

Assessing the Usefulness of GGE Biplot as a Statistical Tool for Plant Breeders and Agronomists

R.O. AKINWALE^{1*}, M.A.B. FAKOREDE¹, B. BADU-APRAKU² and A. OLUWARANTI¹

¹Department of Crop Production and Protection, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria

²International Institute of Tropical Agriculture, PMB 5320, Ibadan, Oyo State, Nigeria

(Received 9 July 2013; Accepted 20 December 2013)

Genotype main effect plus genotype-by-environment interaction (GGE) biplot produces a graphical display of results that facilitates a better understanding of complex genotype-by-environment interaction in multi-environment trials of breeding and agronomic experiments. However, the full potential and weaknesses of this powerful tool are not fully understood by breeders, agronomists, entomologists and pathologists. The objective of this paper was to review the usefulness of this statistical tool and enumerate some of its weaknesses. Its main application has so far been in the analysis of multi-environment data. It has been used to analyze the performance of crop cultivars under multiple stress environments, from which ideal cultivars, mega-environments, and core testing sites were identified. More recently, GGE biplot has been employed in genetic analysis of diallel data to estimate the combining abilities and identify heterotic groups among inbred parents. Genotype-by-trait biplot has also been utilized in trait profile analysis, and in identification of traits that are reliable for indirect selection of a target primary trait. Two major shortcomings of this tool are (i) failure to identify more than two distinct, contrasting groups in diallel studies and (ii) lack of statistical tests for most of its graphical displays. Other aspects of GGE biplot that need further study and development are (i) estimation of genetic variances, covariances, and heritability, including the analysis of data generated from North Carolina Designs I, II, and III as well as other genetic designs, considering their importance in plant breeding programs; (ii) analysis of Quantitative Trait Loci (QTL) data for proper understanding of the genetic constitution of each individual plant or line; and (iii) analysis of Genotype-by-pathogen or insect strain interaction data. Nevertheless, GGE biplot has helped greatly in the accurate analysis and interpretation of data from breeding and agronomic field evaluation experiments.

Keywords: GGE biplot, ideal cultivar, multiple traits, representativeness

Introduction

Statistical analysis is an essential procedure for drawing valid conclusions from the results of biological experiments. Software for statistical methods are now readily available for analyzing data collected from agronomic and plant breeding experiments. Among the very useful statistical methods is biplot analysis, an important multivariate tool that graph-

* Corresponding author; E-mail: akinrichie2002@yahoo.com

ically displays results and therefore, facilitates easy understanding and presentation. A biplot is a scatter plot of values and it graphically displays a rank-2 matrix by rows (entries) and columns (testers). Data read in the rows are referred to as entries while those in the columns are testers. "Entries" as used in this case are not necessarily crop varieties and "testers" are not necessarily the same as used by breeders in combining ability studies. Rather, they refer to any two factors considered in a factorial experiment. The use of biplot was originally proposed by Gabriel (1971) as a graphical tool to present results from principal component analysis (PCA). The PCA is a mathematical process that decomposes a covariance matrix into eigenvalues and column eigenvectors. Gabriel's proposal provided a theoretical basis for a software package named Genotype Main Effect plus Genotype-by-Environment Interaction (GGE) Biplot (Yan 2001). Although the software was developed with the breeders in mind (Yan and Kang 2003), it is capable of analyzing any set of data that may be cast in a 2-way table. The GGE biplot uses the first two principal axes (PC1 and PC2) to analyze the interrelationships among rows (entries) and columns (testers) based on environment-centred PCA. An environment-centred PCA uses within-tester standard deviation-standardized mean values for the PCA (Yan et al. 2000).

Although, the total variation of a phenotype/trait could be partitioned into the environment (E) factor, genotype (G) effect, and genotype-by-environment interaction (GEI) effect, GGE biplot models consider only G and GEI as relevant to cultivar evaluation (Gauch and Zobel 1996) and the two factors are considered simultaneously, removing the environment effect. This is so because the concept of crop performance and stability, as considered by crop scientists when selecting superior genotypes, are centred on G and GEI effects only, irrespective of how large the E effect is. Depending on the type of data and the objective of the analysis, GGE biplot uses four different scaling (singular value partitioning) methods, including genotype-focused scaling, environment-focused scaling, symmetric scaling, and equal-space scaling method. A scaling method describes the type of standardization used for the mean values before the analysis is carried out (Yan 2002). Two of the methods, genotype-focused and environment-focused scaling are most frequently used. In genotype-focused scaling, the genotype scores are in principal coordinates, such that the vector lengths represent the original unit of the genotype data while the environment scores are in standard coordinates (DeLacy et al. 1996). In environment-focused scaling, on the other hand, the environment data assumes the principal coordinates, and the vector lengths represent the original unit of the environment scores while the genotype scores are standardized. Basically, there are three views/displays of the GGE biplot. The first is the polygon view which reveals the winning entry (vertex genotype) in different testers (which may be environments, traits or strains). In this biplot view, the entry at the vertex of the polygon in each sector is considered the best for the testers located within the sector. The second view is the average tester coordination (also known as 'mean-versus-stability') view, which ranks entries according to their performance and stability (consistency) across testers. Entries are ranked along the average-tester axis (ATC abscissa), which has an arrow pointing to a greater value based on their mean performance across all testers. A vertical double-arrowed line (also called average-tester coordinate [ATC] y axis) divides the abscissa into two halves in order to separate entries with be-

