Species- and Moisture-based Sorting of Green Timber Mix with Near Infrared Spectroscopy

Zhu Zhou,[a,b,c,d,*] Sohrab Rahimi,[b] Stavros Avramidis,[b] and Yiming Fang [a,c,d]

Methods suitable for the determination and classification of green timber mix (western hemlock and amabilis fir), with respect to species and moisture content, were developed and tested using near infrared spectroscopy and chemometrics. One thousand two hundred samples were distributed into a calibration set (720 samples) and a prediction set (480 samples). Partial least squares (PLS) and least squares-support vector machines (LS-SVM) for both regression (PLSR and LS-SVR) and classification (PLS-DA and LS-SVC) with different spectral preprocessing methods were implemented. LS-SVM outperformed PLS models for both regression and classification. The coefficient of determination (R^2_p) and root mean square error (RMSEP) of prediction for the best LS-SVR model with spectra pretreated by smooth and first derivative were 0.9824 and 8.7%, respectively, for wood moisture content prediction in the range of 30% to 253%. The best classification model was LS-SVC with spectra pretreated by smooth and second derivative, with overall accuracies of 99.8% in the prediction set, when the samples were divided into four classes. NIRS combined with LS-SVM can be used as a rapid alternative method for qualitative and quantitative analysis of green hem-fir mix before kiln drying. The results could be helpful for sorting green hem-fir mixes with an on-line application.

Keywords: Green timber; Least squares-support vector machines (LS-SVM); Moisture; Near infrared spectroscopy (NIRS); Sorting; Species

*Contact information: a: School of Information Engineering, Zhejiang A&F University, Hangzhou, 311300, Zhejiang, China; b: Department of Wood Science, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada; c: Key Laboratory of State Forestry and Grassland Administration in Forestry Perception Technology and Intelligent Equipment, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China; d: Zhejiang Provincial Key Laboratory of Forestry Intelligent Monitoring and Information Technology, Lin'an 311300, Zhejiang, China; *Corresponding author: zhouzhu@zafu.edu.cn*

INTRODUCTION

Pacific Coast hemlock, commonly known as "hem-fir," is a mixture of two species, namely western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) and amabilis fir (*Abies amabilis* (Dougl.) Forbes). These two species are nearly identical in visual appearance and physical properties and are harvested, processed, and marketed as a species group in lumber packages containing both species (Teal-Jones Group 2019). Due to the high variation in green moisture content (M_g) and basic density (ρ_b), drying hem-fir timber to a uniform final moisture content (M_f) has been recognized by the industry as a key problem (Zhang *et al.* 1996). Species- and/or moisture-based presorting aims to decrease moisture variability between the timber and kiln stacks, thus elevating quality, lowering M_f variation, and reducing drying times. It has been shown to be a viable strategy (Zhang *et al.* 1996; Avramidis *et al.* 2004; Shahverdi *et al.* 2017). Presorting based on species, moisture content, wood zones (heartwood and sapwood), density, or a combination of these

parameters is feasible (Defo *et al*. 2007).

Currently, presorting based on visual, chemical, or mechanical means, such as green weight, density, and timber dimension, has been reported (Defo *et al*. 2007; McClure *et al*. 2015; Deklerck *et al*. 2019). Species-based presorting includes chemical colorimetric methods, electronic nose sensors, DNA markers, and chemical isotopes (Keppler *et al*. 2007; Degen and Fladung 2008; Dawson-Andoh and Adedipe 2012; Yang *et al*. 2015). Presorting based on M_g includes electrical resistance or capacitance, lasers, infrared radiation, and microwaves (Defo *et al*. 2007). Regardless, all options are limited to measuring only a few properties, and some are often laborious, time-consuming, require specialized personnel, or are sometimes hazardous to personnel and the environment (Defo *et al*. 2007; Dawson-Andoh and Adedipe 2012; Yang *et al*. 2015).

Recently, near infrared spectroscopy (NIRS) has been used as a rapid non-destructive technique for qualitative and quantitative analysis of wood and wood by-products (Leblon *et al*. 2013; Tsuchikawa and Kobori 2015). In the field of wood species classification, Adedipe *et al*. (2008) used NIRS and soft independent modelling of class analogy (SIMCA) to separate red (*Quercus rubra*) and white oak (*Quercus alba*) and Dawson-Andoh and Adedipe (2012) used NIRS to separate balsam fir, western hemlock, and white spruce. Yang *et al*. (2012; 2015) tried to classify softwood and hardwood, and wood species from different locations using NIRS and partial least squares discriminant analysis (PLS-DA). Cooper *et al*. (2011) evaluated NIRS with SIMCA and PLS-DA for the separation of the fir from eastern and western spruce-pine-fir (SPF) mix. Sohi *et al*. (2017) successfully determined that NIRS combined with PLS-DA can be used to identify sub-alpine fir from SPF lumber mix in the green chain of sawmills. Lazarescu *et al*. (2017) developed models based on NIRS with PLS-DA and artificial neural networks (ANN) to separate western hemlock from amabilis fir.

The NIRS has also been used to evaluate the moisture content (M) of wood. Hoffmeyer and Pedersen (1995) developed calibrations for the prediction of M below the fiber saturation point (M_{fsp}) in Norway spruce (*Picea abies*) with a coefficient of determination (R^2) higher than 0.90. Cooper *et al*. (2011) predicted the M of western red cedar (*Thuja plicata* Donn ex D. Don). The prediction error for heartwood (M = 0% to 30%) and sapwood (M = 0% to 250%) were approximately ± 2% to 5%, and approximately ± 30%, respectively. Karttunen *et al*. (2008) determined the moisture distribution of Scots pine (*Pinus sylvestris*) green logs *via* NIRS. The models developed by PLS were evaluated using root mean square error of prediction (RMSEP) and found values of 0.8% for heartwood and 10% for sapwood. Adedipe and Dawson-Andoh (2008) examined the feasibility of using NIRS to predict moisture in the range of 0.3% to 80% in yellow poplar (*Liriodendron tulipifera* L.) veneer sheets. The PLS veneer M predictive models showed correlation coefficients of 0.986. Defo *et al*. (2007) evaluated the effects of grain orientation on predicting the moisture content in the range of 68.4% to 99.6% in red oak (*Quercus* spp.). The spectra collected from the transverse and radial surfaces provided better predictions than those collected from the tangential surfaces, with an RMSEP of less than 3.6%. Han *et al*. (2013) successfully estimated the M of frozen and unfrozen logs of black spruce using NIRS. The PLS models showed that the range of correlation coefficients and RMSEP for prediction sets were 0.70 to 0.93 and 7.7% to 16%, respectively, according to log conditions (frozen and unfrozen), wood zones (heartwood and sapwood), and wood surfaces (tangential and transverse). Watanabe *et al*. (2010) used NIRS to detect wet-pockets on the surface of kiln-dried western hemlock (*Tsuga heterophylla*). They also predicted the moisture content of western hemlock with M ranging from 35% to 105% and

