Wood Science and Technology 35 (2001) 475–485 © Springer-Verlag 2001

Originals

Prediction of basic wood properties for Norway spruce. Interpretation of Near Infrared Spectroscopy data using partial least squares regression

J. B. Hauksson, G. Bergqvist, U. Bergsten, M. Sjöström, U. Edlund

Abstract This work was undertaken to investigate the feasibility of using near infrared spectroscopy (NIR) and partial least squares regression (PLS) as a tool to characterize the basic wood properties of Norway Spruce (*Picea abies* (L.) Karst.). The wood samples originated from a trial located in the province of Västerbotten in Sweden. In this trial, the effects of birch shelterwoods *(Betula pendula* Roth) of different densities on growth and yield in Norway spruce understorey were examined. All Norway spruce trees in each shelterwood treatment were divided into three growth rate classes based on diameter at breast height (1.3 m) over bark. Five discs were cut from each tree (i.e. from the root stem, and at 20%, 40%, 60%, and 80% of the total height). The discs from 40% tree height were used (i.e., where the largest variations in annual ring widths and wood density were found). A total of 27 discs were selected. The discs were used for measuring annual ring widths, wood density, average fiber length and the fiber length distributions. Milled wood samples prepared from the discs were used for recording NIR spectra. PLS regression was used to generate prediction models for the wood properties (Y-matrix) and NIR spectra (X-matrix) as well as between the wood properties (Y-matrix) and the fiber length distributions (X-matrix). One set of models was generated using untreated spectra and fiber length distributions. For a second set of models the structure in the X-matrix, which was orthogonal to the matrix described by the wood properties, was eliminated using a soft target rotation technique called orthogonal signal correction (OSC). The PLS model obtained using "raw" untreated NIR spectra and fiber length distributions had a poor modeling power as evidenced by the cumulative Q^2 values. For the PLS models based on untreated NIR spectra the cumulative Q^2 values ranged from a

Received 12 June 1999

Jón B. Hauksson (✉), Michael Sjöström, Ulf Edlund
Department of Chemistry, Organic Chemistry,
Umeå University, 901 87 Umeå, Sweden
e-mail: jon.hauksson@chem.umu.se

Göran Bergqvist, Urban Bergsten
Department of Silviculture,
The Swedish University of Agricultural Sciences,
901 83 Umeå, Sweden

Grants from the Swedish Foundation for Strategic Research are gratefully acknowledged. We also thank Bo Ahlqvist, MoDo Paper R&D, for providing the fiber length measurements.

minimum of 16% (wood density) to a maximum of 46% (no. of annual rings). Orthogonal signal correction of the X-matrix (NIR spectra or fiber length distributions) gave PLS models with a modeling power corresponding to cumulative Q^2 values well in excess of 70%. The improvement in predictive ability accomplished by the OSC procedure was verified by placing four of the 27 observations in an external test set and comparing RMSEP values for the test set observations without OSC and with OSC.

Introduction

The optimal use for a given wood raw material in pulp and paper manufacture is dictated by the underlying wood and fiber properties. Wood density is considered a key property, affecting, for example, pulp yield per unit of wood volume (Zobel and van Buijtenen 1989). A high and uniform wood density is desirable for most products (Olson and Arganbright 1977). For Norway spruce (*Picea abies* (L.) Karst.), a negative correlation has been found between annual ring width and wood density, suggesting that a low growth rate promotes the production of high density wood (Klem 1942; Olesen 1976). However, wood density also exhibits large variations within a tree and between trees of the same species growing at similar rates (1). Growth rate and consequently wood and fiber properties are affected by, for example, site quality and the silvicultural systems applied (Zobel and van Buijtenen 1989).

There is a need for fast and simple methods for the classification of wood raw material with respect to its usefulness for pulp and paper manufacture. A direct measurement of relevant fiber properties using spectroscopy would have many advantages. A study involving modeling of strength properties from fiber properties demonstrates that the strength parameters of hand sheets are nearly completely determined by the fiber properties at a given level of beating treatment (Marklund et al. 1998). The fiber properties, in turn, can be modeled on the basis of spectroscopic measurements.

