Carbohydrate Polymers 82 (2010) 772–778

Contents lists available at ScienceDirect

Carbohydrate Polymers

journal homepage: www.elsevier.com/locate/carbpol

Qualitative and quantitative analysis of wood samples by Fourier transform infrared spectroscopy and multivariate analysis

Huilun Chen [a,b], Carlo Ferrari [b], Marco Angiuli [b], Jun Yao [a], Costantino Raspi [c], Emilia Bramanti [b,*]

[a] *Key Laboratory of Biogeology and Environmental Geology of Chinese Ministry of Education & School of Environmental Studies & Sino-Hungarian Joint Laboratory of Environmental Science and Health, China University of Geosciences, 430074 Wuhan, PR China*
[b] *Italian National Research Council-Istituto per i Processi Chimico-Fisici, Laboratory of Instrumental Analytical Chemistry, Via G. Moruzzi 1, 56124 Pisa, Italy*
[c] *C.G.S. s.a.s. Via del Borghetto, Pisa, Italy*

ARTICLE INFO

Article history:
Received 5 January 2010
Received in revised form 25 May 2010
Accepted 26 May 2010
Available online 4 June 2010

Keywords:
Wood
Classification
Fourier transform infrared spectroscopy
Multivariate analysis

ABSTRACT

Fourier transform infrared (FTIR) spectroscopy, in combination with multivariate analysis, enable the analysis of wood samples without time-consuming sample preparation. The aim of our work was to analysis the wood samples qualitatively and quantitatively by FTIR spectroscopy. A Van Soest method to determine the lignin, cellulose and hemicellulose content, was used as reference method. Multivariate calibration was performed based on first derivative of the FTIR spectra in the wave number range from 1900 to 800 $\mathrm{cm^{-1}}$, using principal component analysis (PCA), hierarchical cluster analysis (HCA) and partial least-squares (PLS) chemometric methods. Multivariate calibration models for FTIR spectroscopy have been developed. Using PCA and HCA approach, wood samples were classified as softwoods and hardwoods while wood samples with and without treatments were labeled as wood, neutral detergent solution fiber (NDSF), acid detergent solution fiber (ADSF) and strong acid solution fiber (SASF). Furthermore, PLS regression method was applied to correlate lignin, cellulose and hemicellulose contents in wood samples with the FTIR spectral data. The models' refinement procedure and validation was performed by cross-validation. Although a wide range of input parameters (i.e., various wood species) was used, highly satisfactory results were obtained with the root-mean-square errors for the contents of lignin, cellulose and hemicellulose are 1.51, 0.96 and 0.62%, respectively. This study showed that FTIR spectra have the potential to be an important source of information for a quick evaluation of the chemical composition of wood samples.

1. Introduction

Wood is an abundant, renewable, and biodegradable composite with many useful applications, such as paper pulp production, building and furniture, natural reinforcement in polymer composites, and it is an important source of energy (Baratieri, Baggio, Fiori, & Grigiante, 2008; Morris, Steward, & Gilmore, 2000). As a renewable source of energy, wood and other forms of biomass offer environmental benefits with easy availability of feedstocks including agricultural and forestry residues and wastes and can be used in various ways to provide energy, such as combustion, gasification, and pyrolysis (McKendry, 2002; Prochnow, Heiermann, Plöchl, Amon, & Hobbs, 2009; Senneca, 2007). The main advantage of wood over fossil fuel is that the amount of carbon dioxide (CO_2) emitted during the burning process is typically 90% less than when burning fossil fuel (Gnansounou, Dauriat, Villegas, & Panichelli, 2009).

The characterization of wood samples is a complex procedure involving several steps wherein wood components are isolated or degraded to monomeric fragments. These procedures destroy the wood matrix and require large sample sizes and long analysis time (Ferraz, Baeza, & Duránt, 1991; Kaar & Brink, 1991).

Fourier transformed infrared (FTIR) spectroscopy has been used as a simple technique for obtaining rapid information about the structure of wood constituents and chemical changes taking place in wood due to various treatments (Michell, 1988). Contrary to conventional chemical analysis, this technique requires small sample sizes and short analysis time. However, wood is a complex material composed of an orderly arrangement of cells with cell walls including varying amounts of cellulose, hemicelluloses, lignin, and extractives (Conrad, Smith, & Fernlund, 2003). The different components are quantified by a set of extractions with alkaline and acid reagents, each able to remove one component, leaving the others intact. Among analytical methods employed for quantification of wood composition the Van Soest method is often used according to its accuracy in quantification and high purity of the fractions obtained (Van Soest, 1963; Van Soest & Wine, 1967).

---

* Corresponding author. Tel.: +39 050 315 2293; fax: +39 050 315 2555.
E-mail address: emilia@ipcf.cnr.it (E. Bramanti).

0144-8617/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carbpol.2010.05.052

Table 1
Summary of papers concerning wood chemical analysis employing FTIR approach.

Analytical approach	Sample material and comments	Ref.
FTIR	Wood: changes in wood properties	Müller, Schöpper, Vos, Kharazipour, & Polle (2009)
	Wood: comparative study in detail	Colom and Carrillo (2005)
FTIR and MLR	Wood: lignin, hemicellulose and cellulose contents	Schultz and Burns (1990)
	Wood: lignin content and the composition of polysaccharides	Schultz, Templeton, & McGinnis (1985)
FTIR and PLS	Acetylated wood: acetyl group content	Stefke, Windeisen, Schwanninger, & Hinterstoisser (2008)
	Eucalyptus globules wood: monosaccharide content	Rodrigues, Puls, Faix, & Pereira (2001)
	Birch stem wood: cellulose, xylan, lignin and extractives contents	Toivanen and Alen (2007)
	Kraft pulps: lignin content	Dang, Bhardwaj, Hoang, & Nguyen (2007)
	Hardwood and softwood pulps: kappa numbers and carbohydrate compositions	Bjarnestad and Dahlman (2002)
	Pine and eucalyptus wood: glucan, polyoses and lignin content	Ferraz, Baeza, Rodriguez, & Freer (2000)
	Pinus radiate wood: chemical composition and the density	Meder, Gallagher, Mackie, Bohler, & Meglen (1999)
	Wood kraft pulps: klason lignin, xylose, glucose contents	Wällbacks, Edlund, Nordén, & Berglund (1991)
FTIR, PLS and PCR	Wood kraft pulps: kappa numbers	Malkavaara and Alen (1998)
	Milled wood lignins: phenolic hydroxyl group contents	Faix and Bottcher (1993)

