Identification and Segregation of High-Value Soybeans at a Country Elevator[1]

Charles R. Hurburgh, Jr.
Agricultural Engineering Department, Iowa State University, Ames, Iowa 50011

For three harvest seasons, 1989–1991, a whole-grain near-infrared transmission analyzer was used at a large Iowa elevator to measure oil and protein content of unground soybeans. In the third year, soybeans were physically sorted on the basis of the sum of protein and oil. The sum is a direct measure of processor revenue potential. The top 23% (approximately 150,000 bu or 4500 MT) of soybeans were isolated from the lower 77%. The "high-value" beans (with sum 0.7 percentage points above the average) were isolated for future sale to a processor. The analyzer performed accurately and required about 1.5 min per test. Compared to the normal 1.0 min per test for moisture measurement, this caused an additional delay of about 30 s per load for farmers, which is a significant cost over the 100–300 loads per day received by the elevator. Because this was a new program for elevator employees, errors in data transcription and communication reduced the value difference of the high-sum beans from its theoretical maximum of 1.2 percentage points, but the high-value beans still had a theoretical 10 cent/bushel advantage over average unsegregated beans delivered to that elevator. Cost accounting showed the testing and segregation to cost about 2–3 cents per bushel. Testing and segregation is feasible in the U.S. grain market.

KEY WORDS: **Composition, grading, marketing, oil, protein, quality, soybeans.**

The Federal Grain Inspection Service (FGIS) introduced soybean oil and protein (basis 13% moisture) as Official Criteria on September 4, 1989 (1). This was done to partly meet the 1986 Congressional mandate to develop standards for end-user-related grain quality factors (2). FGIS evaluated several models of ground-grain near-infrared reflectance analyzers and one whole-grain transmittance instrument. The whole-grain unit was ultimately selected for use at FGIS field offices. FGIS reported standard errors of calibration (SECs) of 0.6 percentage points for protein and oil (3), for both types of instruments. These were both later revised downward to 0.4 percentage points (4).

Protein and oil content have a large impact on end-user (crusher) economics. A computer model of a solvent-extraction soybean plant demonstrated that the typical variations in protein and oil create a $1.00 per bushel range in the meal and oil value of soybeans. The term "estimated processing value per bushel (EPVB)" is defined as the total revenue per bushel to a processor from meal and oil (5).

Because much (but not all) of the variation in composition is varietal (6), the maximum value differentiation exists as farmers deliver to country elevators (7). The country elevator must be able to test inbound grain quickly and accurately if protein and oil become marketing criteria. The actual levels are primarily under the control of the farmer, so any attempts to change the average composition of U.S. soybeans must involve direct incentives to growers. The other major variable affecting composition is environment (7), over which the grower has no control. Segregation by composition creates considerable complexity for the country elevator. Figure 1 gives one schematic diagram of the sorting concept applied to a country elevator (8).

Near-infrared spectroscopy has been used to measure oilseed composition (7,9). The ground-grain reflectance instruments are well documented in their accuracy, but introduce difficulties in sample preparation. Country elevators require rapid and simple equipment that is easily standardized across many buying stations. There are approximately 10,000 country grain elevators in the United States (10), with the majority in areas that handle soybeans.

The primary electronic test instrument in use at country elevators is the capacitance-cell moisture meter. These whole-grain analyzers measure soybean moisture in about 30–60 s per sample to an accuracy (standard deviation) of 0.3 percentage points relative to the air-oven (11). In Iowa State University studies, one whole-grain NIRT analyzer, the Trebor 99 (Zeltex, Inc., Gaithersburg, MD), had a moisture standard deviation of 0.33 percentage points (12). Slightly lower figures have been claimed for another currently available NIRT instrument, the Tecator Infratec 1225 (Perstorp Analytical, Inc., Silver Spring, MD). The usefulness of NIRT in soybean analysis depends on its ability to operate accurately in country elevators (as opposed to grain inspection laboratories), and on the existence of economically relevant variations in composition.

