

Multivariate characterization of pulp using solid-state ^{13}C NMR, FTIR, and NIR

Lars Wallbäcks, Ulf Edlund, Bo Nordin

Department of Organic Chemistry, Umeå University, S-901 87 Umeå, Sweden

Ingemar Berglund

Dept. of Analytical Chemistry, Umeå University, S-901 87 Umeå, Sweden

ABSTRACT *Changes in pulp samples in chemical composition were monitored during the birch kraft pulping process using three different spectroscopic methods— ^{13}C CP/MAS NMR, FTIR, and NIR. The spectroscopic data were evaluated in relation to a matrix composed of chemical descriptors using multivariate methods. All these techniques could readily predict the contents of the main carbohydrate constituents (glucose and xylose) and the Klason lignin content. Although no attempts were made to optimize the analysis methods individually, NIR seems to be the best method. A suitable combination of spectroscopic data preprocessed in a multivariate way was better than any individual spectroscopic method used.*

KEYWORDS

Analysis
Infrared
spectroscopy
Nuclear
magnetic
resonance
Pulps
Spectroscopy

In characterizing complex samples like pulp samples, several difficulties are associated with spectroscopic methods such as ^{13}C cross-polarization/magic angle spinning (CP/MAS) nuclear magnetic resonance (NMR) spectroscopy, Fourier transform infrared (FTIR) spectroscopy, and near-infrared (NIR) spectroscopy. The most important drawback is that signals from the three main polymer constituents (cellulose, hemicellulose, and lignin) partly overlap as a result of the limited resolution of the solid-state techniques.

By using multivariate analysis to characterize each spectrum, it is possible to diminish this problem. In principle, every data point in the spectrum could be used as a descriptor (variable). As a calibration tool and as a means to further increase the structural description of the pulp samples, a model can be derived based on a combined data set from various spectroscopic techniques.

Our purpose in this study was to extract information from the spectral representations and to relate the

information to classical chemical descriptors such as lignin content and carbohydrate content. Information is extracted in this way with the use of a multivariate calibration technique called "partial least squares data analysis" (PLS) (1, 2). Besides being able to resolve information out of overlapping lines, this technique has the advantage of being able to handle cases in which the number of samples is small compared with the number of variables used for the pulp characterization.

Background

The need for indirect or secondary analysis of complicated samples for material like pulp has recently attracted attention (3). Most attempts to extract information from spectroscopic data have been based on stepwise regression in which the compositions of lignin and sugars have been estimated. Recently developed instrumentation, especially for FTIR, uses principal component regression (PCR) for such quantifications. The

components from the orthogonal principal component analysis (PCA) are then used in the stepwise regression description.

In a previous ^{13}C solid-state NMR study of birch pulp samples taken at regular intervals during the kraft pulping process (4), multivariate techniques made it possible to extract the information content in such a way that the amount of lignin could be predicted with great precision. As will be shown here, one can allow several chemical descriptors to enter into the dependent Y matrix.

NIR data obtained from peat have been used earlier in PCA and PLS analyses to extract some information such as water content and calorific value (5). This approach has also been used to determine the protein content of meat (6) and wheat (7).

In a similar fashion, FTIR data at selected wavenumber regions have been used for the multivariate correlation to peat descriptors such as the degree of decomposition and the content of amino acids (8).

Experimental procedures

Samples

We analyzed the kraft pulping process by taking 20 pulp samples, collected during a period of 200 min. The samples were numbered according to minutes elapsed from a zero point, which was set to begin at 80°C. From this point, the temperature was raised at 0.8°C/min to a final temperature of 170°C. The first sample was taken at 60 min, and the last one was taken at 260 min. The monosaccharide constituents (as alditol acetates) were determined by sulfuric acid hydrolysis (72% H₂SO₄, 30°C for 1 h, then diluted to 2.5% H₂SO₄ and autoclaved at 120°C for 1 h) followed by derivatization and gas capillary chromatography (9).

Results were quantified using myoinositol as an internal standard. The remaining material gave the Klason lignin content (corrected for ash content). The contents of monosaccharides and lignin were measured as percent of dry weight (Table I). The analyzed monosaccharides were glucose, xylose, galactose, arabinose, and mannose.

