop calibrations for determining the compressive wood strength (parallel to the grain) using NIR spectroscopy, and to evaluate the predictive ability of NIR spectra with calibrations.

II. EXPERIMENTAL

A. Sample Preparation

*Fraxinus mandshurica* trees grown in Northeast China (44°37' to 44°47' N, 27°35' to 127°55' E) were selected as test samples. Fifteen 20-year-old trees were felled, and the logs were cut at a height of 1.3 m. The logs were then air-dried. After the process of drying, sawing, and sanding, 120 flawless boards with the dimensions of 30 mm (L) × 20 mm (T) × 20 mm (R) were fashioned in accordance to Chinese Standard GB/T 1929. The specimens were numbered from 1 to 120. These labeled specimens were stored at 22 ± 1 °C and 45% ± 5% relative humidity prior to NIR scanning and to compressive strength testing.

B. Determination of Compressive Strength Parallel to the Wood Grain

The MOR and MOE of the specimens were measured in accordance to Chinese Standard GB/T 1935 by applying pressure to the board at a uniform rate (10 mm/min) until rupture (2 to 3 min per specimen) parallel to the wood grain. A total of 120 specimens were divided into two groups: a calibration set (80 specimens) and a prediction set (40 specimens). The specimens with the highest and lowest compressive strength of all the specimens were placed into the calibration set, and the rest of the samples were randomly divided into the calibration set and prediction set. In the experiments conducted, 80 specimens of the calibration set were used to establish the calibration model, and the remaining 40 samples of the prediction set were used to externally validate the model.

C. NIR Spectra Acquisition

It was found that the spectra with wavelength ranging from 1000 to 1600 nm carry important information and can better predict the mechanical properties, density, and other properties of wood [8][9]. A one-chip micro integrated optic fiber spectrometer by INSION Co. GmbH (Heilbronn, Germany) was used to record the NIR spectra for the specimens. The recorded wavelength range was 900 to 1800 nm with a 9 nm resolution (i.e., thermal wavelength stability was less than 0.03 nm/K). The surface NIR stability of the recorded measurements; the instrument was turned on and allowed to warm-up for 10 min to ensure the stability of the recorded measurements; the instrument was calibrated using a commercial PTFE reference tile. Then, the optical fiber probe (5-mm diameter) was fixed onto a bracket, and the specimens were placed below the probe at a gap distance of 1 mm. Each specimen area was automatically scanned 30 times and the collected spectra were averaged to yield a single NIR spectral curve. The spectrum acquisition set-up is shown in Fig 1. The growth characteristics of the wood led to different absorption peaks in the NIR spectra for different sections, but the spectral trends were similar. Consequently, the radial (R) and tangential (T) spectra of the specimens collected in this study were used for analysis and modeling after taking the averaged values.

![Diagram of spectra acquisition](image)

Fig. 1 The diagram of spectra acquisition.

III. METHODS

A. Spectra Data Filtering

In this study, SNV and SG convolution smoothing were used to filter the raw NIR spectra. The spectral $X_{i,k}$ that needed SNV transformation was calculated by (1),

$$X_{i,SNV} = \frac{X_{i,k} - \bar{X}_i}{\sqrt{\sum_{k=1}^{m}(X_{i,k} - \bar{X}_i)^2}}$$

where $k = 1 \ldots m$, $i = 1 \ldots n$

where $\bar{X}_i$ is the average of the spectra of the $i^{th}$ sample, $m$ is the
number of wavelength variables, and \( n \) is the number of specimens in the calibration set. The SG convolution smoothing is calculated as (2),

\[
d^k \left( \frac{x^{i+j}}{d(j)} \right)_{j=0} = k! a_k
\]

where \( x \) is the absorbance, \( i \) and \( j \) are ordinal numbers within the range of wavelength variables, \( k! \) is the factorial of the derivative order \( k \), and \( a_k \) is the weighting coefficient.

B. Spectral Feature Extraction

Due to the wide range of wavelengths contained in the recorded NIR spectra during the experiments, some wavelengths have no correlation or relationship to the compressive strength (i.e., MOR and MOE) of the samples. Additionally, there was collinearity of the NIR spectra that resulted in redundant information that negatively affected the accuracy of the developed regression models. Therefore, it was necessary to eliminate the uninformative wavelengths and to extract the relevant wavelengths in the spectra for regression modeling.

The UVE method was implemented to eliminate the wavelengths in the NIR spectra not related to the compressive strength of the specimens, as well as to remove the influence of various non-target factors [13][14][15].

After data processing the NIR spectra with the UVE method, the genetic algorithm (GA) was then used on the refined data to extract the specific wavelengths related to MOR and MOE, as well as to eliminate irrelevant wavelengths from consideration. Moreover, the process of spectral feature extraction by the combined UVE and GA analyses reduced the complexity of the developed model, which improved the predictive accuracy and stability of the calibration model.