low-average means from those with above-average means. The average performance of an entry is approximated by the projection of its entry's marker on the abscissa while stability is measured by its projection onto the double-arrow line. The greater the absolute length of the projection of an entry onto the double-arrow line, the less stable it is. The third view of a GGE biplot is the vector view that displays the discriminating ability and representativeness of the tester. The discriminating power of a tester, which refers to its ability to differentiate among entries being tested, is displayed by the length of the vector, which approximates the standard deviation of the tester. Testers with shorter vectors provide little or no information about the entries evaluated compared to those with longer vectors. Furthermore, the representativeness of a tester refers to the ability of the tester to represent other testers of its group. The small circle located on the abscissa represents the average-tester axis (AEA) (Yan and Tinker 2005). The cosine of the angle between two vectors with reference to AEA approximates the correlation coefficient between the testers and an ideal tester. Testers that have small angles with AEA are more representative of other testers of the same group than those that have large angles to AEA.

A Windows-driven software, "GGE biplot" that fully automates biplot analysis was developed by Yan (2001). The program is available at www.ggebplot.com (verified 1 December 2012). The basic model for a GGE biplot is based on singular value decomposition (SVD) of the first two principal components; that is:

$$Y_{ij} - Y_j = \lambda_1 \xi_{i1} \eta_{j1} + \lambda_2 \xi_{i2} \eta_{j2} + \varepsilon_{ij}$$

where

Y_{ij} is the measured mean of genotype i in environment j

Y_j is the mean across all genotypes in environment j

λ_1 and λ_2 are the singular values for PC1 and PC2

ξ_{i1} and ξ_{i2} are the PC1 and PC2 scores, for genotype i

η_{j1} and η_{j2} are the PC1 and PC2 scores, for environment j

ε_{ij} is the residual of the model associated with the genotype i in environment j .

More recently, GGE biplot has gained a wide application in many major areas of agronomy, plant breeding and genetic analysis such as genotype selection, genotype-by-environment interaction analysis, test location evaluation, identification of mega-environments, diallel analysis and identification of parental inbreds for hybrid development, evaluation of cultivars for release for commercial production, and a host of others. Most of the uses of GGE biplot has been limited to a relatively few areas of the world, primarily the USA, Canada and a few other developed and developing countries especially by scientists in international research institutes such as International Institute of Tropical Agriculture (IITA) and International Maize and Wheat Improvement Centre (CIMMYT). One reason for the restricted use of this tool could be lack of understanding of its potential capability on the part of breeders, agronomists, entomologists and pathologists. In addition, the major flaws as well as possible useful areas of application of the package that are yet to be developed and tapped are not easily realized by users. For example, most analysis from GGE biplot usually lack discrete statistical test of significance since it only graphically displays results. For this reason, Yan (2002) suggested that the GGE biplot is better used as a hy-

pothesis-generator rather than as a decision-making tool while Yan and Ma (2006) discussed the theoretical fundamental for appropriately using GGE biplot. There is a need to more critically evaluate the theoretical basis and applied results of the software as well as critique the output from the analysis and its interpretation in relation to other statistical analysis tools available to the community of agricultural scientists in the world. An earlier attempt at this was made by Yang et al. (2009) who proposed the use of bootstrap procedure for the test of significance for the genotype by environment biplot pattern displayed by GGE biplot. In contrast, a response from Yan et al. (2010) pointed out the reasons the cumbersome bootstrap procedure is grossly deficient for the purpose of testing significance of biplot displays. The objectives of this paper are to review areas of plant breeding and agronomy where GGE biplot has been used extensively and to highlight the strengths and limitations of this statistical tool. Areas considered herein include multi-environment analysis, trait profile and trait association analysis, and genetic analysis.

Multi-environment Analysis

GGE biplot was originally designed to analyze complex GEI. It complements the results of the ANOVA in that, after the ANOVA had shown significant mean squares for GEI, GGE biplot graphically analyzes the nature of the interactions. The investigation into GEI becomes more interesting with the introduction of the concept of crossover interaction and GGE biplot has increasingly been used in GEI data analysis in agriculture (Yan et al. 2000; Crossa et al. 2002; Badu-Apraku et al. 2010). For this type of analysis, Gauch (1992) reported that GGE biplot is superior to other multivariate statistical tools especially the Additive Main Effect and Multiplicative Interaction (AMMI) in that GGE biplot has many visual interpretations and it particularly allows visualization of crossover GEI. Moreover, Yan et al. (2000) further buttressed this by explaining that GGE biplot is more logical and biological for practice than AMMI in terms of explanation of PC1 score, which represents genotypic effect rather than additive main effect. Conventionally, the appropriate GGE biplot model for the analysis of multi-environment yield data is model 3 where data are not transformed (Transform=0) not standardized (Scale = 0) environment-centred (“Centering_2”) and are based on genotype-focused singular value partitioning (“SVP_2”). The polygon view (which-won-where) of the GGE biplot graphically displays the interaction between genotypes and environments, showing which genotype is best for which environment. The vector view of the GGE biplot analyzes the interrelationships among environments in relation to the average (target) environment and this helps to identify mega-environments, evaluate the discriminating ability and representativeness of the environment with a view to identifying redundant locations that can be culled without sacrificing important information about the genotypes. The average tester coordination (‘mean-vs-stability’) view of the biplot displays the performance as well as the stability of the genotypes across the test location (Yan et al. 2000). For example, when used in the analysis of the performance of tropical early maize cultivars in multiple stress environments, GGE biplot identified four cultivars as ideal for production across drought, *Striga*, and low N environments (Badu-Apraku et al. 2010). Based on GGE biplot analysis, Badu-Apraku et al.

