

# COMPARISON OF GENETIC ALGORITHM OPTIMIZED AND PARTIAL LEAST SQUARES REGRESSION DENSITY MODELS FOR *ACACIA AURICULIFORMIS*

*Laurence R. Schimleck*\*†

Professor  
Department of Wood Science and Engineering  
119 Richardson Hall  
Oregon State University  
Corvallis, OR 97331  
E-mail: laurence.schimleck@oregonstate.edu

*Tu X. Ho*

Associate R&D Engineer  
Simpson Strong-Tie  
5956 W. Las Positas Blvd  
Pleasanton, CA 94588  
E-mail: tuho@strongtie.com

*Doan Van Duong*

Associate Professor  
Thai Nguyen University of Agriculture and Forestry  
Quiet Thang ward, Thai Nguyen city, Vietnam  
E-mail: duongvandoan@tuaf.edu.vn

*Arijit Sinha*†

Professor  
Department of Wood Science and Engineering  
119 Richardson Hall  
Oregon State University  
Corvallis, OR 97331  
E-mail: arijit.sinha@oregonstate.edu

*Ighoyivwi Onakpoma*†

PhD Candidate  
Department of Wood Science and Engineering  
119 Richardson Hall  
Oregon State University  
Corvallis, OR 97331  
E-mail: ighoyivwi.onakpoma@oregonstate.edu

*Galen Fox*

Student  
Department of Wood Science and Engineering  
119 Richardson Hall  
Oregon State University  
Corvallis, OR 97331  
E-mail: foxga@oregonstate.edu

(Received November 2023)

---

\* Corresponding author

† SWST member

**Abstract.** Partial least squares (PLS) regression models based on genetic algorithm (GA) representative near IR (NIR) wavelengths for estimating wood properties provide improved calibration and prediction statistics compared with PLS models based on all available NIR wavelengths. However, the utilization of predicted data, obtained from full NIR wavelength and GA-selected NIR wavelength models, in a practical application, has not been explored. Our application was to examine radial density variation in *Acacia auriculiformis* Cunn. Ex Benth. clones at a resolution of 10 mm. One hundred and forty *A. auriculiformis* samples representing seven clones and two radial positions (adjacent to pith and bark, respectively) had NIR hyperspectral images (wavelength range 931-1718 nm) collected from their transverse surface. Two PLS density models (all NIR wavelengths and GA representative NIR wavelengths) were developed using 134 NIR spectra extracted from the images. The models were then used to predict density in 10 mm increments of 144 radial samples from the same clones. A PLS density model using only 15 representative NIR wavelengths provided a mean ( $0.506 \text{ g/cm}^3$ ) that was not statistically significantly different from the measured density ( $0.503 \text{ g/cm}^3$ ), whereas the mean for the PLS model using all wavelengths was  $0.522 \text{ g/cm}^3$ . However, the PLS model with 15 representative NIR wavelengths had greater variation (standard deviation of 0.060) compared with measured data (0.052) and the full NIR wavelengths PLS model (0.047). Radial density variation was less than  $0.09 \text{ g/cm}^3$  for six of the seven clones examined.

**Keywords:** *Acacia auriculiformis*, density, genetic algorithm, hyperspectral imaging, near infrared spectroscopy, partial least squares regression.

## INTRODUCTION

Near IR (NIR) spectroscopy provides a rapid, nondestructive approach for the assessment of a variety of wood properties (Tsuchikawa and Kobori 2015; Schimleck and Tsuchikawa 2021). Several recent studies have reported the utilization of advanced modeling approaches (eg machine learning and genetic algorithm [GA]) for calibration development (Cogdill et al 2004; Mora and Schimleck 2010; Fernandes et al 2013; Li et al 2019; Nasir et al 2019, 2023; Ayanleye et al 2021; Ho et al 2021, 2022), however, the predictive performance of these models, and their practical significance in terms of an application, have not been assessed and compared with results obtained using predictive models obtained using partial least squares (PLS) regression, a modeling approach used for the large majority of NIR-wood related studies.

One advanced approach involves the utilization of GAs (Bangalore et al 1996; Villar et al 2014; De et al 2017) for model development. Recent papers have investigated the potential of PLS models based on GA-selected NIR wavelengths to estimate pulp yield of Tasmanian blue gum (*Eucalyptus globulus* Labill.) samples (Ho et al 2021) and the estimation of a variety of SilviScan (Evans 1994, 1999, 2006) measured by wood properties (density, microfibril angle, modulus of elasticity and tracheid coarseness, radial diameter,

tangential diameter, and wall thickness) for loblolly pine (*Pinus taeda* L.) (Ho et al 2022). These studies demonstrate that PLS models based on GA-identified NIR wavelengths have improved calibration and prediction statistics. Moreover, band assignments for key wavelengths consistently arise from bond vibrations observed in lignin, cellulose, and hemicellulose, providing a strong fundamental basis for interpretation and explaining improved model performance. As noted, the utilization of predicted data in a practical application has not been explored, or compared with, the performance of predictions obtained using PLS regression models. Hence the aim of this study is to compare the performance of both model types in a practical application.

