

**NEW METHODS TO ASSESS COTTON VARIETAL STABILITY
AND IDENTIFY DISCRIMINATING ENVIRONMENTS**

A Dissertation

Submitted to the Graduate Faculty of
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

in

The Department of Agronomy and Environmental Management

by

Sterling Brooks Blanche, Sr.
B.S., Western Kentucky University, 1999
M.S., Mississippi State University, 2001
August 2005

Copyright by
Sterling Brooks Blanche, Sr.
2005

DEDICATION

This dissertation is dedicated to my wife, Erin, and my son, Sterling, who, while being my foremost priority, gave of themselves so that this project could be completed.

ACKNOWLEDGMENTS

I would like to express appreciation to my advisor Dr. Gerald Myers, whose commitment to be a mentor “on and off the field” speaks volumes about his character. His philosophy, which effectively cultivated friendship with instruction, far exceeded expectations and will be remembered forever. His guidance in my professional and personal life will last a lifetime. I would also like to thank Dr. Tom Kerby with Delta and Pine Land, Inc., and Dr. Roy Cantrell with Cotton Incorporated for providing financial support during the course of this research.

I express my appreciation to the members of my graduate committee, Drs. Manjit Kang, Bobby Harville, Charlie Johnson, and Don Groth, whose suggestions refined the proposed research. I am greatly appreciative to Dr. Myers for his assistance in the statistical analyses of this project.

I could not have completed this research without the help of my fellow graduate students, Scott Fitch, Jimmy Zumba, Baogong Jiang, and Muhanad Akash. I thank them for their help and meticulous attention to detail that I asked of them throughout the course of my graduate study at Louisiana State University.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	ix
ABSTRACT	x
CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW	1
1.1 Introduction.....	1
1.2 Objectives	1
1.3 Genotype by Environment Interaction Evaluation Methods	2
1.4 Evaluating Performance and Genotype by Environment Interactions	4
1.5 Factors affecting Genotype by Environment Interactions	5
1.6 Literature Cited	7
CHAPTER 2. STABILITY COMPARISONS BETWEEN CONVENTIONAL AND NEAR-ISOGENIC TRANSGENIC COTTON CULTIVARS	11
2.1 Introduction.....	11
2.2 Materials and Methods	14
2.3 Results	16
2.3.1 Analysis of Variance.....	16
2.3.2 Biplot Interpretation.....	20
2.3.3 Plant Height	21
2.3.4 Plant Height to Node Ratio	22
2.3.5 Lint Percentage	25
2.3.6 Delinted Seed Index.....	27
2.3.7 Fiber Length and Strength.....	28
2.3.8 Yield.....	30
2.4 Discussion.....	32
2.5 Literature Cited	34
CHAPTER 3. DETERMINING SELECTION GAINS VIA GGE BIPLOT.....	39
3.1 Introduction.....	39
3.2 Materials and Methods	41
3.3 Results and Discussion	43

3.3.1 Mean and Variance Comparisons	43
3.3.2 Heritability Estimates.....	45
3.4 Literature Cited	49
CHAPTER 4. GGE BIPLLOT VS. TRADITIONAL STABILITY MEASURES	51
4.1 Introduction.....	51
4.2 Materials and Methods	54
4.3 Results and Discussion	56
4.3.1 Within-Model Scaling Correlation	56
4.3.2 Correlation Between GGE Biplot Models (GE).....	57
4.3.3 Correlation Between GGE Biplot (GE) and Other Stability Measures	59
4.3.4 Correlation Between GGE Biplot (GGE) and Other Stability Measures	61
4.4 Literature Cited	63
CHAPTER 5. IDENTIFYING DISCRIMINATING LOCATIONS FOR CULTIVAR SELECTION IN LOUISIANA.....	66
5.1 Introduction.....	66
5.2 Materials and Methods	69
5.3 Results and Discussion	72
5.4 Literature Cited	75
CHAPTER 6. SUMMARY AND CONCLUSIONS.....	78
6.1 Summary	78
6.2 Conclusions.....	80
VITA.....	81

LIST OF TABLES

Table 2.1	Degrees of freedom, sums of squares, and percentages of total variation of genotype (G), environment (E), and genotype x environment (GE) interaction by trait in cotton	17
Table 2.2	Mean cotton plant height (inches) by and across seven environments and stability statistic	22
Table 2.3	Mean cotton plant height to node ratio by and across seven environments and stability statistic	24
Table 2.4	Mean cotton lint percentage by and across seven environments and stability statistic.....	25
Table 2.5	Mean delinted seed index of cotton by and across seven environments and stability statistic	27
Table 2.6	Mean cotton fiber strength by and across seven environments and stability statistic.....	29
Table 2.7	Mean cotton fiber length by and across seven environments and stability statistic.....	29
Table 2.8	Mean cotton yield (lb/A) by and across seven environments and stability statistic.....	32
Table 2.9	Cotton yield and stability ranking by and across seven environments.....	33
Table 3.1	Mean standardized values and variances of parents and F _{2:3} cotton plants in six locations in 2004	44
Table 3.2	Linear correlation (r) and regression coefficients (b) between the F _{2:3} and the F ₂ generations in cotton	46
Table 4.1	Correlation coefficients for GGE Biplot within-model scaling options.....	56
Table 4.2	Correlation coefficients between GGE Biplot models and “traditional” stability measures	58

Table 4.3	Stability values and rankings generated by GGE Biplot and other stability measures for seven cotton cultivars.....	61
Table 5.1	Standardized graphical distances between actual and “ideal” locations, standard deviations, and rankings of six locations for cotton yield.....	72
Table 5.2	Standardized graphical distances between actual and “ideal” locations, standard deviations, and rankings of six locations for cotton fiber length.....	73
Table 5.3	Locations ranked according to desirability for cotton yield, fiber length, and simultaneous selection (yield + fiber length) and standard deviations.....	74

LIST OF FIGURES

Figure 2.1	Mean vs. stability coordination biplot for cotton plant height at 60 days after planting.....	23
Figure 2.2	Mean vs. stability coordination biplot for cotton plant height to node ratio at 60 days after planting	23
Figure 2.3	Mean vs. stability coordination biplot for cotton lint percentage	26
Figure 2.4	Mean vs. stability coordination biplot for delinted seed index of cotton.....	28
Figure 2.5	Mean vs. stability coordination biplot for cotton yield	31
Figure 3.1	Regression of selected F _{2:3} cotton plants on F ₂ cotton plants for Population 1	46
Figure 3.2	Regression of selected F _{2:3} cotton plants on F ₂ cotton plants for Population 2	47
Figure 3.3	Regression of selected F _{2:3} cotton plants on F ₂ cotton plants for Population 3	48
Figure 4.1	Biplot showing mean cotton lint yield and yield stability of seven genotypes	60
Figure 4.2	Biplot of concentric circles ranking seven cotton genotypes for yield and yield stability.....	62
Figure 5.1	Example of biplot showing graphical distance (mm) between actual and“ideal” location.....	71

ABSTRACT

Studies were conducted in 2001-2004 evaluating genotype by environment interactions in cotton (*Gossypium hirsutum* L.). Genotype by Environment interactions were characterized using GGE Biplot for conventional cotton cultivars and their transgenic derivatives. Significant interactions existed for several non-target traits. Transgenic cultivars were taller, had greater height to node ratios, larger seed, and lower lint percentages. Transgenic cultivars containing the Bollgard gene yielded more than their conventional parents and STV4691B was the highest yielding, most stable cultivar. In 2002-2004, GGE Biplot was used to identify two levels (high/low) of discriminating locations for three distinct selection criteria. Crosses were made with parents recommended by a least squares means analysis for each population criteria and F₂ plants were planted in the high and low discriminating locations for each population. Gains by selection (h^2) were calculated by regressing the F_{2:3} plants on their F₂ parents. Genotypic variance was greater among F_{2:3} progeny in discriminating environments compared to non-discriminating environments, regardless of population. Heritability was greater in the population containing fiber traits compared to yield. In 2004, GGE Biplot was compared to other widely-accepted stability analysis tools. Correlation coefficients between GGE biplot (stability evaluation) and the Cultivar Superiority Measure, Shukla's Stability Variance, the Eberhart-Russell regression model, Kang's yield stability statistic, and AMMI were 0.54, 0.91, 0.86, 0.63, and 0.55, respectively. Correlation coefficients between GGE biplot (mean performance + stability evaluation) and the Cultivar

Superiority Measure, the Eberhart-Russell regression model, Kang's yield stability statistic, and AMMI were 0.95, 0.60, 0.85, and -0.33, respectively. Based on the results of this study and our experience using GGE Biplot, Model 3 with an entry-focused scaling is the most valuable analysis for breeders engaged in cultivar development. GGE Biplot was used with the 1993-2003 Louisiana Official Variety Trials to identify the most desirable (discriminating and representative) test locations in Louisiana for yield and fiber length. St. Joseph loam was ranked 1st for yield, Winnsboro irrigated was ranked 1st for fiber length, and St. Joseph loam was ranked 1st to simultaneously select for both traits. Winnsboro non-irrigated should not be used to select for yield or fiber length.

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Cotton (*Gossypium* spp.) is grown primarily as a source of fiber and secondarily for its cottonseed (meal, oil), and contributes significantly to the U.S. economy. In 2004, cotton was grown on an estimated 13 million acres from Virginia to California and south to the lower Rio Grande Valley of Texas and total production was estimated at 23 million bales (NASS, 2005). The overwhelming majority of the cotton produced in the U.S. is upland cotton (*G. hirsutum* L.) with only a small amount (<5%) being Pima types (*G. barbadense* L.). Within upland cotton there are many different varieties marketed each year to producers. These varieties are distinguished from each other due to plant type, maturity, fiber properties, added value traits (e.g., insect and/or herbicide resistance transgenes), yield, and environmental adaptation. In an effort to assist producers in their decision making process about which variety to plant, both public and private entities conduct multi-location variety trials to evaluate plant performance. Stepping back even further, both public and private cotton breeders conduct multi-location testing to assess the performance of their materials (genotypes) in specific environments and over a range of environments.

1.2 Objectives

The general objectives of the dissertation were to investigate genotype by environment interactions in cotton using a stability analysis software tool called GGE Biplot (<http://www.ggebplot.com/concept.htm>). More specifically, the objectives were as follows:

- I. To compare conventional and transgenic cotton cultivars with respect to stability and mean performance.
- II. To determine if the use of discriminating environments for a trait or package of traits can increase the heritability over less discriminating environments.
- III. To compare GGE Biplot to other traditionally used stability measures and determine in which cases similarities or dissimilarities exist.
- IV. To evaluate six locations in Louisiana for their desirability as selection environments based simultaneously on discriminating ability and representativeness.

1.3 Genotype by Environment Interaction Evaluation Methods

Cotton researchers and breeders are aware of differences in performance among cotton cultivars, both geographically and yearly, indicating the presence of genotype by environment interactions. The importance of GE interactions has long been acknowledged since that, in the absence of GE interactions, the best cultivar in any one trial would yield more than all cultivars at all locations every year. Historically, various methodologies have been investigated to study GE interaction including linear regression (Mooers, 1921; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966), cluster analysis (Ghaderi et al., 1982; Johnson, 1977), and principal component analysis (Freeman and Dowker, 1973; Mandel, 1971; Williams, 1952).

A linear regression approach, introduced by Mooers (1921), uses the mean performance of all genotypes in an environment as an index of that environment's productivity against which the performance of each genotype was plotted using linear regression where the mean regression slope would be 1.0. Finlay and Wilkinson (1963) further stated that the overall yield should be taken into account in addition to the slope of a genotype such that genotypes with a high mean yield and slope near 1.0 were well

adapted to all environments and that, as mean yield decreased, a higher or lower slope indicated adaptation to favorable or unfavorable environments, respectively. Eberhart and Russell (1966) added that a stable variety would be one with a regression line slope near 1.0 with a small sum of squared deviations.

Many different dissimilarity measures and clustering strategies have been used to study GE interactions. Johnson (1977) used cluster analysis to examine the yield and stability of maize hybrids using a weighted Euclidean distance as a measure of similarity and the maximum distance between clusters as clustering metric. Others used correlation coefficients of genotypes over environments as a similarity measure (Ghaderi et al., 1982).

Williams (1952) showed the equivalency of principal component analysis and linear regression by proving that extracting the first principal component of the genotypic performances was equivalent to the least squares estimation of the regression coefficients. Multiplicative models using principal component analysis have been used to determine the number of dimensions necessary to contain the genotypic variation and give estimates of the corresponding coefficients (Mandel, 1971). According to Hardwick and Wood (1972), this method would be more beneficial than linear regression when deviations from regression on the environmental mean are significant, but no variables were measured.

The importance of GE interactions in both breeding and variety testing has been acknowledged (Campbell and Lafever, 1977; Pederson and Rathjen, 1981; and Wright, 1976). Bridges (1989) stated that GE research can help breeders determine whether to develop cultivars specific to an environment or for a range of environments. The most

effective breeding could be accomplished if high-yielding varieties were selected and bred for each unique testing environment; however, this would be far too costly to be practical. On the other hand, the cost of breeding and cultivar development would be greatly reduced if mean performance was the only criteria used for cultivar selection and breeding; however, this would come at the expense of maximum gains. Yan (2001) added that GE research, using a biplot-oriented software application called GGE Biplot, can be used to evaluate the performance of different cultivars in a given environment, evaluate the relative adaptation of a given cultivar in different environments, compare two cultivars, identify mega-environments, rank cultivars based on average yield and stability, and rank environments based on discriminating ability and representativeness.

1.4 Evaluating Performance and Genotype by Environment Interactions

A statistic that fully represents a genotype's stability and yield potential, a measure of a genotype's desirability, and provides a meaningful selection criteria for plant breeders, geneticists, and production agronomists is needed. When GE interactions are present, selection based solely upon means is insufficient for a single or range of environments. While high-yielding varieties are easily determined for a single environment or a range of environments by calculating the mean yield, stability statistics are more difficult to instantaneously compute. More difficult still are selections based on both stability and mean performance. Stability statistics such as Shukla's stability-variance statistic (Shukla, 1972) or Wricke's ecovalence (Wricke, 1962) have been used to determine stability, but no meaningful assessment of a genotype's mean performance was evaluated simultaneously. Kang and Magari (1995) developed STABLE, a DOS-version computer program that calculates Shukla's stability-variance statistic and Kang's yield-stability

statistic (YS_i) resulting in selection based upon mean performance and stability. In STABLE, an environmental covariate, or the environmental mean if no covariate data are available, may be used to remove heterogeneity from GE interaction. STABLE (i) determines the contribution of each genotype to GE interaction by calculating σ_i^2 ; (ii) assigns ranks to genotypes from highest to lowest yield; (iii) calculates the protected LSD for mean yield comparisons; (iv) adjusts yield rank according to LSD (adjusted rank is labeled Y); (v) determines significance of σ_i^2 , using an approximate F-test; (vi) assigns stability rating (S); (vii) sums adjusted yield rank and stability rating for each genotype to determine the YS_i statistic; and (viii) calculates mean YS_i and identifies genotypes (selections) with $YS_i > \text{mean } YS_i$ (Kang and Magari, 1995). The YS_i statistic provides meaningful insight into a genotype's yield and yield stability rating allowing for simultaneous selection for both criteria.

1.5 Factors Affecting Genotype by Environment Interactions

The performance of a cotton cultivar is dependent on the genetic capacity of the cultivar, the environment where the cultivar is grown, and the interaction between the cultivar and the environment (Kerby et al., 2000; Myers and Bordelon, 1997; Yan, 2001; Yan and Hunt, 2001). The term genotype by environment (GE) interaction refers to variation that cannot be explained by the genotype main effect or the environment main effect (Yan and Hunt, 2001). The term environment refers to all biotic and abiotic factors that influence plant growth at that location including weather (temperature, wind, precipitation, heat, cold, drought), 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.

(Kerby et al., 2000). When a GE interaction is significant, comparisons based strictly upon the mean may not be adequate (McPherson and Gwathmey, 1996). Useful analyses should be made based on the mean performance of a cultivar (genotype main effect) and the stability of that cultivar (GE interaction) over a range of pertinent environments.

Recently, much variability in the performance of cotton cultivars has been attributed to differences in environment (Kerby et al., 2000); however, for some traits, genes have a greater effect than the environment (Meredith and Bridge, 1972). Meredith and Bridge (1972) reported that, within upland cotton genotypes, genes heavily influence fiber length, strength, and fineness. Other studies have suggested that the relative genetic and environmental influences on fiber strength are determined by a few major genes, rather than by variations in the growth environment (May, 1999). However, large differences in environment can inhibit cotton genotypes from reaching their full genetic potential (Green and Culp, 1990). Many environmental factors affect cotton performance; for example, cotton canopy architecture, plant height, and branch formation can be affected by temperature (Hanson et al., 1956; Hodges et al., 1993; Reddy et al., 1990), growth-regulator application (Cadena and Cothren, 1996; Reddy et al., 1990), light intensity (Hanson et al., 1956; Sassenrath-Cole, 1995), and herbivory by insects and other animals (Rosenthal and Kotanen, 1994; Terry 1992). Similarly, Bradow and Bauer (1997a,b) showed that boll retention, an important yield characteristic, was affected by irrigation method. Gipson and Joham (1969) showed that early-stage fiber elongation, which is controlled more by genes than environment, was highly temperature dependent and late-fiber elongation was temperature independent, indicating that environmental variation

can influence traits controlled primarily by genotype. When genes respond differently to environment, a GE interaction is present (Kerby et al., 2000).

The existence of genotype by environment (GE) interactions is a major concern and necessitates the testing of cotton cultivars in multi-environment trials. Various efforts have been made to analyze cotton cultivar stability using multi-environment trial data, and although no method perfectly accommodates GE interactions, most breeders utilize some form of stability analysis in their varietal selections (Lin and Binns, 1988; Pinthus, 1973). GGE Biplot, a windows-based stability analysis software program, has received much attention because it allows researchers to overlook the large degree of environmental variation and focus primarily on the typically obscure genotypic and GE components that are most useful for cultivar evaluation (Yan et al., 2000). In addition to this development, GGE Biplot also allows for an easily interpreted graphical representation of the data using biplot methodology. It would be advantageous for growers if high-yielding cultivars that perform consistently from year to year were available, even if the highest yield was not always attained by that cultivar.

1.6 Literature Cited

Bradow, J.M., and P.J. Bauer. 1997a. Fiber quality variation related to cotton planting date and temperature. p. 1491-1495. *In Proc. Beltwide Cotton Conf., New Orleans, LA. 7-10 Jan. 1997. Natl. Cotton Counc. Am., Memphis, TN.*

Bradow, J.M., and P.J. Bauer. 1997b. How variety and weather determine yarn properties and dye uptake. p. 560-564. *In Proc. Beltwide Cotton Conf., New Orleans, LA. 7-10 Jan. 1997. Natl. Cotton Counc. Am., Memphis, TN.*

Cadena, J., and J. T. Cothren. 1996. Carbon balance of PRG-IV-treated cotton plants grown under two irrigation regimes. p. 1225-1232. *In Proc. Beltwide Cotton Conf., Nashville, TN. 9-12 Jan. 1996. Natl. Cotton Counc. Am., Memphis, TN.*

Campbell, L. G., and H. N. Lafever. 1977. Cultivar x environment interactions in soft red winter wheat yield tests. *Crop Sci.* 17:604-608.

