

Compositional Analysis of Lignocellulosic Feedstocks. 2. Method Uncertainties

DAVID W. TEMPLETON,* CHRISTOPHER J. SCARLATA, JUSTIN B. SLUITER, AND
EDWARD J. WOLFRUM

National Bioenergy Center, National Renewable Energy Laboratory, 1617 Cole Boulevard,
Golden, Colorado 80401-3305

The most common procedures for characterizing the chemical components of lignocellulosic feedstocks use a two-stage sulfuric acid hydrolysis to fractionate biomass for gravimetric and instrumental analyses. The uncertainty (i.e., dispersion of values from repeated measurement) in the primary data is of general interest to those with technical or financial interests in biomass conversion technology. The composition of a homogenized corn stover feedstock (154 replicate samples in 13 batches, by 7 analysts in 2 laboratories) was measured along with a National Institute of Standards and Technology (NIST) reference sugar cane bagasse, as a control, using this laboratory's suite of laboratory analytical procedures (LAPs). The uncertainty was evaluated by the statistical analysis of these data and is reported as the standard deviation of each component measurement. Censored and uncensored versions of these data sets are reported, as evidence was found for intermittent instrumental and equipment problems. The censored data are believed to represent the "best case" results of these analyses, whereas the uncensored data show how small method changes can strongly affect the uncertainties of these empirical methods. Relative standard deviations (RSD) of 1–3% are reported for glucan, xylan, lignin, extractives, and total component closure with the other minor components showing 4–10% RSD. The standard deviations seen with the corn stover and NIST bagasse materials were similar, which suggests that the uncertainties reported here are due more to the analytical method used than to the specific feedstock type being analyzed.

KEYWORDS: Summative biomass compositional analysis; uncertainty; laboratory analytical procedure; LAP; corn stover; bagasse; lignocellulose

INTRODUCTION

The most common procedures for characterizing the chemical components of lignocellulosic feedstocks use a two-stage sulfuric acid hydrolysis to fractionate biomass for gravimetric and instrumental analyses. These methods have been developed and refined since their original use in the pulp and paper industry. For a review of the lineage of these methods, see our companion paper in this issue (1). These analytical methods have enabled complete summative analysis of biomass feedstocks to become a routine part of biofuels research and development.

The data from these methods are used to calculate mass balance and process yields and for technoeconomic analysis. These results affect evaluations of process configuration, reactor design, and process performance. At the National Renewable Energy Laboratory (NREL) we use these data to support technoeconomic analyses of biomass to ethanol processes (2). Therefore, the uncertainty (i.e., dispersion of values from repeated measurement) in the primary data is of general interest to those with technical or financial interests in biomass conversion technology.

*Corresponding author [e-mail David.Templeton@nrel.gov; phone (303) 384-7764; fax (303) 384-6877].

Earlier versions of these methods were used in a collaborative international study of biomass reference materials nearly two decades ago (3). In our previous paper (1), we described our current laboratory analytical procedures (LAPs) that are used to determine the component concentrations in lignocellulosic biomass at NREL.

These biomass compositional analysis methods are empirical in nature; the final results depend on how the method is run. Empirical methods can be compared using reference materials such as the National Institute of Standards and Technology (NIST) biomass reference materials (RMs 8491–4) (4). It is difficult to compare method errors (measured difference from a "true" value) without an accepted true standard. Taylor and Kuyatt state, "A measurement result is complete only when accompanied by a quantitative statement of its uncertainty. The uncertainty is required in order to decide if the result is adequate for its intended purpose and to ascertain if it is consistent with other similar results" (5).

For this work, we measured the composition of a homogenized corn stover feedstock (154 replicate samples in 13 batches, by 7 analysts in 2 laboratories) along with a NIST reference bagasse, run as a control, using our suite of analytical methods (NREL

Table 1. Corn Stover Composition Summary Statistics

	glucan	xylan	galactan	arabinan	acetyl	lignin	ash	protein	sucrose	water extractives (others)	ethanol extractives	total extractives	total structurals	total
Censored Data														
mean	34.0	19.2	1.0	2.5	2.9	12.3	4.7	2.2	7.2	8.0	2.6	22.5	74.4	96.9
SD	0.5	0.3	0.1	0.2	0.3	0.2	0.2	0.1	0.5	1.1	0.1	0.6	1.0	1.0
pooled SD	0.4	0.3	0.1	0.1	0.1	0.1	0.1	NA	0.4	0.9	0.1	0.4	0.6	0.5
N	142	146	154	154	154	146	154	12	154	154	154	129	129	129
Uncensored Data														
mean	34.2	19.3	1.0	2.5	2.9	12.3	4.7	2.2	7.2	8.0	2.6	22.1	74.9	97.0
SD	0.9	0.5	0.1	0.2	0.3	0.5	0.2	0.1	0.5	1.1	0.1	1.2	1.7	1.5
pooled SD	0.8	0.5	0.1	0.1	0.1	0.5	0.1	NA	0.4	0.9	0.1	0.8	1.4	1.3
N	154	154	154	154	154	154	154	12	154	154	154	154	154	154

Table 2. NIST Bagasse (RM 8491) Composition Summary Statistics

	glucan	xylan	galactan	arabinan	acetyl	lignin	ash	protein	sucrose	water extractives (others)	ethanol extractives	total extractives	total structurals	total
mean	39.0	21.8	0.8	1.8	3.3	24.8	3.9	0.5	0.7	2.7	1.9	5.7	95.1	100.8
SD	0.5	0.4	0.2	0.2	0.2	0.5	0.2	0.1	0.3	0.5	0.1	0.4	1.1	1.0
N	14	14	14	14	14	14	14	10	14	14	14	14	14	14

LAPs). We evaluate the uncertainty of our methods by the statistical analysis of these data, and we report the standard deviations of each component measurement.

MATERIALS AND METHODS

The corn stover sample (Pioneer hybrid 33B51) was harvested in the fall of 2003 from a farm in northeastern Colorado. We acquired a large quantity of this material, and it has been used as a feedstock for numerous pretreatment, saccharification, and fermentation experiments at NREL for several years. The stover was knife-milled to pass a $\frac{1}{4}$ in. (6.35 mm) round screen (Reduction Technologies, model 10 × 12, Leeds, AL) and stored in totes. A 5 gal (18.9 L) sample was taken from one of the totes and ground in a Wiley knife mill to pass through a 2 mm round screen. We sieved the milled stover and collected the -20/+80 mesh fraction. We coned and quartered about 900 g of this sieved corn stover three times and then distributed the well-mixed material into 24 glass bottles. To assess the homogeneity of the material, we selected four bottles at random and predicted the composition of eight subsamples of each bottle were using a near-infrared (NIR) calibration model (6). We saw no statistically significant differences in predicted composition among the four bottles. The sugar cane bagasse sample was purchased from NIST as RM 8491 (4) and was analyzed as received. We analyzed the bagasse sample in parallel with the corn stover samples.

