



Multivariate Calibration for Wood Chemistry by Fourier Transform Infrared Spectroscopy and Chemometric Methods

Chengfeng Zhou^{1,3*}, Wei Jiang^{2,1}, Brian K. Via^{1,3}

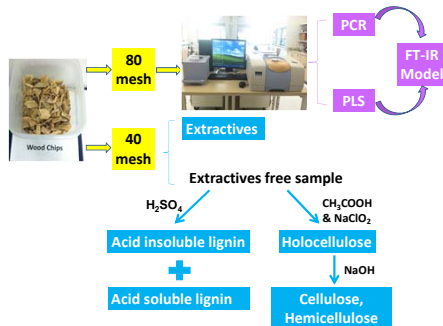
1. Forest Products Development Center, School of Forestry and Wildlife Sciences, Auburn University, 520 Devall Dr., Auburn, AL 36849, USA
 2. College of Textiles, Qingdao University, 308 Ningxia Road, Qingdao 266071, China
 3. Center for Bioenergy and Bioproducts, Biosystems Engineering, Auburn University, 520 Devall Dr., Auburn, AL 36849, USA
- *czz0024@auburn.edu



Abstract

This research addresses a rapid method to monitor hardwood chemical composition by applying Fourier transform infrared (FT-IR) spectroscopy, the precision in factor loadings during partial least squares (PLS) and principal components regression (PCR) analysis was also discussed. Standard laboratory chemistry methods were employed on a mixed genus/species hardwood sample set to collect the original data. PLS and PCR was utilized for model building and loadings investigation. Research identified the wavenumbers of extractives, lignin, cellulose and hemicellulose by analyzing the loading plots, and it also demonstrated that FT-IR has good capability to quantitative analysis wood chemical components.

Methods



Schematic diagram for wet chemistry and FT-IR analysis process.

Sample Preparation and Wood chemistry

All samples were collected from recently harvested hardwood trees. The wood samples were planed down to 3 mm thick wood chips and then stored for 2 weeks at 24 ± 1.5 °C and 45% ± 5% relative humidity. Air dried samples were ground to 40 mesh and 80 mesh for further utilization. Extractives and lignin were measured following National Renewable Energy Laboratory (NREL) standards. The cellulose, hemicellulose and holocellulose contents were also measured by traditional wet chemistry analysis.

NIR Acquisition and Chemometric analysis

A PerkinElmer (Waltham, MA, USA) spectrum 400 FT-IR/FT-NIR spectrometer was utilized for spectra collection. The spectra covered the range of 4,000–650 cm⁻¹ at a spectral resolution of 4 cm⁻¹. Each spectrum was collected from an average of 4 scans and no zero filling. PCR and PLS modules in Spectrum Quant + software was used for model construction. Models were executed on the unprocessed spectra (raw) and first derivative (FD).

Results and Discussion

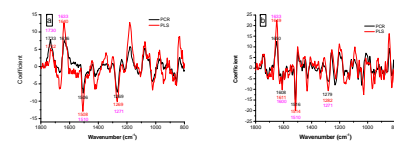


Figure 1. Coefficients by wavenumber for PCR and PLS for extractives prediction (a) when raw spectra was processed and (b) when a first derivative pretreatment was processed.

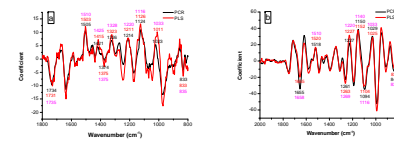


Figure 2. Coefficients by wavenumber for PCR and PLS for lignin prediction (a) when raw spectra was processed and (b) when a first derivative pretreatment was processed.

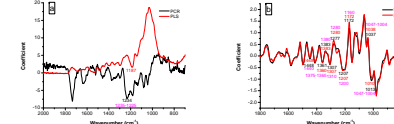


Figure 3. Coefficients by wavenumber for PCR and PLS for cellulose prediction (a) when raw spectra was processed and (b) when a first derivative pretreatment was processed.

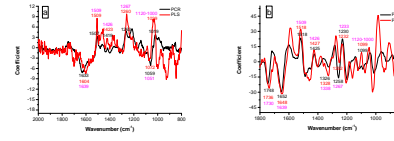


Figure 4. Coefficients by wavenumber for PCR and PLS for hemicellulose prediction (a) when raw spectra was processed and (b) when a first derivative pretreatment was processed.

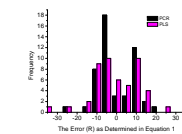


Figure 5. Frequency of R from PLS and PCR loadings of wood chemistry models. (Equation 1: $W - BA_e = R$)

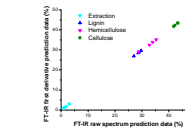


Figure 6. Chemical content (wt%) measured both in the laboratory and that predicted by FT-IR for validation samples.

Table 1. Hypothesis testing of Equations (2) and (3) through the (a) T-Test or (b) F-Test. * means the T-Test or F-Test were significant with 95% confidence.

(Equation 2 $H_0: \sigma_{PLS,R}^2 = \sigma_{PCR,R}^2$ Equation 3 $H_0: \sigma_{PLS,R}^2 > \sigma_{PCR,R}^2$)

	(a) T-Test		(b) F-Test	
	PCR	PLS	PCR	PLS
Mean R	0.06	-0.49	0.06	-0.49
Variance	55.5	92.6	55.5	92.6
Standard deviation	7.45	9.62	7.45	9.62
95% CI	0.06 ± 2.1	-0.49 ± 2.8	0.06 ± 2.1	-0.49 ± 2.8
Observations	49	49	49	49
Degrees of freedom	48	48	48	48
F value	1.67		P (F<I) one tail	0.0398*
P for T-Test	0.0796*		F Critical one-tail	0.62

Table 2. Calibration and predictive results of FT-IR based multivariate models.

Algorithm	Chemistry	Raw spectra			First derivative		
		R ²	RMSEP	RPD	R ²	RMSEP	RPD
PLS	Extraction	73.51	1.19	1.19	86.66	0.34	4.18
	Lignin	87.76	1.05	2.30	90.06	0.50	4.83
	Cellulose	56.74	1.38	1.53	85.62	0.80	1.72
	Hemicellulose	79.31	2.25	1.73	92.90	1.90	2.04
PCR	Extraction	57.93	0.87	1.69	71.04	0.59	2.41
	Lignin	73.70	1.51	1.60	87.36	0.82	2.94
	Cellulose	52.61	1.59	1.32	48.15	0.94	1.34
	Hemicellulose	30.09	4.14	0.94	52.69	3.34	1.16

Conclusions

- PLS was found to provide better predictive diagnostics while PCR exhibited a more precise estimate of loading peaks which makes PCR better for interpretation.
- Application of the first derivative appeared to assist in improving both PCR and PLS loading precision.
- This work is important because it suggests that what is best for prediction is not best for model interpretation.

Acknowledgment

- Thanks to my advisor Dr. Brian K. Via and committee members for their technical and professional supports.
- This work was most supported by the Agriculture and Food Research Initiative (AFRI) CAP – "Southeast Partnership for Integrated Biomass Supply Systems"; the Department of Energy grant titled, "High Tonnage Forest Biomass Production Systems from Southern Pine Energy Plantations".