Estimation of Wood Density and Chemical Composition by Means of Diffuse Reflectance Mid-Infrared Fourier Transform (DRIFT-MIR) Spectroscopy

MARI H. NUOPPONEN,* GILLIAN M. BIRCH, ROB J. SYKES, STEVE J. LEE, AND DEREK STEWART

Scottish Crop Research Institute, Dundee DD2 5DA, Scotland, U.K., and Forest Research, Northern Research Station, Roslin EH24 9SY, Scotland, U.K.

Sitka spruce (*Picea sitchensis*) samples (491) from 50 different clones as well as 24 different tropical hardwoods and 20 Scots pine (*Pinus sylvestris*) samples were used to construct diffuse reflectance mid-infrared Fourier transform (DRIFT-MIR) based partial least squares (PLS) calibrations on lignin, cellulose, and wood resin contents and densities. Calibrations for density, lignin, and cellulose were established for all wood species combined into one data set as well as for the separate Sitka spruce data set. Relationships between wood resin and MIR data were constructed for the Sitka spruce data set as well as the combined Scots pine and Sitka spruce data sets. Calibrations containing only five wavenumbers instead of spectral ranges 4000−2800 and 1800−700 cm^{-1} were also established. In addition, chemical factors contributing to wood density were studied. Chemical composition and density assessed from DRIFT-MIR calibrations had R^2 and Q^2 values in the ranges of 0.6−0.9 and 0.6−0.8, respectively. The PLS models gave residual mean squares error of prediction (RMSEP) values of 1.6−1.9, 2.8−3.7, and 0.4 for lignin, cellulose, and wood resin contents, respectively. Density test sets had RMSEP values ranging from 50 to 56. Reduced amount of wavenumbers can be utilized to predict the chemical composition and density of a wood, which should allow measurements of these properties using a hand-held device. MIR spectral data indicated that low-density samples had somewhat higher lignin contents than high-density samples. Correspondingly, high-density samples contained slightly more polysaccharides than low-density samples. This observation was consistent with the wet chemical data.

KEYWORDS: DRIFT-MIR; wood density; lignin; α-cellulose; extractives; Sitka spruce; tropical hardwoods; PLS

INTRODUCTION

The structure of wood has been the subject of many studies based on vibrational spectroscopic methods such as near-infrared (NIR) (*1−5*), mid-infrared (MIR) (*6*, *7*), and Raman spectroscopies (*8*). All of these methods provide information on the molecular structure of wood and, in many cases, solid or ground wood samples can be analyzed directly without laborious preparation. Hence, it is possible to get information on the molecular level interactions between wood polymers in their native state. Moreover, these methods are both rapid and nondestructive.

Most of the published studies addressing the assessment of solid wood properties by spectroscopic methods have been conducted using NIR spectroscopy. The chemical compositions/components (*3*, *9*, *10*), dry matter content (*11*), and strength (*2*, *12*, *13*), and density (*1*, *2*, *11*, *12*) of solid wood samples have

been estimated from the NIR spectral data combined with multivariate data analysis methods. Furthermore, morphological characteristics were assessed from dry (*4*, *12*) and green (*14*) wood samples.

There are a few publications also on the application of the MIR spectroscopy to determine wood density (*6*), chemical constituents (*6*, *15−17*), and chemical structures (*18*). Raman spectroscopy has been applied to quantify chemical constituents (*19−21*) and density (*22*) of the solid wood samples. Furthermore, calibration between morphological characteristics of the solid wood samples and Raman spectral data has been established (*8*).

These approaches have been shown to be powerful tools to assess, often complex, physical and chemical properties of wood at the laboratory scale. Extension of this to the rapid assessment of the physical and chemical properties, such as density and lignin content, would be of great benefit in tree improvement programs. Additionally, the implementation of these techniques

* Author to whom correspondence should be addressed (e-mail mnuopponen@herc.com; fax +46 42 371151; telephone +46 42 371117).

Wood Density and Chemical Composition

J. Agric. Food Chem., Vol. 54, No. 1, 2006   **35**

Table 1. Tropical Hardwood Samples

common name	scientific name	family
aprokuma	*Antrocaryon micraster*	Anacardiaceae
idigbo	*Terminalia ivorensis*	Combretaceae
afara	*Terminalia superba*	Combretaceae
kokrodua	*Pericopsis elata*	Leguminosae
cobaiba	*Copaifera* spp.	Leguminosae
dahoma	*Piptadeniastrum africanum*	Leguminosae
ogea	*Daniellia ogea*	Leguminosae
awiemfo-samina	*Albizzia ferruginea*	Leguminosae
African mahogany	*Khaya ivorensis*	Meliaceae
avodire	*Turreanthus africanus*	Meliaceae
sapele	*Entandrophragma cylindricum*	Meliaceae
scented guarea	*Guarea cedrata*	Meliaceae
kosipo	*Entandrophragma candollei*	Meliaceae
cedro	*Cedrela Mexicana*	Meliaceae
African walnut	*Lovoa trichilioides*	Meliaceae
iroko	*Chlorophora excelsa*	Moraceae
ekki	*Lophira alata*	Ochnaceae
abura	*Mitragyna ciliata*	Rubiaceae
opepe (kusia)	*Nauclea diderichii*	Rubiaceae
makore (baku)	*Tieghemella heckelii*	Sapotaceae
mansonia	*Mansonia altissima*	Sterculiaceae
nyankom	*Tarrietia utilis*	Sterculiaceae
obeche (wawa)	*Triplochiton scleroxylon*	Sterculiaceae
danta	*Nesogordonia papaverifera*	Sterculiaceae

to postharvest (felling) screening can also be taken advantage of in sawmills and pulp mills.

Here the densities and lignin and cellulose contents of Sitka spruce, Scots pine, and tropical hardwood samples were estimated by means of diffuse reflectance mid-infrared Fourier tranform (DRIFT-MIR) spectroscopy linked with partial least-squares (PLS) analysis. Furthermore, the effects of the reduced spectral range on the predictive abilities of the models were investigated. MIR spectroscopy has been mostly used for qualitative characterization of the solid wood samples. There are only a limited number of MIR-based studies dealing with quantification of the solid wood properties (*6, 15−17*), and to our knowledge MIR-based models including a wide scale of natural variation in density and lignin and cellulose contents have not been published.

