United States
Department of
Agriculture

Statistical
Reporting
Service

Statistical
Research
Division

SRS Staff Report
Number YRB-85-04

June 1985

Evaluation of Sorgf, A Sorghum Crop Simulation Model

William C. Iwig

Benjamin F. Klugh, Jr.

EVALUATION OF SORGF, A SORGHUM CROP SIMULATION MODEL. By William C. Iwig and Benjamin F. Klugh, Statistical Research Division, Statistical Reporting Service, U.S. Department of Agriculture, June 1985. SRS Staff Report No. YRB-85-04.

ABSTRACT

This study examines the performance of the Texas A&M sorghum crop simulation model, SORGF, in randomly selected plots located across Kansas. Many of the model inputs were estimated rather than measured at site. Performance of different feedback options were also examined. Results indicate a large negative bias in the SORGF yield values, most likely due to internal model problems and not to estimation of model inputs.

KEYWORDS

Large area yield forecasting, crop simulation models, model evaluation, feedback, sorghum, SORGF.

ACKNOWLEDGMENTS

Several people made important contributions to the project. Thanks to the office personnel and enumerator staff of the Kansas State Statistical Office for their data collection efforts; to Phil Shideler, Mrs. Jerry Fast, Dr. Dean Bark, and Dr. Dan Tarpley for their cooperation in obtaining the weather data; to Greg Larsen, Dr. Dick Vanderlip, Dr. Gerald Arkin, and Fred Warren for their useful manuscript reviews and discussions; and to Dave Aune for his assistance in the project management and data summarization. Special thanks to Brenda Ganger for her excellent typing of the many manuscript drafts and to Bessie Johnson and Sharon Smith for typing the final printed version.

Washington, DC June 1985

CONTENTS

SUMMARY

Final head weights estimated by the sorghum crop simulation model, SORGF, had a large negative bias in this application. This bias was most likely due to internal model problems. This version of the model would be unacceptable for Agency use.

This study examines the performance of SORGF using estimated, rather than measured, values for many of the input variables. These results were obtained from 79 plots randomly located across Kansas during the 1981 growing season. Modeled final yield averaged 954 kilograms/hectare (kg/ha), or 20 percent, less than the measured yield. Four feedback options averaged 520 to 1219 kg/ha less than measured yield. Feedback refers to using crop measurements during the growing season to adjust modeled crop growth to "match" observed crop growth for the field plot. Although the modeled final yield values were, on the average, too low, large overestimates also occurred for some plots causing large root mean square errors, large relative differences, and moderate correlations between modeled and measured yields for all SORGF options. Correlations ranged from 0.56 to 0.67. Modeled growth and development may have better matched the actual crop growth at each of the 79 plots if measured inputs had been used. In an operational program, where the model is run for many randomly located plots in several States, the cost and time involved to obtain exact inputs would be prohibitive.

EVALUATION OF SORGF, A SORGHUM CROP SIMULATION MODEL

William C. Iwig and Benjamin F. Klugh

INTRODUCTION

This study evaluates the performance of SORGF, a plant simulation model for sorghum, in estimating final grain weight per head in a Statistical Reporting Serivce (SRS) objective yield program. SORGF was developed by G. F. Arkin, R. L. Vanderlip, and J. T. Ritchie in 1975. The model has been modified several times. The version used in this test was completed in 1978. The study was conducted in 79 objective yield sample fields in Kansas in 1981.

A plant process simulation model is a computer program that simulates the daily growth and development of a crop. Daily values are generated by the model for such crop characteristics as leaf number, leaf area, root weight, stalk weight, and grain weight. This type of model generally requires soil data, daily weather data, variety information, and crop management information as inputs. It is quite expensive to obtain exact values at specific locations, such as objective yield plots, for many of these variables. An alternative is to estimate the input values using more cost-efficient data sources.

The first objective of this study is to explore the performance of SORGF when cost-efficient estimates are used as input to the model. The second objective of this study is to evaluate the effect of "feedback" on model performance. "Feedback" refers to using actual plant measurements made during the growing season in field plots to adjust the modeled growth to "match" the field observed growth.

The study is not intended to be a strict validation of the SORGF model, but an evaluation of this type of application.

