

## BREEDING AND GENETICS

### Analysis of Testing Locations in Regional High-Quality Tests for Cotton Fiber Quality Traits

Linghe Zeng\*, Deborah L. Boykin, Jinfa Zhang, Efrem Bechere, Jane K. Dever, B. Todd Campbell, Tyson B. Raper, Calvin Meeks, Wayne Smith, Gerald O. Myers, and Fred M. Bourland

#### ABSTRACT

Determination of an efficient number of testing locations in multiple-location tests for cotton (*Gossypium hirsutum* L.) fiber quality can allow removal of unnecessary locations while maintaining the statistical power in detection of genotype ( $g$ ) by environment ( $e$ ) interactions. Fiber quality data from Regional High-Quality (RHQ) tests from 2011 to 2016 were used to determine an efficient number of locations in the tests for fiber quality and relationships among locations for their representativeness and ability to discriminate among genotypes. Covariance parameters of  $g$ , location ( $l$ ), and  $gl$  in the original RHQ tests were estimated in a random model. The simulating data with varying number of locations omitted from the original tests were created by performing 100 unique simulations. When locations were reduced to five, the standard deviations (std) of  $gl$  increased from 18 to 37% compared to the original tests. Further reduction of locations to four or less increased std of  $gl$  from 30 to 217% compared to the original tests. Therefore, five locations were determined to be an efficient number of locations in tests for fiber quality. The discrimi-

nating ability and representativeness of the eight locations for fiber properties were calculated as their distances to an “ideal environment”, which was designed as a center in GGE biplot graphs for representativeness and discriminating ability. The relationships among locations were different across years. However, by averaging the distances across testing years, the locations of Stoneville, MS; Keiser, AR; Lubbock, TX; and College Station, TX were identified as the most representative testing sites for fiber properties.

Genotype ( $g$ ) by environment ( $e$ ) interactions affect evaluation tests for quantitative traits in crops and reduce breeding efficiency. Different methods have been developed to quantify the  $ge$  effects in crop breeding. The earliest model for quantification of  $ge$  effects can be categorized into a descriptive method (simple ANOVA model) based on variety means across environments to estimate the proportion of  $ge$  variance to genotype and residual (de Leon et al., 2016). This type of model is used by crop breeders to estimate proportions of  $g$  and  $ge$  variance, which can help them choose an optimum breeding and evaluation strategy. The additive mean effects and multiplicative interaction (AMMI) model is a common method to analyze the structure of  $ge$ . This model combines two statistical procedures: analysis of variance (ANOVA) and singular values decomposition (SVD), to analyze interactions by estimates for the multiplicative terms of the AMMI model (Gauch, 2006). GGE biplot is a method to quantify  $ge$  interactions by visualizing the structure of  $ge$  interactions in trial data (Yan, 2001), although there are arguments about its accuracy compared to other models such as AMMI (Gauch, 2006). The use of a mixed linear model in analysis of  $ge$  interactions is relatively recent. In these mixed linear models, the main effects of  $g$  and  $e$  can be treated as random and fixed, respectively, or vice versa, which should have advantages over simple ANOVA in minimizing errors in multi-environment trials (Smith et al., 2005).

---

L. Zeng\* and E. Bechere, USDA, Agricultural Research Service, Crop Genetics Res. Unit, 141 Experiment Station Rd., Stoneville, MS 38776; D.L. Boykin, USDA, Agricultural Research Service, Southeast Area Statistician, 141 Experiment Station Rd., Stoneville, MS 38776; J. Zhang, New Mexico State Univ., Las Cruces, NM 88003; J.K. Dever, Texas A&M AgriLife Center, Texas A&M Univ., Lubbock, TX 79403; B.T. Campbell, USDA, Agricultural Research Service, Coastal Plains Soil, Water, and Plant Res. Center, Florence, SC 29501; T.B. Raper, Univ. of Tennessee, West TN Res. Education Center, Jackson, TN 38301; C. Meeks, Univ. of Missouri, Fisher Delta Res. Center, Portageville, MO 63873; W. Smith, Texas A&M Univ., Dept of Soil and Crop Sci., College Station, TX 77843; G.O. Myers, Louisiana State Univ., LSU Agricultural Center, Baton Rouge, LA 70803; F.M. Bourland, Univ. of Arkansas, Northeast Res. and Ext. Center, Keiser, AR 72351.