MATERIALS AND METHODS

Samples. Sitka spruce (*Picea sitchensis*) cored samples (493; 8 mm in diameter) from 50 different 15-year-old clones exhibiting significant phenotypic variation was cored in Newcastleton (Scotland). Scots pine (*Pinus sylvestris*) wood cored samples (20) were taken at the Forest Research site in Monaughty (Scotland). Coring was performed at breast height (130 cm) of the trees. In addition, 24 blocks of tropical hardwood samples from Ghana were obtained from Forest Research, Scotland. The hardwood samples are listed in **Table 1**. Prior to the chemical analysis, samples were freeze-dried and ground in a Glen Creston laboratory mill to pass a 1000 μm sieve.

Density. Densities of the core samples were calculated by measuring fresh wood dimensions and weights of the freeze-dried samples. Densities of the Sitka spruce and Scots pine samples were in the range of 232−494 and 329−459 kg m^{-3}, respectively. Hardwood samples had densities ranging from 323 to 878 kg m^{-3}. Densities of the hardwood samples were determined by measuring dry wood dimensions and weights.

Wet Chemical Analyses. Wood resin was removed from the ground wood samples (0.5 g) by Soxhlet extraction for 3 h with ethyl acetate. Extracts were evaporated to dryness by rotary evaporation and then weighed. Selected samples were also extracted with hexane and acetone to compare the amounts of extracts removed with all three solvents. The yield of acetone extracts was higher [average difference of 0.35% (w/w)] than that of ethyl acetate soluble substances, whereas hexane removed less extractable substances than ethyl acetate [average difference of 0.22% (w/w)].
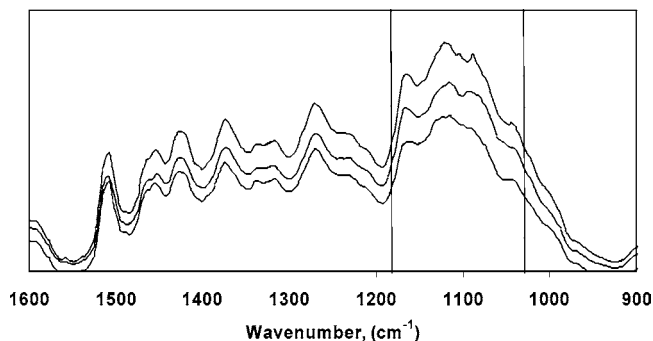

Figure 1. Three parallel DRIFT-MIR spectra of a ground (<1000 μm) Sitka spruce sample. Differences due to particle size variation are observed in the region of 1180–1000 cm^{-1}.

α-Cellulose was isolated from duplicate samples (50 mg of the extractive-free sample) according to the method of Brendel et al. (*23*). Lignin contents of the extractive-free samples (5 mg) were determined in triplicate using an acetyl bromide method (*24*). In this method lignocellulosic samples are dissolved in the mixture of acetyl bromide and glacial acetic acid (1:3, v/v), and the absorbance at 280 nm is measured. Lignin contents of selected samples were also determined as insoluble Klason lignin that gave ∼4.8% (w/w) higher lignin content than the acetyl bromide method.

DRIFT-MIR Spectroscopy. DRIFT-MIR spectra of the ground wood samples were collected by a Bruker IFS 66 spectrometer using a Specac DRIFT accessory over the range from 4000 to 700 cm^{-1} and with a resolution of 4 cm^{-1}. A DTGS detector was used. Four hundred scans were accumulated prior to the Fourier transformation. IR spectra were expressed in the values of the Kubelka−Munk (KM) function. Parallel spectra from each sample were recorded, and an average spectrum was utilized in calibrations. Due to the large amount of samples, IR spectra were recorded from undiluted wood meals (particle size < 1000 μm). Two spectra from each sample were recorded, and an average spectrum was utilized. The particle size of the wood powders and roughness of the solid wood surfaces affect the IR band intensities (*25*). Three parallel DRIFT-MIR spectra of a Sitka spruce sample are presented in **Figure 1**. Spectra exhibit small differences in the region of 1180−1000 cm^{-1} as has earlier been detected by Faix and Böttcher (*25*) and Michell (*26*).

Multivariate Data Analysis. DRIFT-MIR spectra with density and wet chemical data were analyzed using a Simca-P 10.0 software package. The spectral ranges of 4000−2800 and 1800−700 cm^{-1} were used for PLS modeling. In addition, impacts of the reduced wavenumber areas on the reliability of the models were studied. Spectral data were mean-centered (*X* matrix), whereas densities and wet chemical data (*Y* matrix) were both mean-centered and scaled to unit variance. Prior to the calculation of the PLS components, the spectral data were also filtered using an orthogonal signal correction (OSC) procedure. OSC is a spectral filter commonly used with PLS calibration. It removes spectral information unrelated to the response parameters of interest (*27*). OSC pretreated models had better fit and predictive values than models constructed from unfiltered data. Therefore, all of the models shown in this paper have been pretreated with the OSC procedure.

All PLS models were internally and externally validated. Internal validation was based on cross-validation (CV). In this procedure some of the original spectral data were used to construct the model from which the remaining of the spectral data was estimated. For this purpose, the squares of the differences between predicted and observed values were added together to obtain the predictive residual sum of squares (PRESS), which is a measure of the predictive power of the model being tested (*27*). In this paper the PRESS is re-expressed as Q^2, the ability of a model to predict. The accuracy of the calibrations and test set predictions is presented here by root-mean-square errors of estimation (RMSEE) and prediction (RMSEP), respectively. RMSEE and RMSEP are expressed in units of the original measurement.

Sitka spruce (248), pine (20), and 24 tropical hardwood samples were used to establish relationships between wood chemical constituents and density, whereas 243 Sitka spruce samples were used for external

36 *J. Agric. Food Chem.,* Vol. 54, No. 1, 2006

Nuopponen et al.

