A Method to Estimate Soybean Seed Protein and Oil Concentration Before Harvest

R.W. Yaklich[a],* and B.T. Vinyard[b]

USDA-ARS, [a]Soybean Genomics and Improvement Laboratory, and [b]Statistical Consulting Laboratory, Beltsville Agricultural Research Center, Beltsville, Maryland 20705

ABSTRACT: Temperature and precipitation variables during linear seed fill are known to be environmental determinants of protein and oil composition of the soybean [*Glycine max* (L.) Merr.] seed. However, the contribution of other precipitation and temperature events during the growing season and a method that would determine the precipitation and temperature variables most related to protein and oil concentration values of the seed has not been fully explored. The former was evaluated by comparing monthly temperature and precipitation variables of the growing seasons to protein and oil data for the years 1959 to 1996 from three locations listed in the Uniform Soybean Tests, Northern Region. The data set comprised locations from Maturity Groups II and III and consisted of 186 location-years. Classification and regression "tree-based" analysis were conducted to determine the month, environmental variable, and "splitting" points that correctly classified most of the 186 location-years for below-vs.-above-median protein or oil composition. The protein concentrations from the location-years were separated into these two median-boundary categories most readily by temperature variables from the months of April and August. The oil concentrations from the location-years were classified best by August and September temperature variables and precipitation in May and September. The sum of protein and oil concentrations from the location-years were best separated by August and July temperature variables and precipitation in May and July. The protein-to-oil ratios from the location-years were best separated by September precipitation and July and June temperature variables. These data demonstrate that tree-based models can use monthly temperature and precipitation variables during linear seed fill and other specific months of the crop year and relate them to the final protein and oil concentration in the seed. These results could be used by the processing industry to estimate seed composition before harvest.

Paper no. J10812 in *JAOCS 81,* 1021–1027 (November 2004).

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

In a previous paper (1) we tried to determine whether temperature and precipitation variables could be used to estimate the protein and oil concentration of soybean seed before harvest. We selected locations from the Uniform Soybean Tests, Southern Region whose protein and oil concentration data were consecutively listed for the years 1975 through 1983. We obtained the temperature and precipitation data for April through September for these locations from the U.S. National Oceanic and Atmospheric Administration (NOAA) web site and constructed a number of monthly weather variables similar to those presented in Table 1. The results showed that monthly temperature variables could be used to estimate above- and below-median protein, oil, and sum of protein and oil concentration in the seed before harvest. These results were obtained by using nine consecutive years of data from nine southern research station locations.

The understanding of how precipitation and temperature interact to determine seed composition is limited. Yet understanding how these two factors interact is a prerequisite to the challenge of managing and formulating soybean meal with adequate protein concentration for animal feed. Previous research has investigated the effect of temperature and precipitation on protein and oil synthesis during linear seed fill because this is when the protein and oil of the seed are synthesized and stored. However, this is also when the plant remobilizes C and N (used for protein and oil synthesis) formed and stored prior to linear seed fill. Since our first attempt to estimate protein and oil concentration was done with southern germplasm (determinate type growth) and nine years of data, we turned our attention to northern germplasm (indeterminate type growth) to determine whether our method of analysis would work with more years of data.

The purpose of this research is to compare temperature and precipitation values from the months of April through September to a set of yearly protein and oil concentrations obtained from different locations and growing years to determine whether temperature and precipitation values can be used to classify yearly oil and protein concentrations into above- and below-median categories.

MATERIALS AND METHODS

The Uniform Tests, Northern Region (UTs) were established to critically evaluate the best of the experimental soybean lines developed by federal and state research personnel in the United States and Canada. Lines were adapted to Maturity Groups (MGs) OO through IV. The superior soybean lines were evaluated in the UTs at many locations; grown in replicated field plots; and evaluated for yield, maturity, lodging, height, seed size, seed composition, shattering, iron chloro-

*To whom correspondence should be addressed at USDA-ARS, Soybean Genomics and Improvement Laboratory, Bldg. 006, Beltsville Agricultural Research Center, Beltsville, MD 20705. E-mail: yaklichr@ba.ars.usda.gov

