L.R. Schimleck, G.R. Hodge and W. Woodbridge, *J. Near Infrared Spectrosc.* **18**, 355–365 (2010)
Received: 15 April 2010 ■ Revised: 29 July 2010 ■ Accepted: 30 July 2010 ■ Publication: 25 November 2010

JNIRS JOURNAL OF NEAR INFRARED SPECTROSCOPY
Special Issue on Wood and Wood Products

Toward global calibrations for estimating the wood properties of tropical, sub-tropical and temperate pine species

Laurence R. Schimleck,[a,*] Gary R. Hodge[b] and William Woodbridge[b]

[a]Warnell School of Forestry and Natural Resources, The University of Georgia, Athens, GA 30602, USA. E-mail: lschimleck@warnell.uga.edu

[b]Camcore and Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA

Near infrared (NIR) spectroscopy is increasingly being used to replace traditional methods of wood property assessment and, as a result, multi-site, multi-species (or global) calibrations are of interest to organisations who assess wood properties on a large-scale. In this study, the development of global wood property calibrations for tropical, sub-tropical and temperate pines is explored. In a three-stage process, samples provided by ten forest industry companies and representing 14 pine species (two species had two varieties) and one hybrid, grown on 17 different sites in five countries (Argentina, Brazil, Chile, Colombia and South Africa) were used to develop calibrations for air-dry density, microfibril angle (MFA) and stiffness. Initial calibrations, based on samples from individual co-operators, had variable statistics; for example, R^2 for density ranged from 0.51 to 0.89. In the second stage, multi co-operator calibrations were obtained from two independent data sets that included samples from sites in each country. For the three properties, calibration statistics compared well to those obtained in stage 1, with stiffness having the best fit statistics (R^2=0.917, $RPDc$=3.3). MFA and stiffness calibrations showed the presence of nonlinearity in the data. The calibrations performed reasonably well when used to predict the wood properties of the alternate set, with density having the weakest predictions. Stage 3 calibrations were based on all available samples and were similar to those developed in stage 2, demonstrating that it is possible to build multi-site, multi-species calibrations for pines.

Keywords: air-dry density, microfibril angle (MFA), near infrared (NIR) spectroscopy, *Pinus* species, stiffness, tree improvement, wood properties

Introduction

Near infrared (NIR) spectroscopy is increasingly being used to replace traditional methods of wood property assessment in tree improvement programmes where it can rapidly provide wood property data for large numbers of samples. Generally, the goal has been to develop broad-based calibrations that encompass as much of the expected variation as possible. Typically, these broad-based models are referred to as global calibrations. While such calibrations can not be expected to be as accurate those that are site or species specific[1] they can be used for ranking purposes and have wide appeal to many users who would prefer to avoid the need to continually update and refine their predictive models. In addition, the use of large sample sets in agricultural applications indicates that calibrations based on large sample sets are more robust.[2]

While rare in forestry-related studies, multiple-site, multiple-species calibrations are common in agricultural-related research. Over many years, large spectral databases have been created, with an excess of 1000 spectra per commodity.[3] Sharing of samples and NIR spectra between research groups also frequently occurs, facilitating the development

ISSN: 0967-0335
doi: 10.1255/jnirs.900

of extremely large spectral databases. For example, Berzaghi *et al.*[2] reported the development of multiple-site forage calibrations based on a dataset of over 21,000 samples from Australia, Belgium, Canada, Germany, Italy, Sweden and the USA.

The scale of this work has yet to be duplicated in wood-related research and Schimleck[4] attributed this to several factors including the short history of wood-related NIR research, the failure of NIR to be adopted on a large-scale for the estimation of wood properties, the cost and difficulty of determining many wood properties and the lack of a consistent method for the determination of some wood properties, for example, pulp yield. A further barrier that Schimleck[4] did not identify is the difficulty of sharing samples between forestry-related companies (who frequently use NIR spectroscopy in their breeding programmes) and research groups owing to concerns over intellectual property.

However, there are examples of wood chemistry-related calibrations based on multiple sites and multiple species, and while not global in their size, they should be noted. Early examples include Garbutt *et al.*[5] who reported cellulose and lignin calibrations based on mixed species (13 eucalypts, one hybrid) and Michell[6] who developed calibrations for several wood and pulp quality parameters using native forest grown *Eucalyptus globulus* from 10 locations in Tasmania, Australia. Later studies included more species and sites,[7,8] while, more recently, Downes *et al.*[9,10] have reported calibrations for cellulose and pulp yield based on eucalypt samples from across the entire temperate zone of Australia.

Studies have also attempted to develop multiple-site calibrations for estimating the physical–mechanical properties of wood.[11,12] Both studies were based on a single species (*Pinus taeda*) but included samples from several different sites. The study by Jones *et al.*[11] is notable as it was based on samples analysed by the SilviScan instruments. The SilviScan instruments, developed by CSIRO (Australia), utilise X-ray densitometry, diffractometry and image analysis to accurately measure several wood properties at high spatial resolution.[13–15] In addition, the analysis, compared to many other wood properties, is of relatively low cost, making it possible to build large calibration sets based on SilviScan tested samples. This is particularly true if NIR spectra are collected from multiple adjacent sections of the wood strip analysed by SilviScan.

As noted earlier, the development of global wood property calibrations is of great interest, particularly to organisations with tree improvement programmes that frequently have to assess the wood properties of many species, grown on a range of sites. One such organisation is Camcore, a non-profit, international programme based at North Carolina State University (www.camcore.org), which is involved in the conservation of tropical and subtropical forest tree species. Camcore conduct considerable research on *Pinus* species and has established extensive field trials in several countries which aim to evaluate the growth, wood properties and adaptability of many pine species and provide information on their suitability for deployment in plantations. Camcore work with a number of forest industry related co-operators and establish trials on their lands. Hence, plantations can be established with the same species but in several different countries providing a unique opportunity to explore geographically diverse, multiple-species, multiple-site wood property calibrations.