Table 2. Summary of the Chemical and Density Data of the Sitka Spruce, Scots Pine, and Tropical Hardwood Samples Used in Calibrations and External Validation

| | Sitka spruce | | Scots pine (20) | tropical hardwoods (24) |
	calibration set (249)	test set (242)		
lignin, % (w/w)				
av	23.6	23.0	20.8	25.0
min−max	19.8−29.4	20.3−26.9	19.2−23.0	19.1−30.7
SD	1.5	1.4	1.2	2.8
cellulose, % (w/w)				
av	44.3	45.4	44.4	39.9
min−max	35.5−51.3	37.7−51.4	40.5−46.8	32.0−47.8
SD	2.7	2.6	1.9	3.9
density, kg/m³				
av	344	355	400	575
min−max	232−492	267−462	329−459	353−878
SD	43.3	42.2	37.2	115
wood resin, % (w/w)				
av	0.7	0.7	3.0	1.7
min−max	0.3−9.5	0.3−2.9	1.7−5.6	0.2−4.2
SD	0.3	0.3	1.0	1.1
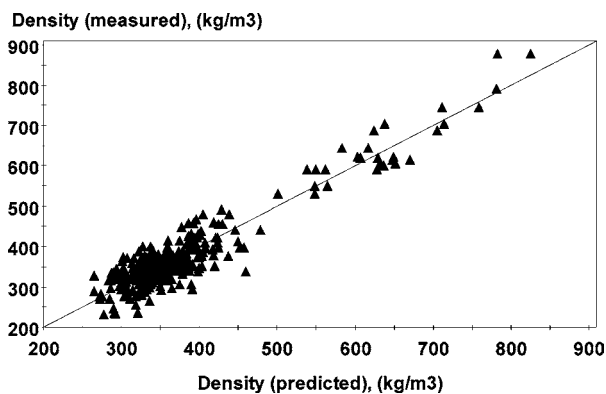
Figure 2. Calibration of Sitka spruce, Scots pine, and tropical hardwood densities and DRIFT-MIR spectral data. R^2, Q^2, and RMSEE for calibration are 0.89, 0.88, and 34.7, respectively.

verification of the models built from Sitka spruce and all of the wood samples. Details of the data sets used to construct and verify the PLS models are shown in **Table 2**.

Loading line plots of the calibrations reveal spectral differences characteristic for each PLS component. Here the loading spectra are plotted from the corresponding unfiltered data, because preprocessing of the spectral data made chemical interpretation of the loading plots (regression coefficients) more complicated.

RESULTS

Density. Density models were constructed from all wood species (Sitka spruce, Scots pine, and tropical hardwoods). In addition, separate calibrations from clonal Sitka spruce samples were built up to compare the predictive abilities of the models. Furthermore, impacts of the reduced wavenumber region on the predictive values of the PLS models were studied.

Wood samples were separated according to their densities in the direction of the first principal components. PLS calibration between density values and DRIFT-MIR data of all wood species is shown in **Figure 2**. The model had R^2 and Q^2 values of 0.89 and 0.88, respectively. RMSEE and RMSEP values for the model were 35 and 56 density units, respectively, which indicated that the model does not provide accurate density values for the samples not included in the calibration. Nevertheless, the accuracy of the IR-fitted density model is good enough to separate between very high- and low-density wood samples.
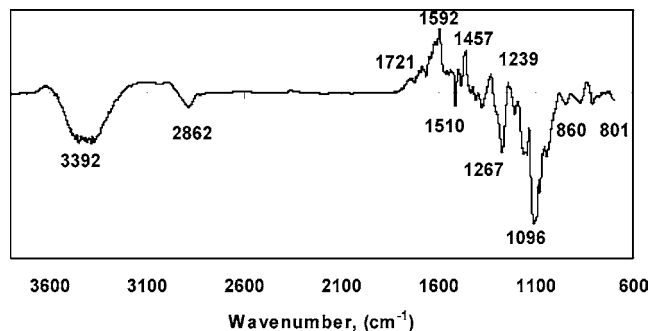
Figure 3. PLS loading plot revealing DRIFT-MIR spectral differences between high- and low-density samples (all samples; Sitka spruce, Scots pine, and tropical hardwood included). Positive and negative bands are associated with high- and low-density samples, respectively.

The spectral ranges of 4000−2800 and 1800−700 cm⁻¹ were narrowed to only a few wavenumbers (3392, 1592, 1267, and 1096 cm⁻¹) to evaluate its impact on the RMSEE, RMSEP, and r^2 values of the density models. The known most significant wavenumbers influencing the density variation were selected from the loading line plot that revealed wavenumbers contributing to the chemical differences between high- and low-density samples (**Figure 3**). Interestingly, the narrower wavenumber range did not have major impacts on the model descriptors and external validation (**Table 3**). These results indicate that it is possible to reduce the number of wavelengths remarkably without decreasing the predictive ability of the model. This reliability is a promising result and vital for the design and development of a cost-effective portable device for screening and analysis.

The loading line plot exhibits chemical structures beyond the wood density variation (**Figure 3**). Positive bands at 1720, 1592, 1457, and 1239 cm⁻¹ in **Figure 3** are related to the high-density samples, whereas those of negative bands at 3392, 2862, 1510, 1267, 1096, 860, and 801 cm⁻¹ characterize low-density samples. Hardwood samples had the highest densities, and therefore the positive bands in the loading line plot reflected structures of the chemical constituents enriched in these samples. The relatively broad MIR absorption range at 1760−1720 cm⁻¹ encompasses signals arising from acetyl groups of xylan and probably to lesser extent from wood resin. The absorption at 1592 cm⁻¹ contains contributions from phenolic compounds, polysaccharides, and lignin. Furthermore, the band at 1239 cm⁻¹ has been assigned to C−O vibrations of the acetyl groups of xylan (*28*). These band assignments indicated that xylan is a major chemical factor contributing to the higher densities of the hardwood species. Of the negative MIR bands in the loading spectrum, the bands at 1510 and 1267 cm⁻¹ have been reported to originate from guaiacyl lignin of softwoods (*29*).

The relationship between density and MIR spectral data for the clonal Sitka spruce samples is shown in **Figure 4**. Calibration had a R^2, Q^2, and RMSEE values of 0.66, 0.65, and 26 respectively, whereas the RMSEP for a test set was 55. The fit for the Sitka spruce samples was inferior to that of the calibration constructed from all wood samples. Standard error of prediction for the test set was 55, being about the same as that for the model built from both hardwood and softwood samples.