TABLE 1
Calculated Monthly Weather Variables and Abbreviations Related to Protein and Oil Concentration of Soybean Seed

Monthly weather variable and abbreviation	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]/days in the month
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
Sum of maximum temperature (Sum Max Temp)	Total accumulated maximum temperature
Minimum temperature (Min Temp)	Lowest minimum temperature for month
Maximum temperature (Max Temp)	Highest maximum temperature for month
Sum of precipitation (SOP)	Total accumulated precipitation for month
Rain-free days (RFD)	Number of days without precipitation for month

sis, emergence, and disease susceptibility. Results from these tests were assembled by the cooperators and sent to a central location, edited by one of the cooperators, and published. Results from 1959 through 1996 were chosen for this study.

The number of testing locations reported in the UTs varied between MG and year. Three testing locations were chosen because they were the most consecutive and were located in some of the principal soybean production states in the United States. These locations were Ames, Iowa (42°02′N, 93°62′W), Urbana, Illinois (40°1′N, 88°2′W), and West Lafayette, Indiana (40°44′N, 86°91′W). These three states produced 42.5% of the soybeans in the United States in 2000. The superior lines were tested at all locations. Protein and oil data were not listed consecutively; therefore, none of the locations had protein and oil data for all of the selected years. The published references to the data used here (1959 to 1996) and the cooperators that gathered the data were listed in a previous publication (2).

Protein and oil concentrations were determined from a composite sample of replications of the line at a location and reported on a moisture-free basis. The methods of analysis for protein and oil concentration changed as newer methods became available and included the Kjeldahl method, NMR, IR reflectance, and NIR transmittance. The chemical analyses were performed at the National Center for Agricultural Utilization Research (changes in the name have occurred), USDA-ARS, Peoria, Illinois. Protein and oil were reported on a percent (dry wt) basis and converted to SI units. Sum of protein and oil concentration was the sum of the individual entry values for protein and oil. Protein-to-oil ratio was derived by dividing the individual entry for protein concentration by its oil concentration.

A different number of elite soybean lines were evaluated each year. The number of lines evaluated was different for each MG but was the same for each location within a MG. The data obtained from these soybean lines were used to calculate the 186 location × MG yearly means (location-years) for each seed composition variable: protein concentration, oil concentration, sum of protein and oil concentration, and protein-to-oil ratio. For ease of analysis, values for each seed composition variable were divided into two "object classes" (above-median or below-median) and each location-year was assigned to the appropriate class. Median values for each seed composition variable were: 406 g kg^{-1} for protein concentration, 215.1 g kg^{-1} for oil concentration, 621.6 g kg^{-1} for sum of protein and oil concentration, and 1.888 for protein-to-oil ratio.

Weather data for these locations were purchased from the NOAA web site (http://www.noaa.gov). The minimum and maximum daily temperature and daily precipitation were obtained for April through September. Weather variables were calculated as in Table 1.

The data were analyzed by the S-Plus software tree-based models implementation, (3) of the Classification and Regression Tree (CART) binary tree-growing algorithm developed by Breiman *et al.* (4). Tree-based analyses were used as a method to determine the months and weather variables most indicative of oil and protein concentration of the seed as presented earlier (1). A separate tree model was fitted to each calculated weather variable (Table 1) for each month. Each tree model's binary split point provided a "boundary" to use in determining how closely related location-year means on opposite sides of this "boundary" were with the two object classes (above- vs. below-median seed composition). For example, in Table 2, April maximum temperature (April Max Temp) had a split-point of 25.8°C. The data indicated that 65 location-years had an April Max Temp less than 25.8°C and 121 location-years had an April maximum temperature above the split-point value. The data also demonstrated that 47 of the 65 location-years had a protein concentration >406.0 g kg^{-1} and 75 of the 121 location-years had a protein concentration <406.0 g kg^{-1}. The tree model chose an April Max Temp of 25.8°C as the best temperature to correctly classify protein concentration for the most location-years and was able to correctly classify (above- or below-median) 122 of the 186 location-years. This procedure was used to identify weather variables and months most directly related to protein, oil, sum of protein and oil concentration, and protein-to-oil ratio of the soybean seed.

