

Stephen S. Kelley · Timothy G. Rials  
Rebecca Snell · Leslie H. Groom · Amie Sluiter

## **Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood**

Received: 31 January 2002 / Published online: 13 May 2004  
© Springer-Verlag 2004

**Abstract** Near infrared (NIR) spectroscopy (500 nm–2400 nm), coupled with multivariate analytic (MVA) statistical techniques, have been used to predict the chemical and mechanical properties of solid loblolly pine wood. The samples were selected from different radial locations and heights of three loblolly pine trees grown in Arkansas. The chemical composition and mechanical properties were measured with traditional wet chemical techniques and three point bending tests, respectively. The microfibril angle was measured with x-ray scattering. These chemical and mechanical properties were correlated with the NIR spectra using projection to latent structures (PLS) models. The correlations were very strong, with the correlation coefficients generally above 0.80. The mechanical properties could also be predicted using a reduced spectral range (650 nm–1150 nm) that should allow for field measurements of these properties using handheld NIR spectrometers.

---

### **Background**

Near infrared (NIR) spectroscopy has been used for the characterisation of different forms of biomass for more than 15 years (Marten et al. 1985). Early work focused on the agricultural and food industries, and remote sensing applications. More recent work has expanded the use of NIR for applications of interest to the forest products industry.

---

S. S. Kelley (✉) · A. Sluiter  
National Bioenergy Center,  
National Renewable Energy Laboratory,  
1617 Cole Blvd. Golden, 80401, CO, USA  
E-mail: stephen\_kelley@nrel.gov

R. Snell · L. H. Groom  
USFS, Southern Research Station,  
2500 Shreveport Highway, Pineville,  
71360, LA, USA

T. G. Rials  
Forest Products Center, University of Tennessee,  
2506 Jacob Dr. Knoxville, 37996, TN, USA

Relative to other spectroscopic techniques NIR has a number of advantages that make it an ideal tool for characterising biomass. These advantages include minimal sample preparation, rapid acquisition times, and non-contact, non-destructive spectral acquisition. Some commercial NIR spectrometers are also lightweight, portable, and battery operated, which make them suitable for field operation.

Some of the initial applications of NIR to forestry and the forest products industry focused on forest health and analysis of leaves and needles (Newman et al. 1994; Aber et al. 1994; Martin and Aber 1994; McLellan et al. 1991a; McLellan et al. 1991b; Couteaux et al. 1997; Hiukka 1998). Together, these researchers showed that NIR could be used to accurately measure nitrogen, cellulose and lignin content of fresh and dried leaves and needles. The NIR spectra could be correlated with the measured compositions using several different statistical techniques including stepwise multiple regression (Newman et al. 1994; Martin and Aber 1994; McLellan et al. 1991a; McLellan et al. 1991b) and projection of latent structures (PLS) analysis (Couteaux et al. 1997; Hiukka 1998). Looking across all these studies, the correlations between NIR spectra and nitrogen content were very strong,  $r > 0.95$ , while the correlations between NIR spectra and the cellulose, starch, or lignin content were very good  $r > 0.90$ . Hiukka found that the correlations between NIR and several specific carbohydrates, e.g., glucose, fructose, pinitol, sorbitol, and inositol, were modest (Hiukka 1998). McLellan et al. (1991b) concluded that there were greater errors for inter-laboratory analysis of foliage samples using traditional wet chemistry compared to NIR based techniques.

NIR has also been used to study cellulose and cellulose derivatives. Sorption of water onto cellulose surfaces is a fruitful area of study due to the strong NIR signal for water and carbohydrate hydroxyl groups (Berthold et al. 1998; Svedas 2000). These studies show that NIR could be used to measure the amount of water absorbed onto the cellulose surface, and distinguish between "free" and "bound" water. Ali et al. (2001) showed that NIR could be used to monitor the aging of paper used for insulation materials. Specific chemical changes of cellulose associated with aging could be detected. Chemical differences of cellulose derivatives have also been monitored with NIR (Svensson et al. 1997).

Recently, the use of NIR to measure wood properties of interest to the pulp and paper industry has received a great deal of attention. Some of the earliest work showed that NIR could be used to measure properties of the wood resource and the resulting paper products (Wright et al. 1990; Michell 1995; Wallbacks et al. 1995). Wright et al. (1990) used NIR to predict both pulp yield and cellulose content of pulps. Michell (1995) looked at both wood and pulp properties. He showed that NIR could be used to measure pulp yield, lignin content, and hot water and alkaline extractable compounds. Wallbacks et al. (1995) demonstrated that NIR spectroscopy could be used to detect day-to-day variations in the characteristics of beaten and unbeaten pulps, and measure paper properties such as tear index, burst index, and elastic modulus. Complete analysis of the chemical composition of wood, e.g., cellulose, lignin, hemicellulose, glucose, xylose, acetyl, can also be conducted with NIR (Schimleck et al. 1997). The correlation coefficients for most of these wood components were very strong,  $r > 0.95$ .

Subsequent work by Schimleck and Michell (1998) and Raymond et al. (2001) highlighted the value of using NIR for predicting properties of interest. This work

shows that NIR models for a property of interest, e.g., pulp yield, can be constructed with a small number of samples, and that this model can then be used as a predictive tool to study a large number of samples with unknown properties. This approach shows how NIR can be used to map relevant properties in trees, e.g., pith to bark or base to crown (Schimleck and Michell 1998; Raymond et al. 2001), or between trees subjected to different silvicultural treatments or growing conditions (Raymond et al. 2001; Schimleck et al. 1996). Recently this approach was extended to drawing correlations between wood composition and genetic characteristics of wood (Greaves et al. 1996).

The majority of work with NIR has focused on chemical composition of wood and paper properties; however, some research has also been done on other wood properties. One of the most extensive studies of wood properties was that of Hoffmeyer and Pedersen (1995) who showed that NIR could be used to predict wood density, and compression and bending strength of dry wood. Tygesen (1994) also showed that NIR could be used to measure the density of wood. Additionally, the use of NIR to predict the stiffness of radiata pine has been reported by Thumm and Meder (2001).

These papers have focused on using the full NIR spectral range (1000 nm–2500 nm) or both the visible and NIR spectral range (400 nm–2500 nm). This is because the range between 1000 nm–2500 nm contains the most distinct spectral information on the first overtone and combination bands that make up NIR spectra. Two recent studies have revealed that even the very subtle second overtone spectral signals contain information that can be used to predict wood properties of interest. Axrup et al. (2000) showed that the full chemical composition of wood chips—glucose, galactose, xylose, mannose, arabinose, and lignin—could be measured using the spectral information between 800 nm and 1100 nm. They also measured the size distribution of bark and the amount of wood in bark. Malkavaara and Alen (1998) measured the lignin content of kraft pulps subjected to different bleaching sequences using a spectral range between 360 nm–740 nm. In both cases the correlation coefficients were generally above 0.90, indicating that large reductions in the spectral range did not have a significant negative effect on the quality of the models. One report on the use of a reduced spectral range (400 nm–1100 nm) for predicting the stiffness of wood showed a substantial decrease in the quality of the models (Thumm and Meder 2001). Defining the effects of reducing the spectral range is very significant since it can be accessed with very inexpensive, lightweight spectrometers that can acquire spectra in fractions of a second. In fact, the results of Axrup et al. were gathered on moving wood chips, highlighting the potential for using NIR for process control applications by the pulp and paper industry.

