APPLICATIONS FOR NIR SPECTROSCOPY IN EUCALYPTUS GENETIC IMPROVEMENT PROGRAMS AND PULP MILL OPERATIONS

by

JUSTIN AUGUST TYSON

(Under the Direction of Laurence R. Schimleck)

ABSTRACT

We tested the applicability of NIR-based wood property calibrations for eucalyptus plantations in central Brazil to trees in southern Brazil. Prediction errors for the southern samples were too high. As we added southern samples to the calibrations, prediction statistics for most wood properties improved considerably. We also collected increment cores and drill shavings from trees in central Brazil and collected NIR spectra from both. The aforementioned calibrations were used to predict the wood properties of the trees, and milled shavings were found to provide biased wood property estimates. Next, we collected bleached eucalyptus handsheets from a pulp mill in Espírito Santo. Mechanical and physical properties of the handsheets were measured and NIR calibrations were created. Calibration and cross-validation statistics were poor for all properties. Finally, we collected unbleached eucalyptus pulp samples, measured their carbohydrates, and created NIR calibrations. Calibration and prediction statistics were excellent for both xylose and rhamnose, but poorer for other carbohydrates.

INDEX WORDS: near-infrared spectroscopy, eucalyptus, genetic improvement, lignin, pentosans, density, pulp yield, pulp property, carbohydrate

APPLICATIONS FOR NIR SPECTROSCOPY IN EUCALYPTUS GENETIC IMPROVEMENTS PROGRAMS AND PULP MILL OPERATIONS

by

JUSTIN AUGUST TYSON

BS, University of Georgia, 2005

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2007

APPLICATIONS FOR NIR SPECTROSCOPY IN EUCALYPTUS GENETIC IMPROVEMENT PROGRAMS AND PULP MILL OPERATIONS

by

JUSTIN AUGUST TYSON

Major Professor: Laurence R. Schimleck

Committee: Richard F. Daniels
Lawrence A. Morris

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2007

DEDICATION

This thesis is dedicated to my parents, Tony and Ramona, who have offered their unwavering support for me during the completion of this document and throughout my life.

ACKNOWLEDGEMENTS

I would like to thank Aracruz Celulose S.A. for funding this research and Aurélio Aguiar, Júpiter Abad, Gabriel Rezende, and Otávio Filho for their support and suggestions. I thank the Warnell School of Forestry and Natural Resources for their assistantship funding. I would also like to express my sincere gratitude to Drs. Laurie Schimleck, Dick Daniels, and Larry Morris for their patient guidance. Finally, I thank my parents, Tony and Ramona, for their wisdom and provision.

TABLE OF CONTENTS

Page

Chapter 1

INTRODUCTION

Aracruz Celulose S.A., a Brazilian forest products company, is the world's leading producer of bleached eucalyptus pulp. They have recently begun using near-infrared (NIR) spectroscopy as a tool for the evaluation of wood properties in their eucalyptus genetic improvement program. Schimleck et al. (2006) previously showed that spectra collected from milled eucalyptus increment cores could be used to estimate several important wood properties, particularly lignin and pentosan content. The samples for this study were collected from the central Brazilian states of Espírito Santo, Bahia, and Minas Gerais. However, Aracruz recently acquired land in the southern state of Rio Grande do Sul, and would like to extend their genetic improvement program to this region. It would be costly and time-consuming to destructively sample the number of trees required to create a new calibration set.

Aracruz's current wood property calibrations are based on milled increment cores. They are considering collecting samples for future analysis using a standard drill bit, as this is a cheaper method. However, an unpublished study (Schimleck and Michell 1999) showed that milled drill bit shavings yielded different NIR spectra than milled increment cores, which led to different predictions of pulp yield. Aracruz would like to see additional research in this area before changing its current approach.

Aracruz are also interested in potential applications of NIR spectroscopy in their pulp mill operations. They currently measure many physical and mechanical traits of bleached pulp handsheets, such as tensile stiffness, air resistance, stretch, and drainability, using labor-intensive laboratory tests. Previous research has indicated that

many of these traits can be accurately estimated using NIR spectroscopy, which requires much less time and equipment (Fardim et al. 2002; Meder et al. 1994; Wallbacks et al. 1991). However, all of these previous studies applied different pulping methods to their samples to introduce variability into their data, which does not reflect the reality of Aracruz's tightly controlled mill operations.

The chemical composition of pulp has a significant effect on the physical and mechanical properties of the pulp. Chemical properties are typically difficult to measure using traditional techniques, but NIR spectroscopy has proven to be useful in this field as well (Aracruz 2006). Aracruz would like to investigate the possibility of using NIR spectroscopy to quantify the composition of various carbohydrates in unbleached pulps, but previous studies have used a variety of sample preparation techniques with varying degrees of success.

Based on the results of prior research and Aracruz's desire to further incorporate NIR spectroscopy into its eucalyptus improvement programs and pulp mill operations, we have developed several objectives for this thesis:

1) Determine if existing calibrations for central Brazil can be used to estimate the wood properties of samples from plantations located in southern Brazil and (if this if found to be unsuccessful);

2) Determine if it is feasible to adjust Aracruz's current calibrations by adding a subset of samples from the southern population to the current calibration set;

3) Determine if milled drill bit shavings yield significantly different wood property estimates than milled increment cores and (assuming differences exist);

4) Determine if the rankings of wood property estimates are similar for drill bit shavings and increment cores;

5) Determine if accurate calibrations for important physical and mechanical characteristics of eucalyptus pulps can be created using NIR spectra from handsheets of mill-line origin;

6) Determine the best method of sample preparation for collecting NIR spectra for chemical analysis of unbleached eucalyptus pulp and;

7) Determine which of the major carbohydrates in eucalyptus pulp may be effectively estimated using NIR spectroscopy.

LITERATURE CITED

Aracruz. 2006. Aracruz Annual Sustainability Report.

Fardim, P., Ferreira, M.M.C., and Duran, N. 2002. Multivariate calibration for quantitative analysis of eucalypt kraft pulp by NIR spectrometry. J. Wood Chem. Technol. 22(1): 67-81.

Meder, R., Gallagher, S., Kimberley, M.O., Uprichard, J.M., and Mackie, K.L. 1994. Prediction of wood chip and pulp and paper properties via multivariate analysis of spectral data. Appita J.: 479-484.

Schimleck, L.R., and Michell, A.J. 1999. Comparison of milled wood from increment cores and drill shavings. CSIRO Divisional Report.

Schimleck, L.R., Rezende, G.D.S.P., Demuner, B.J., and Downes, G.M. 2006. Estimation of whole-tree wood quality traits using near infrared spectra from increment cores. Appita J. 59(3): 231-236.

Wallbacks, L., Edlund, U., and Norden, B. 1991. Multivariate characterization of pulp, Part 1. Spectroscopic characterization of physical and chemical differences between pulps using C CP/MAS NMR, FT-IR, NIR and multivariate data analysis. Nord. Pulp Paper Res. J. 2: 74-80.

Chapter 2

LITERATURE REVIEW

ARACRUZ CELULOSE

Aracruz Celulose S.A., a Brazilian company, is the world's leading producer of bleached eucalyptus pulp, which is used primarily to manufacture printing, writing, and tissue papers. The three million metric tons the company produces annually account for 27% of the global supply of the product (Aracruz 2006). The majority of the pulp produced is exported to Europe (39%), North America (34%) and Asia (25%). The production is distributed among three pulp-making facilities: Barra do Riacho in Espírito Santo (2.1 million tons); Guaíba in Rio Grande do Sul (430,000 tons); and Veracel in Bahia (450,000 tons). The Veracel manufacturing facility is currently operating at only half of its total capacity (Aracruz 2006).

Aracruz Celulose is a vertically integrated forest products company, as they do not simply manufacture pulp and paper; they also own and manage the forests which supply their mills. Aracruz owns approximately 279,000 hectares of eucalyptus plantations and 154,000 hectares of native forest reserves in Brazil. These forests are distributed among the states of Espírito Santo, Bahia, Minas Gerais, and Rio Grande do Sul (Campinhos and Claudio-da-Silva 1990). As a result of their broad spectrum of involvement in the management and utilization of eucalyptus plantations, the company is involved in numerous fields of research, including genetic improvement, wood quality analysis, and pulp quality analysis.

Aracruz has been developing an extensive *Eucalyptus* genetic improvement program since 1973. The process of introducing genetic material, testing, evaluation, and

selection of superior trees within the best families has been repeated for more than three decades, with impressive results. Significant increases in productivity were achieved when Aracruz began to utilize cuttings for the establishment of clonal plantations (Bertolucci et al. 1995). A decrease in wood specific consumption (the volume of wood required to produce one ton of pulp) from 4.9 to 4.1 m^3/ton and an increase in pulp productivity from 5.9 to 10.9 tons/ha/yr were obtained through this program (Bertolucci et al. 1995). Since Aracruz began using cuttings to establish clonal plantations, most of their efforts have been focused on hybrids of *Eucalyptus grandis* Hill ex Maiden × *Eucalyptus urophylla* S.T. Blake, as these hybrids not only have fast growth, excellent form, and good pulping properties; they have also been found to have particularly good rooting ability. Aracruz is also interested in developing *Eucalyptus globulus* Labill. × *E. grandis* and *E. globulus* × *E. urophylla* hybrids to introduce the excellent growth and pulping properties of *E. globulus*, but researchers have found it difficult to root cuttings of these hybrids.

Recently, researchers at Aracruz have begun to consider more wood quality traits, such as basic density, lignin content, and pentosan content, in their genetic improvement program (Bertolucci et al. 1995). These wood quality traits are important for several reasons. Trees with high basic density and low lignin content generally produce the greatest pulp yield, which is defined as the mass of dry pulp produced from a given mass of wood and is expressed as a percentage. Low lignin content is particularly beneficial during chemical and semi-chemical pulping, as the chemicals used to remove the lignin (thereby allowing the separation of fibers) also degrade the cellulose and hemicelluloses (Rhydholm 1965). Wood with lower lignin content can be reduced to a target lignin

content more quickly, thereby reducing carbohydrate degradation and increasing pulp yield. Pentosans (hemicelluloses with a five-carbon ring structure) are important because they contribute to inter-fiber bonding in paper products, improving their tear strength, fracture toughness, and folding endurance (Molin and Teder 2002). High pentosan content, therefore, is another desirable wood quality trait.

NIR SPECTROSCOPY

Unfortunately, traditional methods of wood quality analysis are expensive and laborious, and, therefore, have not been greatly used in genetic improvement programs. Furthermore, the destructive sampling techniques traditionally used, though they may yield accurate measurements, render the sample trees unavailable for breeding purposes. In the last two decades, significant advances have been made in the use of near-infrared (NIR) spectroscopy for the analysis of wood properties. NIR spectroscopic wood quality analyses are typically much cheaper and quicker than traditional analyses, because little sample preparation is required.

The NIR spectrum ranges from 780 to 2500 nm. Absorption bands and overtones within this region are associated with the fundamental vibrations of specific molecular bonds, such as C-H, N-H, and O-H bonds. Although the signals from these vibrations are similar, with many broad and overlapping bands, the application of multivariate analysis to NIR spectra has reduced many of the problems associated with these overlapping bands. One multivariate technique that is widely used for predictive purposes is partial least squares regression. Partial least squares regression can be used to correlate slight changes in NIR spectra with independently measured material properties (Martens and

Naes 1991). It describes the underlying latent structure in the NIR spectra and response variables to generate a calibration that can be used for predictive purposes.

Although the combination of NIR spectroscopy and multivariate analysis has been used in the food industry since the 1960s (Gera and Norris 1968; Norris and Hart 1965), it was not until decades later that this technology was applied to the forest products industry. Birkett and Gambino (1988) were among the first to demonstrate the potential of NIR spectroscopy to rapidly estimate the chemical constituents of wood. Several more studies soon confirmed NIR analysis to be a useful tool in estimating cellulose, hemicellulose, and lignin content, as well as pulp yield (a trait that is strongly correlated with the chemical composition of wood) (Easty et al. 1990; Garbutt et al. 1992; Schultz and Burns 1990; Wright et al. 1990).

Because the effect of NIR radiation on specific chemical bonds is well understood, it is not surprising that the earliest uses of this technology in the forest products industry involved the estimation of chemical characteristics of wood. However, researchers soon began to investigate the utility of NIR spectroscopy for the estimation of a wide variety of physical and mechanical parameters as well. Thygesen (1994) showed that NIR could be used to estimate wood density, while Hoffmeyer and Pedersen (1995) showed that NIR could be used to measure density, compression strength, and bending strength of dry wood. More recently, it has proven to be useful in the prediction of tracheid morphological characteristics, including microfibril angle, and fiber coarseness, and wall thickness (Schimleck and Evans 2002, 2004).

As the range of wood quality traits that have been studied using NIR technology has increased, so, too, has the range of wood sample collection and preparation

techniques. Many of the early studies utilized milled whole-tree composite samples for analyses - i.e., a sample tree was felled, bolts or discs were collected from predetermined heights and then chipped, and a subsample of chips was milled into a fine powder that was used for NIR analysis. Hence, for a tree to be analyzed using NIR spectroscopy for a genetic improvement program, the tree was destroyed to collect a sample. Several recent papers have shown that a single, milled increment core can be used in place of a whole-tree composite to estimate whole-tree wood properties, particularly lignin and pentosan content (Schimleck et al. 2005a; Schimleck et al. 2005b; Schimleck et al. 2006). Thus, use of NIR spectroscopy to estimate wood properties in genetic improvement programs only requires that the trees used to create the original calibration be destroyed.

APPLICATIONS FOR NIR SPECTROSCOPY AT ARACRUZ

Aracruz Celulose began using NIR spectroscopy to analyze wood properties for their genetic improvement program around 2000. They have developed and are currently using NIR calibrations for the prediction of several wood quality traits, including basic density, lignin content, pentosan content, and pulp yield in their plantations in the central Brazilian states of Espírito Santo, Bahia, and Minas Gerais (Schimleck et al. 2005c). These calibrations are based on the use of a milled increment core, so additional sample trees are nondestructively sampled. Recently, Aracruz purchased a large amount of timberlands in the southern Brazilian state of Rio Grande do Sul. Creating an entirely new set of calibrations for their southern plantations is undesirable, as it would require a large number of trees to be destructively sampled. Two recent studies have demonstrated that the wood properties of samples from areas that are not represented in a given calibration set may be accurately measured if a small number of samples from the new

area are added to the calibration (Jones et al. 2005; Schimleck et al. 2005b). This may be an effective method of producing robust calibrations, but the optimal number of samples to collect from a new area to be added to a calibration is not yet known.

Although Aracruz's current NIR calibrations require the collection of an increment core from a standing tree, a more efficient method of sample collection would be to obtain shavings from a standard drill bit. Therefore, they are considering the use of drill bit shavings in place of increment cores in their NIR program. However, it is unlikely that drill bit shavings will yield similar wood property predictions to increment cores when used with Aracruz's current models. Schimleck and Michell (1999) have showed in an unpublished study that microscopic morphological differences existed between milled increment cores and milled drill bit shavings from Tasmanian *Eucalyptus globulus*. These discrepancies resulted in different NIR spectra for the same sample tree, and when these spectra were then used to predict pulp yield using a calibration that was developed using milled increment cores, the different spectra resulted in different predictions of pulp yield. However, it is not known if the spectral changes would significantly affect the predictions of different wood properties, particularly chemical properties such as lignin and pentosan content.

Aracruz Celulose is also interested in potential applications of NIR spectroscopy for the analysis of pulp and paper quality. This includes the estimation of a wide variety of properties, including the following: chemical properties, such as Kappa and carbohydrate content; physical properties, such as bulk, specific volume, and surface area; and mechanical properties, such as air resistance, drainability, stretch, tensile index, and tensile stiffness. Although relatively few studies have investigated the use of NIR

spectroscopy for pulp and paper quality estimation when compared with the number of studies involving wood quality estimation, there have been promising results to indicate that numerous pulp and paper properties may be effectively estimated using this technology. Several papers in the early 1990s demonstrated the potential for this NIR spectroscopy to predict a variety of important physical, mechanical, and chemical pulp properties using spectra collected from prepared handsheets (Edlund et al. 1991; Wallbacks 1993; Wallbacks et al. 1991).

Most of the research involving the use of NIR spectroscopy to predict pulp properties has been conducted in Sweden and utilized softwood pulps (Antti et al. 1996; Edlund et al. 1991; Marklund et al. 1999; Meder et al. 1994; Wallbacks 1993; Wallbacks et al. 1995; Wallbacks et al. 1991). Recent studies have used similar techniques to investigate eucalyptus pulps. Fardim et al. (2002) used NIR spectra from unbleached *E. grandis* pulp handsheets to create good calibrations for several chemical and physical properties, while Fardim et al. (2005) used NIR spectra from bleached *E. grandis* pulp handsheets to create calibrations with good predictive abilities for ten pulp properties, including three properties in which Aracruz has interest—tensile index, drainability, and stretch. However, these studies—like many other studies in this field (Edlund et al. 1991; Wallbacks 1993; Wallbacks et al. 1995; Wallbacks et al. 1991)—used different pulping and beating regimes for their samples to introduce variability into their data set. A high amount of variability in the calibration set improves calibration statistics, but does not reflect the reality of a commercial pulping operation, where the pulping process is tightly regulated and the variability of the final product is minimized. No studies that involve NIR analysis of samples of mill line origin have been reported.

One of the difficulties with NIR analysis within all fields of the forest products industry is the comparison of multiple calibrations for the estimation of a given property. A comparison of two or more calibrations is necessary to determine if, for example, one particular method of sample collection or preparation yields a more accurate model than an alternative method. Almost invariably, the researcher will compare certain statistics, such as coefficients of variation (R^2), standard errors, and ratios of performance to deviation (RPD), then make a subjective judgment to rank the competing models. This method, however, ignores important pieces of information, such as sample size, and may be misleading. A more objective approach would be to use a statistical significance test to determine which models are significantly different from each other.

LITERATURE CITED

Antti, H., Sjostrom, M., and Wallbacks, L. 1996. Multivariate calibration models using NIR spectroscopy on pulp and paper industrial applications. J. Chemom. 10: 591-603.

Aracruz. 2006. Aracruz Annual Sustainability Report.

Bertolucci, F.L.G., Demuner, B.J., Garcia, S.L.R., and Ikemore, Y.K. 1995. Increasing fiber yield and quality at Aracruz. *In* Eucalypt Plantations: Improving Fibre Yield and Quality, Hobart, Tasmania.

Birkett, M.D., and Gambino, M.J.T. 1988. Potential applications for near-infrared spectroscopy in the pulping industry. Pap. S. Afr. November/December: 34-38

Campinhos, E., and Claudio-da-Silva, E. 1990. Development of the eucalyptus tree of the future. *In* ESPRA Spring Conference, Seville.

