Estimating Soybean Seed Protein and Oil Concentration Before Harvest

R.W. Yaklich* and B.T. Vinyard

USDA-ARS, Soybean Genomics and Improvement Laboratory, Biometrical Consulting Service, Beltsville Agricultural Research Center, Beltsville, Maryland 20705

ABSTRACT: Temperature and precipitation during the growing season have been shown to influence the protein and oil composition of the soybean [*Glycine max* L. Merr.] seed. A method based on these parameters was developed to estimate protein and oil concentrations of the seed before harvest. This method was developed with protein and oil data and temperature and precipitation data from the Uniform Soybean Tests, Southern Region, for the years 1975 to 1983. Classification and regression "tree-based" analyses were used to determine the month and numeric value ("splitting point") of the environmental variable that correctly classified the variation from median protein and oil composition for the 126 location-years. Temperature in September was most influential in determining the splitting point for three of the four variables. Oil concentrations from the location-years were separated into low vs. high median-based boundary categories most readily by the September sum of minimum temperatures. Total protein and oil concentrations from the location-years were classified best by September growing degree days. Protein-to-oil ratios were best separated by the September mean minimum temperature. The August mean maximum temperature best separated protein concentration. These data demonstrate that temperature during specific months of the crop year were useful in estimating the final concentration of protein and oil in the seed and could be used by seed processors to estimate seed composition before harvest.

Paper no. J10644 in *JAOCS 81*, 189–194 (February 2004).

KEY WORDS: *Glycine max*, oil, protein-to-oil ratio, seed composition, seed protein, temperature, total protein and oil.

No published methods exist to help the soybean seed processor estimate the protein and oil concentrations of the soybean (*Glycine max* L. Merr.) seed before harvest. Postharvest research has shown that the environment influences seed composition. We know that seed produced in the Corn Belt contains more oil and less protein than that produced in other areas of the soybean-growing region (1) and that protein concentration in soybean meal is influenced by the total protein and oil concentration of the seed (2).

Previously, in a 51-year comparison of the Southern Region (SR) vs. Northern Region (NR) Soybean Uniform Tests (UT) (3), we noticed that the yearly protein and oil concentrations in both tests were similarly influenced by environmental variables as compared with the previous year's mean. In most years, protein and oil concentrations in both regional soybean tests would increase or decrease from the previous year's values. These yearly trends also were found in a comparison of the mean yearly oil and protein concentrations of canola-rape (*Brassica napus* L.) and flax (*Linum usitatissimum* L.) grown in Canada between 1956 and 1998. These long-term data indicated that the growth environment influenced yearly seed composition in three major oilseed crops grown in the United States and Canada. This analysis also indicated that trends in these data sets lasted for a number of years and extended over a large geographical area.

Recently, we attempted to determine whether the oil concentration of soybeans grown in the United States was changing (4), as suggested by one of our largest importers of soybeans (1). Data from the NR and SR UT indicated that oil concentrations in both tests increased from 1948 to 1973 and decreased in both tests from 1974 to 1982 to their lowest value in 1982, then increased dramatically in 1983 in both tests. Since 1983, the oil concentration has continued to decrease in the NR UT. A decline in oil concentration was observed from all maturity group (MG) tests that were evaluated in 1974, and an increase in oil concentration was found in all lines evaluated in 1983. These data indicated that environmental factors were responsible for the annual trends in yearly oil concentration. The long-term decline in oil concentration of the seed between 1974 and 1983 affords an opportunity to determine the effect of monthly temperature and precipitation on this important constituent of the soybean seed. Since protein and oil concentrations are negatively correlated (5), these data sets provide an opportunity to examine weather variables that may have a direct influence on the protein and oil concentrations of the seed at harvest.

In this study, we use UT from SR testing locations that consecutively reported protein and oil data for the years 1975 to 1983 to calculate their yearly protein and oil means. We also obtained the daily maximum and minimum temperature and precipitation data for these testing stations and generated monthly values for April through September. These data were used to determine the monthly weather variables that were most closely related to the protein and oil concentrations of the soybean seed.