subalpine fir (*Abies lasiocarpa*) with *M* ranging from 7% to 210% (Watanabe *et al.* 2011; 2012). Nascimbem *et al*. (2013) used NIRS and chemometrics for classification of the level of moisture in moist wood chips. Four levels of samples with natural moisture (5%), sprinkling (8% to 10%), immersion in water (approximately 50% to 60%), and vaporization of water (approximately 70%) were prepared. A PLS-DA model was developed with a correct rate of 96.6% for the prediction set.

However, there have been few reports on presorting a timber mix with respect to wood species and *M* using NIRS, which combines different chemometrics. In addition, due to spectral variations among wood surfaces and zones (Leblon *et al.* 2013), most of the studies mentioned acquired spectra from specific surfaces and zones. Building a mixed model irrespective of the surface or zone might make the model more robust and reliable. Moreover, the calibration methods for both classification and regression in the literature stated above were either based on PLS or another linear analysis (*e.g.*, SIMCA and PCR). The PLS model only considers the linear relationship between the spectra and wood properties, whereas there might be latent nonlinear information related to wood properties. Only a few previous studies have sought to apply nonlinear analysis for the classification or prediction of the M_g.

Herein, the objectives of this study were (1) to study the ability of NIRS to predict the M_g of hem-fir timber, (2) to sort green hem-fir timber into several classes with respect to species and moisture level, and (3) to obtain the optimal calibration models in comparison with spectral pretreatment methods and calibration methods (partial least squares analysis for regression and classification, and least squares support vector machine for regression and classification).

EXPERIMENTAL

Materials

Two green bolts (1.2 m long) of western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) and amabilis fir (*Abies amabilis* (Dougl.) Forbes) were procured from a local mill in Vancouver, BC, Canada. They were transformed into slabs with a Wood-Mizer sawmilling system (LT15; Wood-Mizer World Headquarters, Indianapolis, IN, USA) and further cut into smaller samples of 100 mm × 100 mm × 10 mm by a sliding table saw (T75; Martin Woodworking Machines Corp., Charlotte, NC, USA).

Each sample was then sectioned into different wood zones, namely sapwood and heartwood, and three wood surfaces (tangential, transverse, and radial) were prepared. The sampling format resulted in 12 combinations (two species by two wood zones by three wood surfaces) and 100 replications of each combination. Therefore, a total of 1200 samples were prepared. Thereafter, all samples were sealed in Ziploc® plastic bags and placed into a cold room at 3 °C to eliminate moisture losses. The samples were left to warm up at room temperature and weighed prior to spectra collection.

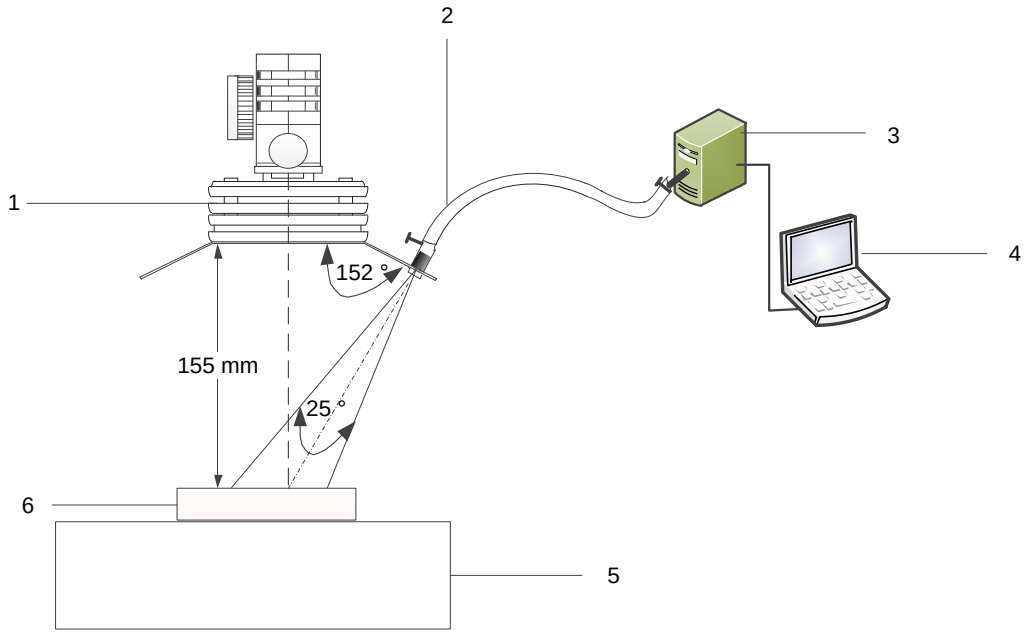
After spectra collection, the samples were oven-dried at 102 °C for 24 h and their moisture content levels were gravimetrically calculated. Consequently, the M_g of the samples ranged from 30% to 253%, with an average of 100.2%.

Samples of each combination were divided into two datasets according to the sample set partitioning based on the joint x-y distance (SPXY) algorithm (Galvão *et al.* 2005). Sixty samples in each combination were used as a calibration set, and the remaining 40 samples were used as a prediction set. Thereafter, the samples in the calibration and

prediction sets of each combination were merged into the final calibration (720 samples) and prediction (480 samples) sets. They were sorted into four classes of two M_g levels, namely fir with low M_g (30% to 100%) (FL), hemlock with low M_g (30% to 100%) (HL), fir with high M_g (100% to 260%) (FH), and hemlock with high M_g (100% to 260%) (HH), according to the average M_g of all the samples.

