Interlaboratory Comparison of Soybean Protein and Oil Determinations

Randy A. Hartwig and Charles R. Hurburgh, Jr. 1.*

Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, Iowa 50011

Iowa State University coordinated an interlaboratory comparison study of Kjeldahl protein and ether oil extraction methods. Blind duplicates of 10 clean, single-variety soybean samples were sent to 30 laboratories grouped in 3 categories of 10 each in public (government and university), commercial and processor facilities. Five of the commercial laboratories were AOCS-certified.

Standard deviations among laboratory means across all samples were 3.87 and 1.82 percentage points (dry basis} for protein and oil, respectively (0.48 and 0.27, respectively, for the AOCS-certified laboratoriesl. The average differences between blind duplicates of a sample were 0.71 percentage points for protein and 0.87 percentage points for oil I0.28 and 0.45, respectively, for the certified laboratories}. Average standard deviations across laboratories on an individual sample were 2.37 and 1.71 percentage points for protein and oil, respectively I1.87 and 0.99, respectively, for the certified laboratories.

KEY WORDS: Ether-extraction, Kjeldahl, oil, protein, soybean.

On September 4, 1989, the Federal Grain Inspection Service (FGIS) included protein and oil analysis of soybeans as Official Criteria on Official USDA inspection certificates (1). Because the FGIS program is based on nearinfrared (NIR) technology, base reference methods were used for calibration. The American Association of Cereal Chemists (AACC), the Association of Official Analytical Chemists (AOAC), and the American Oil Chemists' Society (AOCS) have differing methods for protein, oil, and moisture measurements. In theory, results from all methods should be statistically similar because the methods are all based on Kjeldahl (protein), solvent extraction (oil), and oven drying (moisture). Whether this is so in routine practice is as of yet unknown. The reference methods utilized by the FGIS are hybrids of AACC, AOCS, and AOAC methods.

Kjeldahl protein methods. Although proteins can be analyzed for amino acid content according to methods involving high-performance liquid chromatography (HPLC), ion-exchange chromatography or immunochemistry, it is still difficult to measure crude protein content (2). Kjeldahl protein analysis problems regarding the amount of acid, the digestion temperature, the possible use of oxidizing agents, the digestion time, and the use and type of catalyst occur mostly in the digestion phase.

The aforementioned problems are not mutually exclusive inasmuch as the amount of acid used depends upon the amount of catalyst. The acid quantity in turn affects the required digestion temperature and time. Carpiaux showed that 1 g of protein will consume about 9.0 g of sulfuric acid (3). Addition of catalyst creates an even greater demand for acid (4}. Maintaining the proper acidto-catalyst ratio is crucial to digestion. A low acid-tocatalyst ratio raises digestion temperatures in the early stages of digestion, thereby causing low nitrogen recovery. A high acid-to-catalyst ratio reduces digestion temperatures, thereby causing incomplete digestion in the allotted time period and, once again it causes low nitrogen recovery.

Oxidizing agents such as hydrogen peroxide and perchloric acid have been used successfully by researchers (5,6). Hach *et al.* (5) devised a method in which only sulfuric acid and hydrogen peroxide are used. This resulted in reduction of the digestion time to less than 10 min.

The length of digestion time is a matter of long-standing debate. Beet (7) demonstrated that the length of time required to reach the absolute nitrogen value can be anywhere from 0 to 235 hours. Many analysts use clarification of the sample as the endpoint of digestion, but Crossley's work (8) has shown that there is little relation between clarification and completion of digestion.

There are many types and quantities of catalysts. Osborn and Wilkie (9) found a decreasing N-recovery efficiency with the following catalysts: Hg, Se, Te, Ti, Mo, Fe, Cu, V, W, and Ag. Others have found Hg, Se, and Cu to be effective catalysts (10-13).

