

STUDY ON SEED YIELD STABILITY OF SUNFLOWER INBRED LINES THROUGH GGE BILOT

Pourdad, S.S.* and Moghaddam, M.J.**

*Dryland Agricultural Research Sub-Institute, P. O. Box 67145-1164,
Kermanshah, Iran.*

Received: September 10, 2012

Accepted: June 10, 2013

SUMMARY

The evaluation of GE interaction is an important component of the cultivar selection process in multi-environment trials (METs). GGE biplot analysis is an effective method which is based on the principal component analysis (PCA) in order to fully explore METs. The objectives of the present study were first to analyze the GE interaction in the seed yield of 15 sunflower inbred lines through the application of GGE biplot technique, as well as to identify suitable sunflower inbred lines with both mean performance and high stability. Seed yield of 15 sunflower inbred lines which were derived from 5 basic populations was tested in a randomized complete block design with three replications. These inbred lines were tested from 2008 to 2011 at the Sararood agricultural research station in western Iran, which created 6 environments in total. The results of variance analysis showed the significance of environmental variance compared to the genotype and GE interaction variances. The first two principle components (PC1 and PC2) together explained 68.2% of the total variability, so most of the information could be graphically displayed in the PC1 vs PC2 biplot. Polygon view of GGE biplot revealed that G3 was the inbred line with the highest seed yield in four out of six environments. The average environment coordinate (AEC) biplot showed that G3 with the highest mean yield was a highly stable inbred line as it was positioned close to the AEC abscissa. The second highest yielding and most stable inbred line was G10. The biplot of comparison of the sunflower inbred lines with the ideal genotype revealed that G3 was the closest inbred line to the ideal cultivar. Therefore, this inbred line seems to be widely adapted across several environments. According to the biplot of the ideal environment it was concluded that E1 was the closest to ideal environment and therefore the most desirable of all six environments. So, E1 is the most effective for the selection of superior inbred lines.

Key words: sunflower, graphical analysis, GE interaction

* Corresponding author: Phone: 0831-2333410; e-mail: sspourdada@yahoo.com

** Seyyed Saeid Pourdad, Associate Professor and Mehdi Jamshid Moghaddam, MSc Researcher at Dryland Agricultural Research Sub-Institute

INTRODUCTION

Multi-environment trials (METs) are widely used by plant breeders to evaluate relative performance of genotypes in target environments (Delacy *et al.*, 1996). Recorded yield of each cultivar in each test environment is a mixture of the environmental main effect (E), genotype main effect (G) and genotype $\times$ environment interaction (GE). Yan (2002) declared that typically E explains most of the total yield variations, and G and GE are usually small. However, it is G and GE that are relevant for the cultivar evaluation. Furthermore, G and GE must be considered simultaneously when making cultivar selection. The targeting of cultivars to specific locations is difficult when GE interaction is present, since yield is less predictable and its interpretation cannot be based only on G and E (Ebdon *et al.*, 2002).

Numerous methods have been used trying to understand the causes of GE interaction and they represent three major analytical categories - univariate parametric analysis of the GE matrix, non-parametric methods and the multivariate approach (Flores, 1998). While the univariate analysis (parametric and non-parametric) attempts to define GE interaction by one or two parameters, the objectives of the multivariate approach are to explore the multidirectional aspects of the GE interaction and to attempt to extract the additional information from this component.

Gabriel (1971) developed the biplot that simultaneously displayed the rows and the columns of the matrix. Multiple versions of this graphic multivariate technique have been widely applied and accepted by plant breeders in the analysis of GE tables (Gauch, 2006; Yan *et al.*, 2006). Yan *et al.* (2000) developed the genotype main effect (G) plus genotype $\times$ environment interaction (GGE) biplot methodology used for the graphic analysis of multi-environmental trial data. It is a data visualization tool constructed by plotting the first two principal components (PC1 and PC2) derived from the singular value decomposition of the environment-centered data. Models that decomposed the environment-centered data are commonly referred to as site regression models (Yan *et al.*, 2001).

Model of GGE biplot

We know that observed phenotypic variance (P) consists of variances of the environment (E), genotype (G) and genotype $\times$ environment interaction (GE).

$$P = G + GE + E \text{ or } P - E = G + GE$$

The above formulas are in terms of variance components. When presented as effects, which have the unit of the originally measured values, they become (Yan *et al.*, 2003):

$$y_{ij} = \mu + \alpha_j + \beta_j + \phi_{ij}$$

$$y_{ij} - \mu - \beta_j = \alpha_j + \phi_{ij}$$

Where

y_{ij} = the expected yield of genotype i in environment j ,

μ = the grand mean of all observations,

α_j = the main effect of genotype i,

β_j = the main effect of environment j, and

ϕ_{ij} = the interaction between genotype i and environment j.