Owing to the complex property of wood, most of FTIR observed bands cannot be directly assigned to one single component, and interpretation of isolated bands in wood FTIR spectrum can be misleading (Michell, 1988; Owen & Thomas, 1989). For this reason, multivariate analyses in combination with FTIR spectral data are ideally suited for the analysis of the chemical composition of complex materials such as wood fibers (Bjarnestad & Dahlman, 2002). In a number of investigations, multiple linear regression (MLR), principal component analysis (PCA), hierarchical cluster analysis (HCA), principal component regression (PCR), and partial least-squares regression (PLS), have been used together with FTIR for rapid characterization of lignin and carbohydrates in wood. Some of these studies are summarized in Table 1. To our knowledge, PLS has not been used to evaluate simultaneously lignin, cellulose, and hemicellulose contents in wood samples.

The first purpose of this study is aimed at identifying spectral differences between wood samples by means of routine FTIR in combination with PCA and HCA. For this purpose, the spectra of 12 wood samples and the different detergent fibers obtained from these woods were used as training set. The second purpose of the present study is to predict the contents of lignin, cellulose, and hemicellulose by PLS analysis of FTIR spectra with reasonable accuracy.

2. Experimental

2.1. Experimental samples

Hard- and softwood samples were obtained from Estonia (C, D, E, F), Retford, England (L, M, O) and Norway (BN, LN, RSN, SN, J) (Table 2). The wood samples were ground to a particle size of 200 μm. The wood sampling and preparing procedure is provided in TAPPI T 257 cm (1985).

2.2. Composition analysis

Van Soest analysis was used for raw fiber determination (Van Soest, 1963; Van Soest & Wine, 1967). In the process of Van Soest analysis, biomass was separated progressively into neutral detergent solution fiber (NDSF), acid detergent solution fiber (ADSF) and strong acid fiber (SASF). Generally, NDSF mainly consists of hemicellulose, cellulose, lignin and insoluble minerals, while ADSF is composed of cellulose, lignin and insoluble minerals. SASF contains only lignin and insoluble minerals. The weight of lignin in the SASF fraction was calculated by subtracting the weight of ash (estimated in excess by thermogravimetric analysis, TGA, as <1%). All the results are expressed as percentage of total dry weight (Table 2).

2.3. FTIR spectroscopy

FTIR spectra of wood samples and different detergent fibers were measured by direct transmittance using the KBr pellet technique. Spectra were recorded in the range of 4000–400 cm^{-1} using a Spectrum 100 FTIR spectrometer (PerkinElmer Inc., USA). The pretreatment was carried out by tabletting the mixture of each sample and KBr (where KBr has a proportion of 0.5–1 wt%) into a very thin film with a diameter of 13 mm. All the spectra were measured at a spectral resolution of 4 cm^{-1} and 16 scans were taken per sample. Background spectra were obtained using pure KBr (Evans, 1991; Owen & Thomas, 1989). The FTIR spectra were baseline-corrected

Table 2
Compositional analysis of biomass based on the Van Soest method.

Sample	Component (wt%)	Water and soluble[a]	Hemicellulose	Cellulose	Lignin
Softwood					
C	Pine	19	20	35	25
D	Juniperus communis	23	21	28	27
F	Pine	23	19	27	30
L	Pine chips	15	19	42	23
O	Pines pellets	14	22	41	22
RSN	Residual spruce	22	22	33	22
SN	Spruce	23	17	33	26
Hardwood					
E	Silver birch	20	27	25	27
M	SRC[a] willow chips	22	22	39	16
BN	Birch	26	21	33	20
LN	Lecus	25	25	34	15
Mixture					
J	Pine, larch, birch, and spruce chips	19	21	33	26

[a] Short rotation coppice.

and normalized at 1900 cm^{-1}. The presented FTIR spectra are average value of the respective single spectra.

2.4. *Thermogravimetric analysis (TGA)*

The wood samples were ground to a particle size of 200 μm for thermogravimetric analysis. Approximately 2.5 mg of sample was placed into a ceramic sample pan. The thermogravimetric analysis was performed with a Perkin-Elmer TGA7 thermo-balance. The runs were performed using a constant heating rate of 20 °C/min, starting from 50 up to 800 °C. Each sample was studied in a steady gas flow (90 ml/min) of dry air or nitrogen to discriminate between the weight loss processes due to pyrolysis and those due to combustion. Before use, the TGA was calibrated for temperature and mass according to the procedure outlined by the manufacturer of the equipment.

2.5. *Multivariate data analysis*

All samples were scanned in the mid-infrared range of 4000–400 cm^{-1} but only the region of 1900–800 cm^{-1} that contained the most spectral information was analyzed with multivariate techniques in order to guarantee secure results (Alcantara, Honda, Ferreira, & Ferreira, 2007). For our data set, the first derivative of vector normalized spectra (12 smoothing points) were imported into standard software (SPSS 16.0, SPSS Inc., USA) for PCA and HCA (Chen, Xie, Fu, Yang, & Wang, 2006; Yang et al., 2008), while multivariate data analysis employing PLS was performed with the SIMCA-P demo 11.5 (Umetrics, Sweden) (Eriksson, Johansson, Kettaneh-Wold, & Wold, 1999).