Both AOCS Method Aa 5-38 (14) and AACC Method 46-10 (15) use a mercuric oxide catalyst. The AOCS and AACC methods call for a digestion time of 30 min after the sample has cleared. Both AOAC Method 7.033-7.039 (16) and AACC Method 46-11A (15) use a copper sulfate catalyst and 90 min of heating after the dense white fumes clear the flask. All methods require a heating device that will bring 250 mL of water to boiling in 5 min. The FGIS uses a copper sulfate catalyst and 70 min of digestion with heating units comparable to those of the AOCS, AACC, and AOAC methods (14-16}.

Oil extraction methods. Oil determinations are as difficult as crude protein analysis. The factors affecting oil extraction are sample tempering, choice of solvent, extraction time, and extracted-meal particle size.

A 1-hr tempering period is required by AOCS Method Ac 3-44, but no tempering is required for either AACC Method 30-20 or AOAC Method 14.084-14.085 (14-16). Arnold and Choudhury tested ethanol, benzene, isohexane, pentane, isopentane, and hexane of varying purities (17- 20). Extraction rates increased in this order: isopentane. pentane, isohexane, hexane, and benzene. Extraction rates for hexane and benzene were about equal. The AOCS method calls for petroleum ether $(30-65\degree C)$ boiling point) (14); The AOAC Method 7 calls for diethyl ether (16). The FGIS modified AOCS Method Ac 3-44 incorporates two 5-hr extraction periods instead of one (14). Extraction times ranged from 4-16 hr. Extraction time depends

¹Laboratory Manager, E.I DuPont, Inc., Ames, IA and Associate Professor, Agricultural and Biosystems Engineering Department, Iowa State Unviersity, Ames, IA 50011.

^{*}To whom correspondence should be addressed at Dept. of Agricultural and Biosystems Engineering, Davidson Hall, Ames, IA 50011-3080.

upon the drip rate of condensed vapors, the type of solvent used, and the substrate extracted.

Snyder *et al.* (21) showed that particle size has a significant effect on oil determination. Dehulled soybeans were sieved into three particle sizes--larger than 40 mesh (425) micrometers), 40-100 mesh, and smaller than 100 mesh (150 micrometers). Particles larger than 40 mesh yielded 15.3% oil, 40-100 mesh particles yielded 21.95% oil, and particles smaller than 100 mesh yielde 24.8% oil.

Both AOCS method Ac 3-44 (14) and AOAC method 14.084-14.085 (16) require a particle size of 18 mesh (1 mm) in diameter; AACC method 30-26 requires reduction of the sample to a "suitable size" before extraction, but has no exact size specification (15). Particle size is not mentioned in AACC method 30-20 (15). All these methods apply to ground soybeans.

The AOCS conducts a sample check program, the Smalley Program, in which participants are asked to chemically analyze 10 samples per year for moisture, protein, and oil. Certificates are awarded to the top 10% of participants, based on the square root of the sum of squared observations divided by the number of observations minus one. A maximum of six and a minimum of two certificates are awarded each year. The 1988 season had five certified laboratories with an average combined standard deviation of 0.397%. The average combined standard deviation is a weighted average of the standard deviation of moisture (m), protein (p), and oil (o) determinations $(0.3s_m + 0.3s_o + 0.4s_o)$ (personal communication, AOCS, March, 1990}. The 1988 season had combined standard deviations of 2.219% for the high to 0.373% for the low. The average combined standard deviation for the 51 participating laboratories was 1.017%.

If the grain industry is to use near-infrared instruments for rapid constituent analysis, then calibration methodologies must be consistent (22-26). Protein and oil contents are major factors in processor economics (27). The purpose of this study was to determine the variability within and across laboratories for protein and oil methods as routinely practiced. This study therefore differs from studies of the variability reported on special samples that are known to be check samples.

MATERIALS AND METHODS

Thirty laboratories were chosen to participate in the study. They were grouped in 3 categories of 10 laboratories each: i) public/government/university; ii) commercial; and iii) soybean processing facilities. Five of the commercial laboratories were AOCS Smalley certified labs (1988) for soybean meal analysis (14). The processing laboratories were chosen from the 1988 Soya Bluebook listing of oil extraction plants/refineries (28). Ten samples of the 1987 crop soybeans were used (Table 1). These 10 varieties are commonly grown in Iowa.