Most of the work done on NIR analysis of wood and wood products has used some kind of simple spectral preprocessing such as the first or second derivatives, or multiplicative scattering correction. More recently, other analytical techniques have been used to improve the quality of the models constructed using NIR spectral data (Axrup et al. 2000; Marklund et al. 1999). Axrup et al. (2000) used fixed window evolving factor analysis to identify outliers. This type of analysis would be required to use NIR in a processing environment. Marklund et al. (1999) used orthogonal signal correction (OSC) to improve the correlations and predictive quality of PLS models. They obtained high quality correlations between the NIR spectral properties of pulp and strength properties of paper derived from those pulps. Using OSC they

could extend the correlations back to the wood originally used to produce the pulps.

This previous work has illustrated the potential for using NIR to measure the chemical and physical properties of wood. This paper extends this prior work to include predictions of a variety of chemical and mechanical properties from one uniform set of solid wood samples, and evaluates the impact of reducing the spectral range from 500 nm–2400 nm to 650 nm–1150 nm.

---

## Experimental

### Collection of wood samples

The preparation of the samples is described in greater detail elsewhere and is only briefly outlined here. Three loblolly pine trees were harvested from a plot in southern Arkansas. Two foot long disks were cut every 16 feet along the trees. Slices were isolated from the disks at each compass heading. Samples containing three growth rings were then cut from the slice to produce the bending samples. These samples were oven dried at 105°C, and reconditioned prior to testing. The mechanical properties of the samples from each of the four compass headings were measured and averaged to provide one value for each tree, height and growth ring. Samples from a single growth ring from each tree, height, and compass heading were ground and mixed to form a master batch that was then subjected to chemical analysis.

### NIR measurements

The NIR measurements were made with an Analytical Spectral Devices (ASD) Field Spec at wavelengths between 400 nm–2,500 nm. A fiber optic probe oriented at a right angle to the sample surface was used to collect the reflectance spectra. A piece of commercial, microporous Teflon was used as the white reference material. The samples were illuminated with a DC lamp oriented at 30 ° above the samples and aligned parallel with the longitudinal axis of the sample. Thirty scans were collected and averaged into a single average spectrum. The area that was sampled varied depending on the width of the individual bending specimen, but was generally about 2 cm<sup>2</sup>. Two average spectra were taken from the different locations near the center of both the tension and compression surface of the bending samples, providing four averaged spectra for each sample.

The reflectance spectra were transferred from the ASD to an Unscrambler file. The reflectance spectra were converted to absorbance spectra in the Unscrambler. The four averaged spectra collected on each sample were averaged to provide a single spectrum that was used to predict the mechanical strength or chemical composition of the sample. The data set was further reduced by averaging the spectra that were collected at 1 nm intervals, to a spectral data set at 10 nm intervals. Averaging the spectral data reduces the size of the spectra matrix and significantly reduces the time required to compute the PLS models without decreasing the quality of the models. There are a number of preprocessing techniques that can be used to increase the quality of the PLS models, e.g., second derivatives or orthogonal signal correction (Wold et al. 1998; Fern 2000). However, these techniques greatly complicate ones ability to

provide chemical interpretation of the regression coefficients. In this study no preprocessing techniques were used.

The average spectra were used to perform a principal component analysis of the samples. One sample, tree 1, disk 1, ring 16, was identified as a spectral outlier and was excluded from the subsequent analyses.

#### Chemical analysis

All samples were milled using a standard Wiley knife mill with a 2 mm screen. The -2 mm material was sieved for five minutes using a 20 mesh sieve. Any +20 mesh material was milled until it passed the 20 mesh screen (American Society for Testing Material (ASTM) 1999). All samples were analysed using ASTM standard methods for whole biomass analysis (ASTM 1999). NIST pine standard reference material # 8493 was used as a reference for the determination of extractives, lignin, and carbohydrates. All determinations were performed in duplicate on the prepared biomass samples. Analysis errors for softwood composition using these traditional wet chemical analysis methods are 0.5% for lignin and ash, 1.0% for glucose and extractives, and 1.5% for all other sugars (Milne et al. 1992).

Two samples were identified as outliers with mass closure for all wood components above 105% and were not used in the models for predicting the chemical composition of wood.

#### Measurement of mechanical properties

All the samples were broken in three point bending according to ASTM standards (ASTM 1999). The stiffness (modulus of elasticity (MOE)) was calculated using the linear portion of the load-deflection curve. The ultimate strength (modulus of rupture (MOR)) was the breaking load divided by the cross-sectional area.

#### Microfibril angle measurements

X-ray measurements were conducted on a General Electric XRD-7 diffractometer. A copper target was used at an accelerating voltage of 50 kV. The x-ray spot at the surface of the sample was 3.8 mm. The detector, a sealed xenon gas proportional counter, was positioned to pick up the 002 diffraction arc. Specimens were produced by sampling from one end of the desired growth ring. Sample dimensions were approximately 6.5 mm in the tangential by longitudinal direction, and did not exceed 3 mm thick in the radial direction. The sample was placed with the tangential face perpendicular to the x-ray beam. The diffracted intensity was recorded as the sample was rotated through 360°. Additional procedural details and data analysis methods are available (Huang et al. 1997).

#### Projection to latent structures (PLS) modeling

While a complete description of multivariate analysis can be found elsewhere (Martens and Naes 1991; Vandeginste et al. 1998) the following summary describes the steps used to construct PLS models in this work. Multivariate

analysis was performed using the Unscrambler version 7.6 (CAMO, Corvallis, OR). The package has the capability to perform both principal component analysis (PCA) and projection to latent structures (PLS) (also known as partial least squares) analyses. All of the NIR spectra are combined into a single data matrix (X-matrix) while the chemical and mechanical property data are combined into a response matrix (Y-matrix). The software is used to systematically extract (decompose) variation in the data matrix (X-matrix) while principal component regression is used to regress each response variable (Y-matrix) onto the decomposed spectra (X-matrix), and make a projection to latent structures. This process allows for the simultaneous and independent decomposition of both the X- and Y-matrices and then performs the regression of the Y-matrix onto the X-matrix, e.g., the chemical or mechanical properties onto the spectra.

The Unscrambler also allows for calculation of two different PLS algorithms. A PLS-1 analysis allows only one Y-variable to be projected against the X-matrix at a time, e.g., MFA, while a PLS-2 analysis allows several Y-variables to be projected against the X-matrix, e.g., all of the chemical properties, or MOR and MOE. Both PLS-1 and PLS-2 models were constructed for different response variables. The models were constructed with an X-matrix of 190 points (500 nm–2400 nm) or 50 points (650 nm–1150 nm), and the chemical and mechanical properties as the Y-matrix. Both the X- and Y-matrices were mean centered variance normalised prior to performing the PLS analysis. When PLS-2 models were constructed, weighting the Y-matrix values compensated for the large differences in the magnitude of the MOE and MOR values by the inverse of their standard deviations. The number of principal components (factors) used for a model were selected by observing the response of the residual Y-variance with added factors. When additional factors did not substantially decrease the residual Y-variance, additional iterations were terminated. All of the PLS models are based on either 4 or 5 factors.