The PLS model was developed by utilizing the first derivative FTIR spectra of the wood samples as the *X* matrix and the chemical determinations of the contents of various components (i.e., lignin, cellulose and hemicellulose) as the *Y* matrix. The data of the *X* matrix were mean-centered, while the *Y* data were both mean-centered and variance-scaled prior to the performance of PLS calculations. Due to the small number of samples, the data were not divided into two as the modeling data and the model validation data. Instead, the cross-validation technique was used to assess the model performance (Eriksson et al., 1999).

3. Results and discussion

3.1. *Characteristic bands in the FTIR absorption spectrum*

The FTIR spectra of softwood, hardwood and their mixtures are shown in Fig. 1. Detailed peak positions and assignments of the three types of wood samples are listed in Table 3 (Adler, 1977; Fengel & Shao, 1985). In the “fingerprint” region the spectra are very complex containing many bands assigned to main wood components. The bands at 1595, 1510/1507, 1270 and 1230 cm^{-1} are assigned to characteristic bending or stretching of different groups of lignin. The bands centered at 1463, 1425, 1375, 1160, 1111, and 1030 cm^{-1} are assigned to characteristic bending or stretching vibrations of different groups for lignin and cellulose. The bands at 1739 cm^{-1}, and around 1321–1317 cm^{-1} are assigned to characteristic bending or stretching vibrations of the different groups from cellulose (Popescu et al., 2009). The main characteristic peaks, which are the stretching vibrations of C=O, aromatic and C–O, are around 1640, 1510 and 1030 cm^{-1}. However, in the three types of wood the positions and relative intensity of peaks are different. The band intensity at 1739 cm^{-1} was greater in hardwoods than in softwoods. This band is attributed to the stretching of the free carbonyl groups (Sarkanen, Chang, & Ericsson, 1967). Although carbonyl groups in wood components (cellulose, hemicellulose, and lignin) are mainly in the hemicellulose branched chain component, the intensity of the bands depends on the ratio of holocellulose (cellulose and hemicellulose) content to lignin content in all types of wood (Colom & Carrillo, 2005). Furthermore, the peak around 1270 cm^{-1} seems significantly different in the spectra of hardwoods and softwoods. In hardwoods this peak is intense but in softwoods it appears as little more than a shoulder. This effect has been observed by Evans (1991) but is inconsistent with the result of Colom and Carrillo (2005).

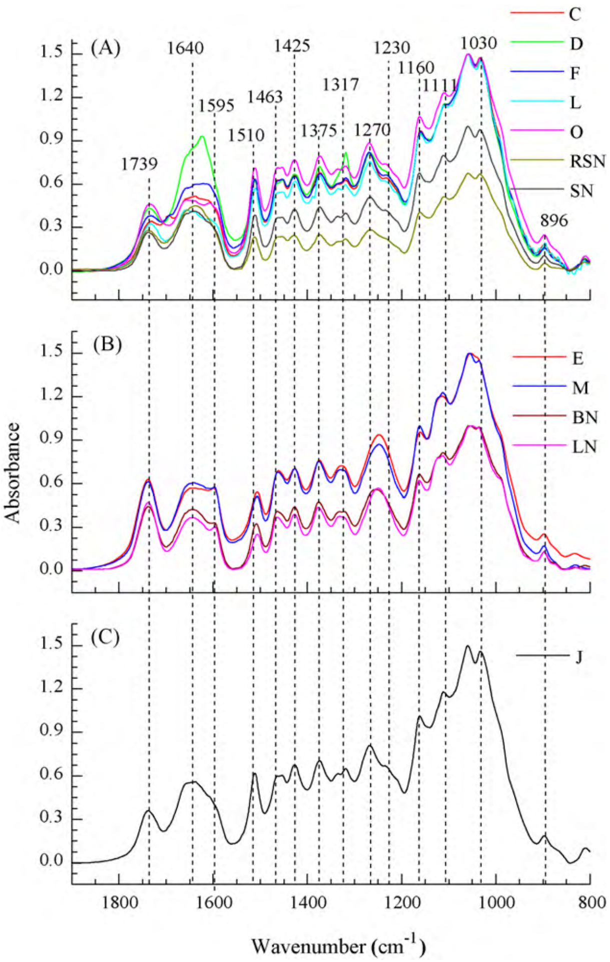

Fig. 1. The 1900–800 cm^{-1} region of the FTIR spectra of (A) softwood, (B) hardwood and (C) mixture species.

Sample *J* is a mixture that contains three-fourths of softwoods and one-fourth of hardwoods. FTIR spectrum of sample *J* (Fig. 1C) is similar to FTIR spectra of softwood (Fig. 1A and Table 3).

Although a part of the origin of FTIR spectral differences is currently addressed, these spectral differences in combination with a multivariate chemometric scheme can be used to better discriminate softwood and hardwood (Toivanen & Alen, 2007).

3.2. *Principal component analysis (PCA)*

The plot of the scores of PCs allowed the qualitative recognition of different groups of samples. Fig. 2A shows the scores scatter plot of the first two PCs for 12 wood samples, which together account for 96.7% of the total variability. Two groups may be suggested based on the distribution of samples along PC2: one group with negative PC2 scores (softwoods), and another group with scores located in the positive side of PC2 (hardwoods). The mixture *J* is grouped into/among softwoods in agreement with its approximate

Table 3
Assignment of bands in the infrared spectra of wood samples.