- Eberhart, S. A., and W. A. Russell. 1966. Stability parameters for comparing varieties. *Crop Sci.* 6:36-40.
- Finlay, K. W., and G. N. Wilkinson. 1963. The analysis of adaptation in a plant breeding programme. *Aust. J. Agric. Res.* 14:742-754.
- Freeman, G. H., and B. D. Dowker. 1973. The analysis of variation between and within genotypes and environments. *Hered.* 30:97-109.
- Ghaderi, a., M. W. Adams, and A. W. Saettler. 1982. Environmental response patterns in commercial classes of common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Gen.* 63:17-22.
- Gipson, J.R., and H.E. Joham. 1969. Influence of night temperature on growth and development of cotton (*Gossypium hirsutum* L.) III. Fiber elongation. *Crop Sci.* 9:127-129.
- Green, C.C., and T.W. Culp. 1990. Simultaneous improvements of yield, fiber quality, and yarn strength in upland cotton. *Crop Sci.* 30:66-69.
- Hanson, R.G., E.C. Ewing, and E.C. Ewing Jr. 1956. Effect of environmental factors on fiber properties and yield of Deltapine cottons. *Agron. J.* 48:546-581.
- Hardwick, R. C., and J. T. Wood. 1972. Regression methods for studying genotype-environment interactions. *Hered.* 28:290-322.
- Hodges, H.F., K.R. Reddy, J.M. McKinion, and V.R. Reddy. 1993. Temperature effects on cotton, p. 1-15. *In* K.H. Remy (ed.). Bull. 990. Mississippi Agric. For. Exp. Stn., Mississippi State, MS.
- Johnson, G. R. 1977. Analysis of genotypic similarity in terms of mean yield and stability of environmental response in a set of maize hybrids. *Crop Sci.* 17:837-842.
- Kang, M. S., and R. Magari. 1995. STABLE: A BASIC program for calculating stability and yield-stability statistics. *Agron. J.* 87:276-277.
- Kerby, T., J. Burgess, M. Bates, D. Albers, and K. Lege. 2000. Partitioning variety and environment contribution to variation in yield, plant growth, and fiber quality. p. 528-532. *In* Proc. Beltwide Cotton Conf. 1997. Natl. Cotton Counc. Am., Memphis, TN.
- Lin, C.S., and M.R. Binns. 1988. A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68:193-198.
- Mandel, J. 1971. A new analysis of variance model for non-additive data. *Technometrics* 13:1-18.

- May, O.L. 1999. Genetic variation in fiber quality. p. 183-229. *In* A.S. Basra (ed.) Cotton fibers, Food Products Press, New York, NY.
- Meredith, W.R. Jr. and R. R. Bridge. 1972. Heterosis and gene action in cotton, *Gossypium hirsutum* L. *Crop Sci.* 12:304-310.
- Mooers, C. A. 1921. The agronomic placement of varieties. *J. Am. Soc. Agron.* 13:337-352.
- McPherson, R., and O. Gwathmey. 1996. Yield and stability of cotton cultivars at three West Tennessee locations. p. 596-598. *In* Proc. Beltwide Cotton Conf., Nashville, TN. 9-12 Jan. 1996. Natl. Cotton Counc. Am., Memphis, TN.
- Myers, G. O., and F. Bordelon. 1997. Stable statistics for cotton varieties grown in the Midsouth, Southeast, North and Texas from 1993-1995. p. 464-466. *In* Proc. Beltwide Cotton Conf., New Orleans, LA. 7-10 Jan. 1997. Natl. Cotton Counc. Am., Memphis, TN.
- National Agricultural Statistics Service (NASS). 2005. Louisiana farm reporter. Volume 5, number 2, 18 Jan, 2005.
- Pederson, D. G., and A. J. Rathjen. 1981. Choosing trial sites to maximize selection response for grain yield in spring wheat. *Aust. J. Agric. Res.* 32:411-424.
- Pinthus, M.J. 1973. Estimate of genotypic value: A proposed method. *Euphytica.* 22:121-123.
- Reddy, V.R., D.N. Baker, and H.F. Hodges. 1990. Temperature and mepiquat effects on cotton canopy architecture. *Agron. J.* 82:190-195.
- Rosenthal, J.P., and P.M. Kotanen. 1994. Terrestrial plant tolerance to herbivory. *Trends Ecol. Evol.* 9:857-859.
- Sassenrath-Cole, G.F. 1995. Dependence of canopy light distribution on leaf and canopy structure for two cotton (*Gossypium*) species. *Agric. For. Meteorol.* 77:55-72.
- Shukla, G. K. 1972. Some statistical aspects of partitioning genotype environmental components of variability. *Heredity* 29:237-245.
- Terry, L.I. 1992. Effect of early season insecticide use and square removal on fruiting patterns and fiber quality of cotton. *J. Econ. Entomol.* 85:1402-1412.
- Williams, E. J. 1952. The interpretation of interactions in factorial experiments. *Biometrika* 39:65-81.

Wricke, G. 1962. Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen Z. Pflanzenzücht. 47:92-96.

Wright, A. J. 1976. The significance for breeding of linear regression analysis of genotype-environment interactions. Hered. 37:83-93.

Yan, W. 2001. GGE Biplot- A Windows application for graphical analysis of multi-environment trial data and other types of two-way data. Agron. J. 93:1111-1118.

Yan, W., and L.A. Hunt. 2001. Interpretation of genotype x environment interaction for winter wheat yield in Ontario. Crop Sci. 41:19-25.

Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnic. 2000. Cultivar evaluation and mega-environment investigation based on the GGE Biplot. Crop Sci. 40:597-605.

CHAPTER 2

STABILITY COMPARISONS BETWEEN CONVENTIONAL AND NEAR-ISOGENIC TRANSGENIC COTTON CULTIVARS

2.1 Introduction

Transgenic cotton (*Gossypium hirsutum* L.) cultivars became commercially available in 1995 with the introduction of bromoxynil [3,5-dibromo-4-hydroxybenzotrile; BXN (Bayer Advance, Peoria, IL)] herbicide-resistant cultivars BXN 57 and BXN 58 (Collins, 1996). Commercial transgenic cotton (*Gossypium hirsutum* L.) cultivars expressing the CryIAc insecticidal protein from *Bacillus thuringiensis* spp. *Kurstaki* [Bt; Bollgard (Monsanto Co., St. Louis, MO)], commonly called '*Bt* cotton', were introduced in the USA in 1996 (Hardee and Herzog, 1997). Glyphosate-resistant cotton cultivars [N-(phosphonomethyl)-glycine; Roundup Ready (Monsanto)] became available in the U.S. in 1997 (Kerby and Voth, 1998). The development of insect- and herbicide-resistant genetically engineered cotton originated as a new approach to control insect pest injury and weeds in production agriculture. In 1995, 82% of US cotton acreage was infested with the tobacco budworm (*Heliothis virescens* L.)/bollworm (*Helicoverpa zea* L.) complex causing 4% yield reductions (Hardee and Herzog, 1996). After the introduction of *Bt* cotton in 1996, US cotton acreage, of which 14% was planted with *Bt* cotton, infested with the tobacco budworm/bollworm complex was only 77% and yield reductions declined to 2.37% (Hardee and Herzog, 1997). In 2000, 28% of the total U.S. cotton acreage was planted with *Bt* cotton and damage caused by the tobacco budworm/bollworm complex was reduced to 1.43% (Hardee and Burris, 2001).

Transgenic cotton acreage continues to increase in Louisiana and the United States. A survey conducted by the National Agricultural Statistics Service (2004) showed that in

2004, 93 and 76% of cotton planted in Louisiana and the United States, respectively, contained transgenes for insect and/or herbicide resistance. The same study found that in Louisiana, stacked-gene cotton (cotton containing 2 or more transgenic traits) acreage increased from 46 to 60% from 2003 to 2004, indicating that Louisiana producers will plant cotton varieties containing multiple pest resistance transgenes. Commercialization of transgenic cottons has enabled producers to increase lint yield and reduce the impact of agriculture on the environment by providing an effective, non- or reduced-chemical means of insect and weed control (Gasser and Fraley, 1989; Benedict and Altman, 2001). In 1994, prior to the introduction of transgenic cultivars, total herbicide and insecticide applications^a in Louisiana were 1.36 and 1.73 million kilograms (kg), respectively, of active ingredient (ai) per year (NASS, 1994). However, in 2001, after the widespread adoption of transgenic cottons, total ai of herbicide and insecticide applications per year in Louisiana declined to 1.15 and 1.00 million kg, respectively (NASS, 2001). In addition to increased weed and insect control and a lower impact on the environment, the ease of use and management of transgenic cottons is desired by cotton producers.

The introduction of only a few transgenic technologies has resulted in many transgenic cotton cultivars, all sharing a genetic base perceived as relatively narrow by many breeders (Benedict and Altman, 2001). Multiple herbicide- and insect-resistant cultivars have been commercialized and there are several transgenic cultivars that, save for the presence of the transgene, are thought to be genetically equivalent to a conventional (non-transgenic) high-yielding cultivar. Historically, the recurrent backcross method has been used for introgression of insect resistance traits from wild species such as high terpenoid concentration (Lukefahr and Martin, 1966; Lee, 1978), absence of leaf nectaries (Meyer

^a Total includes all applications, regardless of target pest or chemical.

and Meyer, 1961), leaf pubescence (Meyer, 1957; Lukefahr et al., 1975), and gossypol level (McMichael, 1960; Lukefahr and Houghtaling, 1969). In this same manner, transgenic cultivars have been developed by backcrossing a transgene-containing line into a high-yielding conventional cultivar until a transgene-containing line is developed that is near-isogenic with the conventional recurrent parent (Bowman, 2000; Benedict and Altman, 2001; Robinson and McCall, 2001). However, the degree of similarity, for both mean performance and stability, between conventional parents and their transgenic derivatives has been questioned (Kerby and Voth, 1998; Raymer and Minor, 1999; Elmore, 2001a; Elmore, 2001b). Each transgenic line results from a separate insertion event, is usually stable, and segregates with Mendelian expectations (Umbeck et al., 1989; Kohel et al., 2000); although, some evidence exists that not all transgenes behave according to strict Mendelian inheritance (Sachs et al., 1998). Researchers have addressed the long-term viability of the recurrent breeding methods for transgene introgression and concluded that there was a need for more forward-breeding in future cultivar development (May et al., 1995; Meredith, 1995; Benedict and Altman, 2001).

Numerous methods for analyzing varietal differences based on their consistency in response to environments have been developed (Pinthus, 1973; Lin and Binns, 1988; Kang and Pham, 1991). One strategy involves factorial regression of the genotype x environment (GE) matrix against environmental factors, genotypic traits, or combinations thereof (Baril et al., 1995). A second strategy, the additive main effects and multiplicative model (AMMI), involves correlation or regression analysis that relates the genotypic and environmental scores derived from a principal component analysis of the GE interaction matrix to genotypic and environmental covariates (Zobel et al., 1988).

GGE Biplot, a recently released Windows-based software package, can be used to perform analyses similar to the popular AMMI model; however, GGE Biplot removes the effect of the environment (E) and focuses on the genotype (G) and GE interaction components relevant to cultivar evaluation. Thus, stability comparisons between conventional parents and their near-isogenic derivatives can be made with higher precision by removing the noise caused by E.

The objective of this research was to use GGE Biplot to compare the mean performance and stability of conventional cotton cultivars and their near-isogenic transgenic derivatives, to determine the non-target traits that are affected by transgene insertion, and the environments that influence that variation

2.2 Materials and Methods

Seven field studies were conducted at four different locations representing major cotton-growing regions in Louisiana. Commercially-available cotton cultivars were planted in 2001 and 2002 at: the Red River Research Station in Bossier City (BC1), the Dean Lee Research Station in Alexandria (A1 and A2), the Northeast Louisiana Research Station in St. Joseph (S1 and S2), and the Macon Ridge Research Station in Winnsboro, Louisiana (W1 and W2). Data taken from Bossier City in 2002 was excluded due to stand loss caused by excessive rainfall and cool early-season temperatures. Each location differs in climate, soil type, geographic location, etc., and management practices differ at all four locations to achieve optimal growth throughout the growing season.

At each location a randomized complete block design was used and four-row plots (4 x 15 m) were replicated three times. Rows were 15 m long and spaced 102 cm apart; all data were collected from the center two rows. Cultivars included in this study were

Stoneville [Stoneville Pedigreed Seed Company, Memphis, TN] (STV) ‘474’, ‘STV 4691B’, ‘STV 4793R’, ‘STV 4892BR’, ‘STV BXN47’, Suregrow [(Delta and Pine Land Company, Scott, MS] (SG) ‘SG 501’, ‘SG 501BR’. The cultivars chosen for this study comprised the range of commercially-available conventional and transgenic cultivars and included at least one cultivar that contained the Roundup Ready[®], Bollgard[®], BXN[®], and the stacked-gene version containing both Roundup Ready and Bollgard technologies.

Data collected during the growing season at 60 and 90 days after planting (DAP) and at harvest included plant height and number of main stem nodes. Plant height was calculated by averaging the distance from the soil surface to the plant terminal for five randomly-selected plants. Main stem nodes between the cotyledonary node and the plant terminal were counted on five randomly-selected plants and averaged. Yield components measured included boll weight, lint per boll, lint percentage (LP), fuzzy seed index (FS), seed index of acid-delinted seed (DS), and lint weight. Fiber properties analyzed were fiber micronaire (MIC), strength (GTEX), length (UHM), uniformity (UI) and elongation (ELON). Yield components were determined from 50 randomly selected bolls taken prior to harvest and cotton fiber quality measurements were obtained using high volume instrumentation (HVI) testing at the Cotton Fiber Testing Laboratory in Baton Rouge, Louisiana.

A SAS analysis, using a proc mixed model (SAS Inst., Cary, NC), was used to create an analysis of variance (ANOVA) table to determine the presence or absence of GE interactions. The percentage of total variation attributed to E, G, or GE interaction for each trait was determined using the sums of squares from the ANOVA table. Response variables that had significant G or GE interactions were analyzed in GGE Biplot and

stability and mean performance of conventional and transgenic cultivars were summarized. Similar to the widely-accepted AMMI model, GGE Biplot is a recently released Windows-based software package designed to examine G main effects and GE interactions using rank-two matrix multiplication and singular value decomposition (Yan et al., 2000; Yan, 2001; Yan and Hunt, 2001). Using the analyses included in GGE Biplot, comparisons of stability and mean performance between transgenic and conventional cultivars were made for plant height, height to node ratio, LP, DS, GTEX, UHM, and lint yield.

2.3 Results

2.3.1 Analysis of Variance

An ANOVA table was generated to indicate the relative magnitudes of G, E, and GE interaction for each trait (Table 2.1). The percentages of the total sums of squares accounted for by G, E, and GE interactions have been used as an indicator of the total variation attributed to each component (Verhalen and Murray, 1970; Kerby et al., 2000). Variation due to G or GE interactions is a measure of how cultivars either respond across environment, or differently according to different environments. The environmental component, E, represents how the cultivar means differ between environments. Studies have shown that environment accounts for >80% of total variation in yield, which is expected considering the large effect that location has on plant growth and morphology; however, traits with high heritability are typically influenced less by environment (Ethridge and Hequet, 2000; Kerby et al., 2000; Epinat-Le Signor et al., 2001). The contribution to the total sums of squares, regardless of trait, for environment ranged from 11 to 92%, genotype ranged from 5 to 55%, and GE ranged 5 to 34% (Table 2.1).

Table 2.1. Degrees of freedom, sums of squares, and percentages of total variation of genotype (G), environment (E), and genotype x environment (GE) interaction by trait in cotton .

Trait	Source	df	ANOVA [†]		% Total variation [#]
			SS	Pr>F [‡]	
Height	E	5	12731	0.0684	70
	G	9	1957	<0.0001*	11
	GE	40	3383	<0.0001*	19
Nodes	E	5	1705	0.0242*	51
	G	9	1249	<0.0001*	37
	GE	40	398	0.0288*	12
Height to node ratio	E	5	12.70	<0.0001*	85
	G	9	0.88	0.0001*	6
	GE	40	1.28	0.3333	9
Boll weight	E	6	4.38	0.0203*	36
	G	9	4.61	0.0186*	38
	GE	26	3.15	0.9177	26
Lint per boll	E	6	1.77	0.0015*	46
	G	9	1.38	<0.0001*	36
	GE	26	0.65	0.3008	17
Lint percent	E	6	132	0.0980	28
	G	9	175	0.009*	38
	GE	26	157	0.6159	34
Fuzzy seed index	E	6	16.2	0.0002*	26
	G	9	29.1	<0.0001*	47
	GE	26	16.4	0.0005*	27
Delinted seed index	E	6	5.41	0.0236*	11
	G	9	26.71	<0.0001*	55

Table 2.1. (cont)

	GE	26	16.16	0.0005*	33
Micronaire	E	6	36009	<0.0001*	49
	G	9	24145	<0.0001*	33
	GE	26	13021	0.0659	18
Fiber strength	E	6	167	<0.0001*	56
	G	9	78	<0.0001*	27
	GE	26	50	0.0907	17
Fiber length	E	6	0.044	<0.0001*	57
	G	9	0.015	0.0011*	19
	GE	26	0.018	0.0802	23
Fiber elongation	E	6	15.8	0.0002*	61
	G	9	6.6	<0.0001*	26
	GE	26	3.4	0.4246	13
Fiber uniformity	E	6	132.2	<0.0001*	82
	G	9	8.4	0.1191	5
	GE	26	21.0	0.8430	13
Yield	E	6	117088771	<0.0001*	92
	G	9	3529769	<0.0001*	3
	GE	49	6148915	0.0008*	5

[†] Analysis of variance (ANOVA) was generated using proc mixed in the SAS System (v. 9).

[‡] Asterisk denotes trait/source combinations significant at ($\alpha=0.05$).

[#] Variation due to each source as a percentage of total sums of squares of E, G, and GE.

Environment accounted for >70% of the total variation for plant height, height to node ratio, and yield, traits expected to be heavily influenced by environment. The relatively small contribution of G to the total sums of squares was 11, 6, and 5% for plant height, height to node ratio, and yield, respectively (Table 2.1). The results of this study generally agree with Kerby et al. (2000), who conducted a study in 1997 and 1998

including nine cultivars at nine locations in North Carolina. They found that contributions by E, G, and GE to the total sums of squares were 94, 1, and 6%, respectively, for yield, and 97, 1, and 2%, respectively, for plant height (Kerby, et al., 2000).

For traits with a higher heritability than yield, the percentage of total variation attributed to G was 33, 27, 19, 26, and 5 for MIC, GTEX, UHM, EL, and UI, respectively, indicating that G was relatively more important concerning fiber characteristics compared to yield or plant height (Table 2.1). GE interaction contributions were 8 to 17% greater for MIC, GTEX, UHM, EL, and UI compared to yield. Traits with the least amount of variation attributed to environment were LP (28%), FS (26%), and DS (11%). For LP, FS, and DS, the primary increase in total contribution to sums of squares was attributed to G (38 to 55%), although contributions from GE interactions were also higher (27 to 34%) than in other traits (Table 2.1). Kerby et al. (2000) found that the contribution of GE interactions for LP, GTEX, UHM, and MIC to the total variation in sums of squares ranged from 9 to 21%, which was much higher than the 6 and 2% for yield and plant height, respectively. This suggests that the study and characterization of GE interactions for fiber characteristics might provide greater gains than in traits typically highly environment-dependent such as yield and plant height. It is within the traits exhibiting the greatest G or GE variation that breeders can most efficiently identify and exploit variation and maximize performance for each environment or mega-environment. For traits in which a large amount of variation is attributed to GE, there are two likely possibilities: 1) the existence of a discriminating, or vastly different environment or range of environments, or 2) traits inherent to a group of cultivars that

respond positively or negatively to a stimulus in the environments tested. If the environments influencing the relatively high GE interactions can be identified and characterized, a mega-environment, in which cultivar rankings differ from other environments, might be identified and exploited.

GGE Biplot is a graphical analysis tool that produces a two-dimensional biplot based upon G and GE information; therefore, only variables that were significant for G or GE at $\alpha=0.05$ were suitable for analysis in GGE Biplot. Table 2.1 lists all variables and their level of significance; all variables were significant for either G or GE indicating that analysis in GGE Biplot was appropriate. In this study, plant height, height to node ratio, LP, DS, GTEX, UHM, and yield were analyzed.