Description of Biomass Compositional Analysis Methods. Complete biomass compositional analysis of the feedstocks was performed using methods described in detail in our companion paper (1), available from ASTM International (7) and on the Web (8). Briefly, batches of 12 replicate biomass samples (approximately 1.5 g in an 11 mL cell) were sequentially extracted with water and ethanol using automated solvent extractors (ASE200, Dionex, Sunnyvale, CA). The batch size (12 corn stover plus 1 bagasse) was set by the available number of cells in the extractor. The insoluble portion from that extraction was air-dried and then subjected to a two-stage sulfuric acid hydrolysis (1 h at 30 °C/72 wt % sulfuric acid, followed by 1 h at 121 °C/4 wt % sulfuric acid in an autoclave). Sulfuric acid (72 wt %, R8191600-1A, Ricca Chemical Co., Arlington, TX) was used as received. Sugar recovery standards (SRS) were included in the second stage of the hydrolysis to account for the loss of carbohydrates to degradation products. We used glass pressure vessels (part 8648-30) and Teflon caps (5845-47, ACE Glass Inc., Vineland, NJ) for both hydrolyses. We separated the lignin-rich residue from the analytical hydrolysate liquor using vacuum filtration through ceramic filtering crucibles (60531, CoorsTek, Golden, CO). The filter crucibles are connected to the filter flask using a rubber gasket (24065-000, VWR, West Chester, PA) after removal of the accompanying glass stem.

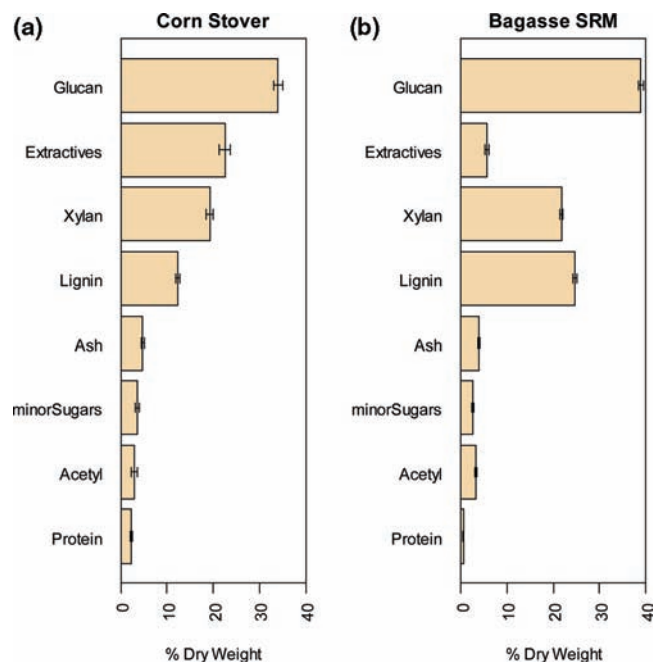


Figure 1. Summary compositional analysis results on a dry weight basis from this study for (a) corn stover and (b) NIST standard reference material (RM) 8491 (bagasse). Data were produced by 7 analysts working in 2 different laboratories. The constituents are shown in decreasing constituent value in corn stover. Error bars show ± 1 standard deviation. The corn stover material is slightly lower in glucan and xylan, much lower in lignin, and much higher in total extractives than the NIST RM. Uncertainties for each constituent are similar between the corn stover and the bagasse RM.

Carbohydrates and acetyl content were determined by high-performance liquid chromatography (HPLC). Lignin was measured by combining a UV measurement for acid-soluble lignin and a gravimetric method for measuring acid-insoluble residue.

HPLC Analytical Conditions. All HPLCs (model 1100, Agilent, Santa Clara, CA) were equipped with an inline degassing unit, chilled autosampler, and refractive index detector (RI). We used a lead cation (Pb^{2+}) exchange column (Shodex SP0810, Showa Denko K.K., Kawasaki,

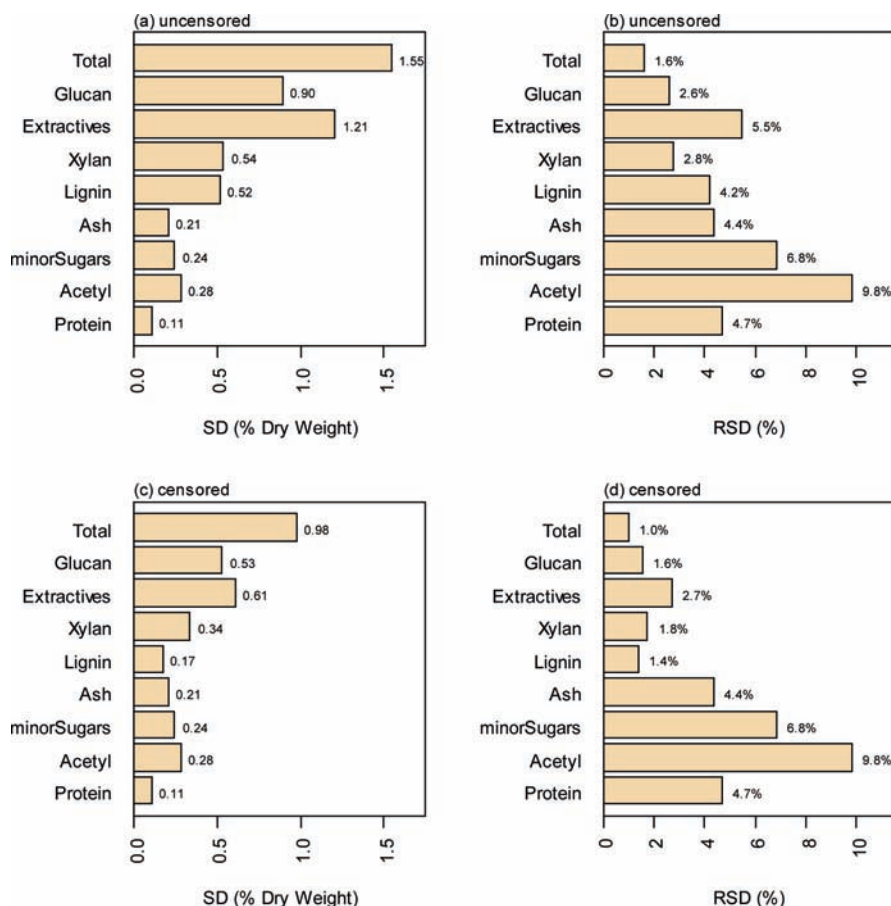


Figure 2. Summary compositional analysis uncertainty results from this study on dry weight basis for the corn stover (a) standard deviation (SD) and (b) relative standard deviation (RSD, SD/mean) for the uncensored data and (c) SD and (d) RSD (SD/mean) for the censored data. The constituents are shown in the same order as in Figure 1. The overall uncertainty is driven to a large extent by total extractives and glucan and drops substantially when outlier data are removed.