Wave number (cm^{-1})			Base group and vibration mode
Softwood	Hardwood	Mixture	
1739	1739	1739	C=O stretching in unconjugated ketone, carbonyl and aliphatic groups xylan
1640	1640	1640	C=O stretching conjugated to aromatic ring
–	1595	–	C=O stretching conjugated to aromatic ring
1510	1507	1510	Aromatic skeletal stretching
1463	1463	1463	CH_2 deformation stretching in lignin and xylan
1425	1425	1425	Aromatic skeletal combined with C–H in-plane deforming and stretching
1375	1375	1375	Aliphatic C–H stretching in methyl and phenol OH
1317	1321	1317	Condensation of guaiacyl unit and syringyl unit, syringyl unit and CH_2 bending stretching
1270	–	1270	Guaiacyl ring breathing plus C=O stretching
–	1230	–	C–C plus C–O plus C=O stretching
1160	1160	1160	C–O–C stretching in pyranose rings, C=O stretching in aliphatic groups
1111	1111	1111	Antisymmetric in-phase ring stretching
1030	1030	1030	C–O deformation in primary alcohols, plus C=O stretch unconj., plus aromatic C–H in-plane deformation
896	896	896	Character of cellulose P-chains, C–H stretching out of plane of aromatic ring

known composition (three-fourths of softwoods and one-fourth of hardwood).

The loads plot using the PC2 *vs* the variable number converted into wave number gave a sub-spectrum of the data set (Fig. 2B). This sub-spectrum gave information on what wave numbers mostly contributed to the separation of wood samples according the PC2 value. The PC2 loadings profile shows negative values at absorption bands characteristic of C=O stretching (1595, 1510, and 1270 cm^{-1}) and Guaiacyl units (1136 cm^{-1}) in lignin, indicating that higher amounts of this functional group characterize the samples with negative PC2 values. Indeed, softwood samples contain more lignin than hardwoods (Pettersen, 1984).

PCA was also performed on the first derivative FTIR spectra (1900–800 cm^{-1}) of 12 wood samples and different detergent fibers (DNSF, ADSF, and SASF) obtained by chemical processing of the same woods (see Section 2). Wood and NDSF differ only for neutral detergent soluble fraction. Hemicellulose, cellulose and lignin are intact in NDSF. NDSF and ADSF differ because of hemicellulose content: indeed, ADS solubilizes hemicellulose component. SASF and ADSF differ for cellulose component. The scores scatter plot of the first two PCs (Fig. 3A) account for 86.4% of the total variability. As expected, the SASF group, which contains only lignin and insoluble minerals, is separated from the others with positive PC2 scores. Fig. 3B shows the scores plot of PC1 and PC5 which describe 71.5 and 1.5% of the total variability, respectively. PC5 is mainly responsible for the separation of wood samples, NDSF and ADSF, from ADSF showing more negative scores than NDSF and NDSF more negative than wood samples (Fig. 3B). Wood samples and NDSF show a reasonable overlap, but the two groups were still discernable. In the PC2 loadings profile (Fig. 3C), positive values are found for broader signals corresponding to lignin (1595, 1270, and 1160 cm^{-1}) and cellulose (1440, 1135, and 997 cm^{-1}), indicating that lignin and cellulose predominate in SASF group and the other groups, respectively (with positive PC2 scores). PC5 loadings (Fig. 3C) may be used to interpret the separation of ADSF group as their scores are strongly negative in the PC5 axis. Negative loadings are found at absorption bands characteristic of C=O stretch on conjugated p-substituted aryl ketones (1670 cm^{-1}) in lignin and C=O and C=C stretching and CH_2 rocking (1100 and 1030 cm^{-1}) in cellulose (Adler, 1977; Dokken, Davis, & Marinkovic, 2005).

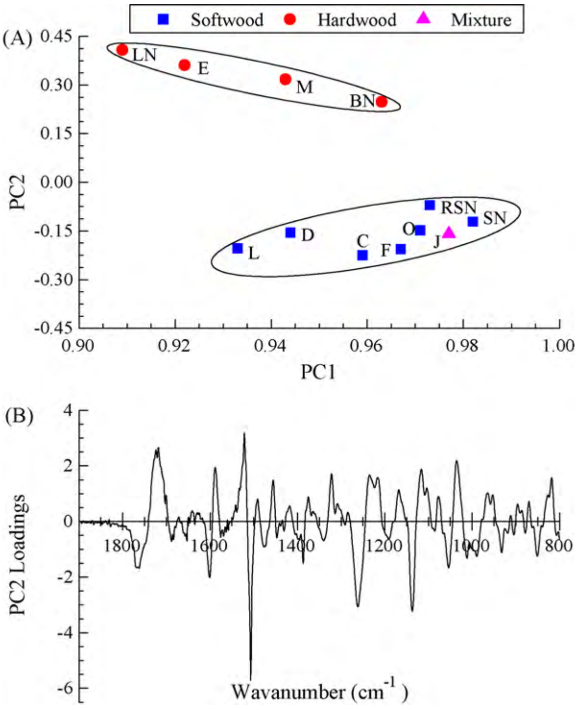

Fig. 2. PCA of the first derivative FTIR spectra of 12 wood samples (1900–800 cm^{-1}): (A) scores scatter of PC1 (91.0%) *vs* PC2 (5.7%) and (B) loadings profile of PC2.

3.3. Hierarchical cluster analysis (HCA)

The primary purpose of HCA is to present data in a manner that emphasizes the natural groupings in the data set. The presentation of HCA results as a dendrogram makes it possible to visualize clustering to show the relationships among different groups (Tan, Steinbach, & Kumar, 2006). In this work we used the Ward's method with the Euclidean distance as the similarity measurement. The resulting cluster for 12 wood samples is shown in Fig. 4, and the samples are grouped into two blocks: softwoods and hardwoods group. Moreover, sample *J* (mixture) appears classified in the softwood group, is agreement with the results obtained by PCA.