2.3.2 Biplot Interpretation

There were significant GE interactions for both plant height and number of nodes, but not for height to node ratio; however, there was a significant G main effect for height to node ratio thus justifying analysis in GGE Biplot (Table 2.1). Figure 2.1, the mean vs. stability coordination biplot for plant height, is a two-dimensional graphical representation of a multi-environment data set with principal component (PC) 1 and PC2, which are unit-less measures, on the x- and y-axis, respectively. For this data set, PC1 and PC2 accounted for 92% of the total variation in G and GE, which suggests that this biplot is a good approximation of mean performance and stability (Figure 2.1). The genotypes are written in lower-case italics and the environments are written in upper-case bold lettering. The average environment, defined by the average PC1 and PC2 scores across all environments and denoted by the circle near W1, is bisected by a line with a single arrow that passes through the biplot origin, the average-tester axis (ATA). The

direction of the arrow on the ATA indicates higher values for the variable measured, in this case plant height, such that moving along the line in the direction of the arrow indicates taller plants. The ten dotted lines are unit-less measures and exist only to rank or evaluate the cultivars for mean performance. The line containing an arrow at each end, called the stability line, which runs perpendicular to the ATA and also passes through the biplot origin, indicates the stability of any given cultivar. A longer projection from a genotype onto the stability line, or an increasing distance from the ATA, indicates a greater tendency for GE interactions for that genotype. In contrast, genotypes with a short projection onto the stability line and clustered on or near the ATA would be highly stable and perform consistently across environments. GGE Biplot also computes a stability statistic for each cultivar, which is interpreted such that cultivars with greater absolute values are less stable and cultivars that have lesser absolute values closer to zero are highly stable (Table 2.2).

2.3.3 Plant Height

Figure 2.1 indicates that the recent STV474 derivatives, 4793R, 4691B, and 4892BR, were the most stable cultivars for plant height, even more so than the conventional parent. In contrast, the first commercially-available transgenic cultivar, STVBXN47, and an older SG501 derivative, SG501BR, were the least stable cultivars for plant height. It is possible that the latter transgenic cultivars have been selected more rigorously and over a broader range of environments than the earlier transgenic releases, thereby increasing their stability. Using the biplot to view the genotype main effect, it is apparent that in all cases but one, that the transgenic derivatives were taller than their conventional parents, regardless of their level of stability, at 60 DAP (Figure 2.1; Table 2.2).

Table 2.2. Mean cotton plant height (inches) by and across seven environments[†] and stability statistic[‡].

Genotype	Environment							Mean	Stability Statistic
	A1	A2	B1	S1	S2	W1	W2		
SG501	27	31	38	30	35	32	26	32	-0.404
SG501BR	28	32	40	36	32	35	28	33	1.38
STV4691B	27	33	40	33	39	35	26	33	-0.044
STV474	26	32	39	32	38	33	27	32	0.176
STV4793R	27	32	38	31	37	33	26	32	-0.143
STV4892BR	28	31	42	34	35	36	27	33	0.051
STVBXN47	30	33	43	32	36	36	27	34	-1.03
Mean	28	32	40	32	36	34	27	33	na

[†] Where: A=Alexandria, B=Bossier City, S=St. Joseph, W=Winnsboro, 1=2001, and 2=2002.

[‡] GGE stability statistic; a higher absolute value suggests greater instability.

2.3.4 Plant Height to Node Ratio

The mean vs. stability coordination biplot for plant height to node ratio explained 85% of the total variation with PC1 and PC2 (Figure 2.2). Figure 2.2 indicates that the two least stable cultivars, SG501BR and STV4892BR, are the two cultivars with the highest average height to node ratio, and that their conventional recurrent parents are more stable, although their height to node ratio is lower (Table 2.3). This suggests that SG501BR and STV4892BR had a very high height to node ratio in some environments, but were subject to rank changes over all seven environments. In fact, SG501BR, which is located near S1

and W2 in the biplot, had the highest height to node ratio in those environments, but not in all environments, hence the instability of this cultivar (Figure 2.2; Table 2.3).

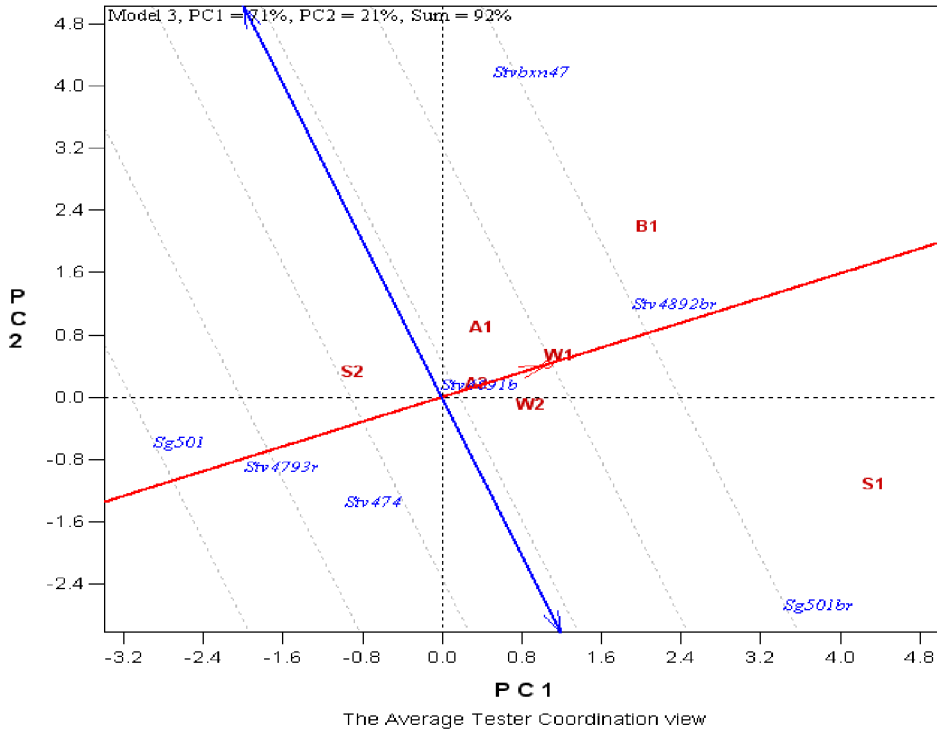


Figure 2.1. Mean vs. stability coordination biplot for cotton plant height at 60 days after planting.

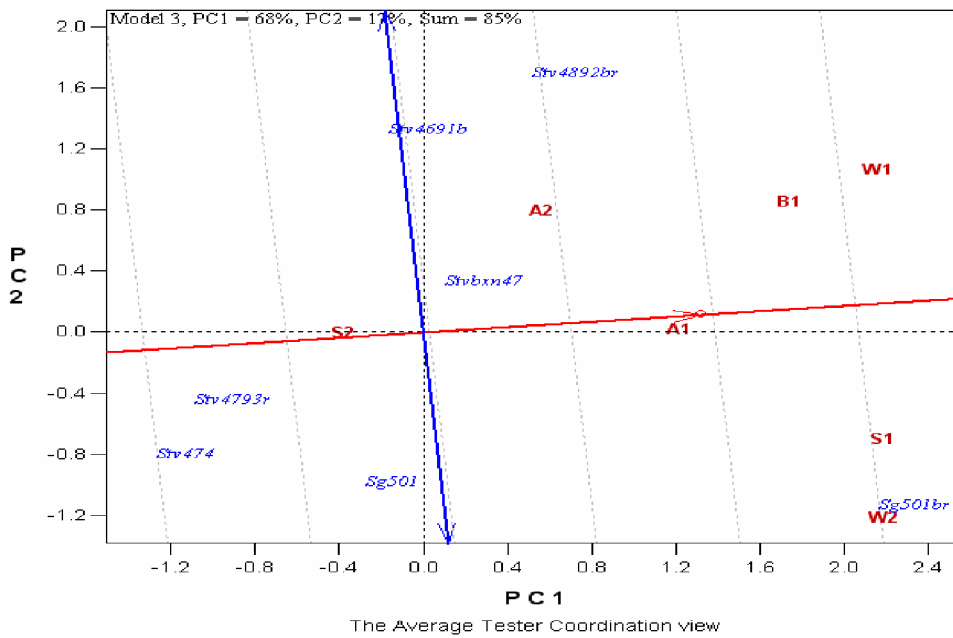


Figure 2.2. Mean vs. stability coordination biplot for cotton height to node ratio at 60 days after planting.

Table 2.3. Mean cotton plant height to node ratio by and across seven environments[†] and stability statistic[‡].

Genotype	Environment							Mean	Stability Statistic
	A1	A2	B1	S1	S2	W1	W2		
SG501	1.95	2.14	2.38	2.41	2.27	2.62	2.03	2.31	0.479
SG501BR	2.17	2.20	2.60	2.66	2.15	2.81	2.26	2.41	0.780
STV4691B	1.92	2.27	2.56	2.41	2.30	2.68	1.85	2.28	-0.697
STV474	1.92	2.14	2.40	2.38	2.24	2.41	1.83	2.19	0.298
STV4793R	2.03	2.08	2.29	2.34	2.23	2.59	1.84	2.20	0.132
STV4892BR	2.11	2.29	2.52	2.41	2.17	2.83	1.90	2.32	-0.829
STVBXN47	2.06	2.11	2.55	2.45	2.15	2.66	1.89	2.27	-0.163
Mean	2.02	2.18	2.46	2.43	2.22	2.66	1.94	2.28	na

[†] Where: A=Alexandria, B=Bossier City, S=St. Joseph, W=Winnsboro, 1=2001, and 2=2002.

[‡] GGE stability statistic; a higher absolute value suggests greater instability.

A similar situation exists for STV4892BR, which had the highest height to node ratio in A2 and W1, but one of the lowest in S2 (Figure 2.2; Table 2.3). Interestingly, the most unstable cultivars, SG501BR, STV4892BR, and STV4691B, were the only cultivars that contained the Bollgard gene, yet they had higher than average height to node ratios. Jones et al., (1996) found that NuCOTN33B was taller and had improved seedling vigor and a greater height to node ratio than its conventional, recurrent parent, DP5415. Considering the insect tolerance imparted by the Bollgard gene, it is reasonable to assume that the relative differences among cultivars might be accentuated over a range of environments in which insect pressure varies, thereby increasing the GE interaction, or

instability, of the cultivars containing insect tolerance while maintaining a high average height to node ratio regardless of the level of insect pressure.

2.3.5 Lint Percentage

There was a significant G main effect for LP and E, G, and GE accounted for 28, 38, and 34% of the total variation in the sums of squares (Table 2.1). Kerby et al., (2000) also found that variation in the sums of squares for LP was influenced more by G and GE than E. Figure 2.3 shows that, averaged across environment, all conventional recurrent parents had higher LP than their transgenic derivatives, but were less stable. The two least stable varieties for LP, evidenced by the greatest stability statistics of 1.579 and -0.763, were the conventional parents SG501 and STV474, respectively (Table 2.4).

Table 2.4. Mean cotton lint percentage by and across seven environments[†] and stability statistic[‡].

Genotype	Environment							Mean	Stability Statistic
	A1	A2	B1	S1	S2	W1	W2		
SG501	40.2	47.3	39.7	42.0	41.1	43.3	41.5	42.0	1.579
SG501BR	36.7	42.0	37.8	40.3	38.8	32.2	39.5	38.2	0.151
STV4691B	39.4	41.2	40.3	42.9	42.6	44.8	43.4	42.1	-0.381
STV474	40.0	40.2	41.0	43.2	42.7	39.2	43.2	41.4	-0.763
STV4793R	39.6	41.0	40.5	42.2	41.7	37.7	41.8	40.6	-0.102
STV4892BR	39.8	40.4	41.3	41.6	41.7	38.6	41.5	40.7	0.243
STVBXN47	39.8	40.8	40.6	42.8	42.0	34.1	42.5	40.4	-0.727
Mean	39.4	41.8	40.1	42.1	41.5	39.0	41.9	40.8	na

[†] Where: A=Alexandria, B=Bossier City, S=St. Joseph, W=Winnsboro, 1=2001, and 2=2002.

[‡] GGE stability statistic; a higher absolute value suggests greater instability.

SG501BR was much more stable than SG501, but had the lowest LP in the test (Figure 2.3).

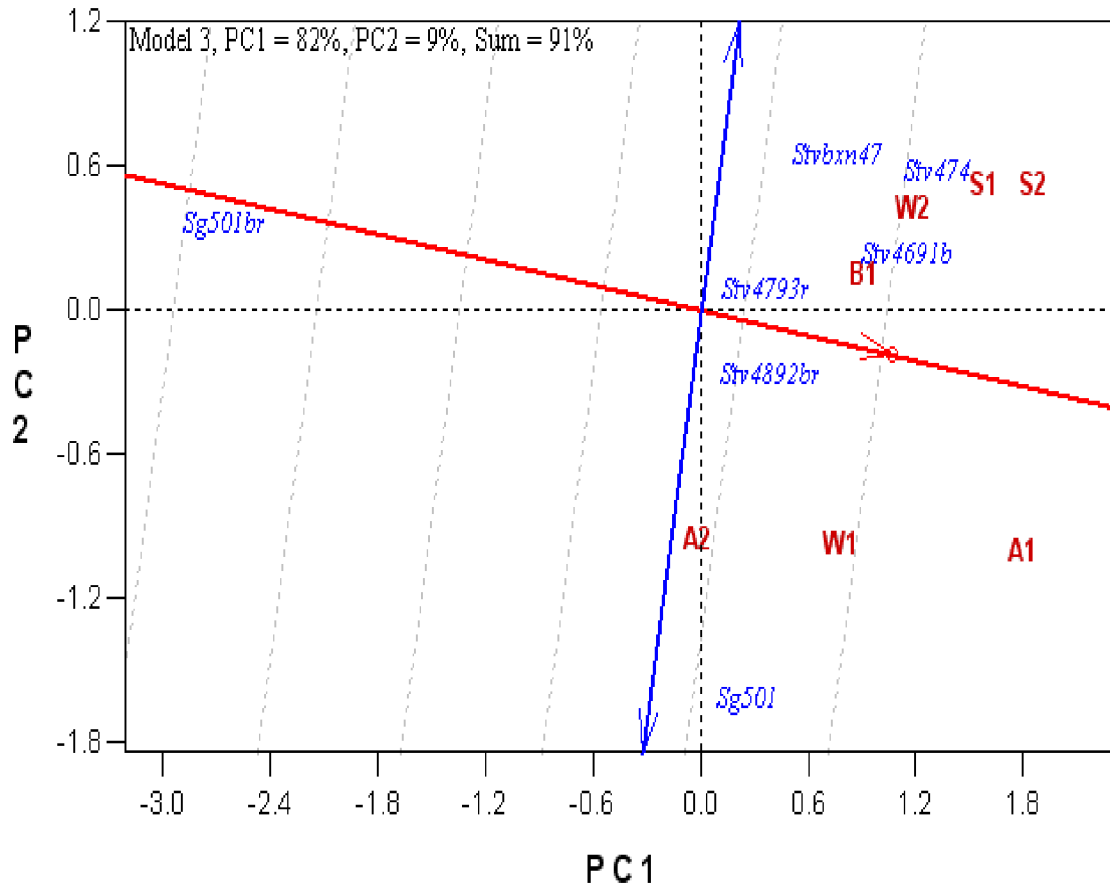


Figure 2.3. Mean vs. stability coordination biplot for cotton lint percentage.

The recent STV474 transgenic derivatives STV4691B, 4793R, and 4892BR were more stable than their conventional parent and had instability values of 0.381, -0.102, and 0.243, respectively, compared to the stability value of 0.763 of STV474 (Table 2.4). These data show that transgenic cultivars have a lower LP, but are more stable than their conventional parents.

2.3.6 Delinted Seed Index

All sources of variation were significant for DS and E, G, and GE contributed 11, 55, and 33%, respectively, to the total variation in sums of squares (Table 2.1). Instability values ranged from -0.857 for STV4793R to 0.735 for SG501 and there were no apparent stability trends for conventional or transgenic cultivars regarding DS (Table 2.5).

Table 2.5. Mean delinted seed index of cotton by and across seven environments[†] and stability statistic[‡].

Genotype	Environment							Mean	Stability Statistic
	A1	A2	B1	S1	S2	W1	W2		
SG501	7.70	8.28	8.65	8.73	8.38	7.88	8.20	8.32	0.714
SG501BR	8.63	8.78	8.73	9.23	9.19	8.27	8.48	8.76	0.735
STV4691B	9.00	8.20	9.63	9.17	9.10	8.90	8.66	8.95	0.164
STV474	8.80	7.83	8.90	8.93	8.81	7.33	8.75	8.48	-0.513
STV4793R	9.20	7.93	8.93	9.27	9.79	8.63	9.18	8.99	-0.857
STV4892BR	9.90	8.47	9.03	9.63	9.57	8.27	8.67	9.08	0.415
STVBXN47	8.37	7.86	8.90	8.37	8.46	7.87	8.68	8.36	-0.658
Mean	8.80	8.19	8.93	9.01	9.04	8.14	8.66	8.69	na

[†] Where: A=Alexandria, B=Bossier City, S=St. Joseph, W=Winnsboro, 1=2001, and 2=2002.

[‡] GGE stability statistic; a higher absolute value suggests greater instability.

However, except for STVBXN47, transgenic cultivars had larger seed than their respective conventional parents averaged across environment (Figure 2.4; Table 2.5). In a study conducted by Jones et al., (1996), seed size was 9% greater for NuCOTN33B compared to DP5415. Summary results of the gin and in-season data show that transgenic cultivars had lower LP, larger seed, and taller, more vigorous plants than their conventional parents.

2.3.7 Fiber Length and Strength

Sources E and G were significant for GTEX and UHM (Table 2.1). The percentage of total variation explained by E, G, and GE was 56, 27, and 17%, respectively, for GTEX and 57, 19, and 23% for UHM, respectively. Kerby et al., (2000) also found that, in each of two studies, G contributed more to the total variation in GTEX compared to UHM.

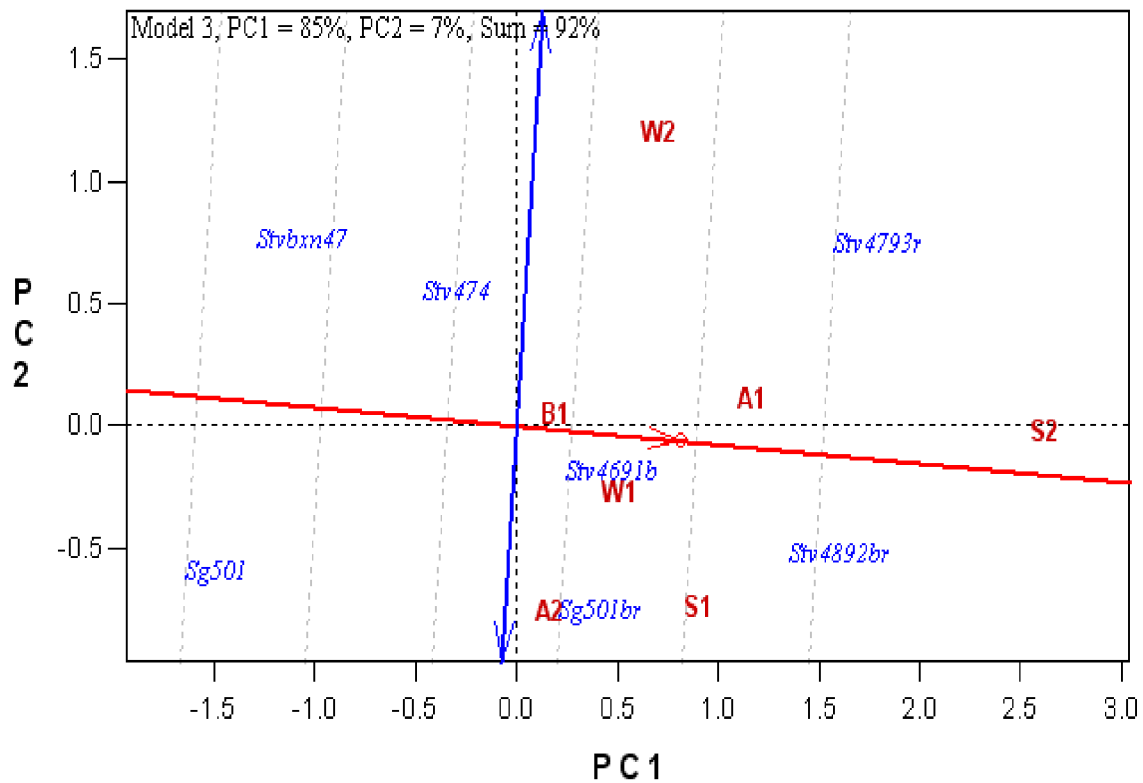


Figure 2.4. Mean vs. stability coordination biplot for delinted seed index of cotton.