Spectral Acquisition

The NIR spectra of the wood samples were acquired by a NIR system that consisted of a light source, an optic fiber, a fiber spectrometer (Quality Spec® Pro; Analytical Spectral Devices Inc., Boulder, CO, USA), sample holder, and a computer, as shown in Fig. 1. The optic fiber, which was connected to the spectrometer, was oriented at 62º above the sample surface. The samples were illuminated with a tungsten halogen bulb (ASD Pro Lamp) oriented perpendicular to the sample surface. The distance between the sample surface and the bulb was 155 mm, resulting in NIR spot areas that were approximately 25 mm in diameter. All samples were scanned in the wavelength range of 350 nm to 2500 nm at intervals of 1 nm. A piece of commercial micro-porous Teflon was chosen as the reference material, and the reference spectra were measured and stored prior to spectra collection. The spectrometer parameters' settings, and spectra data collection and storage were recorded *via* RS3 software (Analytical Spectral Devices Inc., Boulder, CO, USA). Two spectra were collected from both the upper and lower surface of each sample, and they were averaged into a single spectrum.



1.Light source 2.Optic fiber 3.Spectrometer 4.Computer 5.Specimen holder 6.Specimen

Fig. 1. Schematic diagram of the NIRS system

Multivariate Analysis

Partial least squares regression (PLSR) and least squares-support vector machine for regression (LS-SVR) were utilized to construct calibration models for M_g prediction. Partial least squares-discriminant analysis and least squares-support vector machine for classification (LS-SVC) were employed to develop classification models with respect to

the species and M_g level. The principles of the mentioned approaches can be found elsewhere in the literature (Liang and Kvalheim 1996; Suykens and Vandewalle 1999). For PLSR and PLS-DA, leave one out cross validation was used to determine the optimum number of latent variables (LVs). For LS-SVR and LS-SVC, all of the 1000 variables (800 nm to 1800 nm) were used as input, and a radial basis function (RBF) kernel was used as the kernel function. Gridsearch technique and leave one out cross validation were used to find the optimal parameter values including the regularization parameter gamma (γ) and the RBF kernel function parameter sig^2 (σ^2).

The performance of PLSR and LS-SVR models was assessed with the coefficient of determination in the calibration (R^2_c), the coefficient of determination in the prediction (R^2_p), the root mean squared error of calibration (RMSEC), and the root mean squared error of prediction (RMSEP). The performance of PLS-LDA and LS-SVC models were evaluated by classification accuracy, which is defined as the percentage of the correct numbers to the total numbers of calibration set or prediction set.

Prior to modelling, all spectra were preprocessed. The smoothing (SM) method of Savitzky-Golay with a segment size of seven and a default polynomial order of zero was first used to decrease the noise; then, three classical spectral methods including first derivative (FD), second derivative (SD), and standard normal variate transformation (SNV) were used to remove undesirable systematic noise (Chu *et al.* 2004). Both FD and SD can remove baseline shifts and superposed peaks (Nicolaï *et al.* 2007). A SNV can eliminate the multiplicative interferences, such as scatter, particle size, and the change of light distance, by correcting both additive and multiplicative scatter effects (Barnes *et al.* 1989).

All algorithms were implemented with MATLAB R2010a software (MathWorks, Natick, MA, USA) and toolboxes including the LS-SVM toolbox (v.1.7), and PLS Toolbox (v.7.8) (Eigenvector Research Inc., Wenatchee, WA, USA).

RESULTS AND DISCUSSION

Moisture Content Distribution

The statistic values of M_g based on SPXY in the calibration and prediction sets for each combination are presented in Table 1. Regardless of the wood surface, the M_g ranged between 30.1% and 59.9%, and between 30.5% and 94.7% in the heartwood of fir and hemlock, respectively. The M_g varied between 99.4% and 253.4%, and between 56.0% and 158.0% in the sapwood of fir and hemlock, respectively. It was evident that the sapwood had a higher M_g compared to heartwood in both species. The M_g ranged from 30.1% to 253.4%, with the average of 99.7% and standard deviation of 62.6%, in the final calibration set, and from 30.8% to 244.6%, with the average of 100.9% and standard deviation of 65.5%, in the final prediction set. This indicated that both covered a large enough range. From Table 1, it can be observed that the moisture range of the final calibration set was bigger than the corresponding prediction set, and the differences between the means and standard deviations of the two sets were minimal. These features were helpful to develop good models.

The distributions of all samples sorted according to the average M_g are shown in Fig. 2. The numbers of FL, FH, HL, and HH were 182, 178, 194, and 166 in the calibration set and 120, 120, 126, and 114 in the prediction set, respectively. The numbers in each class were almost evenly distributed both in the calibration and prediction sets.

Table 1. M_g (%) for Each Combination

Combination	Calibration Set				Prediction Set			
	Min	Max	Mean	SD	Min	Max	Mean	SD
FHC	30.1	41.3	35.1	2.8	30.8	40.4	34.7	2.6
FHR	31.6	42.5	36.1	2.7	33.1	42.5	36.4	2.4
FHT	32.1	59.9	43.2	9.3	32.2	58.2	39.2	7.3
FSC	149.8	213.1	180.3	10.8	175.6	186.5	180.8	2.7
FSR	99.4	191.7	157.7	21.5	147.4	179.8	162.1	11.1
FST	116.5	253.4	192.3	44.2	121.8	244.6	207.3	36.2
HHC	30.9	42.0	36.1	2.7	31.0	41.8	35.1	2.6
HHR	30.5	63.5	44.9	9.0	31.7	58.6	41.4	9.6
HHT	36.4	111.6	65.2	20.3	36.8	86.2	54.3	11.2
HSC	119.1	154.9	144.6	6.3	136.1	151.7	145.3	3.4
HSR	97.8	158.0	129.9	16.2	111.6	157.2	131.2	11.0
HST	56.0	174.5	131.0	39.6	60.7	172.5	142.5	31.9
Total	30.1	253.4	99.7	62.6	30.8	244.6	100.9	65.5

Note: FHC – Heartwood transverse for amabilis fir; FHR – Heartwood radial for amabilis fir; FHT – Heartwood tangential for amabilis fir; FSC – Sapwood transverse for amabilis fir; FSR – Sapwood radial for amabilis fir; FST – Sapwood tangential for amabilis fir; HHC – Heartwood transverse for western hemlock; HHR – Heartwood radial for western hemlock; HHT – Heartwood tangential for western hemlock; HSC – Sapwood transverse for western hemlock; HSR – Sapwood radial for western hemlock; HST – Sapwood tangential for western hemlock