DATA COLLECTION

Basic data were collected and/or estimated to execute entire crop season runs of the SORGF model. Additional "feedback" data were collected to adjust internal modeled plant variables to improve final model outputs. These data were collected in sorghum objective yield fields in Kansas in 1981. Each sample consisted of three randomly located units (plots). These units were generally in the same quarter of the field. Each unit was three rows wide and nine feet long. Although 141 samples were originally selected, final analysis included only 79 samples. Major reasons for lost samples were that fields were not planted to sorghum, some farmers refused to participate, and data were not complete for some samples. Enumerators first visited each field in late June. At that time the farm operator was interviewed and the sample units were located.

Basic Data for Executing the Model

The data needed to execute the model are listed in table 1. Values for the first seven variables were measured or enumerated for each sample. Weather data were obtained for March 28 - November 13 inclusive, for a total of 231 days. This time frame may be several months before planting. The planting date and planting depth were obtained from the farm operator. Plant population and row spacing were based on measurements obtained in the three units of each sample.

Table 1: SORGF[1/] initial input variables

Variable	Description
1. KI	Number of days of weather data
2. ROSPZ	Row spacing (cm)
3. P	Plant population (plants/hectare)
4. ALT	Latitude (degrees)
5. SDEPTH	Planting depth (cm)
6. MO, ND, IYR	First day of weather data
7. MO, ND, IYR	Planting date
8. SW	Initial plant available soil water in root zone (cm)
9. UL	Maximum plant available soil water in root zone (cm)
10. NTL	Total number of leaves on modeled plant
11. XMAX(J)	Maximum area of leaf J, J=1, ..., NTL
12. RAIN(I)	Precipitation on day I (in)
13. TEMPMX(I)	Maximum temperature on day I (^oF)
14. TEMPMN(I)	Minimum temperature on day I (^oF)
15. SOLRAD(I)	Solar radiation on day I (langleys)
16. IDAYFB	Julian feedback date
17. IRR	Irrigation code (1 or 2)

1/ The version of SORGF listed in (9) in the references section was employed for this study with a few modifications recommended by the model developers (Arkin and Vanderlip, personal communication). These modifications are listed in appendix A.

Input variables 8-15 of table 1 were estimated by use of various data sources and methods because direct measurements of these variables at each sample location would be too costly. The procedures used in estimating these variables are described in pages 3 and 4. If the reader is not interested in these details, skip to the Plot Data for Feedback and Final Yield Section on page 4.

The soil water values SW and UL are calculated in the Palmer Index program by the National Oceanic and Atmospheric Administration

(NOAA) for each Crop Reporting District (CRD). These values for the week ending March 28, 1981, were obtained from Mr. Merle Brown, Kansas Agricultural Experiment Station (KAES) Weather Data Library, Kansas State University (KSU). Maximum available soil water values (UL) for CRD's ranged from 17.8 cm to 27.9 cm. Initial values (SW) ranged from 13 percent to 73 percent of the maximum. All samples within the same CRD were assigned the same initial and maximum soil water levels.

The total number of leaves per plant (NTL) and the maximum area per leaf (XMAX(J)) were assigned from information in table 2. Varieties planted in the sample fields were assigned to one of the four maturity classes based on their average number of days from planting to half-bloom as reported in the 1980 Kansas Sorghum Performance Tests publication. The maturity class was estimated for any variety planted in a sample field but not included in the performance tests. Four NTL values were used that represent four different maturity classes. Total leaf areas were assigned to each class (Bunck, 1977) and were apportioned to individual leaves on a percentage basis after the percentages from available leaf area data were reviewed (Iwig, 1984).

Table 2: Total leaf number and total plant leaf area for maturity classes

Class	Days to half-bloom	Total leaves (NTL)	Total leaf area (cm^2)
1	<= 60	17	2,600
2	61 - 67	18	2,800
3	68 - 74	20	3,600
4	>= 75	22	4,200

Daily precipitation and temperature data were obtained from approximately 100 National Weather Service (NWS) stations across Kansas, normally one per county, courtesy of Mr. Phil Shideler, NWS Meteorologist-in-Charge, Topeka, KS. A nonlinear distance weighting function described by Barnes (1964) was used to interpolate these data to the sites. Daily solar radiation estimates were obtained from NOAA geostationary operational environmental satellite (GOES) data (Tarpley, 1979) provided through the Agricultural and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) program. These values were on a grid basis with each reading centered on a whole degree latitude-longitude intersection. All sorghum samples located within the area represented by a reading were assigned the same daily solar radiation values. Values for 13 days of the GOES data were missing. Solar radiation data from nine Kansas stations of the cooperative USDA and NOAA Touch-Tone system were interpolated to sample locations for those dates. These data are available on this system because of data collection and reporting procedures initiated by Dr. Dean Bark, KAES Climatologist, KSU.