(2011d) identified four mega-environments for evaluating early maize cultivars in West Africa, with Kita and Zaria as the ideal test locations. Yan et al. (2000) classified the winter wheat environment of Ontario Canada into two; western and southern Ontario, a major winter mega-environment and eastern Ontario, a minor mega-environment. Samonte et al. (2005) identified mega-environments, superior rice cultivars and ideal rice cultivars in different mega-environments using site regression GGE biplot. More recently, Yan and Holland (2010) demonstrated the appropriateness of the use of heritability-adjusted (Scale = 2) GGE (HA-GGE) biplot in simultaneous evaluation of genotypes and test locations. Based on HA-GGE biplot, Flores et al. (2013) identified three mega-environments across European climates, with two core testing sites and outstanding Faba bean cultivars.

Furthermore, GGE biplot has been used in assessing the repeatability of test environments using model 2 of the GGE biplot (Scale = 2) and four types of test locations have been identified as follows; type I test locations are ideal or core testing site which is very important in a breeding programme for selection exercise in the early breeding stage when selection is done only at a single location. Type II test locations are also useful, particularly at the multilocation test stage which may not be appropriate for selecting superior genotypes when used alone but are very useful for assessing the stability of genotypes. Type III test locations are of limited use, but if used, can only be used to remove extremely defective genotypes. Type IV are least useful for genotype evaluation because they are not discriminating and not repeatable (Yan et al. 2011).

Trait Profile and Trait Association Analysis

As a multivariate tool for trait analysis, genotype-by-trait (GT) analysis component of GGE biplot integrates some features of AMMI, joint regression, and genetic correlation in its analysis of data, which makes it superior over other multivariate methods. In addition, other multivariate analysis methods such as correlation, multiple regression, path analysis (Wright 1921), factor analysis and cluster models have the common disadvantage of not being capable of identifying genotypes with specific desirable traits that could be used in a selection program. As noted by Fakorede (1979), correlation analysis measures the mutual association between a pair of variables independently of all other variables across all genotypes. Regression analysis, including stepwise multiple regression and path analysis, which is a special case of partial multiple regression analysis, as well as other multivariate techniques examine the association among traits measured on a set of genotypes without identifying individual genotypes superior for specific traits (Fakorede 1979).

The GT biplot was proposed by Yan and Kang (2003) as a tool for evaluating cultivars based on multiple traits and for identifying those that are superior in certain traits and hence could be candidates for use as parents in a breeding programme or directly released for commercial production. The GT biplot analysis allows visualization of the genetic correlation among traits (Yan and Rajcan 2002) and also evaluation of genotypes on the basis of multiple traits (Ober et al. 2005). The use of *Auto Find QTL* option, one of the component functions of the GGE biplot software visually aids in identifying traits that are reliable for indirect selection for a target traits based on their genetic correlations at a level of

probability and also helps to detect less important (redundant) traits. Of the four scaling methods that GGE biplot uses for analysis, the most appropriate for visualizing interrelationships among genotypes as well as traits is standard deviation-standardized ('Scale = 1') and trait-centred ('Centring = 2') model 2. Standardized trait values must be used because the traits are measured in different units. The GGE biplot statistical model equation is modified for this purpose as follows:

$$(\hat{Y}_{ij} - \mu - \beta_j) / d_j = \lambda_1 g_{i1} e_{1j} + \lambda_2 g_{i2} e_{2j} + \varepsilon_{ij}$$

where Y_{ij} is the genetic value of the combination between genotype i and trait j ; μ is the mean of all combinations involving trait j ; β_j is the main effect of trait j ; λ_1 and λ_2 are the singular values for PC1 and PC2; g_{i1} and g_{i2} are the PC1 and PC2 eigenvectors, respectively, for genotype i ; e_{1j} and e_{2j} are the PC1 and PC2 eigenvectors, respectively, for trait j ; d_j is the phenotypic standard deviation; and ε_{ij} is the residual of the model associated with the combination of genotype i and trait j (Yan and Tinker 2006).