Japan), running at 85 °C, using 18 M Ω water as the mobile phase at 0.6 mL/min. We used deashing guard columns (125-0118, Bio-Rad, Sunnyvale, CA) located outside the heating compartment to keep the guard columns within their manufacturer's temperature tolerance. This reduces the appearance of ghost peaks from the monomer sugars that can interfere with quantitation. We purchased the calibration standards and independent calibration verification standards from Absolute Standards Inc. (Hamden, CT). Calibration standards for sugars were run at four levels ranging from 0.01 mg/mL to 6.00 mg/mL for D-(+)-cellobiose, D-(+)-glucose, D-(+)-xylose, D-(+)-galactose, and L-(+)-arabinose, with a 50 μ L injection volume into the HPLC. Acetyl separation was performed using a cation H column (HPX-87H Bio-Rad), running at 55 °C, using 0.01 N sulfuric acid in 18 M Ω water as the mobile phase at a flow rate of 0.6 mL/min. Calibration standards for acetyl ranged from 0.02 to 1.08 mg/mL and were also run at four levels with a 50 μ L injection size.

Description of Sugar Recovery Standards. The measured concentrations of monomer sugars released from the biomass are corrected using SRS autoclaved with the samples. Mixed solutions of biomass sugars, prepared in-house and similar in concentration to the samples, are analyzed before and after the autoclave hydrolysis to account for the fraction of the liberated monosaccharides that are degraded during the hydrolysis. In practice, the glucose concentration in the SRS samples is reduced to approximately 95% of its original value after autoclaving, so the corresponding glucose concentrations in the biomass samples are multiplied by the ratio 1.0/0.95 to correct for this loss. Typical SRS recoveries for xylose are approximately 85%, because xylose is more labile than glucose.

Laboratory Equipment Comparison. We performed these analyses in two similarly equipped laboratories on the NREL campus. The two laboratories (identified below as A and B) were equipped with similar equipment and reagents; with the following exceptions, the materials and

methods should be considered the same for both laboratories. Laboratory A was equipped with a Thermolyne (Barnstead Thermolyne, Hampton, NH) 30400 furnace set to ramp to 575 °C, whereas laboratory B's furnace was set to a constant 575 °C and samples were ignited using a Bunsen burner prior to placement in the furnace. Laboratory A had two large autoclaves with external steam generators (Consolidated SR-24C with Mark II controllers, Consolidated Stills and Sterilizers, Boston, MA) available for the secondary acid hydrolysis. Laboratory B was equipped with a benchtop autoclave that generated steam pressure internally (Sterilmatic model Stm_E, Market Forge, Everett, MA). The only common materials between the laboratories were common batches of SRS solutions and HPLC standards for both carbohydrate and acetyl measurements.

Measured Biomass Components. A total of 11 primary components were measured: glucan, xylan, galactan, arabinan (the structural carbohydrates), acetyl, lignin (the combined acid-insoluble residue and acid-soluble lignin), ash (apportioned between structural and nonstructural), protein (apportioned between structural and nonstructural), water and ethanol extractives, and the sucrose content within the water extractives fraction. On the basis of these primary data, we calculated the values of the following components: water extractives (others), total extractives, total structurals, and total. Water extractives (others) is calculated as the amount of water-soluble material minus the soluble ash, soluble protein, and sucrose. The total extractives value is the sum of the water and the ethanol-soluble material. The structurals value is calculated as the sum of all material not solubilized during extraction; it consists of the glucan, xylan, galactan, arabinan, lignin, acetyl, structural protein, and structural ash. The total value is the sum of all components in the biomass.

To measure the overall method uncertainty, we had a homogenized corn stover sample analyzed in different laboratories, by different analysts, and in multiple batches run by the same analyst. Each analyst performed complete analysis on 12 replicate corn stover samples and 1 sugar cane

