Published May, 1969

Genotype by Environment Interactions in Cotton — Their Nature and Related Environmental Variables[1]

H. A. Abou-El-Fittouh, J. O. Rawlings, and P. A. Miller[2]

ABSTRACT

The performance of four varieties of Upland cotton (*Gossypium hirsutum* L.) in 101 environments was used in (*1*) variance analyses to estimate the components of variance among genotypes, genotype by environment interactions, and experimental error, and (2) multiple regression analyses to relate the environment and the interaction effects to several environmental variables characterizing part of the environmental complex. The analyses were performed on lint yield per hectare, boll size, lint percent, seed index, and five fiber traits. The results of the analyses of variance showed that the interaction components were important for yield, but relatively less important for the other traits studied. In the multiple regression analyses, the independent variables were temperature, elevation, and subjective evaluations of moisture availability, disease condition, insect condition, and soil fertility level. These variables were jointly relatable to a proportion of the interaction sum of squares ranging from .245 for fiber fineness to .382 for fiber length. Of the environmental variables studied, temperature had the highest association with interaction.

Additional index words: Variance components, *Gossypium hirsutum* L.

KNOWLEDGE of the magnitude and the pattern of genotype by environment interaction has become essential in helping the plant breeder reach many of the decisions concerning his breeding programs. This differential response by genotypes when subjected to different environments, has been found to exist in many organisms, and ignoring its existence would lead to unreliable conclusions from breeding experiments. The importance of this interaction to

[1] Paper number 2175 of the Journal Series of the North Carolina State University Agricultural Experiment Station, Raleigh, N. C. 27607. This investigation was supported in part by Public Health Service Research Grant No. GM 11546 from the Division of General Medical Sciences. The support for computing was provided by the National Institutes of Health Grant No. FR 00011. Part of a thesis submitted by the senior author in partial fulfillment of the requirements of the Ph.D. degree. Received for publication Jan. 10, 1969.

[2] Former graduate student, Department of Experimental Statistics (now Statistician, Ministry of Agriculture, Giza, UAR), Professor of Experimental Statistics, and Professor of Crop Science, North Carolina State University, respectively.

quantitative genetics and plant breeding has been considered in detail by Allard and Bradshaw (1964), Comstock and Moll (1963), and Matzinger (1963). These investigators gave considerable attention to the study of the nature and magnitude of the various interaction components in different field crops. Studies on cotton reported by Miller et al. (1958, 1959, 1962) have indicated that genotype by environmental interactions are important for lint yield and less important for yield component and fiber traits and that the second-order interaction among genotypes, years, and locations was predominant compared with the first-order interaction of genotypes and years or genotypes and locations. Miller et al. (1959) postulated that the patterns of rainfall distribution and insect infestation were the primary factors determining the differential varietal response.

The primary objective of this study was to characterize the nature of genotype by environment interactions in cotton over a much wider range of environmental conditions than heretofore reported. A secondary objective was to attempt to identify some of the environmental variables associated with the genotype by environmental interactions. The results reported are a by product of a larger study aimed at defining breeding regions for cotton so as to minimize genotype by environment interactions (Abou-El-Fittouh et al., 1969).

DESCRIPTION OF DATA

Data analyzed in this study were the results of the Regional Cotton Variety Tests for the years 1960-1962. These tests are conducted annually by the United States Department of Agriculture in cooperation with the Agricultural Experiment Stations in the Cotton Belt.[3] Data were available from 101 environments representing 39 locations in the Cotton Belt. These locations are shown in Fig. 1, along with the current zoning system of the Cotton Belt for the Regional Testing Programs.

The sets of varieties tested differed from region to region. There were some varieties common in adjacent regions, but only the four national standard varieties were grown in all regions. The national standards were 'Acala 4-42,' 'Lankart 57,' 'Deltapine 15,' and 'Coker 100-A,' representing the Western, Plains, Delta, and Eastern types of cotton, respectively. The nine characters studied were lint yield (kg/ha), weight per boll (g), lint percent (ratio of lint weight to seed cotton weight), seed index (weight of 100 seeds in grams), fiber length (mm), fiber fineness (Micronaire units), fiber strength (T_1, grams per grex), fiber elongation (E_1, percent elongation), and yarn strength (index for 22 count yarns).

