

Kansas State University Libraries
New Prairie Press

Conference on Applied Statistics in Agriculture 2012 - 24th Annual Conference Proceedings

STATISTICAL TESTS FOR STABILITY ANALYSIS WITH RESAMPLING TECHNIQUES

Jixiang Wu

Karl Glover

William Berzonsky

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

Recommended Citation

Wu, Jixiang; Glover, Karl; and Berzonsky, William (2012). "STATISTICAL TESTS FOR STABILITY ANALYSIS WITH RESAMPLING TECHNIQUES," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1030>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

Statistical Tests for Stability Analysis with Resampling Techniques

Jixiang Wu, Karl Glover, and William Berzonsky

Plant Science Department, South Dakota State University, Brookings, SD 57007

Abstract

Crop trials or crop performance trials (CPT), which are among the most important activities associated with plant breeding programs, are commonly used to measure the performance stability of genotypes. Several methods which include variation, regression, and cluster analyses for determination of crop stability have been proposed and are commonly used. However, many of these approaches require the use of normally distributed data. Thus, commonly used statistical tests, like the t- or F-test may not be appropriate when the assumptions of data are violated. In this study, two resampling techniques (jackknife and bootstrapping) were integrated into several crop stability analyses. An upland cotton data set from China was analyzed to demonstrate the utility of these methods in measuring performance stability.

1. Introduction

Genotype-by-environment ($G \times E$) interactions have been a common issue in developing widely adapted crop cultivars (Gray, 1982; Kang and Miller, 1984). A cultivar associated with high $G \times E$ interactions indicates that it is sensitive to various environmental conditions and it is has low performance stability. Thus, fully investigating $G \times E$ interactions or stability is critical to determining cultivar adaptation to specific markets.

Crop trials or crop performance tests (CPT) at multiple locations and possibly in multiple years are executed to generate experimental data for measuring genotype. Due to the complexity of field trials and various definitions of yield stability (Lin et al., 1986), many different

statistical methods for measuring stability have been proposed. Based on review paper (Lin et al., 1986), these methods can be clustered into two basic categories: variation-based (Francis and Kannenberg, 1978; Plaisted and Peterson, 1959; Shukla, 1972; Wricke, 1962) and regression-based (Eberhart and Russell, 1966; Finlay and Wilkinson, 1963; Perkins and Jinks, 1968). Both variation and regression-based methods are more focused on yield consistency across environments. It appears more appropriate that genotype yield performance and yield stability be considered simultaneously (Fan et al., 2007; Kang and Mangari, 1995; Kang, 1993). Another commonly used approach has been the additive main effects and multiplicative interaction (AMMI) method (Crossa et al., 1990). Based on principal component analysis, as it treats multi-environmental data as multivariate data structures, this method considers these two important parameters as well.

The aforementioned methods use data that are balanced (Crossa et al., 1990; Kang, 1993; Lin et al., 1986); however, many crop trial data may include missing data points due to various reasons (Zhu et al., 1993), or data may be unbalanced due to some genotypes not being trialed in successive years. It is also possible that error variances are heterogeneous among environments (Edwards and Jannink, 2006). In addition, many statistical tests for stability parameters are related to t-tests and/or F-tests (Kang, 1993), which assume data are normally distributed. Additionally, the environmental index is used as an independent variable in regression-based stability analyses, and as a linear function of observations, it is a random variable rather than a fixed one such as is required by a linear regression analysis. Conventional t-tests/F-tests may not be appropriate when an independent variable is also a random variable. Furthermore, breeders and/or farmers are also interested in stability being something that can be compared between

genotypes. Thus, it is necessary that more general approaches and statistical tests be added to the existing group of stability analysis methods.