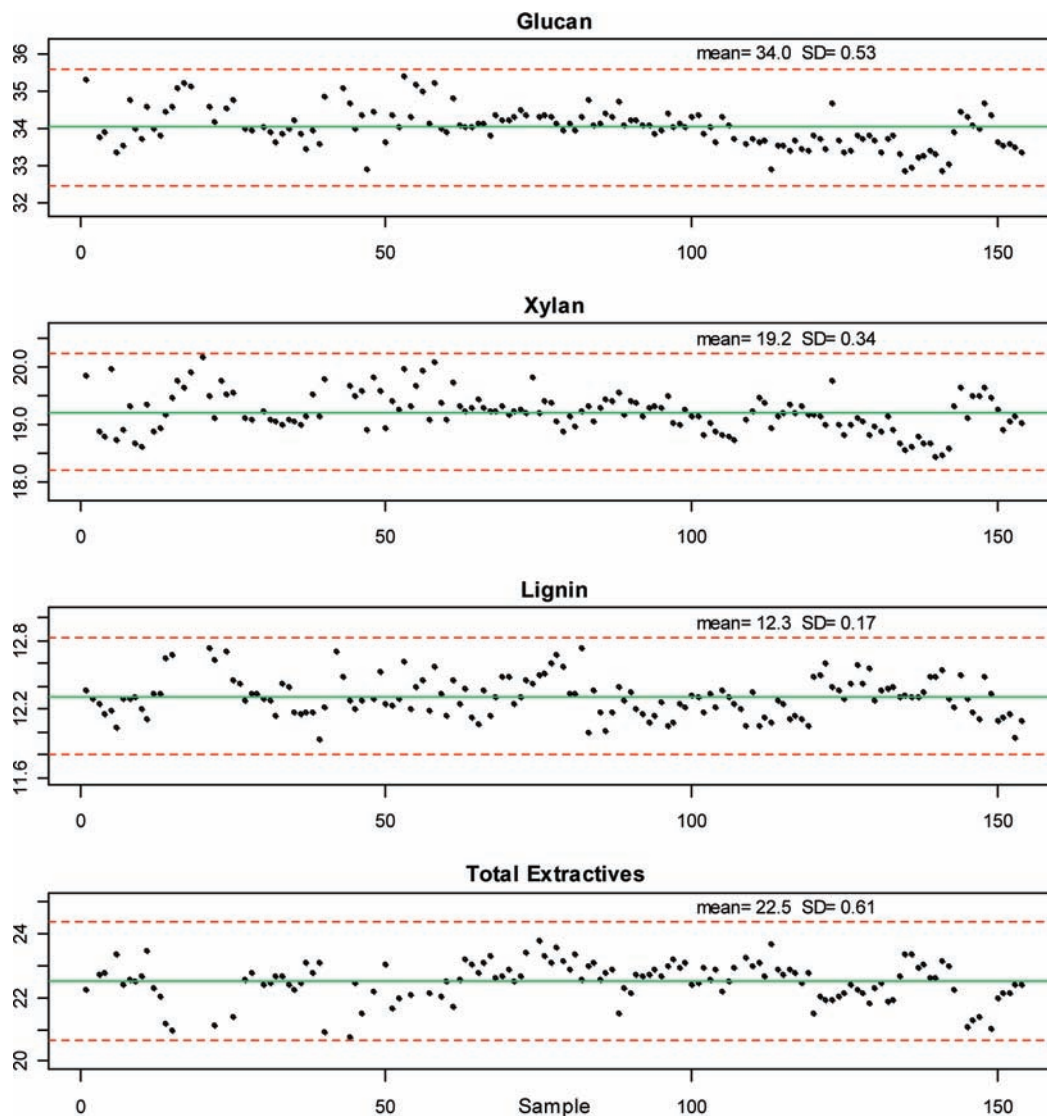


Figure 3. Control chart plots of individual corn stover sample compositional analysis results from this study for glucan, xylan, lignin, and total extractives. These four constituents represent almost 90% of the total mass of the corn stover. Solid lines indicate the mean of each constituent. Dotted lines represent ± 3 standard deviations from the mean.

bagasse sample as a batch. The data presented below are the results of 13 separate batches performed by 7 analysts. Each analyst was assigned a laboratory such that the number of batches was divided approximately evenly between the two laboratories. To evaluate the effect of multiple batches run by a single analyst, two researchers each analyzed four separate batches, two batches in each laboratory. All other analysts performed only one batch in an assigned laboratory.

We used the statistical program “R” to analyze and plot the data (9). All tests of statistical significance were performed at the 95% significance level ($p < 0.05$).

RESULTS AND DISCUSSION

We chose corn stover for this experiment, because it is a widely and currently available herbaceous feedstock. An analyst homogenized and distributed a 900 g sample of corn stover into 24 bottles. We randomly selected four bottles, scanned the contents by near-infrared (NIR) spectroscopy, and predicted the stover compositions. In a one-way ANOVA of these data, no significant compositional differences between the bottles were found. We therefore assumed that the material was homogeneous among the different bottles. A total of 13 batches of 12 corn stover samples were analyzed by 7 analysts, resulting in a total of 156 analyses.

One sample each from two batches was lost during the analysis, leaving 154 complete analyses. As a control, one sample of NIST bagasse reference material (RM 8491) was analyzed with each batch for a total of 13 replicates. The compositional analysis data are included in the Supporting Information.

The analysts performed the compositional analyses in batches (12 corn stover samples and 1 bagasse) over the course of several weeks (Dec 2008–Jan 2009). Here we report both the censored (without Tukey outliers) and uncensored (all data) sets. We removed individual measurements after applying Tukey outlier tests in conjunction with observed experimental anomalies. [We applied the Tukey outlier test to the major components (water extractives, lignin, glucan, and xylan). This test identifies any sample having a value > 1.5 times the interquartile range less than the lower hinge or greater than the upper hinge. A total of 17 samples out of 154 were identified as low outliers for water extractives. All of these low water extractives were run in laboratory B. During this experiment, analysts noted occasional variability in the volume of the solution in the water collection vial after extraction, along with minor instrument warnings from the automated solvent extractor in laboratory B. After all experiments were completed, we discovered and repaired a mechanical

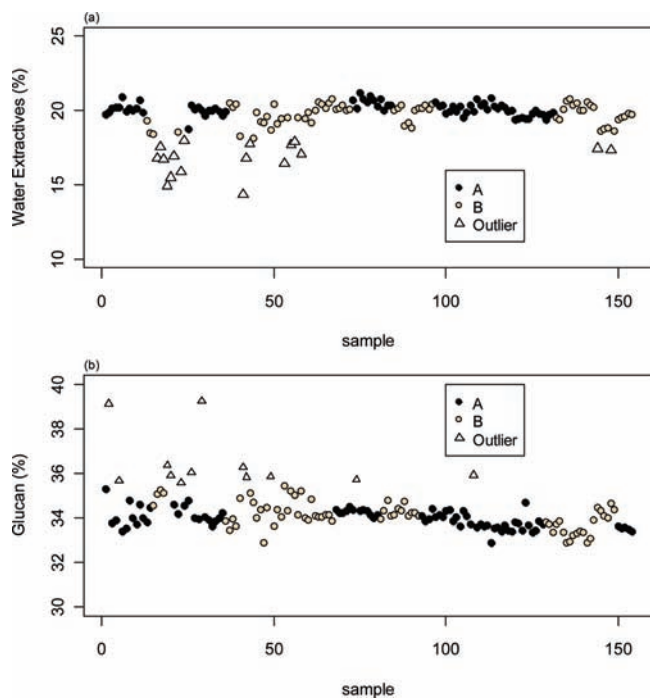


Figure 4. (a) Water extractives and (b) glucan content of 154 samples in the round robin study, generated by seven analysts working in two different laboratories. In both plots the data are separated by laboratory, with Tukey outliers highlighted. All sample results with water extractives data flagged as Tukey outliers ($n = 17$) were from the same solvent extraction instrument in laboratory B and were excluded from the data analysis. A total of 12 carbohydrate outliers were also identified, but 5 of these were flagged as extractives outliers as well. All but 1 of the remaining 7 outliers were analyzed in laboratory A.

problem with the automated solvent extractor in laboratory B that we believe was responsible for the low outlier water extractive values. The exclusion of these 17 samples reduced the water extractives uncertainty considerably, but even after exclusion of these samples, the relative standard deviation of the water extractives value for samples run in laboratory B was still larger than that in laboratory A (3.5 vs 2.5%), a statistically significant difference. We identified a total of 13 Tukey outliers for glucan and/or xylan (7 for both glucan and xylan, 5 for glucan only, and 1 for xylan only). Six of these 13 samples were also outliers for water extractives but were not flagged as carbohydrate outliers on an extractives-free basis. Thus, the underlying carbohydrate data were likely correct, and the low extractives value computationally biased the final carbohydrate result high. Six of the other 7 were from laboratory A. These samples showed high carbohydrate values and normal water extractives values, which suggests a problem in the carbohydrate analysis separate from the water extractives test. We had identified a problem with the gasket seals used for filtration after analytical hydrolysis in laboratory A. We hypothesize that excessive air flow through the flask caused evaporation of the filtrate, leading to artificially high carbohydrate values. The Tukey outlier test also identified 8 lignin outliers, all but 1 of which were associated with an outlier extractives value, which again suggests the low extractives result computationally biased the final lignin result high.] In summary, of 154 analyses, we identified as outliers 17 individual water extractives measurements, 12 glucan measurements, 8 xylan values, and 8 lignin values. We believe the censored data represent the “best case” results of our analyses, whereas the uncensored data show how small method changes can strongly affect the uncertainties of these empirical methods.