The polygon view of the GT biplot helps to identify genotypes that are outstanding for certain traits, which could be used to introgress such trait into population. The mean vs stability view ranks the genotypes based on multiple traits and identify genotypes which combines various good qualities. The vector view of the GT displays the relationship among traits, with a view to identify redundant and less useful traits in selecting genotype for a target trait. Furthermore, Yan and Frégeau-Reid (2008) proposed the use of a comprehensive multitrait selection procedure that combines independent selection, independent culling, and index selection so that all the aspects in breeding line selection are taken into consideration. Similar to the use of base index, the multitrait selection tool of the GGE biplot facilitates assigning appropriate weights to selected important traits and automatically selecting superior cultivars based on the traits. Based on GT biplot, Badu-Apraku and Akinwale (2011) identified four cultivars as ideal genotypes under *Striga* infestation based on multiple traits and also reported that ears per plant (EPP), *Striga* damage at 8 and 10 weeks after planting, and ear aspect (EASP) were the most reliable traits for selecting *Striga*-resistant genotypes. Similar results were obtained when genetic and phenotypic correlations were used (Badu-Apraku 2010). Similarly, Badu-Apraku et al. (2011c) reported that anthesis-silking interval, EPP, ear aspect, and plant aspect were the most reliable secondary traits that could be used for simultaneous selection of early maize cultivars for drought and low-N tolerance. Earlier studies in some other parts of the world (Banziger et al. 1999), similarly identified EPP and anthesis-silking interval as reliable indirect selection criteria for drought and low-N tolerance using other multivariate approaches such as correlation and principal component analyses. Dolatabad et al. (2010) also used GT biplot to examine trait relationships among maize hybrids with a view to identifying which hybrid was best for some particular trait.

The major weakness of GGE biplot in analysis of traits is its insensitivity to multicollinearity and spurious correlations when used to identify reliable traits that could be used for indirect selection for a target trait. Because GGE biplot largely identifies reliable traits based on genetic correlations among traits, the traits that are highly correlated are

displayed as reliable, without considering which of the traits contribute indirectly to the target trait through other traits. This weakness came out clearly when the results of analysis of traits interrelations using GT biplot and sequential path analysis were juxtaposed (Badu-Apraku et al. 2012). In the study, it was observed that GT biplot identified traits that were most reliable for selecting for a target trait irrespective of the high correlations between the traits. On the other hand, if any two traits show high correlation, path analysis considers whether one of them directly contributes to the target trait while the second trait contributes indirectly to the target trait through the first trait. It was concluded from the study that path analysis showed higher sensitivity to multicollinearity and spurious correlations than the biplot analysis.

Genetic Analysis

The phenotypic expression of a plant trait is a product of the genotypic constitution as it is modulated by environmental factors prevalent during the growing season. Breeders are more interested in estimating genetic parameters because they increase the efficiency of prediction of gains from selection to improve crop cultivars. GGE biplot has been extensively applied in the analysis of combining ability and identification of heterotic patterns using diallel data.

Diallel analysis

Of recent, GGE biplot has been applied in genetic analysis of diallel crosses to estimate the combining abilities of inbred parents and classify them into distinct heterotic groups. The general combining ability (GCA) of the inbred parents and specific combining ability (SCA) of the F_1 hybrids are visually displayed by the projections of the entries and testers. The identification of the best testers and classification into heterotic groups obtainable from GGE biplot analysis of diallel data give it an added advantage over the conventional Griffing's methods (Griffing 1956). In biplot analysis of diallel data, the appropriate scaling to use is 'Scale = 0' (i.e., data not standardized). This enables the researcher to estimate the discriminating ability (i.e., the efficiency) of the testers. The polygon view of the biplot displays the best mating partners for a vertex entry, indicating crosses that are capable of producing hybrids of good vigour (Fig. 1A). The vertex inbreds (testers) are the best mating partners for all inbreds (entries) found in their respective sectors provided the same inbred is not found as a tester in the same sector. For instance in Figure 1A, inbred TZEI 23 is the best mating partner for testers TZEI 13 and TZEI 16 and since tester TZEI 23 is not located within the sector, it implies that crosses TZEI 23 \times TZEI 13 and TZEI 23 \times TZEI 16 are expected to exhibit high heterosis. In the average tester coordination view of the GGE biplot (Fig. 1B), the projections of the entries vectors onto the average tester axis (ATC) abscissa approximate their GCA effects while the projections of the entries vectors onto the ATC ordinate approximate their SCA effects (Yan and Kang 2003). The longer the projections, the higher the SCA effects. However, the GCA effects of a set of inbred parents are considered significantly different from one another if the distance from the biplot origin and average tester (small circle on the ATC abscissa) is wide enough such