Near infrared (NIR) spectroscopy has been shown to be an excellent method for the evaluation of wood and fiber properties. It has been used to evaluate such properties as moisture content, density, compression strength and chemical and biological degradation of Norway spruce (Hoffmeyer and Pedersen 1995), as well as the hardwood to softwood ratio in wood chip mixtures (Niemz et al. 1992) and the resin content in wood chips (Niemz et al. 1994).

The combination of NIR spectroscopy and multivariate data analysis leads to improved models. For example, NIR spectroscopy and principal component analysis (PCA) have been used to classify wood samples of different origins (Svensson et al. 1996; Michell and Schimleck 1998; Schimleck et al. 1996). Partial least squares regression (PLS) has been applied to study the correlation between the storage time of wet-stored timber from Scotch pine (*Pinus sylvestris* (L.)) and NIR spectra of the wood, as well as to distinguish between sapwood and heartwood (Borga et al. 1992). A PLS model has been used to determine the chemical contents of wood samples of Eucalypt species (*E. globulus* and *E. nitens*) from NIR spectra (Schimleck et al. 1997). Kraft pulp yields are also modeled in the same work in terms of the glucan and xylan contents of the wood by using PLS regression. The composition of mixtures of wood chips from different sources (Norway spruce from Sweden, Scotch pine from Sweden and Scotch pine from Poland) has been predicted using NIR spectroscopy and PLS regression (Antti et al. 1996). NIR spectroscopy is also used in the same study to characterize a series of pulp samples with respect to seventeen traditionally measured pulp properties.

The objective of this investigation was to evaluate the feasibility of using near infrared spectroscopy (NIR) and partial least squares regression (PLS) as a tool to characterize the basic wood properties of Norway Spruce (*Picea abies* (L.) Karst.).The NIR spectral data were pretreated with the new orthogonal signal correction (OSC) algorithm (Wold et al. 1998) prior to constructing the calibration models. This algorithm removes structure from the NIR spectra (X-matrix) which is unrelated to the y variable being modeled (e.g. viscosity, wood density etc.). Relationships between the spectroscopic characteristics of milled wood samples and the wood and fiber properties (i.e., wood density, annual ring width, no. of annual rings, arithmetic mean fiber length, length weighted mean fiber length and weight weighted mean fiber length) were established. The wood material came from a study which was originally performed in order to evaluate effects of growth rate and birch shelterwood density on wood density traits for Norway spruce understorey in a trial in the boreal coniferous forest (Bergqvist 1998). The OSC pretreatment of NIR spectra has previously been shown to enable the generation of excellent PLS models between NIR spectra of pulps and even of wood raw material and a number of physical and morphological parameters that are commonly used to characterize the properties of wood fibers and pulps (Marklund et al. 1999). The models were validated both internally using Q^2(cum) values and externally by comparing RMSEP values (root mean squared error of prediction) for the y-values obtained for observations in external test sets with and without OSC pretreatment of the X-matrix.

Materials and methods

Wood samples

The wood samples originated from a trial located in the province of Västerbotten, Sweden (64°18′30″N, 19°44′55″E, altitude 260 m) within the middle boreal forest zone (Ahti et al. 1968). In this trial, the effects of birch shelterwoods *(Betula pendula* Roth) of different densities (0, 300 and 600 stems ha^{-1}) on growth and yield in Norway spruce understorey were examined. For a more detailed description see Bergqvist (1998).

Wood sampling took place in October 1996, 48 years after the establishment of the stand and 21 years after the establishment of the trial. Prior to sampling, all Norway spruce trees in each shelterwood treatment were divided into three growth rate classes based on diameter at breast height (1.3 m) over bark (<8 cm (low growth rate), 8–11 cm (intermediate growth rate), >11 cm (high growth rate)). A total of 72 trees, 8 from each growth rate class (3 classes) within each shelterwood treatment (3 treatments) were randomly selected. Thus, 9 batches of trees were created.