\*Corresponding author: [linghe.zeng@ars.usda.gov](mailto:linghe.zeng@ars.usda.gov)

Significant *ge* effects for Upland cotton (*Gossypium hirsutum* L.) lint yield and fiber quality have been reported (Campbell et al., 2012; Meredith et al., 2012), which indicate the necessity of multiple-location tests for evaluation of cotton cultivars for fiber quality. However, it is important to determine an efficient number of testing locations to reduce cost and promote breeding efficiency. The determined efficient number of locations can be employed in the tests while the statistical power of detecting significant *ge* effects is maintained. This type of information is lacking for evaluation tests of fiber quality traits in cotton across the Regional High-Quality (RHQ) tests as conducted by the National Cotton Variety Test (NCVT, 2016). In a study of historical data in RHQ tests for seed quality traits between 2005 and 2013, when one, two, and three locations were omitted from the original RHQ tests, the nonsignificance of *g* by location (*l*) interactions increased by 5.7, 7.3, and 9.1%, respectively, in the analysis by an ANOVA model (Zeng et al., 2018; submitted to JCS). The authors suggested that two to three locations could be reduced from the current RHQ tests for seed quality traits.

Objectives of this study were to (1) determine an efficient number of testing locations for multiple-locations tests of fiber quality traits using a random model and (2) determine relationships among testing locations for their representativeness and ability to discriminate among genotypes for fiber properties. The most recent high volume instrument (HVI) data of fiber quality traits in RHQ tests from 2011 to 2016 were used to determine the efficient number of testing locations.

## MATERIALS AND METHODS

The purpose and regions included in the RHQ tests were described in a previous study (Zeng et al., 2014). The testing locations of the RHQ were assigned to five agricultural regions of the U.S. Cottonbelt identified as Eastern, Delta, Central, Plains, and Western. During the tests between 2011 and 2016, there were three testing locations in the Eastern Region (Belle Mina, AL; Florence, SC; and Jackson, TN), three locations in the Delta Region (Keiser, AR; Portageville, MO; and Stoneville, MS) two locations in the Central Region (College Station, TX and St. Joseph, LA), one location in the Plains Region (Lubbock, TX), and one location in the Western Region (Las Cruces, NM). There were 18 to 22 genotypes

evaluated each year in the RHQ tests during 2011 and 2016. Every three years were set as a testing cycle. Genotypes were different across testing years, but two genotypes were used as standards across testing years within a testing cycle. The cultivars Fibermax 9058F and Phytogen 375WRF were used as standards in 2011, 2012, and 2013, whereas Fibermax 2484B2F and Phytogen 725RF were used as standards in 2014, 2015, and 2016 (Table 1).

**Table 1. Entries and standards in RHQ tests from 2011 through 2016**

year	No. of entries	Standards	PVP
2011	22	FM 9058F PHY 375WRF	200700206
2012	21	FM 9058F PHY 375WRF	200700206
2013	20	FM 9058F PHY 375WRF	200700206
2014	20	FM 2484B2F PHY 725RF	201200291
2015	18	FM 2484B2F PHY 725RF	201200291
2016	18	FM 2484B2F PHY 725RF	201200291

The field test design, harvest, and fiber quality measurements were described in previous reports (Meredith et al., 2012; Zeng et al., 2015). In brief, the tests were conducted in a randomized complete block design with four to six replications. All genotypes were planted in two-row plots of 12 m × 1 m. At harvest, 50 to 150 opened bolls were hand harvested in each plot from two replications for fiber property measurements. Boll samples of individual plots were ginned separately by laboratory saw gins at each location within each year. Lint samples were submitted to Starlab Inc., Knoxville, TN in 2011 and the Cotton Structure and Quality Research Unit, SRRC, USDA-ARS (New Orleans, LA) after 2012. The fiber properties of strength, length, length uniformity, elongation, and micronaire were measured by HVI instrument 1000 from Uster Technologies (Knoxville, TN). In all measurements, the instrument was calibrated using the USDA Spinlab HVI standard cotton.

To determine the efficient number of testing locations in the RHQ tests, varying numbers of locations were omitted from the original tests and new data were created by simulation. First, covariance parameters in the original RHQ tests for fiber properties were estimated by PROC GLIMMIX in SAS (9.4) (SAS Institute, 2013) in a random model:

