

GGE BIPLLOT ANALYSIS USED TO EVALUATE CANE YIELD OF SUGARCANE (*Saccharum* spp.) CULTIVARS ACROSS SITES AND CROP CYCLES

By

J.L. QUEME, H. OROZCO and M. MELGAR

*Centro Guatemalteco de Investigación y Capacitación de la Caña de Azúcar (CENGICAÑA)
Estación Experimental, Finca Camantulul, Km 92.5 Santa Lucía Cotzumalguapa,
Guatemala, Centro América
jlqueme@hotmail.com*

KEYWORDS: GGE Biplot, Genotype-by-Environment (GE) Interaction, Principal Components (PC).

Abstract

MULTI-ENVIRONMENT yield trials (MET) are a series of experiments in which a set of genotypes (G) are evaluated in multiple environments (E). The presence of genotype \times environment (GE) interaction observed in MET complicates the selection and/or recommendation of cultivars. Several statistical methodologies have been developed for the analysis of GE interaction. GGE biplot of the SREG model is a recent methodology based on a graph formed with the first two principal components (PC1 and PC2). GGE represents the G main effect plus the GE interaction effects. Thus, GGE biplot provides an adequate graphical tool for visual analysis of MET data. The objectives of this research were to evaluate sugarcane cultivar responses across environments in order to identify high-yielding cultivars with broad or specific adaptation by applying the GGE biplot analysis. Fourteen sugarcane cultivars were evaluated at three sites in the plant, first and second ratoon crops across the middle land zone of the sugarcane growing area of Guatemala, forming nine environments (three sites \times three crop cycles). Data on tonnes of cane per hectare (TCH) were recorded. The first two principal components (PC1 and PC2) were highly significant ($P < 0.01$) and explained 73% of the GGE. GGE biplot analysis allowed selection of the sugarcane cultivar PR75-2002 as second on average cane yield and had broad adaptation (stable). Two groups of environments (mega-environments) were defined; the first had seven environments and the second one had two environments. The winning cultivars with the highest cane yield were CG00-120 and CG00-092 for each of those groups (specific adaptation), respectively. The G main effect and GE interaction can be exploited by selection of highest cane yield cultivars for each mega-environment.

Introduction

Multi-environment yield trials (MET) are a series of experiments in which a set of genotypes (G) is evaluated in multiple environments (E), usually considered as being the environments as a combination of sites and years. These trials are important because the presence of genotype \times environment interaction (GE) complicates the selection and/or recommendation of cultivars. In order to illustrate this fact, if there is no GE interaction, a single test environment would be sufficient for cultivar testing. Thus, the understanding of the GE interaction observed in MET is very useful in breeding programs in order to identify high-yielding cultivars with broad or specific adaptation (Annicchiarico, 1997; Gauch, 1992; Smith *et al.*, 2001; Yan and Hunt, 2002).

The GE interaction with special interest for breeding programs is that which creates a change in ranking of the cultivars from one environment to another (crossover-interaction). This means that the best cultivar in one particular environment might not be the best in another environment. Several statistical methodologies have been developed for the analysis of GE

interaction (Kang, 2002). Sites (i.e., environment) Regression (SREG) is a linear (additive component) – bilinear (multiplicative component) model which in the bilinear terms absorbs the main effects of G plus the GE interaction (Crossa and Cornelius, 2002; Crossa *et al.*, 2002). Breeders have used the recent methodology GGE biplot from the SREG model to analyse MET data (Burgueño *et al.*, 2009; Yan *et al.*, 2000).

The GGE represents the main effect of genotype plus the genotype by environment interaction (G+GE). GGE biplot of SREG model is based on principal components analysis, and a graph is formed with the scores of the genotypes and the environments of the first principal component (PC1 scores) against their respective scores for the second principal component (PC2 scores). GGE biplot displays the two sources of variation G and GE, and provides an adequate graphical tool for cultivar evaluation (yield and stability), mega-environment analysis (such as ‘which-won-where’ pattern), test-environment evaluation (discriminating among genotypes and representativeness of the mega-environments) and others (Burgueño *et al.*, 2009; Crossa *et al.*, 2002; Ding *et al.*, 2009; Yan *et al.*, 2007).