The application identified is the prediction of radial density variation in 10 mm increments in seven *Acacia auriculiformis* Cunn. Ex Benth. clones from a trial established in north central Vietnam. *A. auriculiformis* was introduced into Vietnam in the 1960s utilizing Australian, Papua New Guinean, and Indonesian provenances (Pinyopusarek et al 1991; Hai 2009) and owing to its fast growth and adaptability has become an important plantation species. It is a source of pulp wood (Jahan et al 2008), sawlogs (Hai et al 2008), and is also used for furniture (Hai 2009). *A. auriculiformis* tree improvement programs have been established with an emphasis on

growth, stem form, and disease resistance but little information is available regarding wood properties. For lumber production, in particular, radial variation in wood properties is important and research (Van Duong et al 2022) has aimed to investigate radial variation amongst clones, however only samples adjacent to the pith and bark were examined. To better understand radial patterns greater resolution is required and NIR-predicted density, obtained from PLS regression models based on all available NIR wavelengths and only those identified by GA, were utilized for this purpose.

## MATERIALS AND METHODS

### A. *auriculiformis* Wood Samples

Sample trees were harvested from an *A. auriculiformis* clonal trial established by the Vietnamese Academy of Forest Sciences to assess the growth rate and stem quality of different clones. The site is located in Cam Hieu commune, Cam Lo district, Quang Tri province, north central Vietnam (16°45'60" N and 107°01'12" E). The plantlets were propagated by tissue culture and planted at the site in December 2015 using a randomized complete block design with four replicates. Each plot comprised 36 ramets from a clone (6 lines × 6 ramets/line). The initial spacing between ramet was 3.0 × 3.0 m<sup>2</sup> (1100 trees/ha). Fertilizer application at planting was 100 g nitrogen (N), phosphate (P<sub>2</sub>O<sub>5</sub>), and potassium oxide (K<sub>2</sub>O) (VADFCO, Hanoi, Vietnam) (elemental ratio 16:16:8) per ramet and 100 g NPK 1 yr later. There are seven clones considered in this study: Clt7, Clt18, Clt19, Clt25, Clt26, Clt57, and Clt43.

### Sampling and Wood Property Assessment

A total of 35 ramets (5 per clone) were chosen based on straightness, branching, and absence of disease or pest symptoms in December 2020. Stress-wave velocity of standing tree was measured using a Fakopp Microsecond Timer for each tree (Serial No.: FN-12/2020; Fakopp Enterprise Bt., Fenyó u.26, Hungary). Prior to destructive sampling, stem diameter at a height of 1.3 m was measured, and the north and south sides were

marked for selected ramets, and once felled, total height was measured.

A 1.0-m log was collected between 0.5 and 1.5 m from each sampled stem. Logs were dried in a room at ambient conditions for approximately 2 mo and after drying, eight 20 (radial) × 20 (tangential) × 300 (longitudinal) mm<sup>3</sup> small wood specimens adjacent to the pith and bark were cut from each log for additional stress wave measurements, gravimetric density determination and destructive evaluation of wood strength and stiffness (note – only the density data are used in this study). Four samples (2 adjacent to the pith and 2 adjacent to the bark) were available for each ramet, giving 140 samples in total for density PLS model development. More detail regarding these samples is reported in Van Duong et al (2022).

### Samples to Test PLS and GA Density Models

From the sampled *A. auriculiformis* trees, a radial strip was also obtained from the 1.0-m log used to analyze density variation at a higher resolution and test the PLS models using all NIR wavelengths and NIR wavelengths identified by GA optimization. The radial strips were split into 10 mm increments to give 4-5 samples per radial strip (total 144 samples) and the density of these samples was determined gravimetrically. As the samples represented the complete radius of the sampled ramets they were distinct, but from the same population, as samples used for calibration that measured 20 mm in the radial direction. Hyperspectral images of examples of samples that formed the calibration and prediction sets are shown in Fig 1.

### NIR-hyperspectral Imaging

NIR-hyperspectral imaging of all samples was done using a Specim FX17 camera (Specim; Spectral Imaging Ltd., Oulu, Finland) fitted with a Specim (OLET 17.5 F/2.1) focusing lens (wavelength range 931-1718 nm, 224 wavelengths). Six tungsten halogen lamps illuminated samples while sample motion and image acquisition (which included a dark current and white reference prior