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 (SAS, Cary, NC), to indicate the strength of the relationships.

RESULTS AND DISCUSSION

The data set represented 38 yr of protein and oil concentrations from two MGs and three locations. The data set consisted of 12 yr where all six location-years were represented; 14 yr where 5 location-years were represented; 10 yr where 4 location-years were represented; and 2 yr where 2 location-years were represented. The calculation of the weather variables was performed as in Table 1.

An odds ratio provides an intuitive measure of likelihood, explained more specifically below, using observed data. The odds ratios for all weather variables (Tables 2–5) were statistically significant because each 95% confidence interval consisted of values above 1 (better than 50:50 odds) and, equivalently, by significance of Fisher's Exact Test for all 2×2 frequency tables, reported under "Number of location-years."

The median protein concentration, 406.0 g kg^{-1}, was used to define low-protein-concentration location-years, 369.4 to 406.0 g kg^{-1}, and high-protein-concentration location-years, 406.1 to 438.2 g kg^{-1}. The weather variables that best classified protein concentration into the high and low categories were primarily derived from the maximum April temperatures and minimum August temperatures (Table 2). For all the April weather variables in Table 2, the odds of low (below-median) protein concentration was significantly greater than the odds of high (above-median) protein concentration when April temperature was warmer. The opposite held for August weather variables where the low classification values (cooler temperatures) were associated with low-protein-concentration years. The odds ratios for these weather variables ranged from 2.72 to 4.26. For example, for April Max Temp >25.8°C, the odds of low- to high-protein concentration was (75:121)/(46:121) = 1.63. For April Max Temp <25.8°C, the odds of low protein concentration were (18:65)/(47:65) = 0.38. The ratio of these odds, 1.63/0.38 = 4.26, is the observed likelihood of low (below-median) protein concentration when April Max Temp >25.8°C. Alternatively stated, the odds ratios (Table 2) indicate that warmer weather in April or cooler weather in August was more likely to yield low (below-median) protein concentration than high (above-median) protein concentration. For August Sum Min Temp, the odds ratio indicates that cooler weather in August (August Sum Min Temp < 500.8 C) was 2.94 =[(65:106)/(41:106)]/[(28:80)/(52:80)] times more likely to yield low protein concentration than high protein concentration.

The median oil concentration, 215.1 g kg^{-1}, was used to define low oil concentration location-years, 174.3 to 215.1 g kg^{-1}, and high oil concentration location-years, 215.2 to 235.4 g kg^{-1}. The weather variables that best classified oil concentration into the high and low categories were derived from the maximum August and September temperatures; minimum May and August temperatures; and September precipitation (Table 3). For all weather variables in Table 3, odds ratios indicate that cooler weather in May, August, or September, or less precipitation in September was each at least 2.95 times more likely to yield

TABLE 2
Monthly Weather Variables and Classification Values That Best Separate Year Locations for Protein Concentration

Weather variable[a]	Classification value for weather variable[b]	Number of location-years for protein concentration <406 g kg^{-1}	>406 g kg^{-1}	Odds ratio[c]	Odds ratio 95% confidence interval[d]
April Max Temp	<25.8	18	47		
	>25.8	75	46	4.26[e]	(2.21, 8.20)
April Sum Max Temp	<501.4	39	67		
	>501.4	54	26	3.57	(1.93, 6.58)
April Av Max Temp	<16.7	39	67		
	>16.7	54	26	3.57	(1.93, 6.58)
April SDD	<–398.6	39	67		
	>–398.6	54	26	3.57	(1.93, 6.58)
August Sum Min Temp	<500.8	65	41	2.94[f]	(1.61, 5.38)
	>500.8	28	52		
April GDD	<21.2	45	68		
	>21.2	48	25	2.90	(1.57, 5.35)
April Av Mean Temp	<10.7	45	68		
	>10.7	48	25	2.90	(1.57, 5.35)
August Min Temp	<16.0	55	32	2.76	(1.52, 5.00)
	>16.0	38	61		
August GDD	<359.2	49	27	2.72	(1.49, 4.99)
	>359.2	44	66		