*To whom correspondence should be addressed at USDA-ARS, Soybean Genomics and Improvement Laboratory, Biometrical Consulting Service, Beltsville Agricultural Research Center, Beltsville, MD 20705.
E-mail: yaklichr@ba.ars.usda.gov

MATERIALS AND METHODS

The SR UT critically evaluate the best of the experimental soybean lines developed by federal and state research personnel in the United States and are lines adapted to MG IVS through VIII. Nine years (1975 to 1983) of data (6–14) were chosen based on findings in a previous study (4). The SR locations were Clinton, North Carolina (34.95°N, 78.31°W); Jay, Florida (30.95°N, 87.15°W); Keiser, Arkansas (35.67°N, 90.09°W); Portageville, Missouri (36.43°N, 89.70°W); Queenstown, Maryland (38.99°N, 76.16°W); Stoneville, Mississippi (33.4°N, 90.9°W); Stuttgart, Arkansas (34.49°N, 91.55°W); Tifton, Georgia (31.46°N, 83.51°W); and Warsaw, Virginia (37.96°N, 76.76°W).

Protein and oil concentrations on a moisture-free basis were determined from a composite sample of replications of each individual line at a location and listed on a moisture-free basis. Chemical analyses were performed at the National Center for Agricultural Utilization Research, USDA-ARS, Peoria, Illinois. The method of analysis for these constituents changed as faster and less expensive methods became available. In 1974 the method of analysis for protein and oil concentrations changed from the Kjeldahl method and NMR, respectively, to IR reflectance. Protein and oil were reported on a percent dry-weight basis and converted to SI units. Total protein and oil concentration was the sum of the individual entry values for protein and oil. The protein-to-oil ratio was derived by dividing the individual entry protein concentration by its oil concentration.

Soybean lines from three MG (V, VI, and VII) were used for further analysis. MG V locations were Keiser, Arkansas; Portageville, Missouri; Queenstown, Maryland; Stoneville, Mississippi; and Warsaw, Virginia. MG VI locations were Jay, Florida; Keiser, Arkansas; Portageville, Missouri; Stoneville, Mississippi; and Stuttgart, Arkansas. MG VII locations were Clinton, North Carolina; Jay, Florida; Stoneville, Mississippi; and Tifton, Georgia. Some of the locations evaluated lines from more than one MG.

Twelve soybean lines were evaluated each year at every location. The 12 lines evaluated were different for each MG but were the same for each location within an MG. These 12 lines evaluated in an MG were used to calculate the location yearly means ("object classes") for protein concentration, oil concentration, total protein and oil concentration, and the protein-to-oil ratio for each MG location. This resulted in 126 location × MG yearly means. For ease of analysis, values for each seed composition variable were divided into two object classes (above or below the median) and each year × location mean was assigned to the appropriate class. Median values for each seed composition variable were 414 g kg^{-1} for protein concentration, 200 g kg^{-1} for oil concentration, 615 g kg^{-1} for total protein and oil concentration, and 2.07 for the protein-to-oil ratio.

Weather data for these locations were downloaded from the National Oceanic and Atmospheric Administration web site (http://www.noaa.gov). The minimum and maximum daily temperatures and daily precipitation were obtained for April through September. Because some of the weather data sets were not complete, weather data for April 1982 from Caruthersville, Missouri, were substituted in the Portageville, Missouri, data set; Brewton, Alabama, weather data were used for Jay, Florida; and Chestertown, Maryland, weather data was used for Queenstown, Maryland. April through September weather data for 1975 through 1983 were used to create monthly weather variables ("predictor variables") for each year and location. The calculated monthly weather variables that separated the most location years were growing degree days, stress degree days, average mean temperature, average maximum temperature, average minimum temperature, and sum of the minimum temperatures (Table 1).

Data were analyzed using the S-Plus software tree-based models (15) of the Classification and Regression Tree (CART) binary tree-growing algorithm developed by Breiman *et al.* (16). Tree-based modeling is a relatively new approach to classification and regression problems in atmospheric science. It has been used, for example, to classify weather types that are most strongly associated with the presence or absence of rainfall in the Columbia River basin and middle-Atlantic regions (17) and with the severity of winters in the Laurentian Great Lakes Basin. (18). A classification tree is a practical way to predict the class of an object (low vs. high protein and oil values in our case) from values of the predictor variables (monthly temperature and precipitation values). The fundamental idea of the CART algorithm is to select each split of a subset so the data in each of the descendant subsets are "purer" than the data in the parent subset. In a completely pure node, all the cases belong to the same class.