Easty, D.B., Berben, S.A., DeThomas, F.A., and Brimmer, P.J. 1990. Near-infrared spectroscopy for the analysis of wood pulp; quantifying hardwood-softwood mixtures and estimating lignin content. Tappi J. 73(10): 257-261.

Edlund, U., Wallbacks, L., Norden, B., Lindgren, T., and Iversen, T. 1991. Pulp characterization and predictions of properties using spectroscopy and multivariate data analysis methods. *In* 6th International Symposium on Wood and Pulping Chemistry.

Fardim, P., Ferreira, M.M.C., and Duran, N. 2002. Multivariate calibration for quantitative analysis of eucalypt kraft pulp by NIR spectrometry. J. Wood Chem. Technol. 22(1): 67-81.

Garbutt, D.C.F., Donkin, M.J., and Meyer, J.H. 1992. Near-infrared reflectance analysis of cellulose and lignin in wood. Pap. S. Afr. April: 45-48.

Ben-Gera, I., and Norris, K.H. 1968. Determination of moisture content in soybeans by direct spectrophotometry. Israel J. of Agr. Res. 18(3): 125-132.

Hoffemeyer, P., and Pedersen, J. 1995. Evaluation of density and strength of Norway spruce wood by near infrared spectroscopy. Holz Roh Werkst. 53: 165-170.

Jones, P.D., Schimleck, L.R., Peter, G.F., Daniels, R.F., and Clark, A.I. 2005. Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. Can. J. For. Res. 35: 85-92.

Marklund, A., Paper, M., Hauksson, J.B., and Sjostrom, M. 1999. Prediction of strength parameters for softwood kraft pulps. Multivariate data analysis based on orthogonal signal correction and near infrared spectroscopy. Nord. Pulp Paper Res. J. 14: 140-148

Martens, H., and Naes, T. 1991. Multivariate Calibration. John Wiley & Sons, Chichester, UK.

Meder, R., Gallagher, S., Kimberley, M.O., Uprichard, J.M., and Mackie, K.L. 1994. Prediction of wood chip and pulp and paper properties via multivariate analysis of spectral data. Appita J. 47(1): 479-484.

Molin, U., and Teder, A. 2002. Importance of cellulose/hemicellulose ratio for pulp strength. Nord. Pulp Paper Res. J. 17(1): 14-19.

Norris, K.H. and J.R. Hart, 1965. Direct spectrophotometric determination of moisture content of grain and seeds. *In*: Principles and Methods of Measuring Moisture Content in Liquids and Solids. Vol. 4. Reinhold, New York: 19-25.

Rhydholm, S.A. 1965. The evaluation of paper permanence and durability. Tappi J. 59(12): 63-69.

Schimleck, L.R., Downes, G.M., and Evans, R. 2005a. Estimation of *Eucalyptus nitens* wood properties by near-infrared spectroscopy. Appita J. 59(2): 136-141.

Schimleck, L.R., and Evans, R. 2002. Estimation of microfibril angle of increment cores by near infrared spectroscopy. IAWA J. 23: 225-234.

Schimleck, L.R., and Evans, R. 2004. Estimation of *Pinus radiata* D. Don tracheid morphological characteristics by near infrared spectroscopy. Holzforshung 58: 66-73.

Schimleck, L.R., Kube, P.D., Raymond, C.A., Michell, A.J., and French, J. 2005b. Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. Can. J. For. Res. 35: 2797-2805.

Schimleck, L.R., and Michell, A.J. 1999. Comparison of milled wood from increment cores and drill shavings. CSIRO Divisional Report.

Schimleck, L.R., Rezende, G.D.S.P., Demuner, B.J., and Downes, G.M. 2006. Estimation of whole-tree wood quality traits using near infrared spectra from increment cores. Appita J. 59(3): 231-236.

Schultz, T.P., and Burns, D.A. 1990. Rapid secondary analysis of lignocellulose: comparison of near-infrared and fourier transform infrared (FTIR). Tappi J. 73(5): 209-212.

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

Wallbäcks, L. 1993. Characterisation of chemical pulps using spectroscopy, fibre dimensions and multivariate data analysis. CTAPI, 7th International Symposium on Wood and Pulping Chemistry, Beijing, China. Proc., 2, 802-809

Wallbäcks, L., Edlund, U., and Lindgren, T. 1995. Multivariate characterization of pulp, Part 3. Evaluation of physical and spectroscopic data on beaten and unbeaten kraft pulp samples. Nord. Pulp Paper Res. J. 10(2): 88-93

Wallbacks, L., Edlund, U., Norden, B., and Burgland, I. 1991. Multivariate characterization of pulp using solid-state ^{13}C NMR, FTIR, and NIR. Tappi J. 74(10): 201-206.

Wright, J.A., Birkett, M.D., and Gambino, M.J.T. 1990. Prediction of pulp yield and cellulose content from wood samples using near-infrared reflectance spectroscopy. Tappi J. 73(8).

Chapter 3

ADJUSTING NEAR INFRARED WOOD PROPERTY CALIBRATIONS FOR CENTRAL BRAZIL TO PREDICT THE WOOD PROPERTIES OF SAMPLES FROM SOUTHERN BRAZIL[1]

[1] Tyson, Justin A., Laurence A. Schimleck, Aurélio M. Aguiar, Jupiter I. Muro Abad, and Gabriel D. S. P. Rezende. Submitted to Appita Journal, 9/26/2007

ABSTRACT

Recent research has shown that accurate whole-tree wood property predictions can be made using calibrations based on near-infrared (NIR) spectra from milled increment cores. In this study, we demonstrate that whole-tree wood property calibrations for lignin and pentosan content, developed with a variety of eucalyptus species and hybrids in central Brazil, may be applied to a separate population of eucalypts in southern Brazil. We randomly selected 2 to 10 samples from southern Brazil and added these to the original calibration set. The ratios of performance to deviation (RPD) for lignin content, pentosans, and pulp yield were significantly improved using this method, but density predictions were unimproved. WinISI software was also used to select the most spectrally representative samples from southern Brazil for addition to the calibration set, but improvements were similar to the randomly selected samples. The addition of four southern samples to the original calibrations did not significantly affect their predictive performance on samples from central Brazil.

INDEX WORDS: near-infrared spectroscopy, eucalyptus, lignin content, pentosans, basic density, pulp yield

INTRODUCTION

During the last decade, the use of near-infrared (NIR) spectroscopy to provide rapid estimates of wood properties has steadily increased. Several studies in the late 1980s and early 1990s indicated that NIR spectroscopy had excellent potential to rapidly estimate a variety of properties related to wood chemistry, including pulp yield and cellulose, hemicellulose, and lignin content (Birkett and Gambino 1988; Easty et al. 1990; Garbutt et al. 1992; Schultz and Burns 1990; Wright et al. 1990). More recently, it has proven to be useful in the prediction of many physical characteristics of wood, such as density, microfibril angle, and fiber coarseness and wall thickness (Schimleck and Evans 2002, 2003, 2004).

Due to the diversity of wood properties that can be predicted using NIR technology, its potential uses in the forest products industry are extensive. Tree breeders, in particular, are beginning to realize the potential of this technology for selecting trees of superior wood quality. Conventional methods of measuring wood quality are relatively expensive and time-consuming, and many require the destruction of the potential breeding candidates. Hence, breeders have traditionally focused on traits that are more easily recognized, such as rapid growth, good form, and drought and disease resistance. NIR spectroscopy offers a method of quickly identifying individuals with superior wood quality traits.

One drawback to this technology is that much of the research in wood property prediction has used whole-tree composites to collect spectra – i.e., a sample tree is felled and chipped, and a subsample of chips is milled to provide the spectra. Obviously, destroying potential breeding candidates is not a desirable outcome of a tree improvement

program. However, several recent papers have shown that a single, milled increment core can be used in place of a whole-tree composite to accurately estimate whole-tree wood properties, particularly lignin and pentosan content (Schimleck et al. 2005a; Schimleck et al. 2005b; Schimleck et al. 2005c). Thus use of NIR technology to estimate wood properties in a tree breeding program only requires that the trees used to create the original calibration be destroyed.

A drawback NIR technology is that the calibrations can only be used with confidence to predict the wood quality of trees from the same general location of the trees in the calibration set. If NIR-based calibrations are used to predict the wood quality of trees of the same species, but from a different location, differences in site conditions may cause the predictions to be biased and either over- or underestimated, although rankings may be less affected (Schimleck et al. 2005b; Schimleck et al. 2000). A resolution to this problem that has only recently been tested is to measure the wood properties and NIR spectra of a small subsample of trees from the new location and add this subsample to the original calibration set.

Jones et al. (2005) created NIR calibrations for density, microfibril angle, and stiffness using loblolly pine (*Pinus taeda* L.) increment cores representing 15 different sites from 3 physiographic regions of Georgia, USA. After dividing the sites into calibration (9 sites) and prediction (6 sites) sets, they found that the calibration statistics – in particular, ratio of performance to deviation (RPD) values – were much better than the prediction statistics. By adding spectra from one core from each site in the prediction set to the calibration set, the subsequent prediction statistics were significantly improved. Schimleck et al. (2005b) used a similar approach for improving *E. nitens* pulp yield

predictions in Tasmania. When they applied a Tasmania-wide calibration to a site that was not represented in the calibration set, the RPD values were very poor. They added 5 samples from the new site to the calibration set, and the prediction statistics were greatly improved. However, neither of these studies investigated the number of samples that needed to be added to the calibration set for optimal performance.

Aracruz Celulose S.A., the world's largest supplier of bleached eucalyptus pulp, recently began to use NIR spectroscopy to screen eucalyptus clones in their tree improvement program for pulping properties. Preliminary work was based on 98 eucalyptus clones representing 7 species and 5 hybrids among these species (Schimleck et al. 2006). The sampled trees were taken from 3 locations (Aracruz, Bahia and São Mateus) near the coast of central Brazil. The 98 samples were split into calibration (60 samples) and prediction (38 samples) sets. The strongest calibrations were obtained for lignin and pentosan content with R^2 of 0.79 and 0.82, respectively, using NIR spectra from milled whole-tree chips, and 0.77 and 0.83, respectively, using NIR spectra from milled 5 mm increment cores collected at a height of 0.65 m.

Aracruz has recently acquired large landholdings in Rio Grande do Sul, located in southern Brazil. Their new landholdings total 106,000 ha, including about 30,000 ha of native forest reserves and 67,000 ha of eucalyptus plantations. The plantations were primarily pastures before Aracruz bought the land, and the major species are now *Eucalyptus dunni* Maiden, *Eucalyptus saligna* Sm., *E. grandis*, and a variety of hybrids. Aracruz plans to utilize NIR spectroscopy for tree improvement purposes in these southern plantations, but developing a completely new set of calibrations for the southern region would require significant expense. Therefore, the objectives of this study are to:

1) determine if existing calibrations, based on samples from Aracruz, Bahia and São Mateus, can be used to estimate the wood properties of samples from plantations located in southern Brazil; 2) determine if it is feasible to adjust Aracruz's current calibrations by adding a subset of samples from the southern population to the current calibration set; 3) determine the optimal number of samples to add to the calibration set to maximize the predictive capabilities of the calibrations while minimizing cost; 4) examine the possibility of using WinISI software as an effective method of selecting samples from the southern population for addition to existing calibrations; and 5) determine if adding samples from the southern population to the existing calibrations significantly affects the prediction of wood properties of trees from the central population.

MATERIALS AND METHODS

Sample origin

Five mm increment cores were removed from 55 trees at a sampling height of 0.65 m from two locations in Rio Grande do Sul—Horto Barbara Negra (51°14'W and 30°26'S) and Colorado (51°51'W and 30°05'S). The trees were then felled and chipped to give whole-tree composites. The trees were chosen to represent a wide range of genetic variability and included six different species and many different hybrids among these species (Table 3.1). Previously, 140 trees were sampled using the same protocols from three locations farther north—Aracruz (40°07'W, 19°42'S), Bahia (39°36'W, 17°52'S) and São Mateus (39°48'W, 18°40'S). Summary statistics for 98 of these species are presented by Schimleck et al. (2005c). Two samples that they discarded due to large residuals are included in this study, as well as 40 additional samples from

Aracruz that were collected at a later date. A comparison of the central and southern population statistics is presented in Table 3.2.

Sample preparation

All of the increment cores and chips were air-dried, and the increment cores were milled in a Cyclone mill. The whole-tree chips were used to determine several wood-quality traits using published standards and methods when possible. These traits include basic density (Tappi standard T258 om-94), screened pulp yield (target Kappa number 18 ± 1), lignin content (Tappi standard T222 om-98), and pentosan content (Tappi standard T223 om-84).

NIR spectroscopy

NIR spectra were measured in diffuse reflectance mode from samples held in a spinning sample holder in a NIRSystems Inc.™ Model 5000 scanning spectrophotometer. The spectra were collected at 2 nm intervals over the wavelength range 1100-2500 nm. The instrument reference was a ceramic standard. Fifty scans were accumulated for each sample, and the results were averaged. After the spectrum was obtained, the sample cup was emptied, repacked, and a duplicate spectrum was obtained. The duplicate spectra were averaged using Vision® software (version 3.1). The data were then imported into the Unscrambler® software (version 9.2) and converted to the second derivative using left and right gaps of 8 nm.

Development of wood property calibrations

Calibrations were made for four properties—density, lignin content, pentosan content, and pulp yield. These calibrations were developed using Partial Least Squares (PLS) regression with four cross-validation segments and a maximum of 10 factors using

the Unscrambler software. Calibrations were developed for each property using the central dataset only and the central dataset plus 2, 4, 6 and 10 samples selected at random from the southern test set. The calibrations were then tested on a prediction set of the remaining 45 samples from the south (i.e., none of the prediction samples were included in the calibration sets). Because the samples in the prediction set remained the same while testing each calibration, meaningful comparisons among the resulting prediction statistics could easily be made.

Additionally, the use of WinISI® software was investigated as a potential tool for improving calibrations by identifying the most spectrally representative samples from the southern set for addition to the calibrations. WinISI uses a neighborhood concept to identify spectrally unique samples. A neighborhood is defined as the space near a sample, and for calibration purposes, only one sample is required per neighborhood. We identified the 4 samples in the southern population with the most neighbors, i.e. the most representative, added these to the central data set, and created new calibrations for each property. We then tested these calibrations on a southern test set of 42 samples. Three samples were omitted from the test set for these predictions to avoid duplicating samples from the calibration set in the prediction set.

Finally, after we determined the optimal number of samples from the southern population to add to the calibration set, we investigated if the adjusted calibrations yielded significantly different predictions of wood properties in the central population. To do this, the central sample set was randomly divided into a new calibration set of 100 samples and a prediction set of 40 samples. The performance of the new calibrations on the 40 samples was calculated. We then added 4 samples (which we determined to be the

optimal number) from the southern set to the new calibration set, and recalculated the performance of the calibrations on the 40 samples.

Statistical analyses

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

The calibrations were used to predict the properties of the various prediction sets. The standard error of prediction (SEP) was used as a measure of how well a calibration predicts the parameter of interest for a set of samples that are different from the calibration set. The predictive ability of calibrations was assessed from the R_p^2 and the ratio of performance to deviation (RPD_p) (ratio of the standard deviation of the reference data to the SEP).

We observed that the optimum number of factors recommended for each wood property calibration was different as the number of samples in the calibration set changed. To facilitate a valid comparison among calibrations and predictions developed using different calibration sets, the number of factors used to develop the calibrations was held to the mode number of recommended factors for each property. The number of factors for pulp yield was particularly variable, as four factors were recommended when

the central set only was used, but nine factors were recommended for all calibrations that included southern samples.

RESULTS

Calibrations for the four wood properties of interest—basic density, lignin content, pentosan content, and pulp yield—were created using the existing set of 140 samples from the central population. Calibrations for density, lignin content, and pentosan content were good (Table 3.3); all had R^2 values greater than 0.75 and RPD_c values greater than 1.7. The calibration for pulp yield was the weakest, with an R^2 value of 0.63 and an RPD_c value of only 1.26.

These calibrations were used to predict the properties of the 45 samples from the southern test set. The lignin and pentosan calibrations provided the strongest predictions, but none of the predictions resulted in an RPD_p value that exceeded 1.5, which is considered sufficient for screening (Williams and Sobering 1993) (Table 3.4).

With the aim of improving the accuracy of the predictions, a subset of samples from the southern population was added to the central set. To determine the optimal number samples to add to the calibration, a random selection of 2, 4, 6, and 10 samples from the southern population was added to the central set, and four new calibrations were created for each wood property. We observed that, if the number of factors used to create the calibrations was held constant for each property, there was very little difference in the calibration statistics (Figs. 3.1-3.3). We used each of the new calibrations to predict the wood properties of the 45 samples in the southern test set. With the exception of density, there was very little change in the R_p^2 values as the southern samples were added to the calibrations (Fig. 3.4). However, the RPD_p values for lignin content, pentosan content,

and pulp yield all increased substantially (Fig. 3.5). The RPD_p value for lignin content increased from 1.19 to 1.68 after only 2 samples were added to the calibration set, while the RPD_p value for pentosan content increased from 1.31 to 1.90 after 6 samples were added. The RPD_p value for pulp yield rose from only 0.49 to 1.04 after 10 southern samples were added to the calibration set, but the adjusted calibrations yielded no improvement in the RPD_p values for basic density.

These statistics demonstrate that the calibrations developed for the central population can effectively predict the lignin and pentosan content of trees from the southern population if 2 to 6 samples randomly selected from the South are included in the calibration set. However, it does not appear that pulp yield or basic density calibrations can be adequately adjusted in the same manner.

To determine if WinISI software can be used to select samples more efficiently for addition to the calibration set, we used it to find the 4 samples that were most representative of the southern population based on their NIR spectra. Once selected, the 4 samples were added to the calibration set, and new calibrations were created. When these calibrations were used to predict the properties of the southern test set (minus the 3 samples that were also included in the calibration set), we observed little change in the prediction statistics as compared to the predictions produced by the calibrations that included 4 randomly selected samples (Table 3.5). Although the RPD_p values for basic density, pentosan content, and pulp yield were all marginally improved by using WinISI software, the RPD_p value for lignin content decreased slightly.

Finally, we tested the adjusted models on a set of samples from the central data set to determine if the addition of the southern samples significantly changed their

predictive capability in the central population. We found that the addition of 4 southern samples had very little influence on prediction statistics, demonstrating that as the central calibrations are adjusted for use on the southern population, the models are still fully applicable to the central population (Table 3.6).