The present MET study was conducted with 14 sugarcane (*Saccharum* spp.) cultivars, which were evaluated across nine environments (combinations of three sites and three crops cycles) in Guatemala. The objectives of this research were to evaluate sugarcane cultivar responses across environments in order to identify high-yielding cultivars with broad or specific adaptation by applying the GGE biplot analysis.

Materials and methods

This study was part of a series of multi-environment yield trials of the Sugarcane Breeding and Development Program of CENGICAÑA where 14 sugarcane cultivars were evaluated in nine environments of the sugarcane production area in the middle land zone of Guatemala (Table 2). The nine environments refer to sites × crop cycle (crop-year) combinations. There were three sites: San Bonifacio (280 masl), Margaritas (116 masl) and Tuluá (220 masl); and three crop cycles: plant cane (harvest season, 2004–05), first (2005–06) and second ratoon crops (2006–07). Of the 14 cultivars tested, 12 are from CENGICAÑA Guatemala (CG and CGSP) and two testers, one cultivar from Canal Point (CP) and one from Puerto Rico (PR). The field experimental design used for each trial was a Randomised Complete Block with four replications and with experimental units of 75 m². Data on tonnes of cane per hectare (TCH) were recorded.

The GGE biplot from the SREG model was constructed according to the manual and SAS program available at the web page of CIMMYT in Biometrics and Statistic Unit (BSU) or at <http://www.cimmyt.org/english/wps/biometrics/>. Following Burgueño *et al.* (2009), Yan *et al.* (2001) and Yan and Hunt (2002), a brief description and interpretation of the GGE biplot is presented below:

PC1 (as x-axis) and PC2 (as y-axis)

The biplot GGE is formed with PC1 scores on the abscissa and PC2 scores on the ordinate for each cultivar and each environment. PC1 represents the average yield of cultivars if the genotype PC1 scores are highly correlated with the genotype main effects; while PC2 is related to the GE interaction.

Cultivar and environment vectors

The cultivars and environments can be visualised as vectors from the origin of axis (0, 0) to the end points determined by their scores. The environment scores have the same orientation to the right (same sign). Cultivars that are far from the origin have positive responses with environments that are far from the origin but in the same direction. Ideal cultivars: large PC1 scores (higher mean yield) and PC2 scores near to zero (more stable). Ideal test environments: Large PC1 scores (high capacity to discriminate among cultivars) and PC2 scores near to zero (‘more representative of an average environment’).

External polygon and biplot sectors

External polygon is formed with the cultivars located farthest from the biplot origin (0, 0). These cultivars are part of the corners of the polygon (cultivar markers) and are connected with straight lines. Cultivar markers are the ones with the highest response and they could be the best or the worst in some or all of the environments. Perpendicular segmented lines to each side of the polygon are drawn to divide the biplot in different sectors of cultivars and environments. The cultivar located in the corner of one sector is the winner cultivar with the highest yield in the environment or environments of this sector.

Results and discussion

The combined analysis of variance (Table 1) showed differences statistically significant ($P < 0.01$) among environments, genotypes and GE interaction. Based on the relative contribution of sum of squares, the environmental effect had the highest contribution, followed by genotypes and lastly the GE interaction. The first two principal components (PC1 and PC2) were highly significant ($P < 0.01$) and explained 73% of the GGE (PC1=61% and PC2= 12%). Cultivar PC1 scores and the cultivar yield means (Table 2) were highly correlated ($r=0.99$), which was expected, since PC1 explained most of the GGE variation (61%). These results are consistent with those of Yan *et al.* (2001), who indicated that a poor correlation occurs when GGE is explained by PC1 and PC2 in similar contribution (complex GE interaction).

