Published January, 1976

Ontogenetic Model of Cotton Yield[1]

Smith Worley, Jr., Harmon H. Ramey, Jr., D. C. Harrell, and T. W. Culp[2]

ABSTRACT

A model of cotton (*Gossypium* sp) lint yield that equated yield to the volume of a parallelepiped with three dimensions was developed by Kerr. The model is extended to basic units by dividing the axes of the parallelepiped and the seed is recognized as the basic unit of yield. Yield per seed, per boll, and per unit area of land can be stated in equation form by using the divisions of the axes. The relative contributions of the primary lint yield components decrease as the complexity of the model increases from yield/seed to yield/unit area of land. The summation of each model has an upper R^2 limit of 1.0. Therefore by the inclusion of additional components, the relative contribution of yield components from the previous models must decrease. Data are presented that illustrate a use of the model to define the component of yield more susceptible to alteration to increase yield.

Additional index words: Stepwise regression, Correlation, Yield Components, Bolls/m², Seeds/boll, Lint/seed, Lint yield/m², Micronaire, Mean length, Fibers/seed, Lint yield/boll, *Gossypium* sp.

RICHMOND (1962) succinctly summarized a philosophy for cotton research: "Solving large complex problems by first solving a number of smaller, technically manageable phases of the problem is a time tested experimental technique of proven value." Also he suggested that this approach has a high probability of success.

Increasing yield of cotton is one of those large, complex problems that has been attacked wholly rather than in a series of smaller parts. Kerr (1966) presented concepts for a cotton yield model that set the stage for attacking the large, complex problem in a series of smaller, more manageable units. He

[1] Cooperative investigations of the ARS-USDA, and the Tenn. and S.C. Agric. Exp. Stns. South Carolina Technical Contribution No. 1213. Received 30 May 1975.

[2] Research agronomist and research geneticist, Cotton Quality Laboratories, ARS-USDA, University of Tennessee, Knoxville, TN 37916; and research agronomists, ARS-USDA, Pee Dee Exp. Stn., Florence, SC 29501.

adapted to cotton the geometric models of yield developed for small grains by Grafius (1956, 1964) and Grafius and Weibe (1959). Maner et al. (1971) showed its usefulness.

The Kerr model (Fig. 1) equates seed cotton yield to the volume of a rectangular parallelepiped with three dimensions: X for bolls/unit area (B/A), Y for seed/boll (SB), and Z for seed cotton/seed (SCS). Seed cotton/seed (Axis Z) is divided into two fractions, fiber weight/seed (F) and weight/seed (S). Lint yield is the volume of XYF. This division of lint yield into three components has been utilized in

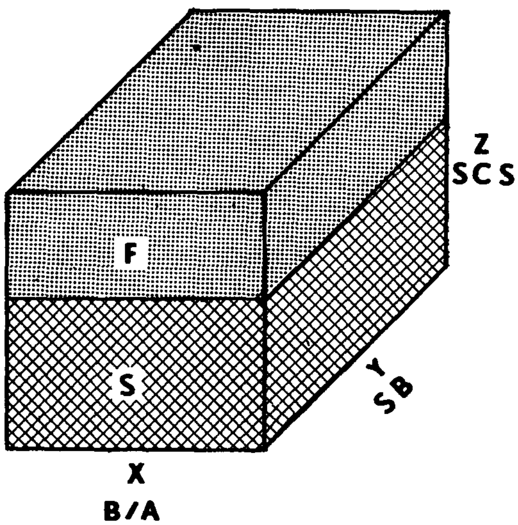

Fig. 1. Kerr model of yield of cotton. X, Y, and Z are the three axis. Other symbols are: B/A for bolls/unit area; S/B for seed/boll; SCS for seed cotton/seed; F for fiber/seed; and S for seed/seed.

the southeastern high-quality cotton-breeding program at Florence, S.C. and has been effective in increasing yield. However the components are still complex and further division could make the model more effective. This paper presents a model for cotton yield that applies to breeding or cultural practice research programs and reduces the components to basic biological entities.

EXTENSION OF MODEL

The rectangular parellelepiped used by Kerr in depicting the model makes the ontogeny of yield rather difficult to visualize. The development of seed and lint yield is more readily seen from a flow chart diagram (Fig. 2). The two sections of the flow chart are additive and the sum equals seed cotton yield. Certain components appear in both sections because the fibers are appendages of the seed surface. By subdividing units of the Kerr model, complexities of cotton yield are more easily visualized. Compartmentalization of the flow chart permits determination of basic units of seed, lint, or seed cotton yield/unit area of land. A model of lint yield can be derived from the basic units and stated in equation form.

