DIFFERENTIATION OF SOME CANADIAN CONIFEROUS WOODS BY COMBINED DIFFUSE AND SPECULAR REFLECTANCE FOURIER TRANSFORM INFRARED SPECTROMETRY[1]

Jason R. Nault[2]

Research Scientist
Forintek Canada Corp.
2665 East Mall
Vancouver, British Columbia, V6T 1W5

and

John F. Manville

Research Scientist
Pacific Forestry Centre
506 W. Burnside Rd.
Victoria, British Columbia, V8Z 1M5

(Received December 1991)

ABSTRACT

Infrared spectroscopy was used to differentiate coniferous woods commonly found in mixtures in lumber producing mills in British Columbia. The method required collection of reflectance Fourier transform infrared spectra of wood samples at a 2 cm^{-1} resolution. From a small subset of spectra, frequencies useful for species differentiation were selected using a combination of correlation analysis and principal component analysis. The selected frequencies were used to develop methods for differentiating species using discriminant analysis. These models were then tested against the remainder of the spectra. This approach was successfully used to classify the same wood samples in freeze-dried and green conditions, but was unsuccessful in classifying extractive free samples.

Keywords: FTIR, reflectance, infrared spectroscopy, species, conifers, classification.

INTRODUCTION

The lumber industry in Canada has expressed a need for rapid methods to sort freshly cut lumber by species. This would enable the industry to take advantage of unique properties of each species, either to increase the value of end products, or to solve problems associated with the processing or use of lumber of a particular species.

Typical species mixtures of western Canada are hem/fir (primarily composed of western hemlock (*Tsuga heterophylla* (Raf.) Sarg.), amabilis fir (*Abies amabilis* (Dougl.) Forbes) and Sitka spruce (*Picea sitchensis* (Bong.) Carr.)), SPF (primarily composed of white/Engelmann-interior spruce (*Picea* spp.), lodgepole pine (*Pinus contorta* Dougl.), and subalpine fir (*Abies lasiocarpa* (Hook.) Nutt.)), and D-fir/larch (Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco.) and western larch (*Larix occidentalis* Nutt.)).

Currently the most reliable method for identification of wood samples from various species is by examination of anatomical features on both a visual and microscopic scale (Strelis and Kennedy 1967).

[1] Taken in part from J. R. Nault's Doctor of Philosophy dissertation, Faculty of Graduate Studies (Department of Forestry), University of British Columbia, Vancouver, British Columbia, Canada, Dec. 1989.

[2] Present address: Pacific Forestry Centre, 506 W. Burnside Rd., Victoria, British Columbia, V8Z 1M5, Canada.

Wood and Fiber Science, 24(4), 1992, pp. 424–431

Another means of species identification of wood samples is by examination of their extractives. Extractives vary considerably in composition and concentration from species to species, tree to tree, heartwood to sapwood, from growth ring to growth ring, and from one type of wood tissue to another. Certain extractive chemicals can be species-specific, and much research has centered upon identifying them and their biological significance.

Species identification by means of extractives has taken several forms. One approach involves using reactions of unique extractives with indicator chemicals to form colored complexes, which can then be used to identify species (Barton 1973; Miller et al. 1985). Another approach is through gas chromatography either alone or in combination with spectroscopic methods (Manville and Tracey 1989; Swan 1966). Ion mobility spectrometry has also been used to study volatile compounds from the wood (Lawrence 1989).

Infrared (IR) spectroscopy has been used extensively to study wood. The majority of the work published on IR spectroscopy of wood or wood components has focused upon the pulp and paper industry, where the main uses are quantitative determination of lignin within the pulp (Berben et al. 1987; Faix 1988; Kolboe and Ellefsen 1962; Saad et al. 1980; Schultz and Glasser 1986), structural studies of lignin (Marton and Sparks 1967; Obst 1982; Sarkanen et al. 1967a; Sarkanen et al. 1967b; Sarkanen et al. 1967c; Schultz and Glasser 1986) and monitoring progress of the pulping process (Faix 1988; Gurnagul et al. 1986; Marton and Sparks 1967; Michell 1988; Michell et al. 1965; Schultz et al. 1985). Very little work has been published on wood extractive chemicals examined *in situ,* probably due to the extreme difficulties in analyzing small amounts of extractives within the complex chemical matrix of wood structural components.