An economic ranking procedure is a necessity for trade. There have been several approaches developed for calculating soybean value ($/bushel) on the basis of protein and oil content. The most inclusive is the Brumm–Hurburgh (5) mass-balance processing plant model SPROC. Updaw *et al.* (13) used simplifications of the mass-balance to reduce calculations.

Although models can be programmed into computers attached to NIRT analyzers, they have more complexity than is easily understood at the country elevator. For example,

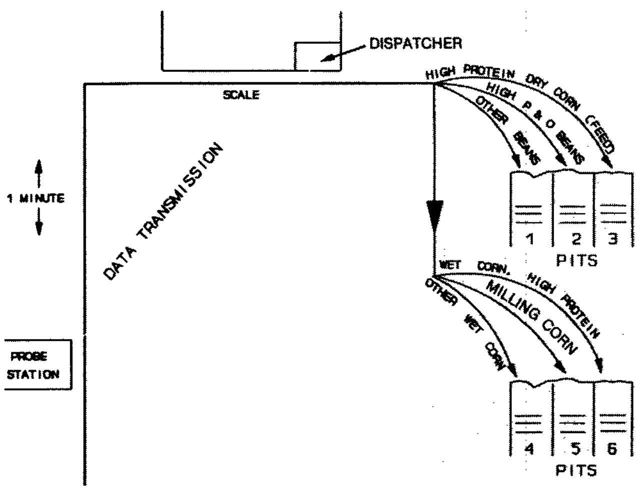

FIG. 1. An example of segregating bulk grains by intrinsic quality parameters (Ref. 8). P, protein; O, oil.

[1]Presented at the 84th AOCS Annual Meeting & Expo, April 27, 1993, Anaheim, California.

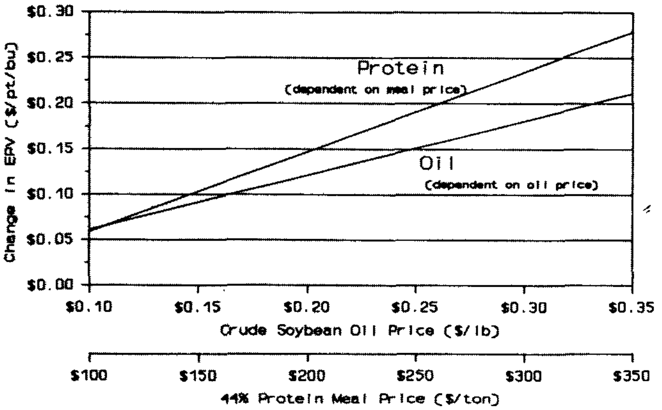

FIG. 2. Dependence of soybean protein and oil incremental values on their respective product prices. EPV, estimated processing value.

all require an estimate of soybean meal and oil prices, which change daily. The relative ranking of EPVB among samples will not shift, but the absolute values of EPVB will. This is confusing to operators who expect to base decisions on some fixed indicator. A simpler criterion is needed for operational decisions, apart from the financial settlement process.

Figure 2, derived from iterative solution of the SPROC model, shows the dependence of soybean protein and oil economics on their respective product prices. Because the incremental values of protein and oil are roughly equal for the price combinations prevalent in recent years, it appeared that the sum (SUM) of protein and oil percentages could be an effective segregation criterion. If the price ratio (per unit weight) of oil to meal varies from the 1.5–2.5 typical of recent years, then the sum sorting strategy will not give correct weighting to either high-oil or high-protein soybeans.

The reported variation in protein and oil content of soybeans delivered to country elevators (standard deviations of 1.0 and 0.5 percentage points, respectively) (6) is theoretically large enough to give value variations of 40–60 cents per bushel. Therefore, a segregation of high-composition soybeans should yield a premium product for a processor willing to pay for it. Segregation generates additional cost, which must be offset by added value. Hurburgh (14) developed a model to estimate testing and segregation costs at elevators. This model requires situation-specific time and economic inputs.