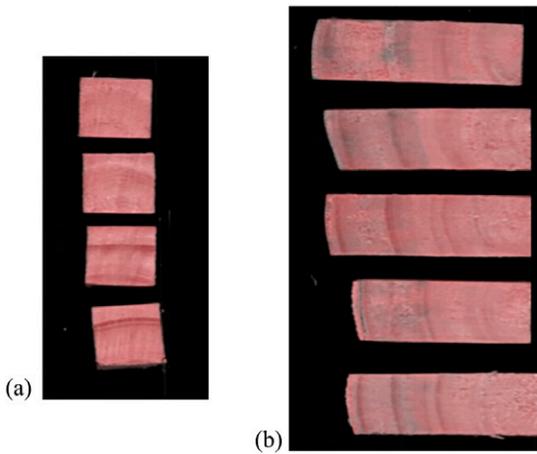


Figure 1. Hyperspectral images of a subset of the *Acacia auriculiformis* samples that formed the calibration and prediction sets. (a) Samples (20 mm radially  $\times$  20 mm tangentially) used for calibration and obtained from near the pith and from near the bark used for calibration. (b) Full radius samples (20 mm tangentially). After hyperspectral images were collected from the full radius samples they were split into sections (10 mm radially  $\times$  20 mm tangentially) and the 10 mm sections were used for density determination and formed the prediction set.

to collecting an image of each set of trays) were controlled using Lumo software supplied with the system. Samples were analyzed on a black background and care was taken to minimize exposure of the samples to the intense light of the lamps.

Relative reflectance values for each image were calibrated with the corresponding dark current and white reference data and then transformed to absorbance ( $A$ ),  $A = \log_{10} 1/R$ . Regions of interest (ROI), corresponding to each 10 mm section for each radial strip, were identified in images of the samples and cropped to provide individual NIR spectra per section. ROI identification was completed using Matlab software (MathWorks, Natick, MA).

Six of the calibration samples had NIR spectra that were quite different from the others having high absorbance values in the region 931–1100 nm. Examination of a leverage plot (Fig 2) obtained from principal components analysis using Unscrambler (version 9.2) software (Camo, Oslo, Norway) indicated that these samples had

an excessive influence on the data set and were excluded from further analysis. As a result, 134 spectra were used to develop PLS models for density using all available NIR wavelengths and only those identified by GA.

### GA Model Optimization

The *A. auriculiformis* data set of 134 samples was separated into two subsets (108 samples in calibration set and 26 samples in prediction set) based on the DUPLEX selection method (Snee 1977). For each sample, the data consists of the NIR reflectance measurement of 224 wavelengths and density. PLS regression was adopted to develop predictive models for density using four cross-validation segments.

The GA was applied to determine the optimal wavelength set, number of wavelengths,  $N_{\text{WVL}}$ , and number of latent variables,  $N_{\text{comp}}$ , through optimizing the predictive performance of PLS models for the *A. auriculiformis* data set's mechanical properties. The optimal wavelengths were determined and varied from 10 to 50. The maximum number of latent variables was selected to be 40. To evaluate the predictive performance of PLS models, the  $R$ -squared of calibration and prediction sets were utilized. The objective function was defined for the optimization process as:  $f_{\text{obj}} = \alpha \times R_c^2 + \beta \times R_p^2$ , in which  $\alpha$  and  $\beta$  are weighted factors for  $R_c^2$  and  $R_p^2$ , respectively. The summation of  $\alpha$  and  $\beta$  equals one. In this study,  $\alpha$  and  $\beta$  were selected to be 0.5.

The constraint conditions for the optimization problem are  $R_c^2$  and  $R_p^2$  values derived from the PLS models using all 224 wavelengths with  $N_{\text{comp}}$  up to 40 to obtain the best objective value. These conditions with their corresponding  $N_{\text{comp}}$  values are presented in Table 1. It is worth noting that PLS models using all wavelengths provide the best prediction performance with  $N_{\text{comp}} = 24$  for density. Higher values of  $N_{\text{comp}}$  caused overfitting on the calibration set which resulted in worse performance on the prediction set. More details on the establishment and optimization process using GAs can be found in Ho et al (2021, 2022).