$$Y_{ijk} = \mu + g_i + l_j + r_k + gl_{ij} + \varepsilon_{ijk}$$

where  $Y_{ijk}$  is the fiber property of genotype  $g$  at location  $l$  for replicate  $r$ ;  $\mu$  is the overall mean;  $g_i$  is the random effect of genotypes ( $i = 1, \dots, 18$  to 22);  $l_j$  is the random effect of locations ( $j = 1, \dots, 8$  to 9);  $r_k$  is the random effect of replicates ( $k = 1, 2$ );  $gl_{ij}$  is the random effect of  $g_i$  by  $l_j$  interaction;  $\varepsilon_{ijk}$  is the residual of normal distribution with zero mean and variance  $\sigma_e$ . Next, the simulating data with varying number of locations (8, 7, ..., 2) were created by random sampling of testing locations from the original RHQ tests 100 times (DO Sim = 1-100). The random model described above was used to estimate covariance parameters of  $g$ ,  $l$ ,  $gl$ , and  $\varepsilon$  in the simulated data. Mean and standard deviations (std) (0, std) of the 100 simulations were calculated for covariance parameters. To determine the efficient number of locations, std of the interaction parameter  $gl$  in simulating data with varying numbers of locations was compared with that of the original RHQ tests.

The testing locations were analyzed for their representativeness of the environments and discriminating ability of fiber properties using GGE biplot. The approach was described by Yan (2001) and Wei and Tinker (2006). The analysis was performed using the module Discriminating Ability and Representativeness by setting data scaling to 0 (unscaled) and data centering method to 2 (tester-centered). In the GGE biplot graphics there is a calculated discriminatory line called the Average-Environment-Axis (AEA). A small circle on the AEA represents the average environment that is calculated by average coordinates of all test locations. The AEA passes through the origin of GGE biplot and the average environment circle. The line from the GGE biplot origin to a testing location is the environment vector, the length of which represents the discriminating ability of that testing location. The environment vector can be calculated from an equation described by Yan and Tinker (2006). A testing location with the smallest angle to the AEA is the most representative environment in a single mega-environment. In this analysis, the testing locations could be viewed for their discriminating ability and representativeness relative to an ideal test environment. This ideal test environment was defined as the center of concentric circles in the graphics. The ideal test environment was a point on the AEA with a distance equivalent to the longest environment vector so that it could be presumed as the most representative and discriminat-

ing environment. The relationships among testing locations were analyzed by their distances to this ideal environment.

## RESULTS AND DISCUSSION

The most recent RHQ tests of 2011 to 2016 were chosen in analysis of the testing locations. Two fiber properties, strength and micronaire, were chosen in this study as examples on how to reduce locations in multiple-locations tests but still maintain the statistical power in detection of  $g \times e$  effects. The random effects of  $l$ ,  $g$ , and  $gl$  were highly significant in the original RHQ tests (Table 2). Generally, the covariance parameters of  $l$ ,  $g$ , and  $gl$  in the simulation with five or more locations were highly significant ( $p < 0.001$ ) (Tables 2, 3). When four locations were used, the parameter of  $l$  was significant at either  $p < 0.05$  or  $p < 0.01$  levels. When three or fewer locations were used, the parameter of  $l$  was less significant ( $p < 0.05$ ) or not significant.

An efficient number for locations was defined in this study as a minimum number beyond which a further reduction of locations would increase std of  $gl$ , say, 40%. In the simulation analysis, when the locations were reduced to five, the std of  $gl$  increased by 18.2 to 36.5% compared to that of the original RHQ tests with eight to nine testing locations (Table 4). When locations were reduced to four, the std of  $gl$  increased by 30.2 to 54.7% compared to that of the original. When the locations were reduced to three or fewer, the std of  $gl$  increased by 66.5 to 216% compared to that of the original. In contrast, when number of locations was reduced to seven from the original amount, the std of  $gl$  either decreased or only increased about 5% (Table 4). Therefore, seven locations should be a maximum beyond which further increase of locations would not improve statistical power in detection of  $gl$ . Five can be an efficient number of locations with a sacrifice of 18 to 37% statistical power in detection of  $gl$  compared to the maximum locations. Replication is another parameter in controlling of experimental errors. In this study, by a given number of replicates as two, varying numbers of locations were analyzed in simulation. Different results in simulations would be expected with increasing number of replications in experiment designs. However, the parameter of  $\varepsilon_{ijk}$  was less than 25% in all simulations (Tables 2, 3) and these results indicated a limited influence of the replication parameter in this study.