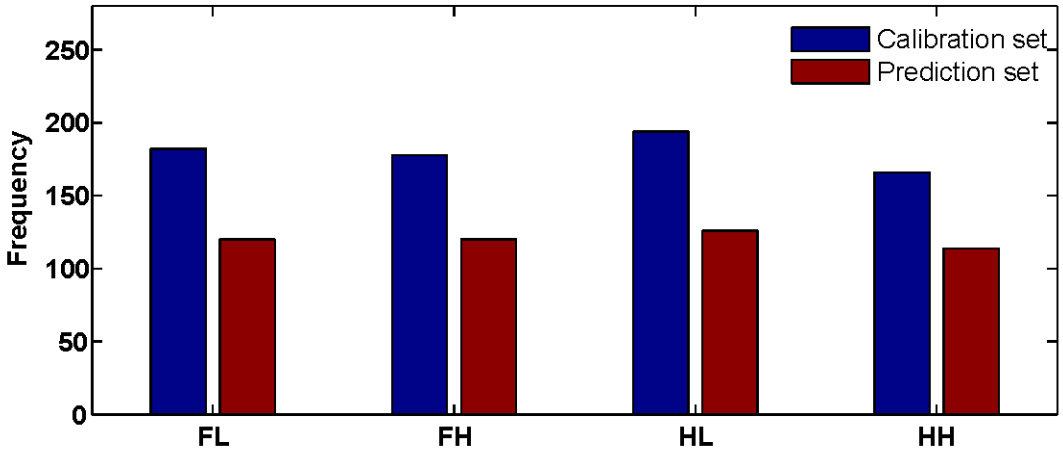


Fig. 2. Distribution of all samples sorted by M_g

Spectral Characteristics

The average preprocessed spectra of the 12 combinations of samples are shown in Fig. 3a to 3d. Due to stationary noise and lack of relevant information in the full range, the wavelength range of the spectra was reduced from 350 nm to 2500 nm to 800 nm to 1800 nm. In the smoothed spectral curve (Fig. 3a), there were some main features of the absorbance spectra around 1200 nm and 1460 nm, which were related to the second overtone of the C-H stretching vibration in cellulose or lignin and the first overtone of O-H stretching in cellulose, hemicellulose, and water (Mehrotra *et al.* 2010). The small peak

centered near 980 nm is attributed to the second overtone of O-H stretch in water (Brøndum *et al.* 2000). The first (Fig. 3b) and second derivative (Fig. 3c) spectra after smooth pretreatment not only reduced the baseline shift but also introduced considerable noise at the beginning and the end of the wavelength range. The great jump and drop of the curves in Fig. 3b to 3c was due to the slope changes of the smoothed spectra. The SNV processing enhanced the features and characteristics of the smoothed spectra (seen in Fig. 3d). The peaks and valleys were almost at the same wavelengths as the smoothed spectra.

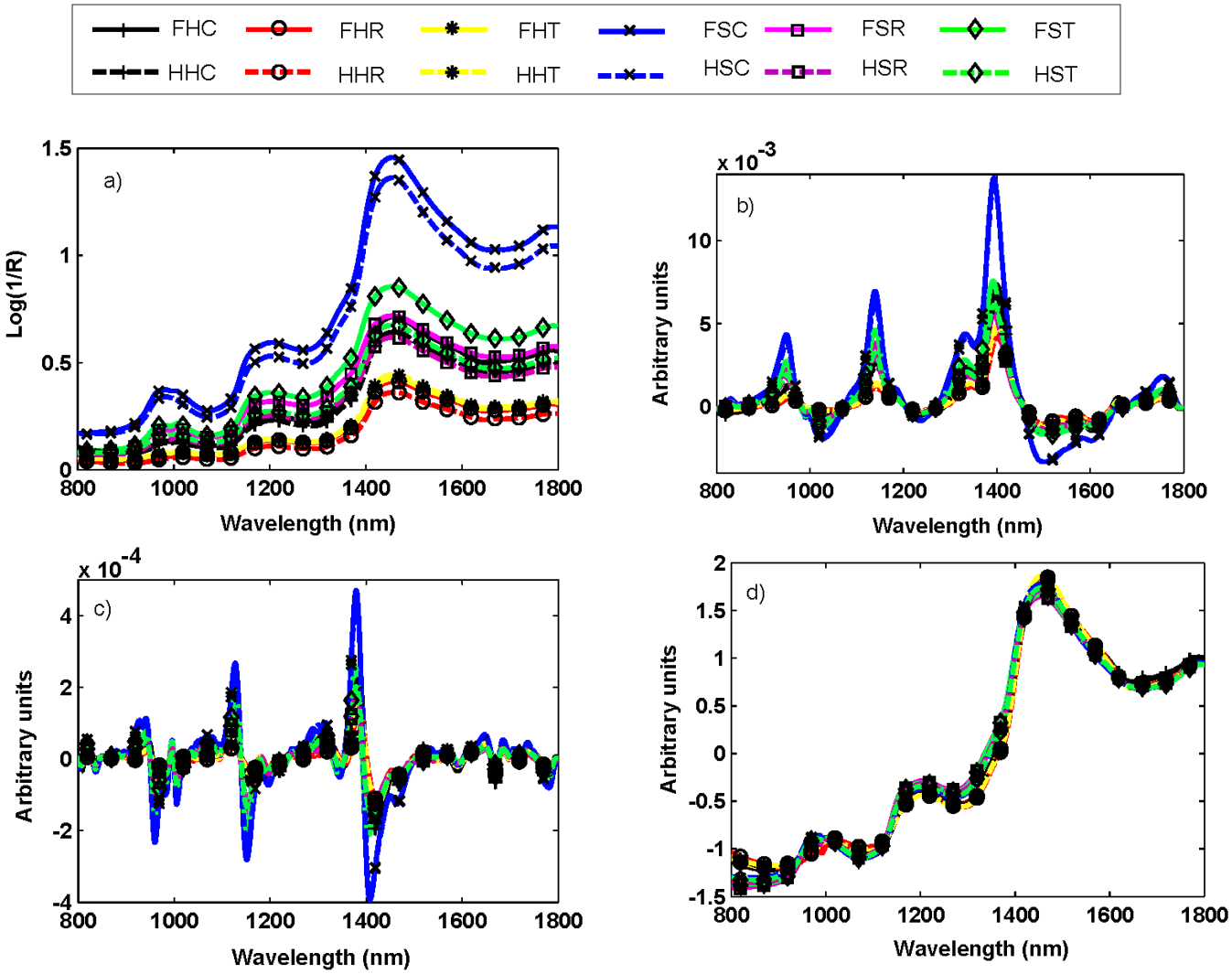


Fig. 3. Mean spectra for each combination with pretreatment: smooth a), smooth + first derivative b), smooth + second derivative c), and smooth + SNV d)

Regarding the wood surface, Fig. 3a shows that the absorbance was the highest in the transverse surface, followed by the tangential and then radial surface for both species. Some researchers reported similar variations and attributed these to the degree of absorption and reflection of NIR by the different wood surface orientations (Defo *et al.* 2007; Adedipe *et al.* 2008). The relatively higher absorbance of the transverse surface may be attributed to the presence of open tracheids and fibers. Light may pass deeper into the transverse surface than the tangential and radial surfaces of wood, leading to less reflectance (Tsuchikawa *et al.* 1996; Defo *et al.* 2007; Adedipe *et al.* 2008). Compared to

radial surfaces, the presence of rays in the tangential surface may explain the relatively high absorbance (Tsuchikawa *et al.* 1996; Defo *et al.* 2007).