Five discs were cut from each tree (i.e. from the root stem, and at 20%, 40%, 60%, and 80% of the total height). In this investigation, the discs from 40% tree height were used (i.e., where the largest variations in annual ring widths and wood density were found). A total of 27 discs were selected (i.e. 3 from each batch) the discs showing the largest, average and smallest average annual ring widths within each batch.

The discs were used for measuring annual ring widths, wood density, average fiber length and the fiber length distributions. Milled wood samples prepared from the discs were used for recording NIR spectra.

Measurements of wood properties

Wood density variations were measured on 1 mm thick samples prepared from the discs (one sample per disc, from bark to pith facing the south aspect of the tree), using a direct scanning micro-densitometer with automatic angle alignment and a resolution of 0.02 mm. Measurement precision was estimated to ±5%. Wood density was measured at 5.97% ± 1.35 (mean ± SD) moisture content and normalised to oven-dry density. The samples were not extracted before measurement. Methods of sample preparation, measurement, and normalisation, as well as the underlying theories and design of the equipment, are described in detail in Jonsson et al. (1990), Larsson et al. (1994), and Pernestål and Jonsson (1996).

Fiber lengths were measured with a Kajaani FS-200 instrument (TAPPI 1991). Three different mean values and fiber length distributions were recorded and used for modeling. The fiber length range, 0–7.20 mm, was divided into 36 classes by the Kajaani instrument. The percentage of fibers in each class was used in the calculations.

NIR measurements

The wood discs were cut into small pieces and dried in an oven at 50 °C for 6 hours. The wood pieces were subsequently milled in a Wiley mill to a particle size $\leq$ 2 mm. Prior to the NIR measurements the milled wood samples were dried in an oven at 60 °C for 24 hours in order to ensure uniform humidity in all the samples. The NIR instrument was a NIR System Model 6500 instrument from Perstorp Analytical equipped with a spinning small ring cup reflectance cell. The sample cup has an inner diameter of 3.8 cm and depth of 0.9 cm. The sample cup was filled to the same level for each sample and the sample was secured with a cardboard back. Care was taken to always place the back at the same level and firmly in place. The instrument, which operates in diffuse reflectance mode, records log(1/R) in the wavelength region 400 to 2500 nm. A 2 nm step size gave absorbance values at 1050 wavelengths. Two spectra per sample were recorded.

Multivariate data analysis

Software

Preprocessing of the data using orthogonal signal correction, OSC, (Wold et al. 1998) was carried out using scripts written in Matlab version 5 (Copyright © The MathWorks Inc.) on a Silicon Graphics O_2 workstation. The multivariate calculations were carried out using the software package Simca-P version 7.01 (Copyright © Umetri AB, Umeå, Sweden).

NIR calculations

Mean values of the two NIR spectra were used in the calculations. The physical variables were mean-centered and auto-scaled prior to the multivariate modeling. The NIR variables were mean-centered, but not scaled. The PLS calculations were based on every data point in the 27 NIR spectra, resulting in a total of 1050 NIR variables per observation. One set of models, i.e. a PLS-model for each y-variable, was calculated using the "raw" untreated NIR spectra. For another set of models, the orthogonal signal correction algorithm (OSC) (Wold et al. 1998) was used to remove structure from the NIR spectra (X-matrix) which was unrelated to the y matrix, in this case the wood properties, prior to generating the PLS-models. This was done for one y-variable at a time, resulting in a separate model for each variable. Cross-validation (Wold 1978) was used to determine the number of significant components in the PLS analysis.

In order to fully determine the predictive ability of a PLS model, it is necessary to have two data sets for the property (y) being studied, i.e. one training set for calculating the model, and another external test set. The model calculated based on the training set is applied to the test set to directly determine the predictive ability of the model.