Mixed linear model approaches have been commonly used to analyze various data structures, including missing and unbalanced data structures (Little et al., 2006; Rao, 1971; Wu et al., 2006; Zhu, 1998). On the other hand, resampling techniques such as those exemplified by the jackknife approach (Miller, 1974; Wu et al., 2008; Wu et al., 2012; Zhu et al., 1993), bootstrapping (Davison and Hinkley, 1997; Efron, 1979), and permutation (Manly, 2006) have been widely used for statistical tests for various parameters. In this study, we focused on variance components, $G \times E$ interactions, genotypic means/ranks, and regression based stability for crop trial data. A mixed linear model approach (Wu et al., 2006; Zhu, 1989; Zhu, 1993; Zhu, 1998) was used to estimate variance components and to predict $G \times E$ interactions. A regression-based method was used to determine the stability of each genotype (Finlay and Wilkinson, 1963). Jackknife and bootstrapping techniques were used to test the significance of each parameter or conduct multiple comparisons among parameters (Davison and Hinkley, 1997; Efron, 1979; Wu et al., 2008; Wu et al., 2012). An upland cotton data set (Zhu et al., 1993) was used to demonstrate the utility of using these methods for stability analysis and statistical tests.

2. Materials, Models, and Methods

2.1. Materials

The data used in the present study was cotton lint yield, reported by Zhu et al. (1993). The data were collected in 1989 and 1990 from eight varieties grown in 24 locations in the Yellow River Region of China. Cultivar Zhong 206 was missing at one location in 1989 and, due to weather problems in 1990, at one location was abandoned and not harvested. Only mean values of each environment for each cultivar were available for analyses.

2.2. Genetic models

As a consequence of different data structures and experimental designs, there are several different models that can be used for analyzing crop trial or multi-environmental data. In this study, we employed four genetic models based on if field plots were replicated (randomized complete block design) and if experiments were repeated between years.

Model 1:

Assuming that g genotypes (cultivars) are grown in multiple years and locations with r replications in each environment (a combination of year and location), a mixed linear model can be expressed as follows

$$y_{hijk} = \mu + Y_h + L_i + G_j + YL_{hi} + GY_{hj} + GL_{ij} + GYL_{hij} + B(YL)_{k(hi)} + e_{hijk} \quad (1)$$

Where y_{hijk} is an observation; μ is population mean, Y_h is year effect; L_i is location effect; G_j is genotypic (variety) effect; YL_{hi} is year-by-location interaction effect; GY_{hj} is genotype-by-year interaction effect; GL_{ij} is genotype-by-location interaction effect; GYL_{hij} is genotype-by-year-by-location interaction effect; $B(YL)_{k(hi)}$ is block effect with year and location; and e_{hijk} is random error.

Model 2:

When only mean plot values without replication in each environment are used for analysis, then each of the block effects should be dropped from *Model 1*. In addition, because each genotype-by-year-by-location interaction effect is confounded with random error, this term should be dropped from *Model 1* too. Thus, *Model 1* can be rewritten as in (2).

$$y_{hij} = \mu + Y_h + L_i + G_j + YL_{hi} + GY_{hj} + GL_{ij} + e_{hij} \quad (2)$$

Definitions of all terms in equation (2) are the same as in *Model 1*. It must be noted that the random error e_{hij} in *Model 2* may contain a genotype-by-year-by-location interaction effect mentioned in *Model 1*, and the data set in this study can be analyzed with this model by use of a mixed linear model approach described in section 2.3.

Model 3:

If g genotypes are grown in multiple locations in one year with plot repetition s , *Model 2* can be simplified and represented as equation (3).

$$y_{ijk} = \mu + L_i + G_j + GL_{ij} + B_{k(i)} + e_{ijk} \quad (3)$$

The terms in equation (3) are defined in equation (1), and this model can be used for a modified Skukla's stability analysis (Kang and Mangari, 1995; Kang, 1993) and AMMI stability analysis (Crossa et al., 1990).