Instead of trying to separate G and GE, GGE biplot model keeps G and GE together and partitions this mixture GGE into two multiplicative terms:

$$y_{ij} - \mu - \beta_j = g_{i1}e_{1j} + g_{i2}e_{2j} + \phi_{ij} + \varepsilon_{ij}$$

Where g_{i1} and e_{1j} are called the primary scores for genotype i and environment j, respectively; g_{i2} and e_{2j} the secondary scores for genotype i and environment j, respectively; and ε_{ij} is the residue not explained by the primary and secondary effects. Actually, a GGE biplot is constructed by plotting g_{i1} against g_{i2} and e_{1j} against e_{2j} in a single scatter plot. The most common way to implement the above formula is by subjecting the GGE data to singular value decomposition (SVD) as below:

$$y_{ij} - \mu - \beta_j = \lambda_1 \xi_{i1} \eta_{1j} + \lambda_2 \xi_{i2} \eta_{2j} + \varepsilon_{ij}$$

Where λ_1 and λ_2 are the singular values of first and second largest principal components, PC1 and PC2, respectively; ξ_1 and ξ_2 are the eigenvectors of genotype i for PC1 and PC2, respectively; and η_1 and η_2 are the eigenvectors of environment j for PC1 and PC2, respectively.

The objective of the present study were:

- 1) to identify the suitable sunflower inbred lines with both mean performance and high stability and
- 2) to demonstrate the application of the GGE biplot technique in visual analysis of yield stability in sunflower inbred lines.

MATERIALS AND METHODS

The sunflower genotypes used in this study were 15 inbred lines derived from 5 basic populations in the sunflower breeding program at the Dryland Agricultural Research Institute of Iran (DARI). These inbred lines were tested during four years (2008 to 2011) at the Sararood Agricultural Research Station. In addition, in 2008 and 2011 two more trials, accompanied by irrigation on one occasion, were carried out, so there were 6 environments in total (Table 1). The soil of this station is deep alluvial with sandy loam texture and moderate high water storage capacity. The station is situated in a valley in the Northern Zagross mountain range at the altitude of 1351 meters (34° 20' N, 47° 19' E). The location typically has 83 days below zero each winter season. The soil has a pH ranging from 7.2 to 7.8. Total precipitation amounts from 2008 to 2011 were 159.1, 288.3, 455.7 and 342.5 mm, respectively. The experimental layout was a randomized complete block design with three replications in each environment.

A plot scheme of 4 rows $\times$ 4 m and inter-row spacing of 0.50 m was used in all trials each year. Tillage and other cultivation practices were similar across all envi-

ronments. Winter ploughing was done to a 40 cm soil depth. Fertilizer applications were 30 kg ha⁻¹ of nitrogen and 15-20 kg ha⁻¹ of phosphorus oxide (P₂O₅) before planting. Sowing was done by hand during March. Seed yield evaluation was performed by measuring the seed mass for each elementary plot, which led to a calculation of the seed yield (t ha⁻¹) with a 10% moisture basis.

Table 1: Mean yield and code of sunflower inbred lines and test environments

Mean yield (Kg/ha)	Code	Year	Environment	Mean yield (Kg/ha)	Code	Inbred lines
485.3	E ₁	2008	Non-irrigation	611.1	G ₁	SIL - 276
690.6	E ₂	2008	1-time irrigation	661.3	G ₂	SIL - 221
699.5	E ₃	2009	Non-irrigation	798.5	G ₃	SIL - 237
405.3	E ₄	2010	Non-irrigation	568.8	G ₄	SIL - 292
608.0	E ₅	2011	Non-irrigation	598.4	G ₅	SIL - 198
766.5	E ₆	2011	1-time irrigation	618.1	G ₆	SIL - 238
				594.0	G ₇	SIL - 215
				553.5	G ₈	SIL - 42
				552.6	G ₉	SIL - 82
				647.1	G ₁₀	SIL - 99
				533.9	G ₁₁	SIL - 140
				570.2	G ₁₂	SIL - 203
				544.1	G ₁₃	SIL - 206
				619.3	G ₁₄	SIL - 224
				666.9	G ₁₅	SIL - 231

The GGE biplot methodology, which is composed of two concepts - the biplot concept (Gabriel, 1971) and the GGE concept (Gauch *et al.*, 1996; Yan *et al.*, 2000) - was used to analyze the data visually. All biplots presented in this paper were generated using the software GGE biplot package that runs in Windows environment (Yan, 2001).

