# Nondestructive Determination of the Compressive Strength of Wood Using Near-Infrared Spectroscopy

Hao Liang,[a] Jun Cao,[a] Wenjun Tu,[a] Xue Lin,[b] and Yizhuo Zhang [a,*]

In this study, *Xylosma racemosum* was selected as the raw material and its compressive strength was predicted through nondestructive methods. The test data consisted of 160 near-infrared (NIR) absorption spectra of the wood samples obtained using an NIR spectrometer, with the wavelength range of 900 to 1900 nm. The original absorption spectra were pre-processed with multiplicative scatter correction (MSC) and Savitzky-Golay (SG) smoothing and divided into several intervals using the backward interval partial least squares (BiPLS) method. The optimal combination of intervals with the smallest root mean square error of cross validation (RMSECV) value was selected, and a genetic algorithm (GA) was used to select featured wavelengths. Finally, a partial least squares (PLS) regression model was established with the featured wavelengths. The BiPLS-GA-PLS model outperformed the other models, resulting in a high prediction correlation coefficient of 0.927 and a root mean square error rate of 4.06. Based on the results, it is feasible to accurately measure the compressive strength of wood processed by different methods using near-infrared spectroscopy.

*Keywords: Compressive strength; Prediction; Near-infrared spectroscopy; BiPLS-GA*

*Contact information: a: College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin, 150040, China; b: Material Science and Engineering College, Northeast Forestry University, Harbin, 150040, China; *Corresponding authors: nefuzyz@163.com*

## INTRODUCTION

Wood is a very important construction material that requires a high degree of structural performance and reliability. Researchers have studied various properties of wood to determine the structural capacity of timber or lumber, *i.e.*, defects, strips, density, moisture content, *etc*.

The compressive strength of wood plays an important role in the sorting degree of timber because it is related to the mechanical and technological properties of wood. It is important to inspect the compressive strength of wood correctly, rapidly, and simply. The determination of the compressive strength is accomplished using standardized methods that are time-consuming, destructive, and costly (Herizo *et al.* 2015). Furthermore, the repeatability of testing lacks consistency. To promote efficiency and complete usage, the application of nondestructive techniques may prove effective in predicting the compressive strength of wood.

Near-infrared (NIR) spectroscopy is an effective tool in wood research. It can be calibrated to predict various physical and chemical properties of wood (Satoru and Hikaru 2015). For example, basic density (Santos *et al.* 2012), moisture content (Eom *et al.* 2013), drying stress level (Ken *et al.* 2013), wood color (Yang *et al.* 2012), surface roughness (Zhang *et al.* 2015), lignin (Rambo *et al.* 2016), cellulose (Li *et al.* 2015), hemicellulose (Li *et al.* 2015), and extractives (He *et al.* 2013) have been determined using NIR

spectroscopy. In addition, NIR spectroscopy has been used to classify wood species and their origins (Adepipe *et al.* 2008; Castillo *et al.* 2008).

To measure the compressive strength of wood, this study determined the effect of different pre-processing methods of the original NIR data on the regression model. Additionally, the results of the wavelength selection, using backward interval partial least squares and genetic algorithm, were investigated, and the predictive effects of the different regression models were compared. The main goal of this study is to develop calibrations for determining the compressive strength of wood using NIR spectroscopy and to evaluate the predictive ability of the calibrations.

## EXPERIMENTAL

### Sampling Preparation

Thirty green logs of *Xylosma racemosum* wood were collected from Chonghe Forest, Heilongjiang province, China. Twelve trees were collected, and disks measuring 5 cm in thick compared to the length of the log were cut from each tree at a height of 1.3 m. After preprocessing, the discs were divided and cut into small samples with dimensions of 20 mm x 20 mm x 20 mm. A total of 180 samples were prepared without defects or significant differences in color.

Figure 1 shows the process of sample collection. The samples were randomly divided into two groups: 120 samples for the calibration set and 40 samples for the prediction set.