Model 4:

If genotypes are grown in multiple locations in a single year, with only genotype mean values recorded, then *Model 3* can be modified as represented in equation (4). The random error term e_{ij} in equation (4) includes a possible genotype-by-location effect.

$$y_{ij} = \mu + L_i + G_j + e_{ij} \quad (4)$$

Though the model in equation (4) cannot be used to detect genotype-by-location effects, it is commonly used for various stability analyses, such as variation-based analyses (Francis and Kannenberg, 1978; Plaisted and Peterson, 1959; Shukla, 1972; Wricke, 1962), regression-based

analyses (Eberhart and Russell, 1966; Finlay and Wilkinson, 1963), modified Skukla's stability analyses (Kang and Mangari, 1995; Kang, 1993), and AMMI stability analyses (Crossa et al., 1990). If multi-year data are to be used, it is necessary that both Models 1 and 2 be converted to *Model 3* or *Model 4* (Lin et al., 1986).

2.3.Methods

Variance component estimation and random effect prediction

It is often the case that crop trial data does not include all genotypes at all environments, or trials are not at the same locations every testing year. It is also possible that some plots in specific environments are missing (Zhu et al., 1993). Thus, many experimental data are not balanced and as a consequence, analysis of variance (ANOVA) methods are not applicable (Zhu, 1998). On the other hand, mixed linear model approaches, such as maximum likelihood (Hartley and Rao, 1967), restricted maximum likelihood (Patterson and Thompson, 1971), and minimum norm quadratic unbiased estimation (MINQUE) (Rao, 1971; Wu et al., 2010; Wu et al., 2006; Zhu, 1989; Zhu, 1998) methods are more generalized and appropriate for various data structures. In this study, a MINQUE approach with all prior variance components equal to 1 was employed for variance component estimation (Zhu, 1989). The reason to use a MINQUE approach is that it does not require data being normally distributed (Rao, 1971). Effects were predicted by applying an adjusted unbiased prediction method (Zhu, 1993). A jackknife technique with a randomized 10-group jackknife can be used to detect significance of each parameter of interest (Miller, 1974; Wu et al., 2008; Wu et al., 2012).

Stability analyses

In stability analysis, we only focused on genotypic mean ranks and means, and slopes (b_1) and coefficients of determination (R^2) related to regression-based analysis for all genotypes

(Finlay and Wilkinson, 1963; Lin et al., 1986). Data were converted by setting the combinations of years and locations as environments so *Model 4* could be employed for these mentioned analyses.

The bootstrapping technique (Davison and Hinkley, 1997; Efron, 1979) was employed to estimate bootstrap means and their corresponding 2.5% and 97.5% percentiles as the low limit (LL) and upper limit (UL) of 95% confidence interval for each parameter of interest by repeating 1000 times. Multiple comparisons for parameters of interest can be made based on 95% confidence intervals among different cultivars. The whole yield data set (sample size=375) was resampled with replacement for estimating genotypic means and ranks, while 47 environmental indexes and the phenotypic means in all 47 environments for each cultivar were resampled with replacement for regression-based stability analysis.

2.4. Data analysis

Model 2 was employed to estimate variance components and to predict genotypic effects and genotype-by-location (cultivar-by-location) interaction effects, which are reported in Tables 2, 3, and 4. Data were also converted so that *Model 4* could be used for estimation of genotypic means/ranks and regression-based stability analysis.

All data analyses were conducted by an R package GenMod developed by Jixiang Wu at South Dakota State University.

3. Results

3.1. Genotypic means and ranks

Genotypic means and rank values and their corresponding 95% confidence intervals for each of eight cultivars are provided in Table 1. The results demonstrated

that there were significant differences among cultivars for cotton lint yield. Cultivar Zhong 206 was significantly higher than cultivars Ji 84-25, Yun1929, and Lu 155, while no significant differences were detected compared with the remaining cultivars. Cultivar Ji 84-25 was significantly lower for lint yield than all other varieties, except Lu 155. The relative genotypic rankings were similar to for those for genotypic means (Table 1). For example, cultivar Zhong 206 was higher than Ji 84-25, Yun 1929, and Lu 155, while not different from cultivars Ping 28, Shi3409, Jizhi 17, or Zhong 12. Cultivar Ji 84-25 was ranked last among all cultivars for lint yield..