The method of validation using an external test set was applied to the PLS models for the physical properties measured in this study in spite of the limited number of observations. Four of the 27 observations from the data set were placed in an "external" test set. OSC was then performed for the remaining 23 observations, whereafter PLS models were calculated. The OSC-PLS parameters from the work set calculations were applied to the NIR spectra from the 4 observations in the test set and PLS predictions were calculated for the test set. For comparison, the same procedure of external validation, using the same 4 observations as the test set, was done without performing orthogonal signal correction prior to the PLS calculations. The RMSEP values (root mean squared error of prediction) for the "raw" PLS models and those for the OSC-PLS models were used to compare the predictive ability of the models. The same procedure of validation using an external test set was also applied to the PLS models that were based on the fiber length distributions. The same four observations constituted the test set in each case.

Results and discussion

The NIR spectra were used only as "fingerprints" of the different wood samples, i.e. no attempts were made to identify certain NIR signals or to interpret the PLS weight vectors. The weight vectors express to what extent the variables in X take part in the modeling. The fiber length distributions were also used as an X matrix to generate PLS models for the wood properties (i.e., the no. of annual rings, annual ring width and wood density).

Models based on the NIR spectra

PLS models for the following wood and fiber properties (y-variables) were generated on the basis of NIR spectra of the wood samples: the number of annual rings, the annual ring width, wood density, arithmetic mean fiber length, length weighted mean fiber length, and weight weighted mean fiber length. The models labeled PLS-R1 etc. were based on untreated (raw) NIR spectra, while the models labeled PLS-O1 etc. were based on OSC-treated NIR spectra. The "raw" models for annual ring width and wood density consisted of one statistically significant component each according to cross-validation, Table 1A. These models used approximately 81% of the variation of the NIR spectra to explain only 32% of the variation of the wood properties (annual ring width and wood density). Q^2 values are a measure of the fraction of the variation of y that can be predicted by the models according to cross-validation. For a model comprising two or more PLS components, the software reports a cumulative Q^2 value for all the components, labeled Q^2(cum). According to this criterion the predictive ability of the models was very poor, i.e. the Q^2(cum) value for the annual ring width model was 27.1%, while the Q^2(cum) value for the wood density model was even lower, 16.1%. For the other four "raw" models, between four and six PLS components were required according to cross-validation. The model for the number of annual rings consisted of 6 PLS components which used 99.9% of the variation of the NIR spectra to explain 69.9% of the variation of the number of annual rings. The predictive ability was low, corresponding to a Q^2(cum) value of only 45.7%. Only PLS components number 5 and 6 had any predictive ability at all. The model for the arithmetic mean fiber length (PLS-R4) consisted of 4 PLS components which used 99.6% of the variation of the NIR spectra to explain only 24.4% of the variation of the mean fiber length values. The Q^2(cum) value was only 0.2%. Thus, this model had no predictive ability at all. The other models for mean fiber lengths, PLS-R5 for the length weighted mean fiber length and PLS-R6 for the weight weighted mean fiber length, consisted of 5 PLS components using more than 99% of the variation of the NIR spectra to explain between 70 and 76% of the variation of the mean fiber lengths. The Q^2(cum) values were quite low, 19.5% for PLS-R5 and 42.9% for PLS-R6. Models PLS-R1 through PLS-R6 are summarized in Table 1A. From these results it is evident that a large percentage of the variation of the NIR spectra was not relevant for modeling the wood and fiber properties. Consequently, it was considered important to use the new orthogonal signal correction procedure to remove variation from X (the NIR spectra) which was unrelated to the physical parameter being modeled. The resulting models (PLS-O1 through PLS-O6) are summarized in Table 1B. The OSC pretreatment of the NIR spectra "cleaned up" the models significantly, lowered the number of significant components and dramatically increased the modeling power as expressed by Q^2(cum)

Table 1A. PLS-models for wood and fiber properties based on untreated NIR spectra of wood samples

