NEAR INFRARED SPECTROSCOPY AS AN ALTERNATIVE TO EVALUATE EUCALYPT WOOD QUALITY AT ARACRUZ

Gabriel Dehon Sampaio Peçanha Rezende [1]
Laurence Schimleck [2]
Braz José Demuner [1]

[1] Aracruz Celulose S.A.-Centro de Pesquisa e Tecnologia, C.P. 33.1011, Aracruz-ES, 29.197-000, Brasil
[2] CSIRO Forestry and Forest Products – Private Bag 10, Clayton South MDC, Victoria 3169, Australia

ABSTRACT: Eucalypt wood quality traits have gained increasing importance in genetic improvement programs, especially those related to pulp production, since they significantly affect possibilities for productivity gains and product differentiation. In order for these characteristics to be effectively considered in the selection process, operational techniques that permit rapid, inexpensive and preferentially non-destructive evaluation of a large number of individuals in the field are fundamental. Near infrared spectroscopy (NIRS) has been shown to be a tool of great potential for this purpose, although it requires an initial effort to develop calibrations. Aware of the potential of NIRS for tree breeding purposes Aracruz began using it in 1998. Efficient calibrations have been obtained for pulp yield, specific wood consumption and pentosans, and it is already possible to predict these characteristics via NIRS. For lignin content, current calibrations are considered satisfactory, but can be improved.

INTRODUCTION

Forest breeding is the application of different quantitative genetic techniques, for the continual generation, evaluation and selection of superior clones and varieties, in terms of productivity and wood quality. In 1973 Aracruz started a eucalypt genetic improvement program that aimed to provide material better adapted to its environmental conditions. The first step was the introduction of several eucalypt species from Australia and Indonesia. After field evaluations, *E. grandis* and *E. urophylla* were identified as the most suitable, owing to their fast growth and adequate pulping properties. Since that time breeding populations of these species have been amplified and by the early 1980s improved pure and hybrid varieties were produced for Aracruz use, and for sale on a commercial basis to other companies. Aracruz also developed cloning on a commercial scale, and great attention was given to the selection and propagation of superior trees, both from the improvement program and commercial plantations. The clonal forests have provided significant improvements in wood productivity, uniformity and quality, and also provided Aracruz with a distinct competitive edge in the worldwide market.

Several recurrent selection methods are now used in order to produce advanced generations of *E. grandis*, *E. urophylla* and their hybrids. Special attention has been devoted to improving wood quality, maximizing yield of pulp per unit area, and allowing differentiation of products. In this context, other species with better pulping attributes have been tested in hybridization with our main genetic materials. Many basic studies have been carried out also, seeking fundamental parameters for the definition of selection methods to be used in the breeding program. Development of operational techniques that permit rapid, inexpensive and preferentially non-destructive evaluation of a large number of individuals in the field has become fundamental.

Near infrared spectroscopy (NIRS) has been shown to be a tool of great potential for this purpose (1, 2, 3, 4). NIRS has been widely used to rapidly estimate parameters which traditionally have been time consuming and difficult to measure. NIR spectroscopic analysis involves measuring the NIR spectra of a large number of samples, developing a regression calibration that links the spectra to the parameter of interest, and then using the calibration and the spectra of a new set of samples to predict that parameter (5, 6).

Aware of this potential, Aracruz has acquired a NIRS instrument and has obtained CSIRO assistance in the development of initial calibrations for application in the breeding program. The objective of this paper is to present the current results of this work.

MATERIALS AND METHODS

Samples

A total of 206 samples (trees) were available for calibration development. The samples were representative of 24 *E. grandis*, *E. urophylla* and hybrids clones, collected at three sites, Aracruz (70 samples), São Mateus (69 samples) and Bahia (67 samples), with ages varying between two and seven years. Air dry chips from each sample were milled in a Wiley mill and then in a Udy mill. Each sample was mixed and a sub-sample removed and placed in a NIRSystems small sample cup. 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 averaged. After the spectrum had been obtained, the sample cup was emptied, repacked and a duplicate spectrum obtained. The duplicate spectra were averaged using Vision® software (version 2.31) and converted to the second derivative. A segment width of 10 nm and a gap width of 20 nm were used for the conversion.

Calibrations development

Constituent values for calibrations development were obtained through mixed chips traditional lab analysis for the following traits: basic density (Tappi T258 om-94), screened pulp yield, specific consumption, total lignin (Tappi T222 om-98) and pentosans (Tappi T223 cm-84).