3.2. Variance components for upland cotton yield

In our study, we treated all effects as random because we were interested in both variance components and effects that can be calculated (Zhu, 1989; Zhu, 1993). Variance components and their corresponding proportions for cotton lint yield were estimated based on Model 2 and the results are summarized in Table 2.

Significant year effects, genotype effects, year-by-location interaction effects, and genotype-by-location interaction effects were detected, while there were no location or genotype-by-year interaction effects for cotton lint yield (Table 2). Among all effects, year-by-location interaction effects contributed the most (60%) to the total variance. Genotypic effects contributed 13.3% to the total variance, and random errors contributed 17.2% or less, because year-location-genotype interaction effects were confounded with random errors.

3.3. Genotypic effects for eight upland cotton cultivars

Predicted genotypic values (population mean $\hat{\mu}$ +genotype effect \widehat{G}_j , first column), genotypic effects \widehat{G}_j (second column), and 95% confidence intervals (3rd and 4th columns) for genotype effects are provided in Table 3. Genotypic values and genotypic effects are equivalent, but genotypic values are more straightforward, while genotype effects are deviations from the population mean (the estimated population $\hat{\mu}=60.53$). Comparing the results in Tables 1 and 3, we observed that predicted genotype values (Table 3) and bootstrapping genotypic means (Table 1) were consistent. In addition, the multiple comparisons among genotype effects showed similar patterns. Thus, both jackknife and bootstrapping tests can be applied to obtain similar results.

3.4. Genotype-by-location interaction effects

Among many G×E interaction effects, genotype-by-location interaction effects are one of the most important since environmental conditions for specific locations are relatively more predictable. For example, soil types, temperatures, and rainfall, the latter two of which can vary significantly year-to-year, are generally consistent between locations within the same year. With predicted genotype-by-location interaction effects, breeders can determine genotypes that are specifically adapted to some locations or similar environments. The interaction effects between eight genotypes and 24 locations are listed in Table 4. The variance of genotype-by-location interaction effects for each cultivar is also provided in this table (the last line of Table 4).

Cultivar Ping 28 had the largest variance for genotype-by-location interaction effects among all eight cultivars, indicating that this cultivar had the lowest stability across these 24 locations. Thus, a breeder might predict that this cultivar would perform

better at several specific locations and/or similar locations. For example, to maximize its yield, Ping 28 could be grown in locations 7, 9, and 9, while not at location 1 or 17.

3.5. Regression-based stability analysis

Regression based stability analysis (Finlay and Wilkinson, 1963) resulted in providing two types of parameters: slopes (b_1) and coefficients of determination (R^2) for all genotypes. A low slope (less than 1) suggests that a genotype has lower than average sensitivity to the environmental indexes; whereas, a slope equal to 1 suggests it has an average response, and a slope greater than 1 indicates that it exhibits higher than average sensitivity to environmental indexes. A coefficient of determination of a genotype is a measure of its performance stability. A high R^2 indicates that a genotype has a high stability of performance across the testing environments. With 1000 times bootstrapping, the low and upper limits (2.5% and 97.5% percentiles, respectively) for each slope and coefficient were generated. These results are summarized in Table 5.

Slopes for cultivars Ping 28 and Lu 155 were significantly less than 1, from which one would predict these they are less sensitivity than average to the environmental index. On the other hand, the slopes for the remaining cultivars were equal to 1, and not significantly different from 1, indicating these cultivars had an average response to environmental indices. Cultivar Shi 3409, Jizhi 17, Zhong 206, and Zhong 12 had higher slopes than Ping 28 and Lu 155, and four cultivars: Shi 2409, Jizhi 17, Zhong 206, and Zhong 12 had similar stabilities, all greater than 0.80. These four cultivars also expressed a higher stability than Ping 28, which was less than 0.70, numerically the lowest stability among all cultivars.