The input variable, IDAYFB, was assigned the Julian value of the feedback date when the model was run in a feedback mode. Otherwise, the variable had a value of zero.

The irrigation code, IRR, was assigned a value of 1 if the field was to be irrigated at least once during the growing season as indicated by the farm operator. Since irrigation facilities were available for these fields, it was assumed that irrigation would be applied appropriately to avoid water stress to the crop. Consequently, the model code was edited so that the water stress coefficient, WATSCO, was held at a constant value of 1 (no water stress) for all samples with an IRR value of 1.

Two soil variables that are not included as inputs but are assigned within the model are U, the upper limit of stage 1 cumulative evaporation, and CONA, a coefficient in the cumulative evaporation equation. Appropriate values for each of these parameters depend on the particular soil properties at the sample location. However, limited resources prohibited measuring these variables. Since the model is fairly insensitive to the variables, constant values of 1.0 cm for U and 0.35 cm day $^{-1/2}$ for CONA were assigned to all samples (Arkin, personal communication).

Obviously, many of the input values are only estimates of the true values at the site of each sample. Some inputs are probably more representative of the average conditions for the field, for the area represented by a latitude-longitude intersection, or for a CRD. Plant population and row spacing were the only inputs measured in the three units of each sample. It was assumed that the soil and weather estimated inputs had a negligible bias over all samples. Any possible bias in the varietal inputs (NTL and XMAX(J)) would be addressed through feedback.

Plot Data for Feedback and Final Yield

Crop simulation models do not always provide accurate simulations of crop growth and development. Feedback refers to the process of utilizing crop data collected during the growing season to adjust the simulated crop growth and development to better match field observations. The purpose of feedback is to improve model accuracy and performance. SORGF was developed with this capability. The approach is more fully described in appendix B.

When the feedback option is executed, it is desirable to have observed values for each of the 17 feedback variables listed in appendix B. This was not possible given our data collection constraints in this study. Only half-bloom date, leaf counts, leaf area, and head weight data were collected for each sample. These data were used in four different combinations to determine their effect on model yield in different feedback applications.

Data were collected in all sample plots as follows:

1. Half-bloom date

   Half-bloom date was estimated from the enumerator count of plants that had bloomed in a 2-row by 3-foot portion of each unit visited at 4-day intervals around flowering.

2. Number of green leaves per plant

   Starting August 1, the number of green leaves were counted monthly on five designated plants located adjacent to unit 1.

3. Area of each individual leaf

   On the first visit after the flag leaf was fully emerged on all five plants counted in (2), the length and width of each green leaf was measured. Individual leaf areas were estimated by use of the following relationship (Stickler et al., 1961):

   $$A = 0.75 * L * W$$

   where:

   A = area of individual leaf (cm^2),
   L = length of leaf (cm), and
   W = width of leaf (cm).

4. Current dry head weight per plant

   Starting on the first monthly visit after half-bloom, heads were clipped from five different plants adjacent to unit 1 and sent to the laboratory for drying and weighing.

5. Final dry head weight per plant

   Final yield measurements were also collected for each sample immediately prior to farmer harvest. All heads in the three units were clipped and weighed in the field. All heads in the 2-row by 3-foot subunits were sent to the laboratory to determine the threshing fraction and moisture content. This crop cutting yield will be referred to in the report as the measured yield. It has an associated sampling error since it was calculated by use of data collected from three separate units.

METHODS

The SORGF model was run for each sample without use of any feedback data. The model was also run by use of four different combinations of feedback variables.