The dendrogram of HCA for 12 wood samples and different detergent fibers (Fig. 5) shows two main clusters, SASF (first cluster) and the other samples (second cluster), in accordance with PCA results (Fig. 3A). It was also observed that the second cluster was separated into two distinct subgroups. The first subgroup contains samples from wood and NDSF, while the second consists entirely of samples from ADSF. It is interesting to note that softwoods and hardwoods are separated in each subgroup. This result was not obtained by PCA. In particular, SASF has two subgroups of softwoods (L, O, C, D, F, J) and hardwoods (E, M, BN, LN), with the exception of RSN and SN that are softwoods, but they resulted included in hardwood subgroup. Wood, NDSF and ADSF cluster has two subgroups: all hardwoods and their respective NDSF (first subgroup, further separated in other two subgroups) and all soft-

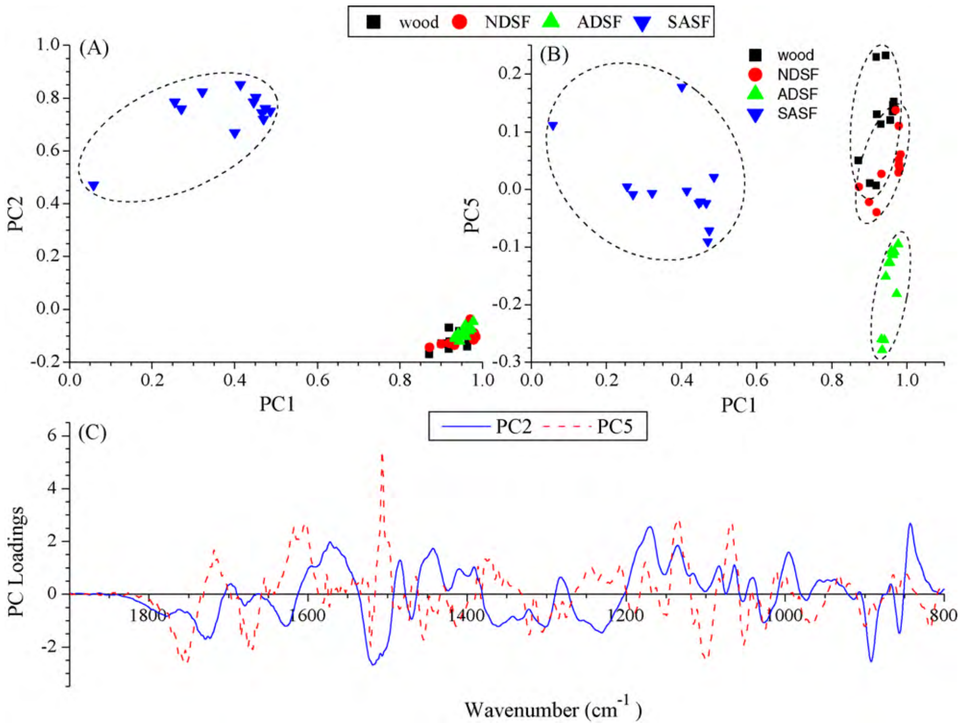

Fig. 3. PCA of the first derivative FTIR spectra (1900–800 cm^{-1} region) of 12 wood samples and NDSF, ADSF and SASF extracted by the Van Soest method: (A) scores scatter of PC1 (71.5%) *vs* PC2 (14.9%), (B) scores scatter of PC1 (71.5%) *vs* PC5 (1.5%) and (C) loadings profile of PC2 (bold line) and PC5 (dotted line).

woods, their respective NDSF and all ADSF (second subgroups). These findings indicate that HCA as an unsupervised technique not only identified clear wood samples and different detergents differences but also showed similarities between each group. Overall, the results of PCA and HCA show that the application of the chemometric methods to the FTIR data is a powerful tool for the discrimination of wood types.

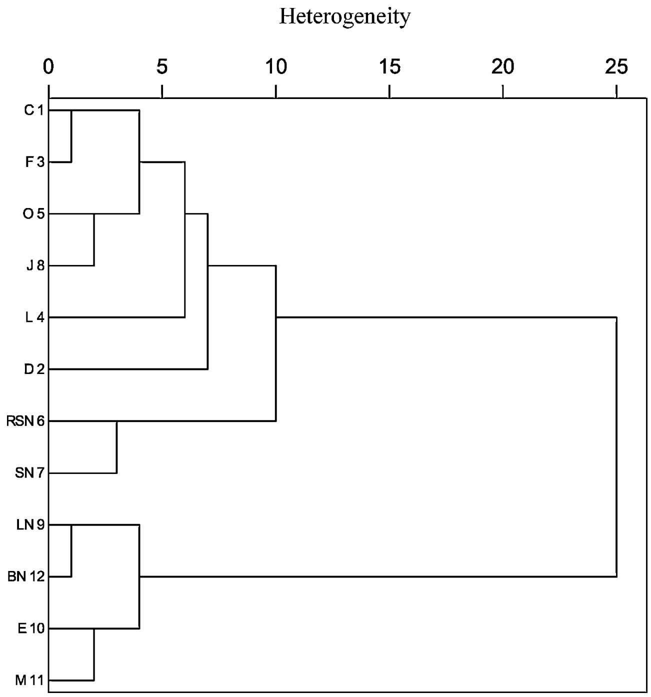

Fig. 4. HCA of FTIR spectra of 12 wood samples (1900–800 cm^{-1}). Spectral data were classified into classes according to their spectral heterogeneity. The first derivation and vector normalization was utilized as pretreatment method of the spectra.

3.4. PLS analysis

PLS was applied to our data set in order to predict wood composition in terms of cellulose, hemicellulose and lignin. To determine the appropriate number of components to include in our PLS model, the approach referred to as cross-validation was used (Eriksson et al., 1999). Here 11/12 of the original spectral data were utilized to construct the model, from which the remaining 1/12 of the data were predicted. Finally, these predicted values were compared with the actual ones obtained by chemical extraction.

