MULTIVARIATE REGRESSION METHODS FOR ESTIMATING BASIC DENSITY IN *Eucalyptus* WOOD FROM NEAR INFRARED SPECTROSCOPIC DATA

Paulo Ricardo Gherardi Hein[1]

ABSTRACT: Near infrared (NIR) spectroscopy is a fast and efficient technique to predict a range of wood traits; however, methods for extracting useful information from the NIR spectra could be improved. Thus, the aim of this study was to evaluate the statistic performance of two regression methods for estimating the basic density in *Eucalyptus urophylla* x *grandis* wood from near infrared spectroscopic data. The predictive models calibrated by principal component regression (PCR) or partial least square regression (PLSR) method provided fine correlations. The coefficients of determination (R^2cv) of the PCR models ranged from 0.78 to 0.85 with standard error of cross-validation (SECV) and the ratio of performance to deviation (RPD) varying from 32.8 to 41.2 kg/m³ and from 1.6 to 1.9, respectively. The PLSR models presented R^2cv with relatively lower magnitude (from 0.65 to 0.78); but also lower SECV (from 29.8 to 38.9 kg/m³) and higher RPD values (from 1.6 to 2.1). In short, PCR method provides higher R^2 between Lab-measured and NIR-predicted values while PLSR produces lower standard errors of cross-validations. For both regression methods, the pre-treatments on NIR spectra, and the wavelength selection improved the calibration statistics, reducing the SECV and increasing the R^2cv and the RPD values. Thus, PCR and PLS regression can be applied successfully for predicting basic density in *Eucalyptus urophylla* x *grandis* wood from the near infrared spectroscopic data.

Key words: Chemometrics, calibration, near infrared spectroscopy, wood, basic density, *Eucalyptus.*

MÉTODOS DE REGRESSÃO MULTIVARIADA PARA ESTIMATIVA DA DENSIDADE BÁSICA DA MADEIRA DE *Eucalyptus* POR ESPECTROSCOPIA NO INFRAVEMELHO PRÓXIMO

***RESUMO**: A espectroscopia no infravermelho próximo (NIR) é uma técnica rápida e eficiente para estimar características da madeira, no entanto, métodos para extrair as informações úteis a partir do NIRS podem ser melhorados. Assim, este estudo foi realizado com o objetivo de avaliar o desempenho estatístico de dois métodos de regressão para estimar a densidade básica da madeira de Eucalyptus urophylla x grandis a partir da espectroscopia no NIR. Os modelos preditivos calibrados pelos métodos da regressão de componentes principais (PCR) ou regressão de mínimos quadrados parciais (PLSR) apresentaram boas correlações. Os coeficientes de determinação (R^2cv) dos modelos PCR variaram de 0,78 a 0,85, com erro padrão na validação cruzada (SECV) e a relação do desempenho de desvio (RPD) variando de 32,8-41,2 kg/m³ e 1,6-1,9 kg/m³, respectivamente. Os modelos PLSR apresentaram R^2cv com magnitude relativamente menor (de 0,65 a 0,78), mas também menores valores de SECV (29,8-38,9 kg/m³) e maiores de RPD (de 1,6 a 2,1 kg/m³). De maneira geral, o método de PCR oferece maior R^2 entre os valores medidos e estimados, enquanto a PLSR produz menor erro padrão de validação cruzada. Para ambos os métodos de regressão, o pré-tratamento dos espectros NIR e a seleção de comprimento de onda melhoraram as estatísticas de calibração, reduzindo o SECV e aumentando os valores de R^2cv e de RPD. Assim, as regressões PCR e PLS-R podem ser aplicadas com sucesso para a estimativa da densidade básica da madeira de Eucalyptus por espectroscopia no infravermelho próximo.*

Palavras-chave: Quimiometria, calibração, espectroscopia no infravermelho próximo, madeira, densidade básica, Eucalyptus.

1 INTRODUCTION

Forest-based industries, such as the ones which produce pulp and paper, steel and timber, in general require fast and accurate methods that could be applied in a large number of samples to characterize their raw materials. Near infrared (NIR) spectroscopy has been widely used on forestry researches. Several studies have showed that this technique can be applied for estimating the main wood characteristics, such as physical (HOFMEYER, 1995), mechanical (VIA et al., 2003), anatomical (SCHIMLECK; EVANS, 2004; VIANA et al., 2009) and chemical traits (BAILLÈRES; DAVRIEUX; HAM-PICHAVANT, 2002), both in solid, as in ground-meal wood (HEIN; LIMA; CHAIX, 2010) with reasonable accuracy.