Each combination was used in a separate set of model runs over all samples. In the first combination of feedback data, the NL option, only the number of green leaves per plant on the date the leaf area data were collected was used as feedback. In the second option, NLLA, the leaf area data were used in addition to the number of leaves. In the third, the HBD option, only the half-bloom date was adjusted. In the fourth option, ALL, the half-bloom date, the feedback leaf number, leaf area, and head weight data were used. Leaf and head measurements were from the first date that head weights were obtained. All other feedback variables that could not be estimated from the feedback data were assigned the internal modeled value as of the feedback date. These four feedback options are discussed in more detail in appendix B. These modeled yields were compared with the measured yields to evaluate performance. Modeled values of the feedback variables were compared with measured values to help interpret the yield results.

We assumed the estimated yield and a SORGF model yield for a particular sample represented the yield for the same area of the field and consequently were comparable. The Dunnett's multiple comparison test (Steel and Torrie, 1960) was used to test whether the estimated yield differed significantly from any of the SORGF final yields. In addition, six statistics, originally suggested by Wilson et al (1980) were calculated for comparing each SORGF option against the measured data.

1. Estimated bias = $\bar{d} = \dfrac{\Sigma d_i}{n}$

   where: d_i = difference between modeled and measured yield for sample i, and

   n = number of samples.

   The bias is the average error and indicates whether the model tends to underestimate or overestimate yield.

2. $SD(d) = \sqrt{\dfrac{\Sigma (d_i - \bar{d})^2}{n}}$

   The standard deviation of the difference indicates the consistency of the modeled yield values compared with the measured values. A small standard deviation implies that the difference between the modeled and measured values is fairly constant with little variability. It can be used to test for significant bias.

3. $RMSE = \sqrt{\dfrac{\Sigma d_i^2}{n}}$

   $$= \sqrt{\frac{\Sigma (d_i - \bar{d})^2}{n} + (\bar{d})^2} = \sqrt{SD\,(d)^2 + (\bar{d})^2}$$

   The root mean square error (RMSE) has two components, the variance of the differences ($SD\,(d)^2$) and the square of the bias $(\bar{d})^2$. A relatively large RMSE would be due to highly variable differences, a large bias, or both. If one of the SORGF options was used on an independent sample, it is about 68 percent probable that the modeled yield would be within one RMSE of the measured yield and about 95 percent probable that the modeled yield would be within two RMSE's.

4. r = Pearson correlation coefficient ($-1 < r < 1$)

   The correlation coefficient, r, is a measure of the degree of linear association between the modeled and measured yields.

5. Maximum $|rd_i|$

   where: $rd_i = 100\ (d_i/Y_i)$, and
   Y_i = measured yield for sample i.

   The maximum absolute relative difference (Max $|rd_i|$) indicates the maximum difference between modeled and measured yields as a percentage of the measured value.

6. % $|rd_i| < 10\%$

   The percentage of the absolute relative differences that is less than 10 percent indicates how often the modeled yield was within 10 percent of the measured yield over all samples.

RESULTS AND DISCUSSION

The mean yield, the largest sample yield, the smallest sample yield, and the six evaluation statistics from the 79 samples are presented in table 3. Measured and modeled yield values are at 15.5 percent moisture. Data in table 4 compare simulated and measured values of the four feedback variables over all 79 samples. This information can be used to help explain yield differences obtained from the different SORGF options.

The Dunnett's multiple comparison test with $\alpha = 0.05$ was used to test for any differences between measured and estimated SORGF yields. Any absolute difference between the estimated yield and the yield from a SORGF option that exceeded 820 kg/ha was significant. The mean yields of all SORGF options except HBD tested to be significantly different from the mean estimated yield. At first it may seem that using only the observed half-bloom date as feedback would produce "satisfactory" yield estimates by SORGF. However, a further review of the results on page 8 indicate that this option provided the right answer (although it still had a relatively large bias) for the wrong reasons. A crop simulation model needs to produce accurate yield estimates, but should also accurately simulate the crop growth and development during the growing season. A review of the data in table 4 indicates that the HBD option does not do an "adequate" job of simulating crop growth. The data also help explain the yield differences among the SORGF options presented in table 3.

Table 3: Model evaluation statistics for SORGF with and without feedback

Statistics	SORGF-no feedback	SORGF feedback options: NL	NLLA	HBD	ALL	Estimated yield
Mean Yield (kg/ha) 1/	3752	3563	3488	4187	3491	4707
Maximum Yield	8197	8138	7788	8368	7729	8590
Minimum Yield	605	462	509	445	476	619
Bias	-954	-1144	-1219	-520	-1216	
SD (d)	1664	1631	1664	1754	1787	
RMSE	1918	1992	2063	1830	2161	
r	.65	.67	.65	.64	.56	
Max |rd|	403	403	407	338	170	
% |rd| < 10%	25	16	15	32	15	

1/ 1 bu/acre = 51.6 kg/ha.