For this purpose, the squares of the differences between predicted and observed values were added together to obtain the predictive residual sum of squares (PRESS), which is a measure of the predictive power of the model being tested. Usually, PRESS is re-expressed as Q^2, which is (1-PRESS/SS), where SS is the sum of squares of Y, after mean-centering. A quantitative measure of the quality of the fit is provided by the parameter R^2 (i.e., the explained variation), and the ability to predict is indicated by the prediction parameter Q^2 (i.e., the predicted variation). Using nine principal components, the PLS model developed here on the basis of our first derivative FTIR spectra could explain 98% of the variance in the X matrix and 98.8% of the variance in the Y matrix.

Fig. 6 shows the content values obtained from PLS model *vs* the actual content of lignin, cellulose and hemicellulose in wood samples. The difference between the actual contents and the predicted values is small, and correlation coefficients, R, for the observed *vs* predicted curves were 0.979 for lignin, 0.993 for cellulose and 0.989 for hemicellulose.

Table 4 shows the performance statistics of cross-validation and summarizes the relationship between the actual and the predicted values modeled. By investigating the R^2Y and Q^2 values of the individual responses, it was possible to determine for which responses the model worked better. R^2 and Q^2 should be as large as possible. In general, a Q^2 value >0.5 may be regarded as good and a Q^2 value >0.9 as excellent, depending on the particular application. It is important that the difference between R^2Y and Q^2 not become

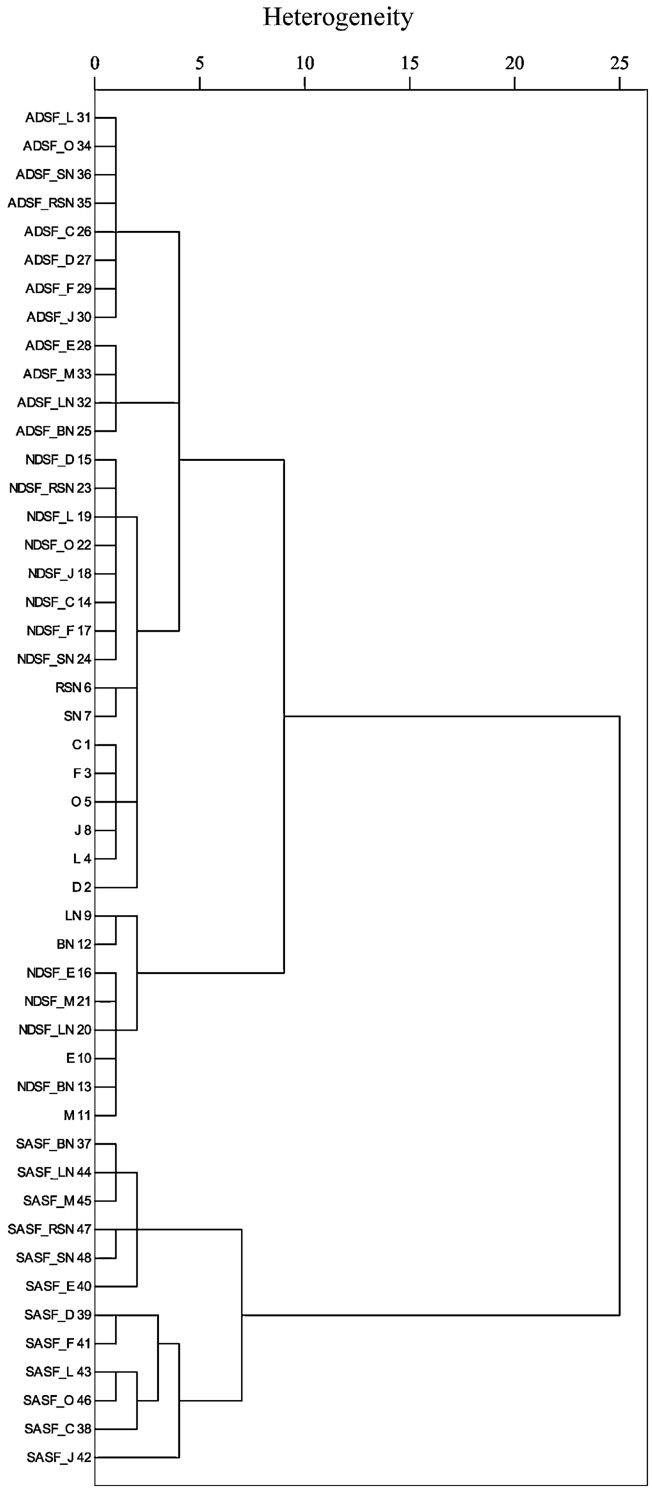

Fig. 5. HCA of FTIR spectra of 12 wood samples and different detergent fibers (NDSF, ADSF and SASF) (1900–800 cm^{-1} region). Spectra data were classified into classes according to their spectral heterogeneity. The first derivation and vector normalization was utilized as pretreatment method of the spectra.

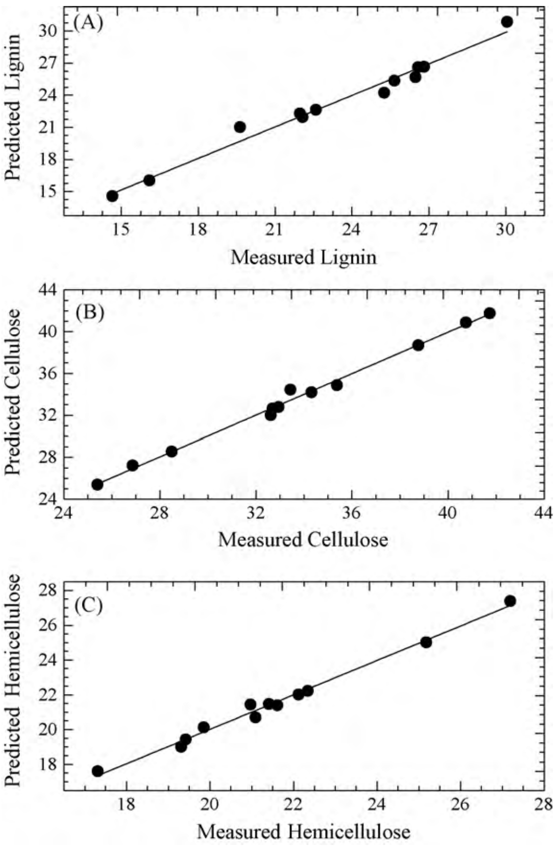

Fig. 6. Scatter plots of analytically determined *vs* predicted values for content of (A) lignin, (B) cellulose and (C) hemicellulose by PLS.