The sum of the percentages of monosaccharides and Klason lignin varied between 80% and 96%. This variation was corrected for by scaling the carbohydrate data in such a way that it added up to 100%, including lignin.

NMR measurements

We obtained ¹³C CP/MAS NMR spectra using a Bruker MSL-100 spectrometer operating at 25.178 MHz. Instrumental parameters were:

Contact time: 1 ms

Repetition rate: 2.5 s

Line broadening: 20 Hz

Rotational speed: 3.5 KHz

Number of scans: 5000

Data points: 700 zero filled to 2000.

Five thousand scans were made using Andrew-Beems rotors made from per-deuterated PMMA (poly methoxymethacrylate). Adamantane (δ -CH₂ = 38.3) was used as an external reference.

For analytical purposes, only magnitude spectra were used in the NMR processing [$R_{\text{new}} = (\text{Real}^2 + \text{Imaginary}^2)^{1/2}$], where the new spectrum, R_{new} , is obtained by squaring the real and imaginary solutions], thus exclud-

I. Chemical composition of the pulp according to the chemical analysis

Sample no.	Pulping time, min	Content, %					Klason lignin
		Arabinose	Xylose	Mannose	Galactose	Glucose	
1	60	0.93	21.89	0.61	0.58	53.45	22.28
2	70	1.12	21.56	0.42	0.51	55.42	20.88
3	80	0.88	21.47	0.39	0.40	55.91	20.94
4	90	1.40	20.77	0.35	0.35	57.50	19.64
5	100	1.31	21.95	0.38	0.37	58.48	17.44
6	110	0.86	22.34	0.33	0.28	60.59	15.60
7	120	1.45	22.44	0.35	0.26	64.48	11.10
8	125	1.07	23.19	0.36	0.21	67.04	8.14
9	130	0.99	22.63	0.36	0.26	67.26	8.52
10	140	1.18	25.15	0.40	0	68.74	4.53
11	145	0.85	24.30	0.45	0	70.29	4.12
12	150	0.65	23.75	0.37	0	70.90	4.33
13	160	0.79	25.17	0.36	0	70.74	2.94
14	170	0.81	24.34	0.36	0	72.25	2.24
15	180	1.11	24.95	0.31	0	70.86	2.77
16	190	1.15	25.13	0.33	0	71.12	2.27
17	200	0.70	24.69	0.39	0	72.03	2.19
18	210	0.98	24.15	0.33	0	72.74	1.79
19	220	0.80	24.32	0.31	0	72.81	1.76
20	260	1.22	25.42	0.33	0	72.17	0.86

ing any error introduced by a manual phase correction.

Diffuse reflectance FTIR measurements

The infrared spectra were recorded at 26°C using a Bruker 113-V FTIR spectrometer equipped with a DTGS (deuterated triglycine sulfate) detector and a Globar source. One thousand scans were combined, and the resultant interferogram was Fourier transformed to obtain a spectral resolution of 4 cm⁻¹ over the spectral range. The samples were not specifically treated before the measurement.

In an independent study by others (10), the effect of milling has been especially studied using diffuse reflectance infrared Fourier transform (DRIFT). A substantial improvement of the standard error of prediction (SEP) was obtained by milling the sample to small particle sizes. However, the optimization of a single method is outside the scope of this study.

For the DRIFT measurements, an automatic sampling accessory from Bruker was used, and pure powdered KBr was used as a reference.

Near infrared reflectance (NIR) measurements

The NIR measurements were made using a Technicon InfraAlyzer 450 operating with 19 different optical filters in the spectral range of from 1445 nm to 2348 nm.