Calibrations for each trait were developed for all samples combined and for individual sites (Aracruz, São Mateus and Bahia) using average second derivative spectra. Calibrations between spectra and constituent values were obtained using Partial Least Squares (PLS) regression, a data decomposition technique that extracts the inherent variation that exists in the data set (7). In the course of calibrations development the sample sets were checked for outliers. Outlier detection was based both on the NIR spectra and on analysis of sample residuals (the residual = NIR fitted value - laboratory determined value). Outliers were detected in all sample sets for all traits and these samples were excluded from the calibration. The number of outliers detected for each situation was variable.

The measures of how well a calibration fits the data are the coefficient of determination (R^2) and the Standard Error of Calibration (SEC) (8, 9, 10), which is given by:

$$\text{SEC} = \sqrt{\frac{\sum_{i=1}^{NC} (\hat{y}_i - y_i)^2}{(NC - k - 1)}}$$

where $\hat{y}_i$ is the value of the constituent of interest for validation sample i estimated using the calibration, y_i is the known value of the constituent of interest of sample i, NC is the number of samples used to develop the calibration, and k is the number of factors used to develop the calibration. The factors (k) are obtained when PLS regression compresses the multidimensional (spectral) variation into a few dimensions (factors) that describe most of the variation present. The selection of how many factors to use for a calibration is very important. If not enough factors are selected, variation in the data will not be sufficiently described and underfitting will occur. Conversely, the selection of too many factors will overfit the data and give a calibration that may only be suited to the data from which it was derived (11).

In this study a measure of how well the calibration predicts the constituent of interest for a set of unknown samples that are different from the calibration test set is given by the Standard Error of Prediction (SEP) (8, 9, 10):

$$\text{SEP} = \sqrt{\frac{\sum_{i=1}^{NP} (\hat{y}_i - y_i)^2}{(NP - 1)}}$$

where $\hat{y}_i$ is the value of the constituent of interest for sample i predicted by the calibration, y_i is the known value of the constituent of interest for sample i, and NP is the number of samples in the prediction set.

RESULTS AND DISCUSSION

Calibrations developed for all samples combined for pulp yield, specific consumption and pentosans were very good, all having coefficients of determination (R^2) greater than 0.8. Calibrations developed for basic density and lignin were not as good, with R^2 of 0.61 and 0.67, respectively (Table 1).

Table 1. NIR calibration statistics for all samples combined for each trait measured by Aracruz.

Trait	# of factors	R^2	SEC
Basic density	3	0.61	21.50
Pulp yield	4	0.90	1.17
Specific consumption	6	0.88	0.19
Total lignin	4	0.67	0.76
Pentosans	5	0.88	0.55

The calibrations developed for pulp yield and pentosans compare well with those reported in other studies (1, 2, 12, 13, 14). As far as the authors aware there are no examples of specific consumption calibrations reported in the literature but Schimleck and French (15) have published calibrations for pulpwood productivity (where pulpwood productivity is equal to pulp yield multiplied by basic density), a comparable trait, and obtained similar R^2. The basic density calibration had a lower R^2 compared to most basic density calibrations published in the literature, but generally strong calibrations have been obtained using solid wood (16, 17, 18). Studies based on milled wood have provided both weaker (19) and stronger calibrations (15). The lignin calibration had a lower R^2 than expected based on calibrations published for this trait (12, 20). The lignin calibration was developed over a narrow range and this coupled with the wide genetic variation in the calibration set may explain the result. Analysis of residuals for the basic density and lignin calibrations showed that both traits are overestimated for low lab data and underestimated for high lab data.

Table 2. Individual site NIR calibration statistics for each trait measured by Aracruz.

Location	Trait	# of factors	R^2	SEC
Aracruz	Basic density	3	0.74	20.55
	Pulp yield	2	0.95	1.21
	Specific consumption	3	0.95	0.20
	Total lignin	3	0.69	0.75
	Pentosans	3	0.80	0.70
São Mateus	Basic density	4	0.71	20.80
	Pulp yield	4	0.91	1.01
	Specific consumption	5	0.80	0.19
	Total lignin	3	0.55	0.80
	Pentosans	3	0.88	0.48
Bahia	Basic density	3	0.65	18.09
	Pulp yield	2	0.83	1.22
	Specific consumption	4	0.89	0.17
	Total lignin	3	0.58	0.76
	Pentosans	3	0.86	0.61

Individual site calibrations for pulp yield, specific consumption and pentosans were again very good, all with R^2 greater than 0.8 (Table 2). The calibrations developed for basic density were also quite good with R^2 ranging from 0.65 to 0.74. Calibrations developed for total lignin content had R^2 that ranged from 0.55 to 0.69.