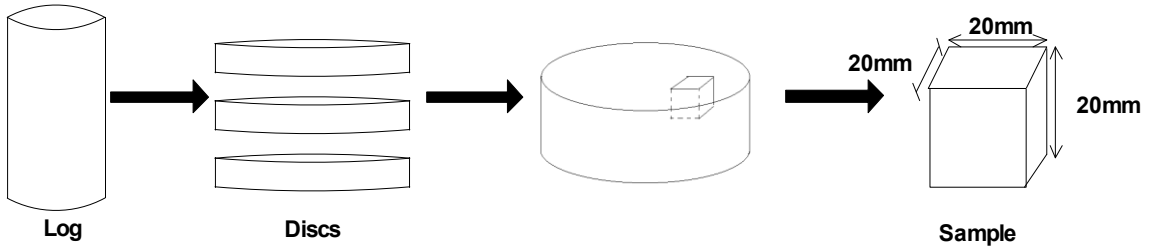


**Fig. 1.** The process of samples collection

### NIR Spectra Measurements

The NIR spectra were acquired using an absorption peak mode with an ultra-compact NIR fiber optic spectrometer (Insion Co., GmbH, Heilbronn, Germany) from 900 to 1900 nm at 9 nm intervals. The spectrometer utilized two bifurcated fiber optic probes to scan the sample surface. The NIR spectra data were obtained from SPEC view 7.1 software (Insion Co., GmbH, Heilbronn, Germany). The temperature and relative humidity were controlled at 22 °C and 50%, respectively. The whole process of NIR spectra measurements was carried out based on "Standard method for near infrared qualitative analysis of wood (LY/T 2053-2012)".

Each facet exhibited different absorption peaks because the growth characteristics of timber differ. Thus, five spectra were collected from every facet, and 30 spectra from each sample were averaged to represent a pooled mean. The wavelength from 1000 to 1600 nm contained the main information about the wood properties (Todorović *et al.* 2015).

### Determination of Compressive Strength

In accordance with national standards, "Method of testing in compressive strength parallel to grain of wood (GB/T 1935-2009)", all the samples were pressed along the grain at a uniform rate until they were destroyed (1.5 to 2 min per sample) to determine the compressive strength of wood.

### Pre-processing of NIR Spectra

After transforming the NIR spectral data into the absorption frequencies, the spectrum was pre-processed to eliminate high-frequency noise, baseline drift, light scattering, and other negative effects. Common pre-processing methods include normalization, smoothing, Fourier transform, *etc*. The multiplicative scatter correction (MSC) and the Savitzky-Golay (SG) smoothing methods were selected for pre-processing, and the results were compared. Multiplicative scatter correction reduces the influence of scattering particle size and uneven distribution from NIR spectroscopy, and SG smoothing eliminates the baseline drift and tilt noise.

### Data Analysis

*Backward interval partial least squares (BiPLS)*

A BiPLS algorithm, developed by Nørgaard *et al*. (2000), was used to divide the spectrum into *n* smaller equidistant regions (sub-intervals), and displayed partial least squares (PLS) regression models to represent each sub-interval. Then, the root mean square error of cross validation (RMSECV) was calculated for each sub-interval. Thereafter, the algorithm reduced the sub-intervals (*m*) that exhibited the largest RMSECV. The result was an optimized PLS regression model and RMSECV for the *n*-*m* sub-intervals. This procedure was continued until only a few sub-intervals remained. When the RMSECV of the sub-intervals was minimal, the optimal combination interval was achieved.

*Genetic algorithm (GA)*

A genetic algorithm simulates Darwin's genetic selection and the natural elimination process of biological evolution. The GA method is an adaptive search inspired by the mechanics of natural genetics and natural selection, and it has been successfully applied to select variables in NIR spectroscopy for building a multivariate calibration model. The fitness of the model was calculated as the inverse of the RMSECV value, and the resulting model was applied to the validation set (Zou *et al.* 2010).

### Model Evaluation Standard

The quality of the models was assessed using several common statistical measures, which included the linear coefficient of determination ($R^2$), the root mean square error of the calibration (RMSEC), and the root mean square error of the prediction (RMSEP). Selection of the final model was based on the predictability, following a procedure by Gierlinger *et al.* (2002).