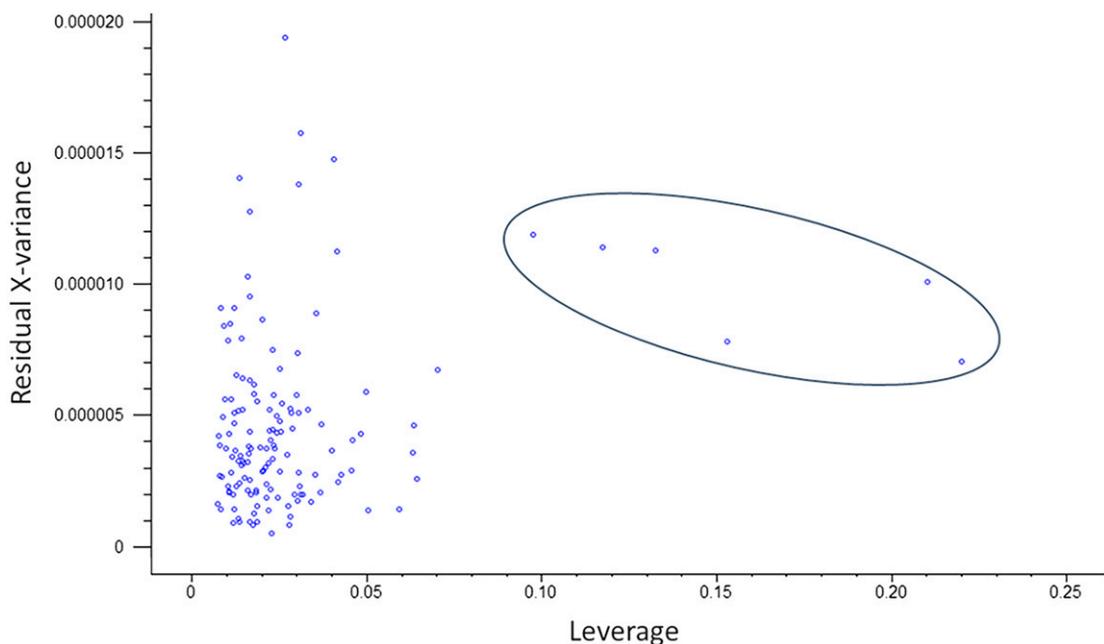


Figure 2. Plot of residual X-variance vs leverage for the 140 *Acacia auriculiformis* calibration samples. The six circled samples had excessive influence relative to the remaining samples and were excluded from genetic algorithm model development.

## RESULTS AND DISCUSSION

### GA Optimization Results for Density

The optimization was run three times which is a combination of density and a specific number of selected NIR wavelengths. The final result was determined as an average of the three optimizations.

Figure 3 shows predictive performance vs the number of wavelengths. The overall trend is that the objective value  $f_{obj}$  increases with  $N_{WvL}$ . When  $N_{WvL}$  is less than 20 for the density model, the predictive performance is weaker than the PLS density model using the full NIR wavelength range. When  $N_{WvL}$  approaches 50,  $f_{obj}$  is improved by about 8% for density, compared with that of the full NIR wavelength PLS model. However,  $R$ -squared of the calibration set,  $R_c^2$  is not improved for the density model with all

Table 1. Constraint conditions for the optimization problem.

Property	$R_c^2$	$R_p^2$	$f_{obj}$	$N_{comp}$
Density	0.9378	0.7958	0.8668	24

investigated  $N_{WvL}$ .  $R_c^2$  of the density model drops to less than 0.8 when  $N_{WvL}$  is smaller than 20. In contrast,  $R$ -squared for the prediction set,  $R_p^2$  improves more than 12% for the density model, even at a low  $N_{WvL}$  value of 10. As  $N_{WvL}$  increases,  $R_p^2$  increases.

The optimization process also determines the optimal  $N_{comp}$  value to provide the best predictive performance corresponding to a specific  $N_{WvL}$ . The results of optimal  $N_{comp}$  are shown in Fig 4 and imply a steady increase in  $N_{comp}$  along with an increase of  $N_{WvL}$  to 25. Afterward, the optimal  $N_{comp}$  fluctuates considerably in the range of 21-38 for the density model. The optimal  $N_{comp}$  value of 40 appears only once and indicates that an upper limit of  $N_{comp}$  of 40 for the optimization process is reasonable.

In each combination of density and  $N_{WvL}$ , the optimization process results in a set of wavelengths that provide the best predictive performance. However, Ho et al (2021) observed that there were few common wavelengths among optimized sets, indicating that not all optimized

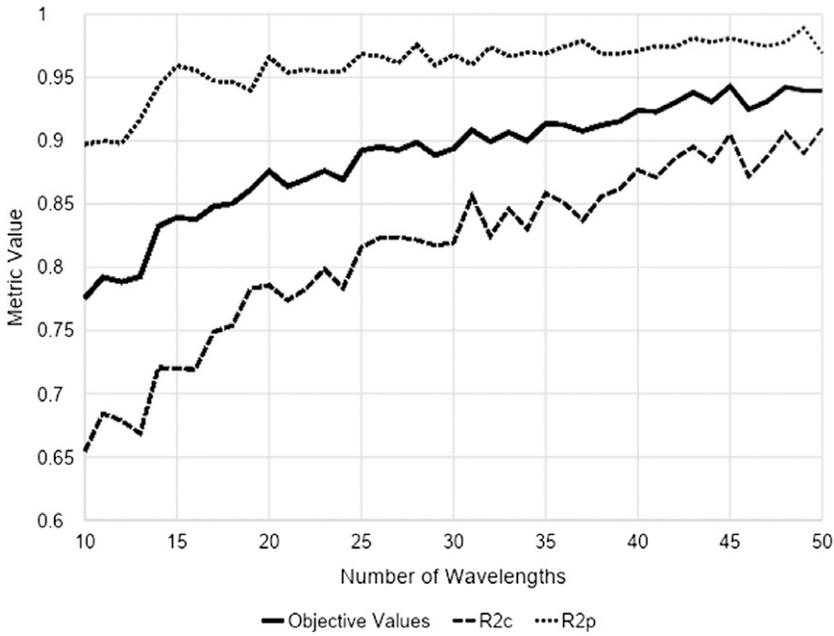


Figure 3. Predictive performance of partial least squares models for density based on an increasing number of genetic algorithm selected near IR wavelengths.