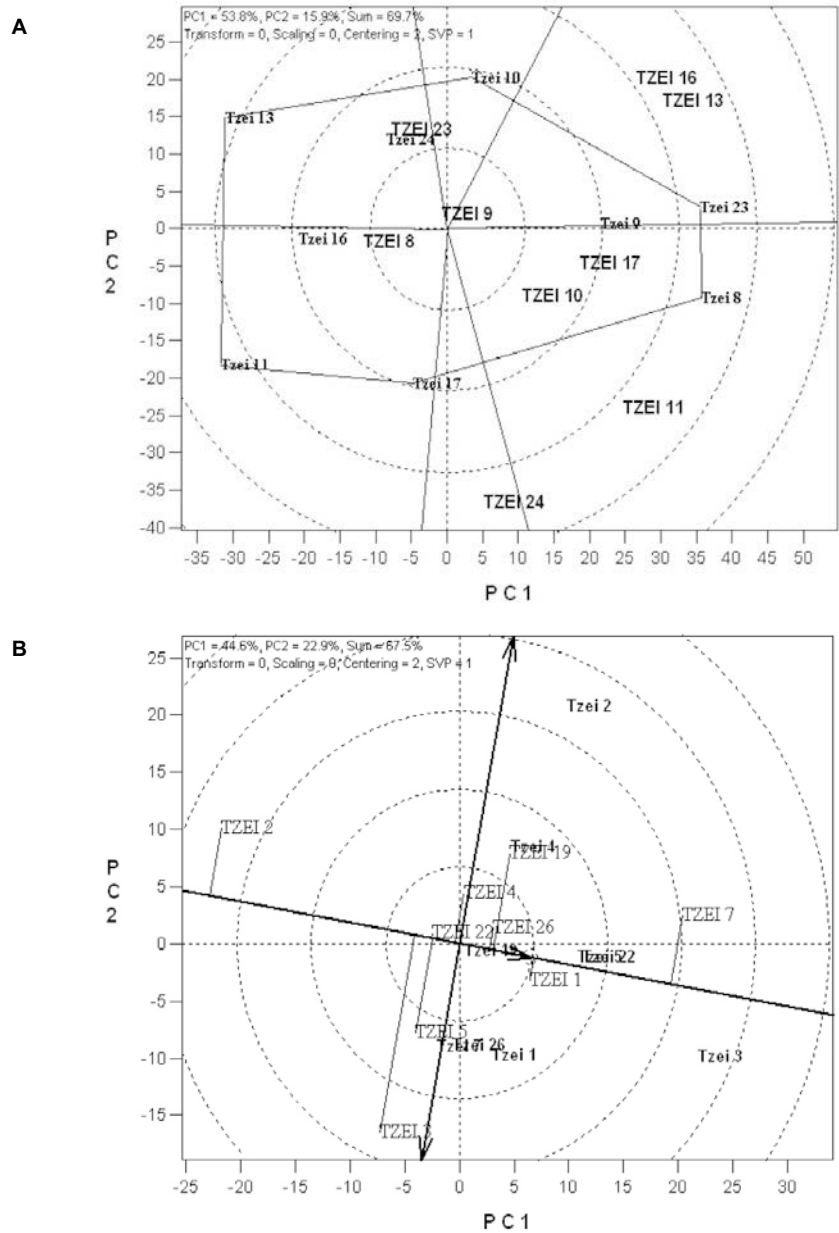


Figure 1. GGE biplot analysis of diallel data of nine early-maturing inbred lines, (A) Polygon view showing best mating partners and (B) Mean-versus-stability view showing the combining abilities of inbred lines under *Striga*-infested environments. Inbred lines are labeled with upper-case letters when viewed as testers and with title-case letters when viewed as entries. Exact positions of the entries and testers are at the beginning of the labels

that all testers are on one side of the ATC ordinate. Premised on significant GCA effects of the inbred parents, distinct heterotic groups can be visually identified using the SCA of the entries (Badu-Apraku et al. 2011a, b). Otherwise, no meaningful grouping can be made.

Line × Tester Analysis

GGE biplot should be a powerful statistical tool for analyzing line × tester data as it has been for diallel data. However, no attempt has been reported in the literature on its use for analyzing the efficiency of testers in classifying genotypes and the resulting heterotic groups identified.

We used the hypothetical values on Line × Tester data contained in Chapter 10 of Singh and Chaudhary (1977) to demonstrate the effectiveness of biplot analysis in this regard. The biplots were constructed using model 3 (Transform = 0; Scaling = 0; Centring = 2). To-