## RESULTS AND DISCUSSION

### Wood Compressive Strength Determination

The compressive strength of the 160 wood samples ranged from 60.31 MPa to 84.85 MPa. The samples were randomly divided into calibration sets and prediction sets.

The statistical characteristics of the compressive strength of the two sets are summarized in Table 1.

## Near-infrared Spectroscopy of Samples and Pre-Processing of Spectra

Figure 2 shows the NIR absorption of all of the samples from 906 to 1864 nm. To verify the superior capability of the calibration models, partial least squares regression (PLSR), multiple linear regression (MLR), and principal component regression (PCR) were used to interpret the relationship between the calibration set and the prediction set. Table 2 shows the prediction results based on different models and pre-processing methods.

**Table 1.** Statistics of the Compressive Strength from the Calibration and Prediction Sets

| Samples | Maximum (MPa) | Minimum (MPa) | Mean (MPa) | Standard Deviation (MPa) |
|---|---|---|---|---|
| Calibration set (*n* = 120) | 84.47 | 60.31 | 71.68 | 7.51 |
| Prediction set (*n* = 40) | 84.85 | 61.51 | 72.30 | 7.16 |

*n*: number of wood samples

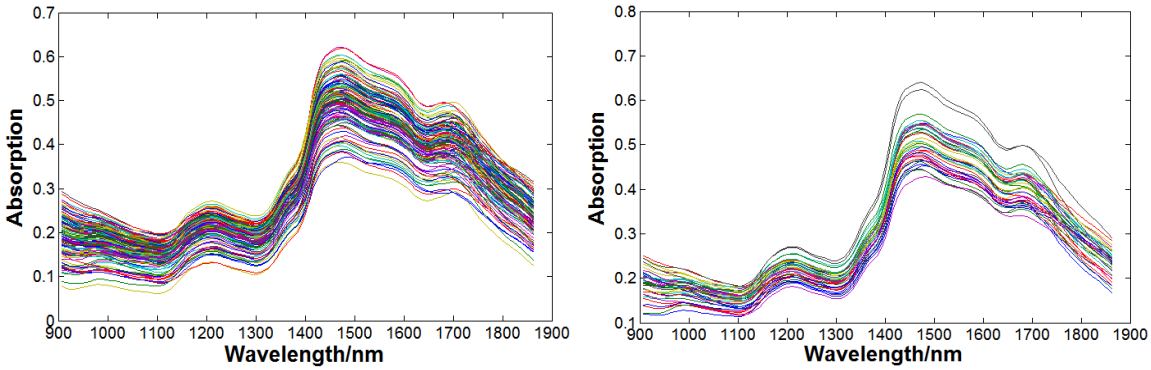


**Fig. 2.** Original spectra of *Xylosma racemosum* in the (a) calibration sets and (b) prediction sets

**Table 2.** Prediction Results of Wood Compressive Strength Based on Various Models and Pre-Processing Methods

| Pre-Processing method | | Unprocessed | MSC | SG | MSC+SG |
|---|---|---|---|---|---|
| PLSR | $R_C^2$ | 0.75 | 0.814 | 0.808 | 0.819 |
| | SEC | 4.692 | 4.352 | 4.41 | 4.227 |
| | $R_P^2$ | 0.667 | 0.805 | 0.817 | 0.824 |
| | SEP | 4.497 | 4.294 | 4.277 | 4.269 |
| MLR | $R_C^2$ | 0.709 | 0.728 | 0.847 | 0.877 |
| | SEC | 4.69 | 4.588 | 4.138 | 4.124 |
| | $R_P^2$ | 0.627 | 0.675 | 0.764 | 0.854 |
| | SEP | 4.707 | 4.508 | 4.38 | 4.275 |
| PCR | $R_C^2$ | 0.716 | 0.806 | 0.743 | 0.877 |
| | SEC | 4.991 | 4.266 | 4.279 | 4.141 |
| | $R_P^2$ | 0.739 | 0.817 | 0.803 | 0.827 |
| | SEP | 4.372 | 4.315 | 4.329 | 4.266 |

