PARTITIONING VARIETY AND ENVIRONMENT CONTRIBUTION TO VARIATION IN YIELD, PLANT GROWTH, AND FIBER QUALITY

Tom Kerby, Janet Burgess, Marc Bates
Dave Albers and Ken Lege
Delta and Pine Land Company
Scott, MS

Abstract

Yield response of a variety in a variety trial is the result of the genetic capacity of the variety and the environment where the trial is grown. Cotton breeders seek to develop genetic materials that will exceed the performance of current varieties. Agronomist, soil scientists, entomologists, pathologists, and physiologists, conduct studies to identify the environment that optimizes cotton production. Many areas of the U.S. cotton belt have experienced environments that are believed to be significant departures from expected norms during recent years. Questions have been raised regarding varietal performance and what percentage of variation in yield, fiber quality, and growth can be accounted for as due to factors associated with varieties or with environments. Delta and Pine Land Company Technical Services conducts wide scale testing of varieties at many locations over multiple years. Nine early picker varieties were compared at nine locations each in 1997 and 1998 from North Carolina to Texas. Yield variation as assessed in large grower strip plot trials or mean values of University Official Variety Trials (OVT's) was conducted for 12 varieties using a total of 785 locations over the period 1996 through 1998. Summary results indicate about 90 % of the total sum of squares could be accounted for by location. Location (environment) accounted for the following percentages of total variation in other variables: fiber length (85 %); micronaire 69 %; fiber strength (48 %); final plant height (90 %), final number of nodes (85 %); cutout node (80 %), and percent retention on the bottom five first position fruiting branches (78 %).

Introduction

Cotton is an indeterminate perennial crop grown as an annual in the U.S. It is a crop that fascinates and captivates agronomists and physiologists. Interest in cotton physiology has grown in recent decades. Significant quantities of data are available to describe cotton growth and development and how cotton responds to various environments. Cotton producers note variation in performance of fields from location to location and between years. Frequently, they assume good as well as poor performance should be attributed to the variety. Verhalen and Murray (1970) grew 11 varieties in three Oklahoma locations over a two year period. Their results indicated that 61 percent of the variation in yield across varieties and environments was associated with environment. While 61 percent of the variation being accounted for by environment may seem high, it is lower than it would be if the same varieties were grown in more environments (more tests over a geographic area greater than just Oklahoma). Bassett and Kerby (1996) reported environmental contribution to varietal yield variation over a nine year period in California. Similar data were available for fiber quality, but not published. They stated that the variation associated with varieties was generally less than the variation associated with locations. Again, this is for similar Acala varieties grown in a similar area, the San Joaquin Valley.

An example of variety by environment measurements for more varieties over a more diverse geographic region is Shafii et al. (1992). This three year study of rapeseed that inolved a total of 45 environments (year x location) and between 26 and 60 varieties per year demonstrated that 85 percent of the total sum of squares associated with yield variation could be accounted for by environment. For cotton, Kerby et al. (1996) reported plant map data to document the variety and environment contribution to earliness from 7 Deltapine varieties grown in 110 different year x location combinations. They reported the following percent of total variation (sum of squares) associated with location (environment): retention of the bottom five first position fruiting positions as bolls (81 %); cutout node (81 %); final plant height (86 %); and final number of nodes (77 %).

Delta and Pine Land Company conducts extensive variety trials across the country. This manuscript partitions the variation in yield, plant growth, and fiber quality associated with variety, environment, and environment by variety. In addition, Delta and Pine Land Company has combined the yield data from all cotton official variety tests conducted by the Universities between 1995 and 1998. Regression analysis of this data will be presented as a possible means of comparing variety by environment estimates to total variation and to plot yield response of some well known varieties over environments (locations x years).

Material and Methods

In this manuscript we use the term environment to mean all the biotic and abiotic factors present at that growing location. It would include all aspects of weather (temperature, wind, precipitation, drought, cold, heat, etc.), impact of planting date, plant stand, disease pressure, soil type, and management factors including items such as irrigation, fertility, use of plant growth regulators, weed control pressure and practices, insect pressure and control, etc.

Reprinted from the *Proceedings of the Beltwide Cotton Conference*
Volume 1:528-532 (2000)
National Cotton Council, Memphis TN

Statistical Analysis

For replicated plots, variation could be partitioned into the following groups:

- Replications nested within locations
- Varieties
- Locations
- Variety x Location
- Error

To keep the comparison equivalent to that developed from strip plots, or data where means of replications were used, data are presented indicating what percentage of the variation from the sum of squares for varieties, locations, and the interaction of variety and location are accounted for by the three variables.