Various sampling techniques have been used, including transmission (Chow 1972; Kolboe and Ellefsen 1962; Michell et al. 1965), multiple internal reflection (Marton and Sparks 1967), diffuse reflection (Owen and Thomas 1989), and photo-acoustic (Kuo et al. 1988) techniques.

The possible contributions of extractive chemicals to overall wood spectrum have been noted, but generally researchers have found such contributions to interfere with their studies (Chow 1972; Owen and Thomas 1989). No author has tried to use infrared spectra for species differentiation except in the most rudimentary fashion, for example to distinguish hardwoods vs. softwoods (Obst 1982; Owen and Thomas 1989; Wood 1988).

It has been hypothesized that differences in the composition of wood extractive chemicals amongst species could be detected by measuring the reflectance spectrum of wood using Fourier transform infrared spectrometry (FTIR). If measurable, such differences could form the basis for a rapid means of species identification. However, given the low concentrations of species-specific extractive chemicals (softwoods are composed of approximately 42% cellulose, 27% hemicelluloses, 28% lignin, and only 1 to 5% extractives on a dry basis), and their inclusion within the strongly absorbing matrix of wood tissue, very small differences were expected. Detection of these differences would thus require high-resolution spectra and a sophisticated algorithm for interpreting the spectra. Prediction of which areas of the spectrum would be useful was difficult because it was not known in advance which extractive compounds would be detected within the wood, nor which specific extractive chemicals would be of most use for species identification. Thus, it was necessary to analyze in minute detail a large portion of the wood IR spectrum; this presented a formidable task in data analyses.

METHODS

Samples (2 × 4 trim ends) from typical production mixtures of freshly cut green lumber (SPF, hem/fir, and D-fir/larch) were obtained from various areas of British Columbia. Identification of each individual sample was confirmed by microscopic examination of anatomical features.

Several thin (5-mm) slices were cut from the edge or face surfaces (selected at random) of each board using a bandsaw. No further preparation of the sample surface was done prior to scanning; our samples therefore varied in surface smoothness, as would be the case in a mill.

Initially, we prepared and scanned samples in conditions similar to those expected in an industrial setting. Thus, samples that were removed from the surface of boards were scanned at existing moisture contents in an open IR bench in which the ambient moisture and CO_2 levels would fluctuate. We also decided to concentrate on classification only within species groups as would commonly occur within a single mill.

Immediately prior to IR scanning, a disk of appropriate size for the reflectance accessory was cut from each sample. In samples where sapwood could be discerned, a separate sample was taken representing sapwood only. To prepare extractive-free samples, another disk was taken from the same slice as the original sample, as close to the first disk as possible. This disk was sequentially extracted in cyclohexane/ethanol, ethanol, and water, and then air-dried before scanning.

A Nicolet model 20SXB FTIR spectrometer equipped with a Spectra-Tech DRIFT cell was used for collection of all spectra. The resulting spectra were a combination of specular and diffuse reflectance. Each sample was scanned 100 times from 4,850 to 400 cm^{-1} at a resolution of 2 cm^{-1} (time to collect = 208 sec).

All samples were scanned at room temperature in an open sample compartment, first green and then after freeze-drying for 24 hours. No attempt was made to orient the samples with regard to grain or growth rings. The open sample compartment was to approximate mill conditions; it would be impractical in a mill to keep conditions constant, and to exclude water and carbon dioxide, which is normally done for FTIR scanning. In total, 740 samples were scanned while green, 738 after freeze-drying, and 264 after extraction.

Each sample spectrum was ratioed against the reflectance spectrum of a KBr pellet run on the same day. A KBr pellet was used to mimic wood surfaces, which were also solid and not powdered as is normally the case for DRIFT experiments. This corrected each spectrum for daily background absorption originating from carbon dioxide and water vapor in the air and daily variations in source intensity and detector sensitivity.

The resulting spectra consisted of 4,615 data points each, representing ratioed signal strengths at evenly spaced points from 4,850 to 400 cm^{-1}, with each interval being 0.964 cm^{-1} (Fig. 1).

RESULTS AND DISCUSSION

This study concentrated on the spectral region from 2,500 to 400 cm^{-1} because it is in this region that the most characteristic "fingerprints" of the wood extractive chemicals are found, and interference from the highly variable moisture content (in the green samples) is minimized in this region. Visual comparison of individual spectra showed that differences existed amongst samples, but these visual differences could not be used to separate species. The recently published work of Anderson et al. (1991a and 1991b), contains infrared spectra of wood of different species. At first glance, these spectra appear different. However, it is possible to find spectra that show obvious visual differences amongst samples from different species, but such differences become obscured when spectra from multiple samples from one species are examined. The visual differences between two samples of one species can be greater than the visual differences between samples from two different species, making visual examination of infrared spectra unreliable for the identification of wood as to species.