C. Model Evaluation Standard

After identifying the relevant NIR wavelengths, a partial least squares (PLS) regression model for the compressive strength (parallel to the wood grain) of *Fraxinus mandshurica* was developed. The regression model factor was determined by a cross-validation method, and the model was validated by the specimens of the prediction set.

The predictability of the regression model was quantified using various statistical analyses: the coefficient of determination of the calibration model (\( R^2_c \)); the coefficient of determination of the prediction model (\( R^2_p \)); the root mean square error of calibration (RMSEC); the root mean square error of prediction (RMSEP); and the relative percent deviation (RPD). The selection of the optimal model was based on its predictability following a procedure described by Gierlinger et al. (2002). A better prediction model generally had larger \( R^2_p \) and \( R^2_p \) values, as well as smaller RMSEC and RMSEP values. Furthermore, if the RPD of a prediction model was between 2.5 and 3.0, the regression model was deemed to have good prediction accuracy.

### Table 1. Measured Compressive Strengths of Specimens (Parallel-to-Grain) of the Data Sets

| Data Set        | Calibration set | Prediction set | All samples |
|-----------------|-----------------|----------------|-------------|
| Number          | 80              | 40             | 120         |
| MOR (MPa)       |                 |                |             |
| Max             | 263.93          | 255.02         | 263.93      |
| Min             | 157.92          | 183.5          | 157.92      |
| Mean            | 211.53          | 206.54         | 209.86      |
| St. Dev         | 25.63           | 21.29          | 23.92       |
| MOE (GPa)       |                 |                |             |
| Max             | 22.34           | 21.09          | 22.34       |
| Min             | 15.83           | 16.17          | 15.83       |
| Mean            | 19.47           | 18.88          | 19.28       |
| St. Dev         | 1.65            | 1.37           | 1.58        |

B. Raw and Filtered NIR Spectra of the Specimens

The near-infrared spectra (900 to 1800 nm) of all 120 specimens are shown in Fig. 2(a). Some high-frequency noise, baseline drift, and other negative effects were observed in the spectra. Hence, the raw NIR data were filtered using SNV and SG convolution smoothing methods. The filtered spectra (Fig. 2(b) and 2(c)) had a clearer outline with more obvious absorption peaks and less noise than the raw data.
The PLS regression models were developed using the original spectra and the spectra filtered by various methods. The effects of different data filtering methods on the MOE modeling accuracy are illustrated in Table 2.

Table 2. Effects of Different Spectral Data Filtering Methods on MOE Modeling Accuracy

| Data Filtering Method | Raw Spectra | SNV | SG | SNV & SG |
|-----------------------|-------------|-----|----|----------|
| Calibration set       | $R_c^2$     | 0.67| 0.79| 0.67     | 0.80     |
|                       | RMSEC       | 0.91| 0.71| 0.90     | 0.70     |
| Prediction set        | $R_p^2$     | 0.64| 0.70| 0.63     | 0.74     |
|                       | RMSEP       | 0.97| 0.84| 0.96     | 0.81     |

The PLS regression model developed using the data filtered by both SNV and SG had the highest MOE prediction accuracy. The $R_c^2$ and $R_p^2$ values of this model were 0.80 and 0.74, respectively, which were higher than the models based on the raw spectra, the SNV filtered spectra, and SG filtered spectra. The RMSEC and RMSEP values in this model were 0.70 and 0.81, respectively, which were higher than the models based on the raw spectra, the SNV filtered spectra, and SG filtered spectra. It was observed that when the smoothing window of the SG method was 11, the filtering effect of the spectral noise was the best. Hence, subsequent raw NIR data were filtered using both SNV and SG.

C. Irrelevant Wavelength Elimination by Uninformative Variables Elimination

After spectral data filtering using both SNV and SG methods, the refined data were processed by the UVE method to eliminate irrelevant wavelengths to predict the specimens’ compressive strength values. The stability distributions of the variables after the UVE method are illustrated in Fig. 3. Figure 3 shows that the original full spectral wavelength variables and the introduced random variables were located on the left and right sides of the vertical line, respectively. The two dotted lines shown in the figure represent the upper and the lower limits for the threshold value of the selected variable. Stability values outside the dotted lines were set to an absolute value of 1, and the stability values inside the dotted lines were set to 0. The wavelength variables corresponding to the stability values within the two dotted lines were considered to be independent of the compressive wood strength. The UVE analysis of the filtered NIR spectra data eliminated 84 of the 117 wavelengths. As a result, 33 wavelengths were identified that were related to the compressive wood strength. The distributions of these correlative wavelengths are shown in Fig. 4. The blue circles indicate the wavelengths that correlated to the specimens’ compressive strength. Some of the retained wavelength regions may be attributed to specific chemical structures in the specimens that are associated with MOR and MOE.