Recently, Camcore initiated a study that aimed to examine the wood properties (air-dry density, microfibril angle (MFA) and stiffness) of 14 pine species (two species had two varieties) and one hybrid, grown on 17 different sites in five countries (Argentina, Brazil, Chile, Colombia and South Africa). Ten different industry co-operators provided core samples for analysis. Owing to the large number of samples involved and the prohibitive cost of analysing all the samples, the wood properties of the samples were to be determined using NIR spectroscopy. The development of wood property calibrations based on the selected cores is described in this study.

The first phase of this study was the establishment of multi-pine species wood property calibrations for each co-operator using a sub-sample of the cores. The calibrations were used to predict the wood properties of the remaining cores and this data was compiled and used to examine radial trends (pith-to-bark variation) in wood properties for the pine species/varieties/hybrids (this component of the study will be the subject of a separate manuscript). Following the development of individual co-operator calibrations, multi co-operator calibrations (sites from each country were included in the calibration) were explored and tested on a separate test set. Finally, samples from all co-operators were combined and used to explore the development of global NIR calibrations for predicting air-dry density, MFA and stiffness.

Material and methods

Sample origin

A total of 3323 breast height cores (12 mm in diameter) were provided by ten forest industry co-operators in five countries: Alto Paraná and Bosques del Plata (Argentina), Klabin and Rigesa (Brazil), Arauco Bioforest and CMPC Forestal Mininco (Chile), Smurfit Kappa Cartón de Colombia (Colombia) and Komatiland Forests, Sappi Forests and Mondi (South Africa). The cores represented 14 pine species (two species had two varieties) and one hybrid and were grown on 17 different sites. A sub-sample of 370 cores was selected for SilviScan/NIR analysis and the development of wood property calibrations. The origin of the selected core species included by location and the number of cores per species are summarised in Table 1.

Preparation of radial strips

Increment cores were dried, glued to core holders and sawn into radial strips using a twin blade saw. Strip dimensions were 2 mm tangentially and 7 mm longitudinally; radial length was determined by the pith-to-bark length of the sample.

Table 1. Summary of samples utilised for the development of wood property calibrations and their origin.

Country and co-operator	Site	Age	Lat.	Long	Elev (m)	Precip (mm)	Sample*																	
							A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Argentina																								
ARG-1	Timbauva	8	27° 59′ S	56° 01′ W	127	1470	5					6				5				5	5			26
ARG-2	El Piray	8	26° 02′ S	54° 37′ W	195	1965			5	5		5				5		6		5	5	5		41
Brazil																								
BRZ-1	Imbauzinho124A	17–18	24° 16′ S	50° 38′ W	780	1473		5								5				5	5			20
BRZ-2	São João da Barra	10	26° 07′ S	50° 10′ W	780	1713			6			5						6		6				23
Chile																								
CHL-1	El Salto	15	39° 59′ S	72° 45′ W	250	1542					5	5			5			5	5		5			30
	Los Ríos	15	37° 30′ S	73° 24′ W	200	1437							5						6					11
CHL-2	Malvén	9	38° 36′ S	72° 24′ W	93	1350						5						5	6					16
	San Juan	14	37° 02′ S	72° 14′ W	160	1100							5		5				5					15
Colombia																								
COL-1	San Benito	14	2° 35′ N	76° 33′ W	1848	2155										5	3	5			6	5		24
COL-2	El Retiro	18	2° 37′ N	76° 33′ W	1854	2176								5		5	5	5						20
COL-3	Sombreros	17	2° 17′ N	76° 34′ W	2648	2105										5	1	7				5		18
South Africa																								
SAF-1	Tweefontein	12–22	25° 02′ S	30° 47′ E	1155	1953										5		5			5		6	21
	Wilgeboom	11–20	24° 58′ S	30° 57′ E	960	1253		5										5				5		15
	Witklip	11–20	25° 11′ S	30° 59′ E	952	1194			6	5						5		5		5	5	5		36
SAF-2	Elburg	10	31° 06′ S	28° 08′ E	1360	712					5	5						4						14
	Langweis	14	31° 11′ S	28° 12′ E	1320	712									6			5						11
SAF-3	Helvetia	10	25° 32′ S	30° 22′ E	1680	770					5	6	6		6			6						29
Total							5	10	17	10	15	37	16	5	22	40	9	69	22	26	36	25	6	370

*Samples were from: A, *P. caribaea*; B, *P. chiapensis*; C, *P. elliottii*; D, *P. ell* × *P. car*; E, *P. greggii* N; F, *P. greggii* S; G, *P. herrerae*; H. *P. kesiya*; I, *P. leiophylla*; J. *P. maximinoi*; K. *P. oocarpa*; L. *P. patula*; M. *P. radiata*; N. *P. taeda*; O. *P. tecunumanii* HE; P, *P. tecunumanii* LE; Q, *P. teocote*; R, total trees.

Note: there are two varieties of *P. greggii* recognised, a northern (N) variety and southern (S) variety. There are also two varieties of *P. tecunumanii* recognised, a high elevation (HE) variety and a low elevation (LE) variety.

Near infrared spectroscopy

Diffuse reflectance NIR spectra were collected from the radial–longitudinal face of each strip in 10 mm increments using a Foss NIRSystems (Laurel, MD, USA) Model 5000 scanning spectrometer. Samples were held in a custom-made holder that had been fitted with a 5 mm × 10 mm mask (the mask was used to ensure that an area of constant size was analysed). For each radial strip the first spectrum was collected from the bark end of the sample. The spectra were collected at 2 nm intervals over the wavelength range 1100–2498 nm. The instrument reference was a ceramic standard. Fifty scans were accumulated for each 10 mm section; these scans were averaged to give a single spectrum per section. All measurements were made in a controlled environment of 40% relative humidity and a temperature of 20°C.