[a]See Table 1 for definitions.
[b]Temperature in Celsius.
[c]Odds ratios were calculated using the 2×2 frequency table counts reported under "Number of location-years."
[d]The exclusion of odds ratio = 1 from all confidence intervals indicates all odds ratios are statistically significant, $\alpha = 0.05$.
[e]Location-years with weather values above the classification value were "odds ratio" times more likely to have below-median protein concentrations than above median.
[f]Location-years with weather values below the classification value were "odds ratio" times more likely to have below-median protein concentrations than above-median.

TABLE 3
Monthly Weather Variables and Classification Values that Best Separate Year Locations for Oil Concentration

Weather variable[a]	Classification value for weather variable[b]	Number of location-years for oil concentration		Odds ratio[c]	Odds ratio 95% confidence interval[d]
		<215 g kg^{-1}	>215 g kg^{-1}		
August SDD	<−76.4	50	18	5.03[e]	(2.60, 9.70)
	>−76.4	42	76		
August Av Max Temp	<27.7	56	24	4.53	(2.43, 8.47)
	>27.7	36	70		
September Max Temp	<31.9	54	28	3.35	(1.83, 6.14)
	>31.9	38	66		
September Av Max Temp	<25.1	50	25	3.29	(1.78, 6.07)
	>25.1	42	69		
May Av Min Temp	<8.2	67	43	3.18	(1.72, 5.87)
	>8.2	25	51		
August Av Min Temp	<16.0	56	31	3.16	(1.73, 5.76)
	>16.0	36	63		
September SOP	<6.58	53	29	3.04	(1.67, 5.56)
	>6.58	39	65		
May MinTemp	<1.9	64	41	2.95	(1.62, 5.40)
	>1.9	28	53		
August GDD	<375.4	64	41	2.95	(1.62, 5.40)
	>375.4	28	53		

[a]See Table 1 for definitions.
[b]Temperature in Celsius and precipitation in centimeters.
[c]Odds ratios were calculated using the 2 × 2 frequency table counts reported under "Number of location-years."
[d]The exclusion of odds ratio = 1 from all confidence intervals indicates all odds ratios are statistically significant, $\alpha = 0.05$.
[e]Location-years with weather values below the classification value were "odds ratio" times more likely to have below-median oil concentrations than above-median.

TABLE 4
Monthly Weather Variables and Classification Values that Best Separate Year Locations for Sum of Protein and Oil Concentration

Weather variable[a]	Classification value for weather variable[b]	Number of location-years for total protein and oil concentration		Odds ratio[c]	Odds ratio 95% confidence interval[d]
		<621.6 g kg^{-1}	>621.6 g kg^{-1}		
August Sum Min Temp	<498	69	33	5.54[e]	(2.94, 10.45)
	>498	23	61		
August GDD	<359.2	56	20	5.76	(3.01, 11.00)
	>359.2	36	74		
August Av Mean Temp	<21.9	62	27	5.13	(2.74, 9.57)
	<21.9	30	67		
August Av Max Temp	<21.7	58	24	4.98	(2.66, 9.32)
	>21.7	34	70		
August SDD	<−68.3	58	24	4.98	(2.66, 9.32)
	>−68.3	34	70		
May SOP	<9.77	32	54		
	>9.77	60	40	2.53[f]	(1.40, 4.58)
July Max Temp	<34.2	40	61		
	>34.2	52	33	2.40	(1.33, 4.34)
July RFD	<21.5	40	60		
	>21.5	52	34	2.29	(1.27, 4.13)

[a]See Table 1 for definitions.
[b]Temperature in Celsius and precipitation in centimeters.
[c]Odds ratios were calculated using the 2 × 2 frequency table counts reported under "Number of location-years."
[d]The exclusion of odds ratio = 1 from all confidence intervals indicates all odds ratios are statistically significant, $\alpha = 0.05$.
[e]Location-years with weather values below the classification value were "odds ratio" times more likely to have below-median sum of protein and oil concentrations than above-median.
[f]Location-years with weather values above the classification value were "odds ratio" times more likely to have below-median sum of protein and oil concentrations than above-median.

low (below-median) oil concentration than high (above-median) oil concentration. The odds ratios for these weather variables ranged from 2.95 to 5.03.