The purposes of this research project were to evaluate the performance of a whole-grain near-infrared transmittance instrument under actual elevator conditions and to assess the practicality of segregating producer-delivered soybeans by composition. The research project was done in the 1989–1991 harvests at the North East Webster (NEW) Cooperative elevator (Vincent, IA). This elevator has a licensed storage capacity of 5.1 million bushels, which ranks eighth largest in the state by 1988 data. Its maximum grain-receiving rate is also among the highest in the state.

MATERIALS AND METHODS

A Tecator Infratec 1225 (Serial *53) was installed at the NEW Cooperative elevator. The Infratec was connected to a portable computer and was operated by elevator personnel. The Infratec had Tecator factory calibration SB7 (in 1991, its successor US22), which measured soybean moisture, protein and oil content. Protein and oil were measured on the as-is moisture basis, then were converted by the computer to a 13% moisture basis. Samples were cleaned over a hand-held 8/64-inch round-hole screen before analysis, but, inevitably, a few sticks or other foreign material remained. The same probe sample used by the elevator for price settlement was used for protein and oil analysis.

Calibration SB7 originated from 1986–1988 crop samples, supplied to Tecator by Iowa State University, and samples from Argentina and Yugoslavia US22 was an update from essentially the same sample set. Some samples were out of range and showed as "outliers." Records were kept of the outliers, but the predicted constituent percentages were used regardless of outlier status. This is the same outlier-handling procedure used by FGIS.

The program SPROC 2.22 (5), which calculates EPVB, was included in the data-acquisition program. Therefore, the soybean oil and meal revenue for a processor was determined for each sample by using the Chicago Board of Trade meal and oil prices as of September 1 of each year. This calculation was done after the conversion to constant-moisture basis sand therefore does not reflect any change in value due solely to low or high moisture (relative to 13%). The EPVB calculation was done for later comparison with SUM.

When possible, varieties were identified. Samples were labeled with the patron's name and customer number. The elevator's settlement moisture content, obtained with a state-approved Dickey-john GACII capacitance moisture meter, was recorded for later comparison with the Infratec moisture.

In 1989–1990, the Infratec was physically located in the elevator's office, adjacent to the scale area. In the first two years, we did not attempt to sort the soybeans but only collect data. The elevator did its weighing and moisture analysis in the office. Farmers reported the variety to the Infratec operator when they picked up their scale tickets after dumping.

In 1991, the elevator opened its remote probing station for soybean harvest. The probe station was located about 200 feet from the scale. Normally, it took farmers about 45 s to travel from the probe station to the scale. This delay was crucial to the segregation process because the dispatcher (scale operator) had to have the Infratec results before telling the driver where to dump.

The 1991 soybeans were physically sorted by SUM. It was decided in advance that the estimated 600,000–700,000 bushels of beans received at harvest would be divided into 1/3 high-value, 2/3 average value. Assuming a normal distribution of SUM, this distribution would be produced if the cutoff for "high value" (HI) were 0.5 standard deviations above the mean. Data from 1989–1990 were used to predict the standard deviation for 1991. The mean for 1991 was first estimated from testing the earliest deliveries, then updated as more data became available. The remaining beans were classified as "low value" (LO).

The segregation procedure required considerable manual labor. The Infratec operator recorded the patron's name, protein, oil and SUM. This information was communicated orally to the scale operator, who recorded the data again, and assigned the load to a dump area. One dump pit was allocated to the high-value soybeans. The scale operator also recorded the variety information if the producer knew it. There was no way to monitor the producers to see if they actually went to the assigned pit.

Data analysis included: (i) regression of SUM against EPVB, to measure the accuracy of SUM as a segregation criterion, (ii) distribution of composition within and across years, (iii) evaluation of varieties for consistency across growers and years, (iv) identification, as possible, of the accuracy with which the loads that should have been declared HI were actually put in the HI storage, and (v) application of cost-accounting information to the segregation cost model. Selected samples were retained for laboratory analysis at Iowa State University, as a check on the NIRT unit.