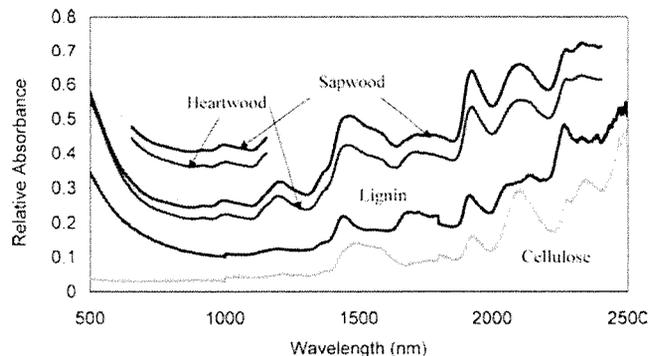
Calibration models (CALB) were constructed with about two-thirds of the samples (45) using full cross-validation. Cross-validation systematically removes a single sample from the data set, constructs a model with the remaining samples, and uses that model to predict the value(s) of the Y-variable(s) for the extracted sample. This process continues until each individual sample has been removed from the data set and a fully cross-validated model is constructed (Martens and Naes 1991; Vandeginste et al. 1998). This fully cross-validated model was then used to predict the response of the test set (TEST) that contains about one-third of the samples (27) that were not included in the original model. This conservative approach insures that the predictive capabilities of the model are reliable.

---

## Results and discussion

### NIR spectra

Representative spectra of wood, cellulose and lignin are shown in Fig. 1. The differences between cellulose and lignin are clear for some regions of the NIR spectra, but there is also considerable overlap. Based on the NIR spectra of the isolated wood components and literature references many of the vibrations in the NIR spectra of wood can be assigned (Ali et al. 2001; Curran et al. 1992;



**Fig. 1** NIR spectra of representative loblolly pine samples, and cellulose and lignin model compounds. The reduced spectral range of the same loblolly pine wood is also shown to highlight this region spectral region

Fourty et al. 1996; Marten et al. 1985). The major vibrations include the yellow-brown color of the wood at 400 nm–700 nm that are primarily due to the presence of lignin and extractives. The first overtone of cellulose and hemicellulose hydroxyls between 1400 nm and 1660 nm, and the interactions between carbohydrate hydroxyls and water between 1890 nm and 2020 nm can be seen. There is also a strong vibration at 2020 nm–2250 nm that has been assigned to the cellulose hydroxyl vibrations. The first and second overtones of the lignin aromatic and aliphatic carbon/hydrogen vibrations are seen between 1635 nm and 1825 nm and 1075 nm and 1250 nm, respectively. Some of the lignin hydroxyl vibrations overlap with the cellulose hydroxyl vibrations, e.g., the first overtone of the lignin hydroxyl vibrations occur between 1400 nm and 1520 nm. These assignments provide some insight into the chemical structures present in the material, but the overlapping bands limit the information available from simple, visual inspection of NIR spectra.

Figure 1 shows the NIR spectra of wood in the visible and short wavelength NIR region (650 nm–1050 nm). These spectra show some features, but these features are very subtle. The peaks in this region are even more difficult to assign to specific wood components but can be attributed to second overtones of hydroxyls and third overtones of C-H stretching vibrations. There is also a contribution from the brown color of lignin in the visible region of the spectra.

### Chemical analysis

Several publications have highlighted the use of spectroscopic techniques, including NIR, to predict the chemical composition of wood (Schimleck et al. 1997; Axrup et al. 2000; Malkavaara and Alen 1998). This work extends these studies by using solid wood samples and investigating the use of a reduced spectral range.

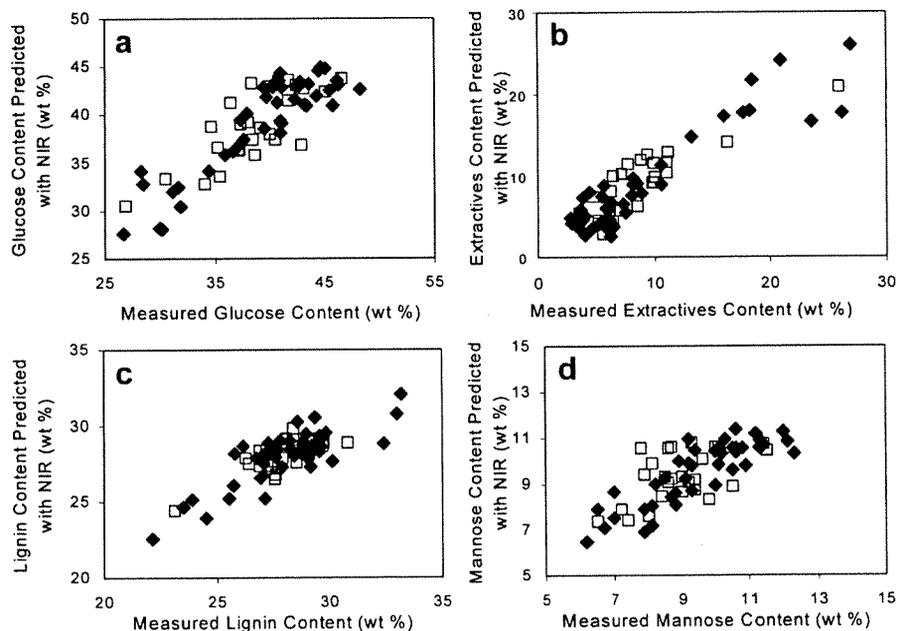
A total of 72 samples representing different tree, height, and growth ring combinations were randomly assigned to either the CALB or TEST set for further analysis. The range of chemical compositions, determined by traditional wet chemistry, of the samples in the CALB and TEST set are shown in Table 1. The two sets are similar in both their medium composition and their composition range. A detailed discussion of the changes in chemical composition with

**Table 1** Compositional range for the 70 loblolly pine samples used for construction of PLS-2 models of wood chemistry. All values are weight percent

	Lignin	Glucose	Xylose	Mannose	Galactose	Extractives	Closure
CALB set							
Min	22.1	26.7	5.0	6.2	1.8	2.8	95.7
Max	33.2	48.2	9.4	12.3	8.3	26.9	104.4
Median	27.9	40.9	7.1	10.0	4.3	6.2	98.1
TEST set							
Min	23.1	26.9	6.1	6.5	2.4	3.6	95.8
Max	30.8	46.5	8.9	11.5	7.7	25.9	103.1
Median	27.9	38.4	7.8	8.7	4.1	6.4	97.5

respect to their location within the tree, and the differences between trees are provided elsewhere (Snell R et al., 2003, submitted).