Table 4: Comparison of simulated and measured feedback variables for the 79 samples

Variable	Simulated: Mean	Simulated: S E 1/	Measured: Mean	Measured: S E	Paired difference = simulated-measured: Min	Max	Mean	Paired t for $H_0: \bar{d}=0$
Leaf Number	11	0	8.41	.23	-2	8	2.59	11.38**
Total Plant Leaf Area (cm^2)	3045	-	2491	91.43	-1691	2281	554	5.88**
Head Weight (g)	21.8	1.95	23.7	1.58	-48.4	89.7	-1.87	<1
Half-Bloom Date	230 2/	1.39	223	1.22	-13	41	6.69	5.91**

** = Paired t-statistic significant at $\alpha = 0.01$.
\- = Not applicable.
1/ S E = Standard Error of Mean.
2/ Reference Julian date: August 16 = 230.

The NL option produced a smaller yield than the no feedback option (table 3). This lower yield was primarily due to a reduction in number of leaves through feedback (table 4). After the flag leaf is fully grown, every modeled plant has 11 leaves. The observed mean number of leaves per plant was 8.41. A smaller number of leaves causes less total plant leaf area available for photosynthesis during grain fill, which tends to cause smaller yields.

The NLLA option produced a slightly smaller yield than the NL option (table 3). The leaf area values used as feedback were slightly smaller than the modeled values, which are actually model inputs (table 1). Table 4 indicates that the measured total plant leaf area on the feedback date averaged over 550cm^2 less than the model value. Most of this difference was due to the difference in number of leaves.

The HBD option produced a substantially higher average yield than the no feedback option (table 3). Evidently this situation resulted from the mean measured half-bloom date's being about 1 week earlier than the simulated date (table 4). The earlier half-bloom date probably provided better grain filling weather and thus higher yields. The simulated half-bloom date is very dependent on NTL, the input total number of leaves per plant (table 1). A variety that produces a larger number of leaves takes longer to reach half-bloom. A separate analysis (Iwig, 1984) indicated that some varieties that were assigned NTL values of 20 and 22 leaves probably produced only 18-19 leaves. Such misclassification would help explain the late simulated half-bloom date.

The mean yield for the ALL option is much lower than the HBD yield (table 3), even though both options have used the measured half-bloom date as feedback. A portion of the reduction is due to using smaller leaf numbers and reduced leaf area as feedback, as indicated by the NLLA results. The remaining yield reduction is evidently due to the head weight feedback variable. The data in table 4 indicate that the mean measured head weight was slightly larger than the simulated mean. However, the simulated mean is from the no feedback option where the half-bloom date was a week late. The mean simulated head weight by the HBD option, with the corrected half-bloom date, would be 30-35 grams.

Comparing the mean measured head weight of 23.7g against a mean simulated head weight of 30-35g shows why using head weight as feedback in the ALL option tends to reduce yields as compared with the HBD option. Simulated head growth is too rapid during the early part of grain fill, prior to the ALL option feedback, and too slow or for too short a period after the feedback date.

The large maximum and minimum differences between simulated and measured head weights indicated in table 4 are mainly due to the differences in the half-bloom date. However, these results could also be partly due to imprecise sample level estimates of mean head weights. The data generated a large variance among the five heads measured from each sample, yielding a sample coefficient of variation (CV) of 56 percent compared with CV's of 26 percent and 13 percent for total plant leaf area and number of leaves per plant, respectively. With this much variability in the measurement data, the estimated mean sample level head weights could easily be in error by relatively large amounts. Assuming a normal distribution, there would be a 32 percent chance of the mean head weight being in error by more than 56 percent.