Model	y-variable	Component	R^2X(cum)	R^2y(cum)	Q^2(cum)
PLS-R1	A	1	0.274	0.134	0.000
		2	0.980	0.151	0.000
		3	0.988	0.271	0.000
		4	0.994	0.374	0.002
		5	0.998	0.478	0.186
		6	0.999	0.699	0.457
PLS-R2	B	1	0.808	0.316	0.274
PLS-R3	C	1	0.810	0.323	0.161
PLS-R4	D	1	0.813	0.038	0.000
		2	0.915	0.102	0.000
		3	0.991	0.173	0.002
		4	0.996	0.244	0.002
PLS-R5	E	1	0.813	0.034	0.012
		2	0.960	0.048	0.012
		3	0.992	0.099	0.012
		4	0.993	0.506	0.012
		5	0.994	0.757	0.195
PLS-R6	F	1	0.814	0.037	0.013
		2	0.953	0.050	0.013
		3	0.991	0.094	0.013
		4	0.996	0.281	0.013
		5	0.997	0.702	0.429

A = No. of annual rings
B = Annual ring width
C = Wood density
D = Arithmetic mean fiber length
E = Length weighted mean fiber length
F = Weight weighted mean fiber length

Table 1B. PLS-models for wood and fiber properties based on OSC-treated NIR spectra of wood samples

Model	y-variable	Component	R^2X(cum)	R^2y(cum)	Q^2(cum)
PLS-O1	A	1	0.600	0.907	0.895
PLS-O2	B	1	0.658	0.900	0.891
PLS-O3	C	1	0.669	0.740	0.736
PLS-O4	D	1	0.770	0.743	0.722
PLS-O5	E	1	0.699	0.852	0.840
PLS-O6	F	1	0.705	0.812	0.806

See Table 1A for an explanation of the meaning of A, B, etc

values. Each model consisted of one PLS component using between 60% (number of annual rings) and 77% (arithmetic mean fiber length) of the variation of the OSC treated NIR spectra to explain between 74% (arithmetic mean fiber length) and 91% (number of annual rings) of the physical parameter being modeled. The modeling power was good, corresponding to Q^2(cum) values from 72.2% (arithmetic mean fiber length) to 89.5% (number of annual rings). However, high Q^2(cum) values could be misleading using the present preprocessing technique. Therefore, validation, using a true external test set, is absolutely essential.

Models based on fiber length distributions

PLS models for the following wood and fiber properties were calculated on the basis of the fiber length distributions: the number of annual rings, annual ring width, and wood density. As with the models based on NIR spectra, two sets of models were calculated, one set based on the untreated fiber length distributions (PLS-R7 through PLS-R9), and a second set based on OSC "corrected" fiber length distributions (PLS-O7 through PLS-O9). The models are summarized in Table 2. The "raw" models for the number of annual rings and the annual ring width consisted of three PLS components, using 71% of the variation of the fiber length distribution to explain 78% of the variation of the two physical parameters being modeled. The modeling power was fairly poor, corresponding to Q^2(cum) values of 40.8% (number of annual rings) and 47.3% (annual ring width). The model for wood density based on the fiber length distribution was poor with a Q^2(cum) value of only 12.2%. It consisted of two PLS components and used 54%

Table 2A. PLS-models for wood and fiber properties based on untreated fiber length distributions of wood samples

Model	y-variable	Component	R^2X(cum)	R^2y(cum)	Q^2(cum)
PLS-R7	A	1	0.493	0.446	0.257
		2	0.618	0.653	0.321
		3	0.713	0.784	0.408
PLS-R8	B	1	0.166	0.452	0.000
		2	0.553	0.603	0.241
		3	0.710	0.775	0.473
PLS-R9	C	1	0.161	0.431	0.000
		2	0.539	0.569	0.122

See Table 1A for an explanation of the meaning of A, B, etc

Table 2B. PLS-models for wood and fiber properties based on OSC-treated fiber length distributions of wood samples

Model	y-variable	Component	R^2X(cum)	R^2y(cum)	Q^2(cum)
PLS-O7	A	1	0.628	0.892	0.878
PLS-O8	B	1	0.648	0.923	0.911
PLS-O9	C	1	0.581	0.730	0.711