PLSR: partial least squares; MLR: multiple linear regression; PCR: principal component regression; $R_c^2$: coefficient of determination of the calibration model; $R_p^2$: coefficient of determination of the prediction model; SEC: the root mean square error of calibration; SEP: the root mean square error of prediction; MSC: multiplicative scatter correction; SG: Savitzky-Golay

For the prediction models, the spectra pre-processed with the combination of MSC and SG provided better results than those pre-processed with MSC, SG, or untreated. Figure 3 shows the wood spectra of samples pre-processed with the combination of MSC and SG. The spectra were smoother, and the noise and baseline drift were mostly eliminated.

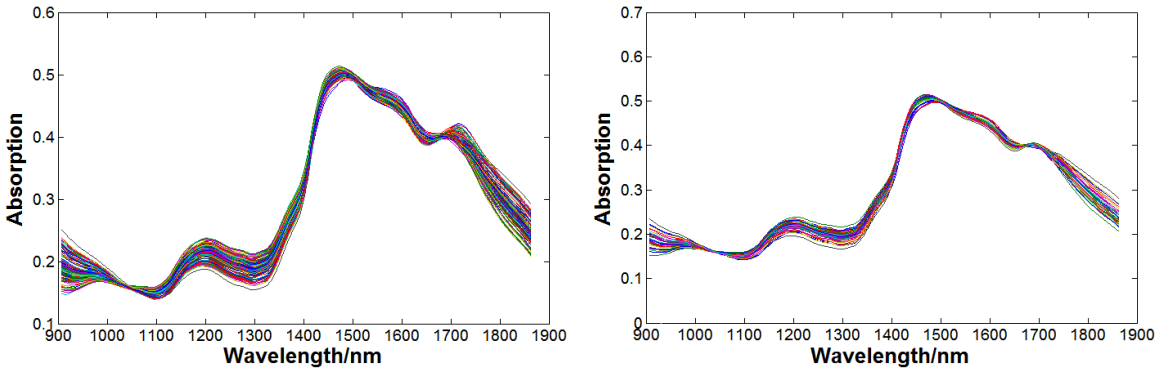


**Fig. 3.** Spectra of wood samples pre-processed with a combination of MSC and SG: (a) calibration sets; (b) prediction sets

## Wavelength Selection

*Optimal spectra intervals selected by backward interval partial least squares (BiPLS)*

After preprocessing with the combination of MSC and SG, BiPLS was applied to select optimal intervals from the spectra. One hundred and seventeen spectrum values were divided into 5 to 20 intervals. Table 3 shows the relationship between the number of wavelength variables selected by BiPLS and the number of intervals.

**Table 3.** Optimal Spectra Intervals by BiPLS

| Number of Intervals | Number of Selected Intervals | RMSECV | Variables |
|---|---|---|---|
| 5 | 3 | 4.3025 | 70 |
| 6 | 4 | 4.3568 | 76 |
| 7 | 4 | 4.1465 | 64 |
| 8 | 4 | 4.3139 | 58 |
| 9 | 4 | 4.2002 | 50 |
| 10 | 5 | 4.1576 | 58 |
| 11 | 5 | 4.1722 | 52 |
| 12 | 6 | 4.5407 | 58 |
| 13 | 8 | 4.5244 | 70 |
| 14 | 8 | 4.3315 | 65 |
| 15 | 7 | 4.4763 | 54 |
| 16 | 8 | 4.5871 | 58 |
| 17 | 9 | 4.5762 | 60 |
| 18 | 9 | 4.3438 | 58 |
| 19 | 9 | 4.3433 | 55 |
| 20 | 10 | 4.4350 | 58 |

RMSECV: the root mean square error of cross validation

The smallest RMSECV occurred when the number of intervals was 7. The optimal combination of intervals, *i.e.*, 2, 3, 4, and 6 corresponded to 1046.23 to 1177.66 nm, 1185.88 to 1317.61 nm, 1325.85 to 1457.88 nm, and 1606.75 to 1739.37 nm, respectively, in the spectral regions shown in Fig. 4 (64 wavelengths total).