It could be argued that we are simply excluding the ~5% of the data that always fall outside the range covered by ~2 standard deviations from the mean. However, we believe these outliers are not representative of the typical results of the compositional analysis procedure for two reasons. First, all of the water extractives outliers were below the mean, and all of the lignin and glucan/xylan outliers were above the mean, suggesting a systematic error in both cases. Second, we have working hypotheses to explain the sources of both errors: a malfunctioning solvent extractor in laboratory B for the water extractives and associated lignin and glucan/xylan outliers, and a leaking gasket in laboratory A during filtration for the independent carbohydrate outliers.

Biomass Compositional Analysis and Uncertainty Data. We show in **Table 1** the censored and uncensored compositional data for the corn stover sample. We report both the overall and the pooled (weighted average within the batches) standard deviations (SD) for this set. The pooled SDs are less than the overall SDs for all components, suggesting that the uncertainty within the batches is lower than the uncertainty between batches. We show in **Table 2** the compositional data for the NIST bagasse (RM 8491) control sample, which we ran once with each batch. No bagasse data were censored. We report only the overall SD, as the pooled SD is not meaningful for single replicates. The average total mass closures were 96.9% (SD 1.0%) and 100.8% (SD 1.0%) for the corn stover and the bagasse sample, respectively. A component closure near 100% suggests that most components are accounted for and little double counting of components is occurring.

Figure 1 shows the average censored compositional data for the corn stover material (**Figure 1a**) and the corresponding data for the NIST sugar cane bagasse (**Figure 1b**). The constituents are shown in the same order in both figures and are sorted by percent dry weight in the corn stover samples. The error bars in this figure represent ± 1 SD about the mean of each constituent. Almost 90% of the total mass is accounted for by glucan, xylan, lignin, and total extractives. The corn stover feedstock contains much more total (water + ethanol) extractives than the sugar cane bagasse (> 22 vs < 6%). The amount of extractives in corn stover depends on the variety, harvest time, and postharvest handling. Sugar cane bagasse is a byproduct of an industrial process to extract the cane juice, and this process leaves behind little extractable material. Thus, these sample materials span a wide range of total extractives content.

The effect of the censoring process on the uncertainty of the measurements is illustrated in **Figure 2**. Panels **a** and **b** of **Figure 2** show the SD and the relative standard deviation (RSD = SD/mean) of the uncensored compositional data for the main constituents of corn stover, whereas panels **c** and **d** show the corresponding censored data. The SDs for the censored components are 30–50% lower than for the uncensored data, which shows the outsized effect of the excluded samples. The extractives term has the highest RSD of the major components. Because the total extractives value is used to correct the measurement of structural components to a whole, dry weight basis, uncertainties in the extractives value affect all of the results. The RSD values are larger for the minor components, due largely to smaller mean values rather than the larger SD values.

Examination of Data Using Control Charts. **Figure 3** shows control charts of the censored corn stover compositional data for each of the individual samples analyzed for glucan, xylan, lignin, and total extractives. In each plot, the solid line indicates the mean and the dashed lines indicate ± 3 SDs above and below the mean. Gaps in the data appear where samples were excluded in the censoring process. Each plot is labeled with the mean and SD for

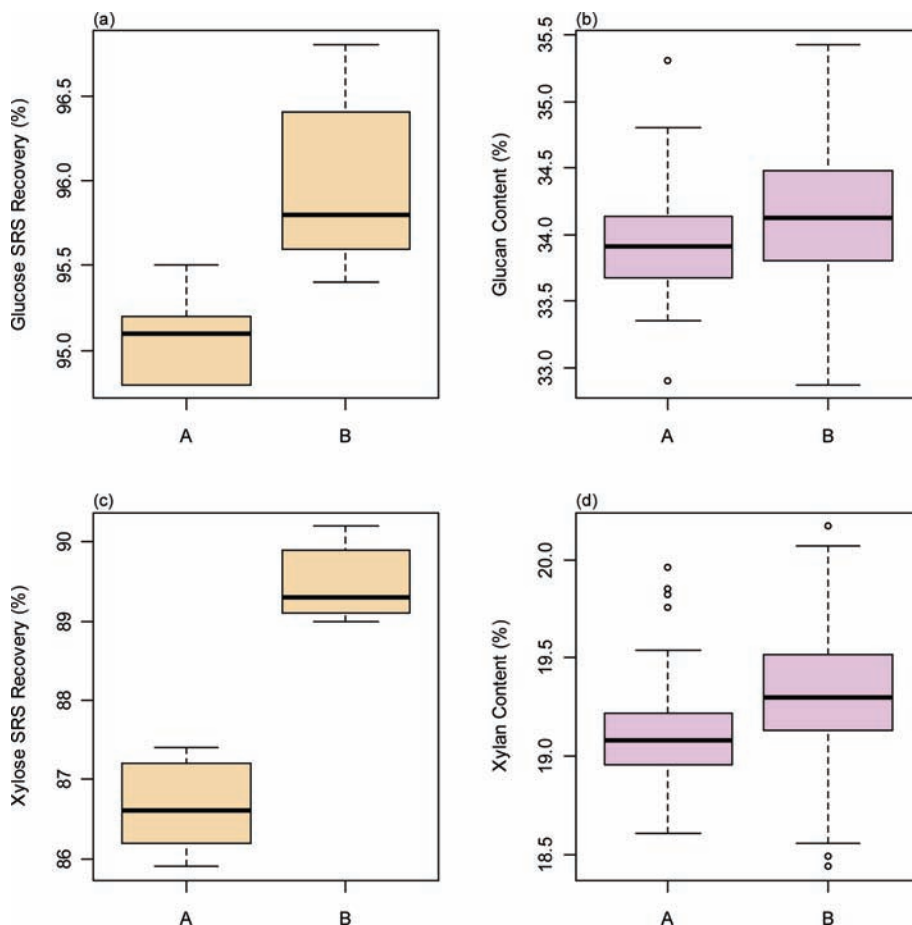


Figure 5. Box plots of the sugar recovery standard (SRS) recoveries for (a) glucose and (c) xylose in the two laboratories (A and B) where the compositional analysis results were generated and the corresponding calculated (b) glucan and (d) xylan composition of the corn stover material. The thick lines show the average values, the boxes cover the interquartile range (IQR, “middle 50”), the whiskers denote the 1st and 4th quartile ranges, and any outliers are noted with circles. Statistically significant differences between the two laboratories were seen in SRS recoveries for both glucose and xylose. When these SRS recovery data were then used to correct the corresponding values for glucose and xylose in the corn stover samples, no statistically significant differences were seen between the two laboratories for either constituent.

each constituent. No trends in constituent values are evident, suggesting that the analyses were performed consistently over time.