RESULTS AND DISCUSSION

Results of analysis of variance (Table 2) showed significant differences between the environments (E) and genotypes (G) at 5% level of probability, as well as the GE interaction at 1% level of probability.

Table 2: Analysis of variance and partitioning of environment, genotype and genotype × environment interaction variances

Percentage of E+ G + GE	MS	df	S.O.V.
---	104987.3 **	89	Treatment
86.4	868019.5 *	5	Environment (E)
8.1	81267.8 *	14	Genotype (G)
5.5	55229.0 **	70	G×E
---	74405.4	269	Total

ns, * and **: Non-significant and significant at the 0.05 and 0.01 probability levels, respectively

Partitioning of variances revealed the significance of the environmental variance, compared to the genotype and GE interaction variances. It shows that in spite of the same locations over the years there were big differences between environments due to different precipitation, temperature and 1-time irrigation in different years. Yan (2002) declared that typically E explains the most (up to 80% or higher) of total yield variation and G and GE are usually smaller. A high environmental variance was reported in soybean (Gauch *et al.*, 1988), cotton (Baxevanos *et al.*, 2008; Kerby *et al.*, 1996; 2001) and safflower (Pourdad *et al.*, 2008; Mohammadi *et al.*, 2008).

The first two principle components (PC1 and PC2), obtained by singular value decomposition, together explained 68.2% of the total variability caused by GE interaction (Figure 1). Therefore, most of the information could be graphically displayed in the PC1 vs PC2 biplot. Yan *et al.* (2005) suggested that the poor explanation of variability by the first two principle components showed the complexity of GE interaction.

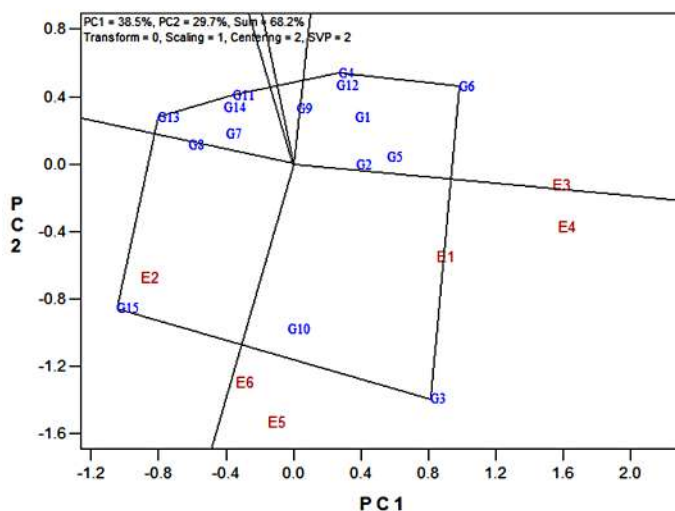


Figure 1: Polygon view of GGE biplot for the 'Which-Won-Where' pattern.

The most responsive inbred lines were G4, G6, G3, G15 and G13 (Figure 1). By connecting the markers of these corner inbred lines a polygon was formed and by drawing perpendiculars to each side of the polygon passing through the origin the environments were divided among several sectors, each with a different corner inbred line (Yan *et al.*, 2000).

The polygon view of the GGE biplot showed that all test environments were divided into 3 groups. Two groups were E3 and E2 environments in G6 and G15 sectors, respectively. Other four environments were in G3 sector. Inbred line G3 had the highest seed yield in E1, E4, E5 and E6 environments. This inbred line as a

vertex cultivar was the one furthest away from the biplot origin, which is an indicator of its responsiveness to environments. G3 had the highest mean yield (798.5 kg/ha) among all inbred lines (Table 1). Inbred lines located near the origin were not responsive to environments and would rank the same in all environments. No environments belonged to the same sectors as G13, as the vertex inbred line. This indicated that this inbred line was the poorest in some or all environments.