See Table 1A for an explanation of the meaning of A, B, etc

of the variation of the fiber length distribution to explain only 57% of the variation of the wood density. Although orthogonal signal correction is a method which has been developed for the purpose of treating spectroscopic matrices (NIR, NMR etc.) prior to PLS modeling, a decision was made to use OSC here to improve the modeling of the three wood and fiber properties based on the fiber length distributions. The resulting OSC models are summarized in Table 2B. A dramatic improvement has been obtained compared with the "raw" models. The one component models used between 58.1% (wood density) and 64.8% (annual ring width) of the OSC treated fiber length distribution to explain between 73% (wood density) and 92.3% (annual ring width) of the physical parameter being modeled. The modeling power was from good (wood density: $Q^2(\text{cum}) = 71.1\%$) to excellent (annual ring width: $Q^2(\text{cum}) = 91.1\%$).

Validation of the predictive ability of the PLS models

The best way to determine the predictive ability of a PLS model is by validation using an external test set. The predictive ability of the PLS models was tested directly by omitting 4 samples and moving them to an "external" test set. For each physical property OSC was performed on the remaining 23 observations in the training set and a PLS model was calculated based on the OSC treated NIR spectra. The same procedure was repeated without the OSC pretreatment. Predictions were calculated, based on both the "raw" NIR spectra and the OSC treated NIR spectra. The RMSEP values for the test sets based on the "raw" PLS models and those based on the OSC-PLS models are shown in Table 3. While it is evident from the RMSEP values that the predictive ability of the PLS models varied considerably depending on which property was being modeled, the OSC pretreatment decreased the prediction errors for the samples in the external tests. The same method of validation using 4 observations as an external test set was also applied to the PLS models that were based on the fiber length distributions.

Table 3. RMSEP-values for the test sets for each of the PLS- models based on NIR spectra of wood samples. Prior to modeling, samples no. 3, 4, 5 and 26 were excluded from the training set and moved to an "external" test set

Physical Property	Raw PLS models	OSC-PLS models
No. of annual rings	15.0	8.51
Annual ring width	1.67	1.04
Wood density	0.94	0.82
Arithm. mean fiber length	2.31	2.13
Length weighted mean fiber length	1.13	0.63
Weight weighted mean fiber length	1.14	0.22

Table 4. RMSEP-values for the test sets for each of the PLS- models based on fiber length distributions of wood samples. Prior to modeling, samples no. 3, 4, 5 and 26 were excluded from the training set and moved to an "external" test set

Physical Property	Raw PLS models	OSC-PLS models
No. of annual rings	5.96	5.10
Annual ring width	0.71	0.53
Wood density	0.044	0.040

The RMSEP values are shown in Table 4. It is evident that the OSC pretreatment of the fiber length distributions improved the predictive ability of the PLS models, but not to the same extent as for the PLS models based on NIR spectra.

Plots showing observed vs. calculated values for two of the PLS models (PLS-R6/PLS-O6, NIR-based models for weight-weighted mean fiber lengths) are shown in Fig. 1. For most of the samples there is a significantly better correspondence between observed and calculated fiber length values after OSC pretreatment of the NIR-spectra (PLS-O6) than without the OSC pretreatment.

Variability in particle size and moisture content between samples can be an important source of error in calibration models based on NIR measurements (Osborne et al. 1993). Since light scattering decreases with increasing mean particle size, the radiation penetrates deeper into the sample, leading to increased log(1/R) values with increasing mean particle size. The effect is to displace the diffuse reflectance spectrum along the vertical axis as a function of the particle size. However, since the scattering power of small particles also varies with wavelength, the displacement due to particle size is not uniform across the entire spectrum. The presence of water in a sample gives rise to characteristic absorption bands but also affects the overall spectrum since the scattering depends on the ratio of the refractive index of the particles to that of the surrounding medium. Methods for compensating for particle size effects include the use of derivatives and various methods for pretreatment of spectral data, such as multiplicative signal correction, MSC (Geladi et al. 1985) and orthogonal signal correction, OSC. Since OSC removes the structure in the spectroscopic data matrix (X) which is orthogonal to the variable being modeled (y), it is the method of choice for removing any unwanted variation in NIR spectra caused by, e.g.