Figure 4 shows the uncensored corn stover compositional data for water extractives and glucan, noted by laboratory, for each of the samples. The triangles depict samples that failed a Tukey outlier test. Many of the glucan outliers (**Figure 4b**) correspond with water extractives outliers (**Figure 4a**). As discussed above, this is due to the computational effect of using the extractives value to convert the primary measurement of extractives-free glucan to a whole, dry weight basis. A few of the glucan high outliers do not correspond with water extractives outliers, and as we mentioned previously, we believe these hydrolysate solutions became concentrated during the vacuum filtration step.

Utility of Sugar Recovery Standards. As discussed above, SRS are included in the secondary (autoclave) hydrolysis step. The use of SRS as a proxy to estimate the extent of degradation of the structural (polymer) carbohydrates may overestimate the amount of sugars lost during hydrolysis, because the structural carbohydrates are present as oligomers (not monomers as in the SRS) at the beginning of the secondary hydrolysis. Although imperfect, until a good model polysaccharide is identified, hydrolyzing monosaccharides simultaneously with biomass samples is the best available method to correct for lost sugars.

Because the SRS procedure accounts for the degradation of structural carbohydrates during a specific batch, it can correct for

slight differences between autoclaves. This is shown in **Figure 5**, which shows box plots of the average SRS recoveries for glucose (**Figure 5a**) and xylose (**Figure 5b**) in the two laboratories as well as the corresponding glucan and xylan composition values for the corn stover material. Analysts working in laboratory A used one of two large autoclaves routinely used by life science researchers to sterilize glassware and biological media, and analysts working in laboratory B used a benchtop autoclave used only for analytical hydrolysis. All autoclaves used ostensibly the same protocol: 121 °C for 60 min. However, the SRS data suggest that the actual conditions in the autoclaves were different. There are statistically significant differences between the two laboratories for SRS recoveries for both glucose and xylose. However, when these SRS recovery data were then used to correct the corresponding values for glucose and xylose in the corn stover samples, no statistically significant differences were seen between the two laboratories for either glucan or xylan composition of the corn stover samples. Thus, a small but consistent systematic bias in the autoclave steps between the two laboratories was eliminated by using the SRS procedure.

Sugar Cane Bagasse Reference Material Compositional Data. **Figure 6** shows control charts of the measured glucan, xylan, lignin, and total extractives content of the sugar cane bagasse RM that was analyzed with each batch of corn stover. The lines in this figure are the mean (solid) and the ± 3 SD (dotted) values. The SDs seen with the corn stover and NIST bagasse materials were

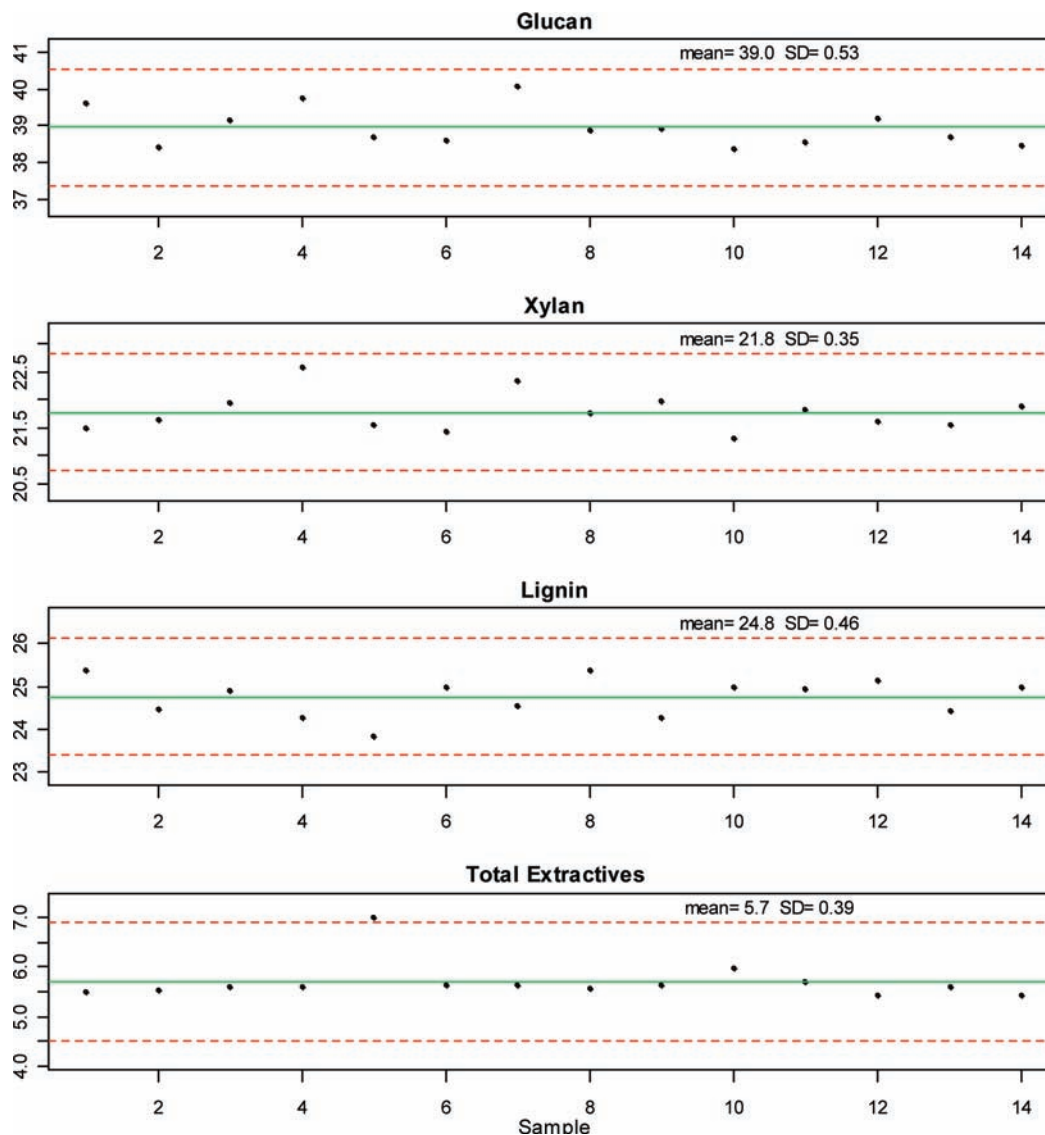


Figure 6. Individual NIST RM 8491 compositional analysis results from this study for glucan, xylan, lignin, and extractives. One NIST RM sample was analyzed with each batch of 12 corn stover samples. Solid lines indicate the mean of each constituent. Dotted lines represent ± 3 standard deviations from the mean.

similar, which suggests the uncertainty reported here is typical of the method and will be similar across different feedstocks. It is useful to track the measured composition of such a control method over time, because large differences between the overall mean and any individual measurement may indicate a problem with the analysis of the associated laboratory samples.