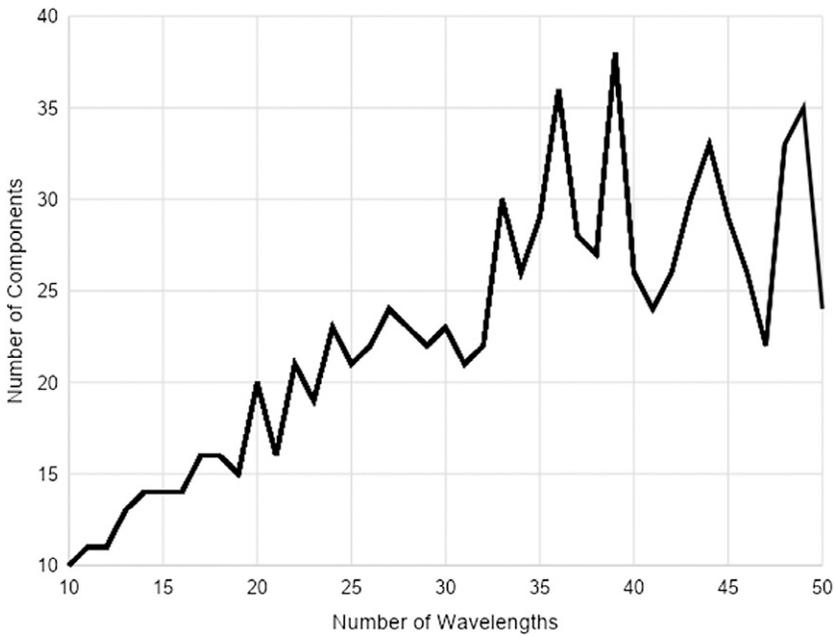


Figure 4. Optimal  $N_{comp}$  values for partial least squares density models based on an increasing number of genetic algorithm selected near IR wavelengths.

wavelengths contribute to global optimization and help to explain the relationship between wavelengths and wood components/properties. Ho et al (2021, 2022) utilized a statistical approach in which the most frequent wavelengths appearing in all optimized sets were considered the representative wavelengths. Band assignment was conducted, and strong agreement between representative wavelengths and specific wood components was observed. The same approach was applied in this study. The utilization of 224 wavelengths (average of three runs) from 931 to 1718 nm for density is shown in Fig 5. To simplify interpretation, the wavelengths are shown in intervals of 10 nm.

### Band Assignments

Wavelengths in increments of 10 nm with an average frequency (determined from the three runs) greater than 20 are summarized for density (Table 2). Corresponding band locations, their bond vibrations, and specific wood component (if available) based on Schwanninger et al (2011) are also included. For nearly all wavelengths listed in Table 2, bond vibrations related to specific wood components. First overtone OH and first overtone CH stretch vibrations were consistently represented.

### Performance of Density PLS Model Using Representative Wavelengths

To further investigate the application of representative NIR wavelengths identified in the GA

optimization process, a PLS model for density was developed to predict the density of the samples cut from the *A. auriculiformis* radial strips. Only the most frequent wavelengths were used, and they varied from 10 to 77.

Figure 6 shows the mean and standard deviation of the measured density of 144 samples in comparison with those from PLS regression models using all wavelengths and GA representative wavelengths with the total number of 10-77 wavelengths. A t-test was implemented to compare the means between prediction models and measured data. There was no statistically significant difference in means between measured density and density predicted using models based on 61, 16, and 15 representative wavelengths ( $p$ -values of 0.135, 0.894, and 0.609, respectively). However, with 15 representative wavelengths, the PLS model provides the closest mean ( $0.506 \text{ g/cm}^3$ ) to that of measured data ( $0.503 \text{ g/cm}^3$ ), which is better than the PLS model using all wavelengths ( $0.522 \text{ g/cm}^3$ ). In addition, the PLS model with 15 representative wavelengths had the lowest variation (standard deviation = 0.060) compared with other prediction models using 61 and 16 representative wavelengths, though it was higher than that of measured density (0.052) and the PLS model based on all wavelengths (0.047).

For each clone, the average density of samples taken from different ramets, but the same radial position, were determined and plotted for six

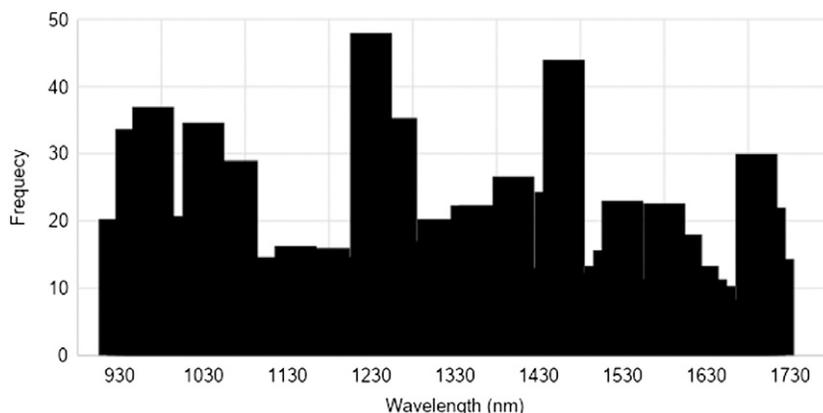


Figure 5. Frequency of optimized wavelengths (average of three runs) for density in increments of 10 nm.