The environmental factors studied were elevation of the test site, moisture availability, insect infestation, disease condition, soil fertility level, and temperature. Moisture, insects, diseases, and fertility were each subjectively scored on a five-stage scale by research workers at the test locations. A score of one represented limited moisture, severe insect and disease damage, and very low level of soil fertility; five represented excess moisture, no damage, and very high soil fertility level. The crop season for each environment was divided into four periods of equal length, and the average maximum and the average minimum temperatures for each period and the squares of each were included in the analyses. Thus period 1 included the germination and seedling growth stage; periods 2 and 3 covered the flowering and fruit development stages; and period 4 the latter stages of fruit development and boll opening. The temperature data were taken from U. S. Weather Bureau records.

[3] Data were kindly provided through the courtesy of Dr. C. F. Lewis, Leader of the Cotton Genetics and Breeding Investigations, Agricultural Research Service, Beltsville, Md. Results of the Regional Cotton Variety Tests are published annually by the Cotton and Cordage Fibers Research Branch, USDA, Beltsville, Md. Appreciation and gratitude are due the research workers who kindly provided information on the environmental conditions at their test locations.

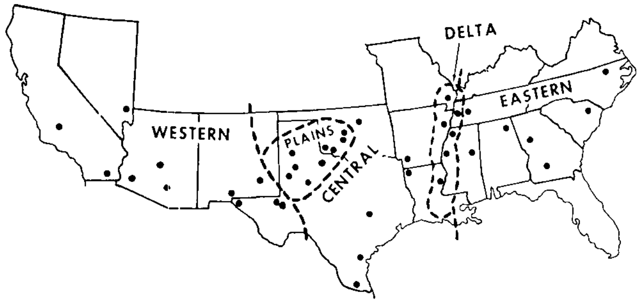

Fig. 1. Test locations of the Regional Cotton Variety Tests.

STATISTICAL ANALYSES

To characterize the nature of the genotype by environment interactions, several different analyses of variance were run on balanced subsets of the data using in each case information on all common varieties. In addition, a least squares analysis for unbalanced data (W. R. Harvey, 1964. Computing procedures for a generalized least-squares analysis program. Paper presented at the Analysis of Variance Conference, Colorado State University, Ft. Collins, Colo.) was used to estimate the pertinent components of variance using data from only the four national standard varieties over all 101 environments. Only the estimates from the last analysis will be presented in detail. Recognizing that the estimates so obtained might be biased due to the nature of the national standard varieties, additional analyses for each region were run omitting the data from the national standards. Some general comments can be made relating the results of the various analyses. Also, due to the difficulty of making unbiased tests of significance with unbalanced data with interaction, statements relating to the significance of the various components are approximations based on the results from the balanced analyses.

To provide an indication of the association of the various environmental variables with the observed environmental and genotype by environment interaction effects, multiple regression analyses were used to partition the corresponding sums of squares into parts due to regression and deviations from regression. From these partitions, the proportions of the total environmental effects and the total variety by environment interaction effects sums of squares relatable to each environmental variable studied, as well as to certain combinations of variables, were calculated.

Significance of the regression of crop performance on the environmental variables was tested by comparing the mean square for regression with the mean square for replications in environments and the interaction mean square was tested for significance against the error mean square.

RESULTS AND DISCUSSION

Estimates of the pertinent components of variance obtained from the general least squares analysis are given in Table 1. For yield, the error component and the interaction components, excepting genotype by year, were from 1¼ to almost 3 times as large as the component due to genotypic differences, and the genotype of location component was the largest of the interaction components. For the remaining traits, the error and interaction components were small and mostly unimportant when compared with the genotypic component. For all traits other than yield, the three-factor interaction was the predominant interaction component and, excepting seed index and lint percent, the genotype by year component was the least important. These results are in general agreement with the findings of similar studies on cotton. The one exception is the genotype by location interaction for yield being larger than the three-factor interaction. In this connection it must be emphasized that the locations involved in this analysis ranged across the entire Cotton Belt, whereas most other estimates have been obtained from studies spanning a much smaller area and, presumably, involving locations much less divergent in the critical environmental factors. The relative importance of the genotype by location component is expected to increase as the reference base of locations is expanded.

The more common pattern with the three-factor interaction being the largest component is obtained if the components are estimated for each region. Table 2 gives the estimates obtained for yield using all varieties common to the locations within each region. With the exception of the Western region all components are significant.