SilviScan analysis

The 370 radial strips were analysed by SilviScan at FPInnovations Paprican (Vancouver, Canada). For each strip, the following wood properties were measured:

- Air-dry density in 25 µm steps using X-ray densitometry
- MFA over 5 mm intervals using scanning X-ray diffractometry[13,14]
- Wood stiffness estimated at the same resolution as MFA by combining X-ray densitometry and X-ray diffraction data[15]

As with NIR analysis, the SilviScan measurements were made in a controlled environment of 40% relative humidity and a temperature of 20°C. The high-resolution SilviScan data were averaged over 10 mm sections from pith-to-bark for correlation with the NIR spectra. A statistical summary of measured wood properties for samples provided by each co-operator is given in Table 2.

Table 2. Summary statistics for each wood property and industrial co-operator. The data presented refers to individual 10 mm increments.

Company	Property	Minimum	Maximum	Average	Standard deviation
AL	Air-dry density (kg m^{-3})	238	723	435	85
	MFA (°)	9.9	44.5	25.4	6.1
	SilviScan stiffness (GPa)	1.2	16.0	6.5	3.0
AR	Air-dry density (kg m^{-3})	312	601	415	48
	MFA (°)	10.6	45.5	27.5	6.9
	SilviScan stiffness (GPa)	1.6	19.1	5.6	2.9
BP	Air-dry density (kg m^{-3})	284	721	425	71
	MFA (°)	9.7	42.4	27.4	6.0
	SilviScan stiffness (GPa)	1.3	15.7	5.6	2.6
KB	Air-dry density (kg m^{-3})	304	806	490	101
	MFA (°)	9.0	40.1	20.8	6.9
	SilviScan stiffness (GPa)	2.0	20.6	9.5	4.1
KLF	Air-dry density (kg m^{-3})	298	724	468	89
	MFA (°)	8.0	44.6	21.8	7.4
	SilviScan stiffness (GPa)	1.1	23.4	8.6	4.4
MC	Air-dry density (kg m^{-3})	304	545	416	46
	MFA (°)	4.8	42.7	28.7	6.0
	SilviScan stiffness (GPa)	1.9	12.8	5.1	2.1
MD	Air-dry density (kg m^{-3})	273	611	410	51
	MFA (°)	11.2	48.0	27.9	7.6
	SilviScan stiffness (GPa)	1.8	17.4	5.3	2.5
RG	Air-dry density (kg m^{-3})	295	613	432	74
	MFA (°)	16.0	42.1	29.6	4.7
	SilviScan stiffness (GPa)	2.0	11.0	4.8	1.8
SC	Air-dry density (kg m^{-3})	266	835	468	107
	MFA (°)	7.4	40.8	21.4	6.6
	SilviScan stiffness (GPa)	2.4	28.3	9.1	5.0
SP	Air-dry density (kg m^{-3})	298	518	415	53
	MFA (°)	1.2	46.0	27.6	6.1
	SilviScan stiffness (GPa)	1.9	12.6	4.9	2.2

Wood property calibrations

Wood property calibrations were developed for three properties: air-dry density, MFA and stiffness. The Unscrambler software (version 9.2) and partial least squares (PLS) regression (with four cross-validation segments) was used to build the calibrations. For each calibration, the number of factors used was recommended by the software.

The calibrations were built in three stages:

- Stage 1. Multiple pine species calibrations were developed for each individual co-operator. For these calibrations a maximum of ten factors were used.
- Stage 2. Multi co-operator calibrations that included a mix of sites from each country were then obtained. There was a range of possible options for building these calibrations but the approach we chose involved splitting the samples into two sets: one for calibration, the other prediction, that included samples from each country but from different co-operators, i.e. CHL-1 and CHL-2 are both from Chile but were in different sets. The exception was Colombia, where all samples were provided by one co-operator. However, plantations were established at different locations and location became the basis for splitting the samples into two sets. Set 1 included samples from ARG-1, BRZ-1, CHL-1, COL-1 and SAF-1, while set 2 included samples from ARG-2, BRZ-2, CHL-2, COL-2 and -3 and SAF-2 and -3. Wood property calibrations were developed for the two sets and tested on each other. Owing to the greater variability represented by the expanded set, 15 factors were used to develop calibrations;
- Stage 3. The final stage involved developing wood property calibrations using all samples, i.e. samples provided by the 10 co-operators in five countries (17 different sites) and representing 14 pine species (two species had two varieties) and one hybrid. Again, owing to the greater variability, 15 factors were used for calibration.

The standard error of calibration (*SEC*) (determined from the residuals from the final calibration), standard error of cross-validation (*SECV*) (determined from the residuals of each cross-validation phase) and the coefficient of determination (R^2) were used to assess calibration performance. The ratio of performance to deviation (RPD_c),[16] calculated as the ratio of the standard deviation of the reference data to the *SECV*, was also used to assess calibration performance. Determination of the RPD_c allowed comparison of calibrations developed for different wood properties that have different ranges in values.

The Stage 2 calibrations were used to predict the wood properties of the alternate sample set, i.e. set 1 calibrations were used to predict the wood properties of set 2 and vice versa. The standard error of prediction (*SEP*) was used to give a measure of how well a calibration predicted the parameter of interest for a set of samples that are different from the calibration set. The predictive ability of calibrations was assessed from the R_p^2 and the ratio of performance to deviation (RPD_p) (ratio of the standard deviation of the reference data to the *SEP*).