Tree-based analyses were used to determine the months and weather variables most indicative of the oil and protein concentrations of the seed. A separate tree model was fit to each individual calculated weather variable (Table 1) for each month. Each tree model's binary split-point provided a "boundary" to use in determining how closely correlated location × MG × year means on opposite sides of this boundary were with the two object classes (above vs. below the median seed composition). For example, in Table 2, August Av Max Temp had a split-point of

TABLE 1
Calculated Monthly Weather Variables Most Closely Related to Protein and Oil Concentrations of the Soybean Seed

Monthly weather variable	Formula
Growing degree days (GDD)	Total accumulated [(maximum daily temperature + minimum daily temperature)/2 – 10°C]
Stress degree days (SDD)	Total accumulated (maximum daily temperature – 30°C)
Average mean temperature (Av Mean Temp)	Total accumulated [(maximum daily temperature + minimum daily temperature)/2]
Average maximum temperature (Av Max Temp)	Total accumulated maximum daily temperature/days in the month
Average minimum temperature (Av Min Temp)	Total accumulated minimum daily temperature/days in the month
Sum of minimum temperature (Sum Min Temp)	Total accumulated minimum daily temperature

32.0°C. Fifty-nine location-years had an August Av Max Temp <32.0°C, and 67 location-years had an August Av Max Temp above the split-point value. The data also demonstrated that 40 of the 59 location-years in the left node had protein concentrations ≤414.0 g kg^{-1} and that 46 of the 67 location-years in the right node had protein concentrations >414.0 g kg^{-1}. The tree model chose an August Av Max Temp of 32.0°C as the best temperature to correctly classify protein concentration for the most location-years and was able to correctly classify (above or below the median) 86 of the 126 location-years. This procedure was used to identify weather variables and months most directly related to protein, oil, total protein and oil concentration, and the protein-to-oil ratio of the soybean seed.

A three-way ANOVA was conducted for each of the four seed composition variables: protein, oil, total protein and oil concentration, and the protein-to-oil ratio using SAS PROC MIXED (19). Bonferroni multiplicity-adjusted mean comparisons were calculated for each level of MG, MG × location, location, and year. The Fisher exact test was used to determine the significance of the relationship between individual weather variables and object classes. Odds ratios and confidence intervals were calculated using exact methods in SAS PROC FREQ to indicate the strength of this relationship.

RESULTS AND DISCUSSION

An odds ratio provides an intuitive measure of likelihood (explained more specifically in the paragraphs below) by using the observed data. The odds ratios for all weather variables (Tables 2–5) were statistically significant; each 95% confidence interval consisted of values above 1, and equivalently, the Fisher exact test was significant for all 2 × 2 frequency tables reported under "number of location-years."

The median protein concentration, 414.0 g kg^{-1}, was used to define low protein concentration location-years, 376.0 to 414.0 g kg^{-1}, and high protein concentration location-years, 415.0 to 455.0 g kg^{-1}. The weather variables that best classified protein concentration into the high and low categories were derived from August temperature data (Table 2). For all weather variables in Table 2, the odds of low (below median) protein concentration were significantly greater than the odds of high (above median) protein concentration when August temperatures were below their respective "classification values." The odds ratios for these weather variables ranged from 4.36 to 4.56. For example, for August Av Max Temp <32.0°C, the odds of low-to-high protein concentration were (40/59)/(19/59) = 2.10. For August Max Temp >32.0°C, the odds of low-to-high protein concentration were (21/67)/(46/67) = 0.46. The ratio of these odds, 2.10/0.46 = 4.56, was the observed likelihood of low (below median) protein concentration when August Av Max Temp was <32.0°C. When August Av Mean Temp was <26.4°C or August growing degree days were <507.8, then protein concentration was 4.36 times more likely to be below rather than above the median protein concentration.

The median oil concentration, 200.0 g kg^{-1}, was used to define low oil concentration location-years, 170 to 199 g kg^{-1}, and high oil concentration location-years, 200 to 228 g kg^{-1}. The weather variables that best classified oil concentration into the high and low categories were derived from September minimum temperature and accumulated heat units calculated by either growing degree days or stress degree days from the months of September, April, July, and May (Table 3). For all weather variables in Table 3, the odds ratios indicated that cooler weather in the months of September, April, July, and May was at least 5.22 times more likely to yield a low (below-median) oil concentration than a high (above-median) oil concentration. The odds ratios for these weather variables ranged from 5.22 to 8.73.