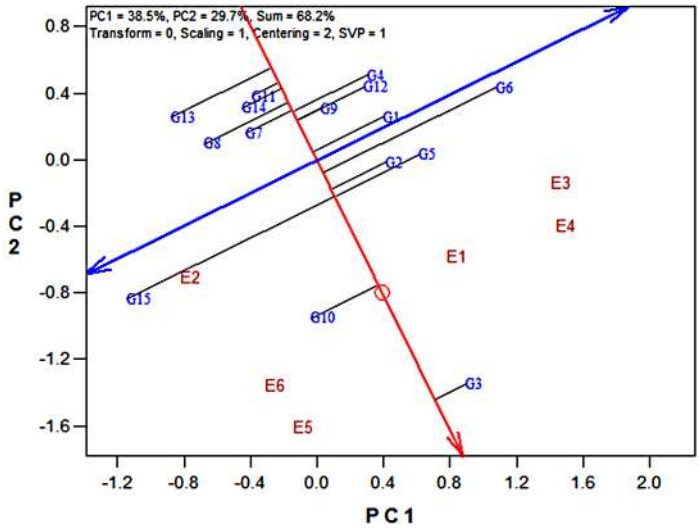


Figure 2: The average environment coordinate biplot to select yield and stability simultaneously in sunflower inbred lines.

To consider the yield and stability simultaneously the average environment coordinate (AEC) biplot was used (Figure 2). It showed the ranking of 15 inbred lines in terms of their mean yield and stability. The average environment, represented by a small circle, is defined by the PC1 and PC2 scores of the environments. The line passing through the biplot origin and average environment is called the average environment axis and serves as the abscissa of the AEC. Projections of inbred lines onto this axis show the approximate mean yield of the inbred lines. The ordinate of the AEC is the line that passes through the origin and is perpendicular to the AEC abscissa. Unlike the AEC abscissa, which has one direction, with the arrow pointing to the greater genotype mean effect, the AEC ordinate is indicated by double arrows, either direction away from the biplot origin indicates a greater GE effect and reduced stability (Yan, 2002). The inbred line G3 was the top yielding genotype, as presented on the front of an average environment towards the pointing arrow of the AEC abscissa. The rank correlation between mean yield of inbred lines and approximate mean yield through projections of inbred line markers onto the average environment axis was 0.786. In addition, the biplot indicated that G3 with the highest mean yield was highly stable, as it is positioned close to the AEC abscissa (Figure 2). The second highest yielding and most stable inbred line was

G10. In contrast, G15 was the most unstable inbred line, as it was away from the AEC abscissa. The inbred line G11 with the lowest yield (533.9 kg/ha) was the most stable inbred line.

An ideal genotype is defined as one that is the highest yielding across all test environments and is absolutely stable in performance, namely one that ranks the highest in all test environments (Yan *et al.*, 2003). Although such an ideal cultivar may not exist in reality, it can be used as a reference for cultivar evaluation. A genotype is more desirable if it is located closer to the ideal cultivar. Thus, using the ideal cultivar as the center, concentric circles were drawn to help visualize the distance between each genotype and the ideal cultivar (Yan, 2002).

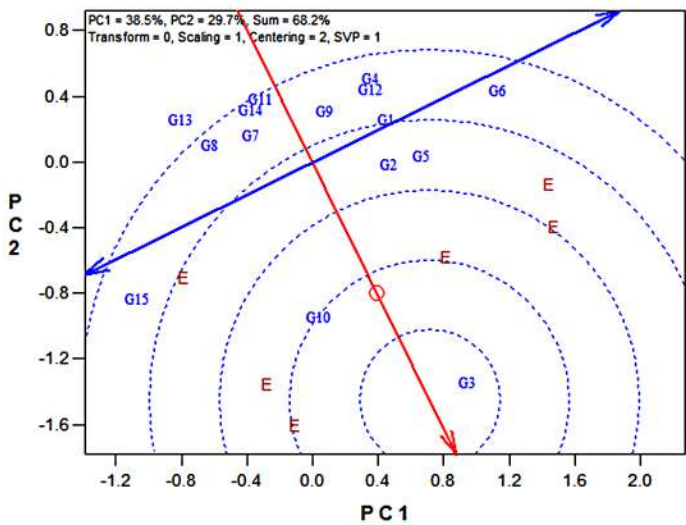


Figure 3: Biplot of comparison of the sunflower inbred lines with the genotype ideal for yield and stability.