Table 2. Band assignments for density optimization for NIR wavelengths with frequency greater than 20 (increments of 10 nm).

Identified wavelengths (nm)	Band location (nm)	Bond vibration	Wood component
930-950	938	3rd OT C-H str.	Not assigned
980, 1000	970-1000	2nd OT O-H str.	Starch (990 nm)
1030-1070	1037, 1053	2xC-H str. + 2xC-H def. + (CH <sub>2</sub> ) <sub>n</sub>	Not assigned
	1030, 1060	2nd OT N-H str.	Not assigned
1230-1240	1212-1225	2nd OT C-H str.	Cellulose
1260		Not assigned	
1310		Not assigned	
1350-1360	1350, 1360	1st OT C-H str. + C-H def.	Hemicellulose/all
1400	1410	1st OT O-H str.	Lignin/extractives
1450-1460	1447	1st OT O-H str.	Lignin/extractives
	1448	1st OT O-H str.	Lignin
1460	1471	1st OT O-H str.	Hemicellulose
	1473	1st OT O-H str.	Cellulose
1530	1534	1st OT O-H str.	Cellulose
1580	1579, 1580, 1586-1596	1st OT O-H str.	Cellulose
1690	1685	1st OT C <sub>ar</sub> -H str.	Lignin
	1695	1st OT C-H str. CH <sub>3</sub> groups	Not assigned
	1698	1st OT C-H str.	Lignin
1700	1703	1st OT C-H str.	Cellulose
	1705	1st OT C-H str.	Hemicellulose

NIR, near IR; OT, overtone; str., stretching vibration; def., deformation vibration; +, combination band; ar, aromatic.

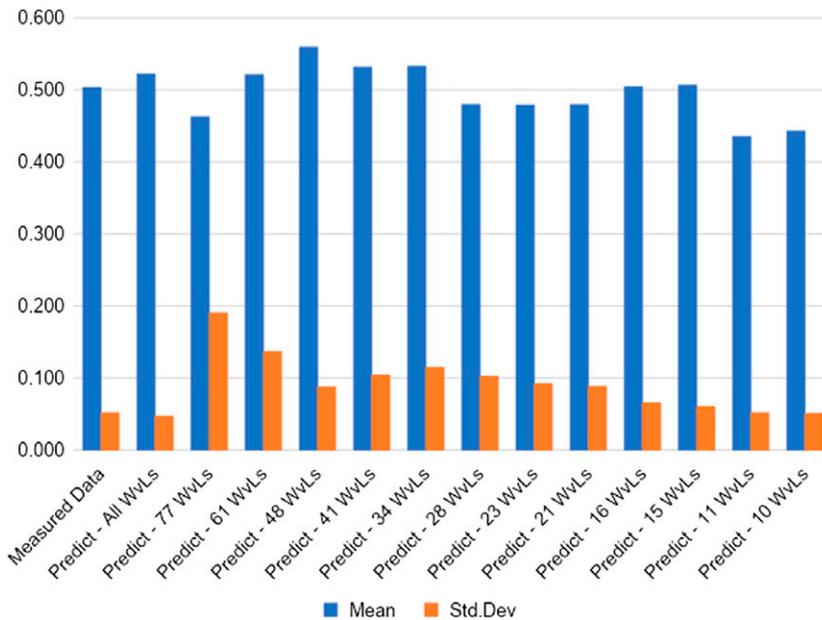


Figure 6. Comparison of mean and standard deviation from measured data, partial least squares prediction using all wavelengths and representative near IR wavelengths.