DISCUSSION

Accurate predictions of wood properties of trees in the southern test set could not be made using calibrations based on samples from the central population. However, by adding a small number of samples from the South to the calibrations, the prediction statistics for lignin and pentosan content and were greatly improved. Although the R_p^2 values changed little, the RPD_p values showed considerable improvements. The RPD_p value for lignin content increased from 1.19 to 1.68 after only 2 samples were added to the calibration set, and the RPD_p value for pentosan content improved from 1.31 to 1.90 after the addition of 6 samples to the calibration set. These statistics indicate that, although these calibrations are not suitable for making accurate predictions of these properties, they are useful for ranking purposes (Williams and Sobering 1993). They compare favorably to the statistics described by Schimleck et al. (2006), who reported RPD_p values for lignin and pentosan content ranging from 1.63 to 2.03.

Although the RPD_p value for pulp yield also displayed marked improvements as southern samples were added to the calibration set, the predictions were still too poor to be useful. This is hardly surprising, as the calibration for pulp yield was also poor, with an RPD_c value of less than 1.3. Although several studies have reported better pulp yield predictions (Michell 1995; Olsson et al. 1995; Wright et al. 1990), few have used increment core samples to make whole-tree predictions. Schimleck et al. (2005b) used

core samples to make whole-tree pulp yield predictions and reported better statistics, but their study investigated only one species. It is possible that our poor results are partially caused by the high genetic variability of the data sets. Another likely cause is that the Kappa numbers of the pulp yield data used in this study varied considerably, from 15.9 to 19.0, meaning that the pulp yield data is not directly comparable between samples. Determination of pulp yield at a consistent Kappa number could greatly improve calibration statistics and the performance of the calibrations when applied to a separate test set.

Predictions for basic density were also poor, despite the reasonable calibration statistics, and they were not improved by adding samples from the southern population the central population basic density calibration. Schimleck et al. (1999) and Schimleck et al. (2006) also investigated the use of milled eucalyptus cores for the prediction of whole-tree basic density, and similarly found that the predictive errors were large. However, recent studies have found that radial strips taken from increment cores may be used to obtain much more accurate basic density predictions (Jones et al. 2005; Schimleck and Evans 2003; Schimleck et al. 2001; Schimleck et al. 2003) This is because these studies used solid wood, and the spectra collected were directly related to the density measurements from each section of wood, not to the whole tree.

We found that adding the 4 most spectrally representative samples from the southern test set to the calibration set, selected with WinISI software, resulted in very little change to the predictive capabilities of the calibrations as compared to a random selection of 4 samples. A possible explanation may be that the central and southern data

sets are so spectrally dissimilar that the addition of a random selection is just as beneficial as the addition of the most representative samples.

Although there has been extensive research into the use of NIR spectroscopy to estimate wood properties related to pulp production, a limited number of studies have examined the use of milled increment cores for estimating whole-tree properties of eucalypts (Raymond and Schimleck 2002; Raymond et al. 2001; Schimleck et al. 2005b; Schimleck et al. 1999; Schimleck et al. 2006). Even fewer have used more than one eucalyptus species to create calibrations (Schimleck et al. 2006). Our research supports the concept that useful whole-tree wood property estimates may be made using increment cores, which is quite important for tree improvement programs. Core sampling is not harmful to the sampled trees, and much less effort and expense is required to obtain the samples. This study also demonstrates that successful calibrations can be created using many different species and hybrids of eucalypts. Many forest products companies plant a variety of species and hybrids to effectively manage the different climates and soil types across their landholdings. Clearly, a single calibration that can be used effectively for trees with widely varying genetics is much cheaper to create and easier to use than a different calibration for each species or hybrid.

Finally, this research demonstrates that NIR-based calibrations can be adjusted for use in different regions simply by adding a small number of samples from the new location to the calibration set. This approach is much cheaper and easier than creating a completely new calibration for a new region, and it is more accurate than using an unmodified calibration from a different region. Because we have also demonstrated that the adjusted calibrations remain fully applicable in the original location, models that may

be used in many regions and for many different species and hybrids can be obtained using this method.

CONCLUSIONS

We tested the applicability of existing NIR-based eucalyptus wood property calibrations from central Brazil to a separate population in southern Brazil and found they could not be used due to large predictive errors. When 2 to 6 samples from the southern population were included in the calibration set, the prediction statistics for lignin and pentosan content were greatly improved. Pulp yield prediction statistics were also improved using this approach, but errors were too large for practical purposes. Prediction statistics for density were not improved. Adding more than 6 samples from the south was unnecessary and did not result in significant improvements in prediction statistics for any property.

When WinISI software was used to select the most spectrally representative samples for addition to the calibration set, prediction statistics were similar to those obtained using a random selection of samples. We also found that the addition of a small number of random samples from the southern test set to the calibrations for the central population did not diminish the predictive capability of the calibrations when applied to a subset of the central population. These findings indicate that NIR-based wood property calibrations can be made more applicable to populations in different regions by adding a small number of samples from the new regions to the original calibration set.

LITERATURE CITED

Birkett, M.D., and Gambino, M.J.T. 1988. Potential applications for near-infrared spectroscopy in the pulping industry. Pap. S. Afr. 34(November/December).

Easty, D.B., Berben, S.A., DeThomas, F.A., and Brimmer, P.J. 1990. Near-infrared spectroscopy for the analysis of wood pulp; quantifying hardwood-softwood mixtures and estimating lignin content. Tappi J. 73(10): 257.

Garbutt, D.C.F., Donkin, M.J., and Meyer, J.H. 1992. Near-infrared reflectance analysis of cellulose and lignin in wood. Pap. S. Afr. April: 45-48.

Jones, P.D., Schimleck, L.R., Peter, G.F., Daniels, R.F., and Clark, A. 2005. Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. Can. J. Forest Res. 35: 85-92.

Michell, A.J. 1995. Pulpwood quality estimation by near-infrared spectroscopic measurements on eucalypt woods. Appita J. 48(66): 425-428.

Olsson, R.J.O., Tomani, P., Karlsson, K., Joseffson, T., Sjöberg, K., and Björklund, C. 1995. Multivariate characterisation of chemical and physical descriptors in pulp using NIRR. Tappi J. 78(10): 158.

Raymond, C.A., and Schimleck, L.R. 2002. Development of near infrared reflectance analysis calibrations for estimating genetic parameters for cellulose content in *Eucaluptus globulus*. Can. J. For. Res. 32: 170-176.

Raymond, C.A., Schimleck, L.R., Muneri, A., and Michell, A.J. 2001. Nondestructive sampling of *Eucalyptus globulus* and *E. nitens* for wood properties. III. Predicted pulp yield using Near Infrared Reflectance Analysis. Wood Sci. Technol. 35: 203-215.

Schimleck, L.R., Downes, G.M., and Evans, R. 2005a. Estimation of *Eucalyptus nitens* wood properties by near-infrared spectroscopy. Appita J. 59(2): 136-141.

Schimleck, L.R., and Evans, R. 2002. Estimation of microfibril angle of increment cores by near infrared spectroscopy. IAWA J. 23: 225-234.

Schimleck, L.R., and Evans, R. 2003. Estimation of air-dry density of increment cores by near-infrared spectroscopy. Appita J. 56: 312-317.

Schimleck, L.R., and Evans, R. 2004. Estimation of *Pinus radiata* D. Don tracheid morphological characteristics by near infrared spectroscopy. Holzforshung 58: 66-73.

Schimleck, L.R., Evans, R., and J., I. 2001. Estimation of *Eucalyptus delegatensis* wood properties by near infrared spectroscopy. Can. J. Forest Res. 31: 1671-1675.

Schimleck, L.R., Kube, P.D., Raymond, C.A., Michell, A.J., and French, J. 2005b. Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. Can. J. For. Res. 35: 2797-2805.

Schimleck, L.R., Michell, A.J., Raymond, C.A., and Muneri, A. 1999. Estimation of basic density of *Eucalyptus globulus* using near-infrared spectroscopy. Can. J. For. Res. 29: 194-201.

Schimleck, L.R., Mora, C., and Daniels, R.F. 2003. Estimation of physical wood properties of green *Pinus taeda* radial samples by near infrared spectroscopy. Can. J. For. Res. 33: 2297-2305.

Schimleck, L.R., Raymond, C.A., Beadle, C.L., Downes, G.M., Kube, P.D., and French, J. 2000. Application of NIR spectroscopy to forest research. Appita J. 53: 458-464.

Schimleck, L.R., Rezende, G.D.S.P., Demuner, B.J., and Downes, G.M. 2005c. Estimation of whole-tree wood quality traits using near infrared spectra from increment cores. Appita J. 59(3): 231-236.

Schimleck, L.R., Rezende, G.D.S.P., Demuner, B.J., and Downes, G.M. 2006. Estimation of whole-tree wood quality traits using near infrared spectra from increment cores. Appita J. 59(3): 231-236.

Schultz, T.P., and Burns, D.A. 1990. Rapid secondary analysis of lignocellulose: comparison of near-infrared and fourier transform infrared (FTIR). Tappi J. 73(5): 209.

Tappi standard T222 om-98. 2004. Acid insoluble lignin in wood and pulps. *Edited by* Tappi revised T222 om-02. TAPPI PRESS, Canada (CD-ROM).

Tappi standard T223 om-84. 2004. Pentosans in wood and pulp. *Edited by* Tappi reaffirmed T223 cm-01. TAPPI PRESS, Canada (CD-ROM).

Tappi standard T258 om-94. 2004. Basic density and moisture content of pulp wood. *In* Tappi revised T258 om-02,. TAPPI PRESS, Canada (CD-ROM).

Williams, P.C., and Sobering, D.C. 1993. Comparison of commercial near infrared transmittance and reflectance instruments for the analysis of whole grains and seeds. J. Near Infrared Spectrosc. 1: 25-33.

Wright, J.A., Birkett, M.D., and Gambino, M.J.T. 1990. Prediction of pulp yield and cellulose content from wood samples using near-infrared reflectance spectroscopy. Tappi J. 73(8).

Table 3.1. Summary statistics of lab-measured properties of basic density, lignin content, pentosans and pulp yield, including the average and standard deviation (St.Dev) and the minimum and maximum observations for each species sampled and the hybrids. The 55 sample trees were located in Horto Barba Negra, RS, Brazil and Colorado, RS, Brazil.

Species	No. of Samples	Statistic	Basic Density (kg/m^3)	Lignin content (%)	Pentosans (%)	Pulp Yield (%)
E. dunnii	5	**Average**	434	29	19.4	51.9
		St.Dev.	33.8	1.36	0.9	1.32
		Minimum	406	27.7	17.9	50.7
		Maximum	479	30.2	20	53.5
E. globulus	5	**Average**	509	28.8	19	52.4
		St.Dev.	42.7	0.13	0.73	0.51
		Minimum	440	28.6	18	51.7
		Maximum	555	29	19.9	52.9
E. grandis	5	**Average**	400	29.8	17	52.2
		St.Dev.	20.8	0.46	0.62	0.85
		Minimum	379	29.3	16.4	51.1
		Maximum	428	30.3	17.7	53.4
E. maidenii	5	**Average**	516	28.8	19.3	52.5
		St.Dev.	30	0.86	0.36	1.27
		Minimum	482	27.5	18.8	51.4
		Maximum	558	29.6	19.7	53.9
E. saligna	5	**Average**	433	31.3	17	51.4
		St.Dev.	42.5	0.11	0.85	1.12
		Minimum	380	31.2	16.1	50.4
		Maximum	490	31.4	18.2	52.9
E. urophylla	4	**Average**	435	31.4	16.8	50.4
		St.Dev.	75.7	0.92	0.83	1.69
		Minimum	328	30.2	15.6	48.2
		Maximum	502	32.4	17.6	52.1
Hybrids*	26	**Average**	463	29.8	18.3	51.6
		St.Dev.	23.5	1.47	1	1.3
		Minimum	412	27.1	16.4	48.3
		Maximum	507	31.2	19.9	53.6

*The hybrids included in this data set are *E. dunnii X grandis* (1), *E. dunnii* X *maidenii* (2), *E. dunnii* X *urophylla* (3), *E. globulus* X *E. grandis* X *urophylla* (1), *E. globulus* X *grandis* (1), *E. globulus* X *urophylla* (6), *E. globulus* X *urophylla* X *E. grandis* X *urophylla* (1), *E. grandis* X *maidenii* (1) *E. grandis* X *urophylla* (3), *E. grandis* X *urophylla* X *E. grandis* X *maidenii* (1), *E. maidenii* X *E. grandis* X *urophylla* (1), *E. maidenii* X *saligna* (1), and *E. maidenii* X *urophylla* (4)

Table 3.2. Comparison of lab-measured properties of basic density, lignin content, pentosans, and pulp yield of sample trees from the central and southern populations. The statistics included are the minimum (Min.) and maximum (Max.) observations for each property, as well as the average (Avg.) and standard deviation (St.Dev).

	Central population (140 trees)				Southern population (55 trees)			
	Min.	**Max.**	**Avg.**	**St.Dev.**	**Min.**	**Max.**	**Avg.**	**St.Dev.**
Basic density (kg/m^3)	327	710	450	54.6	328	558	460	44.2
Lignin content (%)	24.3	33.3	29.2	1.65	27.1	33.6	29.8	1.4
Pentosans (%)	14.6	21.9	16.9	1.45	15.6	20	18.2	1.18
Pulp Yield (%)	45.2	55.9	51.9	1.81	48.2	53.9	51.7	1.27

Table 3.3. Calibration statistics for the central dataset (140 samples). The statistics presented include the coefficient of variation for the calibration (R^2_c), the standard error of calibration (SEC), the standard error of cross-validation (SECV), and the ratio of performance to deviation for the calibration (RPD_c). The number of factors used to create the calibrations are given, as well as the number of factors recommended. The numbers differ because additional calibrations with samples from the south resulted in a different number of recommended factors. The most frequently occurring number of recommended factors was used to create all calibrations.

	R^2_c	SEC	SECV	RPD_c	Factors used (Recommended)
Basic density (kg/m^3)	0.75	27.4	31.9	1.71	7
Lignin content (%)	0.76	0.72	0.78	1.85	3 (2)
Pentosan content (%)	0.73	0.85	0.93	1.78	4
Pulp yield (%)	0.63	1.12	1.44	1.26	9 (4)

Table 3.4. Prediction statistics for the southern dataset (45 samples) using original central calibrations. The statistics include the coefficient of variation for the predicted values (R^2_p), the standard error of prediction (SEP), and the ratio of performance to deviation for the predicted values (RPD_p).

	R^2_p	SEP	RPD_p
Basic density (kg/m^3)	0.14	41.7	1.07
Lignin content (%)	0.64	1.22	1.19
Pentosan content (%)	0.71	0.89	1.31
Pulp yield (%)	0.35	2.80	0.49

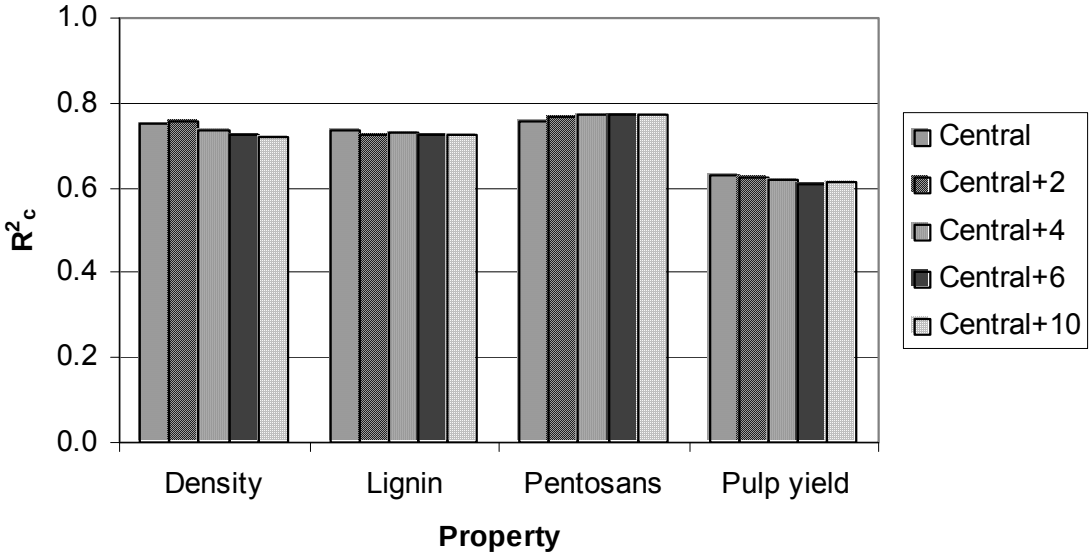


Figure 3.1. Comparison of the coefficients of deviation for calibration (R_c^2) values as southern samples are added to the central calibrations.

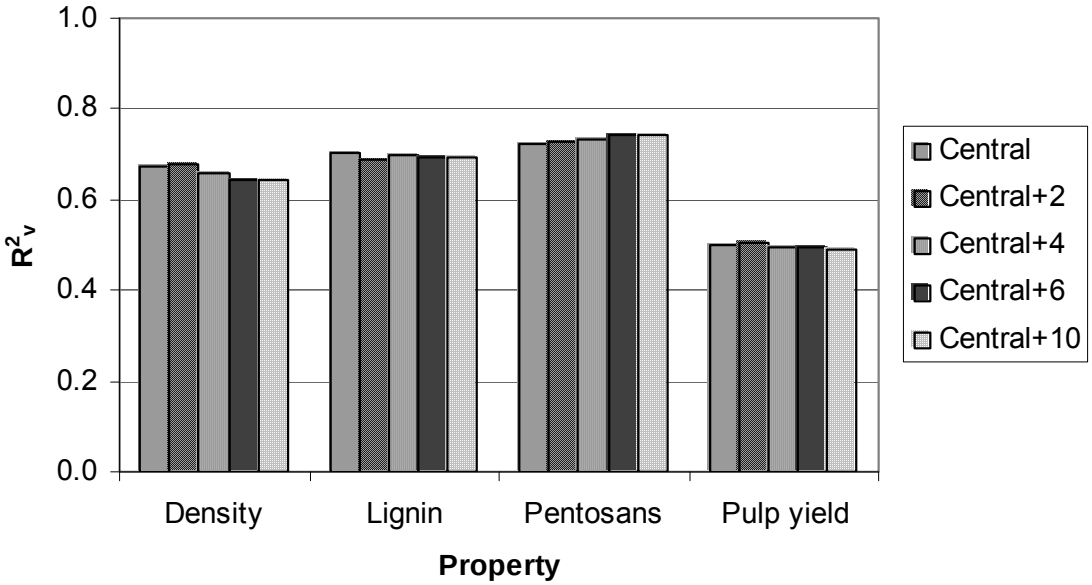


Figure 3.2. Comparison of the coefficients of deviation for cross-validation (R_v^2) values as southern samples are added to the central calibrations.