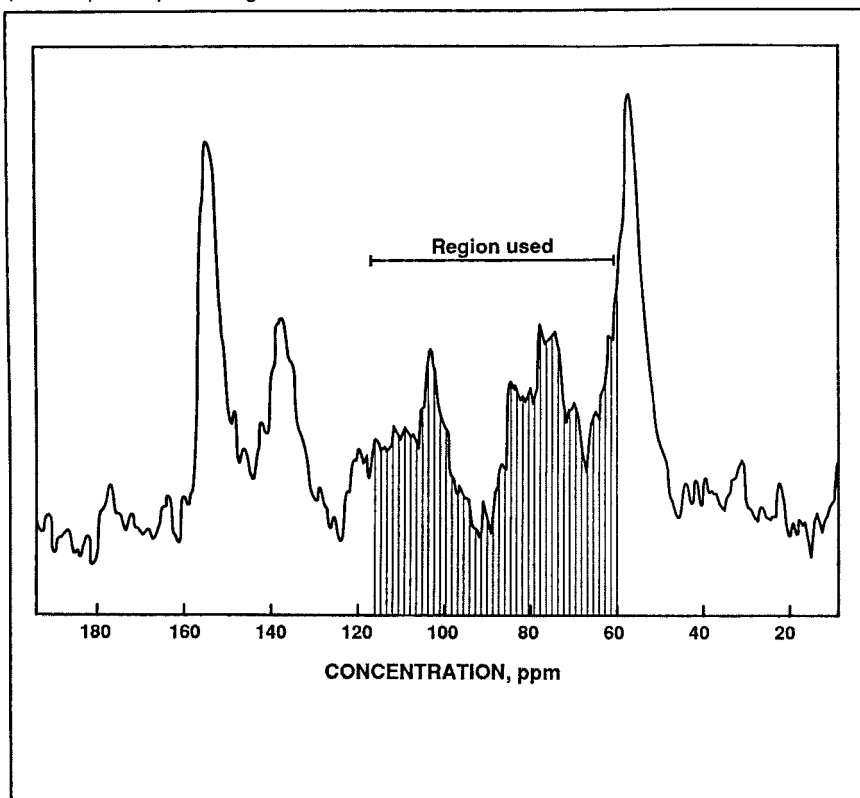
The wavelengths used were (in nanometers): 2348, 2336, 2310, 2270, 2230, 2208, 2190, 2180, 2139, 2100, 1982, 1940, 1818, 1778, 1759, 1734, 1722, 1680, 1445. Reflectance data were recorded (filter factor = 1) and transferred to a computer. Before measurements were made, the samples were milled through a 1-mm sieve by using a KAMAS Slaggy 200 B mill.

Data analysis

From the NMR spectra, 73 evenly distributed data points in the "carbohydrate" region (60–115 ppm, every 4th data point from the original spectrum) were used as variables (Fig. 1). Details about the conversion of raw data into digitized data suitable for principal components or partial least squares analysis can be found elsewhere (11).

From the DRIFT spectra, 75 equally spaced wavenumbers from the

1. The difference between the NMR spectra of the first sample (60 min) and last sample (260 min). The spectral region used in the calculations is indicated.



spectral window 2000–715 cm^{-1} were used as variables. No attempt was made to optimize the selection of data points or spectral regions. Absorbance intensities were used, and a baseline correction was established according to instrumental software before transferring the data to a personal computer.

Multiplicative scatter correction (MSC) (12), which is a method developed for NIR data, was also used because we noticed a slightly improved predictive ability for the IR data with this data preprocessing. We also corrected the NIR data (19 wavelengths) by MSC before these data were used in the PLS calculations.

The chemical analysis data were combined in one dependent matrix (Y) (percentage of 5 monosaccharides and Klason lignin). In the calculations in which principal components were used as variables, these variables were automatically scaled to a variance of 1. When the NIR variables were used together with principal components from NMR and IR data, the NIR variables were scaled to give approximately the same sum of squares as the six principal components together.

During the calculations, “cross-validation” was used to determine the number of significant components (13).

To validate and compare the predictive ability of the models, the samples were divided into a calibration set and a test set.

Reviews about principal components and partial least squares analysis can be found in the relevant literature (1, 2, 14).

Results and discussion

According to the composition of monosaccharides obtained from the chemical analysis (Table I), all samples contain mainly glucose and xylose. The additional sugars are only present in minor amounts.

As mentioned, the spectroscopic methods used for the pulp samples give complicated spectra in terms of overlapping signals from different constituents. To illustrate this complexity, Fig. 1 presents an NMR difference spectrum [first sample (after 60 min) minus the last sample (after 260 min)]. In the region used, signals from lignin, cellulose, and hemicellulose are not resolved. Out-

side this spectral region, the resolved aromatic signals (low field at 115 ppm) and the methoxy carbons of lignin (56 ppm) can be seen.