Comparing several methods, we observed that jackknife and bootstrapping techniques both gave consistent results for genotype means/values, indicating that either resampling method worked well for predicting stability of performance. A regular analysis for estimating genotype means integrated with bootstrapping is a reliable method for predicting stability compared to a complicated mixed linear model analysis. The genotype-by-location interaction (Table 4) for Ping 28 also indicated lower stability and less sensitivity to environmental indexes (Table 5), suggesting G×E interaction effects can be used as another measure of the stability for a genotype.

4. Discussion

Regression-based stability analysis is one of the most commonly used methods (Finlay and Wilkinson, 1963; Lin et al., 1986). Since an environmental index is a linear function of observations and it is a random, independent variable, a traditional t-test for the slope is not appropriate. With bootstrapping, this issue can be easily resolved. In addition, the confidence intervals for parameters of each genotype can be provided such that multiple comparisons among genotypes can be made.

As mentioned by Lin et al. (1986), the results and conclusions are dependent on the environments and genotypes used in trials since environmental indices are used for regression analyses. Having common genotype checks grown in different years will help in making decisions based on performance. In addition, the number of environments used in regression-based stability analysis should be enough so as to provide accurate predictions. If the number of environments is small, examining G×E interaction effects (Gray, 1982; Kang and Miller, 1984; Zhu, 1998) or using the AMMI approach (Cossa et al., 1990) might represent useful alternatives.

However, the AMMI approach is currently limited to the assumption that trial data are balanced. It might be interesting to analyze unbalanced data using an AMMI approach.

As a practical consideration, many trial data are not balanced, and mixed linear model approaches are a generalized way to analyze various data structures (Little et al., 2006; Rao, 1971; Zhu, 1998). With mixed linear model approaches, both genotype effects/values and $G \times E$ interaction effects can be predicted (Zhu, 1993). Genotype effects/values can be used to determine the averaged performance of cultivars over environments. Additionally, with $G \times E$ interaction effects, breeders can determine which genotypes are adapted to specific environmental conditions. For example, a significant positive $G \times E$ interaction effect suggests a genotype is adapted to specific environments or locations. Our results showed that information on $G \times E$ interaction effects can be a useful in determining the yield stability of a cultivar, because they can be predicted by mixed linear model approaches, even from an unbalanced data set.

Environmental conditions can be classified into predictable and unpredictable factors. Soils types can be more predictable in each test location and field practices can be fixed as compared to temperatures and rainfalls between years. If those more predictable environmental factors play a significant role on yield performance, the interactions between genotypes and those factors should be very useful to achieve for maximum yield. If interactions between genotypes and those unpredictable factors are significant, widely adapted genotypes should be considered in breeding selection. For example, two maize hybrids were highly dependent on environmental conditions between two years (Fan et al., 2007) .

There are several ways to boot a multi-environment data set and the results could be different from one to another. One way is to resample all observations with replacement and

another one is within each environment. The first one was used to estimate genotypic means/ranks and their standard errors for these parameters. It appeared that the bootstrapped means were very close to means based on the whole data set. In addition, we employed the second one to estimate these parameters and the results were almost identical (results not shown). The environmental indexes used for regression-based stability analysis are the same for each genotype and environmental indexes and phenotypic means in all environments were bootstrapped. Therefore, the bootstrapping process for regression-based stability analysis was different from the one used for genotypic mean/rank analysis.

In this study, we employed a jackknife technique to estimate standard errors for all variance components and effects while bootstrapping methods for genotypic means/ranks and regression-based stability analysis. Based on our limited experience, it seems that jackknife techniques are more appropriate than bootstrapping methods on observations for estimations of variance components and predictions of random effects. A possible major reason is that each resampled data set will contain some missing identities like genotypes and several repeated observations. Such resampled data could lead to unfavorable biased estimations/predictions and increase large standard errors for those parameters. It might be workable to resample on individual residuals with replacement. On the other hand, we also observed that a 10-fold randomized jackknife approach was an effective way to integrate with a MINQUE approach to reasonably limit Type I error and maintain testing power when (Wu et al., 2012). Multiple comparisons for genotype means and genotypic effects were very consistent, which suggests that the two resampling approaches (jackknife and bootstrapping) can serve equally well for making mean comparisons among genotypes.