This indicated that a multivariate statistical method would be useful in identifying species to species variations as opposed to sample to sample variations.

Starting with the freeze-dried spruce SPF samples (to eliminate variations from moisture content) the region from 2,500 to 400 cm^{-1}

from each ratioed spectrum was analyzed in the following way:

1. Each of the 308 reflectance spectra was centered about its mean (baseline correction). Figure 2 shows sample spectra for spruce, pine and fir.
2. A subset of spectra for each species in its species group was randomly chosen to yield a training set (3 species × 20 samples/species = 60 samples), with the remainder of the spectra used as a test set (248 samples).
3. A Fisher weight was calculated according to Sharaf et al. (1986) for each variable (wavelength).
4. Next, the correlation matrix for the training set was computed. Any variable that was very highly correlated ($r \geq 0.9995$) with the preceding wavelength was eliminated. We hoped that by removing variables that were essentially the same, some degree of colinearity between variables would be removed, and the number of variables would be reduced. At the same time, it was recognized that the spectral differences sought were probably small, and that removal of such variables could remove useful information. This procedure resulted in an abbreviated spectrum of 700 variables for each sample.
5. The calculated Fisher weight was applied to each variable in each abbreviated training set spectrum.
6. The Fisher weighted training set was analyzed using the SAS routine PRINCOMP (SAS 1982).
7. The relative importance of the principal components within each species group was determined by their eigenvalues (Fig. 3). Note that the first five principal components account for virtually all of the variation found in the SPF species group. Thus only the first five principal components were deemed useful and were used for differentiation.
8. From each of the first five principal components, the six variables with the highest loadings (eigenvector value) were selected to perform a discriminant analysis on the training sets. These points are indicated on a typical spectrum in Fig. 4. A discriminant function was calculated for each training set using the SAS routine DISCRIM (SAS 1982), which calculates a linear discriminant function based upon a measure of generalized squared distance between groups.
9. This function was then used to classify all of the samples in both the training set and the test set by species. The results of this discriminant classification are presented in Table 1(a).
10. A discriminant function was then calculated and tested for the spectra from the green samples using the same samples in

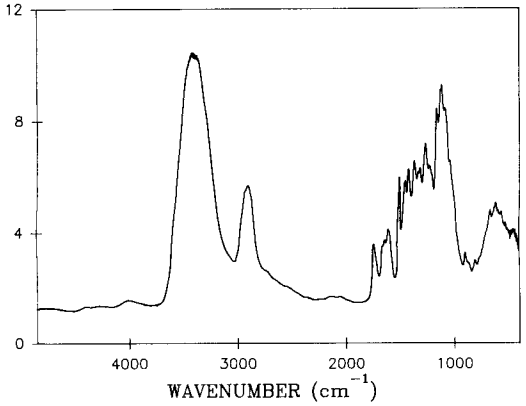

FIG. 1. Typical wood reflectance spectrum.

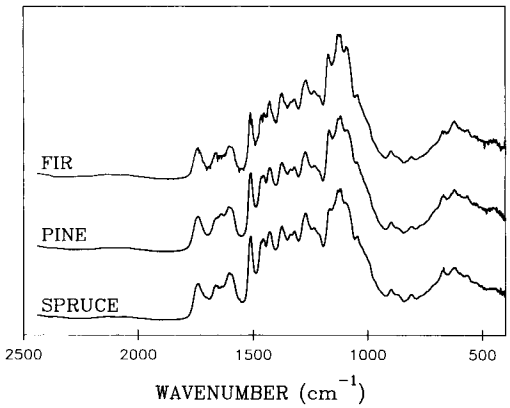

FIG. 2. Typical reflectance spectra.

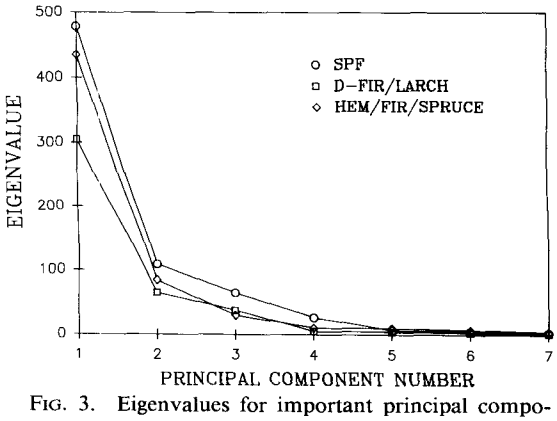

FIG. 3. Eigenvalues for important principal components.

the training set and test set and the same variables (Table 1(b)).

