



American Society of  
Agricultural and Biological Engineers

*An ASABE Meeting Presentation*

*Paper Number: 066155*

## **Fast biomass compositional analysis using Fourier Transform Near-infrared Technique**

<sup>1</sup>Ye, X.P., <sup>1</sup>S. Liu, <sup>1</sup>L. Kline, <sup>1</sup>D.G. Hayes, <sup>1</sup>A.R. Womac, <sup>2</sup>S. Sokhansanj, and <sup>3</sup>S. Narayan.

<sup>1</sup> Department of Biosystems Engineering & Soil Science, The University of Tennessee

<sup>2</sup> Department of Chemical and Biological Engineering, University of British Columbia

<sup>3</sup> First American Scientific Corp.

**Written for presentation at the  
2006 ASABE Annual International Meeting  
Sponsored by ASABE  
Portland Convention Center  
Portland, Oregon  
9 - 12 July 2006**

**Abstract.** The objectives of this research were to determine the variation of chemical composition across botanical fractions of cornstover, and to use Fourier Transform Near-infrared (FT-NIR) techniques to qualitatively classify separated cornstover fractions, and develop calibration model for the quantitative analysis of chemical compositions of cornstover. Large variations of biomass chemical composition for wide calibration ranges were achieved by manually separating the cornstover samples into six botanical fractions, and their chemical compositions were determined by conventional wet chemical analyses, which proved that chemical composition varies significantly among different botanical fractions of cornstover. Husk, followed by rind and pith, has the highest sugar (glucan+xylan) content; node has the lowest sugar content. Based on FT-NIR spectra acquired on the biomass, classification by Soft Independent Modeling of Class Analogy (SIMCA) was employed to conduct qualitative classification of cornstover and Partial Least Square (PLS) regression was used for quantitative chemical composition analysis. SIMCA was demonstrated successfully in classifying botanical fractions of cornstover. The developed PLS models yielded root mean square error of prediction (RMSEP) of 1.058, 1.539, 0.987, and 1.435 for glucan, xylan, lignin, and ash, respectively. The FT-NIR techniques in combination with multivariate analysis are very useful to biomass feedstock suppliers, bio-ethanol manufacturers, and bio-power producers.

**Keywords.** cornstover, botanical fractions, FT-NIR, fast analysis, PCA, PLS, SIMCA

---

The authors are solely responsible for the content of this technical presentation. The technical presentation does not necessarily reflect the official position of the American Society of Agricultural and Biological Engineers (ASABE), and its printing and distribution does not constitute an endorsement of views which may be expressed. Technical presentations are not subject to the formal peer review process by ASABE editorial committees; therefore, they are not to be presented as refereed publications. Citation of this work should state that it is from an ASABE meeting paper. EXAMPLE: Author's Last Name, Initials. 2006. Title of Presentation. ASABE Paper No. 06xxxx. St. Joseph, Mich.: ASABE. For information about securing permission to reprint or reproduce a technical presentation, please contact ASABE at [hq@asabe.org](mailto:hq@asabe.org) or 269-429-0300 (2950 Niles Road, St. Joseph, MI 49085-9659 USA).

---

## Introduction

Biomass feedstock composition determines the theoretical yield from a biochemical conversion process, and can thereby have a significant impact on conversion process economics. Previous work (Thomas et al., 2001) demonstrated that cornstover composition could vary over a wide range that can change the Minimum Ethanol Selling Price (MESP) of bioethanol produced over a range of \$1.04 – 1.36/gallon, assuming all other process variables remain constant across feedstock batches. Thomas et al. (2001) concluded that individual corn plants do not have a uniform composition across different parts of the anatomy, and cornstover composition is highly variable and influenced by a wide variety of genetic and environmental factors. These observations of biomass variability will probably hold true for other feedstock materials, such as wheat straw, switchgrass and hybrid poplars. On the other hand, the major chemical constituents of biomass -- namely cellulose, hemicelluloses, and lignin that have distinctly different uses -- exist in botanical fractions of leaves, nodes and internodes with different concentrations. Physical pre-fractionation of botanical fractions of biomass concentrates higher value components, decreases bulk for wet separation processes, reduces drying costs, and improves transportation and utilization efficiencies of a voluminous, chemically diverse feedstock.

It is difficult if not impossible to control the compositional variability of biomass feedstock. However, it is feasible to monitor the variability and accordingly adjust process parameters. Current methods for chemical characterization of biomass feedstock, process intermediates and residues are not applicable in a commercial setting because they are very expensive (labor intensive) and cannot provide analysis information in a time frame useful for process control. Recently, Hames et al. (2003) demonstrated that dispersive Near Infrared (NIR) Spectroscopy in reflectance mode with “projection to latent structures (PLS)” regression provided a rapid and accurate analysis of cornstover compositions. However, we found few research papers that apply Fourier Transform Near Infrared (FT-NIR) spectroscopy, a newer NIR technology, to biomass feedstock analysis.