There was no clear trend regarding differences between conventional and transgenic cultivars with respect to stability for these two fiber characteristics (Tables 2.6 and 2.7). SG501 was highly stable and had greater values for GTEX and UHM than any other cultivar; STV474 was less stable than any transgenic derivative, and had intermediate values for UHM and GTEX.

Table 2.6. Mean cotton fiber strength by and across seven environments[†] and stability statistic[‡].

Genotype	Environment								Stability Statistic
	A1	A2	B1	S1	S2	W1	W2	Mean	
SG501	30.9	28.3	33.5	33.6	34.8	33.0	36.9	33.1	-0.029
SG501BR	28.1	27.5	28.6	31.1	31.4	32.2	33.4	30.3	-1.254
STV4691B	27.8	28.4	26.5	29.7	31.0	29.9	32.0	29.3	-0.222
STV474	28.6	28.7	29.4	29.6	31.9	29.2	34.0	30.2	1.060
STV4793R	28.8	29.0	28.6	30.5	32.5	30.9	32.7	30.4	-0.032
STV4892BR	29.3	28.0	30.3	30.3	32.9	31.0	33.7	30.8	0.360
STVBXN47	28.1	28.7	29.1	29.7	30.9	30.0	32.5	29.9	0.116
Mean	28.8	28.4	29.9	31.0	32.2	31.1	33.6	30.7	na

[‡] GGE stability statistic; a higher absolute value suggests greater instability.

Table 2.7. Mean cotton fiber length by and across seven environments[†] and stability statistic[‡].

Genotype	Environment								Stability Statistic
	A1	A2	B1	S1	S2	W1	W2	Mean	
SG501	1.12	1.09	1.11	1.13	1.15	1.07	1.11	1.11	-0.397
SG501BR	1.07	1.12	1.08	1.11	1.11	1.07	1.09	1.09	-0.403
STV4691B	1.11	1.10	1.07	1.13	1.15	1.08	1.11	1.11	-0.672
STV474	1.12	1.13	1.07	1.11	1.13	1.03	1.13	1.10	1.255
STV4793R	1.08	1.13	1.06	1.08	1.10	1.07	1.09	1.09	-0.022
STV4892BR	1.11	1.10	1.08	1.11	1.14	1.07	1.10	1.10	-0.347
STVBXN47	1.12	1.12	1.08	1.11	1.13	1.05	1.11	1.10	0.585
Mean	1.10	1.11	1.08	1.11	1.13	1.06	1.11	1.10	na

[†] Where: A=Alexandria, B=Bossier City, S=St. Joseph, W=Winnsboro, 1=2001, and 2=2002.

[‡] GGE stability statistic; a higher absolute value suggests greater instability.

Other studies have yielded similar results in comparisons between conventional and transgenic cultivars (Jones, et al., 1996; Culpepper and York, 1998; Silvertooth and Norton, 1998; Ethridge and Hequet, 2000; Cooke et al., 2001; Moser et al., 2001; Robinson and McCall, 2001; Jordan et al., 2003). In general, their studies showed that differences in fiber quality between conventional and transgenic cultivars were minimal or non-existent and that it was difficult to conclude that a particular transgenic trait resulted in poorer fiber quality in the transgenic cultivars.

2.3.8 Yield

All sources of variation were significant for yield (Table 2.1). E, G, and GE contributed 92, 3, and 5% to the total variation in yield, respectively. Previous studies have shown E to be the predominant source of variation in lint yield (McPherson and Gwathmey, 1996; Myers and Bordelon, 1997; Kerby et al., 2000). In Figure 2.5, the mean vs. stability coordination biplot for lint yield, 81% of the total variation was explained by PC1 and PC2. STV4691B was the highest yielding cultivar in the study and was highly stable (Figure 2.5). Across environments, STV4691B yielded between 205 and 395 kg ha⁻¹ more than all other cultivars and was always ranked in the top three cultivars, hence the high level of stability (Tables 2.8 and 2.9). Across environment, the three highest yielding cultivars contained the Bollgard gene despite variable stability rankings of 2, 5, and 7 (Figure 2.5, Table 2.9). In this study, the only clear indication of differences in mean lint yield or stability between conventional or transgenic cultivars was that cultivars containing the Bollgard gene yielded more, and that transgenic cultivars containing the Roundup Ready or BXN technology yielded less, than their conventional parents, regardless of their level of stability.

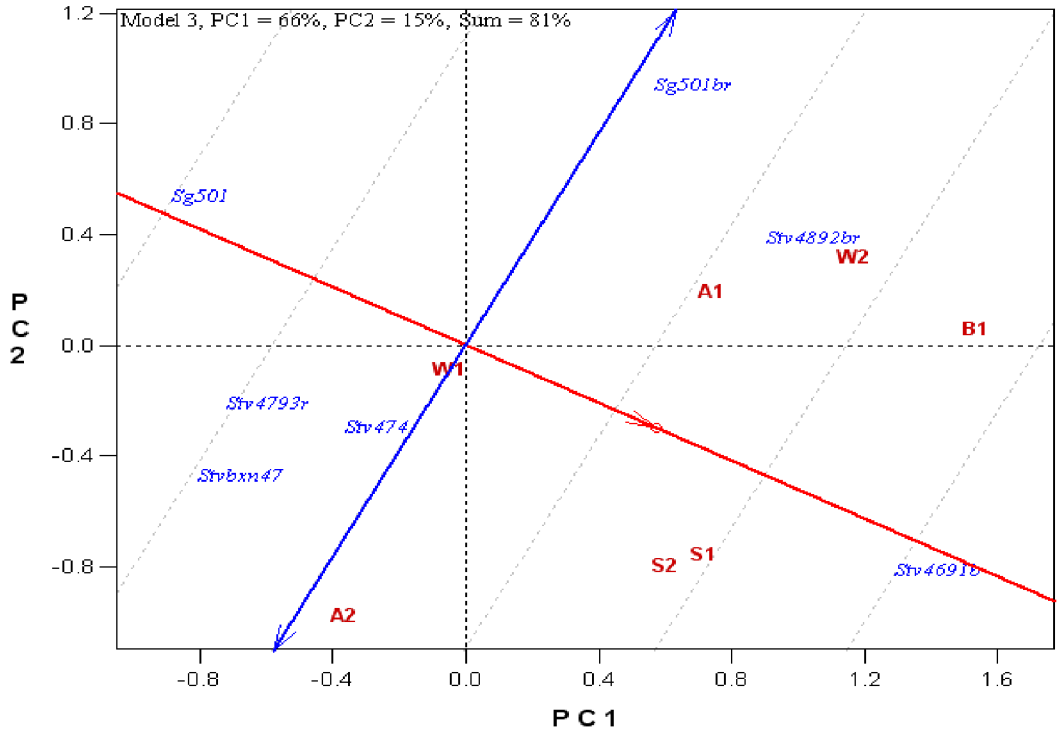


Figure 2.5. Mean vs. stability coordination biplot for cotton yield.

In a similar study, Moser, et al., (2001) found that six of nine Bollgard varieties and four of ten Bollgard/Roundup Ready varieties yielded significantly higher than their conventional parents. However, they also stated that three of ten Roundup Ready varieties yielded similar to or less than their conventional parents indicating that not all transgenes and insertion events affect yield equally. Similarly, Jordan et al., (2003) showed consistent yield advantages for transgenic cultivars containing the Bollgard gene, whether alone or stacked. Other studies have shown little difference in mean lint yield between conventional and transgenic cultivars (Silvertooth and Norton, 1998; Robinson and McCall, 2001).

2.4 Discussion

We hypothesized that non-target traits can be affected by transgene insertion and that the differences between conventional and transgenic cultivars are affected according to the environment in which they are grown. The results of this study indicate that GE interactions are more likely to occur for some traits and that the percentage of total variation attributed to E, G, or GE differs among traits. There were differences in the level of stability between conventional and transgenic cultivars according to the trait analyzed and transgenic technology.

Table 2.8. Mean cotton yield (lb/A) by and across seven environments[†] and stability statistic[‡].

Genotype	Environment							Mean	Stability Statistic
	A1	A2	B1	S1	S2	W1	W2		
SG501	1350	880	2537	2670	2496	1205	1760	1956	-0.06
SG501BR	1603	767	3067	2971	2243	1172	2365	2027	-1.089
STV4691B	1690	984	3271	3206	3084	1189	2400	2261	0.132
STV474	1124	1011	2766	3045	2448	1189	1960	1935	0.44
STV4793R	1150	976	2574	3088	2304	1176	1799	1867	0.532
STV4892BR	1446	802	3202	3206	2317	1128	2444	2078	-0.753
STVBXN47	1333	1202	2605	3027	2200	1237	1751	1908	0.8
Mean	1385	946	2820	2986	2442	1187	2068	2002	na

[†] Where: A=Alexandria, B=Bossier City, S=St. Joseph, W=Winnsboro, 1=2001, and 2=2002.

[‡] GGE stability statistic; a higher absolute value suggests greater instability.

Table 2.9. Cotton yield and stability ranking[†] by and across seven environments[‡].

Genotype	Environment							Mean [#]	Stability Ranking
	A1	A2	B1	S1	S2	W1	W2		
SG501	4	5	7	7	2	2	6	5	1
SG501BR	2	7	3	6	6	6	3	6	7
STV4691B	1	3	1	1	1	3	2	1	2
STV474	7	2	4	4	3	4	4	3	3
STV4793R	6	4	6	3	5	5	5	7	4
STV4892BR	3	6	2	2	4	7	1	2	5
STVBXN47	5	1	5	5	7	1	7	4	6

[†] Ranked according to GGE stability statistic; a higher value suggests greater instability and lower rank.

[‡] Where: A=Alexandria, B=Bossier City, S=St. Joseph, W=Winnsboro, 1=2001, and 2=2002.

[#] Average yield ranking across environments.

We believe that transgenic cultivars, particularly those containing the Bollgard gene, are buffered against variable insect pressure and exhibit a higher level of mean performance and stability as a benefit of that tolerance. Yield increases in Bollgard varieties can be explained by the season-long insect control inherent to that variety, even when insects are present at below-threshold populations and slight yield reductions occur in conventional parents. The level of insect pressure across environments determines the extent of the yield increase and GE interaction. A similar situation might exist for Roundup Ready varieties if a comparable level of weed control was not provided via other chemistries or methods.

Transgenic cultivars were taller, had a greater height to node ratio, had larger seed, and lower LP than their conventional recurrent parents. Large-seeded genotypes generally have lower LP than small-seeded genotypes. It is logical that similarity in seed size to the recurrent parent was not a selection criteria for these transgenic cultivars, and several non-target traits were affected by the differences. This could be due to the fact that transgenic lines are selected for similarities to the recurrent parent and for yield, with yield being the most important characteristic.

2.5 Literature Cited

- Baril, C. P., J. B. Denis, R. Wustman, and F. A. van Eeuwijk. 1995. Analyzing genotype-by-environments for selection and recommendation of common wheat genotypes in Italy. *Plant Breed.* 113:197-205.
- Benedict, J. H., and D. W. Altman. 2001. Commercialization of transgenic cotton expressing insecticidal crystal protein. p. 137-201. *In: J. N. Jenkins and S. Saha (ed.) Genetic Improvement of Cotton.* Science Publ. Enfield, NH.
- Bowman, D.T. 2000. Attributes of public and private cotton breeding programs. *J. Cotton Sci.* 4:130-136.
- Collins, J. R. 1996. BXN Cotton: Marketing plans and weed control programs utilizing buctril. p. 201. *In Proc. Beltwide Cotton Conf., Nashville, TN. 9-12 Jan. 1996.* Natl. Cotton Counc. Am., Memphis, TN.
- Cooke, F. T., Jr., W. P. Scott, S. W. Martin, and D. W. Parvin. 2001. The economics of *Bt* cotton in the Mississippi Delta 1997-2000. p. 175-177. *In Proc. Beltwide Cotton Conf., Anaheim, CA 9-13 Jan. 2001.* Natl. Cotton Counc. Am., Memphis, TN.
- Culpepper, A. S., and A. C. York. 1998. Weed management in glyphosate-tolerant cotton. *J. Cotton Sci.* 4:174-185.
- Elmore, R.W., F.W. Roeth, R.N. Klein, S.Z. Knezevic, A. Martin, L.A. Nelson, and C.A. Shapiro. 2001a. Glyphosate-resistant soybean cultivar response to glyphosate. *Agron. J.* 93:404-407
- Elmore, R.W., F.W. Roeth, L.A. Nelson, C.A. Shapiro, R.N. Klein, S.Z. Knezevic, and A. Martin. 2001b. Glyphosate-resistant soybean cultivar yields compared with sister lines. *Agron. J.* 93:408-412

Epinat-Le Signor, C. S. Dousse, J. Lorgeou, J.-B. Denis, R. Bonhomme, P. Carolo, and A. Charcosset. 2001. Interpretation of genotype x environment interactions for early maize hybrids over 12 years. *Crop Sci.* 41:663-669.

Ethridge, M. D., and E. F. Hequet. 2000. Fiber properties and textile performance of transgenic cotton versus parent varieties. p. 488-494. *In Proc. Beltwide Cotton Conf.*, San Antonio, TX. 4-8 Jan. 2000. Natl. Cotton Counc. Am., Memphis, TN.

Gasser, C. S., and R. T. Fraley. 1989. Genetically engineering plants for crop improvement. *Science* 244:1293-1299.

Hardee, D. D., and Burris, E. 2001. 54th Annual Conference report on cotton insect research and control. p. 741-773. *In Proc. Beltwide Cotton Conf.*, Anaheim, CA 9-13 Jan. 2001. Natl. Cotton Counc. Am., Memphis, TN.

Hardee, D. D., and Herzog, G. A. 1996. 49th Annual Conference report on cotton insect research and control. p. 643-680. *In Proc. Beltwide Cotton Conf.*, Nashville, TN. 9-12 Jan. 1996. Natl. Cotton Counc. Am., Memphis, TN.

Hardee, D. D., and Herzog, G. A. 1997. 50th Annual Conference report on cotton insect research and control. p. 809-834. *In Proc. Beltwide Cotton Conf.*, New Orleans, LA 7-10 Jan. 1997. Natl. Cotton Counc. Am., Memphis, TN.

Jones, K., T. J. Kerby, H. Collings, T. Wofford, M. Bates, J. Presley, and J. Burgess. 1996. Performance of NuCOTN with Bollgard. p. 46-48. *In Proc. Beltwide Cotton Conf.*, Nashville, TN. 9-12 Jan. 1996. Natl. Cotton Counc. Am., Memphis, TN.

Jordan, A. G., P. J. Wakelyn, and O. L. May. 2003. Transgenic cotton and fiber quality Part 1. Effect of transgenic technology. p. 1-10. *In 16th Ann. EFS Conf.* 2003. Natl. Cotton Counc. Am., Memphis, TN.

Kang, M. S., and H. N. Pham. 1991. Simultaneous selection for high yielding and stable crop genotypes. *Agron. J.* 83:161-165.

Kerby, T., J. Burgess, M. Bates, D. Albers, and K. Lege. 2000. Partitioning variety and environment contribution to variation in yield, plant growth, and fiber quality. p. 528-532. *In Proc. Beltwide Cotton Conf.*, New Orleans, LA 7-10 Jan. 2000. Natl. Cotton Counc. Am., Memphis, TN.

Kerby, T. J., and R. Voth. 1998. Roundup Ready – Introduction experiences in 1997 as discussed in the Beltwide Cotton Production Conference Weed Management: Transgenics and new technologies panel. p. 26-29. *In Proc. Beltwide Cotton Conf.*, San Diego, CA 5-9 Jan. 1998. Natl. Cotton Counc. Am., Memphis, TN.

- Kohel, R.J., J.E. Quisenberry, G. Cartwright, and J. Yu. 2000. Linkage analysis of transgenes inserted into cotton via *Agrobacterium tumefaciens* transformation. *J. Cotton Sci.* 4:65-69.
- Lee, J. A. 1978. Allele determining rugate fruit surface in cotton. *Crop Sci.* 18:251-254.
- Lin, C.S., and M.R. Binns. 1988. A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68:193-198.
- Lukefahr, M. J., and J. E. Houghtaling. 1969. Resistance of cotton strains with high gossypol content to *Heliothis* spp. *J. Econ. Entomol.* 62:588-591.
- Lukefahr, M. J., J. E. Houghtaling, and D. G. Cruhm. 1975. Suppression of *Heliothis* spp. With cotton containing a combination of resistance characters. *J. Econ. Entomol.* 68:743-746.
- Lukefahr, M. J., and D. F. Martin. 1966. Cotton plant pigments as a source of resistance to bollworm and tobacco budworm. *J. Econ. Entomol.* 59:176-179.
- May, O. L., D. T. Bowman, and D. S. Calhoun. 1995. Genetic diversity of U. S. upland cotton cultivars released between 1980 and 1990. *Crop Sci.* 35:1570-1574.
- Meredith, W. R., Jr. 1995. Strengths and limitations of conventional and transgenic breeding. p. 166-168 *In Proc. Beltwide Cotton Conf., San Antonio, TX 4-7 Jan. 1995. Natl. Cotton Counc. Am., Memphis, TN.*
- Meyer, J. R. 1957. Origin and inheritance of D₂ smoothness in upland cotton. *J. Hered.* 48:249-250.
- Meyer, J. R., and V. G. Meyer. 1961. Origin and inheritance of nectariless cotton. *Crop Sci.* 1:167-169.
- McMichael, S. C. 1960. Combined effects of the glandless genes, g₁₂ and g₁₃, on pigment glands in the cotton plant. *Agron. J.* 52:385-386.
- Moser, H. S., W. B. McCloskey, and J. C. Silvertooth. 2001. Performance of transgenic cotton varieties in Arizona. p. 420-423. *In Proc. Beltwide Cotton Conf., Anaheim, CA 9-13 Jan. 2001. Natl. Cotton Counc. Am., Memphis, TN.*
- McPherson, R., and O. Gwathmey. 1996. Yield and stability of cotton cultivars at three West Tennessee locations. p. 596-598. *In Proc. Beltwide Cotton Conf., Nashville, TN 9-12 Jan. 1996. Natl. Cotton Counc. Am., Memphis, TN.*
- Myers, G. O., and F. Bordelon. 1997. Stable statistics for cotton varieties grown in the Midsouth, Southeast, North and Texas from 1993-1995. p. 464-466. *In Proc. Beltwide Cotton Conf., New Orleans, LA 7-10 Jan. 1997. Natl. Cotton Counc. Am., Memphis, TN.*

National Agricultural Statistics Service (NASS). 1994. Upland cotton: Chemical usage by State and United States, 1994. Nat. Agric. Stat. Serv., USDA, Washington, D.C. [online]. Available at <http://usda.mannlib.cornell.edu/reports/nassr/other/pcu-bb/agch0395.txt>

National Agricultural Statistics Service (NASS). 2001. Upland cotton: Chemical usage by State and United States, 2001. Nat. Agric. Stat. Serv., USDA, Washington, D.C. [online]. Available at <http://usda.mannlib.cornell.edu/reports/nassr/other/pcu-bb/agcs0501.txt>

National Agricultural Statistics Service (NASS). 2004. Upland cotton: Biotechnology varieties by state and United States, percent of upland cotton planted, 2003-2004. Nat. Agric. Stat. Serv., USDA, Washington, D.C. (online) Available at <http://usda.mannlib.cornell.edu/reports/nassr/field/pcp-bba/acrg0604.txt>.

Pinthus, M.J. 1973. Estimate of genotypic value: A proposed method. *Euphytica*. 22:121-123.