The calculated PLS models readily describe the changes in the main constituents (glucose, xylose, and Klason lignin). For the remaining sugars, which are all present in amounts around 1% each, only the gradual decrease of galactose is described properly. This condition arises because of connection or correlation with the main delignification process. The reason these low-content sugars are not described by any of the methods used is that the analytical sensitivity is limited in the spectroscopic methods and in the chemical analysis.

NMR

Using NMR data, we obtained a three-component model that describes 98.3% of the variance within the chemical composition data (Y). As Fig. 2 shows, the major chemical constituents glucose, xylose, and Klason lignin are rather well accounted for by the model. Also, the galactose content is well described from this model. As mentioned, however, this is most likely an effect attributable to the correlation with the main delignification process.

The changes in mannose and arabinose content are not described by this model, and the situation is similar with DRIFT and NIR. This result is in opposition to some of our preliminary results presented earlier (15).

These results are based on variables using only the unresolved “carbohydrate” region of the spectra (Fig. 1). Even if the signals are overlapping in this limited region, information can be extracted by using the multivariate methods.

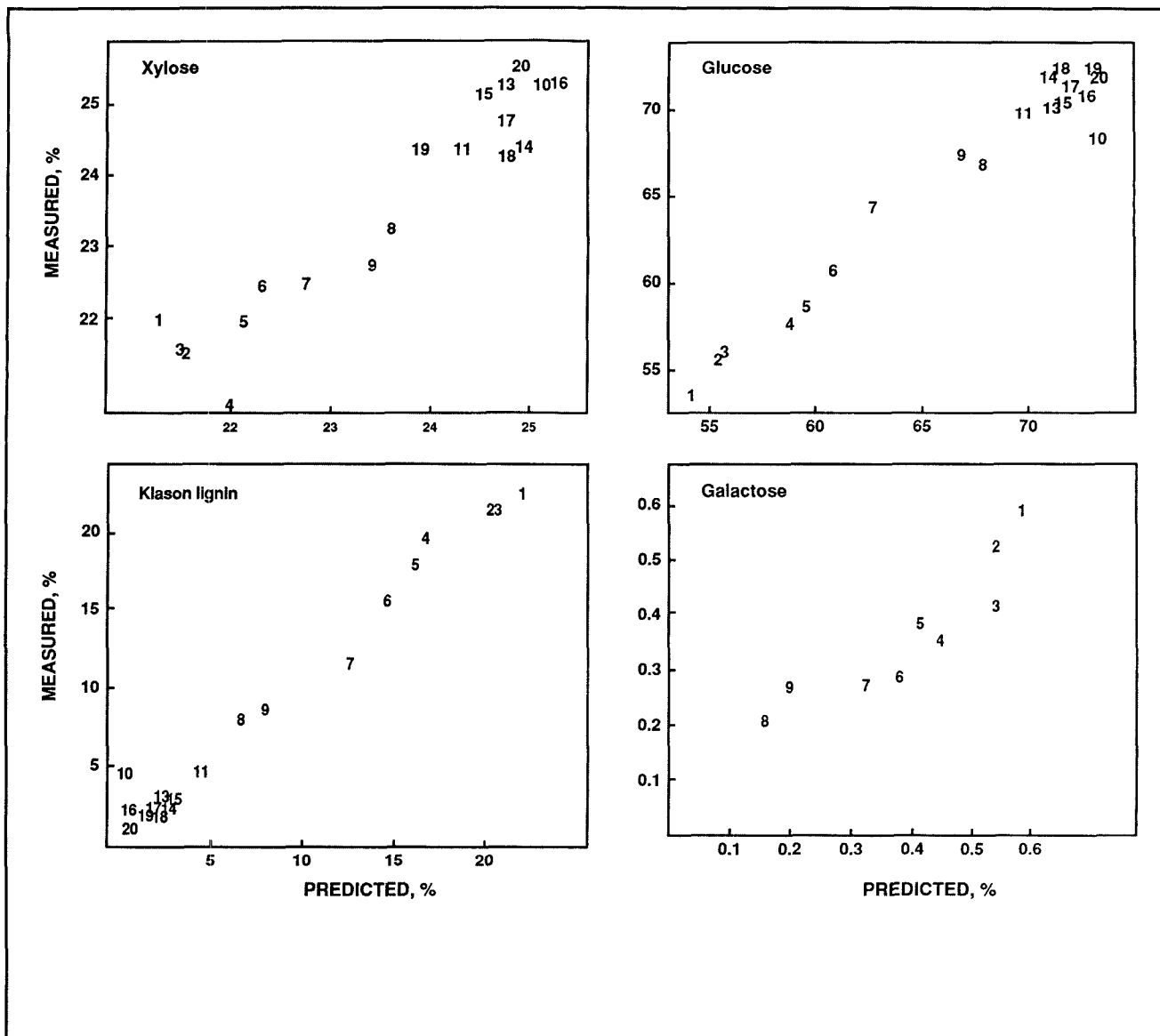
As shown previously (4), an improved description of the Klason lignin content is obtained if the resolved low-field lignin signals from the NMR spectra are included.