The median total protein and oil concentration, 615.0 g kg^{-1} (Table 4), was used to define low total protein and oil concentration location-years, 558 to 614 g kg^{-1}, and high total protein and oil concentration location-years, 615 to 649 g kg^{-1}. Variations of September temperatures that were below the median were 20.35 to 24.12 times more likely to result in low (below-median) total protein and oil concentration than a high (above-median) total protein and oil concentration. The

TABLE 2
Monthly Weather Variables and Classification Values That Best Separate Location-Years for Protein Concentration

Weather variable[a]	Classification value for weather variable (°C)	Number of location-years: Protein concentration ≤414 g kg^{-1}	>414 g kg^{-1}	Odds ratio[b]	Odds ratio (95% confidence interval)[c]
August Av Max Temp	<32.0	40	19	4.56[d]	(2.18, 9.78)
	>32.0	21	46		
August Av Mean Temp	<26.4	43	23	4.36	(2.06, 9.23)
	>26.4	18	42		
August GDD	<507.8	43	23	4.36	(2.06, 9.23)
	>507.8	18	42		

[a]See Table 1 for definitions.
[b]Odds ratios were calculated using the 2 × 2 frequency table counts reported under "number of location-years."
[c]The exclusion of values ≤1.0 from all confidence intervals indicates that all odds ratios are statistically significant (α = 0.05).
[d]Location-years with weather values below the classification value were "odds ratio" times more likely to have below-median protein concentrations than above-median concentrations.

TABLE 3
Monthly Weather Variables and Classification Values That Best Separate Location-Years for Oil Concentration

Weather variable[a]	Classification value for weather variable (°C)	Number of location-years: Oil concentration <200 g kg^{-1}	≥200 g kg^{-1}	Odds ratio[b]	Odds ratio (95% confidence interval)[c]
September Sum Min Temp	<1019.3	50	18	8.73[d]	(3.89, 19.57)
	>1019.3	14	44		
September GDD	<373.9	40	10	8.66	(3.72, 20.18)
	>373.9	24	52		
September SDD	<–192.0	40	11	7.72	(3.39, 17.63)
	>–192.0	24	51		
April SDD	<–338.1	47	17	7.32	(3.33, 16.07)
	>–338.1	17	45		
July SDD	<71.4	49	21	6.38	(2.92, 13.94)
	>71.4	15	41		
May SDD	<–254.2	45	19	5.36	(2.50, 11.47)
	>–254.2	19	43		
April GDD	<177.0	40	15	5.22	(2.42, 11.29)
	>177.0	24	47		

[a]See Table 1 for definitions.
[b]Odds ratios were calculated using the 2 × 2 frequency table counts reported under "number of location-years."
[c]The exclusion of values ≤1.0 from all confidence intervals indicates that all odds ratios are statistically significant (α = 0.05).
[d]Location-years with weather values below the classification value were "odds ratio" times more likely to have below-median oil concentrations than above-median concentrations.

four August temperature variables indicated that cooler August temperatures were 5.78 to 9.45 times more likely to result in a low (below-median) total protein and oil concentration than a high (above-median) total protein and oil concentration. Below-median July Av Mean Temp and June stress degree days were 6.11 and 8.76 times more likely to result in a low (below-median) total protein and oil concentration than a high (above-median) total protein and oil concentration.

TABLE 4
Monthly Weather Variables and Classification Values That Best Separate Location-Years for Total Protein and Oil Concentration

Weather variable[a]	Classification value for weather variable (°C)	Number of location-years: Total protein and oil concentration <615.0 g kg^{-1}	≥615.0 g kg^{-1}	Odds ratio[b]	Odds ratio (95% confidence interval)[c]
September GDD	<385.8	50	8	24.12[d]	(9.32, 62.34)
	>385.8	14	54		
September Av Mean Temp	<22.8	54	13	20.35	(8.19, 50.60)
	>22.8	10	49		
August SDD	<32.2	50	17	9.45	(4.19, 21.34)
	>32.2	14	45		
August GDD	<503.2	46	14	8.76	(3.91, 19.64)
	>503.2	18	48		
August Av Max Temp	<32.2	47	15	8.66	(3.88, 19.35)
	>32.2	17	47		
June SDD	>4.2	53	22	8.76	(3.81, 20.13)
	>4.2	11	40		
July Av Mean Temp	<26.9	41	14	6.11	(2.79, 13.39)
	>26.9	23	48		
August Av Min Temp	<22.1	46	19	5.78	(2.69, 12.46)
	>22.1	18	43		