Raymer, P.L., and H.C. Minor. 1999. Agronomic performance of roundup ready soybean varieties in the southeastern United States. p. 549. *In* H.E. Kauffman (ed.) Proc. World Soybean Res. Conf. VI, Chicago, IL, USA. 4-7 Aug. 1999. Superior Printing, Champaign, IL.

Robinson, M., and L. McCall. 2001. A comparison of transgenic and conventional cotton varieties. p. 419. *In* Proc. Beltwide Cotton Conf., Anaheim, CA 9-13 Jan. 2001. Natl. Cotton Counc. Am., Memphis, TN.

Sachs, E. S., J. H. Benedict, D. M. Stelly, J. F. Taylor, D. W. Altman, S. A. Berberich, and S. K. Davis. 1998. Expression and segregation of genes encoding CryIA insecticidal proteins in cotton. *Crop Sci*. 38:1-11.

Silvertooth, J. C., and E. R. Norton. 1998. Agronomic evaluations of transgenic cotton varieties. [online]. Available at <http://ag.arizona.edu/pubs/crops/az1006/az10063j.html>.

Umbeck, P., W. Swain, and N. S. Yang. 1989. Inheritance and expression of genes for kanamycin and chloramphenicol resistance in transgenic cotton plants. *Crop Sci*. 29:196-201.

Verhalen, L. M., and J. C. Murray. 1970. Genotype by environment interaction study of cotton in Oklahoma. p. 52-54. *In* Proc. Beltwide Cotton Conc., Houston, TX 6-7 Jan. 1970. Natl. Cotton Counc. Am., Memphis, TN.

Yan, W. 2001. GGE Biplot- A Windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agron. J*. 93:1111-1118.

Yan, W., and L.A. Hunt. 2001. Interpretation of genotype x environment interaction for winter wheat yield in Ontario. *Crop Sci.* 41:19-25.

Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnic. 2000. Cultivar evaluation and mega-environment investigation based on the GGE Biplot. *Crop Sci.* 40:597-605.

Zobel, R. W., M. J. Wright, and H. G. Gauch, Jr. 1988. Statistical analysis of a yield trial. *Agron. J.* 80:388-393.

CHAPTER 3

DETERMINING SELECTION GAINS VIA GGE BILOT

3.1 Introduction

Prediction of how a genotype will perform in an environment, or in a cross, is a central issue in plant variety development. In order for cultivars to be successful, they should exhibit a high level of performance for multiple characteristics in numerous environments. Thus, choosing parents that are within acceptable limits for a majority of important characteristics greatly increases the probability of realizing that goal. Maximizing genetic advancement in a plant breeding program has, as its foundation, at least four distinct issues: (1) the number and choice of parental crosses; (2) the trial design and choice of treatments; (3) the number and choice of trial locations; and (4) the intensity and duration of selection (Kempton and Fox, 1997). The focus of this research is to modify some of these processes so important to the development of improved and stable varieties.

Numerous methods exist for choosing parental material to include in a crossing program. Various crossing schemes can prove informative (e.g. topcrosses, diallel), but the limited number of genotypes that can be accommodated and the quality of the data collected has been questioned (Jensen, 1988). Alternative methods such as referencing ancestral records or pedigrees have been used and recently expanded by utilizing measures of genetic similarity based upon molecular genetic marker information. A multivariate approach to selecting parents, based on generalized distances between parents, for use in crossing was proposed by Bhatt (1970). Pederson (1981) outlined a least squares method for identifying parents to be used in a cross and for their proportions

in crosses to be based on a set of parents about which vital information is known. From this information, an ideal value is developed with an acceptable amount of deviation and a linear additive model can be used to determine the best solution based upon crossing members of the set in defined proportions.

The performance of a particular genotype at any one, or a combination, of locations is a function of the grand mean, the genotypic effect, the environmental effect and the interaction between genotype and environment. The identification of superior cultivars for a single or a range of environments has been the focus of GE research. A secondary goal of GE research is to develop a better understanding of target growing regions and the identification of mega-environments (Yan et al., 2000). Using a genotype x environment structure, Yan (2001) developed a methodology for examining GE interactions and identifying mega-environments based upon singular value decomposition of environment-centered data with a primary focus on the first two symmetrically-scaled principal components. Utilizing GGE Biplot, several analyses are possible that are of great interest to plant breeders and cotton producers: (1) cultivar rankings based upon their performance in any given environment, (2) ranking of environments based upon the performance of any given cultivar, (3) grouping of environments based on cultivar performance, (4) evaluating cultivars based on a combination of their average yield and stability, and (5) evaluating environments based on discriminating ability (Yan, 2001).

Estimates of gains by selection in environments varying in discriminating ability can be accomplished through heritability estimates. The heritability of a trait(s) is a dynamic value that can be influenced by the breeder. Factors influencing the magnitude of heritability estimates include: a) the level of genotypic variance in the population

studied, b) the manner in which genotypes are sampled, and c) the method of calculation (Fehr, 1987). The use of discriminating environments to maximize the perceived genotypic variance is one possible way of increasing the heritability of a character. Numerous methods of estimating the heritability of a character exist including parent-offspring regression (Lush, 1940; Fehr, 1987; Frey and Horner, 1957; Robinson, et al., 1949), the variance component method (Fehr, 1987), realized heritability (Falconer, 1981), indirect estimates of environmental variation (Mahmud and Kramer, 1951), and the backcross method (Warner, 1952). One type of parent-offspring regression is the regression of self-pollinated $F_{2:3}$ plants against their parent F_2 since all of the alleles in the selfed $F_{2:3}$ plants are obtained from the parent. In this case, the regression coefficient (b) is equal to the heritability, $b=h^2$ (Fehr, 1987).

The objectives of this study were accomplished primarily through the development of a Regional Breeders Testing Network (RBTN) to encompass a majority of the Upland cotton growing regions in the U.S. The least squares method (Pederson, 1981) was used for identifying superior parental cross combinations, for a single or combination of traits, and GGE Biplot was used to determine highly discriminating locations for each trait or trait combination. The effectiveness of using discriminating environments to maximize breeding progress was accomplished by analyzing the variances of the $F_{2:3}$ generation and gains from selection were calculated using linear regression of selected $F_{2:3}$ plants on the F_2 generation.

3.2 Materials and Methods

Studies were conducted from 2002-2004 to estimate the potential breeding value of cotton cultivars and progress toward breeding objectives. In 2002, three distinct

populations were established: a) one selected for lint yield (POP1), b) one selected for a combination of fiber micronaire, length, strength, uniformity and elongation (POP2), and c) one selected for a combination of lint yield, lint percent, fiber micronaire, length, and strength (POP3). Prior to planting in 2002, data from the multi-location 2001 Mid-South Regional Breeders Tests was used to identify superior parental material. A least squares means analysis (Pederson, 1981) as implemented in Agrobases 20 (Agronomix Software, Inc., Winnipeg, Canada) was used to select the recommended parents for each population. POP1 parents were 'Ark9108-04-17' and 'JAJ08098', POP2 parents were 'LA96110067' and 'LA98404023', and POP3 parents were 'JAJ08098' and 'LA98404028'. During peak-bloom, selected parents were used to make 5 crosses; thus, the seed from 5 crosses comprised the F₁ generation for each population. At harvest, F₁ seed were sent to the Cotton Winter Nursery (Tecoman, Mexico) for generation advancement. The RBTN, 14 locations representing the entire upland cotton growing region, was established in 2002, and GGE Biplot was used to identify two levels of discriminating locations (high and low) for each population (Yan and Kang, 2003). The discriminating ability of each location, specific to the population criteria, was determined by standardizing each genotype's score for a specific trait or combination of traits by the location mean for that trait or trait combination. In populations using multiple traits for selection (POP2 and POP3), the standardized values for all applicable traits were averaged, without weighting, to provide a composite standardized value reflecting that variety's performance for multiple traits. These standardized data from the 2002 RBTN were analyzed by GGE Biplot and for each population, the resulting biplots were used to identify the most and least discriminating locations based on their vector length from the

biplot origin (Yan and Kang, 2003). For all three populations, Bossier City, Louisiana (BC) was determined to be the least discriminating location. St. Joseph, Louisiana (SJ), College Station, Texas (CS), and Starkville, Mississippi (MS) were the most discriminating locations for POP1, POP2, and POP3, respectively.

In 2003, F_2 seed for each selection package were planted in the most and least discriminating locations for that population as determined from the GGE Biplot analysis. One-hundred randomly-selected individual plants were hand-harvested from each F_2 population in 2003 for analysis. Individual plants were processed to obtain gin and fiber data, and were standardized as previously discussed for the criteria for each selection package. Seed from the parents, the top 10%, and the bottom 10% of the individual F_2 plants were planted back in the high and low discriminating locations for their respective selection package in 2004. $F_{2.3}$ seed were planted in progeny rows (40 to 50 ft) and yield data was acquired by machine- or hand-harvesting. Hand-harvesting was accomplished by picking all harvestable bolls from 10 row-feet. Additionally, 25-boll samples were taken to obtain fiber data and lint percentage from each progeny row. Gains from selection were calculated by regressing the $F_{2.3}$ generation against the F_2 generation as per Meredith and Bridge (1973). Additionally, the variances of $F_{2.3}$ progeny were calculated for each location.

3.3 Results and Discussion

3.3.1 Mean and Variance Comparisons

All traits were standardized by the mean for that trait/location combination so multiple trait populations could be established and comparisons between single plants and progeny rows could be made. Table 3.1 shows the mean standardized values for the parents, the

top and bottom 10% of the F₂ population means, and the variances in each location in 2004. The mean of the top 10% was greater than the bottom 10% in four of six locations and two of three populations (Table 3.1).

Table 3.1. Mean standardized values[†] and variances of parents and F_{2:3} cotton plants in six locations in 2004.

Location/Population[‡]	P1	P2	Top 10%	Bottom 10%	Variance
SJ POP1	1.51	0.897	0.992	0.927	9.93 x 10 ⁻²
BC POP1	0.873	1.08	1.07	0.936	1.77 x 10 ⁻²
CS POP2	0.922	0.990	1.04	0.98	1.87 x 10 ⁻³
BC POP2	0.972	0.997	1.02	0.982	1.87 x 10 ⁻³
MS POP3	0.894	1.056	0.992	1.02	9.48 x 10 ⁻³
BC POP3	0.995	0.994	0.990	1.01	2.27 x 10 ⁻³

[†] Data were standardized by the mean of each trait in each location.

[‡] Abbreviations: SJ=St. Joseph, La.; BC=Bossier City, La.; CS=College Station, Tx.; MS=Starkville, Ms.; POP1=population selected for yield; POP2=population selected for fiber micronaire, length, strength, uniformity, and elongation; POP3=population selected for yield, lint percentage, fiber micronaire, length, and strength; P1=Parent 1; P2=Parent 2.

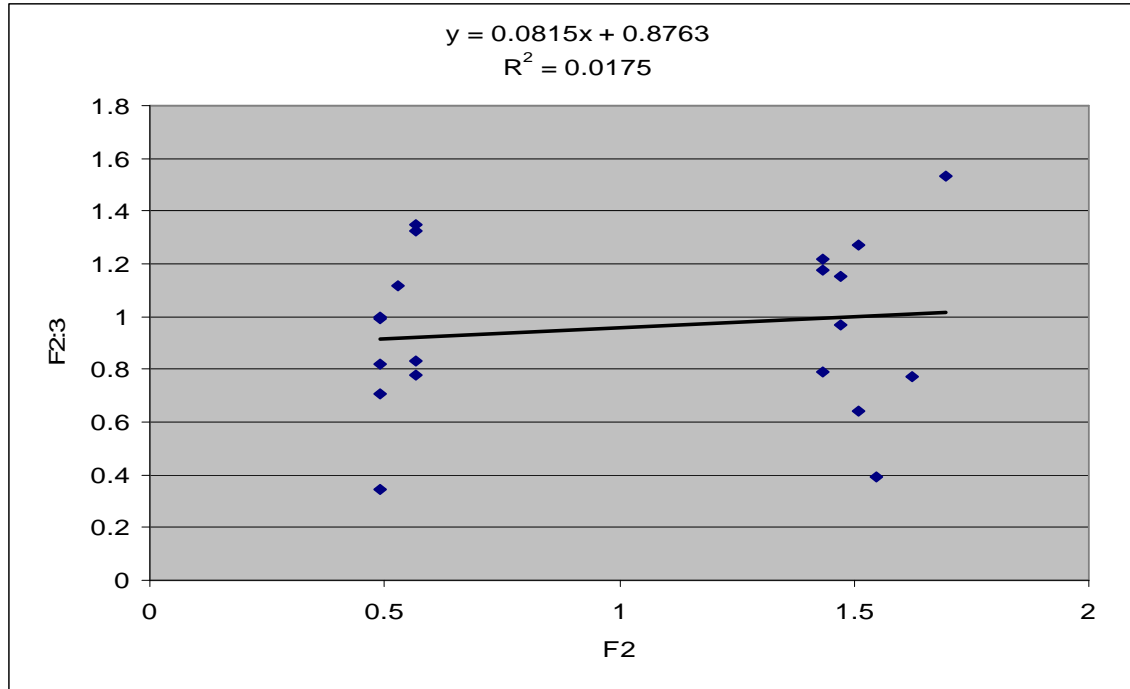
The MS location experienced excess precipitation, in 2003 and 2004 possibly confounding the results in both the F₂ and F_{2:3} generations. The data from that location were included in the study for comparative purposes. For POP1, variance among the F_{2:3} progeny in SJ, the highly discriminating location for yield, was 5.6x that of BC (Table 3.1). The variances among F_{2:3} progeny in CS and BC for POP2 were similar (Table 3.1), possibly due to the fact that the selected traits are largely under genetic control and are less affected by location relative to yield. For POP3, consisting of selections made for yield, lint percentage, fiber micronaire, length, and strength, variance among F_{2:3} progeny in the highly discriminatory MS location was 4.2x that of the BC location (Table 3.1). In

all three populations, the variance among $F_{2:3}$ progeny selected and grown in highly discriminating locations was greater than or equal to those grown in the non-discriminating locations. Variation is the basis for meaningful selection; thus, the use of highly discriminating environments for a trait or trait package should expedite and allow the selection of promising genotypes with increased efficiency. In BC in 2003 there was a stand failure that prevented the F_2 populations from being harvested; thus, the top 10% and bottom 10% selections were made in the respective highly discriminating locations and planted in BC in 2004. We expect that the differences between the variances would be even larger if the $F_{2:3}$ progeny planted in BC in 2004 had been selected there in 2003 rather than in MS.

3.3.2 Heritability Estimates

A measure of heritability for multiple traits was obtained by regressing selected $F_{2:3}$ plants against the selected F_2 plants for all populations. In theory, selection in the F_2 generation for traits with traditionally low heritability, particularly trait packages, should be optimized in a highly discriminating environment. This is due to the fact that in such a location, a greater proportion of F_2 selections are made on genetic differences rather than often-obscuring environmental differences and the associated genes can be transferred to subsequent generations.

For multi-trait populations, F_2 and $F_{2:3}$ data were standardized by the mean of each trait/location combination to yield a standardized value for the traits of interest for each entry.



† Population 1 selected for yield in St. Joseph, La., the highly discriminating location for population 1.

Figure 3.1. Regression of selected F_{2:3} cotton plants on F₂ cotton plants for Population 1[†].

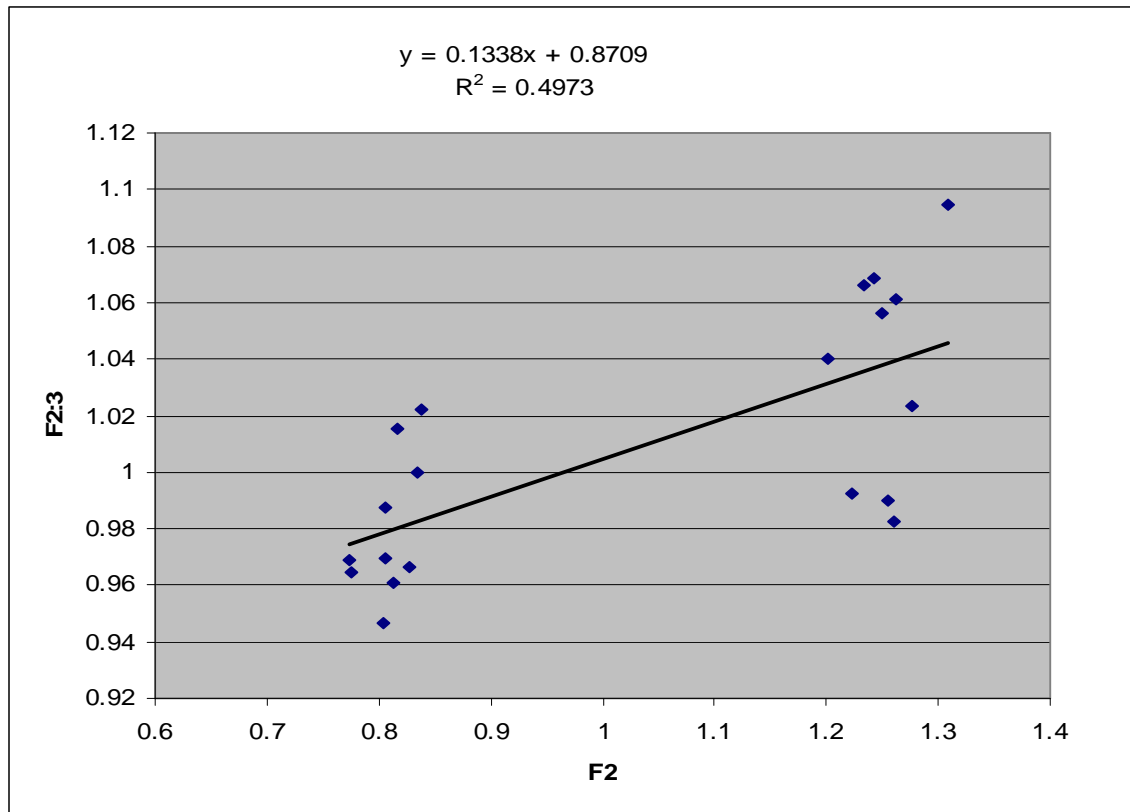
For POP1, the slope of the regression line (b) was 0.082, indicating an 8.2% gain in the F_{2:3} over the F₂ generation (Figure 3.1; Table 3.2). Similarly, Meredith and Bridge (1973) found a 5.7% response to selection for yield compared with random and non-yield selections from an F₂ population.

Table 3.2. Linear correlation (r) and regression coefficients (b) between the F_{2:3} and the F₂ generations in cotton.

Coefficient	POP1 [†]	POP2	POP3
r	0.18	0.50	0.19
b	0.082	0.13	-0.06

† Abbreviations: POP1=population selected for yield; POP2=population selected for fiber micronaire, length, strength, uniformity, and elongation; POP3=population selected for yield, lint percentage, fiber micronaire, length, and strength.

For POP2, the linear correlation (r) was much higher (0.50) than for populations evaluating yield and the regression coefficient was 0.13 (Figure 3.2; Table 3.2).