Table 1—Additive main effects and multiplicative interactions analysis of variance on cane yield (t/ha) of the genotypes across environments.

Source of variation	Degrees of freedom	Sum of squares	Mean squares	F	Prob.
Environment (E)	8	317846	39731	119	0.000
Genotypes (G)	13	77260	5943	18	0.000
GE	104	69722	670	2	0.000
GGE		146982			
PC1	20	90258	4513	14	0.000
PC2	18	17733	985	3	0.000
Error	351	116865	333		

According to the GGE biplot (Figure 1), no cultivar was identified which had high average yield (large PC1 score) and stability in performance (PC2 score near to zero). However the cultivar 13 (PR75-2002) was close to ideal because it had the second highest average cane yield and a PC2 score near to zero (-0.1).

Regarding test environments, S2Y1 (Margaritas-plant cane) and S2Y2 (Margaritas-first ratoon) environments were identified as near to an ideal environment since they presented large PC1 score and PC2 score near to zero (Table 2 and Figure 1). This indicates that this site (Margaritas) would be useful for selecting cultivars which have a high correlation across all environments.

The S3Y1, S3Y2, S3Y3 and S1Y1 environments were near to the biplot origin (0, 0) with short vectors; thus these environments were not useful for cultivar discrimination. The S2Y3 environment had the largest vector with large PC1 and PC2 scores of 5.9 and -3.6 respectively (Table 2).

This environment adequately discriminated the cultivar 5 (CG00-120) with large vector in the same direction. The same interpretation can be given to environments S1Y2 and S1Y3 with the cultivar 8 (CG00-092). These results suggest that the GE interaction in sugarcane increases as the years (crop cycles) also increase.

Table 2—Mean yield of cane (t/ha) of 14 sugarcane cultivars in nine environments and PCA1 and PCA2 scores for each cultivar and environment.

Code	Cultivars	Environments †									Grand Mean	PC1 ‡	PC2 ‡
		S1Y1	S1Y2	S1Y3	S2Y1	S2Y2	S2Y3	S3Y1	S3Y2	S3Y3			
1	CGSP98-08	161	131	129	177	151	162	110	131	105	140	-1.7	-0.1
2	CG00-032	121	124	127	148	129	143	95	119	111	124	-5.5	2.5
3	CGSP98-05	136	122	116	176	179	181	122	150	131	146	0.1	-2.4
4	CGSP98-16	140	117	104	192	170	178	108	132	120	140	-0.7	-3.2
5	CG00-120	159	150	146	212	199	246	111	151	131	167	6.5	-2.8
6	CG00-129	131	138	133	168	143	175	114	113	117	137	-2.1	1.5
7	CG00-001	158	106	107	176	154	153	97	116	95	129	-3.8	-2.6
8	CG00-092	161	169	162	200	193	176	85	144	130	158	3.6	4.1
9	CG99-045	145	114	126	170	155	167	96	135	123	137	-2.1	-0.9
10	CG00-028	123	146	145	206	170	183	122	129	125	150	1.5	1.9
11	CG00-044	122	146	129	192	175	156	113	131	124	143	-0.3	2.4
12	CG-99-014	151	150	142	204	195	191	118	134	136	158	3.5	0.5
13	PR75-2002	174	154	132	231	183	178	123	139	138	161	4.00	-0.1
14	CP72-2086	140	115	115	175	156	159	100	113	119	132	-2.9	-0.9
	Grand Mean	144	134	129	188	168	175	108	131	122	144		
	PC1 ‡	2.4	4.3	3.1	5.7	5.7	5.9	1.3	2.8	2.8			
	PC2 ‡	-2.1	4.9	4.6	-0.3	-0.9	-3.6	-1.1	-0.9	0.7			

† Environments are a combination of sites and crop cycles. Sites: S1=San Bonifacio, S2=Margaritas and S3=Tululá. Crop cycles: Y1=Year 1 (Plant Cane), Y2=Year 2 (First Ratoon) and Y3=Year 3 (Second Ratoon).
‡ PCA1 and PCA2 = First and second principal component respectively.