Our review of the other five summary statistics would suggest that this application of the model tends to produce low yield values as indicated by the bias terms. Biases for the 20 irrigated and 59 nonirrigated fields were not significantly different within each option, so that the bias values in table 3 are representative of all fields regardless of irrigation status. Since the model does not account for such yield-reducing factors as disease, insects, and weeds, a positive bias is expected, at least in the no feedback mode. It is unlikely that using estimated inputs instead of actual measured inputs could cause the large negative bias in the model yields. Internal model problems are indicated. At least four aspects of the model have been identified to be in need of additional experimental work (Arkin, personal communication). These problems involve the calculation of canopy light interception, the conversion of photosynthetically active radiation to the daily increase in total plant dry weight, the calculation of potential evaporation (EO), and the partitioning of dry weight to plant parts, especially to grain. Each of these four aspects of the model has been modified since the original version of the model was validated (Vanderlip and Arkin, 1977). Although that earlier study indicated some large errors in grain weight per plant, there was no consistent bias. Since the current version has large negative bias, it is likely these aspects are contributing factors.

The RMSE values are relatively large, indicating that large errors in final yield could be expected from any of the SORGF options for an individual sample. The large RMSE values are due to large biases and large standard deviations of the differences. The moderate correlation coefficients, 0.56 to 0.67, between the modeled and observed yield values are also a product of the large variance of the differences. A nearly constant difference would provide correlations near 1.0. The maximum $|rd|$ values indicate that the model without feedback was in error by over 400 percent, due to overestimates. The maximum $|rd|$ for an underestimate would be 100. The model did not properly account for the low yield conditions at some sample locations. One possible reason is that some yield-reducing factor such as disease, insects, or weeds, which is not addressed by the model, is dramatically affecting some sites. If such a factor is present, it should display itself through the feedback variables. The lower maximum $|rd|$ values for the ALL option of 170 indicates that, when observed head weights are used as feedback, the model does a better job of estimating yield at the low yield sites. Reviewing the $|rd|$ values over all samples revealed that less than 35 percent of the model yield estimates were within 10 percent of the measured value. Therefore, relatively large errors (greater than 10 percent) occured for a majority of the samples with any feedback option.

Internal model problems certainly contributed to the large RMSE, moderate correlation, and large $|rd|$ values. Their unsatisfactory values are also partially attributed to the use of estimated inputs rather than site-specific measured inputs. Variation about the true yield has to be expected if exact inputs are not used. The estimated input likely to have the greatest effect on model results is daily precipitation. Rainfall varies greatly over a geographic area, and it is unlikely the interpolated estimates will account for this variability. The other meteorological estimates, assuming they are unbiased, should have less of an impact on model results because they have less spatial variability than precipitation. The CRD estimates of UL could also be critical, as they can vary widely within a CRD, within a county, and even within a field. Since the water balance was started approximately 2 months before planting, the effect of not knowing the exact initial SW value was reduced, but may still have been significant in water limited growing conditions (Larsen, 1981). The effect of using estimated values of total number of leaves and individual leaf area on the simulated final yield is probably small. When any errors in the values of these two variables were corrected through feedback in the ALL option, unsatisfactory values of RMSE, correlation, and $|rd|$ were still obtained.

RECOMMENDATIONS AND CONCLUSIONS

We recommend that:

(1) this version of the SORGF model not be implemented in any operational SRS program,

(2) SRS should continue to monitor model development and progress, and

(3) the modeler would repeat the analysis of the 1981 Kansas data with each new version of the model and demonstrate reliable performance before SRS invests any further efforts in this area.

Final head weights estimated by SORGF in this application had a large negative bias. Attempts to employ "feedback" data to improve model estimates were unsuccessful. This bias is most likely due to internal model problems.

Application of any crop simulation model by SRS would involve use of estimated values of model inputs. The cost and time needed to obtain measured values for all inputs at all Objective Yield sites is prohibitive. Accurate and cost-effecient estimates of model inputs are needed as well as a "good" model.

Although sufficient data were not available in this study to investigate the adequacy of each estimated input, it is likely that daily precipitation and maximum available water (UL) have the greatest effect on yield errors due to their variability over a geographic area.

A major benefit of this study was gaining experience in applying the SORGF model in a feedback mode. The half-bloom date is an important feedback item since the modeled date can be in error by 2 weeks or more. This variable would be used to adjust the modeled phenology, or crop stage, if needed. Measured head weights, obtained from a sufficient sample size, should also be an important feedback item.

REFERENCES

(1) Arkin, G.F., R.L. Vanderlip, and J.T. Ritchie, "A Dynamic Grain Sorghum Growth Model," Transactions of the American Society of Agricultural Engineers 19:622-626, 1975.