Table 2. Covariance parameters and standard deviation (std) for fiber strength with different numbers of locations omitted in simulation analysis<sup>z</sup>

Location <sup>y</sup> number	Mean KNm <sup>-1</sup> kg <sup>-1</sup>	$\sigma^2_1$	std ( $\sigma^2_1$ )	$\sigma^2_g$	std ( $\sigma^2_g$ )	$\sigma^2_{gl}$	std ( $\sigma^2_{gl}$ )	$\sigma^2_e$
<b>2011</b>								
8 (original)	342	103****x	62.1	810***	257	101***	24.1	177
7		105***	69.2	826***	285	96.0***	25.4	178
6		107**	87.0	821***	291	104***	29.1	177
5		96.5****	69.7	806***	251	104***	32.2	174
4		111*	97.0	781***	294	103***	36.5	175
3		102	108	819***	295	96.3***	48.0	174
2		70.5	105	780***	277	107***	70.6	172
<b>2012</b>								
9 (original)	328	287***	147	650***	210	56.2***	14.8	121
8		273***	139	681***	232	59.2***	15.0	117
7		279***	169	663***	227	54.0***	15.6	121
6		311***	221	659***	232	59.0***	18.2	120
5		265***	173	648***	201	58.5***	20.2	118
4		292**	235	627***	232	58.2***	22.9	118
3		289*	256	656***	231	55.0***	29.8	118
2		206	277	625***	215	62.5*	46.7	116
<b>2013</b>								
8 (original)	311	78.0***	48.3	139***	48.6	46.8***	11.4	76.0
7		80.0***	54.6	141***	50.8	44.7***	11.4	76.8
6		79.2***	67.8	140***	51.9	48.2***	13.0	76.4
5		73.5**	54.5	137***	45.2	48.1***	14.4	75.2
4		85.2*	75.9	134***	55.5	47.7***	16.5	75.7
3		77.2	86.8	142***	57.6	44.7***	21.4	75.4
2		55.3	79.2	135***	57.9	49.6*	31.9	74.8
<b>2015</b>								
8 (original)	331	397***	217	128***	49.2	46.6***	17.9	143
7		389***	237	130***	48.1	44.0***	16.2	144
6		424***	307	129***	50.1	48.7***	19.0	143
5		367***	239	126***	43.7	48.3***	21.3	140
4		409**	326	123***	54.1	48.3***	23.3	141
3		399*	365	131***	58.6	46.2**	29.8	140
2		286	370	123**	61.7	53.9*	43.5	138
<b>2016</b>								
8 (original)	334	107***	65.1	330***	124	152***	32.9	174
7		107***	70.6	336***	122	146***	32.0	175
6		113***	89.1	332***	124	156***	35.8	174
5		99.6***	71.4	324***	109	157***	39.5	171
4		113*	99.0	319***	136	155***	45.1	172
3		107*	106	337***	141	146***	59.5	171
2		73.4	113	324***	148	157*	89.7	170

<sup>z</sup> The means and covariance parameters were estimated by the model:  $Y_{ijk} = \mu + g_i + l_j + r_k + gl_{ij} + \varepsilon_{ijk}$  (see Materials and Methods for the variable terms).

<sup>y</sup> There were only 6 locations in RHQ tests of 2014 and this year data was not included in the analysis.

x\*, Significant at 0.05 probability level; \*\*, Significant at 0.01 probability level; \*\*\*, Significant at 0.001 probability level.

**Table 3. Covariance parameters and standard deviation (std) for micronaire with different numbers of locations omitted in simulation analysis<sup>z</sup>**