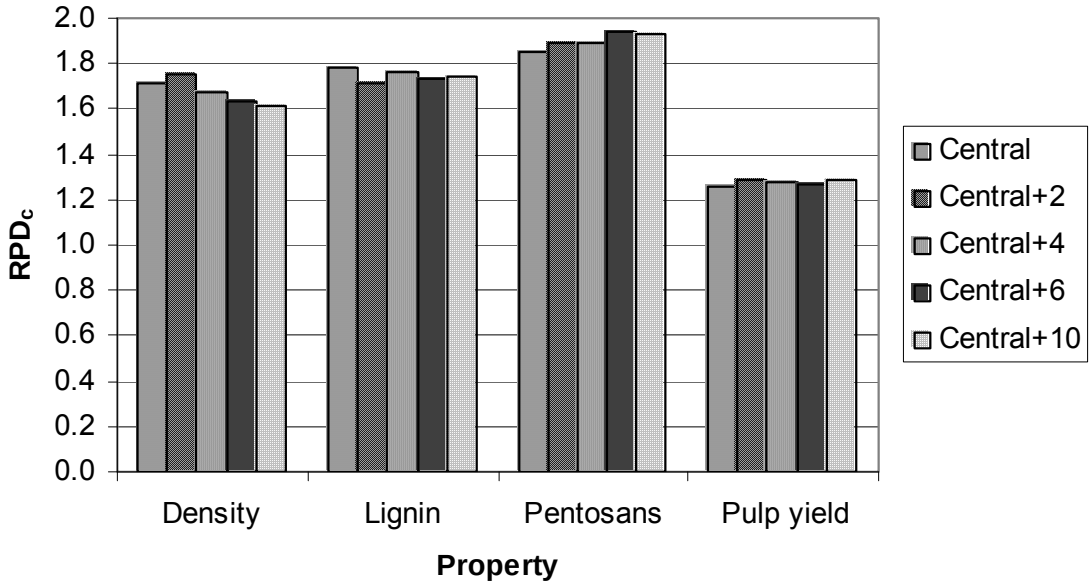


Figure 3.3. Comparison of the ratios of performance to deviation for calibration (RPD_c) as southern samples are added to the central calibrations.

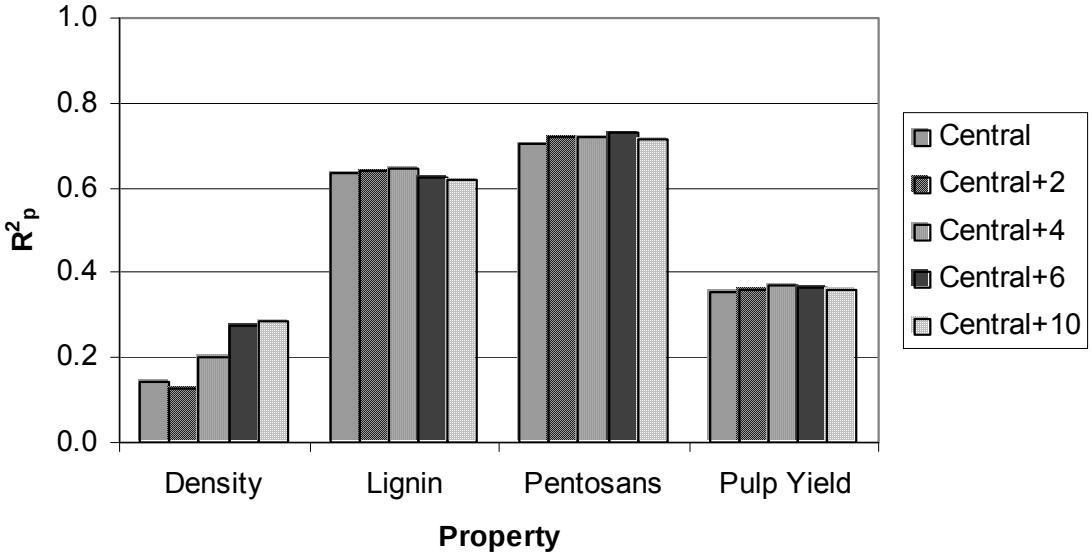


Figure 3.4. Comparison of the coefficients of determination for prediction (R_p^2) values as southern samples are added to the central calibrations.

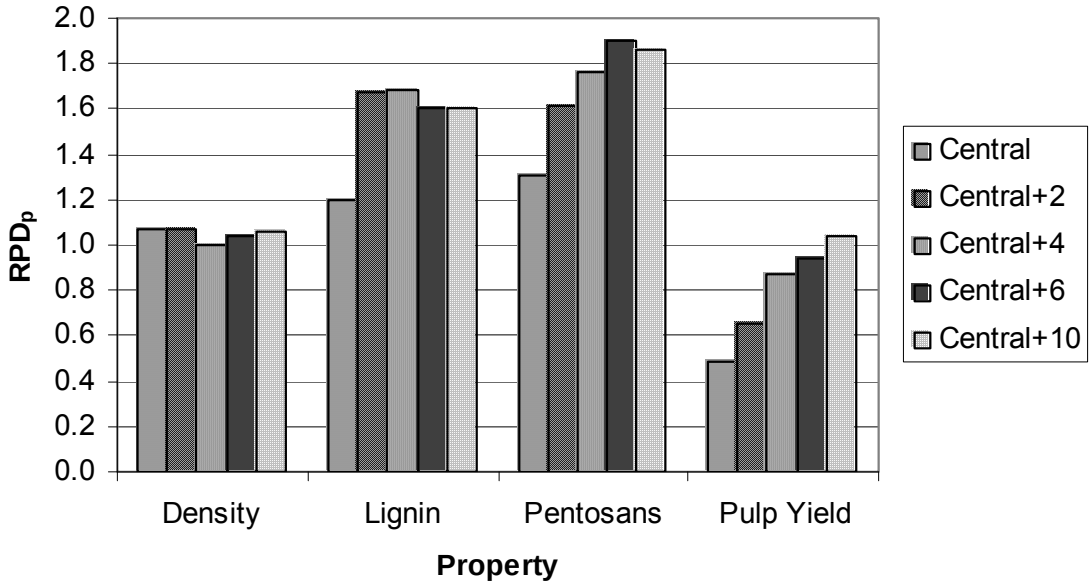


Figure 3.5. Comparison of the ratios of performance to deviation for prediction (RPD_p) as southern samples are added to the central calibrations.

Table 3.5. Comparison of prediction statistics for the southern test set using calibrations made with the central set plus 4 WinISI-selected southern samples and 4 randomly selected southern samples. The prediction set used for the calibration with WinISI-selected samples contained only 42 samples to avoid duplicating samples in the calibration and prediction sets. The statistics presented are the coefficients of determination for prediction (R_p^2), the standard errors of prediction (SEP), and the ratios of performance to deviation for prediction (RPD_p).

	WinISI			Random		
	R^2_p	SEP	RPD_p	R^2_c	SEP	RPD_p
Basic density (kg/m^3)	0.08	43.5	1.03	0.20	44.8	1.00
Lignin content (%)	0.65	0.88	1.65	0.65	0.86	1.69
Pentosan content (%)	0.71	0.65	1.80	0.72	0.67	1.76
Pulp yield (%)	0.36	1.47	0.92	0.37	1.56	0.87

Table 3.6. Comparison of prediction statistics for 40 central samples using calibrations made with only 100 central samples and 100 central samples plus 4 randomly selected southern samples. The statistics presented are the coefficients of determination for prediction (R_p^2), the standard errors of prediction (SEP), and the ratios of performance to deviation for prediction (RPD_p).

	100 Central			100 Central + 4 South		
	R^2_p	SEP	RPD_p	R^2_c	SEP	RPD_p
Basic density (kg/m^3)	0.57	37	1.52	0.57	36.8	1.53
Lignin content (%)	0.69	0.95	1.76	0.70	0.93	1.79
Pentosan content (%)	0.79	0.68	2.09	0.80	0.66	2.16
Pulp yield (%)	0.21	1.73	1.12	0.19	1.76	1.10

Chapter 4

COMPARISON OF THE WOOD PROPERTIES OF EUCALYPTUS INCREMENT CORES AND DRILL BIT SHAVINGS PREDICTED BY NEAR-INFRARED SPECTROSCOPY[2]

[2] Tyson, Justin A., Laurence A. Schimleck, Aurélio M. Aguiar, Jupiter I. Muro Abad, and Gabriel D. S. P. Rezende. To be submitted to Appita Journal.

ABSTRACT

The use of near-infrared (NIR) spectroscopy to rapidly provide estimates of wood properties in the pulp and paper industry has steadily increased over the last two decades. Recently, there has been growing interest in the use of milled increment cores to provide nondestructive samples for NIR analysis. Drill bit shavings provide a quicker and cheaper alternative to increment cores, but it is unknown if the properties of samples collected with a drill bit may be accurately predicted using calibrations based on milled cores. This study compares the estimated values of wood properties of eucalypts from Aracruz, Brazil that were sampled using both collection methods. The calibrations used to provide the estimations were based on NIR spectra from milled increment cores from 195 samples in central and southern Brazil. We plotted predicted values for cores versus shavings for four properties and observed that all plots displayed significant offset. The plots for basic density and pentosans content also had a large amount of bias. The coefficients of determination (R^2) and Pearson correlation coefficients (ρ) for basic density and pentosans were unacceptably low. Though the R^2 values for lignin content and pulp yield were better, the additional error introduced by the using shavings rather than cores would be unacceptable. The ρ-values for these properties were higher, but the rankings produced for neither property (particularly pulp yield) would be useful.

INDEX WORDS: near-infrared spectroscopy, non-destructive sampling, eucalyptus, lignin content, pentosans, basic density, pulp yield

INTRODUCTION

Near-infrared (NIR) spectroscopy is being increasingly used in the forest products industry to quickly estimate a variety of wood properties. Numerous studies have demonstrated the potential for NIR to predict a variety of wood properties important to the pulp and paper industry, including density, lignin and cellulose content, and pulp yield (Easty et al. 1990; Garbutt et al. 1992; Michell 1995; Raymond et al. 2001; Schimleck et al. 2005a; Schimleck et al. 2001; Schultz and Burns 1990; Wright et al. 1990). One method of collecting samples for NIR analysis that has been widely used is to cut down selected trees, collect bolts or discs from predetermined heights and to then chip these samples to give a whole-tree composite sample that is used for analysis. A subsample of chips is then milled into a fine powder, and an NIR spectrum is obtained from the powder (Easty et al. 1990; Garbutt et al. 1992; Michell 1995; Schimleck et al. 2001; Schimleck et al. 2005b; Schultz and Burns 1990; Wright et al. 1990). Although this method is less time-consuming and costly than traditional wet chemistry techniques, it still requires a considerable amount of effort and expense. More importantly, if the samples are being tested for use in a tree improvement program, it is detrimental to kill the trees that one is interested in breeding.

A more rapid and inexpensive approach to this whole-tree method is the nondestructive sampling of trees through the use of increment cores. Several studies have shown that whole-tree wood properties of eucalypts may be estimated with ground increment cores, which allows researchers to sample trees for NIR analysis without destroying them (Raymond et al. 2001; Schimleck et al. 2005a; Schimleck and Evans 2003; Schimleck et al. 2005b; Schimleck et al. 2006). Thus, only the trees used to create

the calibrations must be cut down to determine their whole-tree properties. An even quicker and cheaper method of collecting wood samples from standing trees is to obtain shavings using a standard drill bit.

Aracruz Celulose S.A., a major producer of eucalyptus pulp, has spent considerable effort developing wood property calibrations based on milled increment cores (Schimleck et al. 2006). They are currently using these calibrations to select superior clones for use in their tree improvement programs. Due the inherent benefits of sampling with a drill bit rather an increment borer, their researchers are considering implementing this technique to save time and money. However, it is not clear what effect this sampling method may have on the resultant predictions.

Schimleck and Michell (1999) showed that microscopic morphological differences existed between milled increment cores and milled drill bit shavings from Tasmanian *Eucalyptus globulus*. These discrepancies resulted in differing NIR spectra, which were then used to predict pulp yield using a calibration that was developed using milled increment cores. It was found that the differing spectra resulted in different predictions of pulp yield. However, it is not known if the spectral changes would significantly affect the predictions of different wood properties, particularly chemical properties such as lignin and pentosan content. Furthermore, this study did not investigate the effects that the different sampling techniques had on the rankings of the individual trees based on the wood property estimations. For their breeding program, Aracruz is primarily concerned with the rankings of the predictions to select the best individuals. If the rankings of wood property characteristics that are predicted using drill bit shavings are similar to the rankings predicted by the increment cores, then shavings

may be used in place of cores, regardless whether or not the actual estimations are accurate. Finally, Schimleck and Michell used a Wiley mill and a Newport mill to grind the samples and found that the finer powder produced by the Newport mill resulted in NIR spectra that were more similar to each other than those produced by the Wiley mill. The calibrations used by Aracruz rely on a Cyclotec mill, which has a grinding action similar to the Newport mill and also produces a finer powder, which may result in more similar wood property estimations between correspondent drill bit and increment cores samples.

Thus, this project has three main objectives: 1) to determine if the use of milled drill bit shavings rather than milled increment cores causes a significant difference in the predicted values of four wood properties—basic density, lignin content, pentosan content, and pulp yield; 2) assuming significant differences exist, to determine if the wood property values predicted using drill bit shavings can be converted with a linear equation to more closely correspond to the values predicted using the increment cores; and 3) if the accuracy of the predicted wood property values cannot be sufficiently increased with a linear equation, to determine if the rankings of the candidate trees based on the two sample collection methods are similar enough to permit the use of drill bit shavings for sample collection, regardless of the inaccuracy of the predictions.

MATERIALS AND METHODS

Sample origin

Forty 2-year-old eucalyptus trees were selected from a eucalyptus plantation in northern Espírito Santo, Brazil (40°07'W, 19°42'S). The trees were selected to provide a wide range of genetic variability. Five specimens of *E. grandis* and *E. urophylla* were

sampled, as well as ten *E. grandis* x *urophylla* hybrids. The remaining 25 trees were various species and hybrids among *E. dunnii, E. globulus, E. nitens,* and *E. pellita*.

Sample collection

Two samples (one increment core and one sample of shavings) were taken from each of the selected trees. To collect the samples, first a small patch of bark was removed from the trees to prevent bark contamination of the wood. The increment cores were taken at a height of 1.3 meters with a 5 mm increment borer powered by a motor-driven hand drill. The increment cores were transferred to a controlled environment of 20°C and 40% humidity for drying. The samples of shavings were taken within 5 cm of the increment cores with a 10 mm spade bit powered by a motor-driven hand drill. The shavings were also transferred to a controlled environment of 20°C and 40% humidity for drying.

NIR spectroscopy

After the samples were thoroughly dried to a consistent moisture content, they were ground to a fine powder using a Cyclotec mill. The mill was thoroughly cleaned after grinding each sample to prevent cross-contamination. NIR spectra were obtained from all samples using a NIRSystems Inc. Model 5000 scanning spectrophotometer and Vision® software (version 3.1). The spectra were collected in diffuse reflectance mode in 2 nm intervals over the wavelength range 1100-2500 nm. Two spectra were obtained for each sample, and the data were exported to Unscrambler® software (version 9.2) for data analysis. The duplicate spectra were averaged and a second-derivative mathematical treatment was applied to the spectra with left and right side gaps of 8 nm to minimize the particles size effects on the spectra.

Wood property predictions

The spectra were then used to predict four wood properties of interest (basic density, lignin content, pentosan content, and pulp yield) using NIR calibrations created using milled increment cores. These calibrations were created for use throughout Aracruz's plantations in central and southern Brazil, and include 195 samples of many different species and hybrids from 5 locations. The wood properties of the calibration set varied widely and are assumed to sufficiently embody the variability in the prediction set (Fig. 4.1). The best calibrations have been for lignin content and pentosans, while the calibrations for basic density and pulp yield are quite poor (Fig. 4.2). A more detailed description of the calibrations is provided in Chapter 3 and by Schimleck et al. (2006). Two predictions for each wood property were collected for each of the forty trees sampled—one for the milled increment cores, and one for the milled shavings.

Data analyses

Several methods were used to determine if the wood property predictions provided by the shavings were significantly different from the predictions provided by the increment cores. The core-predicted wood properties were plotted against the properties predicted by the shavings, and a linear regression model for each wood property was used to examine the correlation between the core and shaving predictions. Estimates of slope and y-intercept were calculated for each model, and 95% confidence intervals for these estimates were determined. The Pearson correlation coefficients (R) and coefficients of determination (R^2) were calculated to determine the proportion of variation in the shavings prediction set that was explained by the variation in the core prediction set. Spearman's correlation coefficient (ρ), a non-parametric statistic that is

analogous to the Pearson correlation coefficient, was used to determine if the rankings of the sample trees with regard to predicted wood property values differed significantly between the two techniques.

RESULTS

The 95% confidence intervals for the intercept coefficients did not include 0 for any of the four wood properties, which is an indication that the milled shaving spectra provided biased wood property estimates (Table 4.3). The confidence intervals for the slope coefficients did not include 1 for any property, further indicating that the wood property estimates for milled shavings were biased (Table 4.3). Based on these data, we concluded that shavings were unable to provide accurate predictions of wood property. The confidence intervals for slope coefficients also did not include 0, indicating that there is a statistically significant relationship between the wood property estimates predicted by milled increment cores and milled drill shavings (Table 4.3). However, this does not indicate how strong the relationships between the sets of estimates are.

Plots displaying core-predicted wood properties versus shaving-predicted wood properties show that the data points were scattered irregularly around the regression lines. Basic density (0.13) and pentosans (0.32) had the lowest R^2 values (Fig. 4.1). Although the R^2 for lignin content and pulp yield were higher, at 0.68 and 0.64, respectively (Fig. 4.1), the calibration statistics for pulp yield were already quite poor (Table 4.2), and the moderate R^2 value indicates that using shavings to predict this property would only further reduce the accuracy of the predictions. The statistics also indicate that using NIR spectra from milled shavings in place of NIR spectra from milled increment cores to predict lignin content would not be advisable. Although the calibration for lignin content

was better than the calibration for pulp yield, it still resulted in a sizable amount of error in the predictions of lignin content, and using shavings would add even more error to the predictions.

Thus, the low R^2 for all four properties indicate that it would not be possible to adjust the predictions from the shavings with a linear equation to obtain accurate estimations. However, this does not signify that the rankings of the properties predicted with the shavings are inaccurate. Because the purpose of these predictions is to identify the best clones for use in a tree improvement program, Spearman correlation coefficients (ρ) were calculated to determine if the rankings could be accurately predicted using the shavings. The ρ for all properties were very similar their respective R (Fig. 4.2), so our ability to provide accurate wood property rankings with milled shavings was not significantly better than our ability to provide accurate wood property estimates with milled shavings.