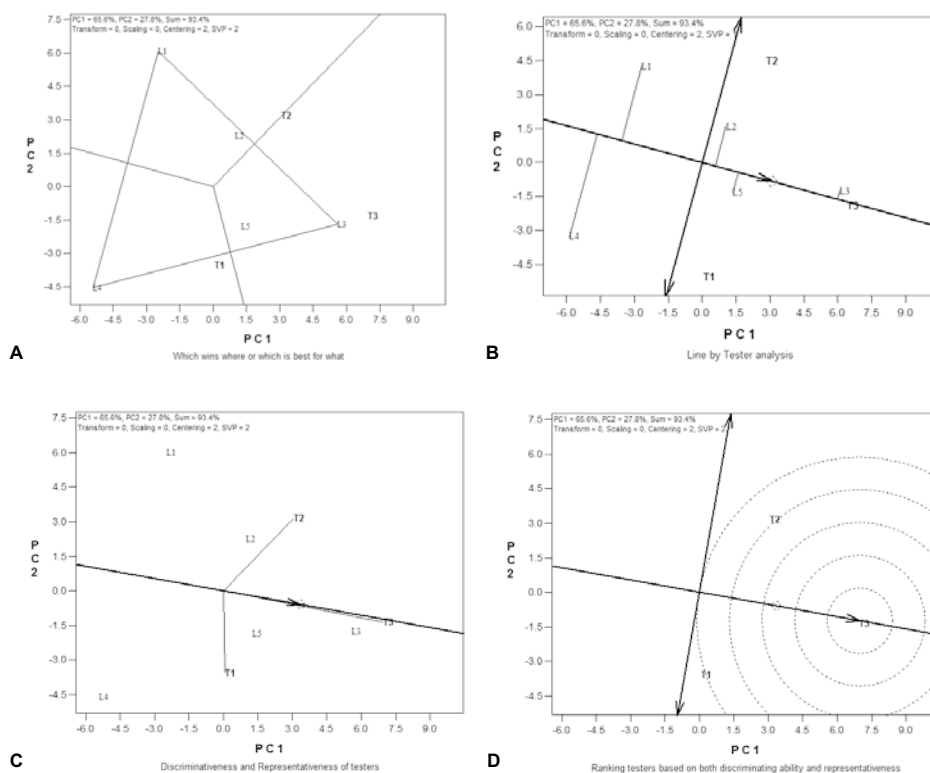


Figure 2. GGE biplot analysis of a line-by tester data (A) Polygon view showing which line best combined with which tester. ‘L’ stands for line and ‘T’ stands for tester (B) Mean-versus-stability view showing the combining abilities of the five lines and three testers (C) Average Tester Coordination view showing the discriminating ability and representativeness of the three testers (D) vector view showing the ranking of the testers based on their discriminating ability and representativeness with the view of identifying the ideal tester for the lines

gether, PC1 and PC2 explained 93.4% of the total variation in the set of data. Figure 2A displays the polygon view, showing which tester combines well with which line. From the biplot, T1 best combined with L3 and L5, and was the best mating partner to L4. Lines L1 and L2 had no tester in their sector, suggesting that they did not combine outstandingly with any of the testers. The vector view of the biplot (Fig. 2B) revealed the combining abilities of the lines. The length of the average tester vector [i.e., distance between the biplot origin and the average tester (small circle on the abscissa)] is a measure of the relative importance of the GCA versus line \times tester interaction. In addition, the fact that the three testers were positioned on the same side of the average tester (ATC) ordination axis (somewhat vertical double-arrow line) signifies that the GCA is significantly large enough and that the testers are capable of classifying the lines into distinct heterotic groups. From the biplot, L3 had the highest GCA, followed by L5 while L4 had the least (Fig. 2B). Furthermore, the projections of the lines onto the ATC ordinate (doubled-arrowed vertical line) approximate their SCA effects, an indication of the tendency of the lines to produce superior hybrids with specific testers. Lines L1 and L4 had the highest SCA. These results agreed with the traditional line \times tester analysis results in Singh and Chaudhary (1977). In addition to the results of the traditional method of analyzing line \times tester data, biplot provides information on the effectiveness of the testers with the view of identifying the ideal (best) tester. An ideal tester should have the largest vector of all testers (i.e., be most discriminating) and have zero projection onto the ATC ordinate (i.e., be most representative of testers). Thus the closer a tester's marker was to the ideal tester, the better it was. Figure 2C displays the performance of the testers in relation to their crosses with the lines. The biplot showed that T3 had the highest discriminating ability and it was the most representative and, therefore, the only tester in the innermost concentric circle in Figure 2D. Thus, T3 was the ideal (best) tester among the three.

Weaknesses of GGE Biplot Analysis

The weaknesses of GGE biplot in analyzing genetic data may be revealed by the proportion of the total variation explained by the PC1 and PC2 of the biplot view. Depending on the complexity in the genetics of the trait under consideration plus the confounding effect of the environments, the biplot view may account for small variation, which makes the results of the analysis less worthy. Badu-Apraku et al. (2011a, b) reported that the biplot generally accounted for small variation under stress environments, especially where the genotypes used were not specifically bred for tolerance to such stress. However, no studies have been carried out to specify when the proportion of variation explained by a biplot becomes too small to make a valid conclusion. However, it is generally assumed that any proportion below 40% is too small.

Another shortcoming is that GGE biplot display can only identify two distinct heterotic groups even when more groups exist. It simply classifies the remaining inbred(s) as the third group if it is only one inbred or as non- or less-responsive inbreds if they are more than one (Yan and Kang 2003). Furthermore, in all diallel studies where GGE biplot was used, classification into heterotic groups was based on the SCA effects only, which is rep-

resented by the projections of the entries vectors onto the ATC ordinate. However, Fan et al. (2009) proposed the use of a heterotic group's general and specific combining ability (HGSCA), a combination of GCA and SCA effects. This proposition was based on the fact that SCA effects were often greatly influenced by the interaction between two inbred lines and between hybrids and environments, which often times lead to assigning the same inbred line into different heterotic groups under different studies. Akinwale (2012) confirmed that HGSCA is more efficient in classifying genotypes into heterotic groups than the use of SCA effects only under both stress and non-stress environments. Therefore, the groups identified by GGE biplot may be as heterotic as when HGSCA is employed. Nevertheless, since both GCA and SCA effects are graphically represented in the GGE biplot, there is possibility of getting a better heterotic grouping using GGE biplot by redefining how heterotic groups are identified.

Up till date, no aspect of GGE biplot can estimate extensively the genetic variances, covariances, and heritability, considering the importance of these genetic parameters in plant breeding programs. In particular, application of GGE biplot to the analysis of data generated from North Carolina Designs I, II, and III as well as other genetic designs, is yet to be conceptualized. Yan and Holland (2010) demonstrated the use of heritability-adjusted GGE biplot in the simultaneous evaluation of test locations and genotypes, where the length of the vector approximates the heritability of a particular trait in the different test locations. The application of this biplot model would have served the purpose of graphically displaying the heritability of traits but no attempt has been made in the literature to explore this possibility of genetic analysis of traits.