Generally, the genotypes used in crop trials are considered as fixed effects. In our data analysis, we treated all components, except the population mean as random, so that both variance components and effects could be calculated (Wu et al., 2006; Zhu, 1989; Zhu, 1993; Zhu, 1998). Although, our statistical statements and conclusions were limited to the use of these specific genotypes, our analysis showed that the predicted genotype effects were consistent to the results when genotypes are treated as fixed effects (Zhu et al., 1993).

References

- Crossa J., Gauch H.G., Zobel R.W. (1990) Additive Main Effects and Multiplicative Interaction Analysis of Two International Maize Cultivar Trials. *Crop Science* 30:493-500.
- Davison A.C., Hinkley D.V. (1997) *Bootstrap methods and their applications* Cambridge University Press, New York, NY.
- Eberhart S.A., Russell W.A. (1966) Stability parameters for comparing varieties. *Crop Science* 6:36-40.
- Edwards J.W., Jannink J.L. (2006) Bayesian modeling of heterogeneous error and genotype X environment interaction variances. *Crop Science* 46:820-833.
- Efron B. (1979) Bootstrap methods: another look at the jackknife. *Ann. Statist.* 7:1-26.
- Fan X.M., Kang M.S., Chen H.M., Zhang Y.D., Tan J., Xu C.X. (2007) Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. *Agronomy Journal* 99:220-228.
- Finlay K.W., Wilkinson G.N. (1963) The analysis of adaptation in a plant-breeding programme. *Aust. J. Agric. Res.* 14:742-752.
- Francis T.R., Kannenberg L.W. (1978) Yield stability studies in short season maize. I. A descriptive method for grouping genotypes. *Can. J. Plant Sci.* 58:1029-1034.
- Gray E. (1982) Genotype \times environment interactions and stability analysis for forage yield of orchardgrass clones. *Crop Science* 22:19-23.
- Hartley H.O., Rao J.N.K. (1967) Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika* 54:93-108.
- Kang K.S., Mangari R. (1995) Stable: A basic program for calculating stability and yield-stability statistics. *Agronomy Journal* 87:276-277.

- Kang M.S. (1993) Simultaneous selection for yield and stability: Consequences for growers. *Agronomy Journal* 85:754-757.
- Kang M.S., Miller J.D. (1984) Genotype \times environment interactions for cane and sugar yield and their implications in sugar breeding. *Crop Science* 24:435-440.
- Lin C.S., Binns M.R., Lefkovitch P.L. (1986) Stability analysis: where do we stand? *Crop Science* 26:894-899.
- Little R.C., Milliken G.A., Stroup W.T., Wolfinger R.D., Schabenberger O. (2006) SAS for Mixed Models. SAS Institute Inc., Cary, NC.
- Manly B.F.J. (2006) Randomization, bootstrap and Monte Carlo methods in biology. 3 ed. Chapman and Hall/CRC, Boca Raton, FL.
- Miller R.G. (1974) The jackknife-a review. *Biometrika* 61:1-15.
- Patterson H.D., Thompson R. (1971) Recovery of inter-block information when block size are unequal. *Biometrika* 58:545-554.
- Perkins J.M., Jinks J.L. (1968) Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. *Heredity* 23:18.
- Plaisted R.L., Peterson L.C. (1959) A technique for evaluating the ability of selections to yield consistently in different locations or seasons. *Amer. Potato J.* 36:381-385.
- Rao C.R. (1971) Estimation of variance and covariance components-MINQUE theory. *J. Multiva. Ana.* 1:19.
- Shukla C.K. (1972) Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* 29:237-245.
- Wricke G. (1962) Uber eine Methode zur Erfassung der okologischen Streubreite in Feldversuchen. *Z. Pflanzenzuchtg* 47:92-96.
- Wu J., Jenkins J.N., McCarty J.C. (2008) Testing variance components by two jackknife techniques. *Proceedings of Applied Statistics in Agriculture*:1-17.
- Wu J., Jenkins J.N., McCarty J.C. (2010) A generalized approach and computer tool for quantitative genetics study. *Proceedings of Applied Statistics in Agriculture*:85-106.
- Wu J., Jenkins J.N., McCarty J.C., Glover K. (2012) Detecting epistatic effects associated with cotton traits by a modified MDR approach. *Euphytica*:In press. DOI: 10.1007/s10681-012-0770-5.
- Wu J.X., Jenkins J.N., McCarty J.C., Wu D.F. (2006) Variance component estimation using the additive, dominance, and additive \times additive model when genotypes vary across environments. *Crop Science* 46:174-179. DOI: DOI 10.2135/cropsci2005.04-0025.