Spectral pretreatment

Stage 1 wood property calibrations were obtained using untreated and multiplicative scatter (MSC) treated spectra. MSC was used to correct for differences in baseline offsets and path length due to differences in particle size,[17] here a function of differences in surface roughness of the radial strips.[18]

For stages 2 and 3 of the study, second-derivative spectra were utilised in addition to the untreated and MSC-treated spectra. NIR data was converted to the second derivative using the Savitzky–Golay approach, with left and right gaps of 8 nm and second-order polynomial.

Results

Stage 1: Individual co-operator wood property calibrations

Wood property calibrations obtained for individual co-operators are summarised in Table 3. Generally, untreated spectra provided the strongest calibrations for density, while MSC-treated spectra provided the strongest statistics for MFA and stiffness and these calibrations are reported.

The calibrations for density were variable with R^2 ranging from 0.51 (SAF-3) to 0.89 (COL1-3) and standard errors of calibration ranging from 27 kg m^{-3} to 43 kg m^{-3}. An important factor in determining the strength of individual density calibrations was the range of the density data. The SAF-3 density calibration had the narrowest range (220 kg m^{-3}), while the COL1-3 calibration had the widest (568 kg m^{-3}). In general, calibrations having a wide range of densities had the strongest statistics, the exception was the SAF-1 calibration that had a range of 436 kg m^{-3} (the third widest) but an R^2 of 0.77 (the sixth best). Five factors were used for this calibration which was less than the higher ranked calibrations. If eight factors were used, the R^2 only improved to 0.80. It was expected that RPD_c would be consistent across co-operators as it is a relative measure not influenced by range; however, this was not observed and varied from 1.4 (CHL-2) to 2.3 (ARG-1).

The MFA calibrations were more consistent than those reported for density, with R^2 ranging from 0.72 (SAF-3) to 0.87 (ARG-1) and standard errors ranging from 2.6° to 3.2°. Values of RDP_c were frequently higher and more consistent than those observed for density; however, the range in RPD_c was still considerable (1.8–2.6). Unlike density, the range in MFA was generally consistent, with MFA for eight of the ten co-operators having ranges of 31.2–38.0°. The exceptions were BRZ-2 (26.2°) and SAF-3 (44.8°). Despite the wide MFA range for SAF-3, it provided the weakest calibration statistics.

Calibrations for stiffness had the strongest statistics of the three properties with R^2 ranging from 0.81 (SAF-3) to 0.920 (ARG-1) and RPD_c ranging from 2.0 (SAF-3) to 3.2 (CHL-1 and SAF-1). Standard errors were variable ranging from 0.6 GPa (BRZ-2) to 1.7 GPa (BRZ-1). The range in stiffness data

Table 3. Summary of wood property calibrations for each industrial co-operator. The number of radial strips analysed and number of NIR spectra collected are also shown.

Company	Property	# factors	R^2	*SEC*	*SECV*	*RPD*	Outliers
AL	Air-dry density (kg m^{-3})	8	0.80	38	40	2.1	0
Strips = 41	MFA (°)	9	0.80	2.7	2.9	2.1	1
Spectra = 296	SilviScan stiffness (GPa)	7	0.90	1.0	1.0	3.0	0
AR	Air-dry density (kg m^{-3})	7	0.68	27	28	1.7	0
Strips = 41	MFA (°)	10	0.86	2.6	2.8	2.5	2
Spectra = 326	SilviScan stiffness (GPa)	9	0.919	0.8	0.9	3.2	0
BP	Air-dry density (kg m^{-3})	9	0.82	30	32	2.2	0
Strips = 26	MFA (°)	9	0.87	2.1	2.3	2.6	0
Spectra = 211	SilviScan stiffness (GPa)	9	0.920	0.7	0.8	3.2	0
KB	Air-dry density (kg m^{-3})	5	0.82	43	44	2.3	0
Strips = 25	MFA (°)	6	0.78	3.2	3.3	2.1	0
Spectra = 298	SilviScan stiffness (GPa)	6	0.82	1.7	1.8	2.3	0
KLF	Air-dry density (kg m^{-3})	5	0.77	43	43	2.1	0
Strips = 72	MFA (°)	7	0.83	3.0	3.1	2.4	2
Spectra = 675	SilviScan stiffness (GPa)	9	0.909	1.3	1.4	3.2	0
MC	Air-dry density (kg m^{-3})	8	0.56	31	33	1.4	0
Strips = 31	MFA (°)	8	0.77	2.8	3.1	2.0	2
Spectra = 234	SilviScan stiffness (GPa)	10	0.89	0.7	0.7	2.8	0
MD	Air-dry density (kg m^{-3})	8	0.68	29	32	1.6	0
Strips = 25	MFA (°)	9	0.82	3.2	3.6	2.1	2
Spectra = 195	SilviScan stiffness (GPa)	8	0.915	0.7	0.9	2.9	0
RG	Air-dry density (kg m^{-3})	10	0.84	29	33	2.2	0
Strips = 23	MFA (°)	8	0.79	2.1	2.3	2.0	3
Spectra = 200	SilviScan stiffness (GPa)	10	0.89	0.6	0.7	2.6	0
SC	Air-dry density (kg m^{-3})	7	0.89	35	36	3.0	0
Strips = 62	MFA (°)	8	0.82	2.8	2.9	2.3	1
Spectra = 605	SilviScan stiffness (GPa)	5	0.903	1.6	1.6	3.2	0
SP	Air-dry density (kg m^{-3})	2	0.51	37	37	1.4	3
Strips = 29	MFA (°)	9	0.72	3.0	3.3	1.8	2
Spectra = 192	SilviScan stiffness (GPa)	9	0.81	1.0	1.1	2.0	0

across sites was also variable; COL1-3 had the widest range (25.9 GPa) and BRZ-2 the narrowest (9.0 GPa). Generally, the data with the widest stiffness range gave the strongest calibration statistics.