Furthermore, Yan and Kang (2003) demonstrated how GGE biplot can be applied to analyze Quantitative Trait Loci (QTL) data in order to (i) generate a linear map among all or a subset of genetic markers; (ii) identify QTL for a trait at a user-specified association level; (iii) identify traits that are influenced by a single locus and (iv) visualize genetic constitution of each individual plant or line relative to the markers among other things. However, evidence of the use of GGE biplot in the analysis of such data is still very scanty in the literature. Yan and Tinker (2005) demonstrated the use of GGE biplot in the analysis of a QTL data (QQE biplot) in which total variation for barley yield was decomposed into QTL main effect and QTL-by environment interaction. The resulting QQE biplot allowed visualization of (i) the magnitude of the effect of a QTL, (ii) the average effect of a QTL and its stability across environments, (iii) the effects of a QTL in individual environments, (iv) the similarity/dissimilarity among QTLs in effect and response to the environments, (v) the similarity/dissimilarity among environments in modulating QTL effects, (vi) any differentiation of mega-environments, and (vii) the combination of QTL alleles for maximum expression of the trait for each environment or mega-environment. Its use has not been widely applied to such analysis in recent literature. This might not be unconnected with the fact that molecular breeders and plant biotechnologist are not aware of the usefulness of this vital tool in analyzing such data. Such information is necessary for genetic study of biotic and abiotic stress. It is believed that the application of GGE biplot to QTL data analysis would facilitate a better understanding of the genetics of polygenic traits, which will in turn help in marker assisted selection.

Another area where the potentials of GGE biplot are still under-explored is in the analysis of Genotype-by-pathogen strain interaction data. Yan and Kang (2003) illustrated how the biplot could be used to determine whether resistance to a pathogen is vertical or horizontal (i.e., whether or not there is race-specific resistance); which genotype is resistant to which strain if there is race-specific resistance; which genotype is the most resistant if there is no race-specific resistance; what relationship exists among genotypes in response to the strain; and what relationship exists among strains with regards to virulence. Yan and Falk (2002) illustrated the usefulness of the biplot technique in the analysis of barley net blotch disease data. However, the application of this technique in studies relating to tropical diseases is very limited in the literature.

In conclusion, among the multivariate analysis tools, GGE biplot has the widest applicability in the analysis of plant breeding data. The interpretation of GGE biplot analysis of genetic data is more extensive with wider applicability than the conventional statistical methods. The results of a GGE biplot can, however, be deficient if the PC1 and PC2 account for only a small proportion of the total GGE. This happens when the genotype main effect is considerably smaller than the $G \times E$ interaction and when the $G \times E$ interaction pattern is complex especially when the genotypes are evaluated under stress environments. Because GGE biplot analysis does not allow a serious statistical testing of hypothesis, the reliability of its results has become questionable in the literature. However, its results have been found to be consistent with that of ANOVA, correlation, regression and multivariate statistical methods.

References

- Akinwale, R.O. 2012. Evaluation of the heterotic groups of *Striga*-resistant early maturing maize (*Zea mays* L.) inbred lines using diallel analysis and molecular markers. Ph.D. Thesis, Department of Crop Production and Protection, Obafemi Awolowo University, Ile-Ife, Nigeria, 125 pp.
- Badu-Apraku, B. 2010. Effects of recurrent selection for grain yield and *Striga* resistance in an extra-early maize population. *Crop Sci.* **50**:1735–1743.
- Badu-Apraku, B., Menkir, A., Ajala, S.O., Akinwale, R.O., Oyekunle, M., Obeng-Antwi, K. 2010. Performance of tropical early maturing maize cultivars in multiple stress environments. *Can. J. Plant Sci.* **90**:831–852.
- Badu-Apraku, B., Akinwale, R.O. 2011. Cultivar evaluation and trait analysis of tropical early maturing maize under *Striga*-infested and *Striga*-free environments. *Field Crops Res.* **121**:186–194.
- Badu-Apraku, B., Lum, A.F., Akinwale, R.O., Oyekunle, M. 2011a. Biplot analysis of diallel crosses of early maturing tropical yellow maize inbreds in stress and nonstress environments. *Crop Sci.* **51**:173–188.
- Badu-Apraku, B., Oyekunle, M., Akinwale, R.O., Lum, A.F. 2011b. Combining ability of early-maturing white maize inbreds under stress and nonstress environments. *Agron. J.* **103**:544–557.
- Badu-Apraku, B., Akinwale, R.O., Ajala, S.O., Menkir, A., Fakorede, M.A.B., Oyekunle, M. 2011c. Relationships among traits of tropical early maize cultivars in contrasting environments. *Agron. J.* **103**:717–729.
- Badu-Apraku, B., Akinwale, R.O., Menkir, A., Coulibaly, N., Onyibe, J.E., Yallou, G.C., Abdullai, M.S., Didjira, A. 2011d. Use of GGE biplot for targeting early maturing maize cultivars to mega-environment in West Africa. *Afr. Crop Sci. J.* **19**:79–96.
- Badu-Apraku, B., Akinwale, R.O., Franco, J., Oyekunle, M. 2012. Assessment of reliability of secondary traits in selecting for improved grain yield in drought and low-nitrogen environments. *Crop Sci.* **52**:2050–2062.
- Banziger, M., Edmeades, G.O., Lafitte, H.R. 1999. Selection for drought tolerance increases maize yields across a range of nitrogen levels. *Crop Sci.* **39**:1035–1040.
- Crossa, J., Cornelius, C., Yan, W. 2002. Biplots of linear-bilinear models for studying crossover genotype \times environment interaction. *Crop Sci.* **42**:136–144.