Approximately 22-28 kg of each sample was cleaned through a Carter-Day dockage tester with a 3.97 \times 19.05 mm (10/64 \times 3/4 inch) slotted screen (Carter-Day Co., Minneapolis, MN}. This is the screen used as an aid in removing splits in the Official Soybean Grades. Any foreign material passing over the screen was removed by hand. Splits and foreign material were thrown out; therefore, the analysis samples contained only whole soybeans. These

TABLE 1

Soybean Sample Variety and Origin

Origin	Variety	Sample number
Iowa State University	Elgin	
Foundation seeds	Century 84	2
(Ames, IA)	Vinton	3
Asgrow, Inc.	Asgrow 2943	4
(Ames, IA)	Asgrow 3127	6
	Asgrow 1937	9
Dekalb, Inc.	Dekalb CX345	7
(Fort Dodge, IA)	Dekalb CX225	8
	Dekalb CX174	10
Farmers Coop Elevator (Farnhamville, IA)	Mixed, bin run	5

TABLE 2

ANOVA Table for Kjeldahl Protein and Ether Oil Extraction Analysis

	Degrees of freedom	Significance test	
Source	Protein Oil	Oil Protein	
Sample	9a 9a	$F_{9,252}$	$F_{9,243}$
Lab Type Lab (type)	28a, b 27a 2a 2 26 ^a 25 ^a	$\mathrm{F}_{2,26}$ $\mathrm{F}_{26,252}$	$\mathrm{F}_{2,25}$ $\mathrm{F}_{25,243}$
Lab \times sample	252 243		
Residual	280 290 579 559		

^{*a*} Variable was significant (P = 0.05).

 b One laboratory was unable to do analysis.

CTwo laboratories were unable to do analysis.

samples were more uniform than the usual inspection sampie. which would contain both foreign material and splits.

Each sample was divided twice with a Boerner divider (Seedburo Co., Chicago, IL}. One of the four subsamples was bagged and stored at 4° C as a file sample. The other three subsamples were split five more times each, which produced a subsample of approximately 150 g for chemical analysis. The 150-g subsamples were coded and stored in plastic bags at 4° C until they were shipped to the laboratories.

Each laboratory received blind duplicates in 2 shipments of 10 subsamples. The second shipment was not sent until the first set of data was returned. A data sheet for analysis accompanied the first set of 10 samples. We requested a method citation for moisture, protein, and oil, as well as any individual variations the laboratory may have made.

A result was considered an outlier for a particular constituent if its average value (over all samples) was more than two standard deviations from the overall mean. Statistical analysis was performed first with the outlier data included, and then with them deleted. The analysis was performed on individual observations, not on the average of duplicate observations.

Table 2 shows the analysis of variance for Tables. The important factors are type of laboratories, laboratory within a type, and residual. Ideally, neither laboratory nor type is significant. If either is, there are serious implications as to standardization of market testing. (If there is no significant difference between laboratories within a type, then a sample sent to any lab within that type would obtain similar results.) The residual estimates intralaboratory error on any given sample.

The average spread between blind duplicates on individual samples was calculated for each laboratory. This gave an indication of how well laboratories could repeat themselves on unknown samples. The average spreads between blind duplicate values were then averaged across laboratory types.

RESULTS AND DISCUSSION

 $Kjeldahl$ protein analysis. Significant differences ($P =$ 0.05) occurred between samples, laboratories, laboratory types, and laboratories within a type. Exclusion of outlier laboratories eliminated the significant effect of laboratory types.