Kerr divided the Z axis into two fractions S and F. Both are complex units and are further divided in the flow chart. S is represented in the flow chart as SWt/S and is the product of the volume of the seed (V/S) and its weight/volume (Wt/V). L is represented in the flow chart by LY/S and is the product of the number of fibers/seed (F/S) and the weight/fiber (Wt/F). Further Wt/F is the product of the mean length of the fiber (ML) and the mean weight/unit length (Mic.). Lint yield/seed (LY/S) can be expressed by the following equation:

$$LY/S = F/S \times ML \times Mic. \qquad [1]$$

The seed is the basic morphological entity of lint yield. Botanically the lint is a by-product of the seed. However this by-product is considerably more important economically than the seed.

The Y axis of the Kerr model is the product of the number of seeds/locule (S/L) and the number of locules/boll (L/B). The seed yield/boll (SY/B) is the product of seed/boll (S/B) and seed weight/seed (SWt/S). In most Upland cotton cultivars there are predominantly four or five locules/boll so S/B can

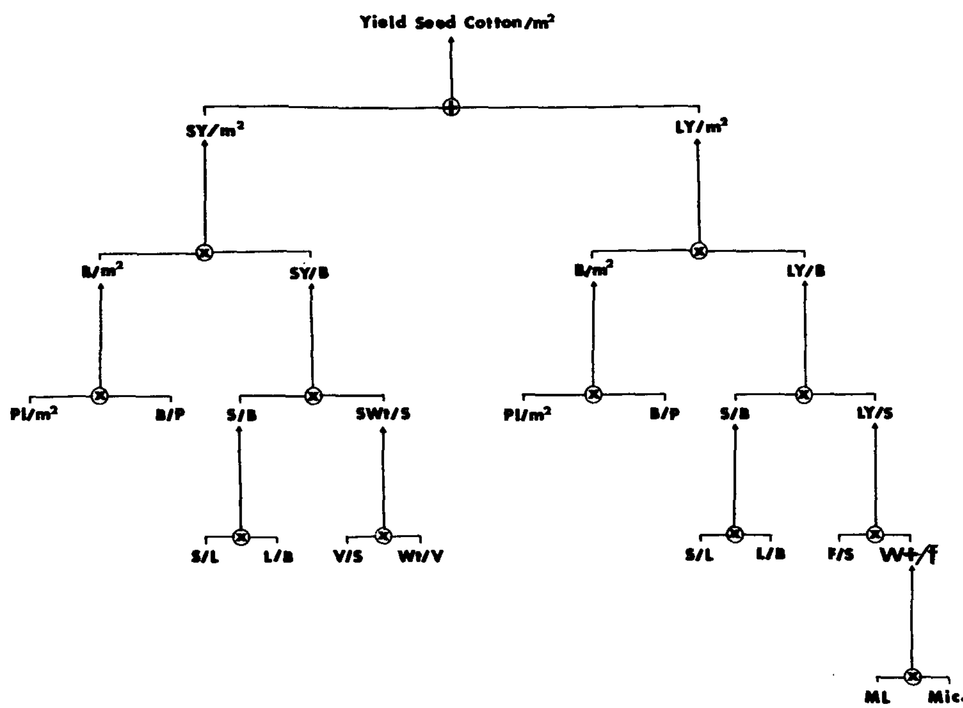

Fig. 2. Flow chart of yield development in cotton. See text for explanation of symbols.

be treated as a unit. In some situations, the seeds per locule and locules per boll may require separate treatment, not discussed here. Also lint yield is more important economically and we will deal only with the lint in further development of the model at this time. Lint yield/boll (LY/B) can be expressed by extention of the equation for lint yield/seed as follows:

$$LY/B = S/B \times F/S \times ML \times Mic. \quad [2]$$

The X axis of Fig. 1 is subdivided into its components, the number of plants/unit area (Pl/m^2) and the number of bolls/plant (B/P). Within certain limits, a compensatory relationship appears between these two variables. The effect of plant spacing on yield is relatively minor when Pl/m^2 is varied threefold (Culp et al. 1974). Therefore we will not use the breakdown but treat bolls/square meter (B/m^2) as one variable. In plant-spacing research, however, these components should be treated individually. Lint yield/unit area (LY/m^2) can be expressed by extension of the equation for LY/B as follows:

$$LY/m^2 = B/m^2 \times S/B \times F/S \times ML \times Mic. \quad [3]$$

This equation becomes the model for lint yield, which can be used in cotton breeding and cultural practice research programs. The first two components (B/m^2 and S/B) may require further subdivision in specific applications.