Three key factors determine a successful application of NIR techniques for fast chemical characterization—accurate and repeatable NIR spectral acquisition, reliable calibration data, and robust chemometric analysis. While FT-NIR has the advantage over dispersive NIR in the aspect of the first factor (Griffiths and Haseth, 1986), more work needs to be done to build up reliable calibration data and robust chemometric methods. Wide ranges of variation of calibrated chemical compositions are needed for a robust chemometric analysis. Diverse and unique calibration dataset will enhance calibration range and depth (Hames et al., 2003).

Using FT-NIR spectroscopy and chemometric analysis, the objectives of this study are to:

1. determine the variation of chemical composition across botanical fractions of cornstover;
2. qualitatively classify separated cornstover fractions;
3. develop calibration model for the quantitative analysis of chemical compositions of cornstover.

## Materials & Methods

Large variation of chemical composition for wide calibration ranges can be achieved by manually separating the samples into specific botanical fractions after their distinction in chemical composition is proven. This alternative solution breaks through the geographical and temporal limitation, and even endows researches some flexibility to prepare certain samples with desired compositional proportion. Furthermore, the research result will also provide justification to physical pre-fractionation of botanical fractions of biomass in order to make full use of agricultural residuals. For example, the botanical

fractions with higher glucan content should be selected as feedstock for fermentation, while those with high lignin content should be good feedstock for co-firing in a power plant.

The cornstover (Dekalb 689) collected at Tennessee Agriculture Experiment Station was manually separated into six different botanic fractions, namely, nodes, leaves, internodal piths, internodal rinds, sheath and husks (Figure 1).

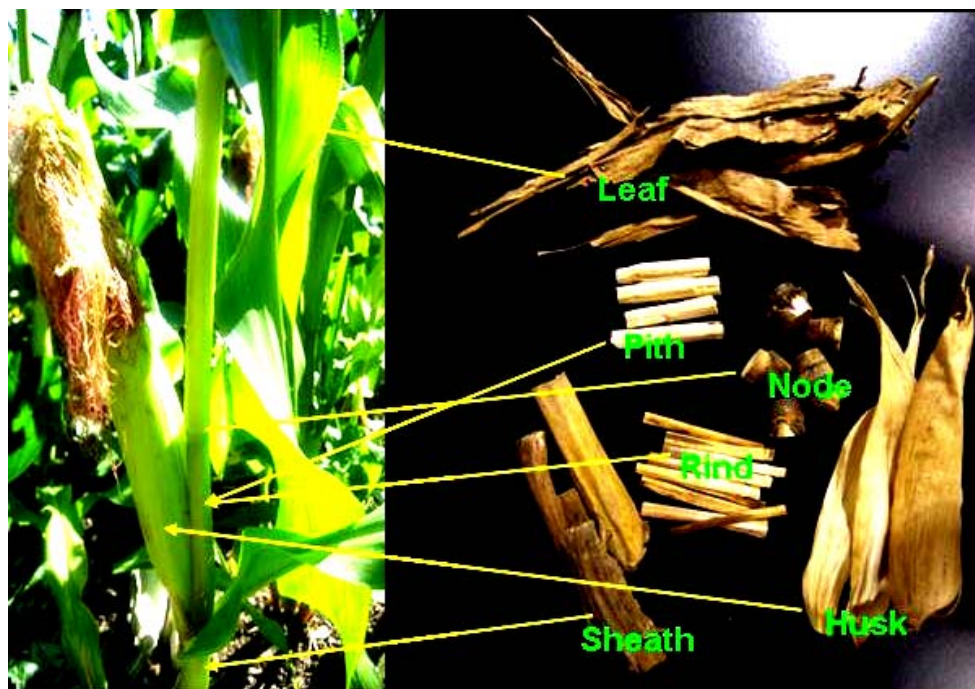


Figure 1. Botanical fractions of cornstover

A total of 29 cornstover samples were prepared. All the samples were ground with a Wiley Mini-Mill (Thomas Scientific, Swedesboro, NJ) through a number 40 mesh sieve and put into a spinning sample cup. FT-NIR spectra were collected in the diffuse reflectance mode on a spectrometer (Excalibur 3100, Varian Inc., Palo Alto, CA) equipped with a NIR integrating sphere (IntegratIR™, PIKE Technologies, Madison, WI). Each spectrum is the average of 256 co-additions of scans while the sample is rotating. Kubelka-Monk conversion was applied automatically to the output spectra by the Varian Resolution Pro® software, since it yields the best result in the mode of diffuse reflectance (Dahm, 1995; Griffiths, 1995). The spectra cover a range of 10000~4000  $\text{cm}^{-1}$  with a spectral resolution of 8  $\text{cm}^{-1}$ , generating 1558 data points.