Infrared spectral differences between high- and low-density Sitka spruce samples are shown in the loading spectrum of the PLS model (**Figure 5**). Characteristic bands for the low-density samples (1580, 1478, 1278, 1171, and 1065 cm⁻¹) result from structures of lignin (*30*). Positive bands in **Figure 5** are associated with high-density samples. A broad absorption region

Wood Density and Chemical Composition

J. Agric. Food Chem., Vol. 54, No. 1, 2006   **37**

Table 3. Summary of PLS Models

|	wavenumbers/regions included in constructing the models	Sitka spruce					all wood samples				
		R^2	Q^2	RMSEE	RMSEP	no. of PCs[a]	R^2	Q^2	RMSEE	RMSEP	no. of PCs[a]
density	4000–2800 and 1800–700	0.66	0.65	25.7	55.3	1	0.89	0.88	34.7	56.1	1
	3392, 1592, 1267, 1096						0.86	0.84	36.8	53.0	1
	1739, 1510, 1375, 1267, 1129	0.62	0.61	27.2	50.0	2					
lignin	4000–2800 and 1800–700	0.74	0.73	0.8	1.8	2	0.78	0.77	1.0	1.6	3
	1600, 1510, 1273, 1220, 1077	0.70	0.69	0.9	1.9	2	0.70	0.69	1.1	1.7	2
cellulose	4000–2800 and 1800–700	0.67	0.66	1.6	3.7	1	0.65	0.65	1.8	3.3	1
	1733, 1373, 1179, 1133, 1116	0.60	0.59	1.7	3.4	2	0.62	0.61	1.9	2.8	2
wood resin	4000–2800 and 1800–700	0.94	0.93	0.2	0.4	1	0.93	0.93	0.3	0.4	1
	2930, 2856, 1698, 1456, 1154	0.77	0.73	0.4	0.5	3	0.84	0.83	0.4	0.5	2

[a] Number of principal components.
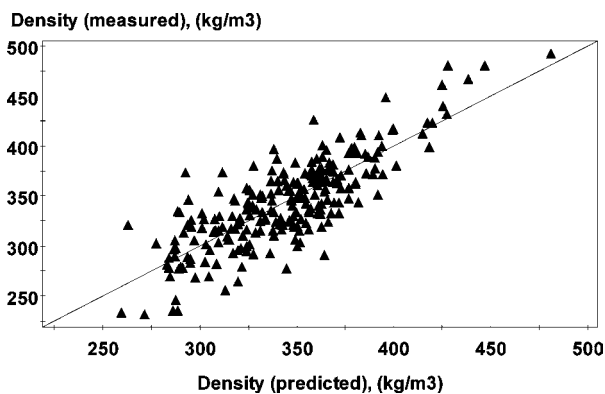
Figure 4. Calibration of Sitka spruce densities and DRIFT-MIR spectral data. R^2, Q^2, and RMSEE for calibration are 0.66, 0.65, and 25.7, respectively.
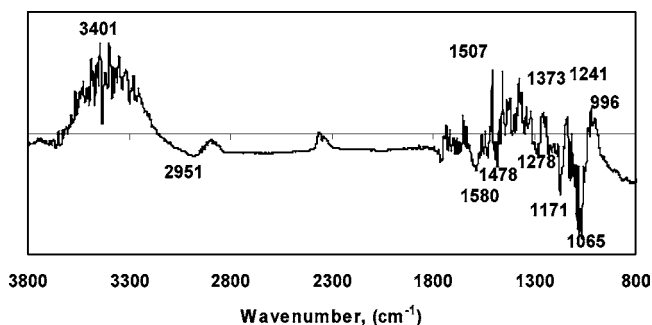

Figure 5. Loading plot revealing DRIFT-MIR spectral differences between high- and low-density Sitka spruce samples. Positive and negative bands are associated with high- and low-density samples, respectively.

showing maximum at 3400 cm^{-1} includes different O—H vibrations of polysaccharides (*31*). Furthermore, high-density samples showed a band at 1507 cm^{-1} that is associated with lignin, whereas the bands at 1456 and 1420 cm^{-1} can encompass vibrations from both from lignin and polysaccharides. MIR absorptions at 1373 and 1241 cm^{-1} have been assigned to C—H vibrations of polysaccharides and C—O stretching of acetyl groups in softwood glucomannans, respectively (*28*). Lignin band absorptions in the loading line plot indicated that high- and low-density samples can have some dissimilarity in their lignin amounts and structures. It has been shown that the condensed aromatic ring has asymmetric aromatic ring vibration at lower wavenumbers than that of the noncondensed ring. Consequently, low-density samples can possibly contain more condensed guaiacyl units than those of high-density samples. However, these results have yet to be confirmed by another method, such as solid-state NMR or wet chemical methods. MIR bands in the loading spectrum also revealed slight differences
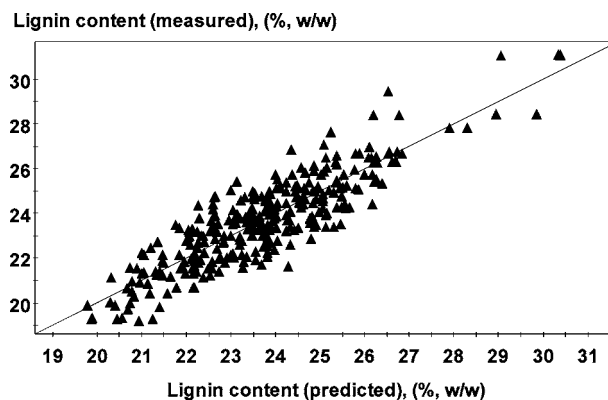

Figure 6. Calibration between lignin content and DRIFT-MIR spectral data of Scots pine, Sitka spruce, and tropical hardwood samples. R^2, Q^2, and RMSEE values for the calibration are 0.78, 0.77, and 1.0, respectively.

in the amount of lignin and polysaccharides contents of the high- and low-density samples. In agreement with earlier results, low-density samples contained somewhat higher amounts of lignin than high-density samples (*6, 32*). Correspondingly, high-density samples included more cellulose than low-density samples.