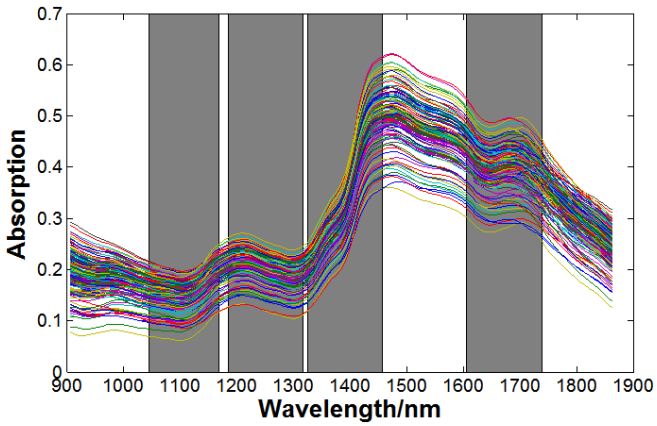


**Fig. 4.** Spectral interval selection by BiPLS

*Feature wavelengths selected by genetic algorithm (GA)*

After the optimal intervals were selected by BiPLS, GA was employed to select 8 wavelengths (1259.95, 1268.18, 1276.42, 1284.65, 1317.61, 1325.85, 1424.85, and 1433.1 nm), corresponding to the optimal points of the RMSECV curve in Fig. 5a. Figure 5b shows the selected wavelengths in the spectra by BiPLS-GA.

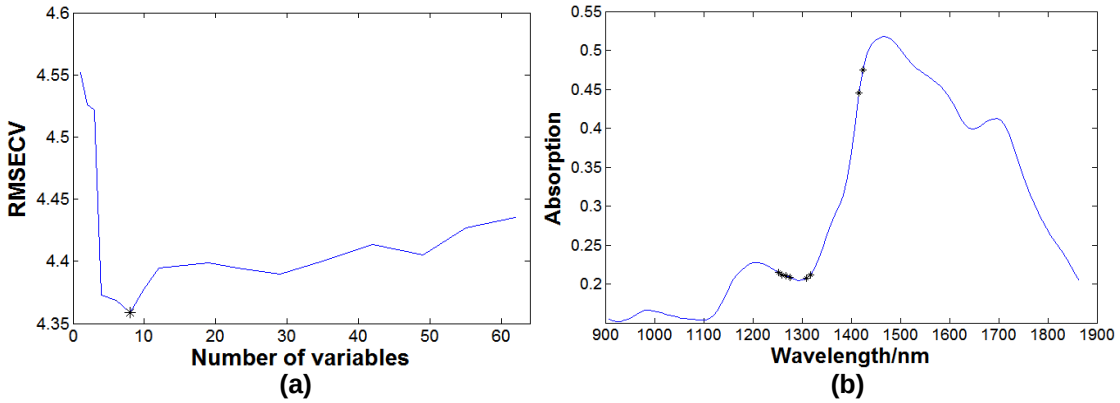


**Fig. 5.** (a) RMSECV curve based on different numbers of variables selected by GA; (b) the location of selected variables in the spectrum

**Table 4.** Comparison of Models

| Models | Selected Variables | Calibration Set | | Prediction Set | |
|---|---|---|---|---|---|
| | | $R_c$ | RMSEC | $R_p$ | RMSEP |
| PLS | 117 | 0.819 | 4.227 | 0.824 | 4.269 |
| iPLS | 22 | 0.844 | 4.201 | 0.836 | 4.252 |
| biPLS | 64 | 0.893 | 4.159 | 0.871 | 4.216 |
| GA-PLS | 16 | 0.917 | 4.077 | 0.904 | 4.198 |
| BiPLS-GA-PLS | 8 | 0.936 | 4.030 | 0.927 | 4.061 |

$R_c$: coefficient of determination of the calibration model; $R_p$: coefficient of determination of the prediction model; SEC: the root mean square error of calibration; SEP: the root mean square error of prediction