Table 3 shows that public and commercial laboratories were not significantly different from each other, but were significantly different from processor laboratories if outlier data were retained. Exclusion of outlier data brought the overall processor mean into close agreement with the public and the commercial means. Public and commercial laboratories also had similar standard deviations across laboratory means. These standard deviations were much smaller than the processor standard deviation (6.90 points) although, with the outlier data excluded, the standard deviation of processor laboratories (0.72 points) was again close to the standard deviations of the other two laboratory types. The overall standard deviation of all laboratory means was 3.87 points (0.74 points, with outliers excluded). Table 3 shows that the least significant difference (LSD) $(P = 0.05)$ was 0.59 percentage points. With the outliers excluded, LSD was 0.55 percentage points.

The AOCS-certified laboratories (Table 4) were less variable than any of the other laboratory types {0.48 percentage points for AOCS laboratories as compared with 0.73 for public, 0.78 for commercial, and $6.90-0.72$ points with the outliers excluded--for the processor laboratories). The AOCS laboratories contributed half of the commercial laboratory data.

Means, by sample (across all laboratory types}, are given in Table 5. The averages shown in Table 5 do not

TABLE 3

Laboratory Means for Protein and Oil (n = 20 per Laboratory)

aLab number used for indexing purposes only. Different labs across rows in Table.

 b All values are on a dry matter basis.</sup>

c Laboratory unable to do analysis.

dConsidered an outlier lab.

e Number in parentheses is value calculated with outlier labs deleted.

TABLE 4

AOCS-Certified Laboratory Means for Kjeldahl **Protein and** for Ether Oil Extraction $(n = 20$ per Laboratory)

aValues are on dry basis.

agree exactly with the averages shown in Table 3, because several processors had missing data points. Public and commercial laboratories had similar standard deviations. Processor laboratories had the greatest average and the largest standard deviation, but when the two outliers were excluded both statistics agreed with those of the other two laboratory types.

The overall average of 41.59% was nearly identical to the AOCS laboratories' value of 41.61%. The overall standard deviation, 2.37 points, was almost five times that of the AOCS laboratories' value of 0.48 points. With the two processor outliers deleted, the average standard deviation of all the laboratories dropped to 0.94 percentage points, or about twice that of the AOCS laboratories. To put this in perspective, a sample with 41.6% protein (dry basis) sent to a randomly chosen laboratory would be analyzed at 39.7% to 43.5% protein in 95% of the cases.

Table 6 shows the average spread (0.71 points) between blind duplicates for all laboratory types. The public laboratories had the greatest spread between blind duplicates, (0.84 points). This was closely followed by the processor laboratories {0.74 points). The AOCS laboratories had the lowest average spread (0.28 points), which was one-third that of the public laboratories. Some laboratories clearly had precision problems.

Ether oil extraction. Significant differences ($P = 0.05$) occurred between samples, laboratories, and laboratories within a type, but not between laboratory types. There were still significant differences between laboratories within all types after the outliers were deleted.

As can be seen from Table 3, the standard deviation across processor laboratories was nearly twice that across either public or commercial laboratories. When the four outliers were deleted (one each from the public and the commercial groups and two from the processor group), standard deviations were 0.77, 0.93, and 0.31 points for public, commercial, and processor laboratories, respectively. Deletion of the outliers raised the overall mean oil content because the outlier laboratories all underestimated oil content. This result suggests problems in either sample preparation or solvent circulation. The LSD ($P =$ 0.05) between laboratories was 0.59 points for all data, and decreased to 0.50 points when the outliers were excluded.

Table 4 shows that, in the oil analysis, the AOCScertified laboratories had less than one-sixth the variability of all laboratories. Even with the outliers removed, the AOCS laboratories were much more consistent than the rest of the laboratories.

Means by sample (across laboratory types) are given in Table 7. The average of all laboratories in Table 7 does not agree with the average value in Table 3 because of missing processor and public data points. Public laboratories and commercial laboratories had similar standard deviations,

TABLE 5

aValues are on a dry matter basis.

bStandard deviation, by type, on each sample.

cWith outliers excluded.

TABLE 6

Individual Sample Means for Oil Extraction Analysis (n = 20 per Sample), all Laboratories Included

aValues are on a dry matter basis.

 b Standard deviation, by type, on each sample.

cWith outlier labs excluded.