FTIR

By using the DRIFT data, we obtained a two-component model that described 78% of the variance in the chemical composition matrix.

As Table II shows, the predictive ability is not as good as with NIR data, but a selection of wavenumbers should improve the model (that is, if

2. The result of the calculation with the NMR data as independent matrix: measured vs. predicted contents of xylose, glucose, Klason lignin, and galactose. (Numerals refer to Table I.)



variables that do not contain information about the actual problem are excluded). It is likely, however, that the observed difference between NIR and DRIFT data can be ascribed to the refining that was performed only on the NIR samples.

NIR

With NIR data, the best model was obtained in terms of predictive ability (Table II). A three-component model described 99.0% of the variance in the chemical composition matrix. As Fig. 3 shows, the predictions are excellent, especially for Klason lignin and glucose.

In Table II, the predictive abilities for the different spectroscopic meth-

ods are compared. NIR gives the best predictions for all of the major constituents (Klason lignin, glucose, and xylose).

NIR, NMR, and FTIR

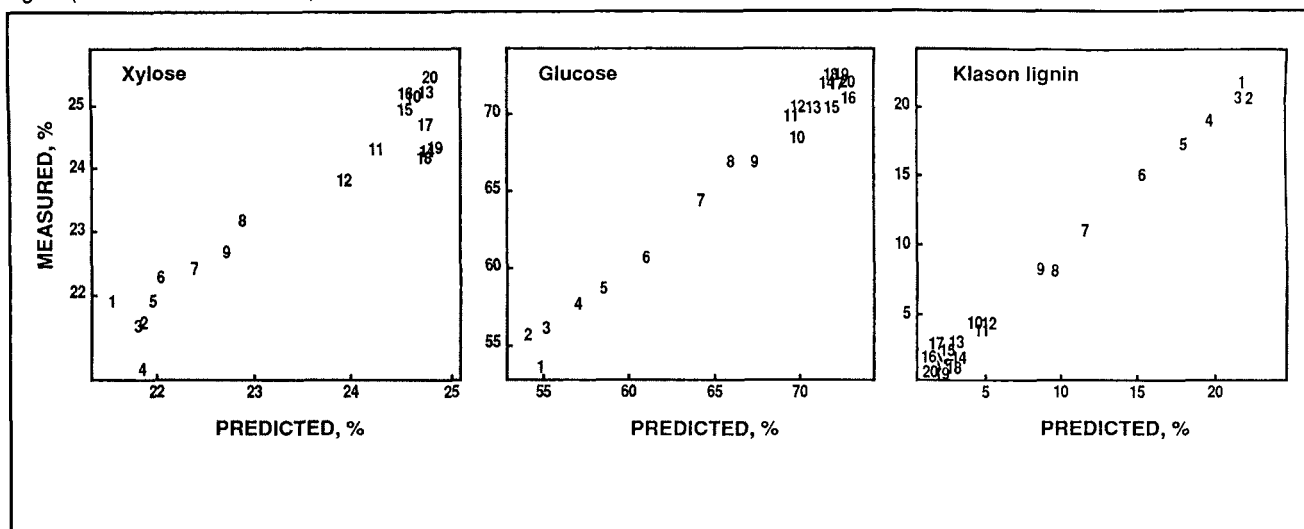
Because all of these spectroscopic methods can be used to probe the chemical changes that occur during the birch kraft pulping process, it is likely that a better description of the process can be derived by using the combined spectroscopic data. There are different ways of combining the spectroscopic data, and we will consider two of these here.