Location <sup>y</sup> number	Mean	$\sigma^2_1$	std ( $\sigma^2_1$ )	$\sigma^2_g$	std ( $\sigma^2_g$ )	$\sigma^2_{gl}$	std ( $\sigma^2_{gl}$ )	$\sigma^2_e$
<b>2011</b>								
8 (original)	4.49	0.239***x	0.131	0.030***	0.011	0.025***	0.59×10 <sup>-2</sup>	0.043
7		0.234***	0.144	0.030***	0.012	0.024***	0.62×10 <sup>-2</sup>	0.044
6		0.253***	0.186	0.030***	0.012	0.026***	0.71×10 <sup>-2</sup>	0.043
5		0.221***	0.144	0.029***	0.011	0.025***	0.79×10 <sup>-2</sup>	0.043
4		0.247**	0.197	0.029***	0.014	0.025***	0.91×10 <sup>-2</sup>	0.043
3		0.240*	0.224	0.031***	0.016	0.024***	1.15×10 <sup>-2</sup>	0.043
2		0.174	0.218	0.029*	0.018	0.026***	1.70×10 <sup>-2</sup>	0.042
<b>2012</b>								
9 (original)	4.50	0.027***	0.015	0.051***	0.018	0.018***	0.37×10 <sup>-2</sup>	0.023
8		0.027***	0.015	0.054***	0.019	0.019***	0.36×10 <sup>-2</sup>	0.023
7		0.028***	0.018	0.052***	0.019	0.018***	0.40×10 <sup>-2</sup>	0.023
6		0.028**	0.023	0.052***	0.019	0.019***	0.45×10 <sup>-2</sup>	0.023
5		0.026***	0.018	0.051***	0.017	0.019***	0.49×10 <sup>-2</sup>	0.023
4		0.029*	0.025	0.050***	0.020	0.019***	0.57×10 <sup>-2</sup>	0.023
3		0.027*	0.028	0.052***	0.021	0.017***	0.74×10 <sup>-2</sup>	0.023
2		0.019	0.027	0.050***	0.021	0.019*	1.12×10 <sup>-2</sup>	0.023
<b>2013</b>								
8 (original)	4.51	0.058***	0.036	0.023***	0.83×10 <sup>-2</sup>	0.78×10 <sup>-2</sup> ***	0.29×10 <sup>-2</sup>	0.025
7		0.059***	0.041	0.023***	0.85×10 <sup>-2</sup>	0.73×10 <sup>-2</sup> ***	0.28×10 <sup>-2</sup>	0.025
6		0.059**	0.050	0.023***	0.89×10 <sup>-2</sup>	0.81×10 <sup>-2</sup> ***	0.33×10 <sup>-2</sup>	0.025
5		0.055**	0.041	0.022***	0.78×10 <sup>-2</sup>	0.80×10 <sup>-2</sup> ***	0.37×10 <sup>-2</sup>	0.025
4		0.063*	0.056	0.022***	0.96×10 <sup>-2</sup>	0.80×10 <sup>-2</sup> ***	0.40×10 <sup>-2</sup>	0.025
3		0.058	0.067	0.023***	1.04×10 <sup>-2</sup>	0.76×10 <sup>-2</sup> **	0.51×10 <sup>-2</sup>	0.025
2		0.043	0.060	0.022**	1.08×10 <sup>-2</sup>	0.90×10 <sup>-2</sup> *	0.74×10 <sup>-2</sup>	0.025
<b>2015</b>								
8 (original)	4.51	0.295***	0.160	0.021***	0.95×10 <sup>-2</sup>	0.024***	0.66×10 <sup>-2</sup>	0.043
7		0.288***	0.176	0.022***	0.90×10 <sup>-2</sup>	0.022***	0.61×10 <sup>-2</sup>	0.044
6		0.315***	0.229	0.021***	0.96×10 <sup>-2</sup>	0.024***	0.70×10 <sup>-2</sup>	0.044
5		0.273***	0.176	0.021***	0.86×10 <sup>-2</sup>	0.024***	0.78×10 <sup>-2</sup>	0.043
4		0.303**	0.241	0.021**	1.11×10 <sup>-2</sup>	0.024***	0.89×10 <sup>-2</sup>	0.043
3		0.298*	0.271	0.022***	1.23×10 <sup>-2</sup>	0.022***	1.12×10 <sup>-2</sup>	0.043
2		0.216	0.271	0.021*	1.53×10 <sup>-2</sup>	0.025*	1.61×10 <sup>-2</sup>	0.043
<b>2016</b>								
8 (original)	4.46	0.100***	0.057	0.062***	0.024	0.046***	0.86×10 <sup>-2</sup>	0.036
7		0.099***	0.062	0.063***	0.024	0.045***	0.86×10 <sup>-2</sup>	0.036
6		0.105***	0.080	0.062***	0.024	0.047***	0.96×10 <sup>-2</sup>	0.036
5		0.093***	0.062	0.060***	0.021	0.047***	1.04×10 <sup>-2</sup>	0.035
4		0.105**	0.087	0.059***	0.027	0.047***	1.21×10 <sup>-2</sup>	0.036
3		0.099*	0.096	0.064***	0.029	0.044***	1.54×10 <sup>-2</sup>	0.036
2		0.071	0.094	0.062**	0.032	0.047**	2.39×10 <sup>-2</sup>	0.035

<sup>z</sup> The means and covariance parameters were estimated by the model:  $Y_{ijk} = \mu + g_i + l_j + r_k + gl_{ij} + \varepsilon_{ijk}$  (see Materials and Methods for the variable terms).

<sup>y</sup> There were only 6 locations in RHQ tests of 2014 and this year data was not included in the analysis.

x\*, Significant at 0.05 probability level; \*\*, Significant at 0.01 probability level; \*\*\*, Significant at 0.001 probability level.