RESULTS AND DISCUSSION

The usefulness of SUM was evaluated first. Table 1 shows the regression equations between SUM and EPVB. Of most note are the low standard deviations, about 1 cent per bushel. This is not large enough to be of concern in using SUM as a segregation criterion. The slope coefficients per bushel (value of one percentage increment of sum) vary with price, but selections based on SUM were consistently correct. The oil-meal price ratio varied from 2.1 in 1989 to 2.7 in 1990 and back down to 2.2 in 1991.

An overall summary of the three years' data is give in Table 2. The protein and oil content varied among years, probably because of weather differences. The 1990 season had considerably hot weather in late August and early September. This provided late-season heat units that were probably utilized to produce extra protein. Typically shorter-season growing areas have lower protein content (7). Clearly, a great diversity of farms and varieties were represented. Each sample represented about 450 bushels. The smaller n (number of samples) value in 1990 was caused because the Infratec was not operation in the first 5 d of harvest.

Figure 3 shows the pattern of daily deliveries in the three years. Soybean harvest lasted approximately 14 d, but the majority of grain was received in a one-week period at the center of the harvest. The data clearly show the dilemma of a country elevator wanting to be more quality-orientated. Their receiving system must be sized for an extreme peak that is hit for only one or two days per year. Investments in automated testing are hard to justify on this pattern of usage, so elevators prefer to deal with quality factors they know will be at acceptable levels in farmer deliveries. In this way, each load will not have to be rigorously tested.

Variety is known to be an important determinant of soybean composition. We were able to collect variety information on 91, 90 and 69% of the samples in the three years, respectively. The separation of test station and scale operation (in 1991) reduced the efficiency of collecting variety information.

TABLE 1

Effectiveness of SUM[a] as an Economic Ranking Criterion for Soybeans

	September 1 product prices		Relationship between EPVB[b] and SUM		
Year	Meal[c] ($/ton)	Oil[d] ($/ton)	Equation	R^2 (%)	Standard deviation ($/bu)
1989	183.10	386.00	EPVB = 0.147 (SUM) − 1.39	97.8	0.013
1990	181.10	486.00	EPVB = 0.122 (SUM) + 0.39	87.8	0.010
1991	185.00	414.00	EPVB = 0.120 (SUM) + 0.23	99.6	0.008

[a]Protein percentage + oil percentage, basis 13% moisture.
[b]Estimated processed value per bushel, SPROC Model (Ref. 5).
[c]44% protein; Chicago Board of Trade.
[d]Crude, degummed; Chicago Board of Trade.

TABLE 2

Composition[a] of Soybeans Tested (at Vincent, IA, 1989–1991)

	1989 (n = 1506)		1990 (n = 1069)		1991 (n = 1428)	
Factor	Avg. (%)	Std. dev. (% pts)	Avg. (%)	Std. dev. (% pts)	Avg. (%)	Std. dev. (% pts)
Protein	34.9	1.0	35.6	0.9	35.3	1.2
+						
Oil	18.5	0.7	19.1	0.6	18.4	0.7
= SUM[b]	53.4	0.8	54.7	0.7	53.7	1.0
EPVB[c] ($/bu)	$6.45	$0.12	$7.08	$0.09	$6.68	$0.13
Number of patrons	232		206		281	
Number of varieties	105		67		98	

[a]Percentages basis 13.0% moisture.
[b]See Table 1 for explanation.
[c]Estimated processed value per bushel ($/bu). Prices: 44% Meal: 1989, $183.10/ton; 1990, $181.10/ton; 1991, $185.00/ton. Oil: 1989, $386.00/ton; 1990, $486.00/ton; 1991, $414.00/ton.

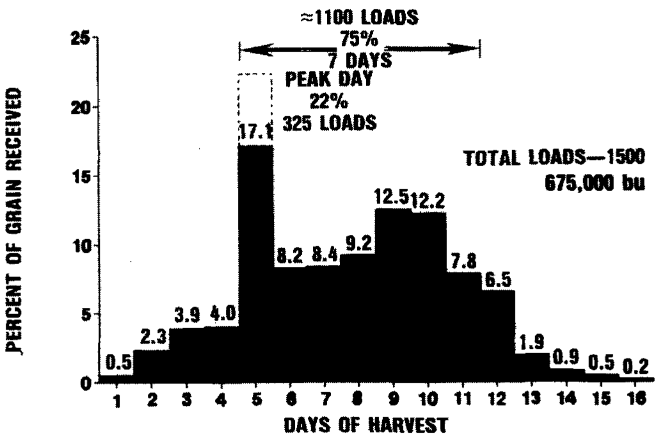

FIG. 3. Daily grain receiving patterns at a country elevator, 1989–1991.