11. The two variables with the lowest weight in each principal component were then dropped (leaving 4 variables × 5 principal components = 20 variables) and the discriminant classification was run again (Table 1(c)). Finally, the next two variables with the lowest weight in each principal component were dropped (leaving 2 variables × 5 principal components = 10 variables), and the discriminant classification was run again (Table 1(d)). Further dropping of variables seriously degraded the classification efficiency. (For a description of principal components analysis and discriminant analysis, see Hope 1968; Johnson and Wichern 1982; Klecka 1980; Malinowski and Howery 1980; Massart et al. 1988; Sharaf et al. 1986).

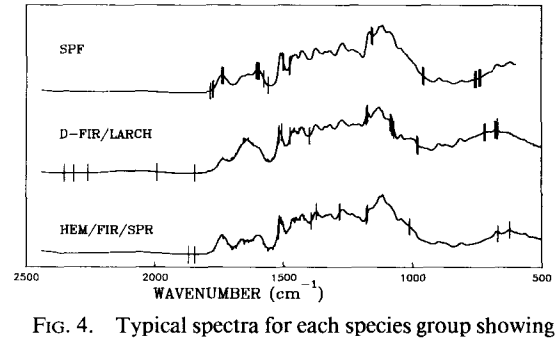

FIG. 4. Typical spectra for each species group showing frequencies used for classification.

To demonstrate the fact that the sorting algorithm relied upon extractive chemicals, a matched set of 264 extractive-free SPF samples were scanned and the spectra were processed by the identical algorithm. Ninety-nine percent of the extracted samples were classified as lodgepole pine. The inability to correctly classify spruce or fir samples confirmed that the presence of extractive chemicals was a prerequisite for the application of this technique.

This series of operations (steps 1 to 10 above) was then applied to the 192 freeze-dried Douglas-fir/western larch samples using 20 samples in the training set (2 species × 10 samples/species). In this case the correlation analysis left 931 variables remaining in the spectra. The principal components analysis showed only four principal components of any importance (Fig. 3), so the 10 variables with the highest weights were taken from each of these principal components (10 variables × 4 principal components = 40 variables) as illustrated in Fig. 4.

Using the 20 sample training set, a discriminant function was calculated and applied to the remaining 172 sample test set. The SAS DISCRIM procedure output for this trial reported a warning that the covariance matrix was not full rank and that only the first 18 variables had been used for the discriminant function. The not full rank warning indicates that some variables are highly correlated and thus redundant. These variables were dropped from the analysis. As shown in Table 2(a), by using only the first 18 variables, 188 of 192 samples were correctly identified as to species, or 98%.

Discriminant analysis using the same 40 variables from the green samples resulted in only 17 variables being used (Table 2(b)), with fewer correctly classified samples (81%). To see if this could be improved upon using other variables, three DISCRIM analyses were performed, utilizing the first 14 variables, the next 14 variables, and the final 12 variables (Table 2 (c), (d) and (e), respectively), with the best classification coming from the last 12 variables (91% correct).

TABLE 1. *Percentage of samples of fir, pine and spruce correctly classified by discriminant analysis.*

Species group	Condition of sample	Variables used	Species			
			Fir	Pine	Spruce	All
			%			
(a) SPF	(freeze-dried)	30	58	90	73	76
(b) SPF	(green)	30	83	78	66	77
(c) SPF	(green)	20	89	76	74	80
(d) SPF	(green)	10	78	93	68	83
(e) SPF	(extractive free)	30	3	99	0	44*

* (99% of all samples were classified as pine, so virtually all pine pieces were classified as pine, while virtually all spruce and fir pieces were misclassified, as pine.)

The same series of operations (steps 1 to 10 above) was next applied to the 166 freeze-dried western hemlock/amabilis fir/Sitka spruce (HFS) mixture using 30 samples in the training set (3 species × 10 samples/species). In this case, the correlation analysis left 1,085 variables remaining in the spectra. The principal components analysis showed only five principal components of any importance (Fig. 3), so the four variables with the highest weights were taken from each of these principal components (4 variables × 5 principal components = 20 variables), as illustrated in Fig. 4.