TABLE 7

Average Spread Between Blind Duplicates for Kjeldahl Protein and Ether Oil Extraction

 a Values are on a dry basis. b No outliers.

but processor laboratories had markedly greater variabil- 6 ity than either. When outliers were excluded, processor laboratories had the lowest standard deviation at 0.76 points, followed by public laboratories at 1.01 points, and commercial laboratories at 1.10 points.

The overall laboratory average, 19.57% , was 0.67 points smaller than the AOCS laboratories' value of 20.24%. The overall standard deviation, 1.71 points, was almost 3.5 times that of the AOCS laboratories' standard deviation of 0.49 points. With the outliers deleted, the average standard deviation of all laboratories dropped to 0.96 percentage points, approximately two times that of the AOCS laboratories. A sample with 19.6% (dry basis) oil sent to a randomly chosen laboratory will receive test values of 17.7-21.5% in 95% of the cases.

Table 6 shows the average spread between blind duplicates by laboratory type. The processor laboratories had the widest spread, 1.31 points, but this was reduced to 0.86 points when the outlier data were excluded. This value was still one-third greater than the outlier-excluded data for the public laboratories and the commercial laboratories (0.63 and 0.60 points, respectively), and nearly twice that of the AOCS laboratories (0.45).

Laboratory method differences. Table 8 has the individual laboratories' moisture methodology information.

TABLE 8

Laboratory Moisture Methodology Information

$Lab^{a,b}$	Type of oven	Drying time (hr)	Drying temperature $(^{\circ}C)$
1	Convection	1	130
2	Convection	1	130
	Convection	$\mathbf{1}$	130
$\frac{4}{5}$	Convection	1	130
	Convection	2	130
$\overline{7}$	Convection	2	135
9	Vacuum	$\overline{\mathbf{4}}$	95-100
10	Convection	18	75–80
11	Convection	3	130
12	Convection	4	105
13	Convection	3	130
14	Convection	3	130
15	Convection	3	130
16	Convection		130
21	Convection	$\frac{3}{2}$	130
22	Convection		135
23	Convection	3	130
24	Convection	$\mathbf{1}$	130
25	Convection	2	130
26	$\operatorname{Convection}$	$\mathbf 1$	130
29	Convection	3	130
30	Convection	2	130

 a Laboratories 1-10 are public, 11-20 are commercial, and 21-30 are processor facilities.

 b Not all laboratories returned methodology data sheets.

Although most laboratories used a drying temperature of 130° C, drying times ranged from 1 to 3 hr. The variable times and temperatures will undoubtly introduce errors when protein and oil values are converted to dry basis. We had to accept each laboratory's moisture data as accurate for this conversion.

Table 9 has the individual Kjeldahl protein methodology information. Inconsistencies showed up in methodology. Catalysts and amounts employed differed, and there were several receiving flask materials $-H_2SO_4$, boric acid, HCl and NaOH. The digestion times ranged from 30 min to 2hr.

Oil methodology information is shown in Table 10. Three solvents were used (petroleum ether, hexane, and ethyl ether). These were combined with four types of equipment (Goldfisch, butt-tube extractor, Soxtec, and Soxhlet). Some laboratories tempered samples, and some did not. Among those that did, tempering times ranged from 30 min to 2 hr.

Tables 8, 9, and 10 indicate the difficulty of comparing values from different labs. The methodology is likely to be different and, therefore, the results are likely to be affected. Although there were too many variant methods to permit meaningful statistical correlations from our data, standardization is obviously needed if analytical laboratories are to agree with a national reference, such as the USDA.

The standard deviations and spreads between duplicates were greater than expected, which explains some of the resistance to NIR analysis in the marketplace. These

TABLE 9

Kjeldahl Protein Methodology Information

aNot all laboratories returned methodology data sheets.