Using the chemical composition information determined by traditional wet chemistry measurements and the NIR spectra collected on the radial face of the corresponding solid wood sample, a series of PLS models were constructed. All of the following results are based on PLS-2 models that predict all six wood components, e.g., lignin, extractives, glucose, xylose, mannose, and galactose, at one time. The PLS-2 models for the glucose, lignin, mannose, and extractives content of these solid wood samples are shown in Fig. 2a–d. These models were

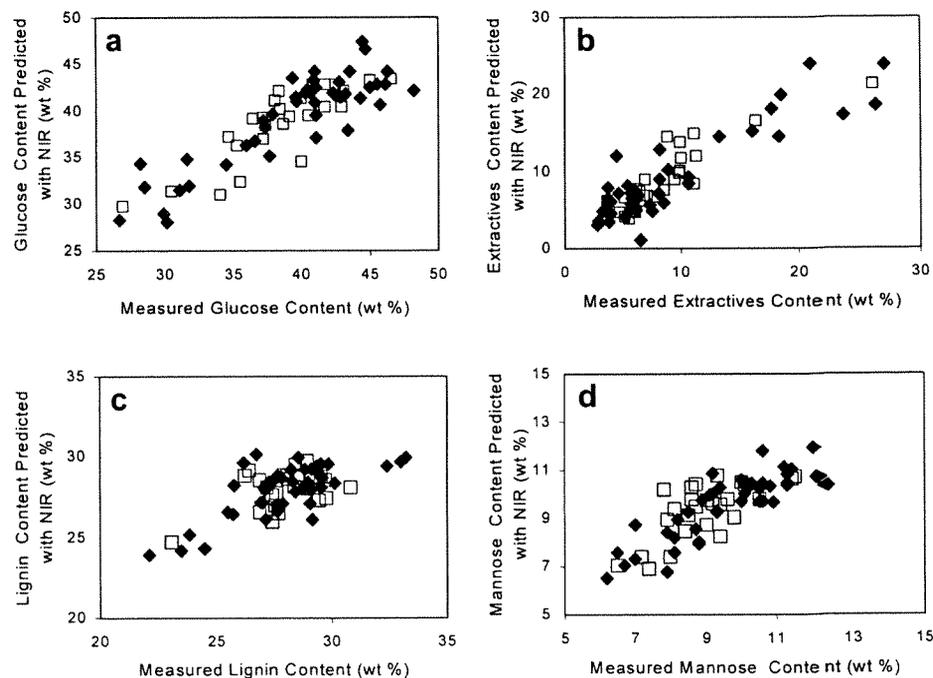


**Fig. 2a–d** The results of PLS-2 models showing the correlation between measured chemical compositions and the chemical composition predicted with NIR spectra collected over the full spectral range (500 nm–2400 nm). a glucose, b extractives, c lignin and d mannose. Solid symbols are the samples included in the CALB set and the open symbols are the samples included in the TEST set

all constructed using the full spectral range (500 nm–2400 nm) and spectra that were averaged over a 10 nm interval.

The 41 samples in the CALB set were used to construct the PLS-2 models for all six of the wood components. These PLS-2 models were then used to predict the chemical composition of samples in the TEST set which had not been included in the original data used to construct the models. Testing the model with samples that were not contained in the original model is an example of how these models might be used to predict the composition of a large number of unknown samples. Examination of the four graphs in Fig. 2 show that the data points for the CALB and TEST set have substantial overlap. With the exception of the lignin model, the values are also uniformly spread across the data range. The values for lignin content are clustered around the mean.

The same wet chemical values, and NIR spectra between 650 nm–1150 nm, were used to construct a second set of PLS-2 models. Again, the 41 samples in the CALB set were used to construct the PLS-2 models for all six of the wood components. These models were then used to predict the chemical composition of samples in the TEST set, which had not been included in the original data used to construct the models. The results of these PLS-2 models are shown in Fig. 3a–d. These models have very similar appearance to the models for the full spectra range. Again, the model for the lignin content has a majority of the data close to the mean.



**Fig. 3a–d** The results of PLS-2 models showing the correlation between measured chemical compositions and the chemical composition predicted with NIR spectra collected over a reduced spectral range (650 nm–1150 nm). **a** glucose, **b** extractives, **c** lignin and **d** mannose. Solid symbols are the samples included in the CALB set and the open symbols are the samples included in the TEST set

The quality of PLS models can be measured in several ways. For this work, the correlation coefficient ( $r$ ) and the root mean square error of prediction (RMSEP) or root mean square error of calibration (RMSEC) were used as a way to compare the PLS models. These data are shown in Table 2.

The results in Table 2 highlight several features of using NIR spectroscopy to predict the chemical composition of solid wood. First, there is very little reduction in  $r$ , RMSEC, or RMSEP when the spectral range is reduced from 500 nm–2400 nm to 650 nm–1150 nm. This is a very significant finding because it opens up these NIR techniques to the use of very low cost and lightweight spectrometers. The only significant reduction in model quality was for the measurement of lignin. This reduction can be attributed to the very low intensity of CH overtone vibrations in this range and the relatively low OH content on the lignin. The second feature that these data highlight is the use of NIR to predict the chemical composition of an unknown sample. In almost all cases the  $r$ -values are slightly lower for the TEST set compared to the CALB set. This is due, in part, to the fact that the TEST set has fewer data points than the CALB set, 26 vs. 44. The RMSEC can also be compared to the RMSEP. In most cases these two measures of the accuracy of the models are comparable.

The number of principal components (PC) can have a significant impact on the quality of a PLS model. Too many PCs can result in overfitting of the model, while too few can result in a less accurate model (REF). In this work four principal components were used to construct the PLS-2 model for the full spectra range and five principal components were used to construct the model for the reduced spectral range. The effects of varying the number of PCs on the quality of the PLS models are shown in Table 3.

As expected, increasing the number of PCs increases  $r$  and decreases the RMSEC or RMSEP for most of the wood components. However, in the case of the models constructed using the full spectrum, there is a notable increase in  $r$  and a decrease in the RMSEC or RMSEP going from three to four PCs, but little improvement in going from four PCs to five PCs. The same trends are seen for the models constructed using the reduced spectral range. Increasing the number of PCs from four to five increases  $r$  and decreases RMSEC or RMSEP, but adding one more PC does not substantially improve the quality of the models.

**Table 2** Correlation coefficients and root mean square error of PLS-2 models for the chemical composition of the CALB and TEST sets. The RMSEC and RMSEP values are weight percent

	Lignin	Glucose	Xylose	Mannose	Galactose	Extractives
500 nm–2400 nm, 4 PCs used for models						
CALB set						
$r$	0.81	0.90	0.80	0.86	0.82	0.93
RMSEC	1.1	2.4	0.6	0.8	1.0	2.3
TEST set						
$r$	0.76	0.78	0.56	0.58	0.80	0.85
RMSEP	1.0	2.7	0.6	1.3	1.0	2.3
650 nm–1150 nm, 5 PCs used for models						
CALB set						
$r$	0.71	0.88	0.80	0.86	0.81	0.93
RMSEC	1.5	2.6	0.8	0.8	1.0	2.4
TEST set						
$r$	0.67	0.84	0.54	0.69	0.83	0.88
RMSEP	1.4	2.3	0.6	1.0	0.8	2.2

**Table 3** Effects of varying the number of principal components on the quality of PLS models for wood chemistry. The RMSEC and RMSEP values are weight percent