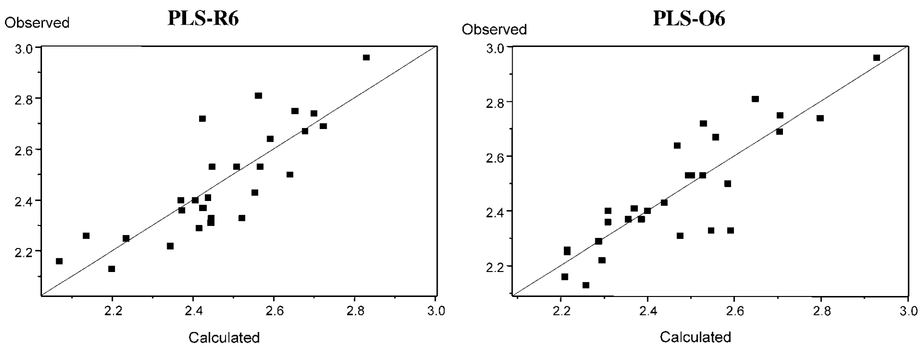

Fig. 1. PLS-R6/PLS-O6: Observed vs. calculated values for weight-weighted mean fiber lengths. R signifies "raw", untreated NIR spectra, and O signifies OSC treated NIR spectra

differences in particle size or different moisture content among the samples being studied. Such variation in spectral data will surely not be related to the physical or chemical parameter being modeled. In spite of the availability of the powerful OSC procedure, care was taken to prepare and treat all the wood samples in exactly the same way (milling, drying, amount of sample, pressure applied to the back of the sample cup, etc.). Any remaining effects likely to adversely affect the quality of the obtained models were for certain removed by the OSC procedure.

Since there will always be variations among different NIR instruments, the models obtained here cannot be expected to be directly transferable to spectra obtained for other similar samples on another instrument in another laboratory. Nevertheless, analogous measurements on these or similar samples done elsewhere should give similar models with approximately equal predictive ability.

These results clearly demonstrate the usefulness of the new OSC procedure for pretreatment of NIR spectroscopic data. It is evident that NIR spectra of milled wood samples carry information which is correlated with all the wood and fiber properties being studied here. For all the properties orthogonal signal correction of the NIR data gave models with a high degree of explanation and with a good predictive ability.

Conclusions

Numerous studies have proven the power of combining spectroscopic data (NMR, NIR etc.) with multivariate data analysis as a tool for characterizing pulps and for building models which enable predictions of pulp properties from spectroscopic data (Wallbäcks et al. 1989; Wallbäcks et al. 1991; Lennholm et al. 1994; Antti et al. 1996; Wold et al. 1998). The results presented here demonstrate good to excellent correlation between NIR spectra of wood meal and a number of wood and fiber properties. It was even possible to correlate fiber length distributions with these wood and fiber properties. The use of a preprocessing method called orthogonal signal correction (OSC) greatly improved the modeling power of the PLS models based on NIR spectra and even fiber length distributions. The improvement in predictive ability was tested using external test sets in all cases.

References

Ahti T, Hämet-Ahti L, Jalas J (1968) Vegetation zones and their sections in Northwestern Europe, Ann Bot Fenn 5: 169–211

Antti H, Sjöström M, Wallbäcks L (1996) Multivariate calibration models using NIR spectroscopy on pulp and paper industrial applications, J Chemometrics 10: 591–603

Bergqvist G (1998) Wood density traits in Norway spruce understorey: effects of growth rate and birch shelterwood density, Annales des Sciences Forestières 55: 809–821

Borga P, Hämäläinen M., Theander O (1992) Correlations between near-infrared spectra of wet-stored timber and the storage time in relation to the water-quality, Holzforschung 46: 299–303

Geladi P, MacDougal D, Martens H (1985) Linearization and scatter correction for near-infrared reflectance spectra of meat. Appl Spectroscopy 39: 491–500

Hoffmeyer P, Pedersen JG (1995) Evaluation of density and strength of Norway spruce wood by near-infrared reflectance spectroscopy, Holz Roh- Werkstoff 53: 165–170