- DeLacy, I.H., Basford, K.E., Cooper, M., Bull, J.K., McLaren, C.G. 1996. Analysis of multi-environment trials – A historical perspective. In: Cooper, M., Hammer, G.L. (eds), *Plant Adaptation and Crop Improvement*. CAB International, Wallingford, UK, pp. 39–124.
- Dolatabad, S.S., Choukan, R., Hervan, E.M., Dehghani, H. 2010. Multienvironment analysis of traits relation and hybrid comparison of maize based on the genotype by trait biplot. *Am. J. Agric. and Biol. Sci.* **5**:107–113.
- Fakorede, M.A.B. 1979. Interrelationships among grain yield and agronomic traits in a synthetic population of maize. *Maydica* **24**:181–192.
- Fan, X.M., Zhang, Y.M., Yao, W.H., Chen, H.M., Tan, J., Xu, C.X., Han, X.L., Luo, L.M., Kang, M.S. 2009. Classifying maize inbred lines into heterotic groups using a factorial mating design. *Agron. J.* **101**:106–112.
- Flores, F., Hybl, M., Knudsen, J.C., Marget, P., Muel, F., Nadal, S., Narits, L., Raffiot, B., Sass, O., Solis, I., Winkler, J., Stoddard, F.L., Rubiales, D. 2013. Adaptation of spring faba bean types across European climates. *Field Crops Res.* **145**:1–9.
- Gabriel, K.R. 1971. The biplot graphic of matrices with application to principal component analysis. *Biometrics* **58**:453–467.
- Gauch, H.G. 1992. *Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs*. Elsevier, Amsterdam, Netherlands.
- Gauch, H.G., Zobel, R.W. 1996. AMMI analysis of yield trials. In: Kang, M.S., Gauch, H.G. Jr (eds), *Genotype-by-Environment Interaction*. CRC Press, Boca Raton, FL, USA, pp. 85–122.
- Griffing, B. 1956. Concepts of general and specific combining ability in relation to diallel crosses system. *Aust. J. Biological Sci.* **9**:463–493.
- Ober, E.S., Bloa, M.L., Clark, C.J.A., Royal, A., Jaggard, K.W., Pidgeon, J.D. 2005. Evaluation of physiological traits as indirect selection criteria for drought tolerance in sugar beet. *Field Crops Res.* **91**:231–249.
- Samonte, S.O., Wilson, L.T., McClung, A.M., Medley, J.C. 2005. Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analyses. *Crop Sci.* **45**:2414–2424.
- Singh, R.K., Chaudhary, B.D. 1977. *Biometrical Methods in Quantitative Genetic Analysis*. Kalyani Publishers, New Delhi, India, 179 pp.
- Wright, S. 1921. Correlation and causation. *J. Agric. Res.* **20**:557–585.
- Yan, W., Hunt, L.A., Sheng, Q., Szlavnic, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* **40**:597–605.
- 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. 2002. Singular-value partitioning in biplot analysis of multienvironment trial data. *Agron. J.* **94**:990–996.
- Yan, W., Falk, D.E. 2002. Biplot analysis of host-by-pathogen data. *Plant Dis.* **86**:1396–1401.
- Yan, W., Rajcan, I. 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. *Can. J. Plant Sci.* **42**:11–20.
- Yan, W., Kang, M.S. 2003. GGE biplot analysis: A graphical tool for breeders. In: Kang, M.S. (ed.), *Geneticists and Agronomist*. CRC Press, Boca Raton, FL, USA, 271 pp.
- Yan, W., Tinker, N.A. 2005. A biplot approach for investigating QTL-by environment patterns. *Mol. Breeding* **15**:31–43.
- Yan, W., Ma, B.L. 2006. *Model Diagnosis and GGE Biplot Analysis*. Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-food Canada, Ottawa, Canada, 39 pp.
- Yan, W., Tinker, N.A. 2006. An integrated biplot analysis system for displaying, interpreting and exploring genotype \times environment interaction. *Crop Sci.* **45**:1004–1016.
- Yan, W., Fréreau-Reid, J. 2008. Breeding line selection based on multiple traits. *Crop Sci.* **48**:417–423.
- Yan, W., Holland, J.B. 2010. A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica* **171**:355–369.
- Yan, W., Glover, K.D., Kang, M.S. 2010. Letter to Editor. *Crop Sci.* **50**:1121–1123.
- Yan, W., Pageau, D., Fréreau-Reid, J., Durand, J. 2011. Assessing the representativeness and repeatability of test locations for genotype evaluation. *Crop Sci.* **51**:1603–1610.
- Yang, R., Crossa, J., Cornelius, P.L., Burgueno, J. 2009. Biplot analysis of genotype \times environment interaction: Proceed with caution. *Crop Sci.* **49**:1564–1576.