For strip trial data, or plots where the mean value of replications were used in the analysis, variation can be partitioned into the following sources:

- Varieties
- Locations
- Variety x Location

For this analysis, variety x location includes the variation of the interaction, and also the random variation that is not partitioned into an "error term". Thus, with the type of analysis available for strip trial data, the percent of variation associated with locations is probably underestimated since variety x location also includes random error that can not be partitioned.

Regression Data

Data were collected from all University Official Variety Trials (OVT's) from 1995 through 1998. Sixteen varieties (described in Fig. 2) were common to most of the these 523 OVT's. Yield of each variety was regressed against the mean yield of all varieties at that location. Mean yield of all varieties should be a reasonable estimate of "environment". Yield of the individual 16 varieties regressed against the yield of all varieties across all 523 locations provided the estimate of percent variation accounted for by "environments".

Small Plot Replicated Tests

Sure Grow conducted small plot (two to four plots 30 to 60 feet in length) tests with four replications at nine locations in 1997 and 9 locations in 1998 where common varieties were present. All nine varieties are considered to be early maturity picker types. Varieties in the test were SG 125, SG 501, SG 747, SG 821, SG 105, SGX 1416, SGX 2725, SGX 890, and ST 474. Test locations were as follows: North Carolina 1, South Carolina 3, Georgia 6, Florida 1, Alabama 3, and Texas 4 locations.

These plots were mostly grown with grower cooperators using management preference of the grower cooperator. Plots were harvested with a small plot spindle picker using a bagging attachment. Fifteen pound seed cotton samples were ginned on a 40 saw experimental gin, and fiber quality determined at Delta and Pine Land Company HVI laboratory.

Strip Plots or Replicated Plots Analyzed using Mean of Replication Data

Yield data are presented from both Delta and Pine Land Company grower cooperator plots, as well as Official Variety Trials (OVT's) from the Universities. Twelve varieties (DP 20 B, DP 32 B, DP 51, DP 5111, DP 5409, DP 5415, DP 5415 RR, NuCOTN 33 B, NuCOTN 35 B, PM 1220 BR, SG 125, and ST 474) were common to most of the test locations used in this analysis. Data were from 1996, 1997, and 1998 and all 16 cotton growing states in the U.S. were represented. While yield data comes from company as well as University OVT's, growth and development data as well as fiber quality data come exclusively from Delta and Pine Land Company field trials.

As could be expected, these data do not represent a balanced design. That is not every one of the 12 varieties occurred in every single test. The twelve varieties were selected because they occurred in most of the variety trials. ANOVA was computed using a general linear model using least square analysis.

Results and Discussion

Small Plot Replicated Tests

Nine varieties were evaluated at nine locations in 1997 and nine locations in 1998. All varieties were present in all tests. Table 1 provides summary data for the partitioning of the sum of squares among variety, location, and variety x location. Variety and the variety by location are all measures of how varieties either respond on average or respond differently according to changing environment. The location component represents how all varieties are responding on average differently at varying locations (environments).

Compared to the results presented by Verhalen and Murray (1970), where 61 percent of the variation was associated with location, these studies of similar maturity cotton across a much more diverse environment resulted in environment accounting for 94 percent of the total variation in yield performance (Table 1). Interestingly, boll size (lint/boll) also showed a very large contribution by location (83 percent). Variety effects had more influence on lint percentage with location accounting for only 55 percent of the total variation.

Fiber length was very much under the control of environment with 85 percent of the variation associated with location (Table 1). Micronaire was influenced a little more by

environment than variety accounting for 59 percent of the total variation. Fiber strength was influenced more by variety than environment with only 29 percent of the variation associated with location. Plant height and number of nodes were strongly under the effect of environment with location accounting for 97 and 93 percent of the total variation, respectively.

Large Grower Plots or University OVT's

This comparison provided for analysis across 16 varieties representing a wide range in plant maturity and growth types ranging from DP 20 to NuCOTN 35 B. Data are also collected over a diverse environment with locations from Virginia to California over a three year period. Number of test locations in the analysis for yield, fiber quality traits, and plant development are shown in Table 2.

Even with a wide range of varieties in the 785 tests for yield, 90 percent of the yield variation was associated with location. Since all 16 states were included in this analysis, we had the question, how much of this high percentage association of variation accounted for by location was due to the fact that locations included Virginia to California? For comparison, we ran the same analysis for 65 tests in Georgia (data not presented). We had almost identical results as the nation wide analysis with location accounting for 91 percent of the variation.