[a]See Table 1 for definitions.
[b]Odds ratios were calculated using the 2 × 2 frequency table counts reported under "number of location-years."
[c]The exclusion of values ≤1.0 from all confidence intervals indicates that all odds ratios are statistically significant (α = 0.05).
[d]Location-years with weather values below the classification value were "odds ratio" times more likely to have below-median total protein and oil concentrations than above-median concentrations

TABLE 5
Monthly Weather Variables and Classification Values That Best Separate Location-Years for Protein-to-Oil Ratio

Weather variable[a]	Classification value for weather variable (°C)	Number of location-years: Protein-to-oil ratio ≤2.07 g kg^{-1}	>2.07 g kg^{-1}	Odds ratio[b]	Odds ratio (95% confidence interval)[c]
September Av Min Temp	<17.0	32	51		
	>17.0	32	11	4.64[d]	(2.05, 10.48)
April SDD	<338.1	23	41		
	>338.1	41	21	3.48	(1.67, 7.24)
September Av Mean Temp	<22.4	19	37		
	>22.4	45	25	3.51	(1.68, 7.33)
July Av Max Temp	<33.1	25	41		
	>33.1	39	21	3.05	(1.47, 6.30)
June SDD	<–72.0	20	36		
	>–72.0	44	26	3.05	(1.47, 6.33)
July SDD	<71.4	28	42		
	>71.4	36	20	2.70	(1.31, 5.58)

[a]See Table 1 for definitions.
[b]Odds ratios were calculated using the 2 × 2 frequency table counts reported under "number of location-years."
[c]The exclusion of values ≤1.0 from all confidence intervals indicates that all odds ratios are statistically significant (α = 0.05).
[d]Location-years with weather values below the classification value were "odds ratio" times more likely to have below-median protein-to-oil concentrations than above-median ratios.

The median protein-to-oil ratio, 2.07 (Table 5), was used to define low protein-to-oil ratio location-years, 1.75 to 2.07, and high protein-to-oil ratio location-years, 2.08 to 2.52. September Av Min Temp and September Av Mean Temp greater (warmer) than the classification values were 3.51 and 4.64 times more likely to yield protein-to-oil ratios below the median value than above the median value. June and April stress degree days above the classification values were 3.05 and 3.48 times more likely to have low than high protein-to-oil ratios. July stress degree days and July Av Max Temp warmer than the classification values (Table 5) were 2.70 and 3.05 times more likely to yield low (below-median) protein-to-oil ratios than high (above-median) protein-to-oil ratios.

The three calculated weather variables that were best at separating oil concentrations for the location-year data used in this research were growing degree days, stress degree days, and the sum of minimum temperatures. These variables all represent cumulative heat units and are related to the maximum temperature (stress degree days), the average temperature (growing degree days), or the minimum temperature (sum of minimum temperatures). Other variables that were related to single weather observations (i.e., date of maximum temperature, number of days of maximum temperature, date of minimum temperature, number of days with rain, consecutive days of rain) did not delineate the location-years as well as the three cumulative variables. Of the seven best weather variables, three were from the month of September, indicating that warmer temperatures in September were associated with higher oil concentrations. This also indicated that a longer growing season would increase oil concentration because of the absence of cooler temperatures or a killing frost (20,21). The other month that was notable was April, listed twice as a month that separated the location-years. A warmer spring generally led to earlier planting, a longer growing season, and somewhat earlier maturity, with resultant warmer temperatures during seed maturation.

Protein concentration for the location years was not as easily separated as that of oil. The best predictors of protein concentration were related to temperature in August, with higher temperatures associated with a higher protein concentration of the seed. Protein concentration is positively related to increased temperatures (21), and in oilseed crops, protein is produced throughout seed development. In soybeans, low day or night temperatures during seed fill produce a lower oil concentration and a higher starch concentration but have no effect on the concentration of protein, sugar, or amino acids in the seed (22,23). Evidently, protein concentration is not as directly affected by temperature as oil concentration. This indicates that other factors, such as nitrogen availability, limit protein concentration.