Moreover, Fig. 3a also shows that the spectra collected from the sapwood had a higher absorbance than those from the heartwood, irrespective of wood surfaces and species. This may have be explained by the fact that the M_g of the sapwood is usually much higher than that of the heartwood for softwood species (Hans *et al.* 2013). Table 1 shows that the average M_g for the two species heartwood was lower than 65.2%, whereas that of sapwood was higher than 129.9%. Therefore, the higher M_g of the sapwood resulted in a higher NIR absorbance.

Wood Moisture Prediction

Both PLSR and LS-SVR models were established to predict wood moisture content. The performance of the PLSR and LS-SVR models with different pretreatments are presented in Table 2.

Table 2. Results of Moisture Models Developed by PLSR and LS-SVR

Model Method	Pretreatment	Latent Variables	Calibration Set		Prediction Set	
			R^2_c	RMSEC (%)	R^2_p	RMSEP (%)
PLSR	SM	18	0.9524	13.6	0.9646	12.3
	SM + FD	20	0.9620	12.2	0.9628	12.7
	SM + SD	13	0.9596	12.6	0.9597	13.2
	SM + SNV	20	0.9591	12.6	0.9684	11.6
LS-SVR	SM	--	0.9906	6.1	0.9810	9.0
	SM + FD	--	0.9912	5.8	0.9824	8.7
	SM + SD	--	0.9842	7.9	0.9779	9.8
	SM + SNV	--	0.9909	6.0	0.9816	8.9

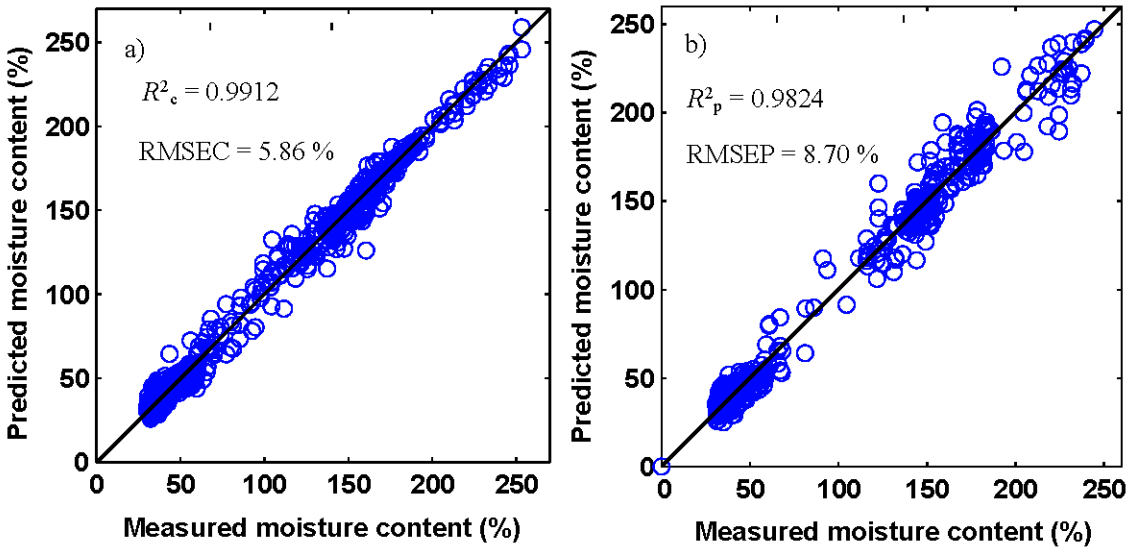


Fig. 4. Scatter plots of measured *versus* predicted M_g by LS-SVR with SM + FD pretreatment: calibration set a) and prediction set b)

From Table 2, it can be observed that although the PLSR model calculated by smooth and first derivative spectra showed the lowest RMSEC value (12.2%) and the highest R^2_c (0.9620), the PLSR model calculated by smooth and SNV spectra had the best predictive ability, given the lowest RMSEP (11.6%) and highest R^2_p (0.9684). Conversely, the LS-SVR model with smooth and first derivative pretreatment performed better than the LS-SVR models with spectra pretreated by other ways in both calibration and prediction sets. The R^2_c, RMSEC, R^2_P, and RMSEP values of the LS-SVR model with the best performance were 0.9912, 5.8%, 0.9824, and 8.7%, respectively. The scatterplots of the predicted *versus* reference values for moisture content prediction by the LS-SVR model with spectra pretreated by smooth and first derivative are shown in Fig. 4a to 4b.

Compared to the PLSR models shown in Table 3, the prediction results obtained by the LS-SVR models outperformed the PLSR models for both the calibration and prediction sets. The reason could have been that some latent nonlinear relationships existed between the spectra and M_g. The PLSR models can only deal with the linear relationships between them, while LS-SVR models could handle certain latent nonlinear information, and the nonlinear information was contributed to the better performance of the LS-SVR models. This result (the LS-SVR model outperforming the PLS model) was in agreement with relevant literature on the prediction of wood parameters, such as quantitative analysis of the density of Chinese Fir using visible/near infrared spectroscopy (Zhu *et al.* 2009), estimation of air-dry density, microfibril angle, stiffness, tracheid coarseness, and tracheid wall thickness in wood radial strip samples based on near infrared diffuse reflectance spectra (Mora and Schimleck 2010), the determination of the modulus of elasticity of *Fraxinus mandschurica* using NIRS (Yu *et al.* 2018), and the detection of microfibril angle using NIRS (Cogdill *et al.* 2004).