DISCUSSION

The results of this study demonstrate that it is inadvisable to sample trees with a drill bit rather than an increment borer when the calibration that is used to predict the sample properties was created using increment cores. This finding is consistent with the findings of Schimleck and Michell (1999), who found that the predicted pulp yields of E. globulus trees were significantly different between these two sampling techniques. When using NIR spectroscopy to estimate any given parameter, it is always desirable to collect and process the prediction samples using methods that are as similar as possible to the methods used to create the original calibrations. This study further substantiates this fundamental characteristic of NIR spectroscopy.

If researchers desire to change the methods for collecting prediction samples, they must collect additional calibration samples using the new methods to create completely new calibrations. However, this is an expensive process that requires a large number of new samples be analyzed to obtain a sufficient number of calibration samples to successfully represent the population. A second option that requires further testing may be to analyze a small number of samples using the new methods and add these samples to the calibration set to adjust the calibrations. This process been used to adjust NIR calibrations for use in a region that is different from the region that the calibration samples were collected in and has proven effective at increasing the accuracy of the predictions (Jones et al. 2005; Schimleck et al. 2005b). There is reason to believe that this technique may be used effectively to adjust for many different types of changes in sample collection procedures, but more research is necessary to test this idea.

One shortcoming of this study is that the actual values for the four properties that were investigated are unknown. Hence, comparisons could only be made between the two NIR-predicted values. If the actual values of these four properties were known, they could be compared to the NIR-predicted values, which would lead to much more conclusive inferences. However, it is nearly certain that the amount of money that would be saved through reduced sampling effort if Aracruz were to begin collecting NIR samples with a drill bit would be much less than the amount of money that may be gained through increased wood quality if they continue to collect samples with an increment borer.

CONCLUSIONS

We used NIR calibrations for eucalyptus wood properties that were created using milled increment cores to predict the wood properties of milled increment cores and drill bit shavings from 40 trees with unknown properties. The predicted properties for the two sample collection methods were used to create linear regressions for the wood properties. The 95% confidence intervals for the slope and intercept estimates indicated that none of the wood properties were estimated accurately by milled shaving spectra. The R^2 for the plots were too low to consider using a linear regression to adjust the predictions. Spearman correlation coefficients (ρ) were very similar to Pearson correlation coefficients (R) for all properties, so the wood property rankings were also inaccurately estimated with spectra from milled shavings. These results indicate that it is inadvisable to use milled drill bit shavings in place of milled increment cores for predicting wood properties when using a calibration that was developed with milled increment cores. The amount of error that would be introduced into the estimations would be unacceptable, even for ranking purposes.

LITERATURE CITED

Easty, D.B., Berben, S.A., DeThomas, F.A., and Brimmer, P.J. 1990. Near-infrared spectroscopy for the analysis of wood pulp; quantifying hardwood-softwood mixtures and estimating lignin content. Tappi J. 73(10): 257-261.

Garbutt, D.C.F., Donkin, M.J., and Meyer, J.H. 1992. Near-infrared reflectance analysis of cellulose and lignin in wood. Pap. S. Afr. April: 45-48.

Jones, P.D., Schimleck, L.R., Peter, G.F., Daniels, R.F., and Clark, A. 2005. Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. Can. J. For. Res. 35: 85-92.

Michell, A.J. 1995. Pulpwood quality estimation by near-infrared spectroscopic measurements on eucalypt woods. Appita J. 48(66): 425-428.

Raymond, C.A., Schimleck, L.R., Muneri, A., and Michell, A.J. 2001. Nondestructive sampling of *Eucalyptus globulus* and *E. nitens* for wood properties. III. Predicted pulp yield using Near Infrared Reflectance Analysis. Wood Sci. Technol. 35: 203-215.

Schimleck, L.R., Downes, G.M., and Evans, R. 2005a. Estimation of *Eucalyptus nitens* wood properties by near-infrared spectroscopy. Appita J. 59(2): 136-141.

Schimleck, L.R., Evans, R., and J., I. 2001. Estimation of *Eucalyptus delegatensis* wood properties by near infrared spectroscopy. Can. J. For. Res. 31: 1671-1675.

Schimleck, L.R., Kube, P.D., Raymond, C.A., Michell, A.J., and French, J. 2005b. Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. Can. J. For. Res. 35: 2797-2805.

Schimleck, L.R., and Michell, A.J. 1999. Comparison of milled wood from increment cores and drill shavings. CSIRO Divisional Report.

Schimleck, L.R., Rezende, G.D.S.P., Demuner, B.J., and Downes, G.M. 2006. Estimation of whole-tree wood quality traits using near infrared spectra from increment cores. Appita J. 59(3): 231-236.

Schultz, T.P., and Burns, D.A. 1990. Rapid secondary analysis of lignocellulose: comparison of near-infrared and fourier transform infrared (FTIR). Tappi J. 73(5): 209-212.

Wright, J.A., Birkett, M.D., and Gambino, M.J.T. 1990. Prediction of pulp yield and cellulose content from wood samples using near-infrared reflectance spectroscopy. Tappi J. 73(8).

Table 4.1. Summary of wood property statistics for the 195 samples in the calibration set. The statistics include the average, standard deviation (St.Dev.), coefficient of variation (C.V.), and the minimum and maximum values for each wood property.

	Average	St.Dev.	C.V.	Minimum	Maximum
Basic density (kg/m^3)	453	52	11.50%	327	710
Lignin content (%)	29.4	1.6	5.40%	24.3	33.6
Pentosans (%)	17.3	1.51	8.70%	14.6	21.9
Pulp yield (%)	51.8	1.67	3.20%	45.2	55.9

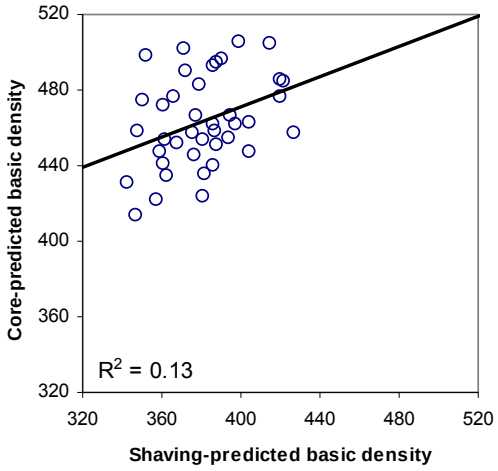
Table 4.2. Statistics for the four calibrations used to predict the wood properties of the 40 sample trees used in this study. Statistics reported include coefficients of determination for calibration (R^2_c), coefficients of determination for cross-validation (R^2_v), standard error of calibration (SEC), standard error of cross-validation (SECV), and ratio of performance to deviation (RPD), which is the ratio of the standard deviation to the SECV. The number of factors used in the calibration is also reported.

	Factors	R^2_c	R^2_v	SEC	SECV	RPD
Basic density	10	0.759	0.565	25.5	35	1.49
Lignin content	5	0.726	0.7	0.834	0.894	1.79
Pentosans	3	0.795	0.787	0.68	0.709	2.13
Pulp yield	5	0.479	0.426	1.21	1.36	1.23

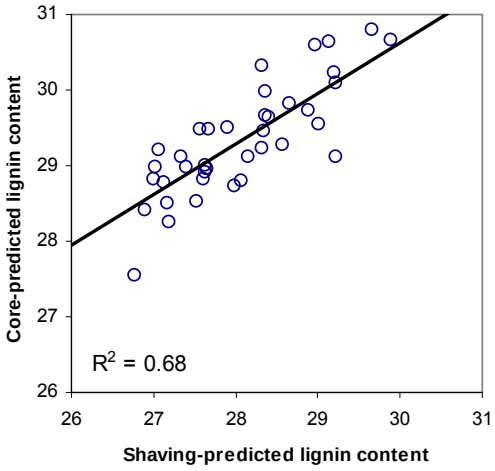
Table 4.3. Comparison of slope and intercept parameters for the relationships between wood property estimates predicted by milled increment cores and milled shavings. A 95% confidence interval that does not include 0 for the intercept coefficient indicates that shavings are providing biased estimates. If the 95% confidence interval for the slope coefficient does not include 0 indicates that there is a statistically significant relationship between the wood property estimates predicted by milled increment cores and milled drill shavings but does not indicate how strong the relationship is. A confidence interval for the slope coefficient that does not include 1 is another indication that shavings are providing biased estimates.

		Estimate of Coefficients	Standard Error	95% Confidence Interval
Basic density	**Intercept**	310.7	63.2	182.9 - 438.6
	Slope	0.400	0.165	0.065 - 0.735
Lignin content	**Intercept**	10.5	2.1	6.3 - 14.7
	Slope	0.670	0.074	0.521 - 0.820
Pentosan content	**Intercept**	6.80	2.41	1.93 - 11.67
	Slope	0.635	0.149	0.332 - 0.937
Pulp yield	**Intercept**	28.6	3.0	22.5 - 34.6
	Slope	0.425	0.052	0.320 - 0.531

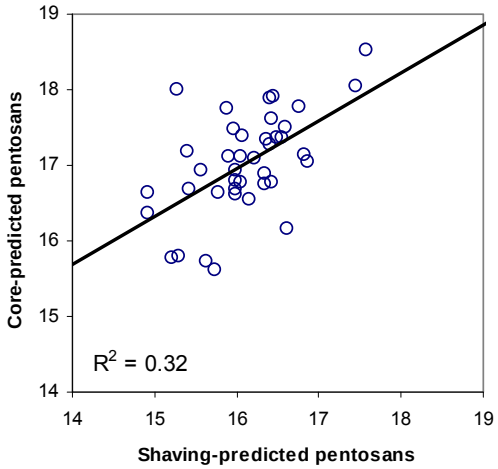
Figure 4.1. Relationships between wood properties estimates predicted using NIR spectra from milled increment cores and milled drill shavings. R^2 for the relationships are presented in the figures and regression lines have been plotted.



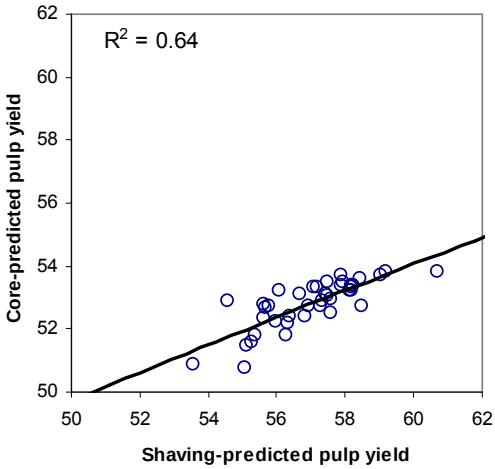
1



2



3



4

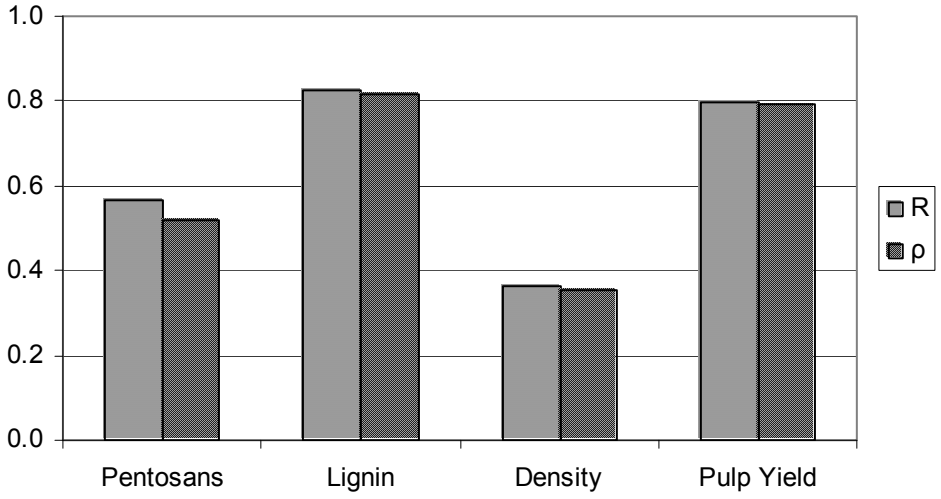


Figure 4.2. Comparison of Pearson correlation coefficients (R) and Spearman correlation coefficients (ρ) for wood property estimates predicted by milled increment cores and milled drill shavings. R is a parametric statistic and was included to enhance the interpretation of ρ, a non-parametric statistic.

Chapter 5

DEVELOPMENT OF NIR CALIBRATIONS FOR PHYSICAL AND MECHANICAL PROPERTIES OF EUCALYPT PULPS OF MILL-LINE ORIGIN[3]

[3] Tyson, Justin A., Laurence A. Schimleck, Aurélio M. Aguiar, Jupiter I. Muro Abad, and Otávio M. Filho. To be submitted to Journal of Wood Chemistry and Technology.

ABSTRACT

Near-infrared (NIR) spectroscopy has been used in several studies to predict the physical and mechanical properties of pulp handsheets. In most of these studies, wood samples were pulped in a laboratory under different regimes to introduce variability into the data set. This study investigates the potential of NIR spectroscopy to create calibrations for eucalyptus pulp properties of mill-line origin. Seven mechanical properties (air resistance, compressibility, drainability, hygroexpansivity, stretch, tensile index, and tensile stiffness) and three physical properties (bulk density, specific volume, and surface area) were investigated. Coefficients of determination (R^2) for all ten properties were poor. The R^2_c value exceeded 0.70 for only one property (tensile index), while the R^2_v values exceeded 0.40 for only two properties (drainability and surface area). Ratios of performance to deviation (RPD) were equally poor, ranging from 0.87 for bulk density to 1.28 for drainability. These statistics indicate that none of the calibrations could be used to accurately predict the properties of unknown samples. The poor performance of the calibrations is likely due to the low variability of our dataset, which is generally inherent in samples of mill-line origin.

INDEX WORDS: near-infrared spectroscopy, eucalyptus, pulp, paper, physical properties, mechanical properties

INTRODUCTION

There has been increasing interest in the use of near-infrared (NIR) spectroscopy in the pulp and paper industry in the last two decades. One of the many potential applications for NIR spectroscopy is the analysis of the pulp properties. Several articles in the early 1990's demonstrated the potential for this technology to predict a variety of important physical and mechanical pulp properties using spectra collected from prepared handsheets (Edlund et al. 1991; Wallbacks 1993; Wallbacks et al. 1991). Other researchers investigated the use of NIR spectra from milled wood chips to predict the properties of the pulps and handsheets (Marklund et al. 1999; Meder et al. 1994). Meder et al. (1994) found this technique to be unsuccessful, but Marklund et al. (1999) reported that the NIR spectra from milled wood chips had nearly the same predictive ability as the spectra from bleached pulps.

Most of the research in the use of NIR technology to predict physical and mechanical pulp properties has been conducted in Sweden and utilized softwood pulps (Antti et al. 1996; Edlund et al. 1991; Marklund et al. 1999; Meder et al. 1994; Wallbacks 1993; Wallbacks et al. 1995; Wallbacks et al. 1991). However, in many temperate and tropical regions of the world eucalypts are the predominant pulpwood species due to their rapid growth rate, ease of propagation, and value for producing high-quality printing and writing papers and tissue products. Eucalypt fibers are relatively short, slender, and thin-walled (Brumby and Maddern 1990; Higgins 1970), and these properties lead to excellent sheet formation and a sheet that has high bulk, excellent surface properties and good density, stiffness, and optical properties (Brumby and Maddern 1990).

Fardim et al. (2005) used NIR spectra from bleached *E. grandis* pulp handsheets to create calibrations with good predictive abilities for ten pulp properties of interest. However, like many other studies in this field, they used different pulping and beating regimes to introduce variability into their sample data set (Edlund et al. 1991; Wallbacks 1993; Wallbacks et al. 1995; Wallbacks et al. 1991). The Kappa (KP) numbers of their calibration set varied from 11.6 to 23.7, while the intrinsic viscosity (IV) ranged from 656-1152 $cm^3 g^{-1}$. A high amount of variability in the calibration set improves calibration statistics, but does not reflect the reality of a commercial pulping operation, where the pulping process is tightly regulated and the variability of the final product is minimized. Thus, the goal of this study was to determine if accurate calibrations for important physical and mechanical characteristics of eucalyptus pulps could be created using NIR spectra from handsheets of mill-line origin.

MATERIALS AND METHODS

Sample origin

Thirty-seven samples of bleached eucalyptus pulp were collected from Aracruz Celulose's elementary chloride-free (ECF) pulp mill at Aracruz, Brazil. The samples were collected at one week intervals over a 37-week period and consisted of several different eucalyptus species and hybrids from the company's plantations in Vitória, Bahia, and Minas Gerais. *E. grandis*, *E. urophylla*, and hybrids between these species made up the majority of the pulp mixture.

Measurement of physical and mechanical properties

The mechanical properties of the pulps were measured on handsheets with a grammage of 60 g/m^2, prepared with a Rapid Köthen apparatus using deionized water.

Tensile index and stretch were measured according to ISO 1924-3:2005. Air resistance was measured using the Gurley method according to ISO 5636-5:2003. Hygroexpansivity was measured according to ISO 8226-1:1994. The degree of refinement as indicated by Schopper-Reigler (SR) drainability was measured according to ISO 5267-1:1999. The compressibility index was also measured.

Three physical properties of the pulps were also quantified, including bulk density, surface area, and specific volume. A summary of each of the measured properties is provided (Table 5.1).

NIR spectroscopy

Two diffuse reflectance NIR spectra were collected from each in a controlled environment of 20°C and 40% humidity. The handsheets were held in a static sample holder in a NIRSystems Inc. Model 5000 scanning spectrophotometer. The spectra were collected using Vision® software (version 3.1) at 2 nm intervals over the wavelength range 1100-2500 nm with a ceramic standard as the instrument reference. The data were then imported into the Unscrambler® software (version 9.2) and the duplicate spectra were averaged. The spectra were converted to the second derivative using left and right gaps of 8 nm.

Data analyses

Calibrations were developed for each property using Partial Least Squares (PLS) regression using full cross-validation and a maximum of 10 factors with the Unscrambler® software. Several statistics were used to assess the performance of the calibrations. The standard error of calibration (SEC), standard error of cross-validation (SECV) and the coefficients of determination (R^2_c and R^2_v) were used to assess

calibration performance. The ratio of performance to deviation (RPD), calculated as the ratio of the standard deviation of the reference data to the SECV, was also used to assess calibration performance. Determination of the RPD allowed direct comparisons of calibrations developed for different pulp properties that have different ranges in values. Because of the small size of the data set (37 samples) and the poor performance of the preliminary calibrations, a prediction set was not created to test the calibrations.