Across years and states, location accounted for 82 percent of the variation in gin turnout percentage. Location accounted for 85, 78, and 66 percent of the variation in fiber length, micronaire, and fiber strength, respectively. Location accounted for the following percentage of variation in various plant growth measures: final plant height (88 %); final number of nodes (87 %); cutout node (80 %); and percent retention in the bottom five first positions (73 %).

Figure 1 provides a summary of the results the replicated small plot data as well as the larger data set with more varieties over more environments. We included the plant map data summary from Kerby et al. (1996). Values presented represent the average value from the different sources of the same data. The yield data from Verhalen and Murray (1970) were not included since it represented only 6 locations from only 2 years all from one state. Location effects have a profound effect on yield, plant growth, and fiber length and micronaire. Of all the factors measured, fiber strength showed the least influence of environment. However, environment had nearly an equal effect as variety on this least influenced factor.

Regression Data

All data presented in this section comes from the Official Variety Test results conducted by the Land Grant Universities in the cotton growing states. In Fig. 2, yield of 16 varieties were regressed against the average yield of all varieties tested in a particular OVT. Mean variety yield at a location should be indicitative of the environment of that location. How yield of the 16 varieties varied at each location should indicate the range in genetic response to the particular environment. Across all environments, R^2 between actual yield of the 16 varieties and mean yield indicates the ratio of variation associated with environment.

Yield response of 16 varieties across 523 University OVT's between 1995 and 1998 demonstrated 90 % of variety yield could be accounted for by average yield of all varieties at that location. This is in close agreement to data developed partitioning sum of squares between varieties, locations, and locations x varieties using ANOV as reported in Tables 1 and 2.

Varietal response with regard to all varieties tested at a location can be characterized by several statistical measures in the regression:

- The variation in response of a variety is quantified by R^2. If all the variation in y (the yield of a specific variety) were accounted for by the average yield of all varieties (x) the value for R^2 would be 1.0. High R^2 values simply suggest the yield level of the variety is stable or deviates in a consistent manner as environment changes. It makes no estimate as to being superior or inferior in yield.
- How the variety responds to changing environment is estimated by the slope. A value > 1.0 indicates the yield of a variety increases more than the average of all varieties as locations move to those that were higher yielding. Conversely, a value < 1.0 indicates the variety is less responsive to yield at higher yield environments.
- The intercept represents the starting point for comparison. It is the "y" value when x is zero. Since no tests contain average yield of all varieties of "0", the regression line (slope) is extended until it crosses the "y" axis. Values > 0 suggest the beginning point for comparison of a variety is greater than the average of all varieties at the lowest yield levels measured. Values < 0 suggest the variety is lower yielding than the average of all varieties at the lowest yielding environments.

Varieties where this type of analysis were conducted are listed in Table 3. N represents the number of OVT locations where the variety was compared. Varieties are listed in descending order for R^2. Linear versus quadratic response across yield levels did not improve R^2 for any variety. Therefore, we can conclude that as varieties varied from the

average of all varieties tested across yield levels, they varied in a linear manner.

DP 5690 is an example of a variety with an intercept near zero (deviation in average yield at low yield), but slope > 1.0 indicating increasing superior performance as yield level increases (Fig. 3). R^2 also suggests it is quite stable. The solid line in Figures 3, 4, and 5 represent a slope of 1.0, or the average yield increase of all varieties across yield environments. The dashed line represents the least square line for response of a particular variety across yield environments. In these 115 University OVT's conducted over a 4 year period, 95 % of the yield variation of DP 5690 could be explained by environment (average yield of all varieties at a location).

DP 50 data are presented in Fig. 4. This variety has an intercept value that is positive, good R^2 suggesting it is stable, but has a slope that is < 1.0. This suggests compared to the average of all varieties tested, DP 50 is above the average at low yields, is less competitive at high yields, but is predicatable in response.

DP 5415 represents a different kind of comparison (Fig. 5). It has an intercept < 0, has a slope slightly > 1.0, but has a fairly low R^2 compared to other varieties in these tests. This suggests it's yield is somewhat lower than average in low yield environments but that yield improves relative to the average as yields increase. The R^2 of 0.895 suggests yield of the variety compared to the average of all varieties is more variable. In Fig. 5 there are many examples of DP 5415 being somewhat above the average (solid line), but there is more deviation of points below the solid line. This indicates there are some OVT's where environment had a larger impact on DP 5415 than the average of other varieties in the test. In these 208 University OVT's conducted over a 4 year period, 89.5 % of the variation in yield could be accounted for by environment. While this is still mostly under the control of environment, DP 5415 has roughtly twice as much variation associated with the variety across environments as did DP 5690 (compare 1- R^2 of Fig. 3 and Fig. 5).