The median sum of protein and oil concentration was 621.6 g kg^{-1}. The median sum of protein and oil concentration, 621.6 g kg^{-1}, was used to define low sum of protein and

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

Weather variable[a]	Classification value for weather variable[b]	Number of location-years for protein-to-oil ratio		Odds ratio[c]	Odds ratio 95% confidence interval[d]
		<1.888	>1.888		
September RFD	<22.5	69	40	3.92[e]	(2.09, 7.33)
	>22.5	24	53		
July GDD	<407	39	67		
	>407	53	27	3.79[f]	(2.03, 7.07)
July SDD	<−31.1	32	58		
	>−31.1	60	36	3.02	(1.66, 5.49)
July Av Max Temp	<28.9	32	58		
	>28.9	60	36	3.02	(1.66, 5.49)
July Av Mean Temp	<23.4	41	65		
	>23.4	51	29	3.02	(1.66, 5.49)
June Av Mean Temp	<21.5	35	58		
	>21.5	57	36	2.62	(1.45, 4.74)
June GDD	<345.8	35	58		
	>345.8	57	36	2.62	(1.45, 4.74)
July Min Temp	<11.4	42	62		
	>11.4	50	32	2.31	(1.28, 4.17)
June Av Max Temp	<27.5	35	54		
	>27.5	57	40	2.20	(1.22, 3.95)

[a]See Table 1 for definitions.
[b]Temperature in Celsius and precipitation in centimeters.
[c]Odds ratios were calculated using the 2 × 2 frequency table counts reported under "Number of location-years."
[d]The exclusion of odds ratios = 1 from all confidence intervals indicate all odds ratios are statistically significant, $\alpha = 0.05$.
[e]Location-years with weather values below the classification value were "odds ratio" times as likely to have below-median protein-to-oil ratios than above-median.
[f]Location-years with weather values above the classification value were "odds ratio" times as likely to have below-median protein to oil ratios than above-median.

oil concentration location-years, 579.7 to 621.2 g kg^{-1}, and high sum of protein and oil concentration location-years, 621.6 to 649.5 g kg^{-1}. The five weather variables most indicative of sum of protein and oil concentration were variations of August temperature (Table 4). Odds ratios indicate that cooler August temperatures were 4.98 to 5.54 times more likely to result in low (below-median) sum of protein and oil concentration than high (above-median) sum of protein and oil concentration. Precipitation in May and July and July temperature were also indicative of sum of protein and oil concentration. Greater May precipitation, more rain-free days in July, or higher temperature maxima in July were each at least 2.29 (2.29 to 2.53) times more likely to yield low (below-median) sum of protein and oil concentration than high (above-median) sum of protein and oil concentration.

The median protein-to-oil ratio, 1.888, was used to define low protein-to-oil ratio location-years, 1.658 to 1.887, and high protein-to-oil ratio location-years, 1.888 to 2.395. The ratio of protein to oil was best classified by September rain-free days (RFD) and temperatures in June and July (Table 5). The ratio of protein to oil was 3.92 times more likely to be low (below-median) than high (above-median) when September RFD were below 22.5 d. June or July temperatures warmer than the classification values (Table 5), were 2.20 to 3.79 times more likely to yield low (below-median) protein-to-oil ratio than high (above-median) protein-to-oil ratio.

The results indicate that monthly weather variables derived from the daily temperature maxima and minima and from precipitation can be used to classify a set of high and low protein and oil concentration years from soybean seed grown at different locations. The locations chosen were similar in latitude and evaluated the same MGs and soybean lines. The soybean lines differed between years as well as the methods of cultivation used at the individual locations.