RESULTS

Coefficients of determination for the calibrations (R^2_c) varied widely for the ten properties (Fig. 4.1). Tensile index (0.70), surface area (0.68), and drainability (0.68) provided the highest R^2_c (Fig. 5.2). Calibrations for these properties are shown in Figure 5.2. The R^2_c values for stretch and hygroexpansivity were extremely low, with respective values of 0.138 and 0.162 (Fig. 5.2). The coefficients of determination for cross-validation (R^2_v) were poor for all ten properties. Only surface area (0.42) and drainability (0.42) had values that exceeded 0.40. The RPD values also reflected the poor performance of the calibrations. The variability of this statistic was significantly lower than the variability of the coefficients of determination, ranging from 0.87 for bulk to 1.27 for surface area (Fig. 5.3). These low values indicate that none of the ten properties of interest can be accurately fitted using these NIR calibrations.

DISCUSSION

Fardim et al. (2005) investigated three of the same properties that are considered in this study–tensile index, stretch, and drainability–and presented much better calibrations for all three properties. Unfortunately, they reported only the root mean square errors of calibration (RMSEC) and the root mean square errors of cross-validation

(RMSECV) for their calibrations. These statistics do not allow direct comparisons between calibrations from different data sets, because their values are dependent upon the range and variability of the individual sets of data. Coefficients of determination and RPD values would allow better comparisons between the models, as they are unitless parameters. However, a simple observation of the measured versus predicted values for the models they presented clearly demonstrates that their calibrations are much more accurate. This is most likely due to the higher variability of their dataset, introduced by the different pulping and beating regimes applied to their samples.

Antti et al. (1996) also investigated several of the properties examined in this study, including bulk density, tensile stiffness, tensile index, and stretch. Although they used mixed softwood pulps, they reported R^2_c and R^2_v for each of their properties, thus facilitating comparisons to our results. Of these four properties, bulk density resulted in the best calibration ($R^2_c > 0.8$, $R^2_v > 0.6$), which contrasts with our results. We found that bulk density provided the worst R^2_v and RPD values of any calibration. Their other findings, however, agree with our results. They found that tensile index yielded better calibration statistics than tensile strength, which likewise provided better statistics than stretch, although none of these properties were well-fitted. Our coefficients of determination are lower for each of these properties, but the ranking is identical. Again, it is likely that our calibrations performed more poorly because of the low variability of our samples. Antti et al. (1996) used 31 different mixtures of three different wood species–Polish pine, Swedish pine, and Swedish spruce–to introduce variability into their data set.

Four of the variables investigated in this study, including compressibility, hygroexpansivity, specific volume and surface area, have not been previously examined using NIR technology. It is likely that, with a more variable data set, our calibrations for these properties–as well as the other six properties of interest–could be significantly improved. However, there are concerns about the applicability of a calibration created using samples pulped under variable conditions in laboratory to samples of mill-line origin. Although it may be possible to use these calibrations to adequately predict the rankings of pulp mill samples, it is unlikely that the calibrations could accurately predict the actual values of the pulp mill samples due to the inherent differences between the laboratory pulps and mill-line pulps.

As we have demonstrated in this study, the low variability of mill-line pulps is not conducive to the development of useful NIR calibrations for mechanical and physical properties. A possible remedy to this problem may be to adjust calibrations created using variable laboratory pulps by adding of pulp samples of mill-line origin. Previous studies have shown that NIR wood property calibrations created for a certain locations (or locations) can be adjusted for use in another location by adding a small number of samples from the new location to the calibration set (Jones et al. 2005; Schimleck et al. 2005). It is possible that this technique may be applied to NIR pulp property calibrations of laboratory pulps, but further research is needed to test this hypothesis.

CONCLUSIONS

We investigated the use of bleached eucalyptus pulps from the Aracruz Celulose's ECF pulp mill at Aracruz, Brazil to create NIR calibrations for ten physical and mechanical pulp properties. All of the calibrations created for these properties were poor

and not useful for prediction purposes. Although previous researchers have had success in estimating several of the same properties, all of these studies used variable pulping methods for their samples to introduce variability into their data set. Because our samples were of mill-line origin, the variability of the samples was quite low for all properties, and this likely contributed to the poor calibration statistics. This study raises serious concerns about the ability to use NIR to create accurate calibrations for physical and chemical pulp properties in an industrial pulp mill setting. It appears that the tightly regulated pulping processes reduce the variability of the pulps to such a degree that NIR analysis is unable to adequately model these properties.

LITERATURE CITED

Antti, H., Sjostrom, M., and Wallbacks, L. 1996. Multivariate calibration models using NIR spectroscopy on pulp and paper industrial applications. J. Chemom. 10: 591-603.

Brumby, P.M., and Maddern, K.N. 1990. The pulping of eucalypts. *In* World Pulp and Paper Technology. *Edited by* F. Roberts. Sterling Publications, London. p. 73-75

Edlund, U., Wallbacks, L., Norden, B., Lindgren, T., and Iversen, T. 1991. Pulp characterization and predictions of properties using spectroscopy and multivariate data analysis methods. *In* 6th International Symposium on Wood and Pulping Chemistry.

Fardim, P., Ferreira, M.M., and Duran, N. 2005. Determination of mechanical and optical properties of eucalyptus kraft pulp by NIR spectrometry and multivariate calibration. J. Wood Chem. Technol. 25: 267-279.

Higgins, H.G. 1970. Technical assessment of eucalypt pulps in the papermaking economy. Appita J. 23: 417-426.

Jones, P.D., Schimleck, L.R., Peter, G.F., Daniels, R.F., and Clark, A. 2005. Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. Can. J. Forest Res. 35: 85-92.

Marklund, A., Paper, M., Hauksson, J.B., and Sjostrom, M. 1999. Prediction of strength parameters for softwood kraft pulps. Multivariate data analysis based on orthogonal signal correction and near infrared spectroscopy. Nord. Pulp Paper Res. J. 14: 140-148

Meder, R., Gallagher, S., Kimberley, M.O., Uprichard, J.M., and Mackie, K.L. 1994. Prediction of wood chip and pulp and paper properties via multivariate analysis of spectral data. Appita J. 47(1): 479-484.

Schimleck, L.R., Kube, P.D., Raymond, C.A., Michell, A.J., and French, J. 2005. Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. Can. J. For. Res. 35: 2797-2805.

Wallbäcks, L. 1993. Characterisation of chemical pulps using spectroscopy, fibre dimensions and multivariate data analysis. *In* 7th International Symposium on Wood and Pulping Chemistry, Beijing, China. Proc. 2, 802-809

Wallbäcks, L., Edlund, U., and Lindgren, T. 1995. Multivariate characterization of pulp, Part 3. Evaluation of physical and spectroscopic data on beaten and unbeaten kraft pulp samples. Nord. Pulp Paper Res. J. 10(2): 88-93

Wallbäcks, L., Edlund, U., and Norden, B. 1991. Multivariate characterization of pulp, Part 1. Spectroscopic characterization of physical and chemical differences between pulps using C CP/MAS NMR, FT-IR, NIR and multivariate data analysis. Nord. Pulp Paper Res. J. 2: 74-80.

Table 5.1. Summary statistics for 37 bleached eucalyptus pulps.

	Air Resistance (Gurley-seconds)	1000/Compressibility (Index)	Drainability (SR)	Hygroexpansivity (%)	Stretch (%)
Average	2.3	67.7	29.5	2.77	2.28
St. Dev.	0.41	9.7	1.6	0.29	0.22
C.V.	17.98%	14.26%	5.43%	10.65%	9.49%
Minimum	1.3	48.9	32.1	3.38	1.84
Maximum	3.2	88	25.9	2.25	2.73

	Tensile Index (kNm/kg)	Tensile Stiffness (kNm/g)	Bulk (cm^3/g)	Specific Volume (m^2/kg)	Surface Area (m^2/kg)
Average	32.53	5.12	1.92	2759	1896
St. Dev.	2.56	0.30	0.04	914	97.37
C.V.	7.87%	5.89%	1.90%	3.31%	5.13%
Minimum	25.62	4.42	1.87	2880	1680
Maximum	37.56	5.65	1.99	2540	2058

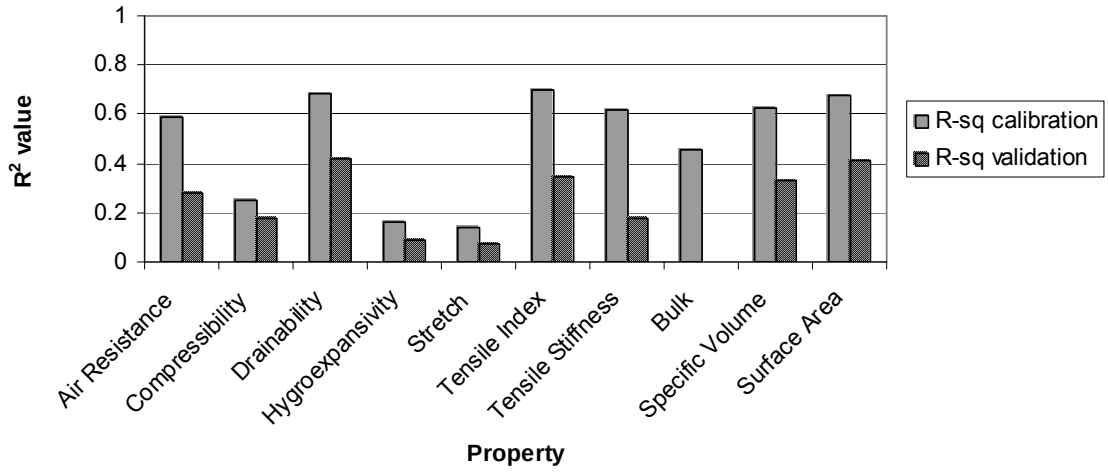


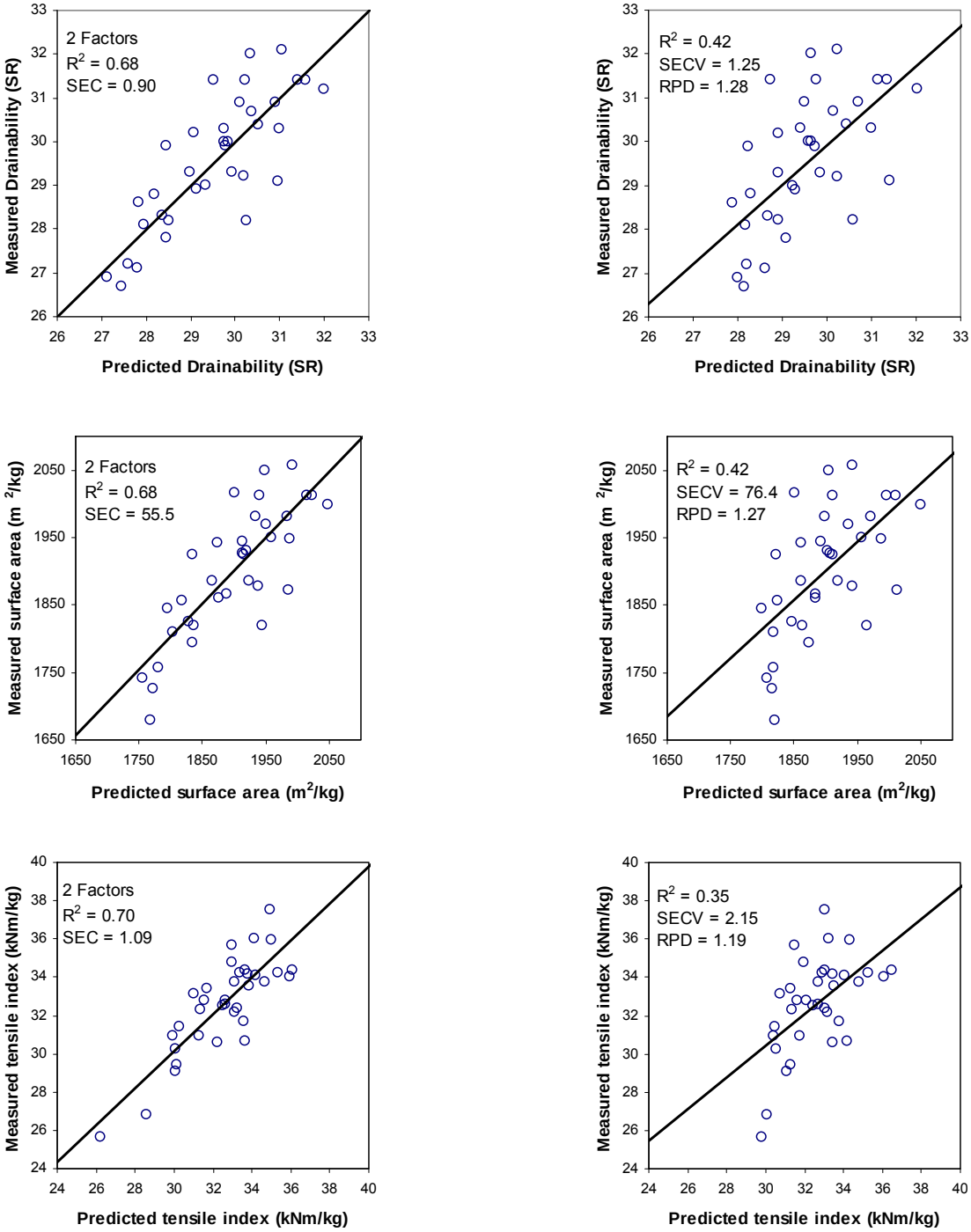
Figure 5.1. Comparison of calibration R^2_c values and R^2_v values for the ten pulp properties.

Table 5.2. NIR calibration statistics for ten pulp properties. Calibrations data were collected from 37 bleached eucalyptus pulps.

	Air Resistance	Compressibility	Drainability	Hygroexpansivity	Stretch
Factors	2	1	2	1	1
R^2_c	0.588	0.230	0.681	0.162	0.138
R^2_v	0.281	0.173	0.422	0.085	0.071
SEC	0.262	8.36	0.904	0.266	0.2
SECV	0.377	8.97	1.25	0.284	0.212
RPD	1.095	1.077	1.280	1.039	1.019

	Tensile Index	Tensile Stiffness	Bulk	Specific Volume	Surface Area
Factors	2	2	1	2	2
R^2_c	0.702	0.619	0.456	0.627	0.675
R^2_v	0.348	0.179	-0.042	0.328	0.415
SEC	1.41	0.189	0.027	0.056	55.5
SECV	2.15	0.295	0.042	0.08	76.4
RPD	1.190	1.020	0.869	1.143	1.274

Figure 5.2. Relationships between lab-measured values and NIR-estimated values for drainability, surface area, and tensile index of 37 handsheets of mill origin. Calibration segments are in the left column and cross-validation segments are in the right column.



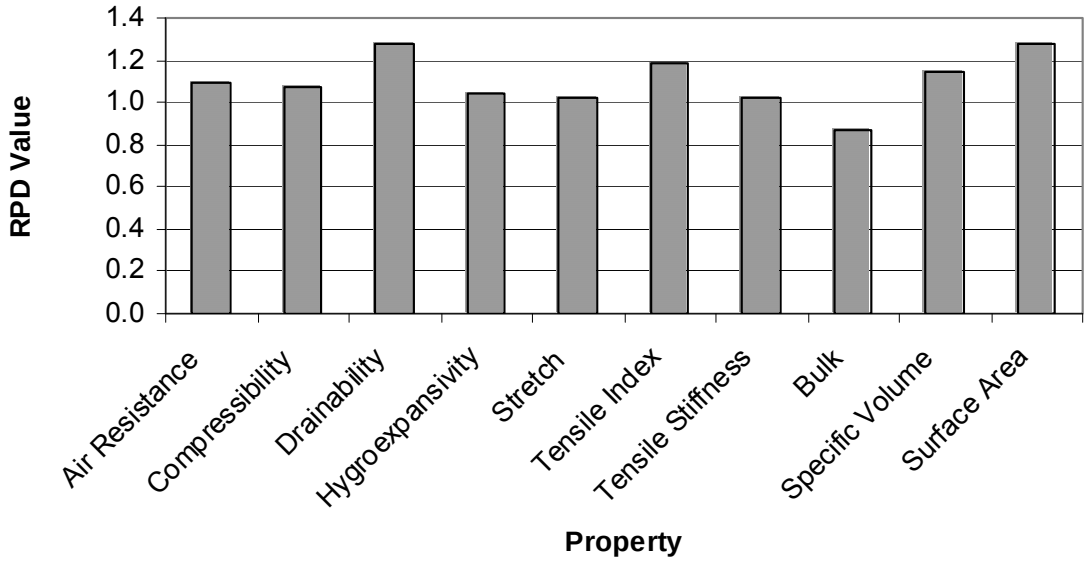


Figure 5.3. Comparison of RPD values for ten NIR-based eucalyptus pulp property calibrations.

Chapter 6

COMPARISON OF SAMPLE PREPARATION TECHNIQUES FOR NIR ANALYSIS OF CARBOHYDRATE CONTENT OF UNBLEACHED EUCALYPTUS PULPS[4]

[4] Tyson, Justin A., Laurence A. Schimleck, Christian R. Mora, Aurélio M. Aguiar, Jupiter I. Muro Abad, and Otávio M. Filho. To be submitted to Journal of Wood Chemistry and Technology.

ABSTRACT

Several studies in the last two decades have investigated the use of NIR spectroscopy for the estimation of pulp carbohydrate content. Very few have used this tool for analysis of eucalyptus pulps, which are increasingly important in the global marketplace. Previous studies used a variety of methods to prepare samples for NIR analysis, but these methods have not been compared. This study compares the quality of carbohydrate content calibrations for four sample preparation methods to identify the most suitable technique. Fifty-nine unbleached pulp samples (52 single-tree samples and 7 composite samples) were collected and prepared using four methods, identified here as coarse pulp, fine pulp, milled pulp, and handsheets. Carbohydrate composition of the pulps (including arabinose, fucose, galactose, glucose, rhamnose, and xylose) was measured. NIR spectra were collected from the samples, which were divided into a calibration set (40 samples) and a prediction set (19 samples). A calibration was created for each combination of sample preparation method and carbohydrate, and these were tested on the prediction set. Calibration statistics were good for rhamnose (RPD_c = 1.79-1.87) and xylose (RPD_c = 2.04-2.38) but poorer for other carbohydrates. Coarse and fine pulp produced the best calibrations, but they were not significantly different than calibrations for handsheets ($p = 0.347$ and 0.484). Milled pulp resulted in the worst calibrations.

INDEX WORDS: near-infrared spectroscopy, eucalyptus, pulp, paper, glucose, xylose, carbohydrates

INTRODUCTION

Near-infrared (NIR) spectroscopy is a tool that is being used increasingly in the pulp and paper industry to rapidly estimate a variety of wood and pulp properties. One possible use of considerable interest is the measurement of the carbohydrate composition of pulp. Wood pulps comprise a variety of different carbohydrates, including arabinose, fucose, galactose, glucose, mannose, rhamnose, and xylose. The relative proportions of these carbohydrates vary widely among the major pulpwood species, but glucose and xylose generally make up the greatest and second-greatest percentages, respectively, of total carbohydrates.