Stage 2: Multi co-operator wood property calibrations

Following the development of individual co-operator calibrations, multi co-operator calibrations that included samples from sites from each country were explored. As described in the methods section, the samples were split into two sets (sets 1 and 2) that were used to develop wood property calibrations that were then tested on the alternate set. Table 4 reports the calibrations that gave the best performance in prediction and the corresponding math treatment.

For the three properties, calibration statistics compared well with those reported in Table 3, with stiffness having the best fit statistics. Calibration statistics were similar regardless of math treatment, while minor differences were observed among calibrations regarding the number of factors recommended (generally calibrations based on second derivative spectra used the fewest factors). Set 1 wood property calibrations had the highest RPD_c for density and MFA [Figure 1(a) and (b)], while set 2 gave the highest for stiffness [Figure 1(c)]. It is apparent from Figure 1(b) and (c) that the calibrations for MFA and, to a lesser degree, stiffness show nonlinearity in the data. For the individual co-operator calibrations, clear nonlinearity was observed for some sets, in particular CHL-2 and SAF-2 for both properties and BRZ-1 for stiffness.

Table 4. Summary of wood property calibrations for multi co-operator data sets (set 1 and set 2) and their predictive performance when used to predict the wood properties of the alternate set.

	Set 1 (1752 spectra)						Tested on Set 2		
Property	# factors	R^2	*SEC*	*SECV*	RPD_c	Outliers	R_p^2	*SEP*	RPD_p
Air-dry density (kg m^{-3}) second derivative	8	0.82	38	39	2.3	0	0.71	46	1.8
MFA (°) raw	12	0.85	2.9	2.9	2.5	2	0.76	3.5	2.0
SS stiffness (GPa) MSC	13	0.906	1.3	1.3	3.2	0	0.87	1.4	2.7
	Set 2 (1480 spectra)						Tested on Set 1		
Air-dry density (kg m^{-3}) second derivative	7	0.77	39	41	2.0	0	0.71	48	1.9
MFA (°) second derivative	10	0.80	3.1	3.3	2.1	3	0.81	3.3	2.2
SS stiffness (GPa) second derivative	10	0.917	1.1	1.2	3.3	0	0.85	1.7	2.6

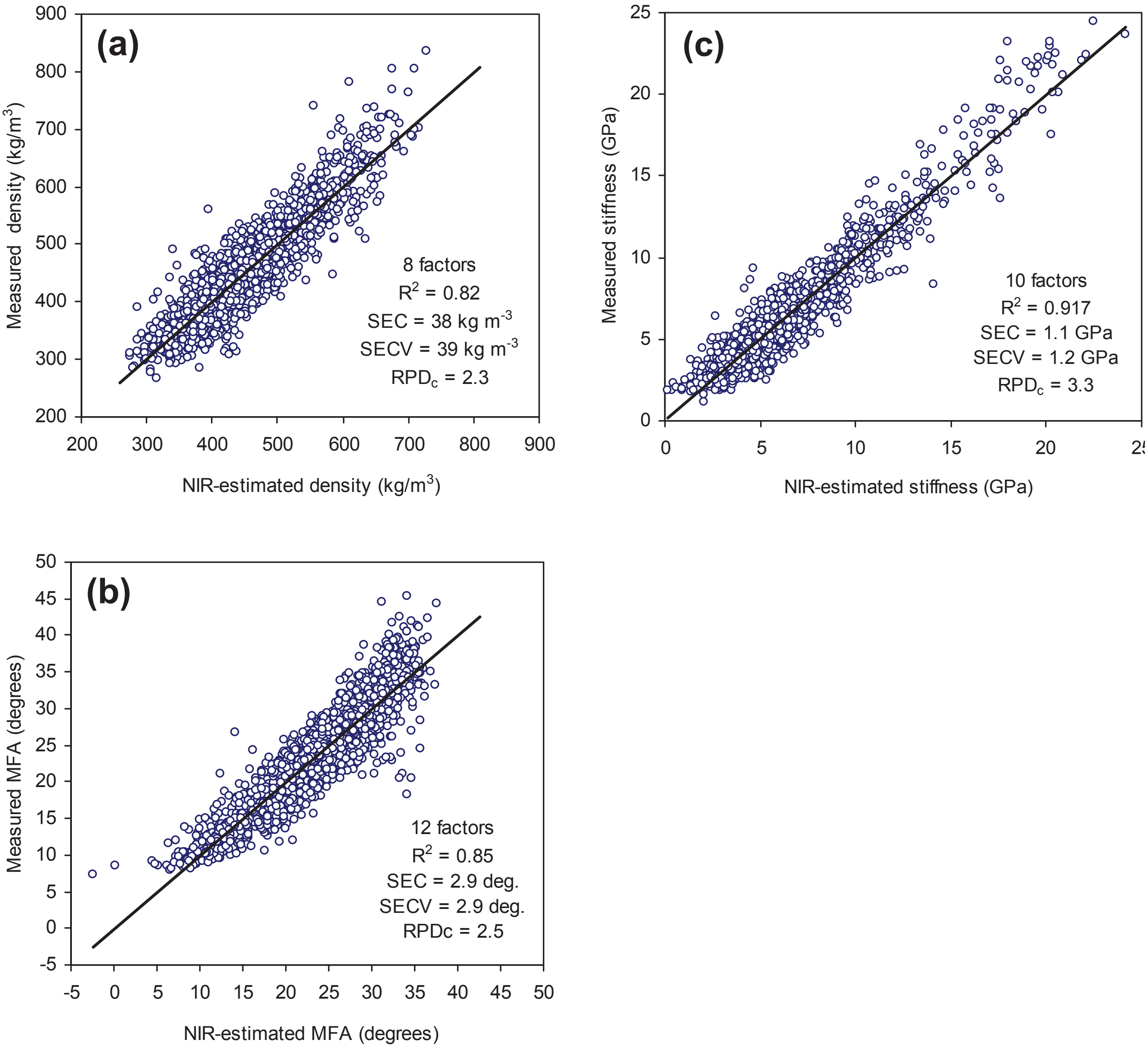

Figure 1. Relationships between SilviScan-measured and NIR-estimated wood properties for (a) density (set 1), (b) MFA (set 1) and (c) stiffness (set 2).