Kerr (1966) regarded fiber length and fineness (weight/unit length) as secondary yield components. Since the seed is the basic unit in the ontogeny of yield, then fiber length and fineness must be primary yield components, whose effects are confounded and obscured as we progress from the basic unit (the seed) to the yield/unit area of land.

VERIFICATION OF MODEL

Data from the 1971 strains tests conducted by the southeastern high-quality cotton breeding program were used for this study. Entries in these tests represented 88 lines of complex pedigrees in the breeding program (Culp and Harrell, 1974; Harrell et al. 1974). Prior to harvesting, a sample of 25 random undamaged bolls was picked from each plot. The 25 bolls from replications 1 and 2 of each line were bulked together and those from replications 3 and 4 of each line were bulked separately before ginning. Boll size (g/boll), lint percentage (LP) and seed index (g/100 seed) were obtained from the two 50 boll bulked samples of each line in each strains test. Seed cotton/plot was the weight harvested by a spindle-type harvester. The seed cotton weight of replications 1 and 2 and those of replications 3 and 4 were combined for the calculations given below. ML and Mic. were determined on lint from the boll samples at the Cotton Quality Laboratories, Knoxville, Tenn. The following equations were used to calculate other yield components:

Seed cotton yield in g/m^2 (SCY/m^2) = Plot yield (g) /plot area (m^2) [4]

Bolls/m^2 (B/m^2) = (SCY/m^2)/(g/boll) [5]

Seeds/boll (S/B) = (g/boll) (100-LP)/Seed index [6]

Seed cotton/seed (SCS) = (g/boll)/ (S/B) [7]

Lint/seed (L/S) = SCS × LP [8]

Seed weight/seed (SWt/S) = Seed index/100 [9]

Lint yield in g/m^2 = (SCY/m^2) × (LP) [10]

Number of fibers/seed (F/S) = (L/S)/(ML) × (Mic.) [11]

Only sound, well-developed, open bolls were included in boll samples, whereas the total seed cotton yield included damaged and partial bolls. Therefore the calculated component values should be viewed as equivalent values.

From the 1971 strains tests, 560 observations, two replications of each entry per test, were obtained from:

(a) 44 entries: two tests; early and late planting, Florence, S.C.
(b) 16 entries: four tests; early, normal, late, and skip-row plantings; Florence, S.C.
(c) 16 entries: two tests; early and late plantings, Florence, S.C.
(d) 12 entries: two tests; early and late plantings, Florence, S.C.
(e) 12 entries: six tests; solid and skip-row plantings, Florence, S.C., and solid plantings at Clemson, S.C.; Griffin, Ga.; Midville, Ga.; and Tifton, Ga.

We used replication values because we were interested in the similarities of the model to the ontogeny of the plant. Also we wanted to use the measured variability, both genetic and environmental.

The net effect of the yield components on lint yield was estimated by stepwise regression analysis (Draper and Smith, 1966). All data were transformed to logarithms for regression analyses because the yield model is multiplicative and the regression model is additive. We used the BMDO2R stepwise regression program (Health Sciences Computing Facility, UCLA). This program computes stepwise a sequence of multiple linear regressions. Added to the regression equation in each step is the one variable that reduces most the error sum of squares. Also it has the highest partial correlation with the dependent variable for fixed values of those variables already added. Further it has the highest F value.

We expect the multiple correlation coefficient for each model to approach one ($R = 1$) if indeed the correct biological entities are used in the regression model.

The relative contributions of lint yield components to lint yield for the population under study are shown in Table 1. The multiple correlation coefficient (R) equals one for each model. This indicates that each model contains appropriate biological entities. Correlations between characters used in the models are shown in Table 2. The partial correlations for each model are shown in Table 3.

The first contribution to LY/S (Model 1) is F/S (Table 1). In the ontogeny of development, the fibers first are initiated, then grow in length, and finally increase in wall thickness (Balls, 1915). About 39% of the variation in lint yield/seed can be attributed to F/S. After removal of the effect of F/S, Mic. contributed 41% of the variation in yield. ML contributed about 20%.

Model 2 further shows the concepts based on the ontogeny of the plant. Here the biological entity is the boll rather than the seed. The first variable that should be removed in the stepwise regression is logically S/B. Yield should be determined first by S/B, then by components of LY/S in the same order as in Model 1. In Table 1, S/B contributed almost 23% of the variation in yield. Since in each model, the summation of all contributions to lint yield should be 1.0, the percentage contributed by the same variables in the different models should decrease as the complexity of the model increases.