It has been proposed that protein concentration is under tighter genetic control than is oil (21). Soybean lines can display a range in protein concentration. Nitrogen availability is inadequate in non-nodulating lines (24) that utilize only soil nitrogen, resulting in seed deficient in protein, whereas high-protein seed lines (25) that have protein concentrations as high as 550 g kg^{-1} accumulate more nitrogen (26) than plants from normal protein lines. Sloger *et al.* (27) reported that the average specific activity of nodules from field-grown soybeans sampled 10 times during the growing season was significantly correlated with the average air temperature but not significantly correlated with the average soil temperature. This suggests that perhaps August temperatures are related to soybean nodule activity.

Weather variables were able to separate more location-years for the total protein and oil concentration than for other seed variables. The temperature variables of growing degree days, stress degree days, average temperature, average minimum temperature, and the mean of maximum temperatures were the variables that classified most of the total protein and oil concentration values. September growing degree days and the September average temperature classified most of the location-years. Like oil concentration, the total protein and oil concentration in the seed was higher with increased temperatures in August and September. A linear relationship has been observed between the total protein and oil concentration and increased temperatures (21).

The protein-to-oil ratio was a very difficult variable to separate in this data set by the weather variables of temperature and precipitation. Soybeans preferentially synthesize two units of protein to one unit of oil. The range in protein-to-oil ratios in this data set was from 1.75 to 2.52. However, the protein-to-oil ratios could be more extreme, like those in non-nodulating lines (i.e., those with the *rj1* allele) such as Clark *rj1*, with a ratio of 1.37, and those in high-protein seed lines such as BARC-8, with a ratio of 4.07 (25). The data evaluated here and the potential range in protein-to-oil ratios in soybean seed indicate that environmental variables such as temperature and precipitation are only a few of the factors responsible. Evidently, nitrogen availability and the genetic potential of the seed to utilize available nitrogen also are important in determining the protein-to-oil ratio of the seed.

These results concur with previous research and verify that increased temperature is influential in the formation of soybean oil and the total protein and oil concentration of the seed. The results also indicate that, in this particular data set, specific months and temperature variables are more important to the oil and total protein and oil concentrations than other factors, and that the CART procedure can be used to define a value for the weather variable that can separate the yearly composition values into high and low categories. This method is an important first step in estimating seed composition before harvest, as in this data set the influence of growing season with respect to soybean seed composition could be estimated as early as June for the total protein and oil concentration. This could be accomplished by using archived company data to construct a similar model to compare protein and oil concentrations with local and regional weather. If successful, the relative composition of the seed could be estimated before harvest.

ACKNOWLEDGMENT

Special thanks to Elissa Petruzzi for her continued and dedicated technical assistance.