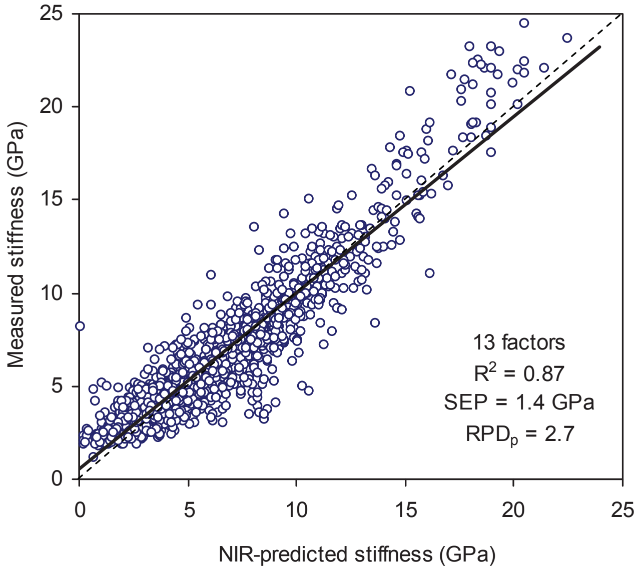

Figure 2. Relationship between SilviScan-measured stiffness and NIR-predicted stiffness. The broken line represents the line of equivalence.

When used to predict the wood properties of the alternate set, calibrations performed reasonably well (Table 4). However, a marked reduction in prediction statistics was observed, particularly for density. The best performing calibration was the set 1 stiffness calibration based on MSC-treated spectra that predicted the stiffness of the set 2 samples with an R_p^2 of 0.87 and a RPD_p of 2.7 (Figure 2). For set 1, raw spectral data consistently gave the best performance, while for set 2 the second-derivative math treatment provided the strongest prediction statistics. The predictive performance of the density calibrations based on raw spectra was similar to those based on second-derivative spectra for both sets, while those based on MSC-treated spectra were clearly inferior. The set 2 calibrations when applied to the set 1 samples predicted the sections of two samples poorly, with extremely large residuals. Both samples were excluded when the wood properties of the set 1 samples were predicted by the set 2 calibrations. It is unknown why the wood properties of these samples were poorly predicted. Both samples were examined for the presence of knots and other defects but appeared normal. Both samples had also been analysed by SilviScan, and examination of the wood property data did not reveal any anomalies; however, one of the excluded samples did have low MFA (maximum 26.5°).

Stage 3: Wood property calibrations based on all samples

Sets 1 and 2 were combined and used to develop calibrations that included samples provided by all co-operators (14 pine species, one hybrid and two species each having two varieties). Untreated, and MSC-treated and second-derivative-treated spectra were used to develop the calibrations (Table 5).

Calibration statistics were generally similar to those observed in stage 2; however, the number of factors recommended for the various properties varied by math treatment making comparison difficult. Calibrations based on second-derivative spectra required the fewest factors, while those based on untreated spectra required the most. Of the three properties the stiffness calibrations, regardless of math treatment, had the strongest statistics, followed by MFA and density. The nonlinearity associated with both MFA and stiffness was again apparent. The relatively strong statistics demonstrate that wood property calibrations can be built using multiple pine species from different countries.

Discussion

Calibrations reported in this study were developed in three stages: the first stage involved developing multi-species

Table 5. Summary of global wood property calibrations.

Property	# Factors	R^2	*SEC*	*SECV*	RPD_c	Outliers
Untreated spectra						
Air-dry density ($kg\,m^{-3}$)	8	0.77	42	42	2.1	0
MFA (°)	12	0.82	3.2	3.2	2.3	5
SS stiffness (GPa)	11	0.901	1.3	1.3	3.1	0
MSC treated spectra						
Air-dry density ($kg\,m^{-3}$)	9	0.71	46	47	1.9	0
MFA (°)	12	0.82	3.1	3.1	2.3	5
SS stiffness (GPa)	9	0.86	1.5	1.5	2.7	0
Second-derivative treated spectra						
Air-dry density ($kg\,m^{-3}$)	8	0.79	40	41	2.1	0
MFA (°)	10	0.82	3.2	3.2	2.3	5
SS stiffness (GPa)	8	0.88	1.4	1.4	2.9	0

calibrations for individual co-operators that may have included samples from more than one site; the second stage involved combining a mix of species and sites from each country; while the final stage involved building wood property calibrations based on samples from several different countries. As the size of the calibration set increased, calibration statistics generally decreased. However, the good performance of the calibrations when used to predict the wood properties of a separate test set (stage 2) demonstrates that it is possible to build multi-site, multi-species calibrations for pines.

The stage 1 calibrations had variable statistics and for many sites were comparable to those reported in the literature (for example, Jones *et al.*[11]), particularly for stiffness. However, statistics for all SAF-3 calibrations were consistently poor having the lowest (or close to the lowest) RPD_c values for each property. The importance of variability (or range) has been discussed in obtaining strong calibration statistics[9] and it is clear that range was important in obtaining good statistics for many of the calibrations. For the SAF-3 samples, density had a range of only 220 kg m^{-3} which may be close to the lower limit (in terms of range) that calibrations for density can be obtained over. The age of the stands is also important in determining variability. Young stands have yet to produce mature wood and will not have the high density wood of older stands. It should be noted that SAF-3 trees were only 10 years old but younger stands such as ARG-1 and ARG-2 (eight years) provided stronger statistics.