Table 4
Statistical data obtained in connection with calibration of the PLS model. Details are given of the explained variation of the *Y* matrix R^2Y, the predicted variation of the *Y* matrix Q^2, the correlation coefficient for the observed value as a function of the predicted value *R*, the root-mean-square error of estimation *RMSEE*.

Component	Range	Mean	R^2Y	Q^2	R	*RMSEE*
Lignin	15–30	23.2	0.980	0.487	0.979	1.51
Cellulose	25–42	33.6	0.994	0.781	0.993	0.96
Hemicellulose	17–27	21.5	0.990	0.829	0.989	0.62

larger than 0.2–0.3 (Bjarnestad & Dahlman, 2002; Eriksson et al., 1999). Table 4 reveals that all of the responses are described well by the model. Additionally, *RMSEE*, which is the root-mean-square error of estimation for the calibration set, is also employed to evaluate the goodness of fit (Aptula, Jeliazkova, Schultz, & Cronin, 2005). The *RMSEE* values can be viewed as a measure of the standard deviation and should therefore be as small as possible. In our case, *RMSEE* for determination of the major components of wood samples in this manner varied between 0.62 and 1.51 wt%, indicating the success of the PLS regression model.

4. Conclusions

In conclusion, FTIR spectroscopy coupled to chemometrics can be used for the qualitative and quantitative analysis of wood samples. A multivariate calibration model for FTIR spectroscopy has been developed. The total time required for analysis is about 15 min per sample (in addition to the time required for wood sample preparation) against 3–4 days required for Van Soest analysis. The use of multivariate analysis of PCA and HCA has enabled us to discriminate between softwoods and hardwoods by their infrared spectral data. A PLS model for FTIR spectroscopy has been developed for the analysis of the contents of lignin, cellulose and hemicellulose in wood samples and it shows that *RMSEE* for the contents of lignin, cellulose and hemicellulose are 1.51, 0.96 and 0.62%, respectively. The calibration model for FTIR spectroscopy is a possible, faster, alternative to Van Soest method and may allow the prediction of these parameters in other wood species and during wood processing.

Acknowledgements

Huilun Chen would like to acknowledge the financial support from the Chinese Scholarship Council (CSC) and Zhiguo Xia from China University of Geosciences (Beijing) for helpful discussions.

References

Adler, E. (1977). Lignin chemistry – Past, present and future. *Wood Science and Technology*, *11*, 169–218.

Alcantara, G. B., Honda, N. K., Ferreira, M. M. C., & Ferreira, A. G. (2007). Chemometric analysis applied in ^{1}H HR-MAS NMR and FT-IR data for chemotaxonomic distinction of intact lichen samples. *Analytica Chimica Acta*, *595*, 3–8.

Aptula, A. O., Jeliazkova, N. G., Schultz, T. W., & Cronin, M. T. D. (2005). The better predictive model: High q^2 for the training set or low root mean square error of prediction for the test set? *QSAR and Combinatorial Science*, *24*, 385–396.

Baratieri, M., Baggio, P., Fiori, L., & Grigiante, M. (2008). Biomass as an energy source: Thermodynamic constraints on the performance of the conversion process. *Bioresource Technology*, *99*, 7063–7073.

Bjarnestad, S., & Dahlman, O. (2002). Chemical compositions of hardwood and softwood pulps employing photoacoustic Fourier transform infrared spectroscopy in combination with partial least-squares analysis. *Analytical Chemistry*, *74*, 5851–5858.

Chen, J. H., Xie, M. Y., Fu, B. Q., Yang, M. F., & Wang, X. R. (2006). Principal component analysis and cluster analysis of inorganic elements in Panax *Quinque folium* L. *Spectroscopy and Spectral Analysis*, *26*, 1326–1329.

Colom, X., & Carrillo, F. (2005). Comparative study of wood samples of the northern area of Catalonia by FTIR. *Journal of Wood Chemistry and Technology*, *25*, 1–11.

Conrad, M. P. C., Smith, G. D., & Fernlund, G. (2003). Fracture of solid wood: A review of structure and properties at different length scales. *Wood and Fiber Science*, *35*, 570–584.

Dang, V. Q., Bhardwaj, N. K., Hoang, V., & Nguyen, K. L. (2007). Determination of lignin content in high-yield kraft pulps using photoacoustic rapid scan Fourier transform infrared spectroscopy. *Carbohydrate Polymers*, *68*, 489–494.

Dokken, K. M., Davis, L. C., & Marinkovic, N. S. (2005). Use of infrared microspectroscopy in plant growth and development. *Applied Spectroscopy Reviews*, *40*, 301–326.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., & Wold, S. (1999). *Multi- and megavariate data analysis: Principles and applications*. Umeå, Sweden: Umetrics Academy.

Evans, P. A. (1991). Differentiating “hard” from “soft” woods using Fourier transform infrared and Fourier transform spectroscopy. *Spectrochimica Acta Part A: Molecular Spectroscopy*, *47*, 1441–1447.

Faix, O., & Bottcher, J. H. (1993). Determination of phenolic hydroxyl group contents in milled wood lignins by FTIR spectroscopy applying partial least-squares (PLS) and principal components regression (PCR). *Holzforschung*, *47*, 45–49.

Fengel, D., & Shao, X. (1985). Studies on the lignin of the bamboo species *Phyllostachys makinoi* Hay. *Wood Science and Technology*, *19*, 131–137.