Simultaneously, the prediction performance by LS-SVR was also comparable with the following similar studies using NIRS combined with PLSR for wood moisture detection. Cooper *et al*. (2011) predicted the M_g of western red cedar (*Thuja plicata* Donn ex D. Don) sapwood (transverse surface) in a range from 0% to 250% with an RMSEP of 30%. Karttunen *et al*. (2008) determined the moisture content of Scots pine logs on sapwood (transverse surface) in an M_g range of 110% to 160% with an RMSEP of 10%. Han *et al*. (2013) monitored the moisture content in black spruce logs according to the log temperature condition (frozen and/or unfrozen), wood zone (sapwood or heartwood), and wood surface (transverse or tangential) in a M_g range of 9.0% to 132%, the ranges of R^2_p, RMSEP for frozen and unfrozen samples were 0.78 to 0.93, 7.7 % to 3.4%, 0.70 to 0.87, and 10.6% to 16%, respectively. The NIR spectrum acquired in the sources cited above were restricted to the specific wood zone and/or surface, which can decrease the effect of spectra variation among different wood zones and surfaces and benefit the building of prediction models. Compared with the above studies, despite samples with three types of wood surfaces and the use of two types of wood zones, the authors' models seemed to have good accuracy. An excellent moisture content prediction with an RMSEP of 5.70% in western hemlock (*Tsuga heterophylla* (Raf.) Sarg) lumber with an M_g ranging from 35% to 105% was obtained by Watanabe *et al*. (2011) using 72 samples with the same wood types as was used in the current study. However, considering the number of species (hemlock and fir) and samples (1200) and large M_g range (30% to 253%) analyzed, the authors' models could be considered very good. Therefore, the results indicated that near infrared spectroscopy combined with LS-SVR could be utilized as a good prediction method for the determination of wood moisture content in the green state.

Performance of Species- and Moisture-based Sorting of Hem-Fir Timbers Mix

The PLS-DA and LS-SVC models were established to classify specimens into different types of groups based on species and M_g level. The overall accuracies of the PLS-DA and LS-SVC models with different pretreatments are shown in Fig. 5.

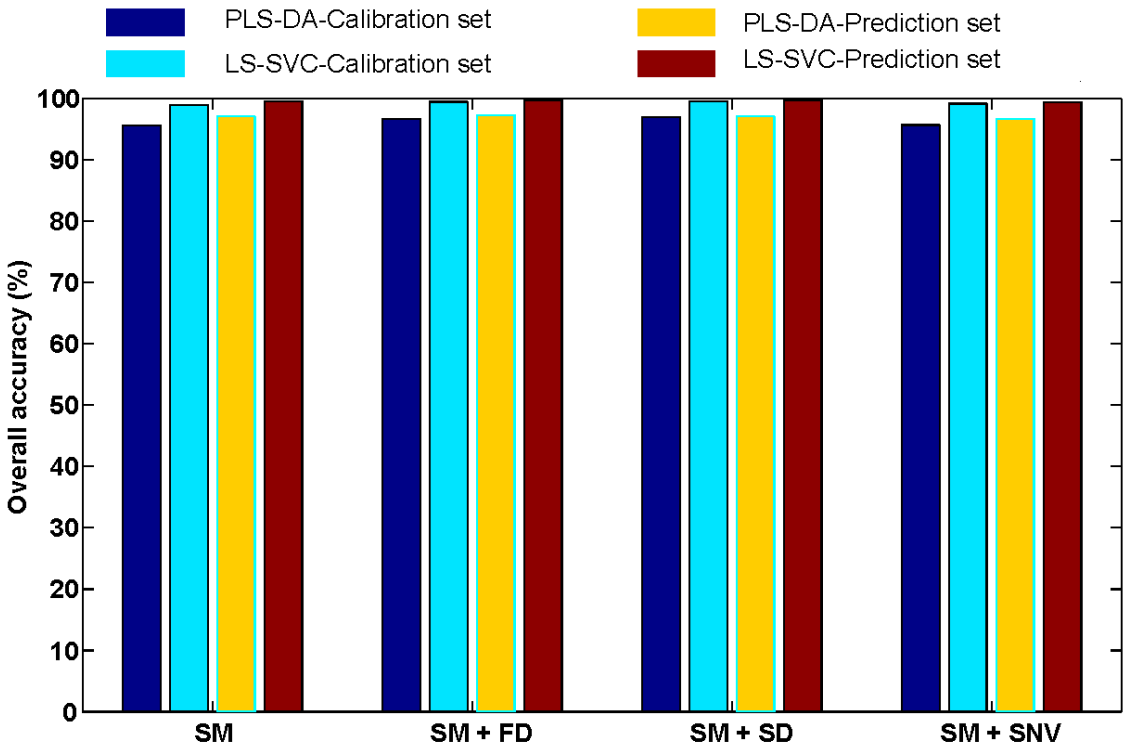


Fig. 5. Performance of sorting models developed by PLS-DA and LS-SVC based on species and M_g level

When the samples were divided into two M_g levels, it was also observed that although the PLS-DA model calculated by smooth and second derivative spectra showed the highest overall accuracy (96.9%) in the calibration set, the PLS-DA model calculated by smooth and first derivative spectra had the best predictive ability given the highest overall accuracy (97.3%) in the prediction set.

The LS-SVC models with smooth and first derivative pretreatment and with smooth and second derivative pretreatment outperformed the other LS-SVC models in the prediction set with an overall accuracy of 99.8%; however, the latter model showed the highest overall accuracy (99.6%) in the calibration set.

Table 3 lists the classification results of the best PLS-DA and LS-SVC models. In both PLS-DA and LS-SVC classification, there was an overlapping classification, resulting in some samples not being uniquely classified, and in fact, some individual samples may have been classified into two or more classes. Therefore, each class has both its right and wrong classified number of samples presented in Table 3. The numbers of unclassified samples in the calibration and prediction sets for the PLS-DA model were 11 and 9, respectively, while there was no sample unclassified for the LS-SVC model. Moreover, only 3 and 1 samples were misclassified in the calibration and prediction sets, respectively.