Protein synthesis starts early during the linear phase of seed growth, is constant throughout seed fill (5), and is not as directly influenced by temperature as oil concentration. This is due to soybean genetics and the ability of the seed to synthesize twice as much protein as oil and to the stronger environmental effect of temperature on oil concentration than protein concentration (6). The data presented in Table 2 showed that variations of August temperature were among the best for classifying location-years for protein concentration; but the best month in this data set was April. What is anomalous about this finding is that soybeans are generally not planted in April at the three locations used in this study, nor is environment considered as strong a determinant of protein concentration as genetics (6,7). The April weather variables for classifying location-years are from an unconventional month for plant growth and therefore may be an anomaly of this particular data set. However, no other published papers have attempted to relate the months of the entire growing season to protein concentration and therefore there is an absence of data for comparison of this anomaly. Another unconventional finding was recently published by Cooper (8). He found in his long-term maximum-yield project that, in years with unusually warm May temperatures, individual lines produced replicated yield in the 6000 to 7000 kg/ha range whereas when May temperatures were normal individual lines produced yields in the 4200 to 5500 kg/ha range. What he observed was that unusually early warm spring temperatures in May resulted in soybean flowering 2 wk earlier than normal. This resulted in the soybeans entering the reproductive cycle earlier, a lengthening of the reproductive cycle, and grain fill occurring during a longer photoperiod because maturity was similar to years of more normal spring temperatures. Unlike yield, the data presented here indicate a negative effect of warm temperatures in April that were associated with decreased protein concentration. Even though we cannot compare April temperatures with some aspect of plant growth related to protein synthesis, April temperatures may be indicative of a warmer and longer growing season that may have a negative effect on protein concentration.

Weather variables derived from August and September temperatures were most influential in the classification of high and low location-years for oil concentration of the seed. Cooler August and September temperatures were associated with low, whereas warmer August and September temperatures were associated with high oil concentration. The linear phase of seed growth of MG II and III soybeans occurs during these 2 mon. Yazdi-Samadi *et al.* (5) observed that oil synthesis in soybean is initially low during the linear phase of seed growth and increases and reaches its maximal level as the seed approaches physiological maturity. Others have observed that high temperatures during the linear phase of seed growth were associated with increased oil concentration (9) whereas delayed planting decreased oil concentration of the seed because of decreased temperatures during maturation (10). Piper and Boote (6) evaluated regression models of oil concentrations vs. temperature and found that oil concentration increased with temperature and approached a maximum at a mean of 28°C. Others (7,11) have shown that increased temperature increases the oil concentration of the seed. However, none of the methods in these publications attempted to use monthly weather variables to classify a set of oil concentration data. The results presented here agree with those presented by others but more importantly show that both August and September temperatures can be used to categorize a set of yearly oil concentration data of the seed. The results also suggest that September rainfall is important to oil concentration and is in agreement with Dornbos and Mullen (12) that drought decreases oil concentration. The minimum temperature in May also classified the location-years and may be related to the planting date. It is known that delayed planting favors protein synthesis (10). It is not clear how temperature affects the synthesis and concentration of oil in the soybean seed (13).

Weather variables classified more of the location-years for sum of protein and oil concentration than the other composition classes. Sum of protein and oil concentration of the seed represents the amount of storage reserves available for industrial use. Hurburgh (14) showed that the higher the sum of protein and oil concentration, the higher the protein concentration of the meal. Unlike oil concentration, where the best weather variables were from the months of August and September, the five best weather variables for classifying sum of protein and oil concentration were from the month of August. The data presented here and by other authors (6,7) indicate that temperature during the linear phase of seed growth is positively related to sum of protein and oil concentration and, as presented here, August temperature is the most influential in the formation of sum of protein and oil concentration in this data set.