Farmers' choices of varieties (probably not based at all on composition) shifted dramatically and quickly across years. Table 3 shows the top five (in number of samples) varieties by year, with their composition compared with the respective year averages. In all years, about 45% of the soybeans were the top five varieties, even though 50–80 varieties were represented.

Consistently, the top five varieties were at least equal in SUM to the average of all samples. These varieties were chosen for yield and agronomic performance, but their selection apparently did not reduce composition. Indeed, the leading variety was always above average in composition.

This leads to the question of varietal consistency within years across growers and within farms within years. An alternative to testing each sample could be to pre-identify varieties and to sort on that basis. The disadvantage of varietal sorting is that there is no direct at-sale check on conformance with specifications. It also requires advance knowledge of which varieties are good and bad.

Table 4 shows the variability (standard deviation) of composition within a growing year. The variation within a variety was about half that of the data as a whole. Because some varieties probably were misidentified, the actual variation within variety may be yet smaller than listed. Farm-to-farm effects contributed some of the within-variety variation because the variability among loads of a single variety from a single producer was about one-third that of the data as a whole. Assuming that producers identified loads correctly by variety, reasonably good sorting could be done by variety or by testing the first few deliveries of a variety by each producer, without testing each load.

Varieties were consistent from year to year. Their relative ranking by SUM did not change. The key point is that the relative position of varieties does not change greatly, even though climatic conditions do change the absolute values of composition.

The effectiveness of the 1991 segregation is shown in Table 5. The actual SUM differential between HI and unsegregated average soybeans was 60% of its theoretical maximum (0.7/1.2). According to Table 1, a 0.7 percentage point difference in SUM should be worth 9 cents per bushel. If segregation occurs, the LO fraction will be of lower value than average, but as long as the

TABLE 3

Composition[a] of Five most Frequently Planted Soybean Varieties (Vincent, IA, 1989–1991)

Year	Rank	Variety	Percent of samples[b]	Protein +	Oil =	SUM[c]
1989	1	Hoffman 8300	13.1	35.0	18.9	53.9
	2	Pioneer 9271	10.6	33.8	18.2	52.0
	3	LOL 2456	9.4	34.7	18.5	53.2
	4	NK 2303	5.3	37.0	17.7	54.7
	5	LOL 2828	5.0	35.6	17.9	53.5
		Total	43.4	34.9	18.4	53.3
Year averages				34.9	18.5	53.4
1990	1	NK 2303	10.9	37.0	18.4	55.4
	2	Hoffman 8300	9.7	35.3	19.4	54.7
	3	LOL 2333	9.6	36.1	18.7	54.8
	4	Pioneer 9272	7.3	35.7	19.0	54.7
	5	LOL 2828	4.9	35.3	18.9	54.3
		Total	42.4	35.9	18.9	54.8
Year averages				35.6	19.1	54.7
1991	1	LOL 233	19.9	36.2	18.0	54.2
	2	Pioneer 9272	8.9	35.2	18.5	53.7
	3	Pioneer 9202	7.5	34.5	18.8	53.3
	4	Hoffman 8300	6.6	35.0	18.6	53.6
	5	Asgrow 2234	3.3	35.1	19.0	54.1
		Total	46.1	35.6	18.4	54.0
Year averages				35.3	18.4	53.7

[a]Basis 13.0% moisture.
[b]With variety known.
[c]See Table 1 for explanation.