	Lignin	Glucose	Xylose	Mannose	Galactose	Extractives
500 nm–2400 nm						
CALB set						
r (3 PCs)	0.81	0.86	0.79	0.79	0.75	0.89
r (5 PCs)	0.82	0.93	0.80	0.88	0.82	0.94
RMSEC (3 PCs)	1.2	2.8	0.6	0.9	1.2	2.8
RMSEC (5 PCs)	1.3	2.0	0.6	0.7	1.0	2.2
TEST set						
r (3 PCs)	0.75	0.74	0.56	0.57	0.77	0.79
r (5 PCs)	0.76	0.86	0.58	0.63	0.82	0.89
RMSEP (3 PCs)	1.0	2.9	0.6	1.3	1.0	2.7
RMSEP (5 PCs)	1.0	2.3	0.6	1.1	1.1	2.0
650 nm–1150 nm						
CALB set						
r (4 PCs)	0.71	0.88	0.80	0.86	0.81	0.89
r (6 PCs)	0.82	0.92	0.81	0.89	0.83	0.93
RMSEC (4 PCs)	1.5	2.6	0.6	0.8	1.0	2.9
RMSEC (6 PCs)	1.3	2.1	0.6	0.7	1.0	2.3
TEST set						
r (4 PCs)	0.60	0.82	0.55	0.66	0.84	0.86
r (6 PCs)	0.43	0.85	0.53	0.66	0.82	0.89
RMSEC (4 PCs)	1.2	2.4	0.7	1.0	0.8	2.2
RMSEC (6 PCs)	1.4	2.3	0.7	1.0	0.8	2.1

These results are comparable to previous work (Schimleck et al. 1997; Hiukka 1998; Axrup et al. 2000). Hiukka (1998) reported r-values above 0.90 for nitrogen and starch content of pine needles, but these correlations decreased to 0.70 to 0.86 for six minor sugars. Using the second derivative of the NIR spectra, but very similar PLS techniques and comparable numbers of PCs, Schimleck et al. (1997) reported r values for a set of *E. globulus* above 0.90 for cellulose, hemicelluloses, and xylan. The correlation for the lignin content was slightly lower. They report RMSEC that are much better for glucose, but comparable for lignin and xylose. Using a reduced spectral range but more elaborate analytical techniques Axrup et al. (2000) reported consistently higher RMSEC and RMSEP values for lignin, extractives, glucose, xylose, galactose, and mannose. However, they collected their NIR spectra and samples on wood chips that were moving on a conveyor belt. Compared to the current results, the prior work shows that very high quality correlations can be obtained for ground samples under ideal laboratory conditions. While the quality of the models may decrease for solid wood or with a reduced spectral range, these NIR techniques have great value for measuring the chemical composition of wood under a wide variety of conditions.

Finally, the accuracy of the NIR models can be compared to the accuracy of wet chemical methods. Based on the work of Milne et al. (1992) NIR analysis is comparable to wet chemistry for the minor sugars, xylose, mannose, and galactose, but 1–2 times less accurate for prediction of glucose, lignin, and extractives. However, given the ease, speed, and low cost of the NIR methods relative to the traditional wet chemistry techniques, this decrease in accuracy is acceptable for many applications.

### Mechanical properties

While it is clear that the NIR spectra contain information on the chemical composition of wood, it is much less obvious that these same spectra contain information on the strength properties of solid wood. There are two reports on the use of NIR of ground wood (Hoffmeyer and Pedersen 1995; Thumm and Meder 2001) to predict the strength properties. It is well-known that the strength properties of wood are related to the density, microfibril angle and slope of grain of wood samples. But, it is not intuitive that the NIR spectra of solid wood will contain information on these wood features.

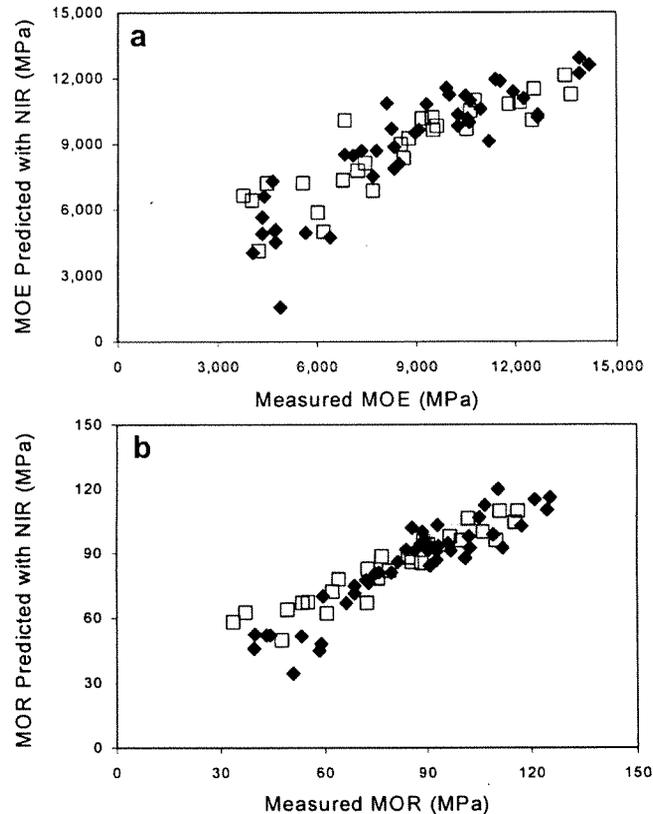
The mechanical properties of the loblolly pine samples were measured in three-point bending. The range of mechanical properties is shown in Table 4. To minimise the relatively large natural variations in the mechanical properties of clear wood, the mechanical properties of the four individual pieces taken at different compass headings at each tree, height, and growth ring location were averaged to provide a single value for that location. As discussed above, the samples were randomly assigned to either the CALB set or TEST set. The same samples were assigned to the CALB and TEST sets for analysis of both chemical and mechanical properties. The average and range of mechanical properties are similar for both the CALB set and TEST set. A detailed discussion of the changes in mechanical properties by location within the tree, and between tree differences are provided elsewhere (Snell R et al., 2003, submitted).

The same NIR spectra collected from the radial face of solid wood samples were used to predict the mechanical strength of these pine samples. The CALB set of samples was used to construct the PLS models while the TEST set was used to validate the models. As mentioned above, the models predicting the mechanical properties of wood are based on PLS-2 models, and the strength properties were normalised by the inverse of their standard deviations.

The results of these predictions using the full spectral range are shown in Fig. 4a and b. The correlations between the measured strength properties and the strength properties predicted with NIR are very good. Both the MOE and MOR can be predicted from NIR spectra. The *r*-values and RMSEC or RMSEP for these models are shown in Table 5. The *r* values for the CALB set models were 0.88 for MOE and 0.92 for MOR. More importantly, the quality of the TEST set was also very good with *r* values of 0.87 and 0.94 for MOE and MOR, respectively. The RMSEP for the TEST set were 1,490 MPa for MOE and 10.4 MPa for MOR. These RMSEP values are about 15% of the mean for

**Table 4** Mechanical properties of the 72 loblolly pine samples used for construction of PLS-2 models of wood strength (MOR) and stiffness (MOE), and microfibril angle (MFA)

	MOR (MPa)	MOE (MPa)	MFA (deg)
CALB set			
Min.	39.6	4,049	6.5
Max.	125.5	15,327	43.0
Median	88.0	8,988	19.5
TEST set			
Min.	33.5	3,791	7.0
Max.	116.3	13,674	36.0
Median	76.6	8,540	19.0



**Fig. 4a,b** The results of PLS-2 models showing the correlation between measured mechanical properties and the mechanical properties (a) MOE and (b) MOR predicted with NIR spectra collected over the full spectral range (500 nm–2400 nm). Solid symbols are the samples included in the CALB set and the open symbols are the samples included in the TEST set

the MOE and 10% of the mean for the MOR. These results are similar to the prior reports (Hoffmeyer and Pedersen 1995; Thumm and Meder 2001) on the use of NIR to measure the bending and compressive strength of wood. These results show that the strength properties of an unknown loblolly pine sample could be accurately predicted from its NIR spectrum.