In near infrared spectroscopy, influence of compounds usually reflects throughout the whole spectrum region instead of only certain specific region. As a result, visual analysis of NIR spectra does not reveal fingerprints of chemicals in the near range (4000-10000  $\text{cm}^{-1}$ ) as they are in the mid range (400-4000  $\text{cm}^{-1}$ ), which means no significant peaks. However, a NIR spectrum is unique representation of a substance or a mixture. Signals from bonds such as C-O, C=O, O-H, and N-H, as well as the overlap of their first and second overtones are included within this region. The composition-informative spectra need to be combined with multivariate mathematical methods so that structural information can be extracted. Therefore, multivariate models are, besides being predictive, also useful in the interpretation of overlapping and wide overtone bands.

After FT-NIR scanning, conventional wet chemistry analyses were conducted following the Standard Biomass Analytical Procedures developed by the National Renewable Energy Laboratory (NREL). Figure 2, which is also a justification of FT-NIR analysis, summarizes and contrasts the wet chemistry analysis and FT-NIR processes.

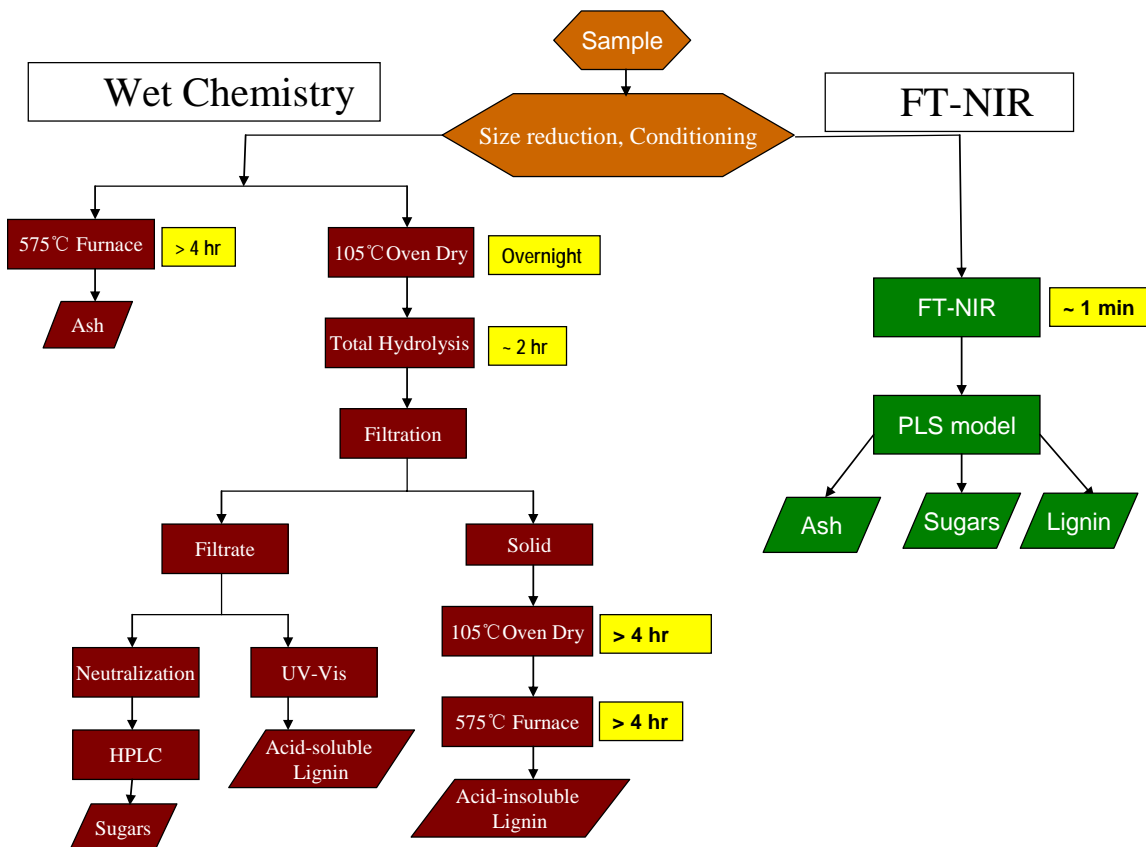


Figure 2. Processes of wet chemistry analysis vs. FT-NIR fast analysis

Carbohydrates (glucan and xylan) were measured using a HPLC system (SD-200, Rainin Instruments, Woburn, MA) equipped with an evaporative light-scattering detector and a Prevail carbohydrate ES 50 column (MKIII, Alltech Associates Inc., Deerfield, IL). The ash content of the acid-insoluble material was determined through combustion at 575°C, and acid-insoluble lignin is found by mass balance. The acid-soluble lignin was determined using UV spectrophotometry (Varian Prostar 325 UV-Vis Detector). Two repetitions were carried out for each corresponding FT-NIR test.