Carbohydrate composition affects many significant physical and mechanical properties of pulp and paper. Hemicelluloses are very important to interfiber bonding allowing cellulose fibers to swell, which improves burst and tensile strength of papers (Molin and Teder 2002). However, hemicelluloses, particularly xylose, can also have a deleterious effect on paper if the brightness of the paper is of primary concern. Xylose binds cellulose and lignin together, so the elimination of xylose through the use of xylanases allows the lignin to be more readily eliminated in subsequent bleaching stages (Pham et al. 1995; Turner et al. 1992).

For many years, the primary method of pulp carbohydrate composition determination has been acid hydrolysis, which requires severe conditions and laborious procedures (Saeman et al. 1954). More recently, pyrolysis gas chromatography has proven to be an effective alternative to acid hydrolysis. In this procedure, a pulp sample is quickly pyrolyzed in an inert atmosphere, and the resultant products are analyzed by gas chromatography (Kleen and Gellerstedt 1991). Although it is accurate and more use-

friendly than acid hydrolysis, pyrolysis gas chromatography requires very expensive equipment.

NIR spectroscopy may provide a more rapid and inexpensive approach for estimating carbohydrate composition. Wallbacks et al. (1991a) investigated the use of spectroscopic methods to predict the chemical compositions—i.e., carbohydrate composition and lignin content—of unbleached kraft pulps. They found that glucose, xylose, and lignin, the primary constituents of the pulps, were well-modeled. Of the remaining pulp constituents—arabinose, galactose, and mannose—only galactose was well-modeled. Another study from the same year focused on the use of NIR spectroscopy to predict many physical and chemical properties of bleached pine pulps, including arabinose, galactose, glucose, mannose, and xylose content (Wallbacks et al. 1991b). The study examined a small number of samples and only reported predictions for glucose and mannose, but it clearly demonstrated the potential of NIR to predict carbohydrate compositions of various pulps.

Although many subsequent studies have investigated the use of NIR spectroscopy to predict the lignin content or kappa numbers of pulps, carbohydrate composition has been largely ignored. An exception is a study published by Fardim et al. (2002) that examined the potential for NIR to predict numerous properties of unbleached kraft *E. grandis* pulps, including carbohydrate composition, uronic acids, lignin content, Kappa number, relative viscosity, and ISO brightness. The carbohydrates that they determined for the study were arabinose, galactose, glucose, mannose, rhamnose, and xylose. However, of these six carbohydrates, only glucose and xylose content were analyzed using NIR spectroscopy. They found that both glucose and xylose were well-modeled

using four and three PLS components, respectively, but the standard errors of cross-validation (SECV) and standard errors of prediction (SEP) were both several times larger than the standard errors of calibration (SEC).

These previous studies have used different sample preparation methods for collecting NIR spectra. Wallbacks et al. (1991b) milled dried pulp samples using a KAMAS Slagy 200 B mill and sieved them through a 1-mm screen, while Fardim et al. (2002) collected spectra from handsheets prepared with a grammage of 75 g/m^2. It is likely that different sample preparation methods could affect the performance of the models, but there have been no studies that compared the effects of different preparation methods applied to the same set of pulp samples. However, there have been many studies that have compared various sample collection and preparation techniques in related areas of NIR research.

Several studies have examined the performance of wood property calibrations based on milled whole-tree samples when applied to milled increment cores to develop a method for nondestructively sampling living trees (Raymond et al. 2001; Schimleck et al. 2005b; Schimleck et al. 2005c). Some articles have compared wood property calibrations based on green wood to those created using dry wood (Schimleck et al. 2003, 2004). Others have investigated whether radial strips cut from increment cores provide better wood property calibrations if the transverse or longitudinal surfaces are exposed to NIR radiation (Schimleck et al. 2005a; Schimleck et al. 2005d). Despite the diversity of these studies, all used similar statistics to compare the effects of the varying sample collection and preparation techniques on the performance of the correspondent calibrations. Correlation coefficients (R^2), standard errors, and ratios of performance to deviation

(RPD) were utilized by the majority of these studies (Schimleck et al. 2005a; Schimleck et al. 2005b; Schimleck et al. 2003, 2004; Schimleck et al. 2005c; Schimleck et al. 2005d), but one study evaluated only correlation coefficients and standard errors (Raymond et al. 2001).

Although these are excellent statistics for evaluating the performances of individual models, there are problems with using them to assess the relative performances of comparable models, as it is inherently subjective. These statistics give no indication of the probability that the models in question are statistically different from each other, as they do not take sample size into consideration. A more objective approach to evaluating NIR calibrations that are created using different sample preparation would be a significant improvement over previous methods.

Therefore, the objectives of this study are to: 1) identify several possible sample preparation techniques for the NIR analysis of unbleached eucalypt pulps, 2) identify statistical procedures to allow for objective comparisons to be made among the models developed using these different sample preparation techniques, 3) use these statistical procedures to identify the most appropriate sample preparation technique for the NIR analysis of pulp carbohydrate composition, and 4) determine which of the carbohydrates are best-modeled using NIR analysis.

MATERIALS AND METHODS

Sample origin

Fifty-two 2- to 3-year old eucalyptus trees were harvested from plantations in two locations of Rio Grande do Sul, Brazil—Colorado (51°51'W and 30°05'S), Horto Barba Negra (51°14'W and 30°26'S). The trees were chosen to represent a wide range of

genetic variability and included six different species and many different hybrids among these species (Table 5.1). The diameters at breast height (DBH) of the trees ranged from 9 to 18 cm. In addition, seven samples were taken from the factory line in of the pulp manufacturing plant in Aracruz, Espírito Santo, Brazil. These were composite samples, as each sample consisted of many different trees, the majority of which were *E. grandis* x *E. urophylla* hybrids. Single-tree samples were used to increase the variability of the dataset without changing the methodology of the pulping process.

Pulping process

The trees for single-tree samples were chipped to form whole-tree composites, and 1000 g of the chips were pulped in a laboratory digester to an intended Kappa number of 18. The actual Kappa numbers of the single-tree samples ranged from 16.6 to 19.0 (Table 6.1). The composite samples were taken from the Elementary Free Chloride (ECF) factory line, before bleaching took place. The Kappa numbers for the composite samples ranged from 15.2 to 16.8.

Pulp carbohydrate analysis

The carbohydrate contents of the pulps were determined in acidic hydrozylates using HPLC-PAD according to the methodology described by Fardim et al. (1999). The carbohydrates measured were arabinose, fucose, galactose, glucose, rhamnose, and xylose (Table 6.1).

Sample preparation

A handsheet with a grammage of 60 g/m^2 was created for each sample, prepared with a Rapid Köthen apparatus using deionized water. The remainder of each sample was dried as loose pulp. Two disks with a diameter of 38 mm were cut from each

handsheet for NIR analysis. These discs are referred to as 'handsheets'. The loose pulp was divided into three groups. The first group, designated 'coarse pulp', was analyzed without any further processing. The second group was sieved through a 7.5 mm screen to separate the finest particles from the coarser material. This material is referred to as 'fine pulp'. The third group, designated 'milled pulp', was milled in a Wiley mill through a 1 mm screen.

NIR spectroscopy

NIR spectra for each of the four sample preparation methods were measured in a spinning sample holder in a NIRSystems Inc. Model 5000 scanning spectrophotometer. The spectra were collected in diffuse reflectance mode using Vision® software (version 3.1) at 2 nm intervals over the wavelength range 1100-2500 nm. Fifty scans were accumulated for each sample, and the results were averaged. After the spectrum was obtained, the sample cup was emptied, repacked, and a duplicate spectrum was obtained. The data were then imported into the Unscrambler® software (version 9.2) and the duplicate spectra were averaged. The spectra were converted to the second derivative using the Savitsky-Golay convolution algorithm with left and right gaps of 8 nm (Savitsky and Golay 1964).

Development of calibrations and predictions

The samples were randomly divided into a calibration set approximately two-thirds (40 samples) and a prediction set of one-third (19 samples) (Table 5.2). Calibrations for each of the four sample preparation methods described were created for each of the six carbohydrates. These calibrations were developed using Partial least

squares (PLS) regression with four cross-validation segments and a maximum of 10 factors.

We observed that the optimum number of factors recommended for each carbohydrate calibration was different as the sample preparation method changed (Table 5.2). To facilitate the comparison among calibrations and predictions developed using different preparation methods, the optimum number of factors was held constant by averaging the optimum number of factors determined for each calibration. These calibrations (based on the average number of factors) were then used to predict the values of the samples in their respective prediction sets.

Data analyses

Several statistics were used to assess the performance of the calibrations. The standard error of calibration (SEC) (determined from the residuals from the final calibration), standard error of cross-validation (SECV) (determined from the residuals of each cross-validation phase) and the coefficients of determination (R^2) were obtained. The coefficient of determination for calibration (R^2_c) was determined from the residuals of the calibration, while the coefficient of determination for cross-validation (R^2_v) was determined using residuals of the cross-validation phases. The ratio of performance to deviation (RPD_c), calculated as the ratio of the standard deviation of the reference data to the SECV, was utilized for direct comparisons of calibrations developed for different carbohydrates that have different ranges in values. Similar statistics were used to assess the predictive capability of the calibrations. The standard error of prediction (SEP), the coefficient of determination for prediction (R^2_p) values, and the RPD_p value (ratio of the standard deviation of the reference data to the SEP) were obtained for the prediction sets.

The statistics mentioned above are widely used in the field of NIR spectroscopy to compare calibration performances. However, one goal of this study was to develop a more impartial approach for comparing the predictive capabilities of different calibrations. Thus, we compared the models using a three-way random-effects ANOVA using SAS® software (version 9.1). The factors included in the analysis were the method of sample preparation, the type of carbohydrate, and sample effect. The sample effect, α_i, was considered a random effect and handles the fact that all of the methods are tested on the same set of samples. The least-squared means for each method were compared using a Tukey test. The following model was used for the analysis:

$$(y - \hat{y})^2 = \mu + \alpha_i + \beta_j + \gamma_k + \beta\gamma_{ik} + \varepsilon_{ijk}$$

Where $(y - \hat{y})^2$ = the square of the difference between the measured carbohydrate value and the predicted carbohydrate value

μ = the overall mean

α_i = the effect of sample i

β_j = the effect of method j

γ_k = the effect of carbohydrate k

$\beta\gamma_{ik}$ = the interaction effect of method j and carbohydrate k

ε_{ijk} = a random effect

RESULTS

Calibrations statistics varied widely for the six carbohydrates of interest. The calibrations for rhamnose and xylose were the best (Figs. 6.1-6.3). The R^2_c ranged from 0.81 to 0.86 for rhamnose and from 0.78 to 0.84 for xylose (Fig. 6.1). R^2_v were only slightly lower—between 0.77 and 0.82 for both rhamnose and xylose (Fig. 6.2). The

RPD_c values were also good for these carbohydrates, but the values for xylose (2.04 to 2.38) were considerably better than those for rhamnose (1.79 to 1.87) (Fig. 6.3). The calibrations for arabinose, fucose, galactose, and glucose were considerably weaker. Although the R^2_c for some of these carbohydrates were reasonable (Fig. 6.1), RPD_c were low, ranging from 0.94 for the coarse pulp fucose calibration to 1.36 for the milled pulp arabinose calibration.

The calibrations were used to predict the carbohydrate contents of the 19 samples in the prediction sets. The prediction statistics were much more variable than the calibrations statistics. Again, rhamnose and xylose provided the best statistics; R^2_p ranged from 0.83 to 0.89 for rhamnose and from 0.79 to 0.82 for xylose (Fig. 6.4). The R^2_p for the remaining carbohydrates were considerably lower. They were as low as 0.01 to 0.23 for fucose, and as high as 0.50 to 0.60 for glucose (Fig. 6.4).

RPD_p showed significant differences among the four sample preparation methods for each carbohydrate. RPD_p for rhamnose from handsheets was only 1.08, but the values for coarse pulp (2.44), fine pulp (2.41), and milled pulp (2.27) were all considerably higher (Fig. 6.5). For xylose, the values from handsheets (0.98) and milled pulp (0.80) were much lower than the values from coarse pulp (2.01) and fine pulp (1.87). Coarse pulp and fine pulp provided the highest RPD_p for all six carbohydrates (Fig. 6.5).

A three-way random effects ANOVA analysis of the predictions revealed that all of the fixed effects—preparation method, type of carbohydrate, and the method*carbohydrate interaction—were statistically significant (Table 6.3). The significant interaction indicates that the methods performed differently among the

carbohydrates. The statistical analyses also tested which sample preparation methods resulted in estimates of least mean squares (LMS) that were significantly different from zero. The null hypothesis (H_0) used for this test was LMS = 0, while the alternate hypothesis (H_A) was that LMS ≠ 0. The analysis determined that the estimates for fine pulp, milled pulp, and handsheets were significantly different from zero, but the estimate of LMS for coarse pulp was not significantly different from zero (α=0.05) (Table 6.4).

Finally, the estimates of LMS for each of the four sample preparation methods were compared, for a total of six comparisons. Differences in the estimates were analyzed with a t-test. No significant differences were found among coarse pulp, fine pulp, and handsheets. However, milled pulp was significantly different from each of these the methods (α=0.05) (Table 6.5).

DISCUSSION

The prediction statistics demonstrate that four of the carbohydrates analyzed in this study—arabinose, fucose, galactose, and glucose—could not be reliably estimated using NIR spectroscopy. However, successful calibrations were created for rhamnose and xylose. These results differ from the findings of Wallbacks et al. (1991b), who found that glucose was very well-modeled by NIR analysis, with xylose and galactose also producing strong correlations. Although they also measured arabinose and mannose, they did not report their findings for these carbohydrates.

Fardim et al. (2002) similarly found that xylose and glucose produced good calibrations with NIR analysis. However, their calibrations performed quite poorly in prediction. The root-mean square error for prediction (RMSEP) for glucose (2.03) was nearly 25 times the root-mean square error for calibration (RMSEC) (0.082). The

predictions for xylose were better, but the RMSEP (0.87) was still more than five times the RMSEC (0.165). Although the researchers also measured the mannose, arabinose, galactose, and rhamnose content of the pulps, they did not report NIR analyses of these carbohydrates.

Based on the previous studies, it was expected that the xylose calibrations would perform reasonably well. However, it is surprising that the glucose calibrations in this study did not perform better, given that glucose accounts for such a large percentage of the total carbohydrates (avg. 78.2%). This is likely due to the small degree of variability contained in the data set. Values for glucose content ranged only from 75.3% to 80.0% (Table 6.2). Perhaps the biggest surprise, though, is the excellent calibration and prediction statistics for rhamnose. Rhamnose composed, on average, only 0.022% of the total mass of the pulps in this study. It was not particularly variable, with values ranging from 0.014% to 0.028% (Table 5.2). Fucose and galactose were both considerably more variable than rhamnose, yet their calibrations were much poorer.

Of the four sample preparation methods that were compared in this study, only two have been reported previously. Wallbacks et al. (1991b) used milled pulp, while Fardim et al. (2002) used handsheets to collect spectra. It was unexpected, therefore, that the other two sample preparation methods – coarse pulp and fine pulp – consistently resulted in the best calibration statistics. Although a statistical significance test did not find these methods to be significantly better than handsheets (Table 6.4), this may be due to the small size of our prediction set (19 samples). Milled pulp, which provided good results for previous researchers (Wallbacks et al. 1991b) was significantly performed significantly worse than both coarse and fine pulp. This result is quite important, given

the additional time and equipment required to mill pulp into a fine powder. However, it is also quite surprising, as milled pulp is much more uniform in texture than coarse or fine loose pulp. It is possible that because the small size of the sieve openings (1 mm) allowed only the finest particles to pass through, the portions that we collected the spectra from were not representative of the entire samples and led to poor estimates. Because coarse pulp requires the least preparation, we recommend that this sample preparation technique be used for pulp carbohydrate analysis by NIR spectroscopy.

CONCLUSIONS

Four different sample preparation techniques were used to analyze the carbohydrate content of unbleached eucalyptus pulps with NIR spectroscopy. Models with excellent predictive capability were obtained for rhamnose and xylose, but models for arabinose, fucose, galactose, and glucose did not perform as well. Coarse and fine pulp preparation methods consistently yielded the best calibration and prediction statistics, although they were not significantly better than prediction statistics for handsheets. Milled pulp produced the worst prediction statistics. Given its ease of preparation, coarse pulp is the recommended method for collecting spectra for carbohydrate content analysis by NIR spectroscopy.

LITERATURE CITED

Fardim, P., and Duran, N. 1999. Wood and pulp carbohydrate analysis using HPLC and electrochemical detection *In* 6th Brazilian Symposium on the Chemistry of Lignins and Other Wood Components. *Edited by* F.T. Silva, Ferraz, A, Paiva, T.C.B. Faculdade de Engenharia Quimica de Lorena, Guarantingueta, SP, Brazil. pp. 167-172.

Fardim, P., Ferreira, M.M.C., and Duran, N. 2002. Multivariate calibration for quantitative analysis of eucalypt kraft pulp by NIR spectrometry. J. Wood Chem. Technol. 22(1): 67-81.

Kleen, M., and Gellerstedt, G. 1991. Characterization of chemical and mechanical pulps by pyrolysis-gas chromatography/mass spectrometry. J. Anal. Appl. Pyrolysis 19(39): 139-144.

Molin, U., and Teder, A. 2002. Importance of cellulose/hemicellulose ratio for pulp strength. Nord. Pulp Paper Res. J. 17(1): 14-19.

Pham, P.L., Alric, I., and Delmas, M. 1995. Incorporation of xylanase in total chlorine free bleach sequences using ozone and hydrogen peroxide. Appita J. 48(3): 213-217.

Raymond, C.A., Schimleck, L.R., Muneri, A., and Michell, A.J. 2001. Nondestructive sampling of *Eucalyptus globulus* and *E. nitens* for wood properties. III. Predicted pulp yield using Near Infrared Reflectance Analysis. Wood Sci. Technol. 35: 203-215.

Saeman, J.F., Moore, W.E., Mitchell, R.L., and Millet, M.A. 1954. Techniques for the determination of pulp constituents by quantitative paper chromatography. Tappi J. 37(8): 336-343.

Savitsky, A., and Golay, M.J.E. 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 38(8): 1627-1639.

Schimleck, L.R., Jones, P.D., Clark, A., Daniels, R.F., and Peter, G.F. 2005a. Near infrared spectroscopy for the nondestructive estimation of clear wood properties of *Pinus taeda* L. from the southern United States. Forest Prod. J. 55(12): 21-28.

Schimleck, L.R., Kube, P.D., Raymond, C.A., Michell, A.J., and French, J. 2005b. Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. Can. J. For. Res. 35: 2797-2805.