With very few exceptions in the balanced analyses run, the genotypic component of variance was highly significant. Also, inferring from individual year-region analyses, ($\sigma^2_{gl} + \sigma^2_{gyl}$) is probably significant for all traits. For yield, the regional analyses of variance suggest that all components are significantly different from zero.

The authors recognize the likelihood of the estimates obtained using only the national standard varieties being biased as compared to the estimates one would obtain with a more random sample of varieties. The series of regional analyses made to study the contribution of the four national standards to the estimates of the components gave varying results. For yield, omission of this set of varieties resulted in changes in the total interaction mean squares ranging from a decrease of 41% to an increase of 33% of the original interaction mean square. For the other traits, no consistent pattern and no major changes were obtained.

Tables 3 and 4 show, for all nine traits studied, the proportions of the total environmental effects and the total variety by environment interaction effects sums of squares, respectively, relatable to each environmental variable with all remaining variables ignored. The tables also present the same information for the set of all variables combined and for the combined temperature variables.

From Table 3, it can be noted that significant proportions of the environmental effects sums of squares can be related to the set of environmental varieties, as measured, with the highest proportion of .591 for fiber elongation and the smallest of .234 for fiber strength. Temperature variables were the most important environmental factors studied. More than half of the environmental effects sum of squares was attributable to the temperature variables, as a set, for fiber elongation, and a minimum proportion of .206 for seed index. Of the temperature variables, the average minimum temperature in the third period was related to a statistically significant proportion of the existing variability for each of the dependent variables. Either the average maximum or the average minimum temperature in the third period of the growing season was the most important variable studied for yield, boll size, fiber fineness, fiber strength, and fiber elonga-

Table 1. Estimates of the genotype, genotype by year, genotype by location, genotype by year by location, and experimental error variance components.

Component estimated	Lint yield	Boll size	Lint percent	Seed index	Fiber length	Fiber fineness	Fiber strength	Fiber elongation	Yarn strength
θ^2_g	2,184	1.278	.605	3.191	61.29	.019	.0514	.728	236.42
σ^2_{gy}	488	.004	.132	.052	.65	.002	.0002	.014	.04
σ^2_{gl}	5,077	.034	.101	.051	7.74	.005	.0006	.042	1.32
σ^2_{gyl}	2,742	.123	.438	.262	15.48	.027	.0019	.113	8.68
σ^2_e	6,182	.096	.575	.256	32.26	.028	.0060	.213	18.55

Table 2. Estimates of pertinent components of variance for yield within the individual regions.

Component estimated	Region				
	Eastern	Delta	Central	Plains	Western
θ^2_g	3,132**	5,166**	2,184**	1,736**	22,832**
σ^2_{gy}	296*	1,110*	457*	492*	2,539 NS
σ^2_{gl}	630**	2,958**	1,417**	1,004*	2,416 NS
σ^2_{gly}	1,516*	2,930**	1,811*	3,734**	5,038*
σ^2_e	6,178	8,199	6,476	8,286	8,759
n†	9	15	12	12	4

NS, *, and ** designate nonsignificance, and significance at the 5% and at the 1% levels of probability, respectively. † n is the number of varieties in the analysis for each region.

Table 3. Proportions of the total environmental effects sums of squares related to each environmental variable with all remaining variables ignored.

Variable	df	Lint yield	Boll size	Lint percent	Seed index	Fiber length	Fiber fineness	Fiber strength	Fiber elongation	Yarn strength
Max 1†	2	.052*	.018*	.027*	.001	.017*	007	.010	.010*	.071*
Max 2	2	.148*	.033*	.013	.021*	.010*	.037*	.024	.092*	.021*
Max 3	2	.282*	.077*	.175*	.004	.020*	.134*	.083*	.247*	.004
Max 4	2	.141*	.058*	.154*	.019	.031*	.130*	.041*	.189*	.003
Min 1	2	.091*	.009	.062*	.010	.015*	.004	.031*	.002	.007
Min 2	2	.024*	.054*	.006	.023*	.076*	.074*	.066*	.105*	.025*
Min 3	2	.159*	.139*	.187*	.021*	.073*	.242*	.111*	.355*	.033*
Min 4	2	.112*	.076*	.193*	.065*	.079*	.187*	.023	.163*	.013
Elevation	1	.017*	.002	.001	.001	.021*	.105*	.063*	.254*	.200*
Moisture	1	.000	.012*	.000	.054*	.047*	.005	.000	.025*	.054*
Insects	1	.005*	.005	.002	.004	.000	.002	.011	.027*	.041*
Diseases	1	.000	.004	.003	.005	.021*	.063*	.007	.002	.004
Fertility	1	.048*	.073*	.012	.004	.020*	.005	.000	.001	.039*
All temp. variables	16	.421*	.229*	.334*	.206*	.297*	.403*	.221*	.550*	.454*
All variables	21	.490*	.281*	.367*	.251*	.330*	.499*	.234*	.591*	.531*