All the multivariate analyses of FT-NIR spectra for qualitative classification and quantitative chemical composition prediction were conducted using a commercial software Unscrambler v. 9.6 (CAMO Software Inc., Woodbridge, NJ). Classification by Soft Independent Modeling of Class Analogy (SIMCA) was employed to do qualitative classification and Partial Least Square (PLS) regression was used for quantitative chemical composition prediction.

## Results & Discussion

### *Wet chemistry analysis*

The chemical compositions analyzed for each botanical fractions of cornstover are presented in figure 3.

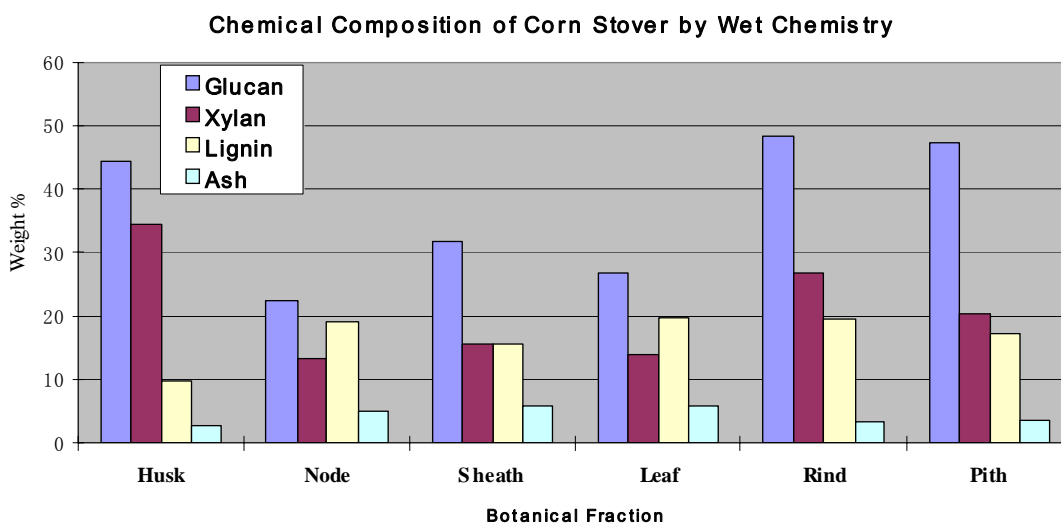


Figure 3. Chemical composition (% w/w) by wet chemistry analyses

It is evident from figure 3 that chemical composition varies significantly among different botanical fractions of cornstover. Husk, followed by rind and pith, has the highest sugar (glucan+xylan) content. On the other end, node has the lowest sugar content. Since husk, rind and pith all have comparatively high sugar content and comparatively easier to be separated, it is feasibly profitable for biomass suppliers to emphasize on these botanical fractions. Since rind and pith have similar chemical composition, treating them as whole inter-nodal fraction would simplify the manufacture process.

### ***Qualitative classification***

Principle component analysis (PCA) of FT-NIR spectra showed promise to qualitatively classify biomass feedstock. Presented in figure 4, the score plot of PCA analysis shows that the six botanical fractions of cornstover can be differentiated by the first two principle components (PC). The repetitions of each botanical fraction cluster together, and the direction of the first PC approximates the increase of glucan concentration.