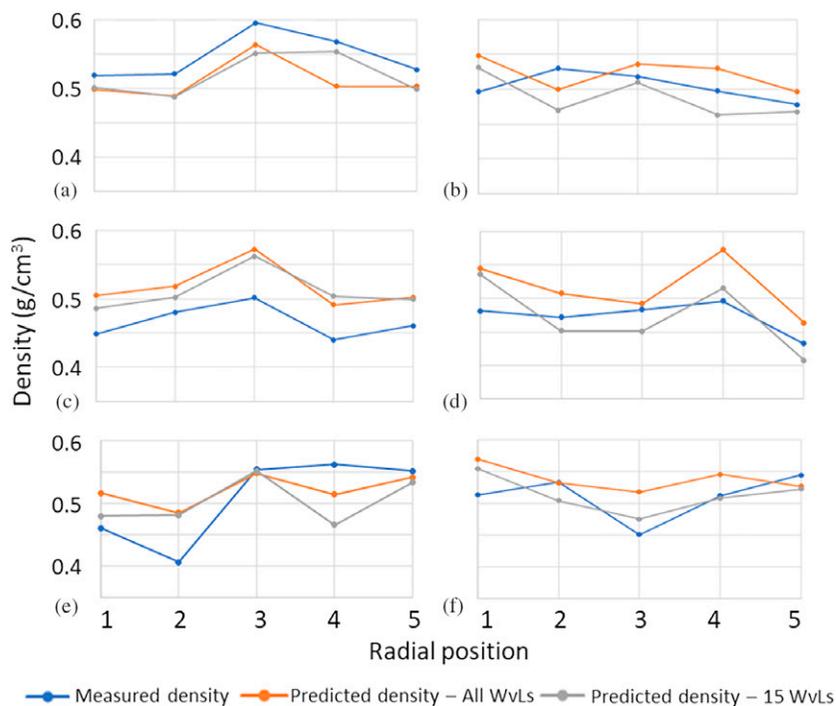


Figure 7. Density variation from pith to bark for six of the seven clones: (a) Clt18, (b) Clt19, (c) Clt25, (d) Clt26, (e) Clt43, and (f) Clt57. Radial density variation for Clt 7 was not plotted as it was missing one position.

clones in Fig 7. There are three density data series, including measured data, prediction data from the PLS model using all NIR wavelengths, and the PLS model using 15 representative NIR wavelengths. The radial location changes from pith (#1) to bark (#5).

Overall, the clones showed low radial density variation with six of the seven having a difference between maximum and minimum gravimetric densities of less than  $0.09 \text{ g/cm}^3$ . The exception was Clt43 with a range of  $0.14 \text{ g/cm}^3$ . Duc Viet et al (2020) showed a similar level of variation with density increasing from  $0.47$  to  $0.54 \text{ g/cm}^3$  as did Hai et al (2010) with heartwood densities of  $0.52 \text{ g/cm}^3$  and sapwood densities of  $0.56 \text{ g/cm}^3$ .

The narrow radial density variation made it difficult to detect a consistent trend in both measured and predicted values amongst *A. auriculiformis* clones. For some clones (Clt18, 19, 25, and 57), predicted densities from both PLS models

displayed similar radial trends to measured values but for others, most notably Clt43, the models were not as successful.

## CONCLUSIONS

To compare the performance of PLS regression models based on GA representative NIR wavelengths and full wavelength PLS models, radial density variation in seven *Acacia auriculiformis* Cunn. Ex Benth. clones at a resolution of 10 mm were examined. Overall, the clones showed low radial density variation with six of the seven having a difference between maximum and minimum measured densities of less than  $0.09 \text{ g/cm}^3$ . A PLS density model using only 15 representative NIR wavelengths provided a mean predicted density ( $0.506 \text{ g/cm}^3$ ) that was closest to the average measured density ( $0.503 \text{ g/cm}^3$ ) and there was no statistically significant difference in mean values between them based on t-tests. In comparison, the

mean density for the full NIR wavelength PLS model was  $0.522 \text{ g/cm}^3$ . The PLS model with 15 representative wavelengths also demonstrated greater variation (standard deviation of 0.060) compared with measured data (0.052) and the full wavelength PLS model (0.047).

Although the GA approach helped to identify representative NIR wavelengths for a PLS model for the density of *A. auriculiformis* clones, further studies should be conducted to generalize the findings in this study. This work should include comparing and validating the results from the GA approach with results from other optimization methods and determining the optimal number of representative NIR wavelengths to be used for the density prediction model regardless of clone.

#### ACKNOWLEDGMENT

This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED), Grant No. 106.06-2019.319.