- Zhu J. (1989) Estimation of genetic variance components in the general mixed model, North Carolina State.
- Zhu J. (1993) Methods of predicting genotype value and heterosis for offspring of hybrids. *Journal of Biomathematics* 8:32-40.
- Zhu J. (1998) *Genetic Models and Analytical Methods* China Agricultural Press, Beijing, China.
- Zhu J., Xu F., Lai M.G. (1993) Analysis methods for unbalanced data from regional trials of crop variety, analysis for single trait. *Journal of Zhejiang Agricultural University* 19:7-13.

Table 1: Mean values and ranks and their 95% confidence intervals for eight cotton cultivars.

Cultivar	Mean					Rank			
	Orig [†]	Boot [‡]	LL [‡]	UL [‡]		Orig [†]	Boot [‡]	LL [‡]	UL [‡]
Ji 84-25	50.34	50.39	46.91	54.13		1	1.00	1	1
Yun 1929	58.54	58.43	54.61	62.05		3	3.22	2	5
Ping 28	64.89	64.87	61.45	68.48		6	6.55	5	8
Shi 3409	60.92	60.90	56.72	64.67		5	4.53	2	6
Jizhi 17	59.54	59.49	55.03	63.43		4	3.80	2	6
Zhong 206	67.00	66.88	62.67	70.85		8	7.43	6	8
Lu 155	57.47	57.43	54.02	60.60		2	2.69	2	5
Zhong 12	65.39	65.34	61.05	69.31		7	6.79	5	8

[†]: Genotypic mean/rank calculated from the original data and [‡]: bootstrapped mean/rank from 1000 resamples.

[‡]: LL=low limit and UL=upper limit

Table 2: Estimated variance components and proportional variance components for cotton yield.

Component	Variance		Proportional variance	
	Estimate	P-value	Estimate	P-value
Year	4.39	0.04	0.02	0.04
Location	7.02	0.22	0.03	0.22
Genotype	27.89	0.00	0.13	0.00
Year×location	125.97	0.00	0.60	0.00
Year×genotype	0.85	0.54	0.00	0.55
Location×genotype	7.95	0.10	0.04	0.10
Residual	36.07	0.00	0.17	0.000

Table 3: Phenotypic values and genotypic effects for eight cotton cultivars over 24 locations and two years.

Cultivar	Phenotypic value	Genotypic effect	LL [‡]	UL [‡]	
Ji 84-25	50.59	-9.94	-11.35	-8.53	D
Yun 1929	58.64	-1.89	-3.06	-0.72	C
Ping 28	64.83	4.30	2.76	5.84	A
Shi 3409	60.84	0.31	-0.67	1.29	B
Jizhi 17	59.54	-0.99	-2.24	0.27	BC
Zhong 206	67.03	6.50	5.27	7.74	A
Lu 155	57.59	-2.94	-3.74	-2.15	BC
Zhong 12	65.18	4.65	3.44	5.85	A

[‡] LL=low limit and UL=upper limit for genotypic effects

Table 4: Genotype-by-location interaction effects between eight cotton varieties and 24 locations