The site 2 (Margaritas) is characterised by adequate soil and weather conditions compared with site 3 (Tululá); therefore, the environments that include the site 2 had high capacity to discriminate among cultivars, as indicated by data in Table 2 and represented graphically in the GGE biplot (Figure 1).

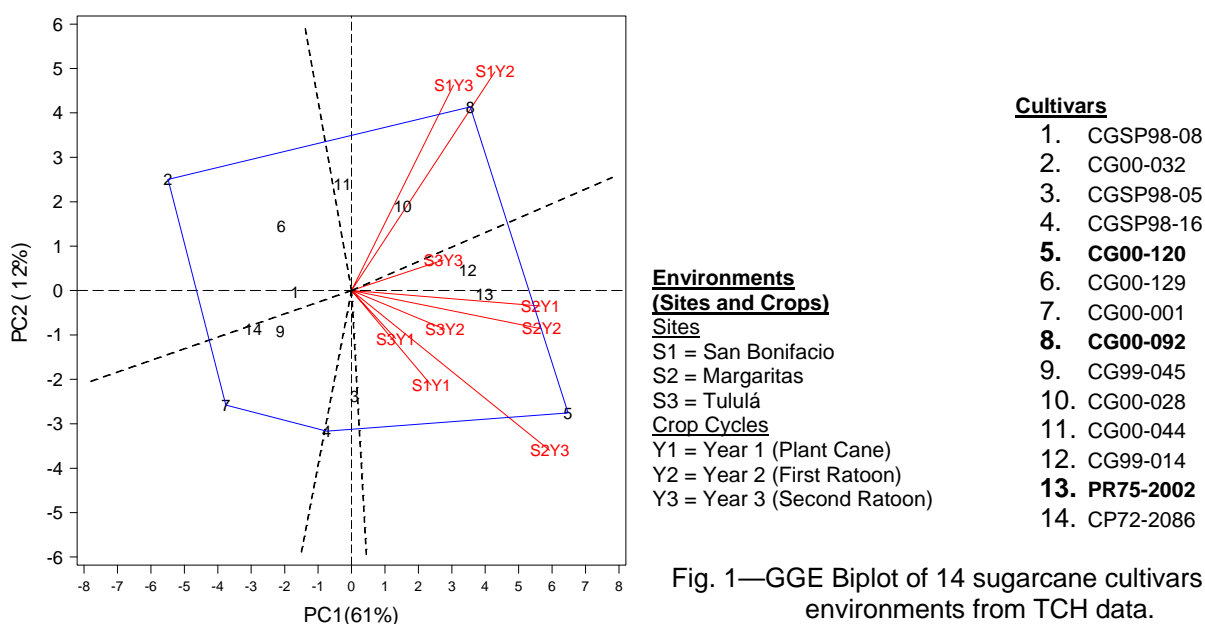


Fig. 1—GGE Biplot of 14 sugarcane cultivars in nine environments from TCH data.

The 'which-won-where' patterns were determined by the polygon (Figure 1). The corners of the polygon were formed with five cultivars from which cultivars 5 (CG00-120) and 8 (CG00-092) showed the best response in TCH. The poorest cultivars were 4 (CGSP98-16), 7 (CG00-001) and 2 (CG00-032) because they are in an opposite direction to the environments. According to the GGE biplot, two sectors (mega-environments) were defined; the first sector made up by seven environments (Margaritas and Tululá with their three crops cycle plus San Bonifacio with its first ratoon); and the second one by two environments (San Bonifacio with its second and third crops). The winning cultivars with the highest yield were 5 (CG00-120) and 8 (CG00-092) for each of the sectors, respectively.

Conclusions

GGE biplot analysis allowed selection of the sugarcane cultivar PR75-2002 as second for average cane yield and with broad adaptation (stable).