(2) Barnes, S.L., "A Technique for Maximizing Details in Numerical Weather Map Analysis," Journal of Applied Meteorology 3:396-409, 1964.

(3) Bunck, J.H., "Effect of Planting Date on Leaf Number and Total Leaf Area of Hybrid Grain Sorghum," Master of Science thesis, Department of Agronomy, Kansas State Univ., 1977.

(4) Cochran, William G., Sampling Techniques, 2nd ed., New York: John Wiley & Sons, Inc., 1963.

(5) Iwig, William C., "Input Leaf Data for the 1981 SORGF Study", unpublished internal memo, Sept. 1984.

(6) Kansas Agricultural Experiment Station, "1980 Kansas Sorghum Performance Tests," Kansas State Univ., 1980.

(7) Larsen, Greg A., "Progress Report on the Evaluation of Plant Growth Models," Staff Report AGESS810716, U.S. Department of Agriculture, Statistical Reporting Service, July 1981.

(8) LeDuc, Sharon, and Tom Hodges, "Large Area Yield Estimates Derived from Plant Simulation Models," JSC-18500, AgRISTARS Mini Symposium, Johnson Space Center, Dec. 12, 1982.

(9) Maas, Stephen J. and Gerald F. Arkin, "User's Guide to SORGF: A Dynamic Grain Sorghum Growth Model with Feedback Capacity," Texas Agricultural Experiment Station, No. 78-1, 1978.

(10) Schaffer, James A., "The Effect of Planting Date and Environment on the Phenology and Modeling of Grain Sorghum, Sorghum bicolor (L.) Moench" Ph.D. dissertation, Department of Agronomy, Kansas State University, 1978.

(11) Sivakumar, M.V.K., "Evaluation of SORGF Subroutines: A Consultancy Report," Blackland Conservation Research Center, Temple, TX, 1981.

(12) Steel, Robert G.D., and James H. Torrie, Principles and Procedures of Statistics, New York: McGraw-Hill Book Company, Inc., 1960.

(13) Stickler, F.C., S. Wearden, and A. Pauli, "Leaf Area Determination in Grain Sorghum," Agronomy Journal 53:187-188, 1961.

(14) Strand, Bruce W., "Spatial Scale of Crop-Yield Models," Staff Report AGESS810320, U.S. Department of Agriculture, Statistical Reporting Service, Mar. 1981.

(15) Tanner, C.B., and W.A. Jury, "Estimating Evaporation and Transpiration from a Row Crop During Incomplete Cover," Agronomy Journal 68:239-243, 1976.

(16) Tarpley, J.D., "Estimating Incident Solar Radiation at the Surface from Geostationary Satellite Data," Journal of Applied Meteorology, 18:1172-1181, 1979.

(17) Vanderlip, R.L., and G.F. Arkin, "Simulating Accumulation and Distribution of Dry Matter in Grain Sorghum," Agronomy Journal 69:917-923, 1977.

(18) Wilson, Wendell W., Thomas L. Barnett, Sharon K. LeDuc, and Fred B. Warren, "Crop Yield Test and Evaluation Criteria," AgRISTARS Yield Modeling Development Project, Document YMD-1-1-2 (80-2.1), 1980.

APPENDIX A: Model Modifications

The version of the SORGF model listed in the User's Guide was used for this study with the following modifications.

1. The subroutine DRWT (Sivakumar, 1981) was substituted for the PHOTO and SYNTH subroutines.

2. The calculation of the potential solar radiation reaching the soil surface, variable RO, was modified so that it is not specific to 30° N latitude (Sivakumar, 1981).

3. The interval from half-bloom to physiological maturity was defined to be 741 heat units (Schaffer, 1978). In the original model version, this interval was a function of the interval from emergence to half-bloom. The heat units accumulated in 1 day were defined as:

   $$HU = \frac{MAXT + MINT}{2} - 1$$

   where:

   MAXT = daily maximum temperature (C°), if MAXT > 38°, MAXT = 38°, and

   MINT = daily minimum temperature (°C), if MINT < 1°, MINT = 1°.

4. A condition that floral initiation, the beginning of stage 2, cannot occur until the fifth leaf is fully grown was added to the phenology calculations.