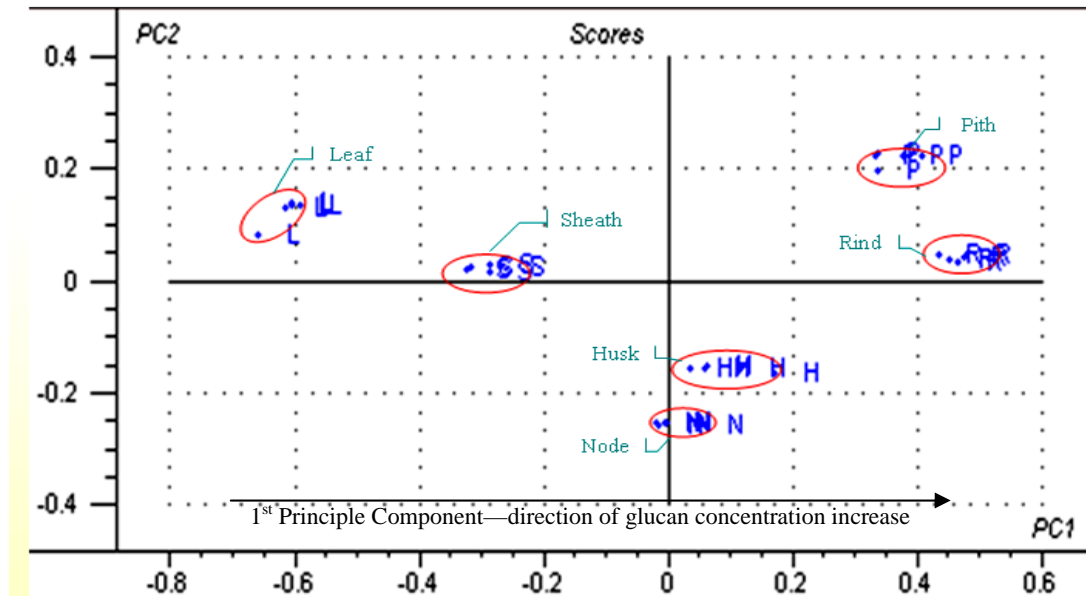


Figure 4. Score plot of PCA on FT-NIR spectra of cornstover fractions

Based on the PCA analysis, a supervised classification method, Soft Independent Modeling of Class Analogy (SIMCA), was used to differentiate botanical fractions of cornstover. SIMCA focuses on modeling the similarities between members of the same class. A new sample will be recognized as a member of a class if it is similar enough to the other members; otherwise it will be rejected. This approach computes the distance of each observation from the model with respect to the known samples used to build the PCA models. Figure 5 shows the average model distances among the six botanical fractions.

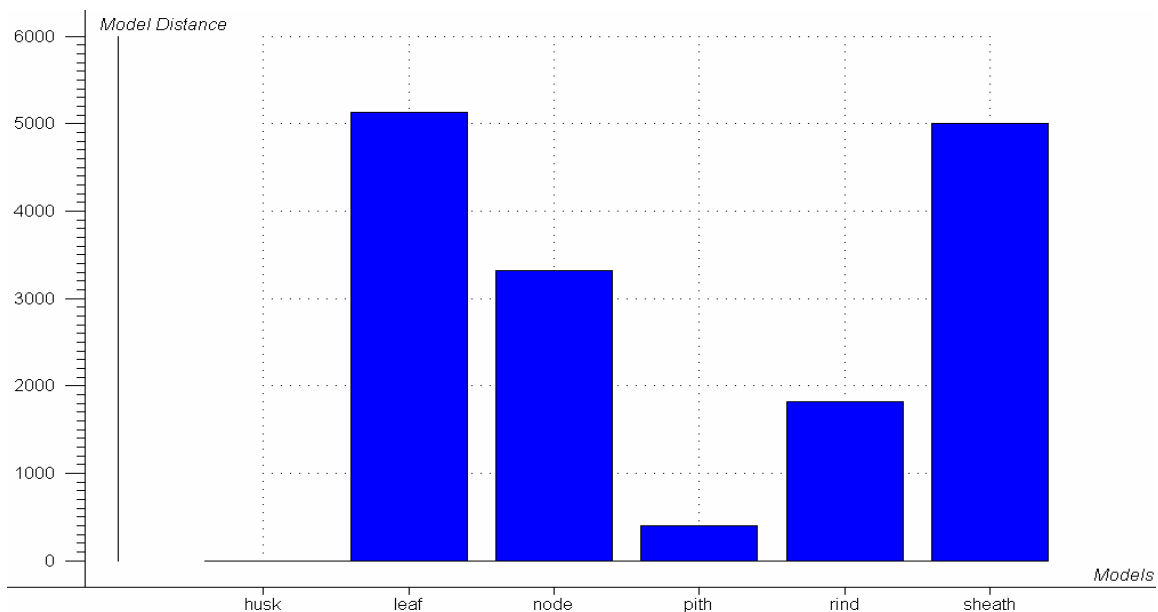


Figure 5. SIMCA model distance among the six botanical fractions with husk as the reference

This model-to-model distance plot gives the degree of dissimilarity among models with husk as the reference. Sixty new FT-NIR spectra were collected, with 10 from each of the 6 botanical fractions to validate the SIMCA methodology. The spectral distinction among the 6 botanical fractions is so significant that without data pretreatment, SIMCA correctly classified all of the 60 new samples ( $\alpha=0.05$ ). Table 1 presents part of the classification table produced by the Unscrambler v. 9.6 (CAMO Software Inc., Woodbridge, NJ), in which a new sample to be classified is assigned a “\*”, corresponding to the predicted botanical fraction as the header of the columns.