Density values for the individual wood components, cellulose, hemicelluloses, and lignin are very close to each other. The density for the pure wood cell wall substances is 1.5 g/cm^3 (*33*). Hence, compositional variation between high- and low-density samples probably arises from the cell wall thickness. It is known that the lignin content of the middle lamella is higher than that of the secondary cell wall (*34*). Therefore, the proportion of the compound middle lamella is higher for the thin-walled tracheids than for thick-walled tracheids.

Lignin Content. Lignin and cellulose calibrations were conducted on all wood species and separately on cloned Sitka spruce samples. Acetyl bromide lignin contents for the Sitka spruce samples ranged from 19.8 to 29.4% (w/w), and the average value was 23.3 (**Table 2**), whereas lignin contents for the Scots pine samples were in the range of 19.2−23.0% (w/w) (**Table 2**). The amounts of lignin for the tropical hardwoods varied from 19.1 to 30.7% (w/w) (**Table 2**).

Samples were segregated according to their lignin contents in the PLS analysis. Calibration between lignin data of all the wood samples is shown in **Figure 6**. The relationship had R^2, Q^2, and RMSEE values of 0.78, 0.77, and 1.0, respectively, whereas RMSEP for the separate test set was 1.6. Calibration of the Sitka spruce lignin contents and DRIFT-MIR data gave R^2 (0.74) and Q^2 (0.73) values comparable to those of the illustrated model (**Figure 7**). The standard error of prediction for the Sitka spruce test set data was 1.8, slightly higher than that for the model above.
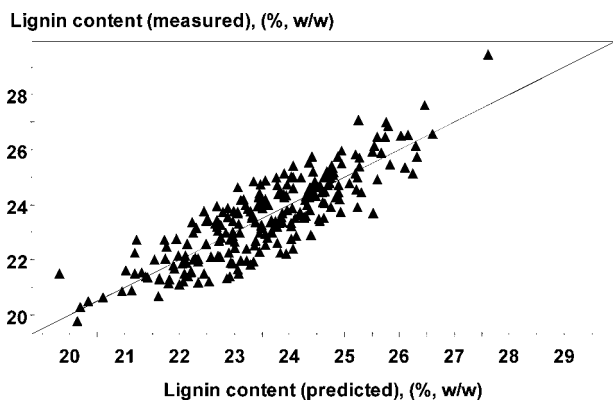
Figure 7. Calibration between lignin content and DRIFT-MIR spectral data of Sitka spruce samples. R^2, Q^2, and RMSEE values for the calibration are 0.74, 0.73, and 0.8, respectively.
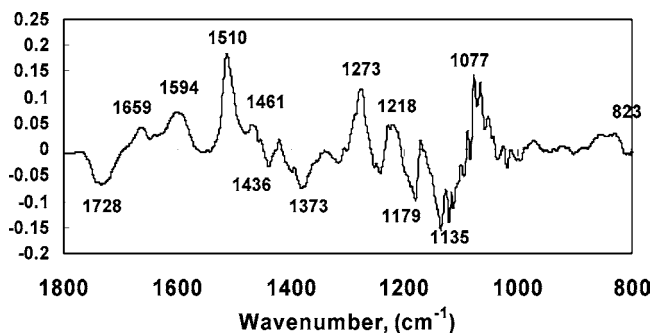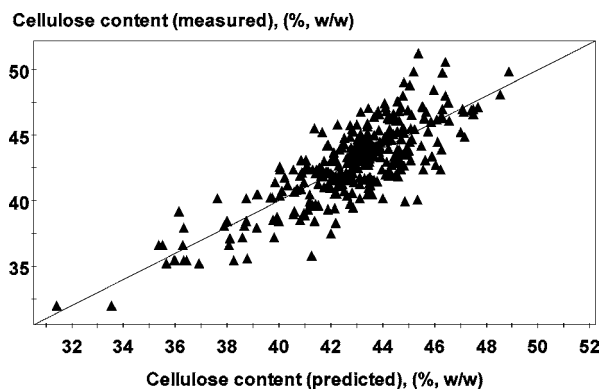

Figure 8. Loading spectrum of the first principal component of Sitka spruce lignin calibration. Positive and negative bands are typical for samples containing high and lower amounts of lignin, respectively.


Figure 9. Calibration between α-cellulose content and DRIFT-MIR spectral data of Scots pine, Sitka spruce, and tropical hardwood samples. R^2, Q^2, and RMSEE values for the calibration are 0.65, 0.65, and 1.8, respectively.

Figure 10. Loading spectrum of the first principal component of Sitka spruce cellulose calibration. Positive and negative bands characterize samples containing higher and lower levels of cellulose, respectively.

Relationships between lignin contents and MIR data were constructed from the characteristic lignin bands (1600, 1510, 1273, 1220, and 1077 cm^{-1}) shown in **Figure 8**. Reduced numbers of specific wavenumbers did not have a significant effect on the error of predictions for calibration and external data sets (**Table 3**).

Both of the models predicted lignin contents of the calibration set reasonably accurately, whereas they provided rough estimates of the lignin amounts for samples not included in the calibration data set. Yeh et al. (*3*) reported similar conclusions in their study on the NIR reflectance spectroscopic determination of the lignin content. However, the prediction accuracy of this magnitude can be useful in applications where a "few percent" precision is acceptable. Of the potential applications, tree-breeding programs can benefit from the fast approximation of the lignin content.