5. The albedo for bare soil was changed from 0.15 to 0.10.

6. The constant, a, used in the following calculation of potential evaporation, EO, was changed from 1.28 to 1.60.

   $$EO = a \frac{s}{(s + \gamma)} HO$$

   where:

   s = slope of saturation vapor pressure curve at mean air temperature,

   γ = psychrometer constant, and

   HO = net radiation above the canopy.

7. The following equation was used to calculate the percent light transmission, LITRAN, regardless of the daily leaf area index value.

   LITRAN = 100. *EXP(X2 * DLAI)

where:

DLAI = leaf area index for current day,

X2 = (.0091 * R) - 0.75, and

R = row spacing (inches).

APPENDIX B: Feedback Options

In the SORGF feedback process, the simulation is started on some given feedback date after emergence. Variables describing the growth status and stage of the crop on that date are required as input, since the model does not simulate the early growth and development as in a regular nonfeedback simulation. The required feedback variables are listed in the table below. The first 14 variables are listed in the "User's Guide to SORGF" as feedback variables. GSUM is needed because of the change in the model for this study that defined the period of half-bloom to maturity to be 741 heat units. SUMES1 and SUMES2 were included as feedback variables in the study for improved accuracy. Values for the initial input variables described in table 1 of the main report are also required, except that the plant available soil water value, SW, should be for the feedback date, not the initial starting date.

Appendix: Feedback variables

Variable	Description
1. NLF	Number of leaves emerged by the feedback date
2. NFULL	Number of fully grown leaves by the feedback date
3. IDAY3	Julian date of floral initiation
4. IDAY6	Julian date of half-bloom
5. IDAY9	Julian date of physiological maturity
6. SPROUT	Julian date of plant emergence
7. RCOUNT(J)	Julian date of emergence of leaf J
8. ACDAYS(J)	Julian date when leaf J attains maximum leaf area
9. GRO(1,J)	Current leaf area (cm^2) for leaf J
10. WR	Current dry weight (g) of roots
11. WL	Current dry weight (g) of leaves
12. WC	Current dry weight (g) of culm
13. WH	Current dry weight (g) of vegetative parts of head
14. WG	Current dry weight (g) of grain
15. GSUM	Heat unit sum from day after half-bloom to feedback date
16. SUMES1	Cumulative Stage 1 evaporation (cm)
17. SUMES2	Cumulative Stage 2 evaporation (cm)

Detailed comments on the four feedback options used in this study follow.

Leaf Counts Only (NL)--The feedback date for each sample was the first visit with leaf counts after it was known that the flag leaf was fully grown. This was also the date that the leaf area data were collected. At that time the model should have all leaves emerged and fully grown. Consequently, the feedback variables NLF and NFULL were assigned to be the total number of leaves produced per plant, which is the input variable NTL. The average number of leaves per plant (NL) for each sample was used to assign values to the feedback variable GRO(1,J) as follows:

GRO(1,J) = 0	J = 1, ..., NTL-NL
GRO(1,J) = XMAX(J)	J = NTL-NL+1, ..., NTL

Only the top NL leaves are present on the plant and have leaf area. The bottom leaves are senesced.

Leaf Counts and Leaf Area (NLLA)--The feedback date for each sample was the date the leaf measurements were collected. The leaf count was used in the same manner as for the NL option. The mean measured leaf areas were used instead of XMAX(J) to assign values to the feedback variable GRO.

Half-bloom Date (HBD)--The feedback date was the observed half-bloom date. The plant feedback variables, such as leaf area and plant part weights, were assigned the modeled values on the modeled half-bloom date. This procedure assures that the plant variables have reasonable values for the half-bloom date, such as no grain weight and all leaves emerged and fully grown.

Half-Bloom Date, Leaf Counts, Leaf Area, and Dry Head Weight (ALL)--The feedback date was the first visit date that head weights were obtained. The observed dry head weight was partitioned to grain and nongrain portions (WG and WH) by use of the modeled ratios of WG and WH to total dry head weight. The modeled ratios were determined from WG and WH values the same number of days after modeled half-bloom as the feedback date was after observed half-bloom. This procedure should provide appropriate partitioning fractions. Leaf area measurements were collected on this feedback date for 69 of the 79 samples and were used to adjust the values in the GRO variable. The observed half-bloom date was used in the calculation of GSUM and for the feedback variable IDAY6.