The chemometrics is a truly interdisciplinary science that combines technical mathematics, statistics and information science (GEMPERLINE, 2006) in order to extract relevant information from NIR spectra. The combination of NIR spectra (instrumentally measured) with the wood properties determined by the conventional method (laboratory analysis) is made through regression techniques (modeling) and makes it possible to predict

[1]PhD candidate at University of Montpellier 2. CIRAD – PERSYST Department – TA A-40/16 – Production and Processing of Tropical Woods, 73 rue Jean-François Breton – 34398 – Montpellier, France – phein1980@gmail.com

these traits on unknown samples (NÆS et al., 2002). The use of mathematical models based on NIR spectroscopic data to quickly and simultaneously predict a range of wood traits, which are generally costly and time-consuming to be measured, results in significant gains in breeding programs.

There are many chemometric tools to examine and to extract information contained within NIR spectra. According to Dardenne, Sinnaeve and Vaeten (2000), among the most currently used calibration (or regression) procedures are Multiple Linear Regression (MLR), Principal Component Regression (PCR), Partial Least Square Regression (PLS), Local Weighted Regression (LWR), Ridge Regression and regression methods based on Artificial Neural Network (ANN) methodology. Pasquini (2003) affirms that the most common methods for NIR spectroscopic data are MLR, PCR and PLS regression. According to this author, all these tools presuppose a linear relationship between the spectra information and the concentration or property to be determined.

The main advantage of PCR and PLS method is their ability to compress the relevant spectral information into a few latent variables; the orthogonality between these variables ensures the stability of the obtained model (JOUAN-RIMBAUD; MASSART; NOORD, 1996). In order to analyze NIR spectra information, PLS regression is of particular interest because, unlike the MLR or PCR, it can analyze the data strongly collinear (correlated), with noise, with many variables-X, and also, simultaneously, several variables can shape-response Y (WOLD; SJÖSTRÖM; ERIKSSON, 2001).

Thus, the aim of this study was to compare the statistic performance of two regression methods for the estimation of wood basic density of *Eucalyptus urophylla* x *grandis* from near infrared spectroscopic data. In addition, we verified the influence of the application of mathematical treatment, the disposal of outlier and the wavelength selection on the models performance. The calculus method of the models and their assumptions are beyond the scope of this study. For further details, we recommend the paper of Geladi and Kowalski (1986) and of Wold, Sjöström and Eriksson (2001), wherein a complete tutorial about PLS regression are presented. The PCR method is widely discussed in many books (GEMPERLINE, 2006).

2 MATERIAL AND METHODS

2.1 Timber used in the simulations of NIR calibration

Simulations of calibration and optimization of the NIR models were performed on 140 samples of 6.5-year-old *Eucalyptus* wood (25 mm x 25 mm x 25 mm) from commercial forests in Curvelo, in the central region of Minas Gerais State, Brazil. The material used in this work was previously characterized by Hein et al. (2009) and presented averaged basic density of 526 kg/m³ ranging from 407 to 708 kg/m³, with coefficient of variation of 11.7%.

2.2 NIR spectra measurement

Bruker spectrophotometer (MPA model, Bruker Optik GmbH, Ettlingen, Germany) was used in the diffuse reflectance mode. A sintered gold standard was used as background. Spectral analysis was performed within the 12,500-3,500 cm^{-1} (or 800-2,850 nm) range at 8 cm^{-1} resolution (each spectrum consisted of 2,335 absorption values). The spectra were measured directly on both transversal faces of each sample, and each spectrum is represented by the average of thirty-two scans. The spectrum used for calibration was obtained by the arithmetic average of two readings (superior and inferior face). All samples were kept in a conditioned room with 60% relative humidity and temperature of 20°C before NIRS analysis. Under these conditions, the equilibrium moisture content was 14%.