## Predictive Effects of the Model

The optimal wavelengths in the calibration sets were utilized to build the calibration models. The experiment compared the prediction effects of the five models: PLS, iPLS, BiPLS, GA-PLS, and BiPLS-GA-PLS (Table 4). The BiPLS-SPA-PLS model obtained the best prediction result. This model yielded the lowest RMSEP of 4.061 and the highest $R_p$ of 0.927. The GA-PLS model was the next-best model, while the performance of the PLS model provided the worst fit. Figure 6 shows the relationship between the measured and the predicted compressive strengths of *Xylosma racemosum* with BiPLS-GA-PLS as the prediction model.

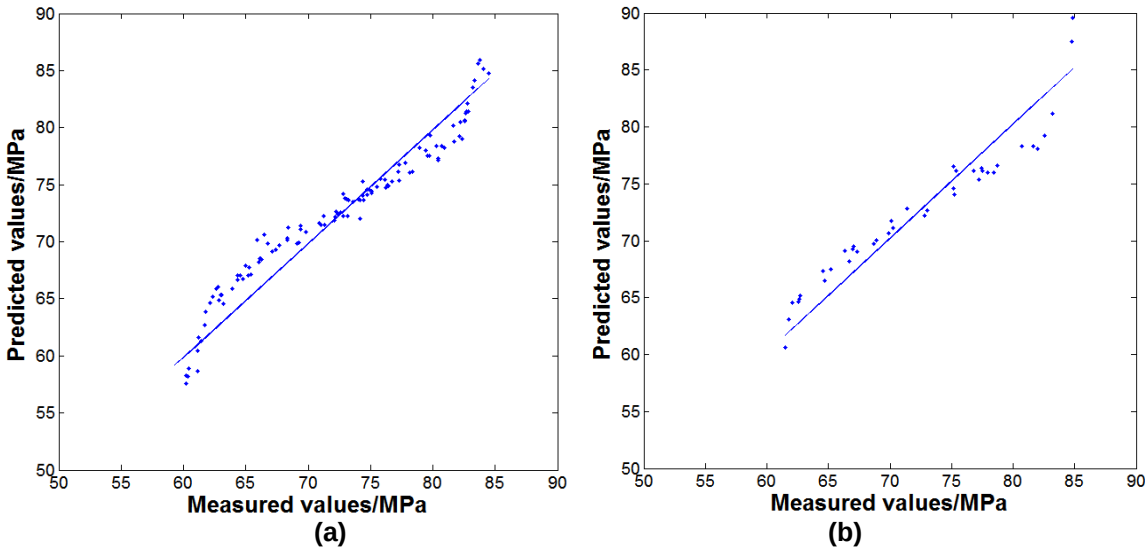


**Fig. 6.** Relationships between the measured and the predicted compressive strengths of *Xylosma racemosum* in the (a) calibration set and (b) prediction set

The BiPLS-GA-PLS model combined the advantages of BiPLS and GA. This model removed the noise and low information region, while retaining useful information and selecting the optimal intervals for calibrating the model. With high accuracy and good predictive ability, the BiPLS-GA-PLS model is robust and simple.

## CONCLUSIONS

1. This study revealed the relationship between the near-infrared (NIR) spectrum and the compressive strength of wood. The near-infrared spectrum can be used to determine the compressive strength of wood.
2. Compared with the prediction results from the other four models, BiPLS-GA-PLS was most capable of determining the compressive strength of wood. An NIR model for the compressive strength determination for multiple species of wood warrants future study.

## ACKNOWLEDGEMENTS

The authors are grateful for the support of National Forestry Bureau of the 948 Project (Grant No. 2015-4-52), and the support of the Fundamental Research Funds for the Central Universities (Grant No. 2572015AB14).