Similarly, PLS-2 models can be constructed with a reduced spectral range, e.g., 650 nm–1150 nm. It should be emphasised that these models were constructed with only 50 individual spectral data points. The results of these PLS-2 models are shown in Fig. 5a and b. In this case the correlation coefficients for the CALB set model were 0.89 for MOE and 0.92 for MOR, and the correlation coefficients for the TEST set were 0.84 for MOE and 0.91 for MOR. The RMSEP were 1,580 MPa for the MOR and 10.2 MPa for the MOE. Again, RMSEP values are about 15% and 10% of the mean for the MOE and MOR, respectively. These results show that the strength properties of an unknown loblolly pine sample could be accurately predicted from its NIR spectra collected over a reduced spectral range. This further highlights the potential for using small lightweight NIR spectrometers for measuring the strength of wood and trees.

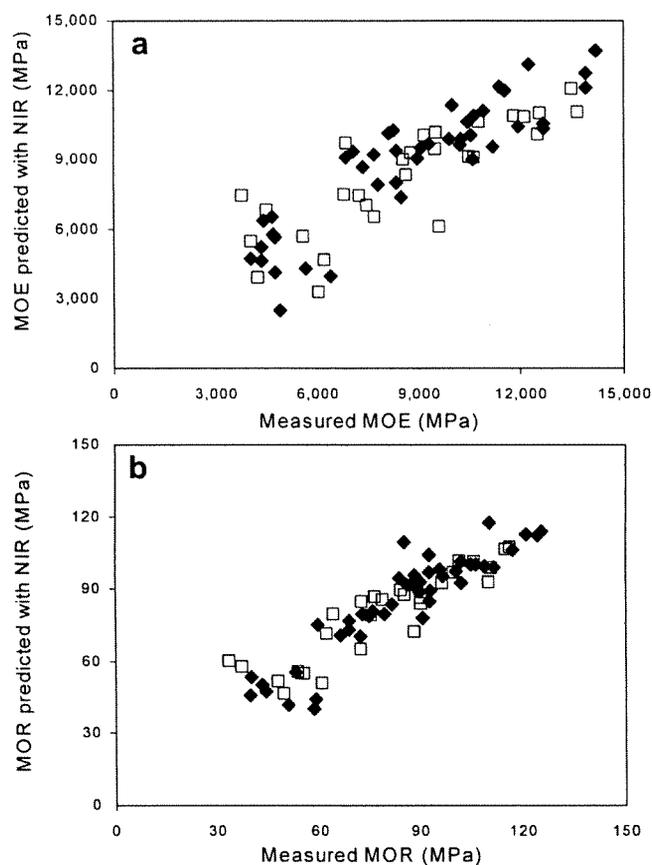
**Table 5** Correlation coefficients and root mean square error of PLS-2 models for the mechanical properties and microfibril angle of the CALB and TEST sets

	MOR (MPa)	MOE (MPa)	MFA (deg)
500 nm–2400 nm, 5 PCs used for models			
CALB set			
r	0.92	0.88	0.82
RMSEC	8.7	1,450	5.7
TEST set			
r	0.94	0.87	0.68
RMSEP	10.4	1,490	6.8
650 nm–1150 nm, 5 PCs used for models			
CALB set			
r	0.92	0.89	0.80
RMSEC	8.8	1,370	6.0
TEST set			
r	0.90	0.84	0.56
RMSEP	10.2	1,580	8.0

The effects of varying the number of PCs on the quality of the mechanical property models are shown in Table 6. As expected increasing the number of PCs generally increases  $r$  and decreases the RMSEC or RMSEP. In most cases increasing the number of PCs from four to five substantially increases the  $r$ -values and decreases the RMSEC or RMSEP. But, further increasing the number of PCs from five to six does not substantially improve the quality of the models.

It is not intuitive that there should be a relationship between reflected NIR radiation and the orientation of the cellulose crystallites, e.g., MFA, in wood. However, this correlation could be a major contributor to the correlation between NIR and mechanical properties. Figure 6a shows the correlation between the MFA measured with x-ray scattering and the MFA predicted from the NIR spectra. The results shown in Fig. 3 are based on PLS-1 models of the MFA using 5 PCs. There is a good correlation between the measured and predicted MFA, with correlation coefficients of 0.82 for the CALB set and 0.68 for the TEST set. These correlations decrease but are still useful when the NIR spectral range is reduced from 500 nm–2400 nm to 650 nm–1150 nm (Fig. 6b). In this case, the correlation coefficients decrease to 0.80 for the CALB set and 0.56 for the TEST set. These correlations are lower than those seen for the chemical or mechanical properties due to the relatively large experimental errors associated with measuring the MFA with x-rays.

The results shown in Fig. 6 indicate that the correlations between mechanical properties and NIR spectra are based, in part, on the relationship between the NIR spectra and MFA. The relationship between mechanical strength and MFA are well-known (Butterfield 1997), and if we can measure MFA with NIR then it is reasonable that we should also be able to measure mechanical strength. There is also a correlation ( $r=0.82$ , not shown) between NIR spectra and specific gravity of solid wood. A strong relationship between MOE and cellulose content has been reported (Tsehaye et al. 1997), and, as shown above, the NIR spectra contain information on the cellulose content of wood. Since molecular information contained in the NIR spectra can be correlated with MFA, specific gravity and cellulose content, it is quite logical that NIR and mechanical properties should be correlated.



**Fig. 5a,b** The results of PLS-2 models showing the correlation between measured mechanical properties and the mechanical properties (a) MOE and (b) MOR predicted with NIR spectra collected over a reduced spectral range (650 nm–1150 nm). Solid symbols are the samples included in the CALB set and the open symbols are the samples included in the TEST set

The fundamental explanation for the relationship between MFA and NIR spectra is not clear at this time. However, an indication of the molecular features that are responsible for the correlations can be obtained by examining the regression coefficients from the PLS models. The regression coefficients for the PLS models of MFA and MOE are shown in Fig. 7. Figure 7a shows the regression coefficients for the PLS models over a spectral range of 500 nm–2400 nm, while Fig. 7b shows the regression coefficients for the PLS models with a spectral range of 650 nm–1150 nm. It is clear from the results in Fig. 7 that there is a strong inverse relationship between the regression coefficients for MFA and MOE. This is expected since there is an inverse relationship between MFA and MOE, e.g., as MFA increases from 0 to 45 degrees the MOE of wood generally decreases. These regression coefficients also provide some information of the chemical features that drive the correlations between NIR spectra, and mechanical properties and MFA.