2.3 Calibration parameters

Partial least squares regression (method PLS-1) and principal component regression were adjusted to describe the relationship between the values of basic density of the wood and near infrared spectra, by means of the software of statistical multivariate The Unscrambler® (v.9.7). The calibrations were adjusted with a maximum of 12 latent variables (LV) for calibration by the partial least squares regression and 12 principal components (PC) for principal components regression. The NIR calibrations were performed in full cross-validation mode. First (13-point filter and a second order polynomial) and second

derivatives (25-point filter and a third order polynomial) were applied on the NIR spectral data by the Savitsky and Golay (1964) algorithm to enhance the quality of the calibrations. The Martens' uncertainty test (WESTAD; MARTENS, 2000) was used to select the wavelength with significant regression coefficients. For comparison purposes, we used five latent variables or principal component for each model. The student residual and leverage values plot was examined in order to detect outlier (HEIN et al., 2009) and the two detected outliers were not included during the calibration of models.

2.4 Calibration routines

The calibrations were performed using three routines: routine 1 - calibration with all samples and with all variables (wavelength); routine 2 - calibration without outlier and with all the variables (wavelength) and routine 3 - calibration without outliers and with automatic selection of variables (wavelength) by Martens' uncertainty test (WESTAD; MARTENS, 2000). For comparative purposes, the results presented in Tables 1 and 2 refer to calibrations processed with five factors (PC's or PLS). It can find calibrations with performance higher than showed; however, to facilitate comparison of predictive models, some parameters were standardized, as the number of latent variables (LV) for PLS-R or principal components (PC) for PCR and the number of outliers.

2.5 Rank of NIR-based models

The criteria adopted to select the model prediction were: (a) coefficient of determination of the model in cross-validation (R^2cv), (b) standard error of cross-validation (SECV), (c) ratio of performance to deviation (RPD) and (d) number of LV for PLS-R or PCR for PCR models. The SECV measures the efficiency of the model calibration in predicting the basic density of the wood in a lot of unknown wood, different from samples in the consignment of calibration. The formula is given by Schimleck, Evans and Ilic (2001). The RPD value is the relationship between the standard deviation of the values of basic average density measured by the conventional method (SD) and the standard error of cross-validation (SECV). For use in the forest science, Schimleck, Evans and Ilic (2003) claim that a RPD greater than 1.5 is considered satisfactory for screenings and preliminary predictions while Schimleck and Evans (2004) consider a RPD of 2.5 sufficient for trees selection in breeding programs.

3 RESULTS AND DISCUSSION

3.1 NIR-based calibrations

Table 1 shows the statistics associated with calibrations adjusted by partial least squares regression for basic density of the wood. Delwiche and Reeves (2004) showed that the application of pre-treatments on the spectra helps to optimize the model. Thus, these models were developed from the raw NIR spectra and after pre-treatment (1st and 2nd derivative). Similarly, Leardi, Seasholtz and Pell (2002) presented a study indicating that the selection of wavelength in calibration is an important procedure to improve the statistic models. Ghasemi, Niazi and Leardi (2003) used genetic algorithms to select the wavelength for PLS regressions. The routine 3 shows the effect of variable selection on the calibration statistics.

The PLS-R calibrations (Table 1) provided models with R^2cv values ranging from 0.65 to 0.78 using 5 latent variables. These statistics are consistent with those reported by other studies. For instance, Thygesen (1994) assessed the basic density of *Picea abies* wood reporting PLS calibrations with R^2 ranging from 0.504 to 0.792 using from six to eight factors. Investigating 8 years-*Eucalyptus globulus* with basic density between 378 and 656 kg/m³, Schimleck, Evans and Ilic (1999) calibrated models with R^2 ranging from 0.384 to 0.593 using from 4 to 10 latent variables. Jiang et al. (2007) evaluated crystallinity in 20 years-*Pinus elliotii* reporting PLS models based on the 1st derivative spectra with R^2p value of 0.86 and RMSEP of 3.9% while the 2nd derivative spectra produced inferior statistics. These authors also used wavelength selection increasing their R^2c and R^2p values and decreasing RMSEC and RMSEP values.

It can be observed in Table 2 the statistics associated with PCR calibrations for basic density of the wood. The PCR models provided R^2cv between Lab-measured and NIR-predicted values ranging from 0.78 to 0.85; SECV values from 32.8 to 41.2 kg/m³ and RPD values from 1.6

Table 1 – NIRS calibrations for basic density of the *Eucalyptus* wood by partial least squares regression from the raw (-) and pre-treated (first -1d and second derivative -2d) NIR spectra.