Figure 3 showed that G3 was the closest inbred line to the ideal cultivar, therefore seems to be widely adapted across several environments. This inbred line was followed by G10, but G13 was the furthest inbred line from the ideal cultivar. It is interesting to note that the inbred line rankings in Figure 2, based on mean performance, and inbred line rankings in Figure 3, based on both mean performance and stability, are almost identical. This is due to the G being greater than GE (Table 2).

The vector view of the GGE biplot (Figure 4) in which the environments are connected with the biplot origin by lines called vectors. It shows the mutual relations among the test environments. The cosine of the angle between the vectors of two environments is approximate to the correlation coefficient between them (Kroonenberg, 1995; Yan, 2002). Based on the fact that the angles of environment vectors, E3 with E4, E4 with E1 and E5 with E6 were small, it was concluded that they were closely correlated. In contrast, the angles between E3 and E5; E3 and E6; E4 and

E5; E4 and E6 and E1 and E2 were about 90°, so there were no correlations between them, in other words, these environments were different. There were negative correlations among E3 and E2 and E4 and E2 due to the angle bigger than 90° between them. Another interesting observation from the vector view of the biplot is that the length of the environment vectors is approximate to the standard deviation within each environment, which is a measure of their discriminating ability (Yan *et al.*, 2003). However, the test environment lacked the discriminating ability so it provided no information about the genotypes and therefore, the test environment was useless.

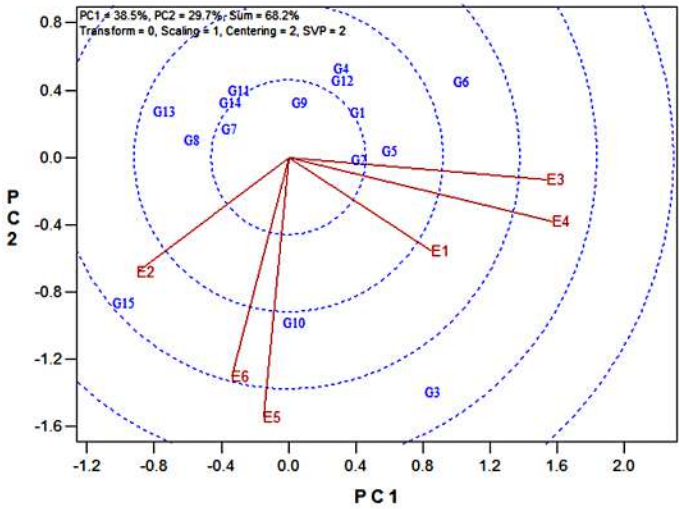


Figure 4: Biplot of correlation map for environments under study.

Thus, three environments including E3, E4 and E5 (2009 to 2011 without irrigation) showed the same vector length and were most discriminating and E1 had the least discriminating ability (Figure 4). Another equally important measure of the test environment is its representativeness of the target environment. The angle between the environment vector and AEC axis is a measure of the representativeness of the environment, where a small angle shows high representativeness and *vice versa*. In this present study E1 was the most representative whereas E2 was the least representative environment (Figure 5).

An ideal environment should be highly differentiating of the genotypes and at the same time representative of the target environment. Ideal test environments should have small (absolute) PC2 scores (more representative of the overall environments) and large PC1 scores (more ability to discriminate genotypes in terms of the genotypic main effect) (Yan *et al.*, 2000; Yan *et al.*, 2002). The ideal environment is represented by the central concentric circle with an arrow passing through it (Figure 5). Figure 5 showed that E1 was the closest to ideal environment and therefore

the most desirable of all six environments. It means that E1 is the most effective for selecting superior inbred lines. In contrast, E2 was the least desirable test environment.

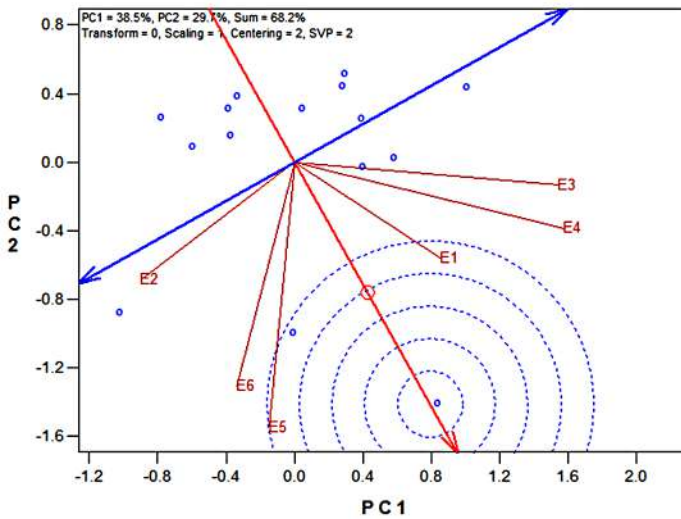


Figure 5: Comparing the environments with the ideal environments based on both discriminating ability and representativeness

ACKNOWLEDGMENT

This work was part of a Regional Sunflower Research Project of Dry-land Agricultural Research Institute (DARI) of Iran and sponsored by the Agricultural Research and Education Organization (AREO). We thank all members of the project for any contribution they may have made towards this work.