JAOCS, VOl. 68, no. 12 (December 1991)

TABLE 10

Oil Extraction Methodology Information

 a Not all laboratories returned methodology data sheets.

 b If yes, numbers correspond to time (hr) and temperature ($^{\circ}$ C) of tempering.

variabilities must be reduced if NIR is to be used effectively for trading purposes. For NIR calibration, the average of several laboratories' data may be a more accurate standard than either one protein or one oil determination.

ACKNOWLEDGMENT

Journal paper J-14490 of the Iowa Agriculture and Home Economics Experiment Station, Iowa State University, Ames, IA 50011. Research supported by the Iowa Soybean Promotion Board and the Iowa Agriculture and Home Economics Experiment Station, Project 2339.

REFERENCES

- 1. United States Department of Agriculture, *Federal Register 54(35):7778* {1989).
- 2. Cherry, J.P., in *Methods for Protein Analysis,* American Oil Chemists' Society, Champaign, IL, 1988, p. 13.
- 3. Carpiaux, E., *Bull. Soc. Chim. Belg. 27:13* (1912).
- 4. Chalmers, R.A., in *Methods of Protein Analysis,* American Oil Chemists' Society, Champaign, IL, 1988, p. 60.
- 5. Hach, C.C., B.K. Bowden, A.B, Kopelove and S.O. Brayton, J. *Assoc. Off. Anal. Chem.* 70:783 (1987).
- 6. Mears, B., and R.E. Hussey, J. *Ind. and Eng. Chem.* 13:1054 (1921).
- 7. Beet, A.E., *Fuel* 11:406 {1932).
- 8. Crossley, H.E., J. *Soc. Chem. Ind.* 51:237 {1932}.
- 9. Osborn, R.A., and J.B. Wilkie, J. *Assoc. Off Anal. Chem.* 18:604 (1935).
- 10. Snider, S.R., and D.A. Coleman, *Cereal Chem.* Ii:414 (1934).
- 11. Poe, C.F., and R.R. Schafer, *J. Dairy Sci. 18*:733 (1935).
- 12. Kane, P.E, J. *Assoc. Off. Anal. Chem.* 69:644 (1986).
- 13. Kane~ P.F., *Ibid.* 67:869 (1984).
- 14. Official and Tentative Methods of the American Oil Chemists' Society, 3rd edn., edited by R.O. Walker, AOCS, Champaign, IL, 1980.
- 15. *Approved Methods of the American Association of Cereal Chemists,* 8th AACC, St. Paul, MN, 1983.
- 16. *Official Methods of Analysis of the Association of Official Analytical Chemists;* 14th edn., edited by S. Williams, AOAC, Arlington, VA, 1984.
- 17. Arnold, L.K., and R.B. Choudhury, J. *Am. Oil Chem. Soc.* 39:296 {1962}.
- 18. Arnold, L.K., and R.B. Choudhury, *Ibid. 37*:4581 (1960).
- 19. Arnold, L.K., and R.B. Choudhury, *Ibid.* 39:379 (1962}.
- 20. Arnold, L.K., and R.B. Choudhury, *Ibid.* 39:378 (1962}.
- 21. Snyder, H.E., G. Shea, P. Clark and K.L. Wiese, *IbicL* 65:255 (1988}.
- 22. Panford, J., in *Methods for Protein Analysis,* American Oil Chemists' Society, Champaign, IL, 1988, p. 1.
- 23. Williams, P.C., *Cereal Chem.* 52:561 {1975}.
- 24. Robertson, J.A., and F.E. Barton, J. *Am. Oil Chem. Soc.* 61:543 {1984}.
- 25. Watson, C.A., *Anal. Chem.* 49:835a {1977}.
- 26. Barton, F.E., II, and D. Burdick, *J. Agric. Food Chem. 27*:1248 (1979}.
- 27. Brumm, T.J., and C.R. Hurburgh, Jr., J. *Am. Oil Chem. Soa* 67.'302 (1990}.
- 28. Soya Blue Book, American Soybean Association, St. Louis, MO, 1984, p. 23.

[Received May 16, 1991; accepted October 21, 1991]