An important aspect in trying to develop broad-based calibrations using a large number of samples is sample quality; ideally, samples as uniform as possible in terms of collection, storage, preparation and testing would be used. When analysing a large number of cores provided by a range of operators who have varying levels of knowledge and skill in sampling trees and preparing samples for analysis, problems related to sample quality can be expected. One of the problems we encountered in this study was that some of the samples were blue-stained owing to poor storage practices. An effort was made to exclude strips with excessive blue stain but samples that had a small amount of blue stain were included, and it became difficult to determine what should or should not be included. Another issue was strips that had excessively tilted tracheids, either due to the corer being at an angle during sampling or the core rolling in the sample holder when glued for sawing. Ideally, tracheid orientation would be consistent with minimal pitch (equivalent to spiral grain) and roll (tracheids tilted toward or away from the pith) but this was not possible in this study. Strips that had tracheids with excessive pitch were not tested but it was difficult to determine severity unless the end of the sample could be examined. As with blue-stain avoidance, determining which samples to exclude owing to excessive pitch was difficult. Such issues can influence calibration quality and reinforces the importance of adherence to sampling protocols and consistency in all aspects of sample collection, storage, preparation and testing.

The stage 2 calibrations were based on a greatly expanded sample set including pine species from five different countries (Argentina, Brazil, Chile, Colombia and South Africa). Despite the wide geographical variation, statistics were comparable to those reported for the individual co-operators and the calibrations performed reasonably well when applied to a separate test set, i.e. set 1 calibrations were used to predict the wood properties of set 2 and vice versa. Considering that the calibrations were being applied to samples from different sites and, in many cases, different pine species, their performance was quite good. For example, for the sites sampled in Argentina, one site had five pine species, while the other site had seven (one species at this site had two varieties) and the two sites only had four species in common.

An important consideration for MFA and stiffness calibrations developed using large data sets is the presence of nonlinearity in the data. For MFA and stiffness calibrations based on data sets of limited variation (for example, Schimleck and Evans[19,20]), this has not been an issue but when samples from multiple sites have been combined, nonlinearity becomes apparent.[21,22] Mora *et al.*[22] observed nonlinearity in *P. taeda* MFA and stiffness calibrations based on samples from sites in Georgia, North Carolina and South Carolina and utilised two kernel regression methods (radial basis functions partial least squares and least squares support vector machines) to accommodate the nonlinear behaviour of the data. Calibrations based on the two kernel regression methods provided stronger calibration statistics than partial least squares (PLS) regression and outperformed PLS regression when tested on a separate set, so it was concluded that both methods gave substantial improvements over PLS.[22] It is possible that the nonlinearity observed for MFA in this study and in others[21,22] is an artefact of its determination by SilviScan, which uses a variance approach.[14] SilviScan cannot give MFA measurements of lower than approximately 7° and the tendency for MFA measurements to cut off at around this value is apparent in Figure 1(b). Removal of these low MFA samples (and three high MFA samples) greatly reduces the appearance of nonlinearity.

It is interesting to consider how many wood properties appear to display nonlinearity when PLS calibrations are developed. In this study and those by Cogdill *et al.*[21] and Mora *et al.*,[22] MFA displayed the greatest nonlinearity followed by stiffness, while nonlinearity was not observed for density. SilviScan was used to determine the stiffness data utilised in this study and used a relationship that combines both X-ray densitometry and X-ray diffraction data[15] and the tendency for the stiffness relationship to be midway between that of density and MFA is consistent with this. Schimleck *et al.*[12] developed calibrations for stiffness using *P. taeda* samples from 81 plantations across the south-eastern US and did not observe nonlinearity in the data. The stiffness data utilised by Schimleck *et al.*[12] were determined using static bending. Other studies have developed broad based calibrations for lignin[7] and pulp yield[9] and neither reported issues related to nonlinearities.

The wood property calibrations reported here are unique in forestry-related research. Efforts to develop multi-site, multiple-species wood property calibrations are generally rare and those that do exist, while having considerable variation in terms of species represented,[8,9] are limited to a single country. Hodge and Woodbridge[7] is an exception, utilising five tropical and sub-tropical pine species grown in Brazil and Colombia. As noted by Schimleck,[4] several barriers exist in creating broad-based wood property calibrations. One of the most important is the difficulty of companies and research organisations (both within a country and among countries) sharing wood samples that can represent considerable investment, both in terms of intellectual property and laboratory analysis. By working closely with a number of industry co-operators in several countries, Camcore is one of the few programmes that can overcome this barrier. Another question relates to the applicability of such calibrations. Generally, the activities of an organisation are tied to a certain country and/or region and there is little need to assess properties on a broad scale. Camcore is an organisation that works in many countries and can benefit from broad-based wood property calibrations for the preliminary assessment of potential plantation species.

Conclusions

Density, MFA and stiffness calibrations were developed for a range of tropical, sub-tropical and temperate pine species. Samples were provided by ten forest industry companies and represented 14 pine species (two species had two varieties) and one hybrid, grown on 17 different sites in five countries (Argentina, Brazil, Chile, Colombia and South Africa). Wood property calibrations were developed in three stages. In the first stage, calibrations were obtained using samples provided by individual co-operators. Calibration statistics varied by co-operator, for example the R^2 for density ranged from 0.51 to 0.89. Possible explanations for the variable statistics include differences in wood property variation, i.e. some sites demonstrated less variation in wood properties than others, owing to differences in age and species composition and variation in sample collection and preparation. In the second stage, multi co-operator calibrations that included samples from sites in each country were obtained. Two data sets were created: set 1 and set 2, and both sets provided calibration statistics for the three wood properties that compared well to those obtained in stage 1 with stiffness having the best fit statistics ($R^2 = 0.917$, $RPD_c = 3.3$). MFA and stiffness, to a lesser degree, showed the presence of nonlinearity in the data. The calibrations performed reasonably well when used to predict the wood properties of the alternate set (set 1 calibrations were used to predict the wood properties of set 2 samples and vice versa) with density having the weakest predictions. Stage 3 calibrations were based on all available samples and were similar to those developed in stage 2, demonstrating that it is possible to build multi-site, multi-species calibrations for pines.