α-Cellulose Content. Sitka spruce clonal samples had a large variation in the α-cellulose content [35.5−51.4% (w/w)], whereas those for the Scots pine and tropical hardwood samples were in the range of 40.5−46.8 and 32.0−47.8% (w/w), respectively (**Table 2**). Samples are distinguished according to their cellulose contents in the direction of the first principal component. The corresponding α-cellulose calibration including all wood samples had fit and Q^2 values of 0.65 and 0.65, respectively (**Figure 9**). RMSEP for the test set data was 3.3, higher than that of the corresponding lignin model. One reason for this can be the several overlapping bands of cellulose and hemicelluloses in the infrared spectrum. Also, the average standard deviation (1.5) for the parallel α-cellulose determination was higher than that for acetyl bromide lignins (0.9), which undoubtedly contributed to the DRIFT-MIR estimation of cellulose contents. Meder et al. (*6*) previously reported R^2 and
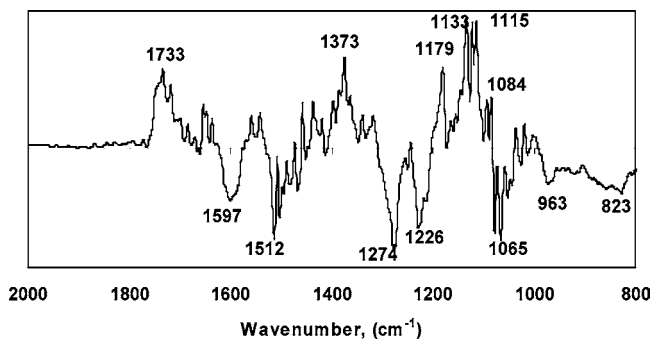
RMSEP values of 0.58 and 6.14, respectively, for the DRIFT-MIR-based calibration on the total carbohydrate content.

Cellulose calibration for the Sitka spruce data set gave R^2 and Q^2 values of 0.67 and 0.66, respectively, whereas RMSEE was 1.7. A test set consisting of the Sitka spruce samples (243) had a RMSEP of 3.7. The loading spectrum of the cellulose model exhibited bands characteristic of the Sitka spruce samples containing high and low amounts of cellulose (**Figure 10**). Positive bands (1373, 1179, 1135, 1115, and 1084 cm^{-1}) have been assigned to cellulose (*34, 35*), whereas negative bands (1594, 1510, 1273, 1218, and 823 cm^{-1}) are distinctive of the guaiacyl lignin (*29*). These MIR vibrations confirm that the calibration is based on the spectral differences due to variation in cellulose content.

Only five characteristic MIR bands for cellulose (1733, 1373, 1179, 1133, and 1116 cm^{-1}) were selected to construct a model to study the effect of the reduced spectral range on the predictive ability of the calibrations. Both cellulose models built from the five cellulose bands had slightly lower RMSEP values (**Table 3**) than the calibrations including the spectral ranges of 4000−2800 and 1800−700 cm^{-1}.

Wood Resin Content. The gravimetric amounts of ethyl acetate extracts in the clonal Sitka spruce samples varied between 0.3 and 9.5% (w/w). The average content of the ethyl acetate soluble wood resins was 0.7% (w/w), which is consistent with the values reported previously for spruce sapwood (*37, 38*). Core samples that contained high levels of extractives included knots. Previous studies showed that softwood knots can contain large amounts of lipophilic extractives (*37, 39*).

Sitka spruce samples were separated according to their extractives content in the PLS analysis of the DRIFT-MIR

Wood Density and Chemical Composition

J. Agric. Food Chem., Vol. 54, No. 1, 2006 **39**
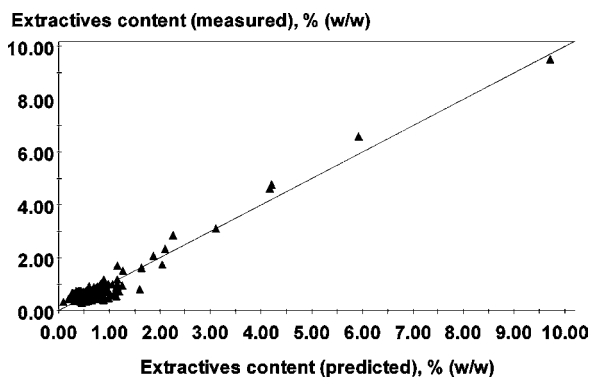
Figure 11. Calibration between extractives content and DRIFT-MIR spectral data of Sitka spruce samples. R^2, Q^2, and RMSEE values for the calibration are 0.94, 0.93, and 0.2, respectively.
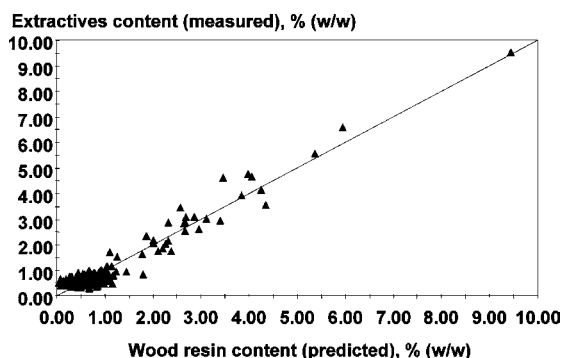
Figure 12. Calibration between extractives content and DRIFT-MIR spectral data of Sitka spruce and Scots pine samples. R^2, Q^2, and RMSEE values for the calibration are 0.93, 0.93, and 0.3, respectively.

spectral data. Calibration for the extractives contents of the Sitka spruce samples had R^2, Q^2, and RMSEP values of 0.94, 0.93, and 0.4, respectively (**Figure 11**). PLS calibration of the amount of extractives that included both Sitka spruce and Scots pine samples is shown in **Figure 12**. The relationship gave R^2 and Q^2 values of 0.93. RMSEE and RMSEP for the calibration data set were 0.3 and 0.4, respectively.

The loading line plot revealed spectral differences due to wood resin content. A characteristic band in the loading line plot for the samples containing higher levels of extractives was seen at 1697 cm^{-1}, which is derived from the C=O vibration of carboxylic acids (*40*). Resin acids are predominating free carboxylic acids in sapwood, because the majority of the fatty acids occur as triglycerides in sapwood (*41*). Furthermore, the bands at 2927 and 1458 cm^{-1} in the loading spectrum have been associated with C—H stretch and bend, respectively (*42*, *43*).

Unlike for the density, cellulose, and lignin models, a reduced number of wavelengths for the wood resin calibrations increased RMSEP and RMSEE values for both softwood models (**Table 3**).