Tabela 1 – *Calibrações NIRS para densidade básica da Madeira de Eucalyptus por regressão dos mínimos quadrados parciais a partir dos espectros originais e pré-tratados (primeira -1d e segunda -2d derivada).*

Treat	Routine	R²c	R²cv	LV	SEC	SECV	Outlier	N	RPD
	1	0.71	0.66	5	34.7	37.3	-	140	1.7
-	2	0.77	0.70	5	30.5	34.8	2	138	1.8
	3	0.77	0.74	5	30.7	32.3	2	138	2.0
	1	0.78	0.65	5	27.8	38.9	-	140	1.6
1d	2	0.82	0.74	5	26.6	32.1	2	138	2.0
	3	0.81	0.78	5	27.3	29.8	2	138	2.1
	1	0.80	0.66	5	28.6	37.5	-	140	1.7
2d	2	0.85	0.72	5	24.1	33.2	2	138	1.9
	3	0.83	0.75	5	25.6	31.1	2	138	2.1

Routine 1 - calibration with all samples and with all wavelength; routine 2 - calibration without outlier and with full wavelength and routine 3 - calibration without outlier and with selection of wavelength by Martens' uncertainty test (WESTAD; MARTENS, 2000)

Table 2 – NIRS calibrations for basic density of the *Eucalyptus* wood by Principal Component Regression from the raw (-) and pre-treated (first -1d and second derivative -2d) NIR spectra.

Tabela 2 – *Calibrações NIRS para densidade básica da Madeira de Eucalyptus por regressão dos componentes principais a partir dos espectros originais e pré-tratados (primeira -1d e segunda -2d derivada).*

Treat	Routine	R²c	R²cv	PC	SEC	SECV	Outlier	N	RPD
	1	0.80	0.78	5	38.3	39.8	-	140	1.6
-	2	0.84	0.82	5	34.3	36.2	2	138	1.8
	3	0.87	0.85	5	31.7	33.2	2	138	1.9
	1	0.79	0.77	5	39.7	41.2	-	140	1.6
1d	2	0.84	0.82	5	34.6	36.0	2	138	1.8
	3	0.84	0.82	5	34.4	36.3	2	138	1.8
	1	0.82	0.80	5	36.6	38.3	-	140	1.7
2d	2	0.86	0.84	5	32.5	33.9	2	138	1.9
	3	0.87	0.85	5	31.7	32.8	2	138	1.9

Routine 1 - calibration with all samples and with all wavelength; routine 2 - calibration without outlier and with full wavelength and routine 3 - calibration without outlier and with selection of wavelength by Martens' uncertainty test (WESTAD and MARTENS, 2000)

to 1.9. These R²cv were higher than those from PLS-R models, but the SECV values also were superior. Higher R² from PCR models when compared to the PLS-R ones were reported in other studies. For instance, Via et al. (2003) evaluated 41-year-old *Pinus* and used multiple linear regression (MLR) and PCR to estimate the basic density of wood. They found calibrations with R² ranging from 0.71 to 0.86 and RMSEC ranging from 43.2 to 51.0 kg m^{-3}.

3.2 What is the best way to improve the statistics model?

Many studies have been made to examine which regression method provides the best predictive performance. For instance, Dardenne, Sinnaeve and Vaenten (2000) studied NIR spectral databases obtained from agricultural products and food products and compared four regression techniques (MLR, PLS, ANN and LOCAL) .According to their results the differences between the methods are very small. Similarly, Blanco and Valdés (2004) used partial least square regression and principal component regression to study the influence of temperature on the predictive ability of near infrared spectroscopy models. They reports that PLS and PCR models provided similar results. However, considering PLS and PCR as standard methods, it was still not possible to determine which is the best one. Henriksen et al. (2004) compared different regression methods to predict important sulphite pulp properties from NIR spectra. Among the different model concepts as least angle regression, ridge regression, partial least square regression and principal component regression, according these authors, PLS is recommended. In the other hand, Chauchard, Roger and Bellon-Maurel (2004) tested MLR and PLS regression to evaluate the effect of the fruit temperature on sugar content prediction in apples. These authors advised that since MLR models use less variables, they appeared more robust than corresponding corrected PLS model and for this reason the stepwise MLR seems to be the most suited for on line process use.