REFERENCES

- Baxevanos, D., Goulas, C., Tzortzios, S., Mavromatis, A., 2008. Interrelationship among and repetitiveness of seven stability indices estimated from commercial cotton (*Gossypium hirsutum* L.) variety evaluation trials in three Mediterranean countries. *Euphytica* 161: 371-382.
- Delacy, I.H., Basford, K.E., Cooper, M., Bull, J.K., 1996. Analysis of multi-environment trials and historical perspective. *Plant Adaptation and Crop Improvement*. Eds. M. Cooper and G.L. Hammer. CAB International.
- Ebdon, J.S., Gauch, Jr.H.G., 2002. Additive main effect and multiplicative interaction analysis of national turfgrass performance trials: I. Interpretation of genotype $\times$ environment interaction. *Crop Sci.* 42:489-496.
- Flores, F., Moreno, M.T., Cubero, J.I., 1998. A comparison of univariate and multivariate methods to analyze environments. *Field Crop Res.* 56: 271-286.
- Gabriel, K.R., 1971. The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58(3): 453-467.

- Gauch, H.G., Zobel, R.W., 1988. Predictive and subsequent success of statistical analyses of yield trials. *Theor. Appl. Genet. (TAG)* 76: 1-10.
- Gauch, H.G., 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Science* 46: 1488-1500.
- Kerby, T., Zelinski, L., Burgess, J., Bates, M., Presley, J., 1996. Genetic and environmental contributions to earliness. In: *Proceedings. Belt wide Cotton Conf. National Cotton Council* pp. 592-594.
- Kerby, T., Burgess, J., Lege, K., Albers, D., 2001. Partitioning variety environment contribution to variation in yield, plant growth and fiber quality. In: *Proceedings. Belt wide Cotton Conf. National Cotton Council* pp. 528-532.
- Kroonenberg, P.M., 1995. Introduction to biplots for $G \times E$ tables. University of the Queensland, Brisbane.
- Mohamadi, R., Pourdad, S.S., Amri, A., 2008. Grain yield stability of spring safflower (*Charthamus tinctorius* L.). *Australian Journal of Agricultural Research* 59: 546-553.
- Pourdad, S.S., Mohamadi, R., 2008. Use of Stability Parameters for Comparing Safflower Genotypes in Multi-Environment Trials. *Asian Journal of Plant Sciences* 7(1): 100-104.
- Yan, W., 2000. Singular-value partitioning in biplot analysis of multi-environment trial data. *Agronomy Journal* 94: 990-996.
- Yan, W., 2001. GGE biplot - A windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agronomy Journal* 93: 1111-1118.
- Yan, W., 2002. Singular-value partitioning in biplot analysis of multi-environment trial data. *Agronomy Journal* 94: 990-996.
- Yan, W., Rajcan, I., 2002. Biplot analysis of sites and trait relations of soybean in Ontario. *Crop Science* 42: 11-20.
- Yan, W., Hunt, L.A., Shen, Q., Szlavnics, Z., 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science* 40: 597-605.
- Yan, W., Cornelius, P.L., Crossa, J., Hunt, L.A., 2001. Two type of GGE biplots for analyzing multi-environment trial data. *Crop Science* 41: 656-663.
- Yan, W., Kang, M.S., 2003. GGE biplot analysis: A graphical tool for breeders, geneticists and agronomists. CRC press, Boca Raton, FL.
- Yan, W., Kang, M.S., Ma, B., Woods, S., Cornelius, P.L., 2007. GGE biplot vs. AMMI analysis of genotype – by- environment data. *Crop Science* 47: 643-655.
- Yan, W., Tinker, N.A., 2005. An integrated biplot analysis system for displaying, interpreting and exploring genotype $\times$ environment interaction. *Crop Science* 45: 1004-1016.