## REFERENCES CITED

Adepipe, O. E., Dawson-Andoh, B., Slahor, J., and Osborn, L. (2008). "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), 49-57. DOI: 10.1255/jnirs.760

Castillo, R., Contreras, D., Freer, J., Ruiz, J., and Valenzuela, S. (2008). "Supervised pattern recognition techniques for classification of eucalyptus species from leaves NIR spectra," *Journal of Chilean Chemical Society* 53(4), 1709-1713. DOI: 10.4067/S0717-97072008000400016

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

Gierlinger, N., Schwanninger, M., Hinterstoisser, B., and Wimmer, R. (2002). "Rapid determination of heartwood extractives in *Larix* sp by means of Fourier transform near infrared spectroscopy," *Journal of Near Infrared Spectroscopy* 10(3), 203-214. DOI: 10.1255/jnirs.336

He, W., and Hu, H. (2013). "Rapid prediction of different wood species extractives and lignin content using near infrared spectroscopy," *Journal of Wood Chemistry Technology* 33(1), 52-64. DOI:10.1080/02773813.2012.731463

Herizo, R., Gilles, C., Loïc, B., Tahiana, R., and Marie, F. T. (2015). "A novel method to correct for wood MOE ultrasonics and NIRS measurements on increment cores in *Liquidambar styraciflua* L," *Annals of Forest Science* 72(6), 753-761. DOI 10.1007/s13595-015-0469-6

Ken, W., Isao, K., Shuetsu, S., Naohiro, K., and Shuichi, N. (2013). "Nondestructive evaluation of drying stress level on wood surface using near-infrared spectroscopy," *Journal of Wood Science and Technology* 47(2), 299-315. DOI: 10.1007/s00226-012-0492-9

Li, X., Sun, C., Zhou, B., and He, Y. (2015). "Determination of hemicellulose, cellulose and lignin in Moso bamboo by near infrared spectroscopy," *Scientific Reports* 5, 1-11. DOI: 10.1038/srep17210

Norgaard, L., Saudland, A., Wagner, J., Nielsen, J. P., Munck, L., and Engelsen, S. B. (2000). "Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy," *Applied Spectroscopy* 54(3), 413-419. DOI: 10.1366/0003702001949500

Rambo, M. K. D., Ferreira, M. M. C., and Amorim, E. P. (2013). "Multi-product calibration models using NIR spectroscopy," *Chemometrics & Intelligent Laboratory Systems* 151, 108-114. DOI: 10.1016/j.chemolab.2015.12.013

Santos, A., Alves, A., Simoes, R., Pereira, H., Rodrigues, J., and Schwanninger, M. (2012). "Estimation of wood basic density of *Acacia melanoxylon* (R. Br.) by near

infrared spectroscopy," *Journal of Near Infrared Spectroscopy* 20(2), 267-274. DOI: 10.1255/jnirs.386

Satoru, T., and Hikaru, K. (2015). "A review of recent application of near infrared spectroscopy to wood science and technology," *Journal of Wood Science* 61(3), 213-220. DOI 10.1007/s10086-015-1467-x

Todorović, N., Popović, Z., and Milić, G. (2015). "Estimation of quality of thermally modified beech wood with red heartwood by FT-NIR spectroscopy," *Wood Science & Technology* 49(3), 527-549. DOI: 10.1007/s00226-015-0710-3

Yang, Z., Lv, B., and Fu, Y. J. (2012). "The relationship between near infrared spectroscopy and surface color of eight rosewoods," *Advanced Materials Research* 479-481, 1772-1776. DOI: 10.4028/www.scientific.net/AMR.479-481.1772

Zhang, M. M., Liu, Y. N., and Yang, Z. (2015). "Correlation of near infrared spectroscopy measurements with the surface roughness of wood," *BioResources* 10(4), 8505-8517. DOI: 10.15376/biores.10.4.8505-8517

Zou, X., Zhao, J., Povey, M. J. W., Holmes, M., and Mao, H. (2010). "Variables selection methods in near-infrared spectroscopy," *Analytica Chimica Acta* 667 (1-2), 14-32. DOI:10.1016/j.aca.2010.03.048

Article submitted: March 18, 2016; Peer review completed: May 2, 2016; Revised version received and accepts: May 4, 2016; Published: July 12, 2016.
DOI: 10.15376/biores.11.3.7205-7213