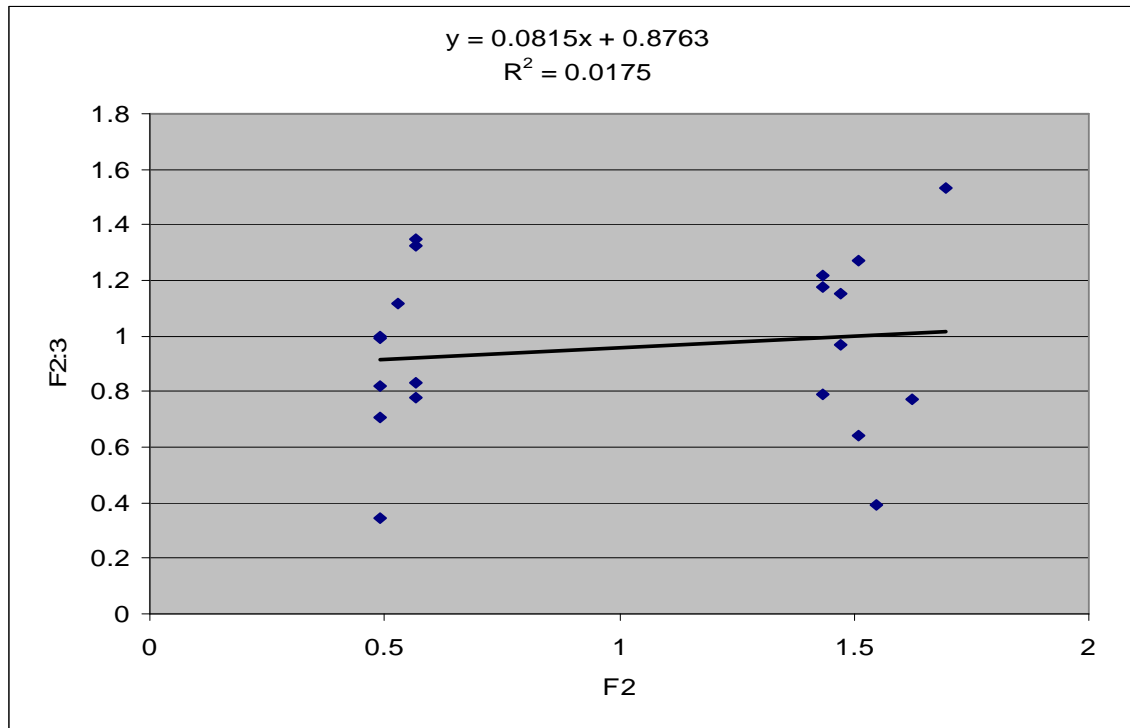


[†] Population 2 selected for fiber micronaire, length, strength, uniformity, and elongation in College Station, Tx., the highly discriminating location for population 2.

Figure 3.2. Regression of selected $F_{2:3}$ cotton plants on F_2 cotton plants for Population 2[†].

The higher regression coefficient associated with POP2, relative to yield containing populations 4 and 5, indicates an increased efficiency, or higher heritability, of transferring multiple fiber quality genes. Meredith and Bridge (1973) regressed selected F_3 plants on the F_2 generation and determined, based on the r and b , that fiber properties (lint percentage, fiber length, strength, and elongation) were more heritable than yield, boll size, seed size, and fiber micronaire. A lower heritability for micronaire compared to other fiber traits is expected because micronaire is heavily dependent on environment.

The heritability of the selection package for yield, lint percentage, fiber micronaire, length, and strength in POP3 was negative ($b=-0.06$) (Figure 3.3; Table 3.2).



† Population 3 selected for yield, lint percentage, fiber micronaire, length, and strength in Starkville, Ms., the highly discriminating location for population 3.

Figure 3.3. Regression of selected F_{2:3} cotton plants on F₂ cotton plants for Population 3 †.

As previously mentioned, MS, the highly discriminating location for POP3, experienced excessive rainfall in 2003 and 2004 and both the F₂ selections and F_{2:3} data could have been affected.

The results of this study show that GGE Biplot can be used to effectively evaluate the discriminating ability of locations giving breeders the ability to make accurate selections from multi-trait F₂ populations. The greater variances observed in discriminating environments allows for the more accurate assessment of genotypic worth; thus, the efficiency of selection is increased facilitating the identification of superior genotypes

and the culling of poor-performing genotypes with a high level of confidence. The regression data was rather limited due to the stand loss in BC, the least discriminating environment for all populations. Theoretically, regression coefficients for each population would have been lower in BC than in the corresponding highly-discriminatory environment. When using standardized data, discriminating environments can be identified for a single or multiple traits encompassing both high and low heritability characteristics, and heritability, i.e., realized gains by selection, can be improved.

3.4 Literature Cited

- Bhatt, G. M. 1970. Multivariate analysis approach to selection of parents for hybridization aimed at yield improvement in self-pollinated crops. *Aust. J. Agric. Res.* 21:1-7.
- Falconer, D. S. 1981. *Introduction to quantitative genetics*. 2nd ed. Longman, New York.
- Fehr, W. R. 1987. Heritability. p. 96-102. *In* W. R. Fehr (ed.) *Principles of cultivar development*. Volume 1, Theory and Technique. Macmillan Publishing Company, New York, N.Y.
- Frey, K. J., and T. Horner. 1957. Heritability in standard units. *Agron. J.* 49:59-62.
- Jensen, N. F. 1988. *Plant breeding methodology*. p. 676. John Wiley & Sons, N.Y.
- Kempton, R. A., and P. N. Fox. 1997. Introduction. p. 191. *In* R. A. Kempton and P. N. Fox (eds.) *Statistical methods for plant variety evaluation*. Chapman & Hall, London.
- Lush, J. L. 1940. Intra-sire correlations or regressions of offspring on dam as a method of estimating heritability of characteristics. *Proc. Am. Soc. An. Prod.* 1:293-301.
- Mahmud, I., and H. H. Kramer. 1951. Segregation for yield, height, and maturity following a soybean cross. *Agron. J.* 43:605-609.
- Meredith, W. R. Jr., and R. R. Bridge. 1973. The relationship between F₂ and selected F₃ progenies in cotton (*Gossypium hirsutum* L.). *Crop Sci.* 13:354-356.
- Pederson, D. G. 1981. A least-squares method for choosing the best relative proportions when intercrossing cultivars. *Euphytica* 30:153-160.

Robinson, H. F., R. E. Comstock, and P. H. Harvey. 1949. Estimates of heritability and the degree of dominance in corn. *Agron. J.* 41:353-59.

Warner, J. N. 1952. A method for estimating heritability. *Agron. J.* 5:595-596.

Yan, W. 2001. GGE Biplot- A Windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agron. J.* 93:1111-1118.

Yan, W. and M. S. Kang. 2003. GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL.

Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnic. 2000. Cultivar evaluation and mega-environment investigation based on the GGE Biplot. *Crop Sci.* 40:597-605.

CHAPTER 4

GGE BILOT VS. TRADITIONAL STABILITY MEASURES

4.1 Introduction

Plant breeders have recently been introduced to GGE Biplot, a stability analysis and variety selection tool, and have begun to incorporate it into their breeding programs (Yan, 2001; Blanche et al., 2002; Myers, 2002; Lubbers, 2003). Research focusing on stability, or genotype (G) x environment (E) interactions, is necessary for plant breeders to develop cultivars that respond optimally and consistently across environments. GE interactions are said to exist when the responses of two genotypes to different levels of environmental stress fail to be parallel (Allard and Bradshaw, 1964). Numerous tools have been developed to measure the response of genotypes to changes in environment (Wricke, 1962; Eberhart and Russell, 1966; Shukla, 1972; Gauch, 1988; Lin and Binns, 1988). However, GGE Biplot offers breeders a more complete and visual evaluation of all aspects of GE interactions by creating a biplot that simultaneously represents both mean performance and stability (Yan, 2001). Widespread acceptance of GGE biplot for its ability to evaluate mean performance and stability, and to identify mega-environments has created a need for research to compare GGE biplot to certain “traditional” stability analysis tools.

The measured performance of a genotype is the result of its genetic make-up, the environment in which the genotype is grown, and the interaction between the genotype and environments. Studies have shown that the environment in which the genotype is grown is the largest source of variation based on the contribution to total sum of squares (Verhalen and Murray, 1970; Kerby et al., 2000; Blanche et al., 2002). In many cases,

the overwhelming effect of environmental variation makes it difficult to glean meaningful information on G and GE interaction. While some classical stability analysis tools, such as ecovalence (Wricke, 1962) and Shukla's stability variance statistic (Shukla, 1972) effectively quantify GE interaction, they do not provide the evaluator with a statistic that includes both mean performance and stability. Shukla's stability variance statistic (σ_i^2) generates values that are estimates of the i^{th} genotype's variance across environments (Shukla, 1972). Shukla (1972) proposed a model that uses an environmental covariate (s_i^2), e.g., difference between the mean of all genotypes at a location ($X_{.j}$) and overall (grand) mean ($X_{..}$), to remove its linear effect so that remaining GE interaction variance (s_i^2) can be attributed to cultivars. Kang proposed methods to integrate yield and stability using Shukla's σ_i^2 (Kang, 1988; Kang 1993).

Others have used a regression model of genotype means on environmental means to model the GE interaction by estimating a set of straight lines (Yates and Cochran, 1938; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966; Tai, 1971). The regression model proposed by Eberhart and Russell (1966) allows for the computation of a complete analysis of variance with individual stability estimates and departure from linearity (sd_i^2) of a regression line. In this model, cultivars with a high sd_i^2 deviate significantly from linearity and have a less predictable response for the given set of environments. Differences in genotype slopes (b_i), along with sd_i^2 , account for GE interactions; however, the validity of these methods has been questioned (Freeman and Perkins, 1971; Shukla, 1972; Freeman, 1973; Vargas et al., 2001).

The cultivar superiority measure (CSM) involves calculations (across environments) of the mean square difference between the performance of a variety and the best variety

within a given environment, measuring mean performance and stability simultaneously (Lin and Binns, 1988). An additive main effect and multiplicative interaction (AMMI) model (Gauch, 1988; Zobel et al., 1988) has been commonly used to analyze multi-environment trial (MET) data. The AMMI model presents a biplot similar to GGE Biplot but does not allow for many of the functions that GGE Biplot provides and can be misleading for identifying which genotypes won in which environment (Yan and Kang, 2003). Kang's yield stability statistic (YS_i) (Kang, 1993), an example of simultaneous selection for mean performance and stability, involves cultivar rankings based on σ_i^2 (Shukla, 1972) and mean performance rankings after a protected LSD adjustment. The most effective method to analyze GE interaction is through the use of GGE Biplot which, using singular value decomposition (SVD). The SVD methodology decomposes GGE into two or more principal components, each of which consists of a set of genotype scores multiplied by a set of environment scores, and generates a two-dimensional biplot. Another significant characteristic of GGE Biplot is its ability to remove noise caused by E, allowing the evaluator to focus on the two components of performance that are meaningful to a breeder, G and GE (Yan and Kang, 2003).

GGE Biplot is equipped with a variety of models, scalings, and data transformations to provide the user with a customized biplot of an MET dataset. A total of 12 different model-by-scaling combinations can be used, each affecting the visual outcome of the biplot and stability-related values that are placed in the GGE log output file. Four models can be used to generate a biplot: 1) Model 0 (M_0) generates biplots based on SVD of the grand-mean centered data and is used only for datasets containing binary data, 2) Model 1 (M_1) generates biplots based on SVD of tester-centered data, commonly used for

datasets in which all testers use the same unit, such as a genotype x environment table of a single trait, 3) Model 2 (M_2) generates biplots based on SVD of within-tester standard deviation-standardized data and is used for datasets in which different units are used for different testers or when all testers are assumed to be equally important, 4) Model 3 (M_3) generates biplots based on SVD of within-tester standard error-standardized data and is used only when replicated data are input and to remove any heterogeneity among the testers. A biplot can be scaled three ways: 1) entry-focused scaling, when the singular values are partitioned entirely into the genotype eigenvectors, is used when the investigator is primarily interested in genotypes; 2) tester-focused scaling, when the singular values are partitioned entirely into the environment scores, is to be used when testers are the primary focus; and 3) symmetrical scaling, which is appropriate when the user wishes to focus equally on testers and entries. The three scaling methods of singular value partitioning do not alter the GE interaction pattern and yield an identical “which-won-where” pattern (Yan and Kang, 2003). The multiple combinations of features that are available increase the potential application of the GGE biplot analysis, but there has been concern over which model-by-scaling combination is most appropriate for variety selection procedures of interest to breeders and producers. Thus, the objectives of this study were to compare the twelve model-by-scaling combinations available within GGE biplot with other stability measures to determine how closely correlated each model-by-scaling combination is with “traditional” stability measures.

4.2 Materials and Methods

This study was conducted to compare the utility of GGE biplot to several traditional stability measures. Cotton yield data from the Louisiana Official Variety Trials (medium

maturity) from 2000, 2001, and 2002 were analyzed. These data were balanced to obtain an equal number of genotypes in all environments. Genotypes included Delta and Pine Land (DP) ‘DeltaPearl’, ‘NuCotn33B’, ‘DP458BR’, ‘DP565’, FiberMax (FM) ‘FM832’, Phytogen (PSC) ‘PSC355’, and Stoneville ‘STV580’. These seven genotypes were planted in 2000 to 2002 at the Dean Lee Research Station in Alexandria (ALEX00-02), the Red River Research Station in Bossier City (BC00-02), the Northeast Louisiana Experiment Station in St. Joseph, and the Macon Ridge Research Station in Winnsboro, Louisiana, yielding a total of 18 environments. In each year, two trials were conducted at the Northeast Louisiana Experiment Station, one on a commerce silt loam (Commerce silt loam; fine-silty, mixed, nonacid, thermic, Aeric, Fluvaquent) (SJL00-02) and the other on a sharkey clay loam (Sharkey clay; very-fine, montmorillonitic, non-acid, thermic Vertic Haplaquept) (SJC00-02) soil type, and at the Macon Ridge Research Station, one under irrigation (WI00-02) and the other non-irrigated (WNI00-02). All experiments were conducted in randomized complete-block designs at each location.

GGE Biplot generates statistics for G, GE, and G+GE allowing the evaluator to assess cultivars for mean performance only, stability only, and desirability based on performance and stability, respectively. For the purpose of this study, the stability statistics evaluating only stability generated by models 1, 2, and 3 will be referred to as M_1GE , M_2GE , M_3GE , respectively. The stability statistic resulting from models 1, 2, and 3 assessing both mean performance and stability (G+GE) and based on the distance of a cultivar from an “ideal” cultivar will be referred to as M_1GGE , M_2GGE , M_3GGE , respectively. Comparisons were made between GGE biplot and seven commonly used stability measures: CSM (Lin and Binns, 1988), σ_i^2 (Shukla, 1972), s_i^2 (Shukla, 1972),

sd_i^2 (Eberhart and Russell, 1966), YS_i (Kang, 1993; Kang and Magari, 1995), and AM_1 (Gauch, 1988). An analysis of variance indicated that the GE interaction was significant ($P < 0.001$). Correlation coefficients were calculated using ‘*proc corr*’ in the SAS System (version 9) (SAS, 2002) between the stability statistics generated by GGE biplot and the traditional measures. The twelve model-by-scaling combinations within GGE biplot were compared with each other and the traditional measures to determine degree of similarity between them and to identify one that is most useful for aiding a breeder in genotype selection.

4.3 Results and Discussion

4.3.1 Within-Model Scaling Correlation

Differences in stability values between entry-focused, tester-focused, and symmetrical scaling options were minimal (Table 4.1) and primarily aided the visual interpretation of the biplot (Figures not shown).

Table 4.1. Correlation coefficients for GGE Biplot within-model scaling options.

Within Model-Scaling	Model-Scaling [†]			
	0-E	1-E	2-E	3-E
T	0.613	1.00	1.00	1.00
S	0.994	1.00	1.00	1.00

[†] Abbreviations: 0, 1, 2, 3=GGE Biplot Model 0, 1, 2, 3; E=Entry-focused scaling; T=Tester-focused scaling; S=Symmetrical scaling.

Except for M_0GE , all within-model scaling combinations were perfectly correlated indicating that different scaling options would be suitable and should depend solely on the focus of the evaluator (Table 4.1). Due to the identical patterns generated by each

scaling option, we chose only to compare the entry-focused scaling method for each model with other stability measures. Using the entry-focused scaling method, all of the singular values were partitioned into the genotype scores, making them much greater than the environment scores, and causing environments to be crowded on the biplot relative to genotypes, thus, allowing for better characterization of genotypes, the primary focus of plant breeders.

4.3.2 Correlation Between GGE Biplot Models (GE)

M_0GE was not correlated with any other GGE biplot model or traditional stability measure. M_0GE uses SVD of the grand-mean-centered data and should only be used for datasets containing binary data (Yan and Kang, 2003). Among GGE biplot models appropriate for the data used in this study, models 1, 2, and 3, correlation coefficients ranged from 0.79 to 0.97 with the strongest correlation being between M_2 and M_3 (Table 4.2).

M_1GE , which had correlations of 0.79 and 0.82 with M_2GE and M_3GE , respectively, is referred to as an environment-centered model because the environmental means are subtracted from each of the observed mean values (Table 4.2). This model is restricted to analyzing MET data in which: 1) all environments are assumed to be homogeneous, 2) a single trait is analyzed, and 3) traits are measured in the same unit for all environments. Breeders, to represent the range of growing regions, commonly use heterogeneous environments; therefore, M_1 might not be appropriate for most MET data collected by cotton breeders. The objective of M_2 , to circumvent the limitations associated with M_1 , is to scale the environment-centered data with the within-environment standard deviation. This procedure eliminates the possibility of detecting any differences among

environments in their discriminating ability by assuming that all environments are equally important.

Table 4.2. Correlation coefficients between GGE Biplot models and “traditional” stability measures.

Model	Model ^{†‡}											
	M ₀ GE	M ₁ GE	M ₂ GE	M ₃ GE	σ_i^2	s_i^2	sd_i^2	CSM	AM ₁	YSi	M ₃ GGE	
M ₀ GE	1.0	-	-	-	-	-	-	-	-	-	-	-
M ₁ GE	-0.48	1.0	-	-	-	-	-	-	-	-	-	-
M ₂ GE	-0.11	0.79*	1.0	-	-	-	-	-	-	-	-	-
M ₃ GE	-0.15	0.82*	0.97*	1.0	-	-	-	-	-	-	-	-
σ_i^2	0.17	0.73*	0.92*	0.91*	1.0	-	-	-	-	-	-	-
s_i^2	0.03	0.81*	0.80*	0.86*	0.91*	1.0	-	-	-	-	-	-
sd_i^2	0.03	0.81*	0.80*	0.86*	0.91*	1.0*	1.0	-	-	-	-	-
CSM	-0.71	0.75*	0.48	0.54	0.38	0.49	0.49	1.0	-	-	-	-
AM ₁	-0.28	-0.41	-0.40	0.55	-0.60	-0.80*	-0.80*	-0.10	1.0	-	-	-
YSi	0.63	0.89*	0.65	0.63	0.55	0.62	0.62	0.91*	-0.12	1.0	-	-
M ₃ GGE	-0.70	0.78*	0.47	0.59	0.39	0.60	0.60	0.95*	-0.33	0.85*	1.0	-

[†] Abbreviations: GGE Biplot model 1=M₁GE, 2=M₂GE, and 3=M₃GE using stability (GE) values, entry-focused scaling used with all models; σ_i^2 =Shukla’s stability variance statistic; s_i^2 =Shukla’s stability variance statistic with location as a covariate; sd_i^2 =Eberhart and Russell regression model deviation values; CSM=Cultivar Superiority Measure; AM₁=AMMI axis 1; YSi=Kang’s yield stability statistic; GGE Biplot model 3= M₃GGE using simultaneous selection for G+GE, entry-focused scaling.

[‡] Correlation coefficients followed by an asterisk are significant at $\alpha=0.05$.