Table 1. Classification result of testing samples by SIMCA

Sample	husk	leaf	node	pith	rind	sheath
H	*					
N			*			
S						*
L		*				
R					*	
P				*		

This SIMCA classification method is useful in feedstock selection for a bioconversion plant, and it can also be a powerful tool in managing large varieties of biomass. SIMCA can also be conducted on the quality classification.

### *Quantitative prediction of chemical composition*

Partial least square (PLS) regression was performed using the average of two repetitions from wet chemistry analysis as calibration data and the average of five corresponding FT-NIR repetitions as predictive variables. Full cross-validation was used to check how well the PLS model would perform for future samples and allow for estimation of the prediction error in future predictions. The full FT-NIR spectra including 1558 data points are shown in figure 6.

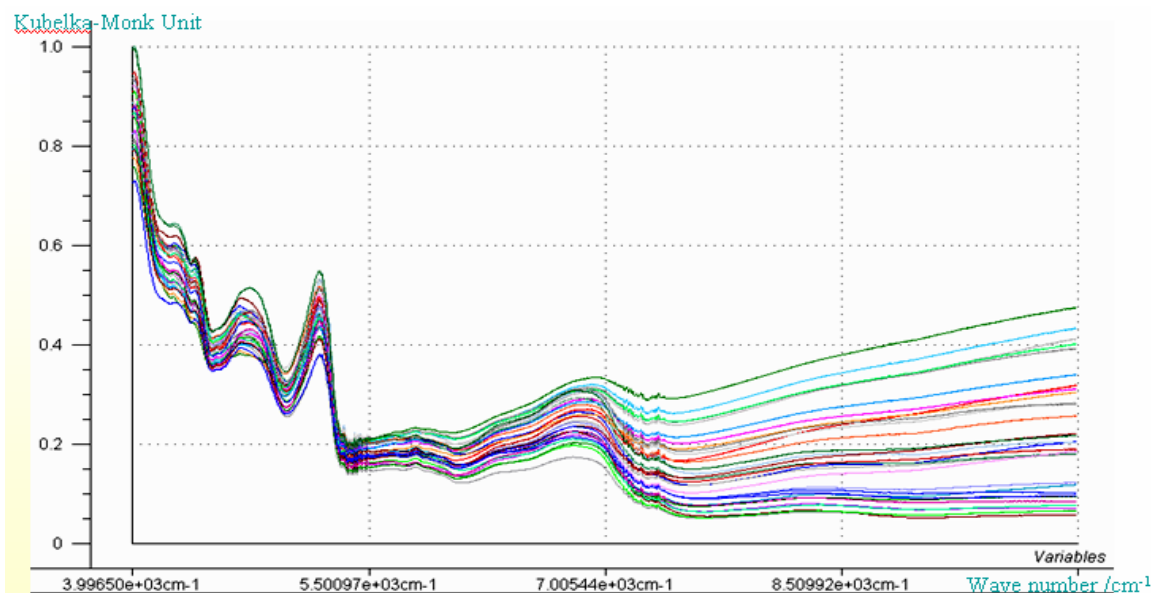


Figure 6. Original FT-NIR spectra of calibration samples.

Water vapor interference and light scattering effect are evident in figure 6. Therefore, data reduction and certain transformations appeared important for the succeeding PLS regression procedure. Two spectral regions ( $5134\text{-}5520\text{cm}^{-1}$ ,  $7063\text{-}7271\text{cm}^{-1}$ ) were first zapped out of the spectra, because the water vapor generates random and irregular spectral signal in the two regions. Several data pretreatment methods were tested on the spectral dataset, including Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV), first derivative and second derivative. The second derivative method generated the best results and is also consistent with several previous studies (Czuchajowska et al., 1992; Tsuchikawa et al., 2003). Root mean square error of prediction (RMSEP) was used to evaluate the model quality. The predicted chemical composition vs. wet chemistry measurements as results generated by the PLS-regression procedure are presented in figure 7, together with PLS model statistics and regression lines. Figure 8 summarizes the results together with target line.