Table 3. Classification Results of the Best PLSR and LS-SVR Models Based on Species and M_g level

Model Method	Pretreatment	Class	Calibration Set			Prediction Set		
			Right	Wrong	Not Classified	Right	Wrong	Not Classified
PLS-DA	SM + FD	FL_2	180	2	0	120	0	0
		FH_2	175	0	3	115	1	4
		HL_2	181	8	5	121	3	2
		HH_2	160	3	3	111	0	3
LS-SVC	SM + SD	FL_2	181	1	0	120	0	0
		FH_2	178	0	0	120	0	0
		HL_2	194	0	0	126	0	0
		HH_2	164	2	0	113	1	0

Compared to the PLS-DA model shown in Fig. 5 and Table 3, the LS-SVC models had a better capability than the PLS-DA models to classify samples based on wood species and M_g level. The probable reason might have been that the LS-SVC took advantage of the latent nonlinear information (wood surfaces, wood zones, *etc*.) of the spectra data, which contributed better results, whereas PLS-DA only dealt with the linear relationship between the spectra and the different wood species and M_g levels. Therefore, the results indicated that near infrared spectroscopy together with LS-SVC could be utilized as a good method for the classification of green hem-fir timber mix based on species and M_g level.

CONCLUSIONS

1. Near infrared spectroscopy combined with chemometrics of PLSR, LS-SVR, PLS-DA, and LS-SVC models was found to be an efficient technique for predicting the moisture content of green hem-fir timber mix and classifying green hem-fir timber mix with respect to species and M_g level in the wavelength range of 800 nm to 1800 nm.
2. The LS-SVR models performed better than PLSR models for the prediction of M_g. The best model for M_g prediction was LS-SVR with spectra pretreated by smooth and first derivative. The R^2_p and RMSEP were 0.9824 and 8.7%, respectively, in the M_g range of 30% to 253%.
3. The LS-SVC models performed better than the PLS-DA models for the classification of green hem-fir timber mix. The best model for classification was LS-SVC with spectra pretreated by smooth and second derivative with an overall accuracy of 99.8% in the prediction set.

ACKNOWLEDGMENTS

This work was supported by the State Scholarship Fund of China Scholarship Council (No. 201708330485), the foundations from Zhejiang Provincial Science and Technology Department (No. LGG18F010006), the Postdoctoral Program Foundation of

Zhejiang Province (No. ZJ20180156). Many thanks to Brandon Chan, Pablo Chung, and Joseph Kim of the UBC-Centre for Advanced Wood Processing for wood sample preparation and Professor Shawn Mansfield for allowing the authors to use his NIR spectrometer.

REFERENCES CITED

Adedipe, O. E., and Dawson-Andoh, B. (2008). "Predicting moisture content of yellow-poplar (*Liriodendron tulipifera* L.) veneer using near infrared spectroscopy," *Forest Products Journal* 58(4), 28-33.

Adedipe, O. E., Dawson-Andoh, B., Slahor, J. 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

Avramidis, S., Aune, J., and Oliveira, L. (2004). "Exploring pre-sorting and re-drying strategies for pacific coast hemlock square timbers," *Journal of the Institute of Wood Science* 16(4), 189-198.

Barnes, R. J., Dhanoa, M. S., and Lister, S. J. (1989). "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Applied Spectroscopy* 43(5), 772-777. DOI: 10.1366/0003702894202201

Brøndum, J., Munck, L., Henckel, P., Karlsson, A., Tornberg, E., and Engelsen, S. B. (2000). "Prediction of water-holding capacity and composition of porcine meat by comparative spectroscopy," *Meat Science* 55(2), 177-185. DOI: 10.1016/S0309-1740(99)00141-2

Chu, X., Yuan, H. F., and Lu, W. Z. (2004). "Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique," *Progress in Chemistry* 16(4), 528-542. DOI: 10.1016/j.jco.2003.08.015

Cogdill, R. P., Schimleck, L. R., Jones, P. D., Peter, G. F., Daniels, R. F., and Clark, A. (2004). "Estimation of the physical wood properties of *Pinus taeda* L. radial strips using least squares support vector machines," *Journal of Near Infrared Spectroscopy* 12(4), 263-269. DOI: 10.1255/jnirs.434

Cooper, P. A., Jeremic, D., Radivojevic, S., Ung, Y. T., and Leblon, B. (2011). "Potential of near-infrared spectroscopy to characterize wood products," *Canadian Journal of Forest* R*esearch* 41(11), 2150-2157. DOI: 10.1139/x11-088

Dawson-Andoh, B. E., and Adedipe, O. E. (2012). "Rapid spectroscopic separation of three Canadian softwoods," *Wood Science and Technology* 46(6), 1193-1202. DOI: 10.1007/s00226-012-0468-9

Defo, M., Taylor, A. M., and Bond, B. (2007). "Determination of moisture content and density of fresh-sawn red oak lumber by near infrared spectroscopy," *Forest Products Journal* 57(5), 68-72.

Degen, B., and Fladung, M. (2008). "Use of DNA-markers for tracing illegal logging," *Proceeding of the International Workshop Fingerprinting Methods for the Identification of Timber Origins* 321, 6-14.

Deklerck, V., Mortier, T., Goeders, N., Cody, R. B., Waegeman, W., Espinoza, E., Acker, J. V., Bulcke, J. V. d., and Beeckman, H. (2019). "A protocol for automated timber species identification using metabolome profiling," *Wood Science and Technology* 53(4), 953-965. DOI: 10.1007/s00226-019-01111-1

Galvão, R. K. H., Araujo, M. C. U., José, G. E., Pontes, M. J. C., Silva, E. C., and Saldanha, T. C. B. (2005). "A method for calibration and validation subset partitioning," *Talanta* 67(4), 736-740. DOI: 10.1016/j.talanta.2005.03.025

Hans, G., Leblon, B., Stirling, R., Nader, J., LaRocque, A., and Cooper, P. A. (2013). "Monitoring of moisture content and basic specific gravity in black spruce logs using a hand-held MEMS-based near-infrared spectrometer," *The Forestry Chronicle* 89(5), 607-620. DOI: 10.5558/tfc2013-112

Hoffmeyer, P., and Pedersen, J. G. (1995). "Evaluation of density and strength of Norway spruce wood by near infrared reflectance spectroscopy," *Holz als* R*oh- und Werkstoff* 53(3), 165-170. DOI: 10.1007/BF02716418

Karttunen, K., Leinonen, A., and Sarén, M.-P. (2008). "A survey of moisture distribution in two sets of Scots pine logs by NIR-spectroscopy," *Holzforschung* 62(4), 435-440. DOI: 10.1515/HF.2008.060