Ferraz, A., Baeza, J., & Duránt, N. (1991). Softwood biodegradation by an ascomycete *Chrysonilia sitophila* (TFB 27441 strain). *Letters in Applied Microbiology*, *13*, 82–86.

Ferraz, A., Baeza, J., Rodriguez, J., & Freer, J. (2000). Estimating the chemical composition of biodegraded pine and eucalyptus wood by DRIFT spectroscopy and multivariate analysis. *Bioresource Technology*, *74*, 201–212.

Gnansounou, E., Dauriat, A., Villegas, J., & Panichelli, L. (2009). Life cycle assessment of biofuels: Energy and greenhouse gas balances. *Bioresource Technology*, *100*, 4919–4930.

Kaar, W. E., & Brink, D. L. (1991). Summative analysis of nine common North American woods. *Journal of Wood Chemistry and Technology*, *11*, 479–494.

Malkavaara, P., & Alen, R. (1998). A spectroscopic method for determining lignin content of softwood and hardwood kraft pulps. *Chemometrics and Intelligent Laboratory Systems*, *44*, 287–292.

McKendry, P. (2002). Energy production from biomass (part 3): Gasification technologies. *Bioresource Technology*, *83*, 55–63.

Meder, R., Gallagher, S., Mackie, K. L., Bohler, H., & Meglen, R. R. (1999). Rapid determination of the chemical composition and density of Pinus radiata by PLS modelling of transmission and diffuse reflectance FTIR spectra. *Holzforschung*, *53*, 261–266.

Michell, A. J. (1988). Infra-red spectroscopy transformed – New applications in wood and pulping chemistry. *Appita Journal*, *41*, 375–380.

Morris, D. R., Steward, F. R., & Gilmore, C. A. (2000). Comparative analysis of the consumption of energy of two wood pulping processes. *Energy Conversion and Management*, *41*, 1557–1568.

Müller, G., Schöpper, C., Vos, H., Kharazipour, A., & Polle, A. (2009). FTIR-ATR spectroscopic analysis of chages in wood properties during particle- and fibreboard production of hard- and softwood trees. *Bioresources*, *4*, 49–71.

Owen, N. L., & Thomas, D. W. (1989). Infrared studies of “hard” and “soft” woods. *Applied Spectroscopy*, *43*, 451–455.

Pettersen, R. C. (1984). The chemical composition of wood. In R. M. Rowell (Ed.), *The chemistry of solid wood, advances in chemistry*. Washington, DC: American Chemical Society.

Popescu, C. M., Singurel, G., Popescu, M. C., Vasile, C., Argyropoulos, D. S., & Willfor, S. (2009). Vibrational spectroscopy and X-ray diffraction methods to establish the differences between hardwood and softwood. *Carbohydrate Polymers*, *77*, 851–857.

Prochnow, A., Heiermann, M., Plöchl, M., Amon, T., & Hobbs, P. J. (2009). Bioenergy from permanent grassland – A review. 2. Combustion. *Bioresource Technology*, *100*, 4945–4954.

Rodrigues, J., Puls, J., Faix, O., & Pereira, H. (2001). Determination of monosaccharide composition of *Eucalyptus globulus* wood by FTIR spectroscopy. *Holzforschung*, *55*, 265–269.

Sarkanen, K. V., Chang, H. M., & Ericsson, B. (1967). Species variation in lignins. I. Infrared spectra of guaiacyl and syringyl models. *Tappi Journal*, *50*, 572–575.

Schultz, T. P., Templeton, M. C., & McGinnis, G. D. (1985). Rapid determination of lignocellulose by diffuse reflectance Fourier transform infrared spectrometry. *Analytical Chemistry*, *57*, 2867–2869.

Schultz, T. P., & Burns, D. A. (1990). Rapid secondary analysis of lignocellulose: Comparison of near infrared (NIR) and Fourier transform infrared (FTIR). *Tappi Journal*, *73*, 209–212.

Senneca, O. (2007). Kinetics of pyrolysis, combustion and gasification of three biomass fuels. *Fuel Processing Technology*, *88*, 87–97.

Stefke, B., Windeisen, E., Schwanninger, M., & Hinterstoisser, B. (2008). Determination of the weight percentage gain and of the acetyl group content of acetylated wood by means of different infrared spectroscopic methods. *Analytical Chemistry*, *80*, 1272–1279.

Tan, P. N., Steinbach, M., & Kumar, V. (2006). Cluster analysis: Basic concepts and algorithms. In P. N. Tan, M. Steinbach, & V. Kumar (Eds.), *Introduction to data mining*. Boston, MA: Pearson Addison Wesley.

T 257 cm (1985). Sampling and preparing wood for analysis, TAPPI Standards.

Toivanen, T. J., & Alen, R. (2007). A FTIR/PLS method for determining variations in the chemical composition of birch (Betula pendula/B-publescens) stem wood. *Appita Journal*, *60*, 155–160.

Van Soest, P. J. (1963). Use of detergents in the analysis of fibrous feeds. II. A rapid method for the determination of fiber and lignin. *Journal of the Association of Official Analytical Chemists*, *46*, 829–835.

Van Soest, P. J., & Wine, R. H. (1967). Use of detergents in the analysis of fibrous feeds. IV. Determination of plant cell-wall constituents. *Journal of the Association of Official Analytical Chemists*, *50*, 50–55.

Wällbacks, L., Edlund, U., Nordén, B., & Berglund, I. (1991). Multivariate characterization of pulp using solid-state13 C NMR, FTIR, and NIR. *Tappi Journal*, *74*, 201–206.

Yang, Z. J., Zhang, L., Yang, R. W., Ding, C. B., Zhou, Y. H., & Wan, D. G. (2008). Principal component and cluster analysis of trace elements in Chinese herb *Salvia miltiorrhiza* and its relative species. *Spectroscopy and Spectral Analysis*, *28*, 2441–2445.