The SAS (DISCRIM procedure) output for this trial again reported a warning that the covariance matrix was not full rank and that one variable had been dropped from the discriminant function. As shown in Table 3(a), by using 19 variables, 138 of 166 samples (83%) were correctly identified as to species.

The same method was then applied to the green samples. Again the rank warning occurred, this time dropping five variables and retaining 15 variables. The results were poor; 68% of the samples were correctly classified (Table 3(b)). The discriminant analysis was then attempted using only the variables with the highest and lowest loadings for the first five principal components, resulting in 10 variables (five principal components with two variables each). The results were somewhat poorer than for 20 variables, with only 63% of all samples being correctly classified, representing a decrease in effectiveness for each species (Table 3(c)).

In an attempt to develop a sorting method applicable to separating only western hemlock and Sitka spruce, all amabilis fir samples were dropped from the data and classification was attempted again. Using the same 20 variables as selected for the sort of all three species (with the same five dropped by SAS), the results were promising; 82% of all samples were classified correctly as were 88% of Sitka spruce samples, and 78% of western hemlock samples (Table 3(d)).

When only 10 variables were used (five principal components with two variables each), the results were almost the same (Table 3(e)) as for 20 variables; 82% of samples were correctly

TABLE 2. *Percentage of samples of Douglas-fir and larch correctly classified by discriminant analysis.*

Species group	Condition of sample	Variables		Species		
		Selected	Used	D-fir	Larch	All
				%		
(a) DFL	dry	40	18	88	100	98
(b) DFL	wet	40	17	59	86	81
(c) DFL	wet	14	14	88	90	90
(d) DFL	wet	14	14	68	88	84
(e) DFL	wet	12	12	85	92	91

TABLE 3. *Percentage of fir, hemlock and spruce correctly classified by discriminant analysis.*

Species group	Condition of sample	Variables		Species			
		Selected	Used	Fir	Spruce	Hemlock	All
				%			
(a) HFS	dry	20	19	81	84	83	83
(b) HFS	wet	20	15	63	69	67	68
(c) HFS	wet	10	10	56	60	65	63
(d) HS	wet	20	15	—	88	78	82
(e) HS	wet	10	10	—	83	82	82

classified. Sitka spruce was slightly poorer (83%) and western hemlock slightly better (82%) than before.

To conclude, it has been demonstrated that the combination of Fisher weighting, principal components analysis, and discriminant analysis with a linear discriminant function can be used to sort hem/fir, SPF, and D-fir/larch with a degree of success.

This demonstrates the validity of the original hypothesis that the reflectance FTIR spectra of wood samples contain the information needed to determine species of samples within the common industrial species groups. Further, we have shown that the classification criteria are based upon the presence of extractable chemicals. This technique has proven successful for both heartwood and sapwood, and for green and freeze-dried samples.

The techniques presented in this paper are biased towards using the fewest possible number of frequencies, mainly for the purpose of designing a simple identification procedure which would be of use to the wood industry. Further investigation of other techniques that make use of different statistical techniques or a different subset of variables could yield more accurate sorting algorithms.

ACKNOWLEDGMENTS

The authors wish to acknowledge those Forintek member companies which generously supplied samples of lumber and J. Gonzalez of Forintek for the meticulous microscopic species identifications of each piece used in this study.

REFERENCES

ANDERSON, E. L., Z. PAWLAK, NOEL L. OWEN, AND WILLIAM C. FEIST. 1991a. Infrared studies of wood weathering. Part I: Softwoods. Applied Spectroscopy 45:641–647.

———, ———, ———, AND ———. 1991b. Infrared studies of wood weathering. Part II: Hardwoods. Applied Spectroscopy 45:648–652.

BARTON, G. M. 1973. Chemical color tests for Canadian woods. Canadian Forest Ind. 93(2):57–60,62.

BERBEN, S. A., J. P. RADEMACHER, O. S. LOWELL, AND D. B. EASTY. 1987. Estimation of lignin in wood pulp by diffuse reflectance Fourier transform infrared spectrometry. Tappi 70(11):129–133.

CHOW, S. Z. 1972. Infrared spectral study of woody tissues from four conifers. Wood Sci. 5(1):27–33.

FAIX, O. 1988. Practical uses of FTIR spectroscopy in wood science and technology. Mikrochimica Acta 1988(I):21–25.