* Designates significance of regression mean squares at the 5% level of probability. † Max i designates average maximum temperature in the ith period of the growing season, and Min i designates average minimum temperature in the same period.

Table 4. Proportions of the total variety by environment interaction effects sums of squares related to each environmental variable with all remaining variables ignored.

Variable	df	Lint yield	Boll size	Lint percent	Seed index	Fiber length	Fiber fineness	Fiber strength	Fiber elongation	Yarn strength
Max 1†	6	.011*	.090*	.041*	.074*	.050*	.012	.022	.017	.023
Max 2	6	.064*	.058*	.078*	.109*	.101*	.033*	.062*	.033*	.072*
Max 3	6	.060*	.054*	.127*	.107*	.125*	.044*	.079*	.045*	.087*
Max 4	6	.018*	.023*	.095*	.034*	.028*	.037*	.036*	.031*	.035*
Min 1	6	.081*	.080*	.013*	.019*	.039*	---‡	.010	.030*	.023
Min 2	6	.009*	.015*	.057*	.023*	.045*	.031*	.031*	.037*	.024
Min 3	6	.017*	.033*	.153*	.062*	.076*	.043*	.032*	.061*	.058*
Min 4	6	.035*	.049*	.066*	.034*	.043*	.051*	.011	.034*	.021
Elevation	3	.014*	.114*	.055*	.047*	.029*	.023*	.011	.008	.023*
Moisture	3	.015*	.004	---‡	.003	.016*	.007	---	.004	.023*
Insects	3	.009*	.032*	.023*	---	---	---	.003	.036*	.005
Diseases	3	.011*	.006	---	.016*	.010	.003	---	.016*	.018*
Fertility	3	.007*	.034*	.012*	.047*	.003	.014*	---	.012*	.021*
All temp. variables	48	.306*	.283*	.305*	.323*	.322*	.206*	.222*	.209*	.222*
All variables	63	.364*	.373*	.339*	.349*	.382*	.245*	.286*	.289*	.302*

* Designates significanee of regression mean squares at the 5% level of probability. † Max i designates average maximum temperature in the ith period of the growing season, and Min i designates average minimum temperature in the same period. ‡ Where no data are presented regression mean squares were nonsignificant for each of the four varieties at the 5% level of probability.

tion. For lint percent, seed index, and fiber length, the average minimum temperature in the fourth period was the most important variable among those included in the analyses. The highest proportion of the environmental effects sum of squares for yarn strength and significant proportions for all other fiber traits were relatable to elevation. The variable describing disease conditions accounted for a nonsignificant proportion for each character except fiber length and fiber fineness. Moisture scores had almost no association with yield, lint percent, or fiber strength.

Table 4 shows that the set of environmental variables, as estimated, has accounted for significant proportions of the interaction effects sums of squares ranging from .245 for fiber fineness to .382 for fiber length. The temperature variables combined accounted for most of this association. They could explain, as a set, a minimum proportion of .206 for fiber fineness and a maximum of .323 for seed index. The average maximum temperature in the second, third, and fourth periods, and the average minimum temperature in the third period were relatable to significant proportions of the interaction sum of squares for each of the nine traits studied. The variable having the highest association with the differential varietal yield performance was the average minimum temperature in the first period. Average maximum temperature in the second period accounted for the highest proportion for seed index. For lint percent, fiber length, fiber strength, fiber elongation, and yarn strength, the highest proportion in each case was attributable to either the average maximum or the average minimum temperature in the third period. The largest proportion of the interaction sum of squares associated with only one variable for fiber fineness was .051 and was due to the average minimum temperature in the fourth period of the crop season. Of all the variables studied, elevation of the test site accounted for the maximum proportion of the interaction for boll size.