**Table 4.** Percentage increase<sup>z</sup> of standard deviation (std) for fiber strength and micronaire on each testing location omission

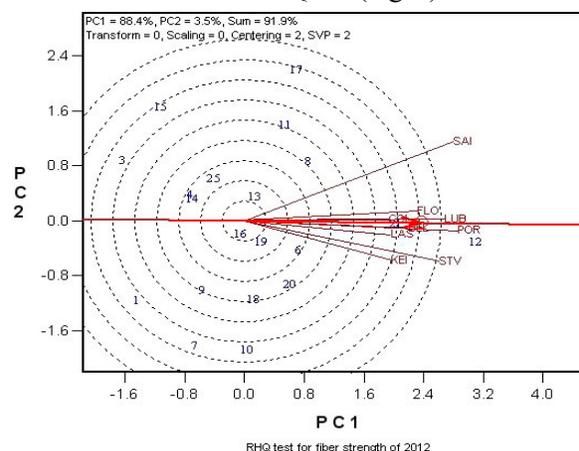
Location number <sup>y</sup>	Strength	Mic
<b>2011</b>		
7	5.39	5.09
6	20.7	20.3
5	33.6	33.9
4	51.5	54.2
3	99.2	94.9
2	193	188
<b>2012</b>		
8	1.35	-2.70
7	5.41	8.11
6	22.9	21.6
5	36.5	32.4
4	54.7	54.1
3	101	100
2	216	203
<b>2013</b>		
7	0	-3.45
6	14.0	13.8
5	26.3	27.6
4	44.7	37.9
3	87.7	75.9
2	180	155
<b>2015</b>		
7	-9.50	-7.58
6	6.15	6.06
5	19.0	18.2
4	30.2	34.9
3	66.5	70.0
2	143	144
<b>2016</b>		
7	-2.74	0
6	8.82	11.6
5	20.1	20.9
4	37.1	40.7
3	80.9	79.1
2	173	178

<sup>z</sup> Comparisons were made between the std of *gl* for each location omission in the simulation and the std of the original test.

<sup>y</sup> There were only 6 locations in RHQ tests of 2014 and this year data was not included in the analysis.

It would be desirable to know which testing locations should remain in the RHQ tests if the number of locations in the tests were reduced. This requires a determination of relationships among testing locations to identify locations that are most representative and

informative for selecting genotypes. In GGE biplot analysis, the testing locations with shortest distances to the ideal environment can be identified as most representative and discriminating for HVI fiber properties. Fiber properties such as strength, length, length uniformity, elongation, and micronaire were analyzed for the relationships among testing locations. The distances between testing locations and the ideal environment were obtained from the log sheets in GGE biplot analysis (Table 5). The testing locations were ranked from the top as the one with the shortest distances to the bottom as the one with the longest distances. The locations of Belle Mina and Jackson were not included in the final rankings because their limited participation during the period analyzed. Among the remaining eight locations, College Station, Keiser, Lubbock, and Stoneville had the shortest distances to the ideal environment for different fiber properties. Stoneville was ranked at least third for all five fiber properties and Keiser was ranked first for strength and fiber length and second for micronaire. Lubbock was ranked first and second for elongation and fiber length, respectively. College Station was ranked first for length uniformity and third for micronaire. Among these locations, Stoneville and Lubbock also were identified as most representative for evaluation of nitrogen content in a previous study of cotton seed traits using GGE biplot (Zeng et al., 2018). Other locations such as St. Joseph, LA can be a site with good discriminating ability among genotypes for fiber strength under a target environment as shown in the 2012 RHQ test (Fig. 1).



**Figure 1.** GGE biplot view of testing locations ranking based on discriminating ability and representativeness in the 2012 RHQ test for fiber strength.

Arabic numbers are genotypes under evaluation. Uppercase letters are testing locations: BEL, Belle Mina, AL; COL, College Station, TX; FLO, Florence, SC; KEI, Keiser, AR; LAS, Las Cruces, NM; POR, Portageville, MO; LUB, Lubbock, TX; SAI, Saint Joseph, LA; STV, Stoneville, MS. The point at the center of the concentric circles is the ideal environment that is the assuming environment with the shortest distance to the AEA line (representativeness) and the longest environment vector (discriminating ability). The more detailed interpretation of the graph was described in Materials and Methods.