Effect of Multiple Analysts on Uncertainties. We show the aggregate censored corn stover statistics among analysts in **Figure 7**. This figure shows box plots of each analyst's data with the overall median, shown as a bold line, and the overall interquartile range (IQR, "middle 50") marked by the shaded region. Analysts 1 and 5 ran four batches each, and their IQR boxes are the widest among analysts. The rest of the analysts ran only one batch each and show narrower IQRs. The wider IQRs for analysts 1 and 5 are consistent with the data in **Table 1**, showing that the uncertainty between batches is higher than the uncertainty between analysts. We saw some small yet statistically significant differences among some of the analysts. For example, the median glucan values (**Figure 7a**) determined by analysts 2 and 6 were 0.5 and 0.8 percentage points (% dry weight) below the median value of all analysts. These biases are small and based on a single batch, so we cannot draw conclusions about possible

biases among analysts. Most of the component values outside the overall IQR were biased low.

Compositional Analysis Data Comparisons. The corn stover used for this study was not chosen to have a representative composition, but rather it was prepared to be compositionally homogeneous. The composition seen here is within the range of stover values reported elsewhere. For a detailed discussion of corn stover compositional variability, see Templeton et al. (10).

We included the NIST RM 8491 as a control sample with each analytical batch. The data we report here agree reasonably with previous data. A round robin analysis sponsored by the IEA, NREL, and NIST reported values (all % dry weight) of 38.6, 23.1, 20.4, and 94.3 for glucan, lignin, xylan, and total component closure, respectively (3). The authors did not report uncertainties with their whole biomass compositional data as they did with their extractives-free compositional data. We report bagasse compositions (all % dry weight) of 39.0 (SD 0.5), 24.8 (0.5), 21.8 (0.4), and 100.8 (1.0) for glucan, lignin, xylan, and total component closure, respectively. Our component closure is higher than the IEA data because we report acetyl [3.3% (0.2%)] and protein [2.2% (0.1%)], whereas the IEA round robin did not, although the IEA reported glucuronic acid (1.3%), whereas we

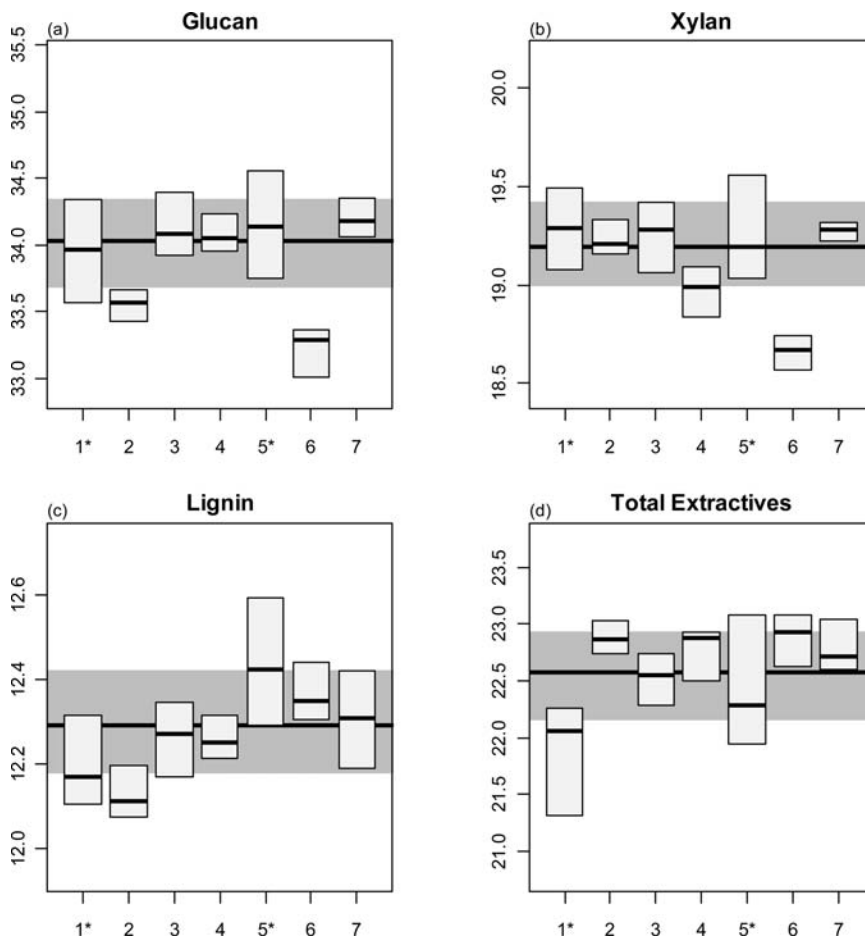


Figure 7. Box plots of the glucan, xylan, lignin, and extractives content of corn stover analyzed during this study, separated by analyst. The median value of the entire data set is shown as a solid horizontal line, and the interquartile range (IQR, “middle 50”) of the entire data set is shown in gray. Analysts 1 and 5 ran four batches each, whereas the other analysts ran a single batch.

did not. A new interlaboratory study is currently being performed to recertify the compositions of the four NIST biomass RMs.

Method Quality Control Considerations. These empirical methods require attention to detail and can be problematic, with many sample manipulations and therefore opportunities for mistakes. We have shown that it is possible to get reproducible feedstock compositional data among different analysts and laboratories using these compositional analysis methods. We have also shown, in the censored and uncensored data (**Figure 2**), that small differences in the analytical technique can have a great effect on the sample uncertainty.

To generate high-quality data, it is important to follow the methods closely, although simply following the written protocol does not guarantee good data. Additional quality control measures are needed. In addition to good laboratory practices (pipet and balance calibration, proper HPLC standards, etc.), we ensure quality compositional data by analyzing SRS with each sample batch, running a reference material with each batch (NIST bagasse, in this case), and analyzing samples in duplicate (or more). Evidence of good compositional analyses includes replicate data within these reported method uncertainties, good analytical component closure (95–105%), and good component closure around a given unit operation (pretreatment, conditioning, saccharification, fermentation, etc.).

On the basis of our experience, there are several common mistakes in technique that can affect the results. These include improperly dried samples and incomplete wetting or sample stirring during the primary hydrolysis, leading to incomplete

hydrolysis and biasing the carbohydrate measurement low. Other technique mistakes included raising the pH of the analytical hydrolysate to >6 during neutralization with CaCO_3 and difficulties in interpreting and integrating the HPLC chromatograms, because baseline resolution of all sugars is not possible with current columns. These mistakes could bias the carbohydrate results higher or lower. We have also seen concentration of the hydrolysate during filtration of the acid-insoluble residue and incomplete extraction of samples prior to analytical hydrolysis, which tend to bias the carbohydrate results high. In general, there are many possible causes to bias the carbohydrate measurements low and fewer causes for high biases.