Because of the important association between the temperature variables and the differential varietal responses, another set of multiple regression analyses was performed using only temperature variables in a more complicated model. In addition to the 16 variables representing the linear and the squared deviations from the means of the average maximum and minimum temperatures in each of the four periods of the crop season, the model was expanded to also include the four products between the average maximum and the average minimum temperatures within each period, the three products between the average maximum readings from consecutive pairs of periods, and the three corresponding products between the average minimum readings. For yield, this augmented model increased the proportion of the environmental effects sum of squares accounted for by regression from .421 to .481 and the proportion of the interaction sum of squares from .306 to .451. The modified model was also used for analyses on lint percent. For this trait, the proportion of the environmental effects sum of squares due to regression increased from .334 to .410 and the proportion of the interaction sum of squares from .305 to .419.

Soil survey bulletins, published by the United States Department of Agriculture, give predicted average lint yields per acre by counties for some of the major soil types. The predicted values for the appropriate soil types were studied by the authors in another series of regression analyses. The predicted yield levels of the soils did not account for an appreciable proportion of either the total environmental effects or the interaction effects sums of squares for yield. In a similar manner, the mean yield level of each environment as measured by varieties other than the four national standards in the Regional Cotton Variety Tests failed to show an appreciable association with the genotype by environment interactions for lint yield.

It must be emphasized that the data on environmental factors are very limited, particularly those variables which are subjective evaluations of the environmental conditions. The failure of a particular variable, e.g. moisture score, to show a significant association with either environmental effects or genotype by environment interaction effects should not be taken to imply that moisture *per se* is necessarily unimportant. The apparent lack of importance may be due to the inadequacy of the variable as defined herein. From the results, it appears that with better representation of these factors a relatively high association with interaction can be secured. Special reference is made to insect and disease conditions which were expected to account for a larger proportion of the interaction. However, the care practiced by the research personnel in keeping their research material free from insects and diseases probably limited, to a great extent, the direct and consequently the differential effects of insects and diseases on the performance of the varieties studied.

The temperature variables as a set were related to a proportion ranging from .206 to .323 of the total interaction effects sums of squares for the traits analyzed. Differences in temperatures of a given period of the growing season over sites can be attributed to differences in latitude, elevation, planting date, or combinations of the three. The high negative correlations between elevation and temperature that might be expected *a priori* were not observed in the data used in this study. The simple correlation coefficients between elevation and the different temperature variables were positive for the average maximums in the first and second periods and negative for the others. They ranged from —.510 for the correlation with the average minimum in the third period to .461 for the average maximum in the first period. The average minimum in the first period, which showed the highest proportion of the interaction sum of squares for yield relatable to one variable, had a relatively low correlation of —.120 with elevation of the test site.

Temperature readings were recorded at the nearest United States weather station which was sometimes located near the experimental site and sometimes tens of miles away. It is therefore believed that with more accurate temperature observations for shorter periods of the crop season and with more objective scoring systems for the other environmental factors studied, sizeable proportions of the environmental effects and the interaction effects sums of squares could be accounted for. Increasing the number of temperature variables greatly would require an increase in the number of observations recorded to provide sufficient data for fitting the regression model.

LITERATURE CITED

Abou-El-Fittouh, H. A., J. O. Rawlings, and P. A. Miller. 1969. Classification of environments to control genotype by environment interactions with an application to cotton. Crop Sci. 9:135-140.

Allard, R. W., and A. D. Bradshaw. 1964. Implications of genotype environmental interactions in applied plant breeding. Crop Sci. 4:503-508.

Comstock, R. E., and R. H. Moll. 1963. Genotype-environment interactions. p. 164-196. *In* Statistical genetics and plant breeding. Nat. Acad. Sci. Pub. 982.

Matzinger, D. F. 1963. Experimental estimates of genetic parameters and their applications in self-fertilizing plants. p. 253-279. *In* Statistical genetics and plant breeding. Nat. Acad. Sci. Pub. 982.

Miller, P. A., H. F. Robinson, and O. A. Pope. 1962. Cotton variety testing: additional information on variety × environment interactions. Crop Sci. 2:349-352.

Miller, P. A., J. C. Williams, and H. F. Robinson. 1959. Variety × environment interactions in cotton variety tests and their implications on testing methods. Agron. J. 51:132-134.

Miller, P. A., J. C. Williams, Jr., H. F. Robinson, and R. E. Comstock. 1958. Estimates of genotypic and environmental variances and covariances in Upland cotton and their implications in selection. Agron. J. 50:126-131.