On the other hand, Dardenne, Sinnaeve and Vaeten (2000) states the following: "instead of testing various regression methods, more work and time should be devoted to measure the spectra and in determining the reference values". This important statement sums up the message of the present study: to improve the statistic of the models, we have to increase the accuracy of the measurements of the NIR spectra and, mainly, of the reference values.

4 CONCLUSIONS

In this study, the statistic performance of two multivariate regression methods for estimating the basic density in *Eucalyptus urophylla* x *grandis* wood from near infrared spectroscopic data was evaluated. For PCR models, the R^2 between Lab-measured and NIR-predicted values ranged from 0.78 to 0.85, the SECV from 32.8 to 41.2 kg/m³ and the RPD from 1.6 to 1.9. The PLSR models presented Rcv^2 with relatively lower magnitude (from 0.65 to 0.78); but also lower SECV (from 29.8 to 38.9 kg/m³) and higher RPD values (from 1.6 to 2.1). In short, PCR method provides higher R^2 between Lab-measured and NIR-predicted values while PLSR produces lower standard errors of cross-validations. For both regression methods, the pre-treatments on NIR spectra and the wavelength selection improved the calibration statistics, reducing the SECV and increasing the R^2cv and RPD values. Thus, PCR and PLS regression can be applied successfully for predicting basic density in *Eucalyptus urophylla* x *grandis* wood from the near infrared spectroscopic data. However, more work and time should be devoted to measure the spectra and in determining the reference values in order to improve the model statistics.

5 ACKNOWLEDGEMENT

The author thanks to José Tarcísio Lima (UFLA) and Plantar Company S.A. for providing wood samples; to V&M Tubes Company for the assistance for NIR spectra measurements and to UPR39 and UPR40 of CIRAD (France) for providing technical support. The author was supported by the National Council of Technological and Scientific Development (CNPq, Brazil - process no. 200970/2008-9).

6 REFERENCES

BAILLÈRES, H.; DAVRIEUX, F.; HAM-PICHAVANT, F. Near infrared analysis as a tool for rapid screening of some major wood characteristics in a *Eucalyptus* breeding program. **Annals of Forest Science**, Les Ulis, v. 59, n. 5/6, p. 479-490, July/Oct. 2002.

BLANCO, M.; VALDÉS, D. Influence of temperature on the predictive ability of near infrared spectroscopy models. **Journal of Near Infrared Spectroscopy,** Sussex, v. 12, n. 2, p. 121-126, 2006.

BRINKMANN, K.; BLASCHKE, L.; POLLE, A. Comparison

of different methods for lignin determination as a basis for calibration of near-infrared reflectance spectroscopy and implications of lignoproteins. **Journal of Chemical Ecology**, New York, v. 28, n. 12, Dec. 2002.

CHAUCHARD, F.; ROGER, J. M.; BELLON-MAUREL, V. Correction of the effect on near infrared calibration - application to soluble solid content prediction. **Journal of Near Infrared Spectroscopy,** Sussex, v. 12, n. 3, p. 199-205, 2004.

DARDENNE, P.; SINNAEVE, G.; VAETEN, V. Multicalibration and chemometrics for near infrared spectroscopy: which method? **Journal of Near Infrared Spectroscopy,** Sussex, v. 8, n. 4, p. 229-237, 2000.

DELWICHE, S. R.; REEVES, J. B. The effect of spectral pretreatments on the pls modeling of agricultural products? **Journal of Near Infrared Spectroscopy,** Sussex, v. 12, n. 3, p. 177-182, 2004.

GELADI, P.; KOWALSKI, B. R. Partial Least-Square regression: a tutorial. **Analytica Chimica Acta,** Amsterdam, v. 185, n. 1, p. 1-17, 1986.

GEMPERLINE, P. **Practical guide to chemometrics.** 2. ed. Boca Raton: Taylor and Francis, 2006. 520 p.

GHASEMI, J.; NIAZI., A.; LEARDI, R. Genetic-algorithm-based wavelength selection in multicomponent spectrophotometric determination by PLS: application on copper and zinc mixture. **Talanta,** London, v. 59, n. 2, p. 311-317, Feb. 2003.

HEIN, P. R. G. et al. Near infrared spectroscopy for estimating wood basic density in *Eucalyptus urophylla* and *Eucalyptus grandis*. **Cerne,** Lavras, v. 15, n. 2, p. 133-141, abr./jun. 2009.