TABLE 4

Variability in Soybean Composition[a] Within a Growing Season (1989–1991, Vincent, IA)

Source of variability	Average standard deviation, 3 yr pooled		
	Protein (%)	Oil (%)	SUM[b] (%)
Overall	1.0	0.7	0.8
Within a variety	0.5	0.4	0.4
Within a variety on a single farm	0.4	0.3	0.4

[a]Basis 13% moisture.
[b]See Table 1 for explanation.

predominant resale market does not analyze, then these beans will receive average price.

We lost 187 HI loads to the LO storage, and we put 104 LO or unknown loads in the HI storage. This was a 20% error rate. Principal causes for error were incorrect addition of protein and oil (5%), misapplication of the cutoff criterion (5%) and miscommunication of data from test operator through scale operator to farm (10%).

HI value was defined as SUM = 54.0 or greater. This was set on the first day as 0.5 standard deviation over the mean and was not changed through harvest. This method of setting the cutoff was accurate, as 32% of loads tested as HI.

The 1991 beans, when sold, were tested for protein and oil by an official FGIS-designated inspection agency. The agency used an Infratec 1225, but with the official USDA calibrations as described in References 1, 3 and 4. The soybeans tested 35.5% protein and 18.1% oil (average of all samples) by the official test. This is close to the average reported in Table 2 for the original test.

Pertinent cost variables were evaluated as inputs to the cost model (14), as summarized in Table 6. The principal item to note is the 30-s additional wait imposed on each farmer. This may not seem great, but was objectionable to many producers rushing to return to harvesting operations.

These inputs were used in the segregation cost model to estimate per-bushel costs for sorting. The results in Table 6 should be interpreted as a case study, for specific conditions and assumptions only. Costs were included as if the elevator had purchased equipment and was operating on its own without assistance from the University. However, the costs to the elevator were considerably lower than most observers predict for differentiated marketing. A major component, underutilized space, was zero because grain facilities currently have excess storage space. The overall cost estimate, 3.7 cents per bushel, was considerably less than the theoretical additional value of the sorted beans.

Check sample data are reported in Table 7. The primary check was our ground-grain reflectance unit (7) because laboratory chemistry was expensive. Some samples were analyzed by chemistry, however. The agreement among methods was good. The standard deviations between near-infrared analyzers and between the Infratec and chemistry were similar to those reported for the official USDA soybean program (4).

The testing and segregation project was a success. The NIRT analyzer operated in elevator conditions, although it was slightly slow for this elevator. SUM is a simple criterion for ranking soybeans by processor value. The grain market can isolate soybeans by composition, but automated data handling will be needed to reduce classification errors and to increase efficiency.

TABLE 5

Pit Classification (Vincent soybean segregation,[a] 1991)

Pit[b]	Number of loads	Percent of loads	Protein (%)	+	Oil (%)	=	SUM[c] (%)
3 (HI)	323[d]	22.6	35.8		18.3		54.1
2 (LO)	1050	73.5	35.1		18.3		53.4
Not recorded[e]	55	3.9	34.9		18.2		53.1
Totals	1428	100.0	35.3		18.4		53.7
2, With not listed[d]	1105	77.4	35.1		18.3		53.4
Theoretical 3[f]	455	31.8	35.9		18.5		54.4

[a]1428 Loads, 98 different varieties from 281 customers. 1274 tested—406 high-value (HI), 868 low-value (LO). 154 were not tested. Protein, oil and SUM percentages basis 13.0% moisture.
[b]Assigned at office.
[c]See Table 1 for explanation.
[d]Of 406 loads tested as HI, 219 were correctly assigned (53.9%), and 104 LO or not tested were put in HI (pit 3).
[e]Includes pit 1 probe shack assignments.
[f]Actually tested HIs (406) plus 32% of loads not tested. LO would be SUM = 53.2 (a 1.2 point spread).