#### REFERENCES

- Ayanleye S, Nasir V, Avramidis S, Cool J (2021) Effect of wood surface roughness on prediction of structural timber properties by infrared spectroscopy using ANFIS, ANN and PLS regression. *Eur J Wood Wood Prod* 79(1):101-115.
- Bangalore AS, Shaffer RE, Small GW, Arnold MA (1996) Genetic algorithm-based method for selecting wavelengths and model size for use with partial least-squares regression: Application to near-infrared spectroscopy. *Anal Chem* 68(23):4200-4212.
- Cogdill RP, Schimleck LR, Jones PD, Peter GF, Daniels RF, Clark A (2004) Estimation of the physical wood properties of *Pinus taeda* L. radial strips using least squares support vector machines. *J Near Infrared Spectrosc* 12(4):263-270.
- De A, Chanda S, Tudu B, Bandyopadhyay RB, Hazarika AK, Sabhapondit S, Baruah BD, Tamuly P, Bhattacharyya N (2017) Wavelength selection for prediction of polyphenol content in inward tea leaves using NIR. Pages 184-187 in *IEEE 7th International Advance Computing Conference (IACC)*, January 5-7, 2017, Hyderabad, India.
- Duc Viet D, Ma T, Inagaki T, Tu Kim N, Quynh Chi N, Tsuchikawa S (2020) Physical and mechanical properties of fast growing polyploid acacia hybrids (*A. auriculiformis* × *A. mangium*) from Vietnam. *Forests* 11:15.
- Evans R (1994) Rapid measurement of the transverse dimensions of tracheids in radial wood sections from *Pinus radiata*. *Holzforschung* 48:168-172.
- Evans R (1999) A variance approach to the X-ray diffractometric estimation of microfibril angle in wood. *Appita J* 52:283-289, 294.
- Evans R (2006) Characterization of the cellulosic cell wall. Pages 138-146 DG Stokke and L Groom, eds. Blackwell Publishing, Ames, IA.
- Fernandes A, Lousada J, Morais J, Xavier J, Pereira J, Melo-Pinto P (2013) Measurement of intra-ring wood density by means of imaging VIS/NIR spectroscopy (hyperspectral imaging). *Holzforschung* 67(1):59-65.
- Hai PH (2009) Genetic improvement of plantation-grown *Acacia auriculiformis* for sawn timber production. PhD dissertation, Swedish University of Agricultural Sciences, Uppsala, Sweden.
- Hai PH, Hannrup B, Harwood C, Jansson G, Ban DV (2010) Wood stiffness and strength as selection traits for sawn timber in *Acacia auriculiformis* A. Cunn. ex Benth. *Can J Res* 40(2):322-329.
- Hai PH, Jansson G, Harwood C, Hannrup B, Ha TH (2008) Genetic variation in growth, stem straightness and branch thickness in clonal trials of *Acacia auriculiformis* at three contrasting sites in Vietnam. *For Ecol Mgmt* 255(1):156-167.
- Ho TX, Schimleck LR, Sinha A (2021) Utilization of genetic algorithms to optimize *Eucalyptus globulus* pulp yield models based on NIR spectra. *Wood Sci Technol* 55(3):757-776.
- Ho TX, Schimleck LR, Sinha A, Dahlen J (2022) Utilization of genetic algorithms to optimize loblolly pine wood property models based on NIR spectra and SilvScan data. *Wood Sci Technol* 56:1419-1437.
- Jahan MS, Sabina R, Rubaiyat A (2008) Alkaline pulping and bleaching of *Acacia auriculiformis* grown in Bangladesh. *Turk J Agric For* 32(4):339-347.
- Li Y, Via BK, Cheng Q, Zhao J, Li Y (2019) New pretreatment methods for visible-near-infrared calibration modeling of air-dry density of *Ulmus pumila* wood. *Forest Prod J* 69(3):188-194.
- Mora C, Schimleck LR (2010) Kernel regression methods for the prediction of wood properties of *Pinus taeda* using near infrared (NIR) spectroscopy. *Wood Sci Technol* 44(4):561-578.
- Nasir V, Danish Ali S, Mohammadpanah A, Raut S, Nabavi M, Dahlen J, Schimleck L (2023). Fiber quality prediction using NIR spectral data: tree-based ensemble learning vs. deep neural networks. *Wood Fib Sci* 55(1):100-115.
- Nasir V, Nourian S, Zhou Z, Rahimi S, Avramidis S, Cool J (2019) Classification and characterization of thermally modified timber using visible and near-infrared spectroscopy and artificial neural networks: A comparative study on the performance of different NDE methods and ANNs. *Wood Sci Technol* 53(5): 1093-1109.

- Pinyopusarek K, Williams E, Boland D (1991) Geographic-variation in seedling morphology of *Acacia auriculiformis* A. Cunn. ex Benth. Aust J Bot 39(3): 247-260.
- Schimleck LR, Tsuchikawa S (2021) Application of NIR spectroscopy to wood and wood derived products (Chapter 37). Pages 759-780 in E Ciurczak, B Igne, J Workman, D Burns, eds. The Handbook of Near-Infrared Analysis, Fourth Edition. CRC Press, Boca Raton, FL.
- Schwanninger M, Rodrigues JC, Fackler K (2011) A review of band assignments in near infrared spectra of wood and wood components. J Near Infrared Spectrosc 19:287-308.
- Snee R (1977) Validation of regression models: Methods and examples. Technometrics 19:415-428.
- Tsuchikawa S, Kobori H (2015) A review of recent application of near infrared spectroscopy to wood science and technology. J Wood Sci 61(3):213-220.
- Van Duong D, Schimleck LR, Tran DL, Vo HD (2022) Radial and among-clonal variations of stress-wave velocity, wood density, and mechanical properties in 5-year-old *Acacia auriculiformis* clones. BioResources 17(2):2084-2096.
- Villar A, Fernandez S, Gorritxategi E, Ciria JI, Fernandez LA (2014) Optimization of the multivariate calibration of a Vis-NIR sensor for the on-line monitoring of marine diesel engine lubricating oil by variable selection methods. Chemom Intell Lab Syst 130:68-75.