Summary

While cotton breeders have data for many varieties over many locations, partitioning of variation of yield, fiber quality, and plant growth traits has seldom been reported. Data from small plot replicated variety trials, data from Delta and Pine Land Company grower strip trials, and yield data from University OVT tests plots all grown over many environments including years indicate that approximately 90 % of the yield of a variety can be related to environmental conditions where the test was grown. Environment was shown to account for 69 % of the variation in micronaire, and 48 % of the variation in fiber strength. Environment accounted for between 80 and 90 % of the variation in plant growth factors measured.

References

Bassett, D.M. and T.A. Kerby. 1996. Variety Selection. pp. 123 to 133. *In* Cotton Production Manual. (eds.) S. Johnson Hake, T.A. Kerby, and K.D. Hake. Univ. California Publication 3352.

Kerby, Tom, Lowell Zelinski, Janet Burgess, Marc Bates, and Jim Presley. 1996. Genetic and environmental contributions to earliness. pp. 592-594. Beltwide Cotton Conferences. Nat'l Cotton Council, Memphis, TN.

Shafii, Bahman, Karen A. Mahler, William J. Price, and Dick L. Auld. 1992. Genotype x Environment interaction effects on winter rapeseed yield and oil content. Crop Sci. 32:922-927.

Verhalen, L.M. and J.C. Murray. 1970. Genotype by environment interaction study of cotton in Oklahoma. pp. 52-54. Proc. Beltwide Cotton Prod. Res. Conf., Nat'l Cotton Council, Memphis, TN.

Table 1. Partitioning of sum of squares (SS) into that associated with variety, location, and variety x location (V x L) for nine early picker varieties grown at nine locations each in 1997 and 1998.

	% of SS				
Factor	Var.	Loc.	V x L	Model R^2	N
Yield (#lint/A)	1	94	6	0.903	618
% Lint	24	55	21	0.896	635
g lint / boll	4	83	13	0.905	324
Fiber Length	6	85	9	0.942	324
Fiber Strength	54	29	17	0.881	324
Micronaire	21	59	20	0.835	324
Final Height	1	97	2	0.941	522
Final # Nodes	3	93	4	0.926	522

Table 2. Partitioning of sum of squares (SS) into that associated with variety, location, and variety x location (V x L) for 12 diverse varieties grown throughout the 16 U.S. cotton growing states in 1996, 1997, and 1998.

	% of SS					
Factor	Var.	Loc.	V x L	Model R^2	No. Loc.	N
Yield (#lint/A)	1	90	9	0.911	785	5705
% Gin Turnout	6	82	12	0.881	430	3677
Fiber Length	6	85	9	0.815	405	3236
Fiber Strength	15	66	19	0.801	405	3236
Micronaire	6	78	16	0.847	405	3236
Final Height	1	88	11	0.894	213	1722
Final # Nodes	4	87	9	0.911	213	1722
% Ret. Bot. 5	1	73	25	0.75	213	1707
Cutout Node	4	80	16	0.85	213	1722

Table 3. Summary regression data for selected varieties from Unviersity OVT data from states where varieties were entered in tests from 1995 through 1998. Yield of the variety (y) is regressed the average yield of all varieties at a test location (x).

Variety	N	Intercept	Slope	R^2 Linear	R^2 Quadratic
DP 5690	115	- 12.2	1.021	0.951	0.952
DP 50	210	62.3	0.913	0.942	0.943
DP 90	141	- 38.6	1.016	0.935	0.936
DP 51	269	1.0	0.993	0.931	0.931
NuCOTN 35 B	101	55.8	1.004	0.927	0.927
SG 125	378	23.7	1.028	0.925	0.926
DP 5415	208	- 30.2	1.016	0.895	0.895
NuCOTN 33 B	268	30.4	1.001	0.888	0.888

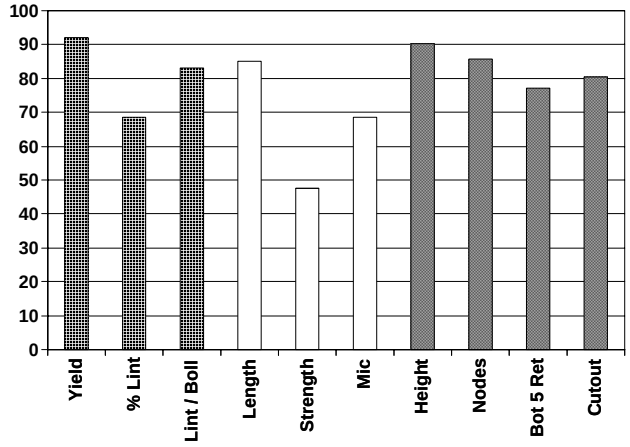

Figure 1. Average of data from Kerby et al. (1996) and data presented in this paper from small plot replicated studies with nine early maturity varieties grown from North Carolina to Texas over a 2 year period and 12 varieties grown in Delta and Pine Land Company or University OVT's from Virginia to California over a 4 year period.

