2005 Beltwide Cotton Conferences, New Orleans, Louisiana - January 4 - 7, 2005

USING GGE BIPLOT TO IDENTIFY IDEAL TEST SITES FOR EVALUATING FIBER QUALITY

Dawn E. Fraser and Cynthia C. Green
Delta and Pine Land Company
Hartsville, SC

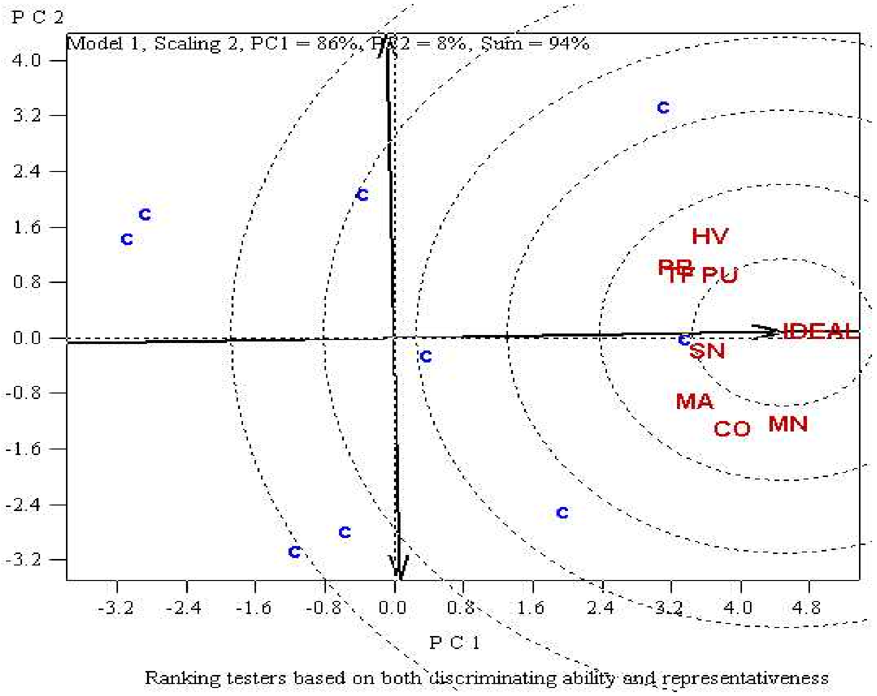
Abstract

Each year thousands of genotypes are evaluated in hundreds of different locations across the Cotton Belt in the hope of identifying and selecting genotypes with improved yield and fiber quality. Genotypes with broad adaptation and stability are preferred over genotypes with more specific local adaptation since this enables seed to be grown and sold over a wider market. Observed phenotypic variation is the result of the variation due to the genotype plus that of the environment as well as the interaction of these two components. If certain environments could be identified where the genotypic performance is expressed most clearly, and if these environments were representative of the target market area, then these environments would be closest to ideal.

Materials and Methods

A study was conducted to determine whether ideal selection environments could be identified regardless of the genotypes tested using GGE Biplot analysis (Yan & Kang, 2002). GGE Biplot is a statistical software package based on principal component analysis. A biplot is a visual representation of the two main components explaining the majority of the variation in any given data set. Figure 1 shows an illustration of a biplot where the ideal selection environment is identified. Environments closest to this ideal are the most discriminating and representative for any given trait and become less ideal the further they are from the designated ideal selection environment. In this particular example, the environment labeled SN was the most ideal selection environment. For the purpose of this study, each variable was analyzed using GGE Biplot and the visual interpretation of the biplot provided information on environments that were closest to the ideal selection environment for that particular variable.

Figure 1 An example of a GGE Biplot indicating the ideal environment for selecting for genotypes with superior fiber qualities. Environments distant from the ideal environment are less discriminating and less representative of the target region.



Two different data sets were used in this study. The first data set (data set A) consisted of nine conventional genotypes grown together in tests at the same eight test locations in 2002 and 2003 (Coolidge, AZ (CO), Hartsville, SC (HV), Maricopa, AZ (MA), Manilla, AR (MN), Panther Burn, MS (PB & PU), Scott, MS (SM), Tifton GA (TF)). The second data set (data set B) was divided into sub-groups based on maturity (early-mid or full season), and technology (conventional or transgenic). All sub-groups of data set B were grown in eight tests environments in 2002 and the same eight test environments in 2003.The early-mid conventional and transgenic tests were grown in nine test environments in 2002 and 2003 (CO, HV, Haskell, TX (HI), MA MN, PB, PU, SM, TF). The early-mid conventional group consisted of 55 genotypes in 2002 and a different set of 59 genotypes in 2003. The early-mid transgenic group consisted of 20 genotypes in 2002 and a different set of 24 genotypes in 2003. The full season conventional and transgenic tests were grown in nine environments in 2002 and 2003 (Bainbridge, GA (BB), CO, HI, HV, MA, Newellton, LA (NE), PB, PU, TF). The full season conventional group consisted of 41 genotypes in 2002 and a different set of 44 genotypes in 2003. The full season transgenic group consisted of 20 genotypes in 2002 and 15 different genotypes in 2003.

Data on lint yield (LYLD), fiber length (LEN), fiber strength (T1), elongation (E1), micronaire (MIC), and fiber uniformity (UR) were collected on each genotype in each test. The data were then analyzed using the graphical statistical software GGE Biplot. GGE Biplot ranks environments based on their closeness to ideal. Figure 1 shows an example of a biplot. An ideal environment is one, which shows the most discrimination among the genotypes, but is representative of the other environments in the target area. The rank of each environment for each variable in data set A is shown in Table 1. Similarly the rank of each environment for each variable in data set B is shown in Tables 2 (early-mid tests) and 3 (full season tests).