Location	Ji 84-25	Yun 1929	Ping 28	Shi 3409	Jizhi 17	Zhong 206	Lu 155	Zhong 12
1	4.89**	0.41	-8.39**	0.15	3.75	0.63	-0.92	2.75*
2	-1.39	1.07	3.22*	-2.70	-0.13	3.96*	-0.15	2.37*
3	5.02	-0.79	-1.03	-0.84	-0.48	2.46	0.89	1.15
4	2.43*	-4.56*	0.10	-2.60*	-1.53	0.39	0.19	0.25
5	2.65	1.00	-5.89	0.30	2.65	3.64*	-3.77	1.90
6	-6.96	3.34*	-0.96	-2.60	3.20*	0.81	-0.60	2.24
7	-1.53	-1.69	5.02**	4.74*	-0.38	0.82	-0.90	-2.88*
8	-3.68*	4.40	-2.21	2.66*	0.29	4.09*	-3.87**	3.15
9	-3.95*	0.45	6.61**	1.96	0.94	-1.36	-2.46	1.97
10	1.15	-0.44	-0.59	-0.30	0.24	0.59	0.32	-0.63
11	-0.58	-0.04	-0.92	0.52	2.05	-3.00	-1.03	-0.95
12	-2.95	2.31	2.63	-3.83	-4.98	-1.66	9.32**	-3.11
13	1.44	-0.21	-0.82	0.89	0.93	-1.40	-1.30	2.29*
14	-0.03	-0.91	0.36	1.62	0.93	-1.56	1.37	0.57
15	-2.65	3.21	3.04	-1.15	-0.83	-1.89	-1.23	-3.74*
16	-0.58	0.46	1.64	1.07	0.09	-2.25	-3.82*	3.67*
17	1.61	-0.55	-8.04**	2.70	-0.47	-1.36	1.31	-0.44
18	-0.62	1.01	3.80	-1.16	0.73	2.03	0.02	-2.32
19	1.45	-4.91*	3.97	-4.13*	-5.36**	4.21*	0.24	-3.71*
20	0.74	-2.05	5.76**	-0.50	-2.09*	1.91	-1.62	-3.93*
21	0.17	-2.80	1.14	-3.11*	-1.49	-1.73	-1.19	1.98*
22	-1.19	0.55	-1.85	0.94	0.97	-1.56	3.51	0.44
23	-1.48	-0.22	-2.63*	2.96*	-0.78	-4.74*	3.46*	-1.73
24	0.30	-0.18	-1.53	2.69*	1.21	0.68	0.61	1.38
Variance [†]	<u>7.47</u>	<u>4.87</u>	<u>15.38</u>	<u>5.58</u>	<u>4.59</u>	<u>5.90</u>	<u>7.65</u>	<u>5.77</u>

[†]: Variance of genotype-by-location interaction effects for each genotype; * and ** are significant from zero at probability levels of 0.05 and 0.01, respectively.

Table 5: Estimated slopes and coefficients of determinations (R^2) and their 95% confidence intervals (LL: low limit and UL: upper limit).

	Slope					R^2			
	Orig [†]	Boot [‡]	LL [‡]	UL [‡]		Orig [†]	Boot [‡]	LL [‡]	UL [‡]
Ji 84-25	0.91	0.91	0.76	1.07		0.74	0.74	0.59	0.86
Yun 1929	1.01	1.01	0.86	1.13		0.83	0.83	0.73	0.91
Ping 28	0.72	0.72	0.50	0.96		0.49	0.49	0.29	0.67
Shi 3409	1.13	1.13	0.99	1.26		0.89	0.89	0.81	0.94
Jizhi 17	1.12	1.12	0.99	1.26		0.88	0.89	0.82	0.93
Zhong 206	1.14	1.14	0.99	1.28		0.88	0.88	0.81	0.93
Lu 155	0.84	0.84	0.67	0.98		0.72	0.72	0.53	0.85
Zhong 12	1.14	1.14	0.99	1.27		0.88	0.88	0.81	0.93

[†]: Genotypic mean/rank calculated from the original data and [‡]: bootstrapped mean/rank from 1000 resamples.

[‡]: LL=low limit and UL=upper limit