Jonsson B, Dissing E, Pernestål K (1990) A high resolution and high speed annual ring analyser with automatic angle alignment, Swedish University of Agricultural Sciences, Dept. of Biometry and Forest Management, Umeå Report 19

Klem GG (1942) Effect of planting space on the quality of spruce wood and sulphite pulp, Comm. Norwegian Forest Research Institute 28: 257–293 (in Norwegian with English summary)

Larsson B, Pernestål K, Jonsson B (1994) A wood sample preparation method for direct scanning X–ray microdensitometry, Swedish University of Agricultural Sciences, Dept. of Forest Mensuration and Management, Umeå, Report 29
Lennholm H, Larsson T, Iversen T (1994) Determination of cellulose I_α and I_β in ligno-cellulosic materials, Carbohydrate Research 261: 119–131
Marklund A, Hauksson JB, Edlund U, Sjöström M (1998) Prediction of strength parameters for softwood kraft pulps. Multivariate data analysis based on physical and morphological parameters, Nord Pulp Paper Res J 13: 30–38
Marklund A, Hauksson JB, Edlund U, Sjöström 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: 12–20
Michell AJ, Schimleck LR (1998) Further classification of eucalypt pulpwoods using principal components analysis of near-infrared spectra, Appita J 51: 127–131
Niemz P, Korner S, Wienhaus O, Flamme W, Balmer M (1992) Applying NIR spectroscopy for evaluation of the hardwood softwood ratio and resin content in chip mixtures, Holz Roh- Werkstoff 50: 25–28
Niemz P, Dutschmann F, Stolken B (1994) Using NIR-spectroscopy for evaluating resin content in chip mixtures, Holz Roh- Werkstoff 52: 6–8
Olesen PO (1976) The interrelation between basic density and ring width of Norway spruce, Danish Forest Research Institute 34: 341–359
Olson JR, Arganbright DC (1977) The uniformity factor – a proposed method for expressing variations in specific gravity, Wood Fiber Sci 9: 202–210
Osborne BG, Fearn T, Hindle PH (1993) Practical NIR spectroscopy with applications in food and beverage analysis, Longman Scientific & Technical, UK
Pernestål K, Jonsson B (1996) A method for automatic collimator alignment in microdensity measurements, Wood Sci Tech 30: 91–97
Schimleck LR, Michell AJ, Vinden P (1996) Eucalypt wood classification by NIR spectroscopy and principal components analysis, Appita J 49: 319–324
Schimleck LR, Wright PJ, Michell AJ, Wallis AFA (1997) Near-infrared spectra and chemical compositions of *E-globulus* and *E-nitens* plantation woods. Appita J 50: 40–46
Svensson O, Josefson M, Langkilde FW (1996) Classification of chemically modified celluloses using a near-infrared spectrometer and soft independent modeling of class analogies. Applied Spectroscopy 51: 1826–1835
TAPPI (1991) Fiber length of pulp and paper by automated optical analyzer. TAPPI, T 271 pm-91. 4 pp
Wallbäcks L, Edlund U, Nordén B (1989) Multivariate Data Analysis of in Situ Pulp Kinetics Using ^{13}C CP/MAS NMR, J Wood Chem Technol 9: 235–249
Wallbäcks L, Edlund U, Nordén B, Iversen T (1991) Multivariate Characterization of Pulp. Part I. Spectroscopic characterization of physical and chemical differences between pulps using ^{13}C CP/MAS NMR, FT-IR, NIR and multivariate data analysis. Nord Pulp Paper Res J 6: 74
Wold S (1978) Cross-Validatory Estimation of the Number of Components in Factor and Principal Component Models, Technometrics 20: 397–405
Wold S, Antti H, Lindgren F, Öhman J (1998) Orthogonal Signal Correction of NIR spectra, Chemometrics and Intelligent Laboratory Systems 44: 175–185
Zobel BJ, van Buijtenen JP (1989) Wood variation. Its causes and control, Springer–Verlag, Berlin, 1989