REFERENCES

1. Hurburgh, C.R., Jr., T.J. Brumm, J.M. Guinn, and R.A. Hartwig, Protein and Oil Patterns in U.S. and World Soybean Markets, *J. Am. Oil Chem. Soc. 67*:966–973 (1990).
2. Hurburgh, C.R., Jr., Long-Term Soybean Composition Patterns and Their Effect on Processing, *Ibid. 71*:1425–1427 (1994).
3. Yaklich, R.W., J.K. Daun, and D.R. DeClercq, Historical Trends in the Composition of Soybean, Canola-Rape, and Flax Seeds Grown in the United States and Canada, *Ibid. 79*:207–208 (2002).
4. Yaklich, R.W., B. Vinyard, M. Camp, and S. Douglass, Analysis of Seed Protein and Oil from Soybean Northern and Southern Region Uniform Tests, *Crop Sci. 42*:1504–1515 (2002).
5. Burton, J.W., Quantitative Genetics: Results Relevant to Soybean Breeding, in *Soybeans: Improvement, Production, and Uses*, edited by J.R. Wilcox, Agronomy Society of America, Crop Science Society of America, and Soil Science Society of America, Madison, WI, 1987, pp. 211–247.
6. Hartwig, E.E., and K.W. Jamison (eds.), The Uniform Soybean Tests, Southern States: 1975, USDA-ARS, Stoneville, MS, 1975.
7. Hartwig, E.E., and H. Lappas (eds.), The Uniform Soybean Tests, Southern States: 1976, USDA-ARS, Stoneville, MS, 1976.
8. Hartwig, E.E., and H. Lappas (eds.), The Uniform Soybean Tests, Southern States: 1977, USDA-ARS, Stoneville, MS, 1977.
9. Hartwig, E.E., and H. Lappas (eds.), The Uniform Soybean Tests, Southern States: 1978, USDA-ARS, Stoneville, MS, 1978.
10. Hartwig, E.E., and H. Lappas (eds.), The Uniform Soybean Tests, Southern States: 1979, USDA-ARS, Stoneville, MS, 1979.
11. Hartwig, E.E., and H. Lappas (eds.), The Uniform Soybean Tests, Southern States: 1980, USDA-ARS, Stoneville, MS, 1980.
12. Hartwig, E.E., and H. Lappas (eds.), The Uniform Soybean Tests, Southern States: 1981, USDA-ARS, Stoneville, MS, 1981.
13. Hartwig, E.E., and H. Lappas (eds.), The Uniform Soybean Tests, Southern States: 1982, USDA-ARS, Stoneville, MS, 1982.
14. Hartwig, E.E., and C.J. Edwards, Jr. (eds.), The Uniform Soybean Tests, Southern States: 1983, USDA-ARS, Stoneville, MS, 1983.
15. Chambers, J.M., and T.J. Hastie (eds.), *Statistical Models in S*, Chapman & Hall, London, 1992.
16. Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, *Classification and Regression Trees*, Wadsworth, Belmont, CA, 1984, 358 pp.
17. Zorita, E., J.P. Hughes, D.P. Lettemaier, and H. Von Storch, Stochastic Characterization of Regional Circulation Patterns for Climate Model Diagnosis and Estimation of Local Precipitation, *J. Clim. 8*:1023–1042 (1995).
18. Rodionov, S., and R. Assel, Atmospheric Teleconnection Patterns and Severity of Winters in the Laurentian Great Lakes Basin, *Atmospere-Ocean 38*:601–635 (2000).
19. Littell, R.D., G.A. Milliken, W.W. Stroup, and R.D. Wolfinger, *SAS System for Mixed Models*, Cary, NC, 1996.
20. Leffel, R.C., Planting Date and Varietal Effects on Agronomic and Seed Compositional Characters of Soybeans, Bulletin A-117, University of Maryland, College Park, 1961, 22 pp.
21. Piper, E.L., and K.J. Boote, Temperature and Cultivar Effects on Soybean Seed Oil and Protein Concentrations, *J. Am. Oil Chem. Soc. 76*:1233–1241 (1999).
22. Sato, K., and T. Ikeda, The Growth Responses of Soybean to Photoperiod and Temperature. IV. The Effects of Temperature During the Ripening Period on the Yield and Character of Seeds, *Jpn. J. Crop Sci. 48*:283–290 (1979).
23. Wolf, R.B., J.F. Cavins, R. Kleiman, and L.T. Black, Effect of Temperature on Soybean Seed Constituents: Oil, Protein, Moisture, Fatty Acids, Amino Acids, and Sugars, *J. Am. Oil Chem. Soc. 59*:230–232 (1982).
24. Williams, L.F., and D.L. Lynch, Inheritance of a Non-nodulating Character in the Soybean, *Agron. J. 46*:28–29 (1954).
25. Yaklich, R.W., β-Conglycinin and Glycinin in High-Protein Soybean Seeds, *J. Agric. Food Chem. 49*:729–735 (2001).
26. Leffel, R.C., P.B. Cregan, A.P. Bolgiano, and D.J. Thibeau, Nitrogen Metabolism of Normal and High-Seed-Protein Soybeans, *Crop Sci. 32*:747–750 (1992).
27. Sloger, C., D. Bezdicek, R. Milberg, and N. Boonkerd, Seasonal and Diurnal Variations in N_2 (C_2H_2)-Fixing Activity in Field Soybeans, in *Nitrogen Fixation by Free-Living Micro-organisms*, edited by W.D.P. Stewart, Cambridge University Press, London, 1975, pp. 271–284.

[Received May 16, 2003; accepted December 4, 2003]