HEIN, P. R. G.; LIMA, J. T.; CHAIX, G. Effects of sample preparation on NIR spectroscopic estimation of chemical properties of *Eucalyptus urophylla* S.T. Blake wood. **Holzforschung,** Berlin, v. 64, n. 1, p. 45-54, 2010.

HEIN, P. R. G.; LIMA, J. T.; CHAIX, G. Robustness of models based on near infrared spectra to predict the basic density in *Eucalyptus urophylla* wood. **Journal of Near Infrared Spectroscopy,** Sussex, v. 17, n. 3, p. 141-150, 2009.

HENRIKSEN, H. C. et al. Prediction of important sulphite pulp properties from near infrared spectra: a feasibility study and comparison of methods. **Journal of Near Infrared Spectroscopy,** Sussex, v. 12, n. 5, p. 279-285, 2004.

HOFFMEYER P.; PEDERSEN, J. G. Evaluation of density and strength of Norway spruce by near infrared reflectance spectroscopy. **Holz als Roh und Werkstoff,** Berlin, v. 53, n. 1, p. 165-170, Jan. 1995.

JIANG, Z. H. et al. Rapid prediction of wood crystallinity in *Pinus elliotii* plantation wood by near-infrared spectroscopy. **Journal of Wood Science,** Madison, v. 53, n. 5, p. 449-453, 2007.

JONES, P. D. et al. Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. **Canadian Journal of Forest Research,** Ottawa, v. 35, n. 2, p. 85-92, Feb. 2005.

JOUAN-RIMBAUD, D.; MASSART, D. L.; NOORD, O. E. Random correlation in variable selection for multivariate calibration with a genetic algotithm. **Chemometrics and Intelligent Laboratory Systems,** Amsterdam, v. 35, n. 2, p. 213-220, Dec. 1996.

LEARDI, R.; SEASHOLTZ, M. B.; PELL, R. J. Variable selection for multivariate calibration using a genetic algorithm: prediction of additive concentrations in polymer films from Fourier transform-infrared spectral data. **Analytica Chimica Acta**, Amsterdam, v. 461, n. 2, p. 189-200, June 2002.

NÆS, T. et al. A user-friendly guide to multivariate calibration and classification. Chichester: NIR Publication, 2002.

PASQUINI, C. Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, São Paulo, v. 14, n. 2, p. 198-219, 2003.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least-squares procedures. **Analytical Chemistry,** Washington, v. 36, n. 8, p. 1627-1639, July 1964.

SCHIMLECK, L. R.; EVANS, R. Estimation of *Pinus radiata* D. Don tracheid morphological characteristics by near infrared spectroscopy. **Holzforschung**, Berlin, v. 58, n. 1, p. 66-73, June 2004.

SCHIMLECK, L. R.; EVANS, R.; ILIC, J. Estimation of *Eucalyptus delegatensis* wood properties by Near Infrared Spectroscopy. **Canadian Journal of Forest Research**, Ottawa, v. 31, n. 10, p. 1671-1675, Oct. 2001.

SCHIMLECK, L. R. et al. Estimation of basic density of *Eucalyptus globulus* using near-infrared spectroscopy. **Canadian Journal of Forest Research,** Ottawa, v. 29, n. 2, p.

194-201, Feb. 1999.

THYGESEN, L. G. Determination of dry matter content and basic density of Norway spruce by near infrared reflectance and transmission spectroscopy. **Journal of Near Infrared Spectroscopy,** Sussex, v. 2, n. 3, p. 127-135, 1994.

VIA, B. K. et al. Multivariate modelling of density, strength and stiffness from near infrared spectra for mature, juvenile and pith wood of longleaf pine (*Pinus palustris*). **Journal of Near Infrared Spectroscopy,** Sussex, v. 11, p. n. 4, p. 365-378, 2003.

VIANA, L.C. et al. Predicting morphological characteristics and basic density of *Eucalyptus* wood using NIRS technique. **Cerne,** v. 15, n. 4, p. 421-429, out./dez. 2009.

WESTAD, F.; MARTENS, H. Variable selection in near infrared spectroscopy based on significance testing in partial least square regression. **Journal of Near Infrared Spectroscopy,** Sussex, v. 8, n. 2, p.117-124, 2000.

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, Amsterdam, v. 58, n. 2, p. 109-130, Oct. 2001.