Assigning molecular features to NIR spectra is difficult due to the extensive overlap of the fundamental molecular vibrations in the NIR range. Some tentative assignments can be made to the chemical features associated with these

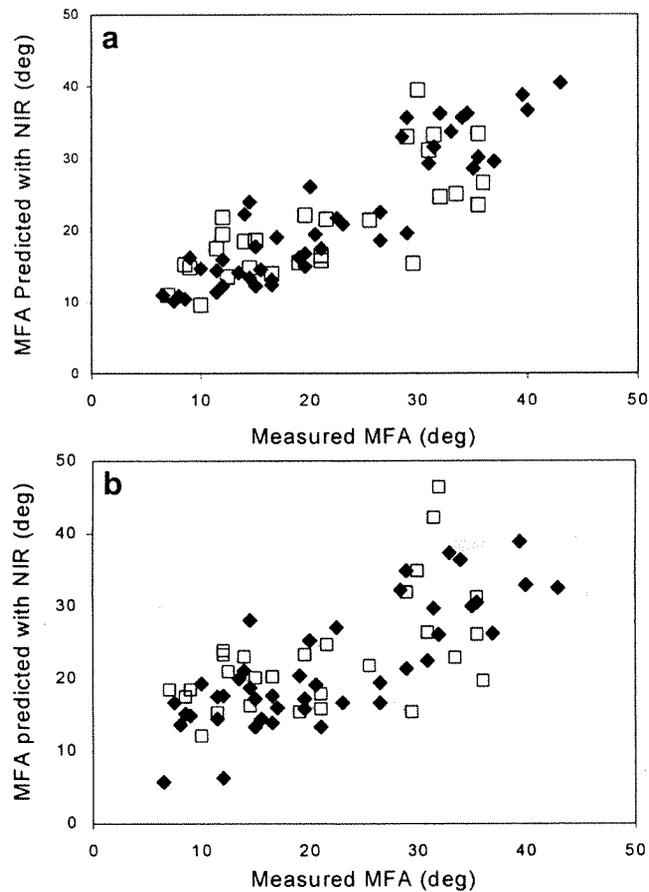
**Table 6** Effects of varying the number of principal components on the quality of the PLS-2 models for the mechanical properties and the PLS models for microfibril angle

	MOR (MPa)	MOE (MPa)	MFA (deg)
500 nm–2400 nm, 5 PCs used for models			
CALB set			
r (4 PC)	0.88	0.86	0.79
r (6 PC)	0.93	0.91	0.85
RMSEC (4 PC)	10.6	1,550	6.1
RMSEC (6 PC)	8.3	1,280	5.3
TEST set			
r (4 PC)	0.82	0.77	0.54
r (6 PC)	0.95	0.88	0.56
RMSEP (4 PC)	13.6	2,040	8.0
RMSEP (6 PC)	9.7	1,400	7.9
650 nm–1150 nm, 5 PCs used for models			
CALB set			
r (4 PC)	0.86	0.89	0.72
r (6 PC)	0.93	0.91	0.82
RMSEC (4 PC)	11.4	1,400	6.9
RMSEC (6 PC)	8.5	1,300	5.8
TEST set			
r (4 PC)	0.88	0.83	0.48
r (6 PC)	0.90	0.84	0.63
RMSEP (4 PC)	11.5	1,640	8.6
RMSEP (6 PC)	10.2	1,580	7.4

regression coefficients (Ali et al. 2001; Curran et al. 1992; Fourty et al. 1996; Marten et al. 1985). The peak between 1090 nm and 1210 nm can be assigned to the second overtone of cellulose hydroxyls. The importance of these second overtone vibrations is highlighted by the regression coefficients for the reduced spectral range shown in Fig. 7b. These second overtone vibrations are the major contributors to the correlations between NIR spectra, mechanical properties, and MFA. The peak between 1450 nm and 1630 nm is associated with the first overtone of cellulose and hemicellulose hydroxyls. The peak between 1630 nm and 1800 nm appears to be associated with the first overtone of aliphatic and aromatic CH vibrations. Finally, the peak between 1890 nm and 2000 nm is related to OH combination bands. These different vibrations are closely associated with the major peaks in the regression coefficients, showing that the molecular features of wood and the NIR predictions of wood strength and MFA are clearly related.

## Conclusions

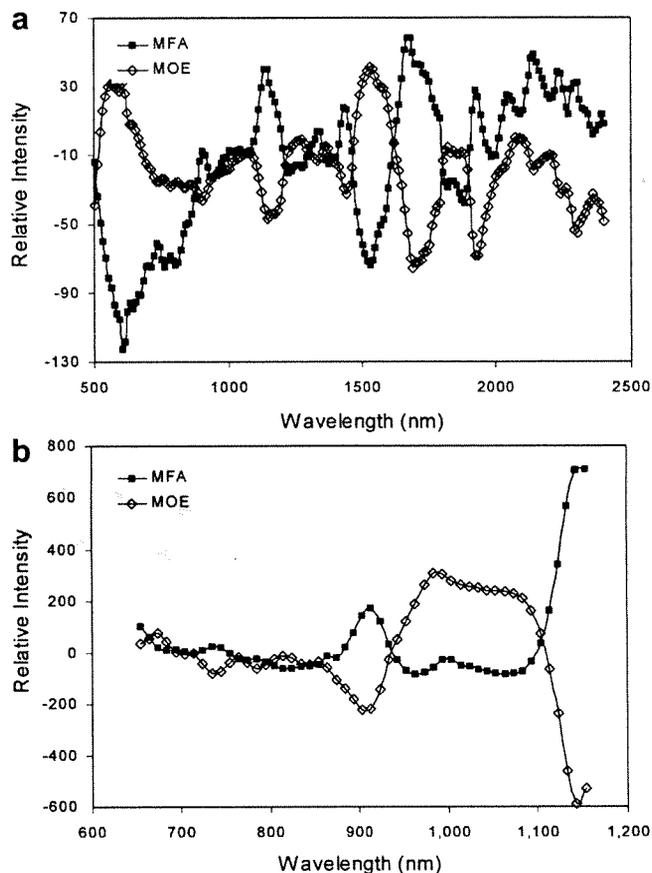
NIR spectra of solid wood samples were collected and related to the chemical and mechanical properties, and microfibril angle of the samples. NIR spectra of solid wood could be correlated with the chemical composition of wood. The correlation coefficients were generally above 0.80 for lignin, extractives, glucose, xylose, mannose, and galactose. The correlation coefficients remained high even when the spectral range was reduced from 500 nm–2400 nm to 650 nm–1150 nm. These same NIR spectra were also correlated with the mechanical properties of solid wood. The correlation coefficients between NIR spectra and



**Fig. 6a,b** The results of PLS models showing the correlation between measured microfibril angle and the microfibril angle predicted with NIR spectra collected over (a) the full spectral range (500 nm–2400 nm) and (b) a reduced spectral range (650 nm–1150 nm). Solid symbols are the samples included in the CALB set and the open symbols are the samples included in the TEST set

the mechanical properties of wood were generally above 0.85 and above 0.90 for some of the models. Again, both the full spectral range and the reduced spectral range could be used to predict the mechanical properties of wood. Finally, these same NIR spectra could be correlated with the microfibril angle of wood. This result is unexpected but helps explain the relationship between NIR spectra and mechanical properties.

The strength of the correlations between a reduced NIR spectral range, and the chemistry, mechanical properties, and microfibril angle is particularly significant since it enables the use for small, inexpensive, lightweight NIR spectrometers. These small spectrometers could be conveniently used in field applications or for process control applications. Foresters for a variety of applications, including genetic screening, timber management, harvesting and sorting decisions could use these NIR tools.