The uncertainties associated with these methods can be reduced by having well-trained analysts running the methods; because the methods are so manually intensive, it is easy for an analyst to fall out of practice. We recommend having one analyst work up each sample from beginning to end, which allows the analyst to spot problems in the data and correct his or her technique.

Practical Consideration 1: Extractions. Even with all of these safeguards in place, we noted larger compositional uncertainties in the uncensored data due to problems in the extraction step. Poor extractions led to poor corrections from extractives-free data to a whole, dry weight basis, and we excluded the poorly extracted samples from the censored data set. We noted but did not appreciate differences in the volume of water extract solution and minor instrumental errors. In hindsight, this should have alerted us to mechanical problems with the instrument. We now

measure the amount of water extract solution to ensure more consistent extractions.

The outlier water extractives were easy to spot with so many replicates run in this data set (Figure 4a). Even if few replicates were run, an alert analyst should note unusually low total mass closure, unusually low extraction volumes, or unusually high carbohydrate values, which indicate problems with the analytical data.

Practical Consideration 2: Concentrated Analytical Hydrolysates. We found a few samples with high carbohydrate values but normal extractives values. We believe these are caused by concentrating the analytical hydrolysate during the solid/liquid vacuum filtration step, because these samples tended to have high concentrations for all of the sugars and acetyl simultaneously. We have replaced the filtration gaskets used in that laboratory to limit the amount of air drawn past the analytical hydrolysate.

Practical Consideration 3: Sugar Recovery Standards. As discussed previously, we recommend that a SRS be run with each batch. This allows for a direct correction of the samples with the specific hydrolysis conditions and can be used to track changes in autoclave performance. Because the SRS recovery is specific to each autoclave, we do not recommend using SRS values from one autoclave to correct for data run in a different autoclave.

Overall Method Uncertainties. The uncertainty values presented here combine the uncertainty contributions from laboratory operations (e.g., weighing, dilutions), instrumental operations (ASE, HPLC), different analysts, and different laboratories. Using a homogenized corn stover sample, we report censored RSDs in Figure 2 and Table 1. Estimates of errors, in the form of 98% confidence intervals, for these methods have been published elsewhere, and are reported to range from 0.5 to 1.5% (absolute basis) for all components (11, 12). In contrast, we report the method uncertainties of replicate analyses on a homogenized corn stover sample. We report method uncertainties for the major biomass components (all % RSD) as glucan (1.6%), lignin (1.4%), xylan (1.8%), and extractives (2.7%). Not surprisingly, we found higher uncertainties for the lesser components including protein (4.7%), whole ash (4.4%), minor sugars (6.8%), and acetyl (9.8%). These uncertainties are based on results from 7 analysts, run in 13 batches, analyzed in 2 laboratories. We believe these data demonstrate typical uncertainties associated with these methods. We recognize that these feedstock compositions are used for conversion and yield calculations, and the uncertainties in these primary measurements propagate into these calculations. We are currently examining the effect of this uncertainty propagation (manuscript in preparation).

ABBREVIATIONS USED

IEA, International Energy Agency; HPLC, high-performance liquid chromatography; LAP, laboratory analytical procedure; NIST, National Institute of Standards and Technology; NREL, National Renewable Energy Laboratory; RM, reference material; RSD, relative standard deviation; SRS, sugar recovery standard.

ACKNOWLEDGMENT

We thank Crissa Doepcke, Erik Fisk, Deb Hyman, Ryan Ness, Courtney Payne, Kristen Reichel, and Jeff Wolfe for running these analyses. We thank Al Darzins, Stewart Black, Sara Havig, Ling Tao, Dan Schell, and Amie Sluiter for reviewing the manuscript.

Supporting Information Available: CSV file containing all compositional results and “R” codes used to analyze the data and generate the figures and tables. This material is available free of charge via the Internet at <http://pubs.acs.org>.

LITERATURE CITED

- (1) Sluiter, J. B.; Ruiz, R. O.; Scarlata, C. J.; Sluiter, A. D.; Templeton, D. W. Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *J. Agric. Food Chem.* **2010**, doi: 10.1021/jf1008023.
- (2) Aden, A.; Ruth, M.; Ibsen, K.; Jechura, J.; Neeves, K.; Sheehan, J.; Wallace, B.; Montague, L.; Slayton, A.; Lukas, J. *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover*; National Renewable Energy Laboratory; Harris Group: Golden, CO, and Seattle, WA, June 2002.
- (3) Milne, T. A.; Chum, H. L.; Agblevor, F.; Johnson, D. K. Standardized analytical methods. *Biomass Bioenergy* **1992**, 2 (1–6), 341–366.
- (4) NIST Standard Reference Materials; <http://ts.nist.gov/measurementservices/referencematerials/index.cfm> (Dec 2009).
- (5) Taylor, B. N.; Kuyatt, C. E. *NIST Technical Note 1297: Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*; U.S. Government Printing Office: Washington, DC, 1994.
- (6) Wolfrum, E. J.; Sluiter, A. D. Improved multivariate calibration models for corn stover feedstock and dilute-acid pretreated corn stover. *Cellulose* **2009**, 16 (4), 567–576.
- (7) ASTM Standards; <http://www.astm.org/Standard/index.shtml> (Dec 2009).
- (8) NREL Laboratory Analytical Procedures; http://www.nrel.gov/biomass/analytical_procedures.html (Dec 2009).
- (9) *R: A Language and Environment for Statistical Computing*; R Development Core Team: Vienna, Austria, 2006; R Foundation for Statistical Computing, ISBN 3-900051-07-0, <http://www.R-project.org>.
- (10) Templeton, D. W.; Sluiter, A. D.; Hayward, T. K.; Hames, B. R.; Thomas, S. R. Assessing corn stover composition and sources of variability via NIRS. *Cellulose* **2009**, 16 (4), 621–639.
- (11) Decker, S. R.; Sheehan, J.; Dayton, D. C.; Bozell, J. J.; Adney, W. S.; Hames, B.; Thomas, S. R.; Bain, R. L.; Czernik, S.; Zhang, M.; Himmel, M. E. Biomass conversion. In *Kent and Riegel's Handbook of Industrial Chemistry and Biotechnology*, 11th ed.; Kent, J. A., Ed.; Springer Verlag: New York, 2007; Vol. 2, pp 1449–1548.
- (12) Hames, B. R. Biomass compositional analysis for energy applications. In *Biofuels, Methods in Molecular Biology*; Mielenz, J. R., Ed.; Humana Press: Totowa, NJ, 2009; pp 145–167.

Received for review February 26, 2010. Revised manuscript received May 26, 2010. Accepted June 2, 2010. This work was supported by the U.S. Department of Energy Office of the Biomass Program.