Results

Test results for data set A suggested that no single environment in either 2002 or 2003 was completely ideal for selecting for all fiber traits tested in this study (Table 1). In 2002 CO was the only environment that was ideal for selecting for more than a single trait i.e. LYLD, MIC and UR. Other environments in the study were considered ideal selection environments for only a single trait. In 2003, only MN was ideal for selecting for more than one trait (LYLD and LEN); other environments were only ideal selection environments for a single trait. However, there was a high level of consistency between the two years in the environments most ideal for selection of certain traits. MN was ideal for selecting for LEN in both years, CO for UR, HV for T1 and SM for E1. Therefore with set of genotypes that remain constant when tested over multiple years, it may be possible to identify selection environments that are consistently ideal to test in.

Tests results for data set B showed much less consistency than those of data set A. For the early-mid conventional test in 2002 (Table 2), SM was ideal for the selection of multiple traits including LYLD, LP and T1. Another environment, MN, was ideal for selecting for the traits LEN, UR and E1. However, for the early-mid conventional tests in 2003, the most ideal selection environments changed from the 2002 data except for LYLD where the environment that were most ideal in 2002 (SM) remained most ideal in 2003. Similarly for the early-mid transgenic tests, CO was the most ideal selection environment for LYLD, MIC and UR and SM the most ideal selection environment for other traits such as LEN, T1 and E1, but these same environments were not the most ideal in 2003. Only SM was consistently ideal for selecting for E1 in both years.

For the full season conventional tests (Table 3) HV was the only environment that was ideal for selecting for multiple traits in 2002 and BB was the only environment ideal for selecting for multiple traits in 2003. In addition to this, BB and HV, were consistently the most ideal environments for the selection of genotypes with superior elongation (E1) and fiber strength (T1) in both years, respectively. For the full season transgenic tests, BB was the ideal environment for the selection of genotypes with long fiber in both 2002 and 2003.

The results suggest that in an "artificial" data set where lines are constant across multiple years, for some traits such as UR, LEN, T1 and E1 it is possible to identify selection environments that are the most discriminating and representative of a target region. However, in a commercial breeding program where the genotypes evaluated change each year, identifying these selection environments is dependent on the maturity group of the test and the technology incorporated as well as the particular trait to be evaluated. In these circumstances testing genotypes over as many environments as possible is still preferable. However, the data provided by GGE Biplot could be used *a* *posteriori* to place more or less emphasis on data from particular environments in relation to how close to the ideal selection environment they are.

Summary

GGE Biplot is a useful tool for visualizing patterns in data sets and can be used to identify ideal selection environments for certain fiber traits when data sets remain constant. However when data sets change each year or as a result of different maturity or technology, GGE Biplot was not found to be as useful in consistently identifying ideal environments for selecting for yield and fiber quality traits.

Table 1 Environments identified using GGE Biplot analysis as the most ideal for selecting for superior genotypes using the same set of genotypes evaluated in eight different environments in 2002 and 2003

	Ideal Selection Environments	
Variable	2002	2003
Lint yield	CO[1]	MN
Fiber length	MN	MN
Micronaire	CO	PU
Uniformity	CO	CO
Strength	HV	HV
Elongation	SM	SM

[1] SM Scott, MS, MN Manilla, AK, CO Coolidge, AZ, HV Hartsville, SC, PU Panther Burn, MS

Table 2 Environments identified using GGE Biplot analysis as the most ideal for selecting for superior genotypes using different early-mid maturing genotypes evaluated at eight different environments in 2002 and 2003

	Conventional		Transgenic	
Variable	2002	2003	2002	2003
Lint yield	SM	SM	CO	PU
Lint percent	SM	HV/ HI	HI	SM
Fiber length	MN	HV	SM	PB
Micronaire	CO	MA	CO	MA
Uniformity	MN	SM	CO	TF
Strength	SM	TF	SM	PU / PB
Elongation	MN	HV / CO	SM	SM

[1] SM Scott, MS, MN Manilla, AR, CO Coolidge, AZ, HV Hartsville, SC, HI Haskell, TX, MA Maricopa, AZ, TF Tifton, GA, PB Panther Burn, MS, PU Panther Burn, MS

Table 3 Environments identified using GGE Biplot analysis as the most ideal for selecting for superior genotypes using different full season genotypes evaluated at eight different environments in 2002 and 2003

	Conventional		Transgenic	
Variable	2002	2003	2002	2003
Lint yield	TF	BB	TF	BB
Lint percent	HV	NONE	PU	HV
Fiber length	PU	PB	BB	BB
Micronaire	MA	NE	HV	CO / PU
Uniformity	PB	HI	MA	HI / TF
Strength	BB	BB	NE	PU
Elongation	HV	HV	MA / CO	HI

[1] SM Scott, MS, MN Manilla, AR, CO Coolidge, AZ, HV Hartsville, SC, HI Haskell, TX, MA Maricopa, AZ, TF Tifton, GA, PB Panther Burn, MS, PU Panther Burn, MS, BB Bainbridge, GA, NE Newellton, LA

References

Yan, Weikai and Kang, Manjit S., 2002. GGE Biplot Analysis: a graphical tool for breeders, geneticists, and agronomists.