GURNAGUL, N., F. G. T. ST-GERMAIN, AND D. G. GRAY. 1986. Photoacoustic Fourier transform infrared measurements on paper. J Pulp Paper Sci. 12(5):J156–J159.

HOPE, K. 1968. Methods of multivariate analysis. University of London Press Ltd.

JOHNSON, R. A., AND D. W. WICHERN. 1982. Applied multivariate statistical analysis. Prentice-Hall, Englewood Cliffs, N.J.

KLECKA, W. R. 1980. Discriminant analysis. Sage Publications, New York NY.

KOLBOE, S., AND O. ELLEFSEN. 1962. Infrared investigations of lignin: A discussion of some recent results. Tappi 45(2):163–166.

KUO, M., J. F. MCCLELLAND, S. LUO, P. CHIEN, AND R. D. WALKER. 1988. Applications of infrared photoacoustic spectroscopy for wood samples. Wood Fiber Sci. 20(1):132–145.

LAWRENCE, A. H. 1989. Rapid characterization of wood species by ion mobility spectrometry. J Pulp Paper Sci. 15(5):J196–J199.

MALINOWSKI, E. R., AND D. G. HOWERY. 1980. Factor analysis in chemistry. John Wiley & Sons, New York NY.

MANVILLE, J. F., AND A. S. TRACEY. 1989. Chemical differences between alpine firs of British Columbia. Phytochemistry 28(10):2681–2686.

MARTON, J., AND H. E. SPARKS. 1967. Determination of lignin in pulp and paper by infrared multiple internal reflectance. Tappi 50(7):363–368.

MASSART, D. L., B. G. M. VANDEGINSTE, S. N. DEMING, Y. MICHOTTE, AND L. KAUFMAN. 1988. Chemometrics: A textbook. Elsevier Science Publishers, New York, NY.

MICHELL, A. J. 1988. Usefulness of Fourier transform infrared difference spectroscopy for studying the reactions of wood during pulping. Cellulose Chem. Technol. 22:105–113.

———, A. J. WATSON, AND H. G. HIGGINS. 1965. An infrared spectroscopic study of delignification of Eucalyptus regnans. Tappi 48(9):520–532.

MILLER, R. B., J. T. QUIRK, AND D. J. CHRISTENSEN. 1985. Identifying white oak logs with sodium nitrite. Forest Prod. J. 35(2):33–38.

OBST, J. R. 1982. Guaiacyl and syringyl lignin composition in hardwood cell components. Holzforschung 36: 143–152.

OWEN, N. L., AND D. W. THOMAS. 1989. Infrared studies of "hard" and "soft" woods. Appl. Spectroscopy 43(3): 451–455.

SAAD, S. M., R. M. ISSA, AND S. FAHMY. 1980. Infrared spectroscopic study of bagasse and unbleached high-yield soda bagasse pulps. Holzforschung 34:218–222.

SARKANEN, K. V., H. CHANG, AND B. ERICSSON. 1967a. Species variation in lignins. I. Infrared spectra of guaiacyl and syringyl models. Tappi 50(11):572–575.

———, ———, AND G. G. ALLAN. 1967b. Species variation in lignins. II. Conifer lignins. Tappi 50(12):583–587.

———, ———, AND ———. 1967c. Special variation in lignins. III. Hardwood lignins. Tappi 50(12):587–590.

SAS INSTITUTE. 1982. SAS user's guide: Statistics. SAS Institute Inc., Cary, NC.

SCHULTZ, T. P., AND W. G. GLASSER. 1986. Quantitative structural analysis of lignin by diffuse reflectance Fourier transform infrared spectrometry. Holzforschung 40(Supp):37–44.

———, M. C. TEMPLETON, AND G. D. MCGINNIS. 1985. Rapid determination of lignocellulose by diffuse reflectance Fourier transform infrared spectrometry. Anal. Chem. 57:2867–2869.

SHARAF, M. A., D. L. ILLMAN, AND B. R. KOWALSKI. 1986. Chemometrics. John Wiley and Sons, Inc., New York, NY.

STRELIS, L., AND R. W. KENNEDY. 1967. Identification of North American commercial pulpwoods and pulp fibers. University of Toronto Press, Toronto.

SWAN, E. P. 1966. Chemical methods of differentiating the wood of several western conifers. Forest Prod. J. 16(1):51–54.

WOOD, D. J. 1988. Characterisation of charcoals by DRIFT. Mikrochimica Acta 1988, II:167–169.