**Table 5. Distances<sup>z</sup> between testing locations and the 'ideal environment' in the RHQ tests for fiber properties**

<b>Strength</b>							
<b>Locations<sup>y</sup></b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2015</b>	<b>2016</b>	<b>Avg.</b>	<b>Rank</b>
BEL	0.312	0.830	---- <sup>x</sup>	----	----	0.571	NA <sup>w</sup>
COL	0.659	1.10	1.10	0.504	1.04	0.881	5
FLO	1.91	0.766	0.923	0.965	1.71	1.25	7
JAC	----	----	----	----	0.913		NA
KEI	0.550	1.15	0.334	0.714	0.661	0.682	1
LAS	0.467	1.06	1.48	0.739	----	0.937	6
LUB	0.377	0.381	0.968	1.35	0.859	0.787	4
POR	1.25	0.161	1.04	1.18	0.100	0.746	2
SAI	----	1.31	1.01	2.07	2.83	1.81	8
STV	0.353	0.601	0.652	1.31	0.892	0.762	3
<b>Length</b>							
BEL	0.137	0.115	----	----	----	0.126	NA
COL	0.152	0.122	0.007	0.204	0.004	0.098	4
FLO	0.491	0.006	0.005	0.009	0.084	0.119	6
JAC	----	----	----	----	0.003		NA
KEI	0.141	0.009	0.008	0.009	0.003	0.034	1
LAS	0.150	0.005	0.139	0.175	----	0.117	5
LUB	0.007	0.140	0.065	0.165	0.005	0.076	2
POR	0.176	0.120	0.161	0.123	0.136	0.143	7
SAI	----	0.108	0.101	0.285	0.629	0.281	8
STV	0.156	0.110	0.002	0.123	0.006	0.079	3
<b>Uniformity</b>							
BEL	0.235	0.463	----	----	----	0.349	NA
COL	0.183	0.801	0.006	0.821	1.16	0.594	1
FLO	1.19	0.505	1.04	0.691	0.006	0.640	3
JAC	----	----	----	----	0.523		NA
KEI	0.222	1.04	1.21	0.866	1.06	0.880	6
LAS	0.615	0.507	1.20	0.927	----	0.812	5
LUB	0.168	0.689	0.560	0.762	1.63	0.762	4
POR	0.921	0.729	0.787	0.788	1.57	0.959	7
SAI	----	0.768	0.660	1.74	1.97	1.29	8
STV	0.507	0.591	0.551	1.07	0.422	0.628	2
<b>Elongation</b>							
BEL	0.223	0.300	----	----	----	0.262	NA
COL	0.351	0.947	0.483	0.973	0.603	0.671	7
FLO	0.971	0.453	0.341	0.397	0.505	0.533	5
JAC	----	----	----	----	0.391		NA
KEI	0.550	0.522	0.408	0.374	0.575	0.486	4
LAS	0.119	0.718	0.446	0.299	----	0.396	3
LUB	0.267	0.450	0.042	0.373	0.009	0.228	1
POR	0.498	0.514	0.376	0.589	0.707	0.537	6
SAI	----	0.789	0.236	1.79	1.94	1.19	8
STV	0.323	0.586	0.338	0.297	0.364	0.382	2
<b>Micronaire</b>							
BEL	0.529	0.378	----	----	----	0.454	NA
COL	0.203	0.413	0.354	0.690	0.142	0.360	3
FLO	0.495	0.307	0.310	0.229	0.781	0.424	6
JAC	----	----	----	----	0.337		NA
KEI	0.231	0.534	0.290	0.337	0.341	0.347	2
LAS	0.349	0.377	0.454	0.401	----	0.395	5
LUB	0.312	0.487	0.623	0.009	0.487	0.384	4
POR	0.865	0.406	0.447	0.724	0.458	0.580	7
SAI	----	0.669	0.377	0.988	1.41	0.861	8
STV	0.450	0.430	0.325	0.262	0.200	0.333	1

<sup>z</sup>The distances between testing locations and the 'ideal environment' were obtained from log files in Discriminating Ability vs. Representativeness module of GGE biplot analysis<sup>1</sup>.

<sup>y</sup>BEL, Belle Mina, AL; COL, College Station, TX; FLO, Florence, SC; JAC, Jackson, TN; Keiser, AR; Las Cruces, NM; Lubbock, TX; Portageville, MO; Saint Joseph, LA; Stoneville, MS. The testing locations of Belle Mina, AL and Jackson, TN were not included in final rankings because they only participated the tests one to two years during the analyzing period. There were only six locations in 2014 and the test of this year was not included in GGE analysis to avoid possible bias.

<sup>x</sup>Data are not available.

<sup>w</sup>The testing locations of Belle Mina, AL and Jackson, TN were not included in the final rankings because of their limited participations during the analyzing period. The remaining locations were ranked from 1 to 8 with those of the smallest averages as 1.