Acknowledgements

The authors would like to thank all of the Camcore member organisations for their financial support of this research, and Coby Smith for collecting the NIR spectra utilised in this study.

References

1. I. Murray, "Scattered information: philosophy and practice of near infrared spectroscopy", in *Near Infrared Spectroscopy: Proceedings of the 11th International Conference*, Ed by A.M.C. Davies and A. Garrido-Varo. NIR Publications, Chichester, UK, p. 1 (2004).
2. P. Berzaghi, P.C. Flinn, P. Dardenne, M. Lagerholm, J.S. Shenk, M.O. Westerhaus and I.A. Cowe, "Comparison of linear and non-linear near infrared calibration methods using large forage databases", in *Near Infrared Spectroscopy: Proceedings of the 10th International Conference*, Ed by A.M.C. Davies and R.K. Cho. NIR Publications, Chichester, UK, p. 107 (2002).
3. P. Dardenne, "Calibration transfer in near infrared spectroscopy", in *Near Infrared Spectroscopy: Proceedings of the 11th International Conference*, Ed by A.M.C. Davies and A. Garrido-Varo. NIR Publications, Chichester, UK, p. 19 (2004).
4. L.R. Schimleck, "Near infrared spectroscopy: A rapid, non-destructive method for measuring wood properties and its application to tree breeding", *N.Z. J. For. Sci.* **38,** 14 (2008).
5. D.C.F. Garbutt, M.J. Donkin and J.H. Meyer, "Near-infrared reflectance analysis of cellulose and lignin in wood", *Pap. S. Afr.* **April,** 45 (1992).
6. A.J. Michell, "Pulpwood quality estimation by near-infrared spectroscopic measurements on eucalypt woods", *Appita J.* **48,** 425 (1995).
7. G.R. Hodge and W.C. Woodbridge, "Use of near infrared spectroscopy to predict lignin content in tropical and sub-tropical pines", *J. Near Infrared Spectrosc.* **12,** 381 (2004). doi: 10.1255/jnirs.447
8. L.R. Schimleck, G.D.S.P. Rezende, B.J. Demuner and G.M. Downes, "Estimation of whole-tree wood quality traits using near infrared spectra collected from increment cores", *Appita J.* **59,** 231 (2006).
9. G.M. Downes, R. Meder, C. Hicks and N. Ebdon, "Developing and evaluating a multisite and multispecies NIR calibration for the prediction of Kraft pulp yield in eucalypts", *Southern Forests* **71,** 155 (2009).
10. G.M. Downes, R. Meder and C. Harwood, "A multi-site, multi-species calibration for the prediction of cellulose content in eucalypt woodmeal", *J. Near Infrared Spectrosc.* **18,** 381–387 (2010). doi: 10.1255/jnirs.910
11. P.D. Jones, L.R. Schimleck, G.F. Peter, R.F. Daniels and A. Clark, "Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites

in Georgia", *Can. J. For. Res.* **35,** 85 (2005). doi: 10.1139/x04-160

12. L.R. Schimleck, P.D. Jones, A. Clark, R.F. Daniels and G.F. Peter, "Near infrared spectroscopy for the nondestructive estimation of clear wood properties of *Pinus taeda* L. from the Southern USA", *For. Prod. J.* **55,** 21 (2005).

13. R. Evans, "Rapid scanning of microfibril angle in increment cores by x-ray diffractometry", in *Microfibril Angle in Wood*: *Proceedings of the IAWA/IUFRO International Workshop on the Significance of Microfibril Angle to Wood Quality*, November 1997, Westport, New Zealand, Ed by B.G. Butterfield. University of Canterbury Press, New Zealand, pp. 116–139 (1998).

14. R. Evans, "A variance approach to the x-ray diffractometric estimation of microfibril angle in wood", *Appita J.* **52,** 283 (1999).

15. R. Evans, "Wood stiffness by X-ray diffractometry", in *Characterization of the Cellulosic Cell Wall*, Ed by D.D. Stokke and L.H. Groom. Blackwell Publishing, Ames, Iowa, p. 138 (2006).

16. P.C. Williams and D.C. Sobering, "Comparison of commercial near infrared transmittance and reflectance instruments for the analysis of whole grains and seeds", *J. Near Infrared Spectrosc.* **1,** 25 (1993). doi: 10.1255/jnirs.3

17. P. Geladi, D. McDougall and H. Martens, "Linearization and scatter-correction for near-infrared reflectance spectra of meat", *Appl. Spectrosc.* **39,** 491 (1985). doi: 10.1366/0003702854248656

18. C. Mora and L.R. Schimleck, "Determination of specific gravity of green *Pinus taeda* samples by near infrared spectroscopy: Comparison of preprocessing methods using multivariate figures of merit", *Wood Sci. Technol.* **43,** 441 (2009). doi: 10.1007/s00226-008-0235-0

19. L.R. Schimleck and R. Evans, "Estimation of wood stiffness of increment cores by near infrared spectroscopy: the development and application of calibrations based on selected cores", *IAWA J.* **23,** 217 (2002).

20. L.R. Schimleck and R. Evans, "Estimation of microfibril angle of increment cores by near infrared spectroscopy", *IAWA J.* **23,** 225 (2002).

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

22. C. Mora, and L.R. Schimleck, "Kernel regression methods for the prediction of wood properties of *Pinus taeda* using near infrared (NIR) spectroscopy", *Wood Sci. Technol.* **44,** 561 (2010). doi: 10.1007/s00226-009-0299-5