DISCUSSION

RMSEP values of the PLS models shown in this paper were of the same magnitude as reported earlier for MIR-based calibrations of the wood properties. Schulz and Burns (*44*) reported analogous standard error of prediction (SEP) values for the lignin and cellulose predictions, whereas Meder et al. (*6*) determined wood densities using the DRIFT technique and reported a RMSEP value of 37 for the density model. Standard error values for the test set data were higher than those for the calibration sets. This can partly result from the heterogeneity

of the wood samples, because the analysis area is relatively small (∼8 mm^2) in DRIFT measurements. In addition, density was measured as an average over the several annual rings, which probably affected the results. It is known that the density of wood varies between annual rings as well as within latewood and earlywood. The particle size of the samples also affects the DRIFT-MIR spectra in the region of 1200−950 cm^{-1} (*25*, *26*) (**Figure 1**). PLS calibrations were also performed by excluding the spectral range affected by the particle size variation, and no significant differences in the predictive values compared to the models including the spectral range of 1200−900 cm^{-1} were observed. For example, R^2 and Q^2 values for the density model that included all of the wood species were 0.89 and 0.88, respectively, which was the same as for the model with the broader spectral range (**Table 3**). OSC pretreatment and calculation of the average spectra enhanced predictive values of the models considerably. In addition to the particle size variation, the precision of the wet chemical methods possibly contributed to the standard error values of the calibration and test sets. Some of the models contained outliers that were mostly due to the poor spectral quality. We did not detect outliers, for example, resulting from wood resin, in the PLS models built from Sitka spruce.

For the models, the first principal component was the most prominent one explaining the majority of the variation between spectral and other data. For the calibrations constructed from selected wavenumbers, two or three components needed to be calculated (**Table 3**), although the first principal components accounted for the majority of the variation between spectral data and other parameters.

Reduced numbers of wavelengths from 2386 to only the 4 or 5 most significant ones resulted in slightly lower RMSEP values for the calibrations of cellulose and density, whereas that for the lignin calibration was a little higher. The robustness of the model with vastly reduced wavenumbers will be of great benefit for the development of the low-cost hand-held device. In addition, calibrations including all wood samples had RMSEP values similar to those of the models constructed only from Sitka spruce samples, which means that one calibration can be utilized for predicting properties of several wood species.

ACKNOWLEDGMENT

LITERATURE CITED

(1) Haukson, J. B.; Bergqvist, G.; Bergsten, U.; Sjöström, M.; Edlund, U. Prediction of basic wood properties for Norway spruce. Interpretation of near infrared spectroscopy data using partial least squares regression. *Wood Sci. Technol.* **2001**, *35*, 475−485.

(2) Schimleck, L. R.; Evans, R.; Ilic, J. Estimation of *Eucalyptus delegatensis* wood properties by near infrared spectroscopy. *Can. J. For. Res.* **2001**, *31*, 1671−1675.

(3) Yeh, T.-F.; Chang, H.-M.; Kadla, J. F. Rapid prediction of solid wood lignin content using transmittance near-infrared spectroscopy. *J. Agric. Food Chem.* **2004**, *52*, 1435−1439.

(4) Schimlek, L. R.; Evans, R. Estimation of *Pinus Radiata* D. Don tracheid morphological characteristics by near infrared spectroscopy. *Holzforschung* **2004**, *58*, 5866−73.

(5) Gierlinger, N.; Schwanninger, M.; Hinterstoisser, B.; Wimmer, R. Rapid determination of heartwood extractives in *Larix* sp. by means of FT-NIR. *J. Near Infrared Spectrosc.* **2002**, *10*, 203−214.

40   *J. Agric. Food Chem.,* Vol. 54, No. 1, 2006

Nuopponen et al.

(6) Meder, R.; Gallagher, S.; Mackie, K. L.; Bohler, H.; Meglen, R. R. Rapid determination of the chemical composition and density of *Pinus radiata* by PLS modelling of transmission and diffuse reflectance FTIR spectra. *Holzforschung* **1999**, *53*, 261−266.

(7) Nuopponen, M.; Wikberg, H.; Vuorinen, T.; Maunu, S.; Jämsä, S.; Viitaniemi, P. Heat-treated wood exposed to weathering. *J. Appl. Polym. Sci.* **2004**, *91*, 2128−2134.

(8) Ona, T.; Sonoda, T.; Ito, K.; Shibata, M.; Ootake, Y. *In situ* determination of proportion of cell types in wood by Fourier transform Raman spectroscopy. *Anal. Biochem.* **1999**, *268*, 43−48.

(9) Kelley, S. S.; Rials, T. G.; Snell, R.; Groom, L.; Sluiter, A. Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood. *Wood Sci. Technol.* **2004** *38*, 257−276.

(10) Gierlinger, N.; Schwanninger, M.; Hinterstoisser, B.; Wimmer, R. Rapid determination of heartwood extractives in *Larix* sp. by means of FT-NIR. *J. Near Infrared Spectrosc.* **2002**, *10*, 203−214.

(11) Thygesen, L. Determination of dry matter content and basic density of Norway spruce by near infrared reflectance and transmittance spectroscopy. *J. Near Infrared Spectrosc.* **1994**, *2*, 127−135.

(12) Schimleck, L. R.; Mora, C.; Daniels, R. F. Estimation of the physical properties of green *Pinus taeda* radial samples by near infrared spectroscopy. *Can. J. For. Res.* **2003**, *33*, 2297−2305.

(13) Schimleck, L. R.; Evans, R.; Ilic, J.; Matheson, A. C. Estimation of wood stiffness of increment cores by near infrared spectroscopy. *Can. J. For. Res.* **2002**, *32*, 129−135.

(14) Schimleck, L.; Mora, C.; Daniels, R. F. Estimation of tracheid morphological characteristics of green *Pinus* taeda L. radial strips by near infrared spectroscopy. *Wood Fiber Sci.* **2004**, *36*, 527−535.

(15) Silva, J. C.; Nielsen, B. H.; Rodriques, J.; Pereira, H.; Wellendorf, H. Rapid determination of the lignin content in Sitka spruce (*Picea sitchensis* (Bong.) Carr.) wood by Fourier transform infrared spectrometry. *Holzforschung* **1999**, *53*, 597−602.

(16) Rodrigues, J.; Faix, O.; Pereira, H. Determination of lignin content of *Eucalyptus globulus* wood using FTIR spectroscopy. *Holzforschung* **1998**, *52*, 46−50.