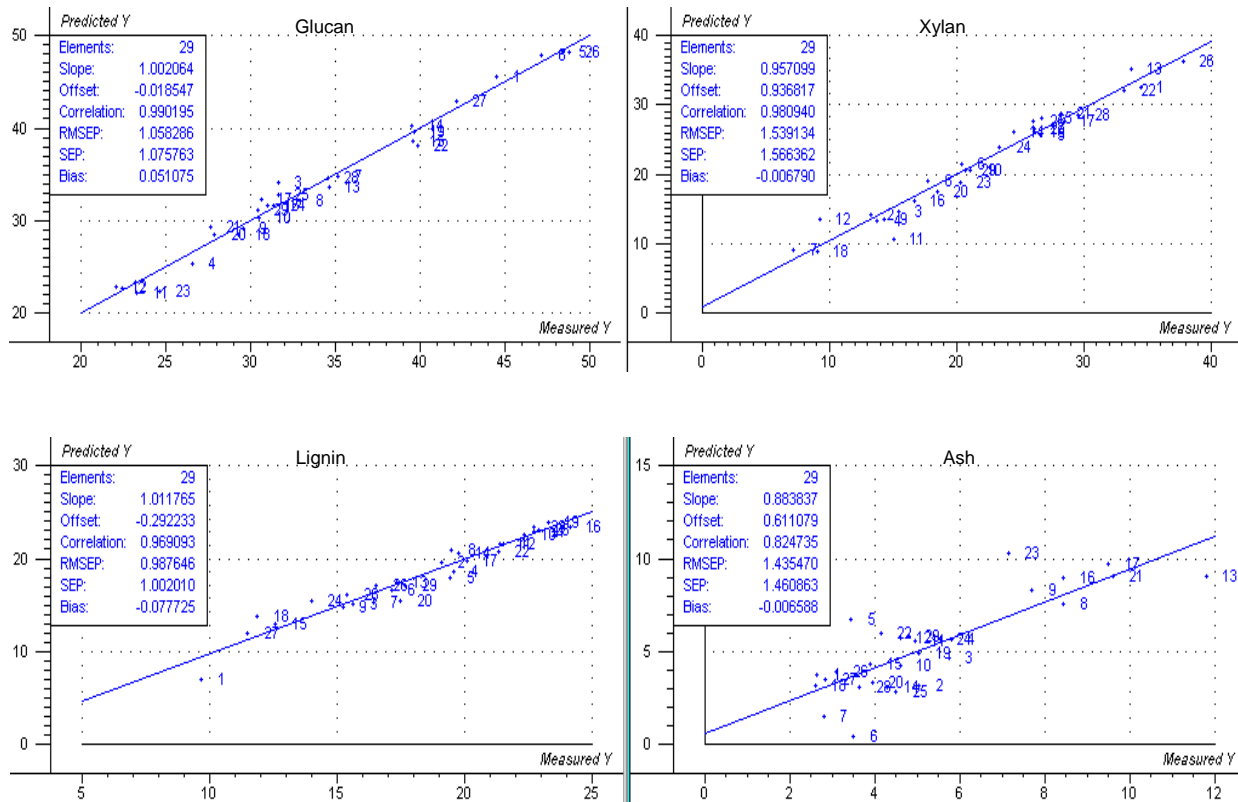


Figure 7. Chemical composition (% w/w) predicted by FT-NIR techniques vs. measured by wet chemistry



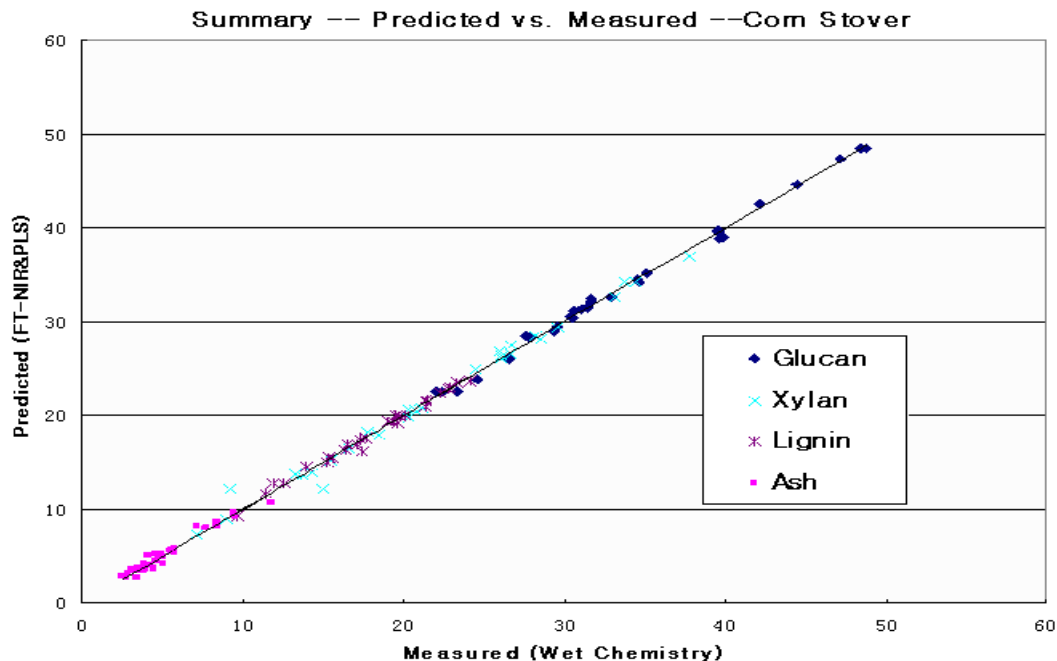


Figure 8. Overall results of FT-NIR/PLS analysis (predicted vs. measured in % w/w)