The protein-to-oil ratio represents the relationship of the units of protein relative to those of oil. Values above 2.0 indicate more units of protein synthesized relative to oil than the theoretical 2:1, and values below represent the opposite. Two examples taken from the location-years for a protein-to-oil ratio of 2.00 are 433.4/216.0 g/kg^{-1} and 415.2/207.1 g/kg^{-1} protein concentration divided by oil concentration, respectively, which illustrate that it was the amount synthesized that determined the ratio of the protein and oil concentration. There were only 39 location-years when the protein-to-oil ratio was equal to or greater than 2.00, indicating that the data set consisted mainly of location-years with more oil units than the theoretical relationship of 2 units of protein to 1 of oil. The September days without precipitation would indicate that more days without precipitation favor a higher protein-to-oil ratio. This would be similar to the observations of Dornbos and Mullens (12) who observed that drought decreased oil concentration. The remaining weather variables were from the months of June and July, and all indicated that temperatures below the classification values were associated with high protein-to-oil ratios whereas temperatures above the classification values were associated with low protein-to-oil ratios. For MG II and III soybean, June would represent part of the vegetative growth stage. Since the growth type of these soybean lines would be indeterminate, growth would continue during July and flowering would ensue with cell division in the embryo. Why June and July temperatures classified the protein-to-oil ratio best invites speculation. However, cooler temperatures would favor protein synthesis and a normal to high protein-to-oil ratio (6,11).

Comparing the results presented here and from our previous paper (1) using southern germplasm shows that protein concentration was estimated by August growing degree days (GDD) in both studies; sum of protein and oil concentration was estimated by August stress degree days (SDD), August GDD, and August Av Max Temp; and protein-to-oil ratio was estimated by July SDD. September temperature variables estimated oil concentration in both studies but with different variables.

Finally, the data presented suggest a specific month that may be important to managers in the processing industry and soybean breeders. Since sum of protein and oil concentration is directly related to the protein concentration of soybean meal (14), August temperatures could be used by the processing industry to formulate models that would estimate sum of protein and oil concentration prior to harvest, allowing a strategy for the harvested crop and its use in formulating and marketing soybean meal. Further, August temperatures could be used by soybean breeders to search for germplasm whose sum of protein and oil concentration is insensitive to a range in August temperatures.

ACKNOWLEDGMENT

Special thanks to Elissa Petruzzi for her dedicated technical assistance.

REFERENCES

1. Yaklich, R.W., and B.T. Vinyard, Estimating Soybean Seed Protein and Oil Concentration Before Harvest, *J. Am. Oil Chem. Soc. 81*:189–194 (2004).
2. Yaklich, R.W., B.T. 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).
3. Chambers, J.M., and T.J. Hastie (eds.), *Statistical Models in S,* Chapman & Hall, London, 1992.
4. Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, *Classification and Regression Trees*, Wadsworth, Belmont, California, 1984, 358 pp.
5. Yazdi-Samadi, B., R.W. Rinne, and R.D. Seif, Components of Developing Soybean Seeds: Oil, Protein, Sugars, Starch, Organic Acids, and Amino Acids, *Agron. J. 69*:481–486 (1977).
6. 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).
7. Gibson, L. R., and R. E. Mullen, Soybean Seed Composition Under High Day and Night Growth Temperatures, *Ibid. 73*:733–737 (1996).
8. Cooper, R.L., A Delayed Flowering Barrier to Higher Soybean Yields, *Field Crops Res. 82*:27-35 (2003).
9. Kane, M.V., C.C. Steele, L.J. Grabau, C.T. MacKown, and D.F. Hildebrand, Early-Maturing Soybean Cropping System: III. Protein and Oil Contents and Oil Composition, *Agron. J. 89*:464–469 (1997).
10. 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.
11. Howell, R.W., and J.L. Cartter, Physiological Factors Affecting Composition of Soybeans: II. Response of Oil and Other Constituents of Soybeans to Temperature Under Controlled Conditions, *Agron. J. 50*:664–667 (1958).
12. Dornbos, D.L., Jr., and R.E. Mullen, Soybean Seed Protein and Oil Contents and Fatty Acid Composition Adjustments by Drought and Temperature, *J. Am. Oil Chem. Soc. 69*:228–231 (1992).
13. Egli, D.B., *Seed Biology and the Yield of Grain Crops*, CAB International, New York, 1998, 178 pp.
14. Hurburgh, C.R., Jr., Long-Term Soybean Composition Patterns and Their Effect on Processing, *J. Am. Oil Chem. Soc. 71*:1425–1427 (1994).

[Received March 1, 2004; accepted November 8, 2004]