## CONCLUSIONS

Results in this study showed that seven locations can be a maximum and every single location omitted from that would reduce statistical power in detection of *gl*. Thus, more testing locations below that number would be better for evaluation of fiber quality. However, five testing locations can be considered as a minimum in the RHQ tests for fiber quality and the actual number of testing locations can be decided by breeders as a balance between accuracy in breeding and cost. From a statistical perspective, the variable *l* was treated as a random effect in the random model in this study, thus, the conclusion of five locations as a minimum number of locations can be generalized to other multiple-location tests for cotton fiber quality. In actual breeding, considering the diverse environments in the U.S. Cotton Belt, the testing locations selected in the RHQ tests would not represent all regional environments. However, this result can be used as a reference in the experimental designs to determine the efficient number of locations in the regional tests under specific target environments.

The testing locations of Stoneville, Keiser, Lubbock, and College Station were identified as most representative and discriminating of genotypes for fiber properties. Considering the nature of *ge* interactions, it would be unrealistic to expect a consistent ranking of environments across testing years and different fiber traits. However, it is concluded that Stoneville, Keiser, Lubbock, and College Station were the four most representative and discriminating sites for fiber properties in the recent testing years between 2011 and 2016. Therefore, under the situations when breeders need to limit testing locations to a minimum efficient number, these four locations should not be among the list to eliminate.

## ACKNOWLEDGMENTS

This research was funded by USDA-ARS, Project No. 6066-21000-052-00D. USDA is an equal opportunity provider and employer. We thank Patricia Maugh, Crop Genetics Res. Unit at Stoneville, MS for her excellent technical assistance.

## DISCLAIMER

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation by the US Department of Agriculture.

## REFERENCES

- Campbell, B.T., P.W. Chee, E. Lubbers, D.T. Bowman, W.R. Meredith, Jr., J. Johnson, D. Fraser, W. Bridges, and D.C. Jones. 2012. Dissecting genotype  $\times$  environment interactions and trait correlations present in the Pee Dee cotton germplasm collection following seventy years of plant breeding. *Crop Sci.* 52:690–699.
- De Leon, N., J. Jannink, J.W. Edwards, and S.M. Kaeppler. 2016. Introduction to a special issue on genotype by environment interaction. *Crop Sci.* 56:2081–2089.
- Gauch, H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Sci.* 46:1488–1500.
- Meredith, W.R. Jr., D.L. Boykin, F.M. Bourland, W.D. Caldwell, B.T. Campbell, J.R. Gannaway, K. Glass, A.P. Jones, L.M. May, C.W. Smith, and J. Zhang. 2012. Genotype  $\times$  environment interactions over seven years for yield, yield components, fiber quality, and gossypol traits in the Regional High Quality tests. *J. Cotton Sci.* 16:160–169.
- National Cotton Variety Test [NCVT]. 2016. 2016 National Cotton Variety Test. Yield, boll, seed, spinning, and fiber data. Available online at <https://www.ars.usda.gov/south-east-area/stoneville-ms/crop-genetics-research/docs/national-cotton-variety-test> (verified 18 Nov. 2019).
- SAS Institute, 2013. SAS/STAT. User's guide. Version 9.4. SAS Inst., Cary, NC.
- Smith, A.B., B.R. Cullis, and R. Thompson. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *J. Agric. Sci.* 143:449–462.
- Yan, W. 2001. GGEbiplot – A window application for graphical analysis of multi-environment trial and other types of two-way data. *Agron. J.* 93:1111–1118.
- Yan, W., and N.A. Tinker. 2006. Biplot analysis of multi-environmental trial data: Principles and applications. *Can. J. Plant Sci.* 86:623–645.
- Zeng, L., W.R. Meredith, Jr., B.T. Campbell, J.K. Dever, J. Zhang, K.M. Glass, A.S. Jones, G.O. Myers, and F.M. Bourland. 2014. Genotype-by-environment interactions effects on lint yield of cotton cultivars across major regions in the U.S. Cotton Belt. *J. Cotton Sci.* 18:75–84. ]
- Zeng, L., B.T. Campbell, E. Bechere, J.K. Dever, J. Zhang, A.S. Jones, T.B. Raper, S. Hague, W. Smith, G.O. Myers, and F.M. Bourland. 2015. Genotypic and environmental effects on cottonseed oil, nitrogen, and gossypol contents in 18 years of regional high quality tests. *Euphytica* 206:815–824.
- Zeng, L., W.C. Bridges, Jr., and F.M. Bourland. 2019. Testing locations in regional high quality tests for cotton seed quality traits. *J. Cotton Sci.* 23:1–12.