In Model 3, the preponderance of the variation in LY/m^2 comes from B/m^2. The range in LY/B is necessarily limited by its components.

The correlation coefficients of lint yield components are shown in Table 2. These data indicate that an increase in F/S is associated with a decrease in both ML and Mic. LY/S shows strong positive association with F/S and a weaker negative association with Mic. S/B shows low (but highly significant) negative correlations with LY/S and F/S. Generally as S/B increases, the size of the seeds decreases, so these relationships reflect the ontogenetic development of the boll. LY/B shows a high positive relationship to LY/S and a smaller positive relationship to S/B, F/S, and Mic. B/m^2 shows low, but highly significant, relationships with all the components of LY/S. The highest correlation is that of B/m^2 with F/S. This negative relationship with F/S indicates that increasing number of fibers/seed may lead to fewer bolls/unit area; however, the correlation appears too low to create insurmountable problems. Also this relationship probably arises from the negative association between S/B and F/S.

LY/m^2 and B/m^2 show a very strong positive relationship, as would be logically expected.

The partial correlation coefficients of lint yield components used in modeling LY/S, LY/B, and LY/m^2 are shown in Table 3. These coefficients explain the order of variable removal in the stepwise regression program.

With LY/S (Model 1), F/S was the variable with the highest correlation coefficient. When the effect of the variation due to F/S was removed (held constant), then Mic. showed the highest partial correlation coefficient.

The partial correlation coefficients in Models 2 and 3 followed a similar pattern.

Table 1. Relative contributions (cumulative) of lint yield components to lint yield.

Variable entered	R	R^2	Increase in R^2
Model 1-Lint yield/seed			
Fibers/seed	0.6220	0.3869	0.3869
Micronaire	0.8947	0.8004	0.4134
Mean length	1.0000	1.0000	0.1996
Model 2-Lint yield/boll			
Seeds/boll	0.4759	0.2265	0.2265
Fibers/seed	0.7100	0.5041	0.2776
Micronaire	0.9138	0.8352	0.3310
Mean length	1.0000	1.0000	0.1648
Model 3-Lint yield/m^2			
Bolls/m^2	0.9210	0.8483	0.8483
Seeds/boll	0.9401	0.8839	0.0356
Fibers/seed	0.9615	0.9245	0.0406
Micronaire	0.9873	0.9748	0.0504
Mean length	1.0000	1.0000	0.0251

Table 2. Correlation coefficients of lint yield and its components.

	ML	F/S	LY/S	S/B	LY/B	B/m^2	LY/m^2
Mic.	0.176**	-0.384**	-0.357**	0.058	0.292**	0.125**	0.232**
ML	1.000	-0.626**	-0.081	0.010*	-0.066	0.157**	0.135**
F/S		1.000	0.610**	-0.206**	0.407**	-0.270**	-0.123**
LY/S			1.000	-0.300**	0.700**	-0.161**	0.095*
S/B				1.000	0.468**	0.068	0.240**
LY/B					1.000	0.099*	0.267**
Bm^2						1.000	0.929**

*,** Significant at 5 and 1% levels of probability, respectively.

Table 3. Partial correlation coefficients of lint yield and its components.

	Variables correlated	Variable (s) constant	Partial correlation coefficient
Model 1	LY/S vs. Mic.	F/S	0.821**
	LY/S vs. ML	F/S	0.499**
	LY/S vs. ML	F/S + Mic.	1.000**
Model 2	LY/B with F/S	S/B	0.5991**
	LY/B with Mic.	S/B	0.3533**
	LY/B with ML	S/B	-0.0807
	LY/B with Mic.	S/B + F/S	0.8167**
	LY/B with ML	S/B + F/S	0.4831**
	LY/B with ML	S/B + F/S + Mic.	0.9999**
Model 3	LY/m^2 with S/B	B/m^2	0.4845**
	LY/m^2 with F/S	B/m^2	0.4128**
	LY/m^2 with Mic.	B/m^2	0.2983**
	LY/m^2 with ML	B/m^2	-0.0551
	LY/m^2 with F/S	B/m^2 + S/B	0.5911**
	LY/m^2 with Mic.	B/m^2 + S/B	0.3782**
	LY/m^2 with ML	B/m^2 + S/B	-0.0616
	LY/m^2 with Mic.	B/m^2 + S/B + F/S	0.8162**
	LY/m^2 with ML	B/m^2 + S/B + F/S	0.4938**
	LY/m^2 with ML	B/m^2 + S/B + F/S + Mic.	0.9998**

** Significant at 1% level of probability.

DISCUSSION

The adoption of Richmond's (1962) philosophy for cotton research has led to the evolution of a model for cotton lint yield. This model is based on the ontogeny of the cotton plant. The seed is the basic biological unit in the ontogeny of yield and in the model. However the appendage of the seed (lint or fiber) is the unit of most economic importance.