For glucan, xylan and lignin, all the calibration data points scattered around the regression lines whose slope is very close to one. The RMSEP for these three chemicals are small, demonstrating the good predicting capability for these three major constituents. However, the prediction of ash is not as accurate as the other three, with a slope of 0.8838 and RMSEP of 1.4354, relatively large compared to the average ash content. Controversial conclusions have been found in our literature review concerning the utilization of NIR in predicting ash content. The major concern is that NIR does not directly interact with most of the ash forming atom species (Lestander and Rhén, 2005), although successful application has been found in the literature for an organic substance (Woo et al., 1999). More research is needed to address this issue so that FT-NIR can be used to accurately predict ash content in biomass.

## Conclusion & Summary

Fourier Transform Near Infrared spectroscopy together with multivariate analysis was applied to qualitatively classify separated cornstover fractions and quantitatively analyze chemical compositions of cornstover. It was found in this study that manual separation of botanical fractions of cornstover sufficiently created representative variation of chemical concentrations for multivariate analysis, which bypassed expensive and time-consuming collection of diverse sample throughout years and nationwide locations. This alternative sampling method makes biomass research more flexible and accessible to all the research groups. This study shows great potential of the application of FT-NIR spectroscopic technique in combination with chemometrics methodology in fast classifying biomass and measuring its chemical constituents.

SIMCA methodology was demonstrated as a reliable method in cornstover classification, even without FT-NIR data pretreatment. Although this study emphasized on classifying the botanical fractions of cornstover, the method can be extended to categorize biomass feedstocks by their heating value and fermentable sugar content, etc.. Rapid analysis method through PLS showed potential in predicting carbohydrates and lignin. At this era when cellulosic ethanol and biomass power are striding into the energy mainstream, a rapid biomass compositional analysis method that instantly provides information of

heating value, and contents of sugars, lignin, and ash, is in great demand for biomass feedstock suppliers, bio-ethanol manufacturers, and bio-power producers.

## References

- Hames, Bonnie R., Steven R. Thomas, Amie D. Sluiter, Christine J. Roth, and David W. Templeton (2003). *Applied Biochemistry and Biotechnology*, 105-108, 5-16
- Dahm, Donald J., Kevin D. Dahm, (1995). Letter: math pretreatment of NIR reflectance data:  $\log(1/R)$  vs  $F(R)$ . *Journal of Near Infrared Spectroscopy*, 3(1), 53-59.
- Griffiths, P.R. and de Haseth, J.A. (1986). *Fourier Transform Infrared Spectrometry*, J. Wiley & Sons, New York
- Griffiths, P.R. (1995). Letter: practical consequences of math pre-treatment of near infrared reflectance data:  $\log(1/R)$  vs  $F(R)$ . *Journal of Near Infrared Spectroscopy*, 3(1): 60-62
- Tsuchikawa, S., A.Murata, M.Kohara and K.Mitsui (2003), *Near Infrared spectrosc.* 11, 401-405.
- Lestander, T.A. and C. Rhe'n (2005). Multivariate NIR spectroscopy models for moisture, ash and calorific content in biofuels using bi-orthogonal partial least squares regression. *Analyst*, 130, 1182-1189.
- Thomas S, Ruth M, Macfarlane B, and Hames, B, 2001. Estimating the Effects of Compositional Variation on Biomass Conversion Process Economics. Biofuels Program FY01 C Milestone report #FY01-273, National Renewable Energy Lab.  
<http://devafdc.nrel.gov/bcfc/oraclefind3.cgi?5821>
- Woo Y.-A., Kim H.-J., Park S.-Y., Chang S.-Y., and Chung H (1999). Feasibility of Rapid and Nondestructive Determination of Ash Content in Cornu Cervi Parvum Using Near-Infrared Spectroscopy. *Microchemical Journal*, Volume 63, Number 1, September 1999, pp. 154-160.
- Czuchajowska, Z., J.Szczodrak, and Y.Pomeranz (1992). Characterization and estimation of barley polysaccharides by near-infrared spectroscopy. I. Barleys, starches, and  $\beta$ -D-glucans. *Cereal Chemistry*, 69(4), 413-418