As a first attempt to improve the predictive ability of our model, three principal components from each spec-

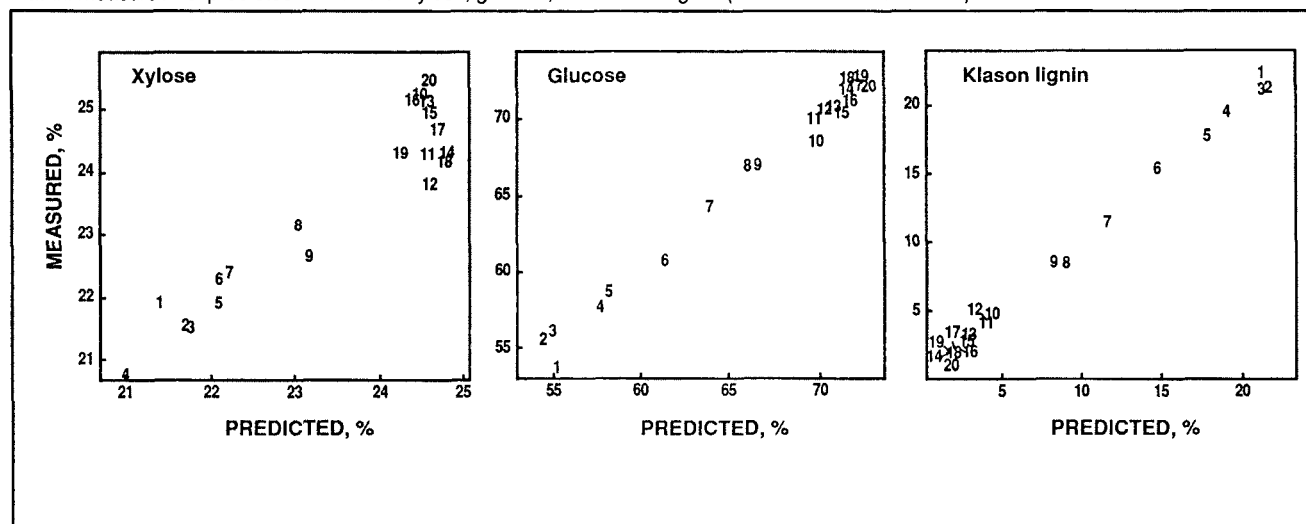
II. The standard error of prediction for the different models

	Klason lignin, %	Glucose, %	Xylose, %
NMR	1.82	1.89	0.57
FTIR	3.31	3.09	0.65
NIR	0.70	1.03	0.57
3 PC NIR	1.15	1.03	0.73
+ 3 PC NMR			
+ 3 PC FT-IR			
NIR	0.61	0.74	0.54
+ 3 PC NMR			
+ 3 PC FT-IR			

3. The result of the calculation with the NIR data as independent matrix: measured vs. predicted contents of xylose, glucose, and Klason lignin. (Numerals refer to Table I.)



4. The result of the calculation with the combined matrix NIR and three principal components from each NMR and FTIR data as independent matrix: measured vs. predicted contents of xylose, glucose, and Klason lignin. (Numerals refer to Table I.)



trosopic data matrix are combined to an X matrix containing nine variables. After automatically scaling the principal components in the X matrix, we performed a PLS calculation, with the composition of chemical constituents as the dependent matrix (Y).

As Table II shows, this approach of combining experimental data does not improve the standard error of prediction for any of the constituents, compared with analysis based on NIR data only. Because NIR was the preferred method in the separate calculations, we combined the original 19 NIR variables with three principal components from NMR, as well as FTIR, to form an X matrix totally composed of 25 variables.