Three components of LY/S are F/S, ML, and Mic. Kerr (1966) regarded the latter two as secondary yield components. In our model, they are primary components. Both ML and Mic. have relatively narrow ranges of acceptability in the textile industry. Generally effort is expended to maintain these components at their present level. Thus the main way to increase lint yield/seed is to increase the number of fibers/seed.

The simple correlations within these data indicate, however, that increasing F/S may not be the best way to increase LY/m^2. In these data an extremely high simple correlation of LY/m^2 with B/m^2 was found and the second highest correlation was with LY/B. Of the four components of LY/B, two—ML and Mic.—are subject to commercial limitations. Therefore S/B appears to be the most likely component for alteration to increase yield in the populations evaluated, if efforts to increase B/m^2 are maintained.

We have also analyzed line means within tests. The results were so similar to the data presented here that the analyses of line means are not given. The conclusion given above that S/B is the most likely component to alter was deduced independently from each of our analyses of these populations, and also of other populations (Worley et al., 1974).

Common elements were used in the estimation of values for some pairs of traits. However the observed correlation between each pair of traits in which a common element was used in estimating values was not altered appreciably when the common element was treated in the analysis and held constant. Partial correlation coefficients between the pair of traits with the common elements constant were not unlike the simple correlations. Thus there was sufficient independence among the estimated traits for the regression analysis to be valid.

Our model of cotton lint yield is useful in pinpointing the yield components most susceptible to alteration. Application of these models in the southeastern high-quality cotton breeding program has resulted in significant yield increase.

The relative contributions of the basic yield components would be expected to vary considerably, depending upon the ranges in the genetic populations and environments. Because of differences in environments and cultural practices, the optimum combination of yield components for the eastern edge of the Cotton Belt evaluated here should be quite different than that of, say, the Texas High Plains. Therefore we expect the regression coefficients from different research programs to vary considerably.

Furthermore the yield component most susceptible to alteration may differ from program to program depending upon the diversity of the germplasm and environments.

Yield is the result of a series of concommitant, difficult to define, response surfaces. The yield models presented herein are basic biological entities. Alteration of these entities, within commercial limitations, as determined from the model can result in increased yield.

ACKNOWLEDGMENTS

We thank Miss Nellie Acres for her careful attention to detail in handling the mass of data analyzed and for her help in preparing the tabulations, and the Univ. of Tennessee Computing Center for the analyses.

REFERENCES

Balls, W. L. 1915. The development and properties of raw cotton. A. C. Black, Ltd., London 221 p.

Culp, T. W., and D. C. Harrell. 1974. Breeding quality cotton at the Pee Dee Experiment Station, Florence, S.C. ARS-USDA S-30.

———, ———, and J. B. Pitner. 1974. (*Gossypium hirsutum* L.). South Carolina Agric. Exp. Sta. Bull. 575.

Draper, N. R., and H. Smith. 1966. Applied Regression Analysis. John Wiley, N.Y., p. 407.

Grafius, J. E. 1956. Components of yield in oats: A geometrical interpretation. Agron. J. 48:519-523.

———. 1964. A geometry for plant breeding. Crop Sci. 4:241-246.

———, and G. A. Weibe. 1959. Expected gain in yield in small grain. A geometrical interpretation. Agron. J. 51:506-562.

Harrell, D. C., T. W. Culp, W. E. Vaught, and J. B. Blanton. 1974. Recent breeding progress in improving lint yield and fiber quality in PD lines of Upland cotton (*Gossypium hirsutum* L.) South Carolina Agric. Exp. Stn. Tech. Bull. 1052.

Kerr, T. 1966. Yield components in cotton and their interrelationships with fiber quality. Proc. 18th Cotton Improvement Conference.

Maner, B. A., S. Worley, D. C. Harrell, and T. W. Culp. 1971. A geometrical approach to yield models in Upland cotton *Gossypium hirsutum* L. Crop Sci. 11:904-906.

Richmond, T. R. 1962. How crop climate research contributes to cotton research. Econ. Bot. 16:217-220.

Worley, S., T. W. Culp, and D. C. Harrell. 1974. The relative contributions of yield components to lint yield of yield of Upland cotton, *Gossypium hirsutum* L. Euphytica 23:399-403.