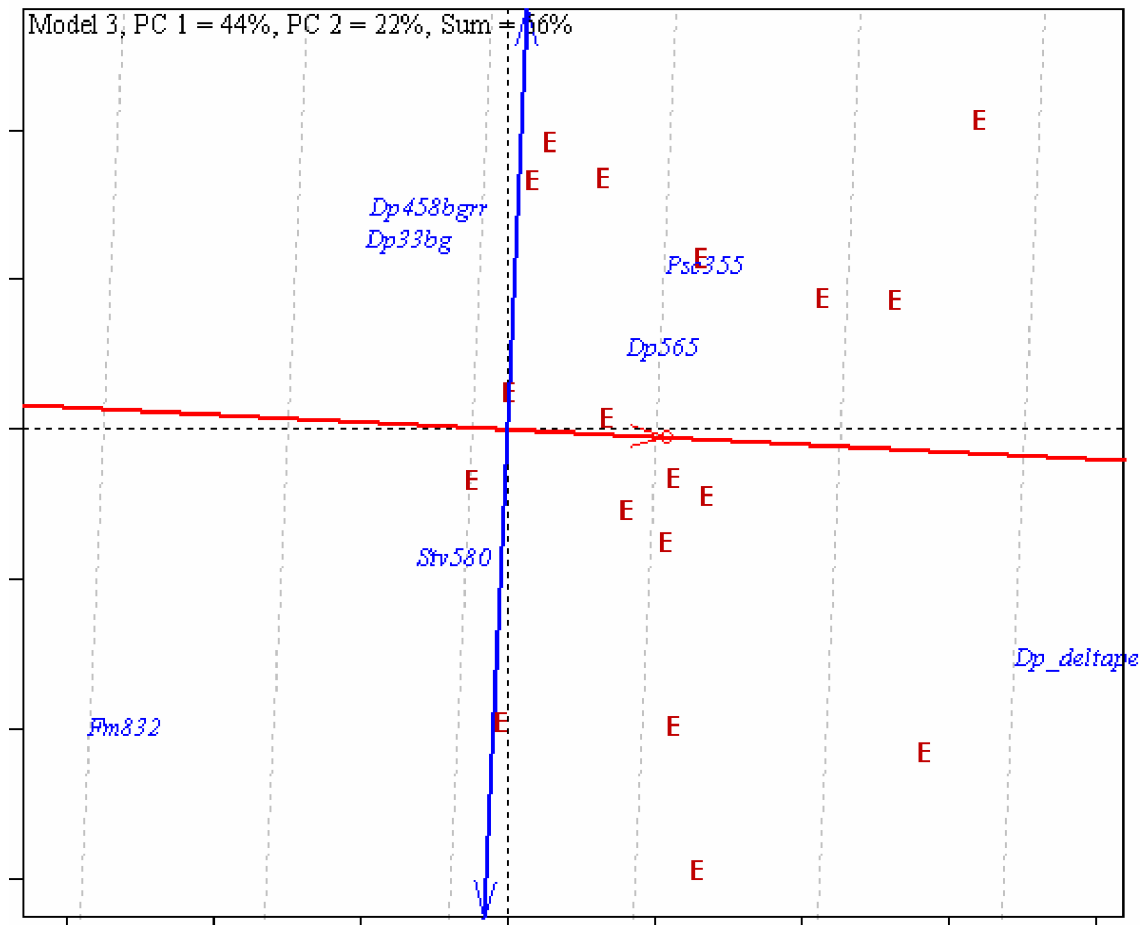
M₂GE was correlated with M₃GE (r=0.97) due to the similar methods for scaling the environment-centered data (Table 4.2). M₃, which should only be used for replicated data, scales the environment-centered data with the within-environment standard error

accounting for the heterogeneity among environments (Yan, 2001).

4.3.3 Correlation Between GGE Biplot (GE) and Other Stability Measures

Since M_2GE and M_3GE were highly correlated ($r=0.97$); future discussion will focus only on M_3GE . Correlation coefficients between M_3GE and σ_i^2 , sd_i^2 , CSM, AM_1 , YSi and M_3GGE were 0.91, 0.86, 0.54, -0.55, 0.63, and 0.59, respectively (Table 4.2). M_3GE , unlike estimates given by CSM, AM_1 , YSi , and M_3GGE , approximates only the level of stability of each cultivar, not the desirability of each cultivar based on $G+GE$. Hence, the high correlations between M_3GE and the traditional stability analyses focus solely on GE interaction, σ_i^2 and sd_i^2 . This indicates that the biplot and stability statistics generated by GGE biplot allow for an effective evaluation of genotype stability. Figure 4.1 is the mean vs. stability coordination biplot, which was generated by M_3 in GGE Biplot. A solid line with a single arrow, called the average tester axis (ATA), passes through the biplot origin and the average environmental coordinate (AEC). Ten dotted lines intersect the ATA and are used for categorizing cultivar mean performance such that cultivars further along the line, away from the biplot origin and in the direction of the arrow exhibit a higher level of mean performance. In this manner, G can be assessed with the biplot. GE interaction can be determined as cultivar distance, in either direction, from the ATA such that cultivars closer to the ATA are more stable than cultivars with a longer distance from the ATA. A line bisecting the biplot origin and perpendicular to the ATA with arrows at each end, called the stability line, aids in the interpretation such that longer projections onto the stability line indicate greater cultivar instability. The projections can be determined as the point where a perpendicular line from the cultivar marker to the stability line intersects. Using the biplot to make stability determinations,

the most stable cultivars would be DP565 and STV580 and the least stable cultivar would be FM832 (Figure 4.1).



† Biplot generated using model 3 with entry-focused scaling.

‡ Cultivars further to the right in the direction of the single arrow yield more.

Cultivars further from the biplot origin, in either direction of the arrows on the vertical line, are less stable.

Figure 4.1. Biplot[†] showing mean cotton lint yield[‡] and yield stability[#] of seven genotypes.

The correlation coefficient between M_3GE and σ_i^2 was 0.91 (Table 4.2), a fact easily validated by a visual assessment of the biplot (Figure 4.1) and cultivar rankings (Table 4.3). Although M_1GE was significantly correlated with σ_i^2 and sd_i^2 , correlation coefficients were lower compared with M_2 or M_3 (Table 4.2). M_1GE was correlated with

CSM ($r=0.75$), although correlations were not significant between CSM and M_2GE or M_3GE . A stronger correlation between CSM and M_1GE is probably due to the fact that both assume homogeneity of environments and use replicate means subtracted from a benchmark, but do not standardize environments as in M_2GE or M_3GE .

Table 4.3. Stability values[†] and rankings[‡] for seven cotton cultivars generated by GGE Biplot and other stability models for seven cotton cultivars.

Cultivar	Model [#]						
	M_3GE	M_3GGE	CSM	σ_1^2	sd_1^2	AM_1	YSi
	2.028	2.20	5624	162882	1.16	2.02	0
DPDeltapearl	(5)	(1)	(1)	(6)	(4)	(2)	(1)
	-1.829	7.40	22327	125980	0.925	6.76	9
DP33B	(4)	(5)	(5)	(4)	(5)	(4)	(5)
	-2.128	7.40	23326	136130	0.871	12.1	10
DP458BR	(6)	(6)	(6)	(5)	(3)	(6)	(6)
	-0.879	4.50	13096	106654	1.07	3.4	4
DP565	(1)	(3)	(3)	(2)	(2)	(3)	(4)
	3.208	10.30	27943	222539	0.995	-17.98	12
FM832	(7)	(7)	(7)	(7)	(7)	(7)	(7)
	-1.723	4.40	6814	122432	1.01	-6.91	2
PSC355	(3)	(2)	(2)	(3)	(6)	(5)	(3)
	1.323	6.70	16302	78703	0.964	0.6	1
STV580	(2)	(4)	(4)	(1)	(1)	(1)	(2)

[†] Cultivars with greater absolute values are considered to be less stable.

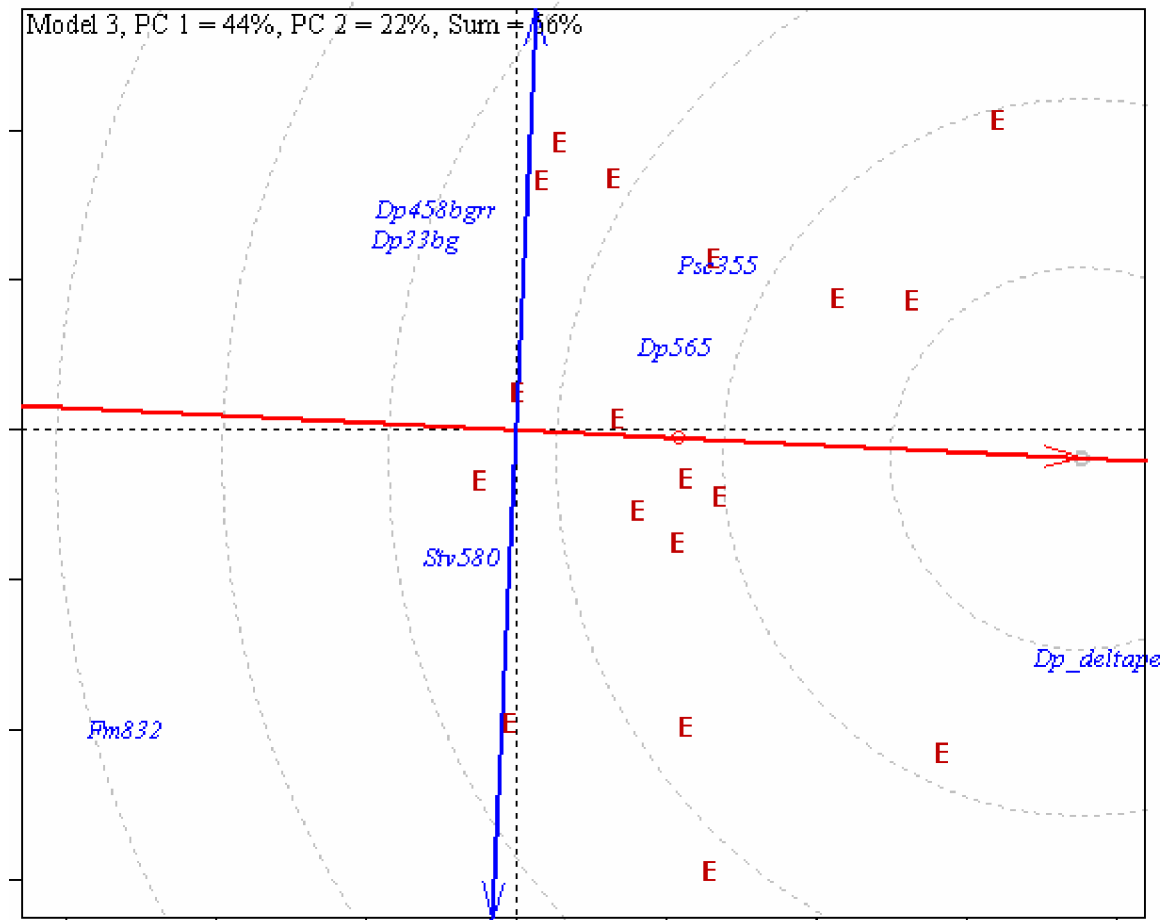
[‡] Cultivar rankings in parentheses are based on stability values such that 1 is highly stable (desirable) and 7 is less stable.

[#] Abbreviations: GGE Biplot model 3 (GE)= M_3GE ; GGE Biplot model 3 (G+GE)= M_3GGE ; CSM=Cultivar Superiority Measure; σ_1^2 =Shukla's stability variance statistic; sd_1^2 =Eberhart and Russell regression model deviation values; AM_1 =AMMI axis 1; YSi=Kang's Yield Stability Index.

4.3.4 Correlation Between GGE Biplot (GGE) and Other Stability Measures

Simultaneous selection for G and GE was accomplished using the concentric-circle biplot (Figure 4.2) and distances from an "ideal" genotype (Table 4.3). In the concentric-circle biplot, an "ideal" genotype, denoted by the smallest circle on the ATA, is created and concentric circles are drawn such that cultivars closer to the center are more desirable relative to mean performance and stability (Figure 4.2). This property of the biplot

greatly simplifies the process of identifying stable, high-yielding cultivars in MET datasets.



† Biplot generated using model 3 with entry-focused scaling.

‡# Cultivars closer to the ideal genotype, denoted by the small circle on the single-arrow line, are more desirable considering yield and yield stability.

Figure 4.2. Biplot[†] of concentric circles ranking seven cotton genotypes for yield[‡] and yield stability[#].

Distances between cultivar markers and the “ideal” genotype are printed to a log file similar to that for the GE stability statistics. Cultivar GE and G+GE statistics and rankings generated by GGE Biplot and other stability measures are listed in Table 4.3.

We believe that between models 2 and 3, the latter is more appropriate for replicated MET datasets; thus, henceforth, comparisons will only be made with M₃GGE.

Correlation coefficients between M_3GGE and CSM, AM_1 , and YS_i were 0.95, -0.33, and 0.85, respectively (Table 4.2). A visual inspection of Figure 4.2 clearly indicates that Deltaparl was the most desirable and FM832 was the least desirable cultivar based on G+GE. This is validated by the strong correlation coefficients (Table 4.2) and similar cultivar rankings (Table 4.3) between M_3GGE , CSM, AM_1 , and YS_i .

The results of this study indicate that GGE Biplot can be used to analyze MET datasets for G, GE, or G+GE and provide results similar to other popular stability analysis tools. However, the multiple options and ease of visual interpretation clearly make GGE Biplot the preferred tool for cotton breeders interested in cultivar stability evaluation. In addition to the analyses performed in this study, GGE Biplot can be used to identify discriminating environments, partition multiple environments into mega-environments, and identify winning genotypes for each mega-environment. An exhaustive list of GGE Biplot properties and functions, many of which are not available in other stability models, has been previously published (Yan et al., 2000; Yan and Kang, 2003).

4.4 Literature Cited

Allard, R. W., and A. D. Bradshaw. 1964. Implications of genotype-environmental interactions in applied plant breeding. *Crop Sci.* 4:503-507.

Blanche, S. B., G. O. Myers, M. Akash, and B. Jiang. 2002. Transgene effect on the stability of cotton cultivars in Louisiana. *In Proc. Beltwide Cotton Conf., Atlanta, GA.* 9-13 Jan. 2002. Natl. Cotton Counc. Am., Memphis, TN.

Eberhart, S. A., and W. A. Russell. 1966. Stability parameters for comparing varieties. *Crop Sci.* 6:36-40.

Finlay, K. W., and G. N. Wilkinson. 1963. The analysis of adaptation in a plant breeding programme. *Aust. J. Agric. Res.* 14:742-754.

- Freeman, G. H. 1973. Statistical methods for the analysis of genotype-environment interactions. *Heredity* 31:339-354.
- Freeman, G. H., and J. M. Perkins. 1971. Environmental and genotype-environmental components of variability. VIII. Relation between genotypes grown in different environments and measures of these environments. *Heredity* 27:15-23.
- Gauch, H. G., Jr. 1988. Model selection and validation for yield trials with interaction. *Biometrics* 44:705-715.
- Kang, M. S. 1993. Simultaneous selection for yield and stability in crop performance trials: Consequences for growers. *Agron J.* 85:754-757.
- Kang, M.S. 1998. A rank-sum method for selecting high-yielding, stable corn genotypes. *Cereal Res. Commun.* 16:113-115.
- Kang, M. S., and R. Magari. 1995. STABLE: A BASIC program for calculating stability and yield-stability statistics. *Agron. J.* 87:276-277.
- Kerby, T., J. Burgess, M. Bates, D. Albers, and K. Lege. 2000. Partitioning variety and environment contribution to variation in yield, plant growth, and fiber quality. p. 528-532. *In Proc. Beltwide Cotton Conf., San Antonio, TX. 4-9 Jan. 1997. Natl. Cotton Counc. Am., Memphis, TN.*
- Lin, C.S., and M.R. Binns. 1988. A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68:193-198.
- Lubbers, E. L. 2003. Evaluation of cotton breeding test environments in the southeast United States. p. 791-807. *In Proc. Beltwide Cotton Conf., Nashville, TN. 6-10 Jan. 2003. Natl. Cotton Counc. Am., Memphis, TN.*
- Myers, G. O. 2002. Biplot analysis of the National Cotton Variety Test. *In Proc. Beltwide Cotton Conf., Atlanta, GA. 9-13 Jan. 2002. Natl. Cotton Counc. Am., Memphis, TN.*
- SAS Institute. 2002. SAS user's guide: Statistics. v. 9.0. SAS Inst., Cary, NC.
- Shukla, G. K. 1972. Some statistical aspects of partitioning genotype environmental components of variability. *Heredity* 29:237-245.
- Tai, G. C. C. 1971. Genotypic stability analysis and its application to potato regional trials. *Crop Sci.* 11:184-190.
- Vargas, M., J. Crossa, F. van Eeuwijk, K. D. Sayre, and M. P. Reynolds. 2001. Interpreting treatment x environment interactions in agronomy trials. *Agron. J.* 93: 949-960.

Verhalen, L. M., and J. C. Murray. 1970. Genotype by environment interaction study of cotton in Oklahoma. p. 52-54. *In* Proc. Beltwide Cotton Prod. Res. Conf., Houston, TX. 6-7 Jan. 1970. Natl. Cotton Counc. Am., Memphis, TN.

Wricke, G. 1962. Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen *Z. Pflanzenzücht.* 47:92-96.

Yan, W. 2001. GGE Biplot- A Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agron. J.* 93:1111-1118.

Yan, W., and M. S. Kang. 2003. *GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists.* CRC Press, Boca Raton, FL.

Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnic. 2000. Cultivar evaluation and mega-environment investigation based on the GGE Biplot. *Crop Sci.* 40:597-605.

Yates, F., and W. G. Cochran. 1938. The analysis of groups of experiments. *J. Agric. Sci.* 28:556-580.

Zobel, R. W., M. J. Wright, and H. G. Gauch, Jr. 1988. Statistical analysis of a yield trial. *Agron. J.* 80:388-393.

CHAPTER 5

IDENTIFYING DISCRIMINATING LOCATIONS FOR CULTIVAR SELECTION IN LOUISIANA

5.1 Introduction

Genotype (G) x environment (E) interactions (GEI) have been studied regarding cultivar stability (Wricke, 1962; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966; Baker, 1988; Lin and Binns, 1988; Kang, 1993; Yan, 2001) and environment groupings (Gauch and Zobel, 1997; Atlin et al., 2000; Trethowan, et al., 2003; Yang et al., 2005). However, relatively few researchers have studied GEI to determine the desirability of test locations. Yan and Kang (2003) proposed using GGE Biplot Pattern Explorer to examine GEI with respect to discriminating ability and representativeness of test locations as a measure of desirability.

A highly discriminating location is one that maximizes the observed genotypic variation among genotypes for a given trait. The efficiency and accuracy of cultivar selection for a given trait is greatly enhanced in highly discriminating locations compared with non-discriminating locations. Therefore, the identification of highly discriminating locations, for a single or combination of traits, should be of paramount concern to breeders. The discriminating ability of a location is comprised of a variety of factors, including soil type, pest pressure, field drainage, temperature, precipitation, soil fertility, and management practices. Some of these factors such as soil type are static and indigenous to each location. For example, Winnsboro, La. is located on the Macon Ridge and is characterized as a slightly acid light-textured soil with a high aluminum content. An array of genotypes exhibiting any degree of variation in aluminum toxicity would be highly discriminated against at this test site compared to a random test site. St. Joseph,

La., while geographically close (30 miles) to Winnsboro, is characterized as an alluvial, deep, highly fertile silt loam and a high-yield environment. Any number of the characteristics inherent to a soil type (texture, fertility, organic matter content, etc.) could affect the discriminating ability of a location. Alternatively, dynamic factors such as pest pressure, precipitation, temperature and management practices fluctuate yearly, although some trends are evident over years. A discriminating location should portray a favorable array of both static and dynamic factors with reasonable repeatability. Ideally, a plant breeder would conduct the selection and early testing phase of the breeding program in the location that provides the most information regarding cultivar separation for each trait. However, limited resources often inhibit that detail and most plant breeders use few test locations for selection (Lubbers, 2003).

In addition to exhibiting a high level of discrimination, an ideal test location should also be representative of the target growing region, or mega-environment (Lubbers, 2003). Yan (2001) and Yan et al. (2001) discussed the use of GGE Biplot Pattern Explorer to categorize locations into mega-environments. Traditionally, cotton (*Gossypium hirsutum* L.) breeding companies have used test locations in various mega-regions, e.g., the Mid-South, Southeast, Southwest, Far West, and cultivar selection at those sites is primarily targeted for that region. Thus, to identify an ideal test location for a region, an average, or representative, location should be used to reflect all locations in the mega-region (Yan, 2001). The shifted multiplicative model (SHMM) has been used to observe the associations among locations and their similar tendencies to differentiate among genotypes, and to identify locations with a high degree of representativeness (Trethowan et al., 2003; Lillemo et al., 2004). However, a measure of discriminating

ability was not determined. In their studies, sites with a low level of association with other global sites indicated that those sites were irrelevant for predicting global yield performance and key locations with a high level of association with other global sites would be considered key sites and good predictors of global performance. Glaz et al., (1985) used Shukla's stability-variance parameter (Shukla, 1972) to identify similar location pairs and single degree of freedom interactions to determine which of the location pairs identified contained the most similar locations.