Calibrations developed for a specific site were used to predict traits on the other sites. For example, Aracruz calibrations were used on São Mateus and Bahia spectra. Estimates of pulp yield, specific consumption and pentosans all gave very good correlations. Correlations for total lignin and basic density were generally poor. An important criteria for assessing estimates made by a calibration is the SEP. An SEP close to the SEC of a calibration indicates that the estimates have errors similar in magnitude to the calibration. Generally pentosans was the only trait that had SEP results similar to the SEC. Some calibrations gave good R^2 results but high SEP results. For example when the Bahia calibration was used to estimate the pulp yield of the Aracruz samples an R^2 of 0.84 was obtained but the SEP (2.71) was higher than the Bahia calibration SEC (1.49). The Bahia calibration overestimated the pulp yields of the Aracruz samples because the Bahia calibration was developed with samples that had a higher average pulp yield.

These results suggest that once good calibrations are developed for a trait for a specific site of the company, it is not necessary to develop new calibrations for predictions on samples from other sites, particularly if the predictions will be used to rank trees. The similar genetic stock and environmental conditions across the sites make this possible. The good correlations obtained between sites may also justify the use of a single calibration for ranking clones across sites in the breeding program. Further studies on the matter are needed.

FINAL REMARKS

Results of calibrations for important wood characteristics, such as pulp yield, specific wood consumption, wood density, pentosans and lignin content, collected for different specific environment conditions and genetic materials of Aracruz Celulose S. A., have confirmed the wide potential of NIRS to be used as a rapid, inexpensive and non-destructive tree evaluation tool.

Additional effort is being applied in order to improve calibrations before starting routine predictions.

REFERENCES

1. 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), 164-166.

2. Michell, A. J. 1995. Pulpwood quality estimation by near-infrared spectroscopic measurements on eucalypt woods. *Appita J.* 48(6), 425-428.

3. Clarke, C. R. E.; Shaw, M. J. P.; Wessels, A. M.; Jones, W. R. 1999. Effect of differences in climate on growth, wood, and pulp properties of nine eucalypt species at two sites. *Tappi J.* (82)7: 89-99.

4. Schimleck, L. R., Raymond, C. A., Beadle, C. L., Downes, G. M., Kube, P. D. and French, J. 2000. Applications of NIR spectroscopy to forest research. *Appita J.* 53(6), 458-464.

5. Martens, H. and Næs, T. 1984. Multivariate calibration. 1. Concepts and distinctions. *Trends Anal. Chem.* 3: 204-210.

6. Thomas, E. V. 1994. A primer on multivariate calibration. *Anal. Chem.* 66(15), 795A-804A.

7. Esbensen, K., Schönkopf, S. and Midtgaard, T. 1994. Multivariate Analysis in Practice. Camo AS, Trondheim.

8. Miller, C. E. 1989. Analysis of synthetic polymers by near-infrared spectroscopy. Ph.D. Thesis, University of Washington, USA.

9. Anon. 1990. Manual for NIR spectral analysis software. NIRSystems, Inc., Silver Spring, Maryland.

10. Workman, J.J., Jr 1992. NIR spectroscopy calibration basics. *In* Handbook of Near-Infrared Analysis. *Edited by* D.A. Burns and E.W. Ciurczak. Marcel Dekker Inc., New York. pp. 274-276.

11. Martens, H. and Næs, T. 1991. Multivariate Calibration. John Wiley and Sons, Chichester.

12. Mroczyk, W. B., Kasprzyk, H. and Gawecki, T. 1993. Quantitative near infrared analysis of wood and the use of spectrophotometer infraanalyzer 450. *Rocz. Akad. Roln. Poznaniu* 250, 15-20.

13. Schimleck, L. R., Wright, P. J., Michell, A. J. and Wallis A. F. A. 1997. Near-infrared spectra and chemical compositions of *Eucalyptus globulus* and *E. nitens* plantation woods. *Appita J.* 50(1), 40-46.

14. Michell, A. J. and Schimleck, L. R. 1998. Developing a method for the rapid assessment of pulp yield of plantation eucalypt trees beyond the year 2000. *Appita J.* 51(6), 428-432.

15. Schimleck, L. R. and French, J. 2001. Application of NIR spectroscopy to plantation eucalypt samples with low pulp yield variation. 55th Appita Gen. Conf., Hobart, Australia, 30 April - 2 May, 2001, 59-65.

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

17. Hoffmeyer, P., and Pedersen, J.G. 1995. Evaluation of density and strength of Norway spruce by near infrared reflectance spectroscopy. *Holz Roh Werkst.* 53: 165-170.

18. Schimleck, L. R., Evans, R. and Ilic, J. 2001. Estimation of *Eucalyptus delegatensis* clear wood properties by near infrared spectroscopy. Accepted *Can. J. For. Res.*

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

20. 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.