TABLE 6

Segregation Cost Analysis[a]

Cost category	Cost basis	Item	Amounts ($/bu) By item	Amounts ($/bu) Cumulative
Grading	V	Tester cost	0.0071	
	B	Operate tester	0.0003	
	V	Data equipment	0.0019	
	B	Wait time	0.0008	
	V	Sample storage	0.0006	
	B	Accounting	0.0001	
	B	Standardization	0.0004	
	V	Software	0.0004	
	B	Disputes	0.0025	
	Subtotal grading			$0.0141 (51%)
Handling	A	Waiting time	0.0028	
	A	Pit labor	0.0019	
	V	Modifications	0.0000	
	V	Underutilized storage	0.0000	
	A	Misgrades	0.0050	
	V	New storage	0.0000	
	V	Loss in receiving	0.0042	
	Subtotal handling			$0.0139 (49%)
	Total cost per bushel			$0.0280

[a]V—Volume-based ($0.0142; 52%); B—load size-based. ($0.0041; 15%); A—across-the-board ($0.0092; 33%).

TABLE 7

Comparisons of Soybean Protein and Oil Contents Measured in a Country Elevator and in the Iowa State University Laboratory (standard deviation in parentheses)

Year	n	Protein (%)[a] Infratec[b]	− Instalab[c]	= Difference	Oil (%)[a] Infratec[b]	− Instalab[c]	= Difference
1989	22	34.6	34.7	−0.1 (0.47)	18.7	18.8	−0.1 (0.34)
1990	33	35.5	35.7	−0.2 (0.52)	19.1	19.1	0.0 (0.37)
1991	30	35.3	35.4	0.1 (0.42)	18.5	18.5	−0.0 (0.35)
			Kjeldahl[d]			Extraction[d]	
1989–1991	14	35.0	35.2	−0.2 (0.42)	18.3	18.2	0.1 (0.40)

[a]Basis 13.0% moisture.
[b]Tecator Infratec 1225, SB7 (US22 in 1991) manufacturer calibration, operated at New Cooperative Elevator (Vincent, IA).
[c]Dickey-john Instalab 800, S5 calibration (S6 in 1991), calibrated and operated at Iowa State University (Ames, IA), Magic Mill III+grinder.
[d]Chemistry data from Woodson-Tenant, Inc. (Des Moines, IA).

ACKNOWLEDGMENTS

Journal Paper J-15324 of the Iowa Agriculture and Home Economics Experiment Station. Research supported by the Iowa Soybean Promotion Board and the Iowa Agriculture and Home Economics Experiment Station, Project 2339.

REFERENCES

1. Anon., *Federal Register 54(157)*:33702 (1989).
2. Grain Quality Improvement Act, Public Law 99-641, U.S. Congress, Washington, D.C., 1986.
3. United States Department of Agriculture, Federal Grain Inspection Service, Washington, D.C., Certificate of Conformance FGIS 89-101.
4. Watson, C.A., and M. Bashanti, Soybean Protein and Oil Pilot Field Test, USDA-FGIS, Washington, D.C., 1990.
5. Brumm, T.J., and C.R. Hurburgh, Jr., *J. Am. Oil. Chem. Soc 67*:302 (1990).
6. Hurburgh, C.R., Jr., and T.S. Brumm, *Appl. Eng. Agric. 6(1)*:65 (1990).
7. Hurburgh, C.R., Jr., T.S. Brumm, J.M. Guinn and R.A. Hartwig, *J. Am. Oil Chem. Soc. 67*:966 (1990).
8. Hurburgh, C.R., Jr., *Proceedings of GEAPS Exchange '90*, Grain Elevator and Processing Society, Minneapolis, 1990, p. 72.
9. Panford, J.A., P. Williams and J. deMan, *J. Am. Oil Chem. Soc. 65*:1627 (1988).
10. *Sosland Publishing Co. Grain Guide 1992*, Sosland Publishing Co., Shawnee Mission, 1992.
11. Hurburgh, C.R., Jr., *Trans. ASAE 30(2)*:582 (1987).
12. Hurburgh, C.R., Jr., *INFORM 4*:353 (1993).
13. Updaw, N.J., J.B. Bullock and T.E. Nichols, Jr., *Southern J. Agric. Econ. 8*:129 (1976).
14. Hurburgh, C.R., Jr., *Appl. Eng. Agric.*, in press.

[Received October 21, 1993; accepted April 18, 1994]