D. Feature Wavelength Selection by Genetic Algorithm

The UVE data analysis reduced the irrelevant NIR wavelengths over the 900 to 1800 nm range from 117 to 33 candidates. The genetic algorithm (GA) method was employed to optimize how these chosen wavelengths related to the compressive strength of *Fraxinus mandshurica*. In this study, the root mean square error of cross-validation (RMSECV) with
partial least squares regression (PLSR) was used to develop the correlation models. The experiment repeated a total of 5 screening processes to take the intersection to obtain the final feature wavelength variables, and six wavelength variables (1482.67 nm, 1490.93 nm, 1507.46 nm, 1532.27 nm, 1623.31 nm, 1664.74 nm) were selected. Figure 5(a) shows the optimal points of the RMSECV curve. Figure 5(b) illustrates the selected feature wavelength variables in the spectra by UVE-GA.

It was deduced from the observations listed in Table 3 that the calibration model based on the UVE-GA approach increased the $R_c^2$ and decreased the RMSEC (relative to PLS only model), which indicated that the prediction performance of the model had been improved. The predictability of the UVE-GA-PLS model was the best of all the modeling approaches examined. The $R_c^2$ and RMSEC of the MOR calibration model based on this approach were 0.92 and 7.01, respectively, whereas the $R_c^2$ and RMSEC values were 0.92 and 0.52, respectively. Furthermore, the spectral dimension was reduced from 117 to 6 candidate wavelengths when both UVE and GA methods were used. This approach appreciably reduced the data dimension, which reduced the modeling time.

F. Predictability of the Regression Models

After the calibration model was developed, the prediction models for MOR and MOE based on PLS, UVE-PLS, GA-PLS, and UVE-GA-PLS approaches were examined to determine their ability to predict compression strength parameters accurately. The statistical analyses of the models to represent the measured results from the experimental set are shown in Table 4.

| Compressive Strength | Models       | $R_p^2$ | RMSEP | RPD  |
|----------------------|--------------|---------|-------|------|
| MOR                  | PLS          | 0.70    | 13.09 | 1.62 |
|                      | UVE-PLS      | 0.77    | 10.46 | 2.03 |
|                      | GA-PLS       | 0.75    | 9.97  | 2.14 |
|                      | UVE-GA-PLS   | 0.88    | 7.37  | 2.89 |
| MOE                  | PLS          | 0.74    | 0.81  | 1.69 |
|                      | UVE-PLS      | 0.74    | 0.58  | 2.36 |
|                      | GA-PLS       | 0.73    | 0.62  | 2.21 |
|                      | UVE-GA-PLS   | 0.89    | 0.49  | 2.80 |

The results of Table 4 indicated that the range of $R_p^2$, RMSEP, and RPD for the MOR prediction model was 0.70 to 0.88, 7.17 to 13.09, and 1.62 to 2.96, respectively, and the range of $R_c^2$, RMSEP, and RPD of MOE prediction model was 0.73 to 0.89, 0.47 to 0.81, and 1.69 to 2.91, respectively. The RPD of PLS, UVE-PLS, and GA-PLS were all lower than 2.5, which showed that their predictions of MOR and MOE based on selected NIR wavelength measurements were not acceptable. Among the models, the UVE-GA-PLS model had the highest $R_p^2$, the smallest RMSEP, and the highest RPD. Hence, the UVE-GA-PLS model should quantitatively predict the MOR and MOE of Fraxinus mandshurica wood samples; however, the accuracy of the model’s predictions needs to be improved.

V. CONCLUSIONS

This study reported the relationship between the NIR spectra and the compressive strength of Fraxinus mandshurica (parallel to the wood grain). The NIR technique can be used to estimate the MOR and MOE non-destructively.

After data filtering by SNV and SG convolution smoothing methods, the noise in the raw spectra was effectively eliminated.
Based on the filtered spectra, the wavelengths related to the compressive strength of *Fraxinus mandshurica* were identified by UVE and GA analyses. These approaches effectively reduced the spectral dimension and the computational complexity of the developed regression model while improving the model's accuracy.

The UVE-GA-PLS model for predicting the compressive strength of *Fraxinus mandshurica* was developed using candidate wavelengths from the data filtering methods. The correlation coefficient for the MOR and the MOE prediction model were 0.88 and 0.89, respectively; the RPD for this model was higher than 2.5.

Observed results showed that the model can predict the compressive strength of *Fraxinus mandshurica* samples (parallel to the wood grain) without having to conduct destructive sample testing with a universal testing machine; however, the accuracy of the developed models needed to be further improved.

**REFERENCES**

[1] Kelley, S. S., Rials, T. G., Snell, R., Groom, L. H., and Sluiter, A. “Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood,” Wood Science and Technology 38(4), 2004, 257-276.