Schimleck, L.R., Mora, C., and Daniels, R.F. 2003. Estimation of physical wood properties of green *Pinus taeda* radial samples by near infrared spectroscopy. Can. J. For. Res. 33: 2297-2305.

Schimleck, L.R., Mora, C., and Daniels, R.F. 2004. Estimation of tracheid morphological characteristics of green *Pinus taeda* L. radial strips by near infrared spectroscopy. Wood Fiber Sci. 36(4): 527-535.

Schimleck, L.R., Rezende, G.D.S.P., Demuner, B.J., and Downes, G.M. 2005c. Estimation of whole-tree wood quality traits using near infrared spectra from increment cores. Appita J. 59(3): 231-236.

Schimleck, L.R., Sturzenbecher, R., Mora, C., Jones, P.D., and Daniels, R.F. 2005d. Comparison of *Pinus taeda* L. wood property calibrations based on NIR spectra from the radial-longitudinal and radial-transverse faces of wooden strips. Holzforshung 59: 214-218.

Turner, J.C., Sherker, P.S., Burns, B.J., Howard, J.C., Alonso, M.A., and Andres, J.L. 1992. Bleaching with enzymes instead of chlorine-mill trials. Tappi J. 75(12): 83-89.

Wallbäcks, L., Edlund, U., and Norden, B. 1991a. Multivariate characterization of pulp, Part 1. Spectroscopic characterization of physical and chemical differences between pulps using C CP/MAS NMR, FT-IR, NIR and multivariate data analysis. Nord. Pulp Paper Res. J. 2: 74-80.

Wallbäcks, L., Edlund, U., Norden, B., and Burgland, I. 1991b. Multivariate characterization of pulp using solid-state ^{13}C NMR, FTIR, and NIR. Tappi J. 74(10): 201-206.

Table 6.1. Statistics for the carbohydrate content and Kappa numbers of the pulps used, separated by origin

Species	No. of Samples	Statistic	Arabinose (%)	Fructose (%)	Galactose (%)	Glucose (%)	Rhamnose (%)	Xylose (%)	Kappa
		Average	0.0321	0.0078	0.152	76.6	0.0264	18.5	18.1
E. dunnii	**5**	**St.Dev.**	0.00308	0.00105	0.0360	1.47	0.00167	0.77	0.52
		Minimum	0.0280	0.0064	0.119	75.3	0.0239	17.3	17.2
		Maximum	0.0352	0.0092	0.201	79.1	0.0280	19.4	18.5
		Average	0.0389	0.0085	0.257	77.1	0.0239	18.2	17.6
E. globulus	**5**	**St.Dev.**	0.00348	0.00191	0.0554	0.56	0.00120	1.21	0.41
		Minimum	0.0370	0.0061	0.183	76.3	0.0225	16.5	17.2
		Maximum	0.0420	0.0110	0.310	77.6	0.0255	19.5	18.0
		Average	0.0327	0.0068	0.165	79.0	0.0208	15.4	17.8
E. grandis	**5**	**St.Dev.**	0.00337	0.00131	0.0408	0.69	0.00263	1.27	0.45
		Minimum	0.0293	0.0050	0.121	78.1	0.0184	13.4	17.3
		Maximum	0.0372	0.0081	0.206	79.8	0.0239	16.6	18.4
		Average	0.0378	0.0074	0.200	76.9	0.0256	18.7	17.6
E. maidenii	**5**	**St.Dev.**	0.00353	0.00222	0.0273	1.39	0.00270	1.44	0.67
		Minimum	0.0330	0.0048	0.161	75.3	0.0210	16.3	16.6
		Maximum	0.0414	0.0107	0.227	78.9	0.0282	20.2	18.1
		Average	0.0321	0.0066	0.192	79.0	0.0202	16.2	18.1
E. saligna	**4**	**St.Dev.**	0.00367	0.00178	0.0234	0.97	0.00145	1.00	0.48
		Minimum	0.0274	0.0046	0.162	77.7	0.1810	14.9	17.5
		Maximum	0.0361	0.0087	0.214	80.0	0.2150	17.4	18.5
		Average	0.0291	0.0088	0.237	78.7	0.0193	16.7	17.9
E. urophylla	**4**	**St.Dev.**	0.00194	0.00186	0.0270	0.47	0.00153	0.45	0.72
		Minimum	0.0265	0.0065	0.221	78.1	0.0174	16.0	17.5
		Maximum	0.0310	0.0110	0.277	79.2	0.0209	17.1	19.0
		Average	0.0319	0.0068	0.177	78.0	0.0231	17.1	17.7
Hybrids	**24**	**St.Dev.**	0.00429	0.00168	0.0362	1.04	0.00263	1.03	0.42
		Minimum	0.0248	0.0046	0.113	76.0	0.0183	15.2	16.6
		Maximum	0.0398	0.0095	0.258	80.0	0.0273	19.3	18.5
		Average	0.0274	0.0051	0.165	79.1	0.0162	15.8	16.2
Composite samples	**7**	**St.Dev.**	0.00191	0.00152	0.0325	0.69	0.00145	0.30	0.66
		Minimum	0.0247	0.0036	0.133	78.3	0.0140	15.4	15.2
		Maximum	0.0301	0.0071	0.225	80.0	0.0181	16.2	16.8

Table 6.2. Statistics for the carbohydrate content and Kappa number of the pulps used, separated by calibration and prediction sets.

	No. of Samples	Statistic	Arabinose (%)	Fructose (%)	Galactose (%)	Glucose (%)	Rhamnose (%)	Xylose (%)	Kappa
Calibration	**40**	**Average**	0.033	0.0071	0.192	78.2	0.022	17.0	17.6
		St. Dev.	0.0048	0.00200	0.0411	1.12	0.0035	1.40	0.70
		Minimum	0.025	0.0036	0.121	76.1	0.014	13.4	15.4
		Maximum	0.042	0.0110	0.306	80.0	0.028	19.5	19.0
Prediction	**19**	**Average**	0.032	0.0067	0.175	77.7	0.023	17.3	17.7
		St. Dev.	0.0043	0.00151	0.0518	1.44	0.0039	1.33	0.74
		Minimum	0.025	0.0043	0.113	75.3	0.016	15.2	15.2
		Maximum	0.042	0.0097	0.310	80.0	0.028	20.2	18.4

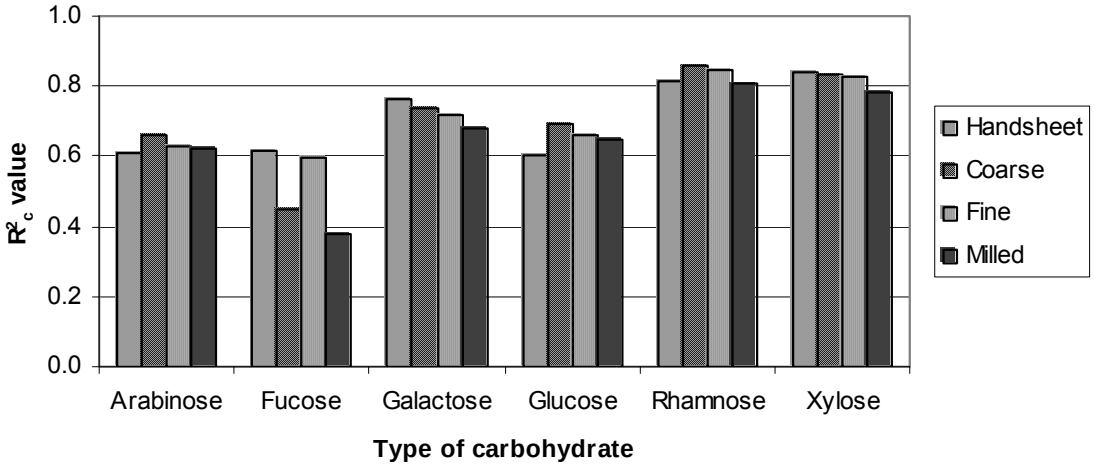


Figure 6.1. R^2_c values by carbohydrate type, grouped by sample preparation method.

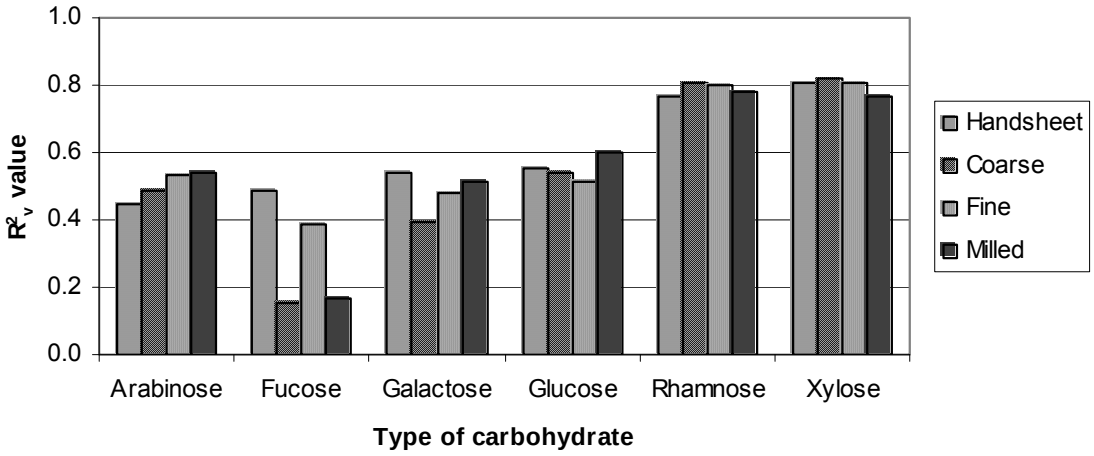


Figure 6.2. R^2_v values by carbohydrate type, grouped by sample preparation method

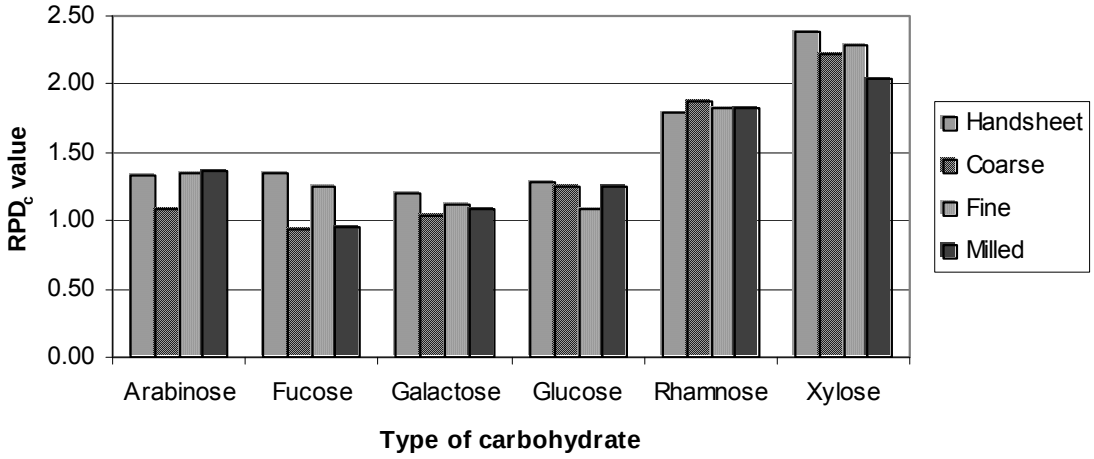


Figure 6.3. RPD_c values by carbohydrate type, grouped by sample preparation method.

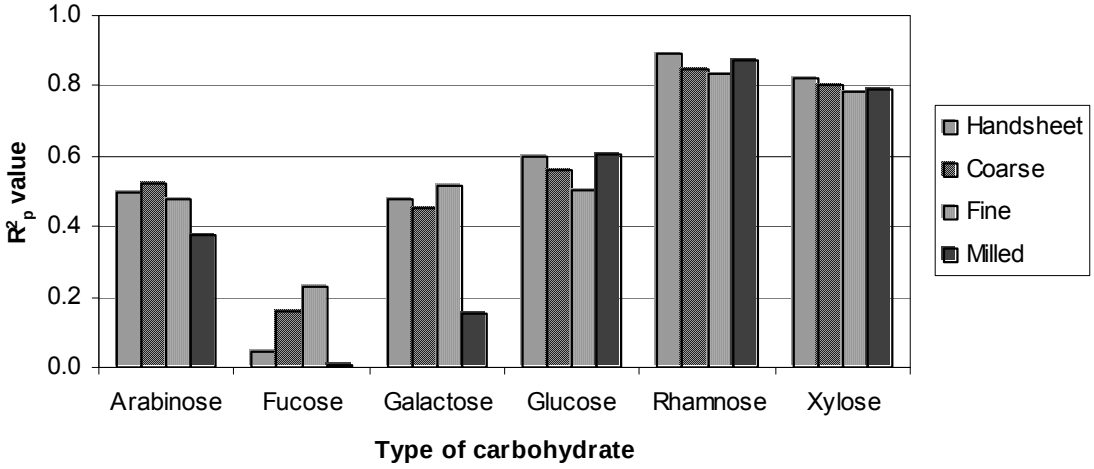


Figure 6.4. R^2_p values by carbohydrate type, grouped by sample preparation method.

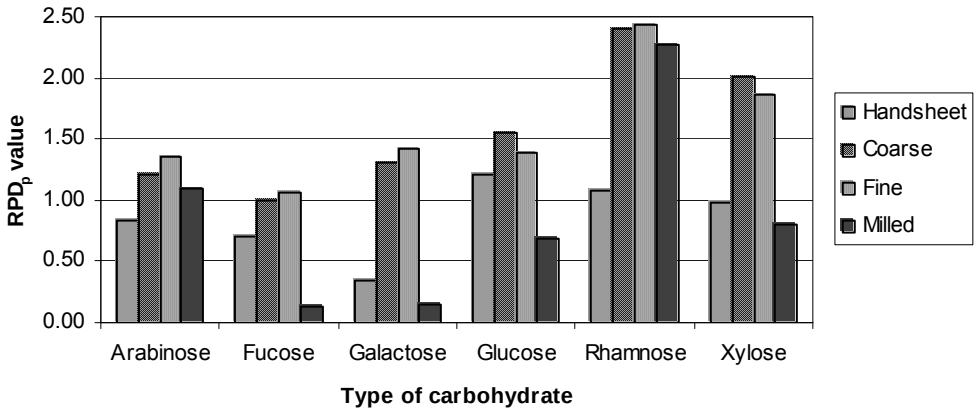


Figure 6.5. RPD_p values by carbohydrate type, grouped by sample preparation method.

Table 6.3. ANOVA table for the fixed effects of the model. All of the fixed-effect factors were statistically significant.

	Numerator degrees of freedom	Denominator degrees of freedom	F-value	P-value
Method	3	54	10.55	<0.0001
Carbohydrate	5	360	45.10	<0.0001
Method*Carbohydrate	15	360	7.92	<0.0001

Table 6.4. Estimates of the least mean squares (LMS) of the four sample preparation methods. A t-test was conducted to test the null hypothesis (H_0) that LMS = 0; the alternate hypothesis (H_A) is that LMS ≠ 0. Fine pulp, milled pulp, and handsheets produced least-squares means that were significantly different from zero (α=0.05), but coarse pulp did not.

Method	Estimate of LMS	T-value	P-value
Coarse	0.218	1.65	0.1044
Fine	0.265	2	0.0492
Milled	1.196	9.05	<0.0001
Handsheet	0.546	4.13	<0.0001

Table 6.5. Comparison of the estimates of LMS among the four sample preparation methods. There are a total of six comparisons. The t-values and p-values (α=0.05) are presented for each comparison.

Comparison	Difference between LMS's	t-value	p-value
coarse-fine	-0.047	-0.24	0.995
coarse-milled	-0.979	-4.99	<0.0001
coarse-handsheet	-0.328	-1.67	0.347
fine-milled	-0.931	-4.75	<0.0001
fine-handsheet	-0.281	-1.43	0.4843
milled-handsheet	0.650	3.31	0.009

Chapter 7

CONCLUSIONS

In our first study, we tested the applicability of existing NIR-based eucalyptus wood property calibrations from central Brazil to a separate population in southern Brazil and found they could not be used due to large predictive errors. When 2 to 6 samples from the southern population were included in the calibration set, the prediction statistics for lignin and pentosan content were greatly improved. Prediction errors were too large for density and pulp yield after adjustment. Adding more than 6 samples from the south was unnecessary and did not improve prediction statistics for any property. We also found that adjusting the calibrations did not diminish the predictive capability of the calibrations when applied to a subset of the central population. These findings show that NIR-based wood property calibrations can be made applicable separate populations by adding a small number of samples from the new regions to the original calibration set.

Next, we used NIR calibrations for eucalyptus wood properties that were created using milled increment cores to predict the wood properties of milled increment cores and drill bit shavings from 40 trees with unknown properties. The predicted properties for the two sample collection methods were used to create linear regressions for the wood properties. Confidence intervals for the slope and intercept estimates indicated that none of the wood properties were estimated accurately by milled shaving spectra. The R^2 for the plots were too low to consider using a linear regression to adjust the predictions. Spearman correlation coefficients (ρ) were very similar to Pearson correlation coefficients (R) for all properties, so the wood property rankings were also inaccurately estimated with spectra from shavings. This indicates that it is inadvisable to use milled

drill bit shavings in place of milled increment cores for predicting wood properties when using a calibration that was developed with milled increment cores. The error that would be introduced into the estimations would be unacceptable, even for ranking purposes.

In the next study, we used bleached eucalyptus pulps from the Aracruz's ECF pulp mill at Aracruz, Brazil to create NIR calibrations for ten physical and mechanical pulp properties. The calibrations created for these properties were poor and not useful for prediction purposes. Although previous researchers have had success in estimating several of the same properties, all of these studies used variable pulping methods for their samples to introduce variability. Because our samples were of mill-line origin, the variability of the samples was quite low, and this likely contributed to the poor calibration statistics. This study raises serious concerns about the ability to use NIR to create accurate calibrations for physical and chemical pulp properties in an industrial pulp mill setting. The tightly regulated pulping processes may reduce the variability of the pulps to such a degree that NIR analysis is unable to adequately model these properties.

In our final study, four different sample preparation techniques were used to analyze the carbohydrate content of unbleached eucalyptus pulps with NIR spectroscopy. Models with excellent predictive capability were obtained for rhamnose and xylose, but models for arabinose, fucose, galactose, and glucose did not perform as well. Coarse and fine pulp preparation methods consistently yielded the best calibration and prediction statistics, although they were not significantly better than prediction statistics for handsheets. Milled pulp produced the worst prediction statistics. Given its ease of preparation, coarse pulp is the recommended method for collecting spectra for carbohydrate content analysis by NIR spectroscopy.