Keppler, F., Harper, D. B., Kalin, R. M., Meier-Augenstein, W., Farmer, N., Davis, S., Schmidt, H.-L., Brown, D. M., and Hamilton, J. T. G. (2007). "Stable hydrogen isotope ratios of lignin methoxyl groups as a paleoclimate proxy and constraint of the geographical origin of wood," *New Phytologist* 176(3),600-609. DOI: 10.1111/j.1469-8137.2007.02213.x

Lazarescu, C., Hart, F., Pirouz, Z., Panagiotidis, K., Mansfield, S. D., Barrett, J. D., and Avramidis, S. (2017). "Wood species identification by near-infrared spectroscopy," *International Wood Products Journal* 8(1), 32-35. DOI: 10.1080/20426445.2016.1242270

Leblon, B., Adedipe, O., Hans, G., Haddadi, A., Tsuchikawa, S., Burger, J., Stirling, R., Pirouz, Z., Groves, K., Nader, J., *et al.* (2013). "A review of near-infrared spectroscopy for monitoring moisture content and density of solid wood," *The Forestry Chronicle* 89(5), 595-606. DOI: 10.5558/tfc2013-111

Liang, Y.-Z., and Kvalheim, O. M. (1996). "Robust methods for multivariate analysis — a tutorial review," *Chemometrics and Intelligent Laboratory Systems* 32(1), 1-10. DOI: 10.1016/0169-7439(95)00006-2

McClure, P. J., Chavarria, G. D., and Espinoza, E. (2015). "Metabolic chemotypes of CITES protected Dalbergia timbers from Africa, Madagascar, and Asia," *Rapid Communications in Mass Spectrometry* 29(9), 783-788. DOI: 10.1002/rcm.7163

Mehrotra, R., Singh, P., and Kandpal, H. (2010). "Near infrared spectroscopic investigation of the thermal degradation of wood," *Thermochimica Acta* 507-508, 60-65. DOI: 10.1016/j.tca.2010.05.001

Mora, C. R., and Schimleck, L. R. (2010). "Kernel regression methods for the prediction of wood properties of *Pinus taeda* using near infrared spectroscopy," *Wood Science and Technology* 44(4), 561-578. DOI: 10.1007/s00226-009-0299-5

Nascimbem, L. B. L. R., Rubini, B. R., and Poppi, R. J. (2013). "Determination of quality parameters in moist wood chips by near infrared spectroscopy combining PLS-DA and support vector machines," *Journal of Wood Chemistry and Technology* 33(4), 247-257. DOI: 10.1080/02773813.2013.783075

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., and Lammertyn, J. (2007). "Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review," *Postharvest Biology and Technology* 46(2), 99-118. DOI: 10.1016/j.postharvbio.2007.06.024

Shahverdi, M., Oliveira, L., and Avramidis, S. (2017). "Kiln-drying optimization for quality pacific coast hemlock timber," *Drying Technology* 35(15), 1867-1873. DOI:

10.1080/07373937.2017.1283324
Sohi, A., Avramidis, S., and Mansfield, S. (2017). "Near-infrared spectroscopic separation of green chain sub-alpine fir lumber from a spruce-pine-fir mix," *BioResources* 12(2), 3720-3727. DOI: 10.15376/biores.12.2.3720-3727
Suykens, J. A. K., and Vandewalle, J. (1999). "Least squares support vector machine classifiers," *Neural Processing Letters* 9(3), 293-300. DOI: 10.1023/A:1018628609742
Teal-Jones Group (2019). "Hem-fir," (https://tealjones.com/our-products/whitewood-lumber/hem-fir/overview/), Accessed 1 Jan 2019.
Tsuchikawa, S., Hayashi, K., and Tsutsumi, S. (1996). "Nondestructive measurement of the subsurface structure of biological material having cellular structure by using near-infrared spectroscopy," *Applied Spectroscopy* 50(9), 1117-1124. DOI: 10.1366/0003702963905114
Tsuchikawa, S., and Kobori, H. (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
Watanabe, K., Hart, F., Mansfield, S. D., and Avramidis, S. (2010). "Detection of wet-pockets on the surface of *Tsuga heterophylla* (Raf.) Sarg. by near infrared (NIR) spectroscopy," *Holzforschung* 64(1), 55-60. DOI: 10.1515/hf.2010.014
Watanabe, K., Mansfield, S. D., and Avramidis, S. (2011). "Application of near-infrared spectroscopy for moisture-based sorting of green hem-fir timber," *Journal of Wood Science* 57(4), 288-294. DOI: 10.1007/s10086-011-1181-2
Watanabe, K., Mansfield, S. D., and Avramidis, S. (2012). "Wet-pocket classification in *Abies lasiocarpa* using spectroscopy in the visible and near infrared range," *European Journal of Wood and Wood Products* 70(1-3), 61-67. DOI: 10.1007/s00107-010-0490-2
Yang, Z., Liu, Y., Pang, X., and Li, K. (2015). "Preliminary investigation into the identification of wood species from different locations by near infrared spectroscopy," *BioResources* 10(4), 8505-8517. DOI: 10.15376/biores.10.4.8505-8517
Yang, Z., Lü, B., Huang, A.-M., Liu, Y.-N., and Xie, X.-Q. (2012). "Rapid identification of softwood and hardwood by near infrared spectroscopy of cross-sectional surfaces," *Spectroscopy and Spectral Analysis* 32(7), 1785-1789. DOI: 10.3964/j.issn.1000-0593(2012)07-1785-05
Yu, H., Liang, H., Lin, X., and Zhang, Y. (2018). "Nondestructive determination of the modulus of elasticity of *Fraxinus mandschurica* using near-infrared spectroscopy," *Optical Engineering* 57(4), Article ID 046103. DOI: 10.1117/1.OE.57.4.046103
Zhang, Y., Oliveira, L., and Avramidis, S. (1996). "Drying characteristics of hem-fir squares as affected by species and basic density presorting," *Forest Products Journal* 46(2), 44-50.
Zhu, X., Shan, Y., Li, G., Huang, A., and Zhang, Z. (2009). "Prediction of wood property in Chinese Fir based on visible/near-infrared spectroscopy and least square-support vector machine," *Spectrochimica Acta Part A*: *Molecular and Biomolecular Spectroscopy* 74(2), 344-348. DOI: 10.1016/j.saa.2009.06.008

Article submitted: May 29, 2019; Peer review completed: Oct. 1, 2019; Revised version received: November 14, 2019; Accepted: November 15, 2019; Published: November 18, 2019.
DOI: 10.15376/biores.15.1.317-330