[2] He, W., and Hu, H. “Rapid prediction of different wood species and lignin content using near infrared spectroscopy,” Journal of Wood Chemistry and Technology 33(1), 2013,52-64.

[3] Tsuchikawa, S., and Kobori, H. “A review of recent application of near infrared spectroscopy to wood science and technology,” Journal of Wood Science 61(3), 2015,213-220.

[4] Yang, Z., Zhang, M., Li, K., and Chen, L. “Rapid detection of knot defects on wood surface by near infrared spectroscopy coupled with partial least squares discriminant analysis,” BioResources 11(1),2016, 2557-2567.

[5] Cao, J., Liang, H., Lin, X., Tu, W., and Zhang, Y. “Potential of near-infrared spectroscopy to detect defects on the surface of solid wood boards,” BioResources 12(1),2017, 19-28.

[6] Adépie, O. E., Dawson-Andoh, B., Slahor, J., and Osbom, L. “Classification of red oak (Quercus rubra) and white oak (Quercus alba) wood using a near infrared spectrometer and soft independent modelling of class analogies,” Journal of Near Infrared Spectroscopy 16(1),2008, 49-57.

[7] Castillo, R., Contreras, D., Freer, J., Ruiz, J., and Valenzuela, S. “Supervised pattern recognition techniques for classification of Eucalyptus species from leaves NIR spectra,” Journal of Chilean Chemical Society 53(4), 2008,1709-1713.

[8] Schimleck, L. R., Mora, C., and Daniels, R. F. “Estimation of the physical wood properties of green, Pinus taeda, radial samples by near infrared spectroscopy,” Canadian Journal of Forest Research 33(12), 2003,2297-2305.

[9] Thumm, A., and Meder, R. “Stiffness prediction of radiata pine clearwood test pieces using near infrared spectroscopy,” Journal of Near Infrared Spectroscopy 9(2), 2001,117-122.

[10] Tong, L., and Zhang, W. “Using Fourier transform near-infrared spectroscopy to predict the mechanical properties of thermally modified southern pine wood,” Applied Spectroscopy 70(10), 2016,1676-1684.

[11] Yu, H., Liang, H., Lin, X., and Zhang, Y. “Nondestructive determination of the modulus of elasticity of *Fraxinus mandshurica* using near-infrared spectroscopy,” Optical Engineering 57(4), 046103,2018.

[12] Centner, V., and Massart, D. L. “Elimination of uninformative variables for multivariate calibration,” Analytical Chemistry 68(21), 1996,3851-3858.

[13] Shao, X. G., Wang, F., Chen, D., and Su, Q. D. “A method for near-infrared spectral calibration of complex plant samples with wavelet transform and elimination of uninformative variables,” Analytical and Bioanalytical Chemistry 378(5), 2004,1382-1387.

[14] Wu, D., Chen, X. J., Shi, P. Y., Wang, S. H., Feng, F. Q., and He, Y. “Determination of α-linolenic acid and linoleic acid in edible oils using near-infrared spectroscopy improved by wavelet transform and uninformative variable elimination,” Analytica Chimica Acta 634(2), 2009,166-171.

[15] Eom, C. D., Park, J. H., Choi, I. G., Choi, J. W., Han, Y., and Yeo, H. “Determining surface emission coefficient of wood using theoretical methods and near-infrared spectroscopy,” Wood and Fiber Science 45(1),2013, 76-83.

[16] GB/T 1929 “Method of sample logs sawing and test specimens selection for physical and mechanical tests of wood,” Standardization Administration of China, Beijing, China,2009.

[17] GB/T 1935 “Method of testing in compressive strength parallel to grain of wood,” Standardization Administration of China, Beijing, China,2009.

[18] Gietlinger, N., Schwanninger, M., Hinterstoisser, B., and Wimmer, R. “Rapid determination of heartwood extractives in Larix sp. by means of Fourier transform near infrared spectroscopy,” Journal of Near Infrared Spectroscopy 10(3),2002,203-214.

[19] LY/T 2053 “Standard method for near-infrared qualitative analysis of wood,” Standardization Administration of China, Beijing, China,2012.

[20] Schimleck, L. R., Mora, C., and Daniels, R. F. “Estimation of the physical wood properties of green, Pinus taeda, radial samples by near infrared spectroscopy,” Canadian Journal of Forest Research 33(12), 2003,2297-2305.

[21] Tong, L., and Zhang, W. “Using Fourier transform near-infrared spectroscopy to predict the mechanical properties of thermally modified southern pine wood,” Applied Spectroscopy 70(10), 2016,1676-1684.

[22] Yu, H., Liang, H., Lin, X., and Zhang, Y. “Nondestructive determination of the modulus of elasticity of *Fraxinus mandshurica* using near-infrared spectroscopy,” Optical Engineering 57(4), 046103,2018.