**Fig. 7a,b** Regression coefficients for PLS models predicting the MOE and MFA from NIR spectra over (a) the full spectral range (500 nm–2400 nm) and (b) a reduced spectral range (650 nm–1150 nm)

**Acknowledgements** Supported for this work from the United States Department of Energy Offices of Industrial Technologies Pulp and Paper Agenda 2020 Initiative, the United States Department of Agriculture National Research Initiative Wood Utilisation Program and the National Renewable Energy Laboratory Director's Discretionary Research and Development fund.

## References

- Aber J, Bolster K, Newman S, Soulia M, Martin M (1994) Analyses of forest foliage II: measurement of carbon fraction and nitrogen content by end-member analysis. *J. Near Infrared Spec* 2:15–23
- Ali M, Emsley A, Herman H, Heywood R (2001) Spectroscopic studies of the ageing of cellulose paper. *Polym* 42:2893–2900
- American Society for Testing Materials (1999) Standard test methods for evaluating properties of wood-base fiber and particle panel materials. ASTM D 1037–99. West Conshohocken, PA
- Axrup L, Markides K, Nilsson T (2000) Using miniature diode array NIR spectrometers for analyzing wood chips and bark samples in motion. *J Chemo* 14:561–572

- Berthold J, Olsson R, Salmen L (1998) Water sorption to hydroxyl and carboxylic acid groups in Carboxymethylcellulose (CMC) studies with NIR-spectroscopy. *Cellulose* 5:281–298
- Butterfield, B. (1997) Microfibril Angle in Wood, International Association of Wood Anatomist
- Couteaux M, McTernan K, Berg B, Szuberla D, Dardenne P, Bottner P (1997) Chemical composition and carbon mineralization potential of Scots Pine needles at different stages of decomposition. *Soil Biol Biochem* 30(5):583–595
- Curran P, Dungan J, Macler B, Plummer S, Peterson D (1992) Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. *Remote Sens Environ* 39:153–166
- Fern T (2000) On orthogonal signal correction. *Chemo Intell Lab Sys* 50:47–52
- Fourty T, Baret F, Jacquemoud S, Schmuck G, Verdebout J (1996) Leaf optical properties with explicit description of its biochemical composition: direct and inverse problems. *Remote Sens Environ* 56:104–117
- Greaves B, Schimleck L, Borralho N, Michell A, Raymond C (1996) Genetic control and repeatability from Eucalyptus nitrens woodmeal. *Appita J* 49(6):423–426
- Hiukka R, (1998) A multivariate approach to the analysis of pine needle samples using NIR. *Chemo Intell Lab Sys* 44:395–401
- Hoffmeyer P, Pedersen J (1995) Evaluation of density and strength of Norway spruce wood by near infrared reflectance spectroscopy. *Holz Roh Werkst* 53:165–170
- Huang C-L, Kutscha N, Leaf G, Megraw R (1997) Comparison of microfibril angle measurement techniques. In: Microfibril angle in wood, BG Butterfield (ed). *Int Assoc Wood Anatomists*, pp 177–183
- Malkavaara P, Alen R (1998) A spectroscopic method for determining lignin content of softwood and hardwood kraft pulps. *Chemo Intell Lab Sys* 44:287–292
- Marklund M, Hauksson J, Sjoström M (1999) Prediction of strength parameters for softwood kraft pulps: Multivariate data analysis based on orthogonal signal correction and near infrared spectroscopy. *Nord Pulp Pap Res J* 14(2):140–148
- Marten G, Shenk J, Barton III F (1985) Near infrared reflectance spectroscopy (NIRS): analysis of forage quality. USDA Agriculture Research Service, Handbook No 643
- Martens H, Naes T (1991) Multivariate calibration. Wiley, pp 419ff
- Martin M, Aber J (1994) Analyses of forest foliage III: determining nitrogen, lignin and cellulose in fresh leaves using near infrared reflectance data. *J Near Infrared Spec* 2:25–32
- McLellan T, Martin M, Aber J, Melillo J, Nadelhoffer K, Dewey B (1991a) Comparison of wet chemistry and near infrared reflectance measurements of carbon-fraction chemistry and nitrogen concentration of forest foliage. *Can J For Res* 21:1689–1693
- McLellan T, Aber J, Martin M, Melillo J, Nadelhoffer K (1991b) Determination of nitrogen, lignin, and cellulose content of decomposing leaf material by near infrared reflectance spectroscopy. *Can J For Res* 21:1684–1688
- Michell AJ (1995) Pulpwood quality estimation by near infrared spectroscopic measurements on eucalypt woods. *Appita J* 48(6):425–428
- Milne TA, Chum HL, Agblevor F, Johnson DK (1992) Standardized analytical methods. *Biomass Bioenergy* 2(1–6):341–366
- Newman S, Soulia M, Aber J, Dewey B, Ricca A (1994) Analyses of forest foliage I: laboratory procedures for proximate carbon fractionation and nitrogen determination. *J Near Infrared Spec* 2:5–14
- Raymond C, Schimleck L, Muneri A, Michell A (2001) Nondestructive sampling of Eucalyptus globules and E. nitens for wood properties. III. Predicted pulp yield using near infrared reflectance analysis. *Wood Sci Tech* 35(3):203–215
- Schimleck L, Michell A (1998) Determination of within-tree variation of kraft pulp yield using near-infrared spectroscopy. *Tappi J* 81(5):229–236
- Schimleck L, Michell A, Vinden P (1996) Eucalypt wood classification by NIR spectroscopy and principal component analysis. *Appita J* 49:319–324
- Schimleck L, Wright P, Michell A, Wallis A (1997) Near Infrared spectra and chemical compositions of E. globules and E nitens plantation woods. *Appita J* 50(1):40–46
- Svedas V (2000) Transformation of the near-infrared bands of cellulose surface hydroxyls under the influence of adsorbed water molecules. *Appl Spec* 54(3):420–425
- Svensson O, Josefson M, Langkilde F (1997) Classification of chemically modified cellulose using a near-infrared spectrometer and soft independent modeling of class analogies. *Appl Spec* 51(12):1826–1835
- Thumm A, Meder R (2001) Stiffness prediction of radiata pine clearwood test pieces using near infrared spectroscopy. *J Near Infrared Spec* 9:117–122

- Tsehaye A, Buchanan A, Meder R, Newman R, Walker J (1997) Microfibril angle: determining wood stiffness in radiate pine. In: Microfibril angle in wood, BG Butterfield (ed). Int Assoc Wood Anatomists, pp 323–336
- Tygesen L (1994) Determination of dry matter content basic density of Norway spruce by near infrared reflectance and transmittance spectroscopy. *J Near Infrared Spec* 2:127–135
- Vandeginste B, Massart D, Buydens L, De Jong S, Lewi P, Smeyers-Verbeke J (1998) Handbook of chemometrics and qualimetrics: Part B. Elsevier, New York, pp 349–379
- Wallbacks L, Edlund U, Lindgren T, Agnemo R (1995) Multivariate characterization of pulp, Part 3. Evaluation of physical and spectroscopic data on unbeaten and beaten kraft pulp samples. *Nord Pulp Pap Res J* 2:88–93
- Wold S, Antti H, Lindgren F, Ohman J (1998) Orthogonal signal correction of near-infrared spectra. *Chemo Intell Lab Sys* 44:175–185
- Wright J, Birkett M, Gambino M (1990) Prediction of pulp yield and cellulose content from wood using Near Infrared Reflectance spectroscopy. *Tappi J* 73:164–166