(17) Rodrigues, J.; Puls, J.; Faix, O.; Pereira, H. Determination of monosaccharide composition of *Eucalyptus globulus* wood by FTIR spectrsocopy. *Holzforschung* **2001**, *55*, 265−269.

(18) Zanuttini, M.; Citroni, M.; Martínez, M. J. Application of diffuse reflectance infrared Fourier transform spectroscopy to the quantitative determination of acetyl groups in wood. *Holzforschung* **1998**, *52*, 263−267.

(19) Ona, T.; Sonoda, T.; Ito, I.; Shibata, M.; Kato, T.; Ootake, Y. Non-destructive determination of hemicellulosic neutral sugar composition in native wood by Fourier transform Raman spectroscopy. *J. Wood Chem. Technol.* **1998**, *18*, 27−41.

(20) Ona, T.; Sonoda, T.; Ohshima, J.; Yokota, S.; Yoshozawa, N. A rapid quantitative method to assess *Eucalyptus* wood properties for kraft pulp production by FT-Raman spectroscopy. *J. Pulp Pap. Sci.* **2003**, *29*, 6−10.

(21) Bergström, B.; Gustafsson, G.; Gref, R. Seasonal changes of pinosylvin distribution in the sapwood/heartwood boundary of *Pinus sylvestris. Trees* **1999**, *14*, 65−71.

(22) Ona, T.; Sonoda, T.; Ito, K.; Shibata, M.; Kato, T.; Ootake, Y. Determination of wood basic density by Fourier transform Raman spectroscopy. *J. Wood Chem. Technol.* **1998**, *18*, 367−379.

(23) Brendel, O.; Iannetta, P. P. M.; Stewart, D. A rapid and simple method to isolate pure α-cellulose. *Phytochem. Anal.* **2000**, *11*, 7−10.

(24) Johnson, D. B.; Moore, W. E.; Zank, L. C. The spectrophotometric determination of lignin in small wood samples. *Tappi* **1961**, *44*, 793−798.

(25) Faix, O.; Böttcher, J. H. The influence of particle size and concentration in transmission and diffuse reflectance spectroscopy of wood. *Holz Roh Werkst.* **1992**, *50*, 221−226.

(26) Michell, A. J. An anomalous effect in the DRIFT spectra of woods and papers. *J. Wood Chem. Technol.* **1991**, *11*, 33−40.

(27) Eriksson, L.; Johansson, E.; Kettaneh-Wold, N.; Wold, S. PLS. In *Multi- and Megavariate Data Analysis: Principles and Applications*; Umetrics: Umeå, Sweden, 2002; p 533.

(28) Harrington, K. J.; Higgings, H. G.; Michell, A. J. Infrared spectra of *Eucalyptus regnans* F. Muell and *Pinus radiata* D. Don. *Holzforschung* **1964**, *18*, 108−113.

(29) Faix, O. Classification of lignins from different botanical origins by FT-IR spectroscopy. *Holzforschung* **1991**, *45* (Suppl.), 21−27.

(30) Collier, W. E.; Schultz, T. P.; Kalasinsky, V. F. Infrared study of lignin: reexamination of aryl−alkyl ether C−O stretching peak assigments. *Holzforschung* **1992**, *46*, 523−528.

(31) Hinterstoisser, B.; Salmén, L. Application of dynamic 2D FTIR to cellulose. *Vib. Spectrosc.* **2000**, *22*, 111−118.

(32) Savidge, R. A.; Barnett, J. R.; Napier, R. *Cell and Molecular Biology of Wood Formation*; BIOS Scientific Publishers: Oxford, U.K., 2000; p 530.

(33) Kellogg, R. M.; Satry, C. B. R.; Wellwood, R. W. Relationships between cell wall composition and cell wall density. *Wood Fiber* **1975**, *71*, 170−177.

(34) Saka, S. Chemical composition and distribution. In *Wood and Cellulosic Chemistry*; Hon, D. N.-S., Shiraishi, N., Eds.; Dekker: New York, 1991; pp 59−88.

(35) Åkerholm, M.; Hinterstoisser, B.; Salmén, L. Characterization of the crystalline structure of cellulose using static and dynamic FT-IR spectroscopy. *Carbohydr. Res.* **2004**, *339*, 569−578.

(36) Kaèuráková, M.; Smith, A. C.; Gidley, M. J.; Wilson, R. H. Molecular interactions in bacterial cellulose composites studied by 1D FT-IR and dynamic FT-IR spectroscopy. *Carbohydr. Res.* **2002**, *337*, 1145−1153.

(37) Willför, S.; Hemming, J.; Reunanen, M.; Eckerman, C.; Holmbom, B. Lignans and lipophilic extractives in Norway spruce knots and stemwood. *Holzforschung* **2003**, *57*, 27−36.

(38) Ekman, R. Analysis of the non-volatile extractives in Norway spruce sapwood and heartwood. *Acta Acad. Abo, Ser. B* **1979**, *39*, 1−20.

(39) Willför, S.; Hemming, J.; Reunanen, M.; Holmbom, B. Phenolic and lipophilic extractives in Scots pine knots and stemwood. *Holzforschung* **2003**, *57*, 359−372.

(40) Holmgren, A.; Bergström, B.; Ericsson, A.; Gref, R. Detection of pinosylvins of Scots pine using Fourier transform Raman and infrared spectroscopy. *Wood Chem. Technol.* **1999**, *19*, 139−150.

(41) Sjöström, E. Extractives. In *Wood Chemistry Fundamentals and Applications*; Academic Press: San Diego, CA, 1993; pp 92−107.

(42) Nuopponen, M.; Vuorinen, T.; Viitaniemi, P.; Jämsä, S. Effects of heat treatment on the behaviour of extractives in softwood studied by FTIR spectroscopic methods. *Wood Sci. Technol.* **2003**, *37*, 109−115.

(43) Williams, D. H.; Fleming, I. Infrared spectra. In *Spectroscopic Methods in Organic Chemistry*, 5th ed.; McGraw-Hill: Glasgow, U.K., 1995; pp 28−62.

(44) Schulz, T. P.; Burns, D. A. Rapid secondary analysis of lignocellulose: comparison of near infrared (NIR) and Fourier transform infrared (FTIR). *Tappi* **1990**, *73*, 209−212.