Identification of an ideal test location based on discriminating ability and representativeness implies that selections made at that site would have the highest probability of representing truly superior genotypes that perform well in all locations in the growing region. Major benefits to breeders would include the increased efficiency of selecting in discriminating locations and the discontinued use of poorly discriminating locations. Thus, cultivar development can be achieved most efficiently within the limited resources available to breeders. Multiple trait selection is important in plant breeding because ideal cultivars must exhibit acceptable performance for multiple characteristics such as yield, quality, maturity, pest resistance, etc. Yan and Kang (2003) identify a method for cultivar evaluation based on multiple traits; however, one limitation of their method is that multiple trait selection can be applied for cultivar selection but not determinations of discriminating ability and representativeness of test locations. Thus, a measure of a test location's desirability based on multi-trait selection is needed. Our objective was to use the method presented by Yan and Kang (2003) and present a method for determining an ideal test location based on weighted simultaneous selection for multiple traits.

5.2 Materials and Methods

Early and medium maturity groups of the 1993 to 2003 Louisiana Official Cotton Variety Trials (LAOVTs) (LCES, 1993-2003) were used to construct 21 datasets for each year x maturity combination for analysis via GGE Biplot Pattern Explorer. In 1993, the LAOVTs were not separated by maturity group so the 1993 LAOVT was analyzed as a medium maturity trial resulting in a total of 10 early maturity and 11 medium maturity datasets. The fact that the genotypes were not constant throughout the 10-year period was irrelevant; genotypes were only used to calculate the desirability of the locations for each biplot.

The LAOVTs have traditionally been conducted at six test locations in Louisiana: Alexandria (ALEX), Bossier City (BC), St. Joseph loam (SJL) (Commerce silt loam; fine-silty, mixed, nonacid, thermic, Aeric, Fluvaquent), St. Joseph clay (SJC) (Sharkey clay; very-fine, montmorillonitic, non-acid, thermic Vertic Haplaquept), Winnsboro irrigated (WIR), and Winnsboro non-irrigated (WNI). While these six test locations do not encompass the entire range of potentially desirable selection environments in Louisiana, they have historically been used for variety testing and are intended to represent the major cotton-growing regions in Louisiana. These six test locations were analyzed to determine which test location of those in the study was most desirable for enhancing germplasm selection. Traits analyzed were lint yield and fiber length, alone and as a component of an indexed value. Yield data were obtained by weighing machine-picked seedcotton and multiplying by lint percentage and were transformed into kilograms of lint per hectare prior to standardization. Fiber length determinations were

made at the Louisiana State University Cotton Fiber Testing Laboratory using HVI instrumentation and reported as the upper-half mean length prior to standardization.

GGE Biplot Pattern Explorer generates a biplot of an “ideal” tester, which is highly discriminating and an average of the locations in the dataset (representative). Therefore, the 2-dimensional graphical distance between the actual test location (ALEX, BC, SJL, SJC, WIR, WNI) and the “ideal” tester is an indication of the desirability of that location with respect to discriminating ability and representativeness for that trait. Yan and Kang (2003) provide a detailed explanation of the biplot calculations and “ideal” test site determinations. The distance (in mm) between each location marker and the “ideal” test location marker was determined (Figure 5.1) and that distance was then standardized by the mean distance of all locations for each biplot. The standardized distances for each test location were averaged across the 21 datasets to obtain the mean distance from the “ideal” tester and standard deviation for each test location. The standard deviation was the deviation of the standardized distance of each test location from the “ideal” test location across 21 year by maturity biplots. Since standardized data was used, the 21 year x maturity biplots were treated as replications. These data are presented for lint yield, fiber length, and combined into a single selection index value.

For the combined selection index value, the graphical distance between each test location and the “ideal” location for lint yield was measured and given a 60% weight for each year x maturity biplot (i.e., 1996 early). For the same year x maturity biplot (i.e., 1996 early), the graphical distance between each test location and the “ideal” location for fiber quality was measured and given a 40% weight. The mean distance of each test location’s combined selection index value is the weighted average of the distance

between the actual and “ideal” test location for lint yield (60%) and fiber length (40%) for each of the 21 biplots. Standard deviations were calculated as previously described. Generally, breeders select primarily for yield and secondarily for numerous quality components assigning weights to each trait based on personal conviction. The weights given to each trait were assigned arbitrarily to show the method; in reality, the procedure is easily customizable to various scenarios. The resulting selection index value represents the distance from the “ideal” test location weighted 60:40 for lint yield and fiber length, respectively.

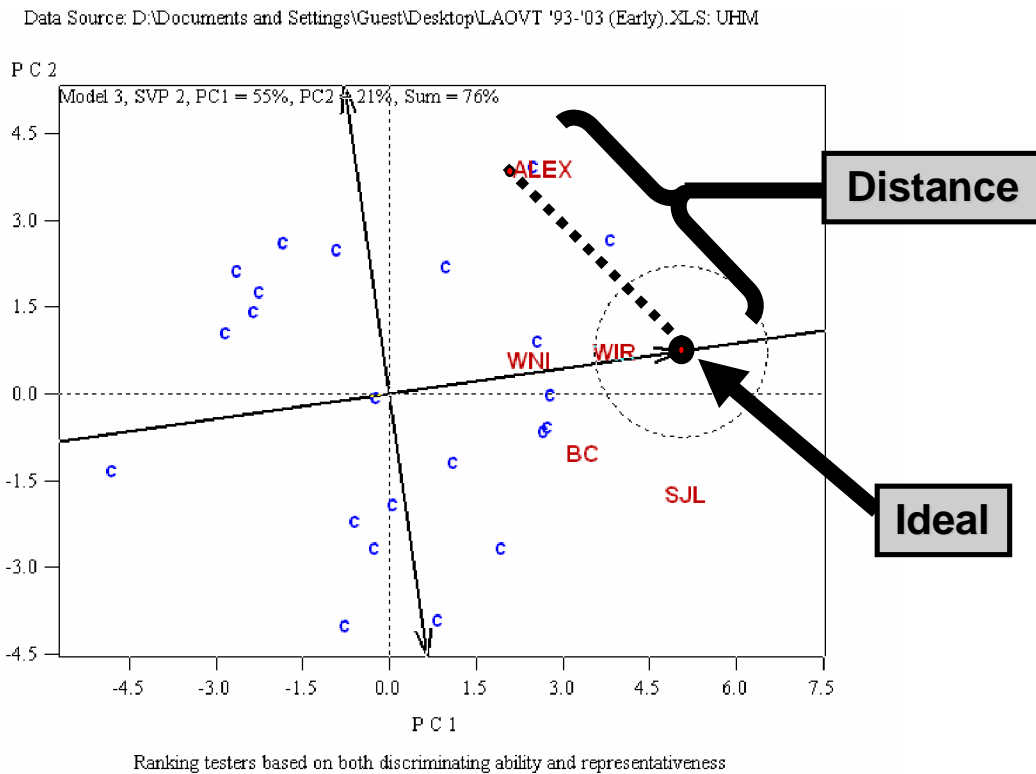


Figure 5.1. Example of biplot showing graphical distance (mm) between actual and “ideal” location.

Test locations with shorter graphical distances relative to the “ideal” test location are regarded as the most suitable for maximizing selection progress. Mean separation was done using an F-protected Duncan’s Multiple Range Test in the SAS System v. 9.0 (SAS, 2002)

5.3 Results and Discussion

The order of location desirability was SJL, SJC, ALEX, WIR, BC, and WNI for lint yield (Table 5.1).

Table 5.1. Standardized graphical distances between actual and “ideal” locations, standard deviations, and rankings of six locations for cotton yield.

Location	Rank	Distance	SD
St. Joseph loam	1	0.769 a [†]	0.30
St. Joseph clay	2	0.932 ab	0.49
Alexandria	3	0.933 ab	0.43
Winnsboro irrigated	4	0.974 ab	0.44
Bossier City	5	1.141 b	0.35
Winnsboro non-irrigated	6	1.214 b	0.36

[†] Means followed by the same letter are not significantly different at $\alpha=0.05$.

Among the test locations included in this study, SJL, SJC, ALEX, and WIR would be equally sufficient test locations for lint yield whereas BC and WNI were less effective, either by failing to discriminate among cultivars or not representing the other test sites or growing regions. SJL had the lowest standard deviation of the six test locations indicating that it was consistently close to the “ideal” test location and fluctuated less across years (Table 5.1). Lubbers (2003) conducted a GEI study including 16 locations spanning the Southeastern cotton belt from lower Alabama to just south of the North Carolina/Virginia border and west to Louisiana to identify ideal test locations for PhytoGen Seed Company, L.L.C. He reported location groupings based on maturity and

separated the two maturity groups (early and late) into mega-environments. Lubbers (2003) reported that out of the 7 test sites covering Mississippi, Arkansas, Missouri, Tennessee, and Louisiana in the late maturity group, St. Joseph, Louisiana was the ideal test location to select for lint yield.

Selection for fiber length alone would be most effective in WIR, followed by BC, SJL, ALEX, SJC, and WNI (Table 5.2).

Table 5.2. Standardized graphical distances between actual and “ideal” locations, standard deviations, and rankings of six locations for cotton fiber length.

Location	Rank	Distance	SD
Winnsboro irrigated	1	0.669 a [†]	0.48
Bossier City	2	0.871 ab	0.42
St. Joseph loam	3	0.887 ab	0.64
Alexandria	4	1.016 ab	0.65
St. Joseph clay	5	1.113 bc	0.51
Winnsboro non-irrigated	6	1.451 c	0.69

[†] Means followed by the same letter are not significantly different at $\alpha=0.05$.

Among the test locations included in this study, WIR, BC, SJL, and ALEX would be equally sufficient test locations for fiber length, but SJC and WNI would be less effective. It is not likely that a separate test location would be used to select only for a quality trait; however, the test location distances for each individual quality trait is needed to create the multi-trait selection index.

Selection of an ideal genotype is seldom based on any single criterion but rather on a composite of attributes. To investigate the ability to use a selection index to identify an ideal test location at which to make selections, we formulated one in which the major emphasis (60%) was given to lint yield and lesser emphasis (40%) was given to fiber

length. For the individual traits, WIR and SJL ranked 4th and 1st for lint yield and 1st and 3rd for fiber length, respectively, indicating that both were desirable test locations to select for each trait individually (Table 5.3).

Table 5.3. Locations ranked according to desirability for cotton yield, fiber length, and simultaneous selection (yield + fiber length) and standard deviations.

Location	Rank (yield)	Rank (length)	Avg. Dist.‡ (Yield +UHM)	SD	Avg. Rank
St. Joseph loam	1	3	0.810 a [†]	0.30	1
Winnsboro irrigated	4	1	0.859 ab	0.23	2
Alexandria	3	4	0.967 ab	0.46	3
St. Joseph clay	2	5	0.977 ab	0.33	4
Bossier City	5	2	1.04 b	0.29	5
Winnsboro non-irrigated	6	6	1.31 c	0.26	6

[†] Means followed by the same letter are not significantly different at $\alpha=0.05$.

[‡] Distance for yield and UHM weighted at 60/40 resulting in weighted average distance from the ideal.

Table 5.3 also contains the average ranking of each test location for both traits and the combined indexed value representing simultaneous selection for lint yield (60%) and fiber length (40%). The composite indexed values indicate that SJL would be the most desirable test location for cultivar selection for both traits in Louisiana (Table 3). WIR, ALEX, and SJC were also acceptable test locations whereas BC and WNI were not desirable either because they provided few meaningful selections or did not well represent the other test sites (growing regions) in Louisiana. It is possible that an exhaustive study using many more test locations than are included in this study would yield different results; however, the subset of potential test locations included in this study are the only selection sites that realistically lend themselves to cultivar evaluation due to available resources and expertise at these locations.

Plant breeders are usually restrained by resource limitations, and conducting selections in the most desirable test location for each individual trait may not be realistic. However, in some cases the economic potential for improving selection efficiency for a secondary trait may warrant trait-specific selection locations. It should be noted that the test locations included in this study were assumed to represent all possible regions in Louisiana, which was assumed to comprise a single mega-environment. There may be various mega-environments within the state in certain years; however, the state cotton breeder is responsible for servicing all of the growing regions in the state with limited resources and it would be impractical to divide Louisiana into multiple selection environments. In comparison, Lubbers (2003) uses multi-state data for identifying desirable selection environments using a regional perspective. The multiple-trait selection techniques employed by the authors in this study are adaptable to many different interests and the number of traits used for determinations and the weights given to each trait are subject to the convictions of the researcher. Certainly test locations that are desirable selection environments for a combination of traits can prove beneficial to breeders interested in cultivar development. Therefore, the most ideal test location for breeders to use as a selection environment is one in which they can select with reasonable effectiveness for ancillary traits without compromising the ability to effectively select for

5.4 Literature Cited

Atlin, G. N., K. B. McRae, and X. Lu. 2000. Genotype x region interaction for two-row barley yield in Canada. *Crop Sci.* 40:1-6.

Baker, R. J. 1988. Tests for crossover genotype-by-environment interactions. *Can. J. Plant Sci.* 68:405-410.

Eberhart, S. A., and W. A. Russell. 1966. Stability parameters for comparing varieties. *Crop Sci.* 6:36-40.

- Finlay, K. W., and G. N. Wilkinson. 1963. The analysis of adaptation in a plant breeding programme. *Aust. J. Agr. Res.* 14:742-754.
- Gauch, H. G., and R. W. Zobel. 1997. Identifying mega-environments and targeting genotypes. *Crop Sci.* 37:311-326.
- Glaz, B., J. D. Miller, and M. S. Kang. 1985. Evaluation of cultivar testing locations in sugarcane. *Theor. Appl. Genet.* 71:22-25
- Kang, M. S. 1993. Simultaneous selection for yield and stability in crop performance trials: Consequences for growers. *Agron. J.* 85:754-757.
- Lillemo, M., M. van Ginkel, R. M. Trethowan, E. Hernández, and S. Rajaram. 2004. Associations among International CIMMYT bread wheat yield testing locations in high rainfall areas and their implications for wheat breeding. *Crop Sci.* 44:1163-1169.
- Lin, C. S., and M. R. Binns. 1988. A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68:193-198.
- Lubbers, E. L. 2003. Evaluation of cotton breeding test environments in the Southeast United States. p. 791-807. *In Proc. Beltwide Cotton Conf., National Cotton Council of America, Memphis, TN.*
- SAS Institute. 2002. SAS user's guide: Statistics. v. 9.0. SAS Inst., Cary, NC.
- Shukla, G. K. 1972. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity.* 29:237-245.
- Trethowan, R. M., M. van Ginkel, K. Ammar, J. Crossa, T. S. Payne, B. Cukadar, S. Rajaram, and E. Hernandez. 2003. Associations among twenty years of international bread wheat yield evaluation environments. *Crop Sci.* 43:1698-1711.
- Wricke, G. 1962. Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen *Z. Pflanzenzücht.* 47:92-96.
- Yan, W. 2001. GGE Biplot- A Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agron. J.* 93:1111-1118.
- Yan, W. P. L. Cornelius, J. Crossa, and L. A. Hunt. 2001. Two types of GGE biplots for analyzing multi-environment trial data. *Crop Sci.* 41:656-663.
- Yan, W. and M. S. Kang. 2003. *GGE Biplot Analysis: A graphical tool for breeders, geneticists, and agronomists.* CRC Press.

Yang, R., S. F. Blade, J. Crossa, D. Stanton, and M. S. Bandara. 2005. Identifying isoyield environments for field pea production. *Crop Sci.* 45:106-113.

CHAPTER 6

SUMMARY AND CONCLUSIONS

6.1 Summary

GGE Biplot was used to examine genotype by environment (GE) interactions in cotton in several different respects. First, comparisons were made between conventional and transgenic cotton varieties using GGE Biplot to elucidate any nontarget trait variation resulting as a consequence of transgene insertion., Secondly, GGE Biplot analysis was used to identify the best, most discriminating environment in which to practice single or multi-trait selection in cotton using historical data from the Louisiana Official Variety Tests. GGE Biplot was also compared to traditional stability measurements. Finally, variance and regression analysis was used to estimate the heritability of multi-trait selection packages and how these estimates responded to selection in locations identified by GGE Biplot as having either high or low levels of discriminating ability.

In the first experiment, environment (E) was responsible for >70% of the variation in total sums of squares for traits with a low heritability such as plant height, height to node ratio, and yield. However, genotype (G) and GE heavily influenced yield components and fiber traits. Transgenic cotton varieties were different from their conventional parents regarding stability and performance for a number of traits including plant height, number of nodes, fuzzy seed index, and lint percentage.

A separate study was conducted using the Louisiana Official Variety Trials to determine the locations that optimize genotype selection based on discriminating ability and representativeness. The most “ideal” location to conduct selection for yield was in St. Joseph loam. Winnsboro non-irrigated and Bossier City were not good selection

locations for lint yield. For fiber length, Winnsboro irrigated was the most “ideal” selection location followed by St. Joseph loam. When data were combined to determine the most “ideal” location considering simultaneous selection for yield and fiber length, the best location was St. Joseph loam. Winnsboro non-irrigated should not be used to select for yield and fiber length due to a low level of discriminating ability and unique behavior.

Different scaling combinations available in GGE Biplot were completely correlated; thus, the scaling options affected the visual appearance of the biplot, but yielded the same results. Models 2 and 3 in GGE Biplot for stability (M_2GE and M_3GE) were 0.97 correlated. GGE Biplot was highly correlated with other stability analysis tools. M_2GE and M_3GE were 0.80 to 0.92 correlated with traditional GE interaction analyses based exclusively on GE interaction. The main advantage in using GGE Biplot is the ability to simultaneously evaluate mean performance and stability. Models 2 and 3 in GGE Biplot simultaneously evaluating mean performance and stability (M_2GGE and M_3GGE) were correlated 0.91 and 0.95 with the Cultivar Superiority Measure and Kang’s Yield Stability Index, respectively.

Two levels of discriminating locations were determined for each of three populations varying in number and choice of selected traits. Regardless of the population, or selection package, Bossier City, La., was the least discriminatory location. Population 1, selected for lint yield, was most discriminated in St. Joseph, La. For population 2, selected for fiber micronaire, length, strength, uniformity and elongation, College Station, Tx., was the most discriminatory location. For population 3, selected for a combination of lint yield, lint percent, fiber micronaire, length, and strength, Starkville, Ms., was the

most discriminatory location. Regardless of population, genotypic variance was greater in the highly discriminating population compared to Bossier City, the least discriminating location. Heritability was greater in population 2 compared to populations 1 and 3, which evaluated yield. Yield is known to be highly responsive to environment; however, the characters in population 2 are less affected by environment, hence the higher heritability.

6.2 Conclusions

GGE Biplot can be used to obtain a variety of information regarding GE interactions. It was well correlated with other traditional stability measures, but is particularly useful as an easily-interpreted, visual tool to assess complex, multi-environment trial data. M_3 GGE, with entry-focused scaling, is the most appropriate model for breeders to analyze replicated, multi-environment trial data. As a tool for breeders, it has much potential for selecting promising genotypes in a single or range of environments, subdividing or grouping various environments into mega-environments, determining the stability of genotypes, and identifying environments that effectively discriminate among genotypes based on a single trait or combination of traits.

VITA

Sterling Brooks Blanche, Sr., was born September 18, 1974, in Monroe, Louisiana. He was raised in St. Joseph, Louisiana, by his parents, Winn and Lynn Blanche of St. Joseph, and graduated from Tensas Academy in the spring of 1992. He received his undergraduate degree in agronomy (soils) in May 1999 from Western Kentucky University. He began work on his master of science degree at Mississippi State University in weed science and finished May 2001. He enrolled at Louisiana State University in June 2001 to pursue a doctorate in agronomy with Dr. Gerald Myers (cotton breeding). He and his wife, Erin Lynne Blanche, have two children, Sterling and Ellie Blanche.