After scaling the NIR data to give approximately the same sum of squares as the six principal components together, we performed a PLS calculation. The resulting three-component model describes 98.8% of the variance in chemical composition. As can be seen in Table II, the standard error of prediction is reduced for all of the major constituents (Klason lignin, glucose, and xylose).

The resulting plots are presented in Fig. 4.

Chemical analysis

The chemical analysis of the pulp samples suffers from some complicating effects. First, the samples measured do not contain monosaccharides

or Klason lignin, and the relationship between these and actual constituents (cellulose, hemicellulose, and lignin) might be complex.

In addition, because the amounts are given as percent of dry substance, a "closure" effect is present in the chemical analysis matrix (16). This effect arises because the variation of the concentration of one constituent causes variations for others. In this case, a decrease in Klason lignin will induce a relative increase in the amount of carbohydrates. This condition will make it difficult to interpret the chemical significance for the calculated PLS components.

Finally, the amounts of the less

abundant sugars approach the errors associated with the chemical analysis.

Conclusions

Any of these spectroscopic methods (CP/MAS NMR, FTIR, and NIR) can be used to describe the major changes in chemical composition that occur during the kraft pulping of birch. The data were evaluated by using multivariate data analysis methods.

In this study, NIR was the spectroscopic method that gave the best model in terms of predictive ability. The NIR method is both fast and easy to use, and it represents an attractive alternative method for obtaining information about chemical constituents. In combination with fiber optics, NIR might also be used for "on-line" measurements (17). However, no attempts have been made to optimize the individual spectroscopic analysis

in terms of sample handling (the sample holders in DRIFT, the rotors in NMR, etc.) or in terms of sample pretreatments (such as sample dilution and refining).

The addition of principal components from the NMR and DRIFT data to the NIR variables resulted in a model with better predictive ability than obtainable with the single NIR measurements only. □

Literature cited

1. Wold, S., Albano, C., Dunn, W. J. III, Edlund, U., *et al.*, *Chemometrics Mathematics and Statistics in Chemistry* (B. R. Kowalski, Ed.), D. Reidel Publ. Co., Dordrecht, Holland, 1984, p. 17 (ISBN 90-277-1846-6).
2. Geladi, P. and Kowalski, B. R., *Analytica Chimica Acta* 185: 17(1986).
3. Schultz, T. P. and Burns, D. A., *Tappi J.* 73(5): 209(1990).
4. Wallbäck, L., Edlund, U., and Nordén, B., *J. Wood Chem. Tech.* 9(2): 235(1989).
5. Johansson, E., Persson, J.-A., and Albano, C., *Fuel* 66: 22(1987).
6. Martens, H. A., *Multivariate calibration*, thesis, Technical University of Norway, Trondheim, 1985.
7. Martens, H. A. and Naes, T., *Trends in Analytical Chemistry* 3(8): 204(1984).
8. Holmgren, A. and Nordén, B., *Applied Spectroscopy* 42(2): 255(1988).
9. Theander, O. and Westerlund, E., *J. Agr. Food Chem.* 34: 330(1986).
10. Backa, S. and Brolin, A., personal communication.
11. Nordén, B. and Albano, C., *Fuel* 68: 771(1989).
12. Geladi, P., MacDougall, D., and Martens, H., *Applied Spectroscopy* 39(3): 491(1985).
13. Wold, S., *Technometrics* 20: 397(1978).
14. Wold, S., Esbensen, K., and Geladi, P., *Chemometrics and Intelligent Laboratory Systems* 2: 37(1987).
15. Nordén, B., Edlund, U., and Wallbäck, L., *Cellulose and Wood: Chemistry and Technology* (C. Schuerch, Ed.), Wiley, 1989, p. 1023.
16. Johansson, E., Wold, S., and Sjodin, K., *Analytical Chemistry* 56: 1685(1984).
17. Weyer, L. G., *Applied Spectroscopy* 41(5): 786(1987).

We are grateful to MoDo AB for providing pulp samples and to Christina Rockner for making the chemical analysis. We thank Allan Holmgren for assistance with the FTIR measurements. We are grateful to Nils and Dorthii Troedsson for financial support.

Received for review Dec. 7, 1990.

Accepted April 8, 1991.