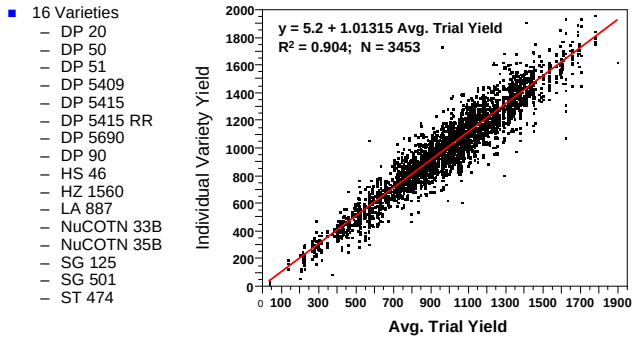

Figure 2. Relationship between individual yield of 16 varieties and the average yield of all varieties regressed across 523 Official Variety Trials conducted between 1995 and 1998. Data from 16 cotton growing states from Virginia to California.

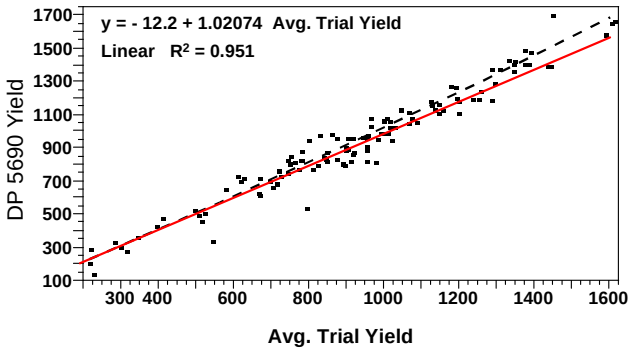

Figure 3. Yield of DP 5690 as related to average yield of all varieties in 115 University OVT trials conducted in 12 U.S. cotton growing states between 1995 through 1998. Solid line represents a slope of 1.00 while the dashed line represents the response line for DP 5690.

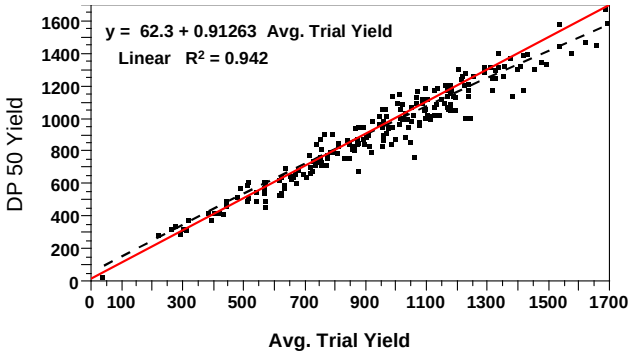

Figure 4. Yield of DP 50 as related to average yield of all varieties in 210 University OVT trials conducted in 15 U.S. cotton growing states between 1995 through 1998. Solid line represents a slope of 1.00 while the dashed line represents the response line for DP 50.

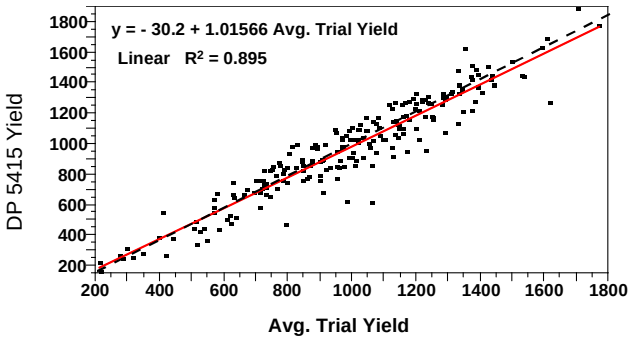

Figure 5. Yield of DP 5415 as related to average yield of all varieties in 208 University OVT trials conducted in 14 U.S. cotton growing states between 1995 through 1998. Solid line represents a slope of 1.00 while the dashed line represents the response line for DP 5415.