Two groups of environments (mega-environments) were defined; the first made up of seven environments and the second one of two environments. The winning cultivars with the highest cane yield were CG00-120 and CG00-092 for each of the groups (specific adaptation) respectively.

The G main effect and GE interaction can be exploited by selection of highest cane yield cultivars for each mega-environment.

REFERENCES

- Annicchiarico, P.** (1997). Additive main effects and multiplicative interaction (AMMI) analysis of genotype-location interaction in variety trials repeated over years. *Theor. Appl., Genet.*, 94: 1072–1077.
- Burgueño, J., Crossa, J. and Vargas, M.** (2009). SAS programs for graphing ge and gge biplots. Biometrics and Statistics Unit. CIMMYT. México. 16 Mar. 2009. <http://www.cimmyt.org/english/wps/biometrics/>
- Crossa, J., Cornelius, P.L. and Yan, W.** (2002). Biplot of linear-bilinear models for studying crossover genotype \times environment interaction. *Crop Sci.*, 42: 619–633.
- Crossa, J. and Cornelius, P.L.** (2002). Linear-bilinear model for the analysis of genotype-environment interaction. In: Kang, M.S. ed. *Quantitative Genetics, Genomics and Plant Breeding*. CABI Publishing, 305–321.
- Ding, M., Tier B. and Yan W.** (2009). Application of GGE biplot analysis to evaluate Genotype (G), Environment (E) and G \times E interaction on *P. radiata*: a case study. 05 Jul. 2009. [http://www.cdesign.com.au/proceedings_afgc2007/papers%20\(pdf\)/thurs%20ding.pdf](http://www.cdesign.com.au/proceedings_afgc2007/papers%20(pdf)/thurs%20ding.pdf)
- Gauch, H.G.** (1992). *Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs*. Amsterdam, Elsevier.
- Kang, M.S.** (2002). Genotype-environment interaction: progress and prospects. In: Kang, M.S. ed. *Quantitative Genetics, Genomics and Plant Breeding*. CABI Publishing, 221–243.
- Smith, A., Cullis, B. and Thompson, R.** (2001). Analysing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57, 1138–1147.
- Yan, W., Cornelius P. L., Crossa, J., and Hunt, L. A.** (2001). Two types of GGE biplots for analysing mulyi-environment trial data. *Crop Sci.*, 41: 656–663.
- Yan, W. and Hunt, L.A.** (2002). Biplot analysis of multi-environment trial data. In: Kang, M.S. ed. *Quantitative Genetics, Genomics and Plant Breeding*. CABI Publishing, 289–303.
- Yan, W., Hunt, L.A., Sheng, Q. and Szlavnics, Z.** (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.*, 40: 597–605.
- Yan, W., Kang, M.S., Ma, B., Woods, S. and Cornelius, P. L.** (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.*, 47: 643–655.

**L'UTILISATION DE L'ANALYSE BIPLLOT GGE POUR L'ÉVALUATION
DU RENDEMENT DES CULTIVARS DE LA CANNE À SUCRE
(*Saccharum spp.*) À TRAVERS LES SITES ET LES CYCLES DE CULTURE**

Par

J.L. QUEME, H. OROZCO et M. MELGAR

*Centro Guatemalteco de Investigación y Capacitación de la Caña de Azúcar (CENGICAÑA)
Estación Experimental, Finca Camantulul, Km 92.5 Santa Lucía Cotzumalguapa, Guatemala,
Amérique Centrale*

jlqueme@hotmail.com

**MOTS CLÉS: Biplot GGE, Interaction Génotype × Environnement (GE),
Composantes Principales (PC).**

Résumé

LES ESSAIS multi-locaux (MET) constituent une série d'essais variétaux où un groupe de génotypes (G) sont évalués dans divers environnements (E). La présence de l'interaction génotype x environnement (GE) complique la sélection et/ou les recommandations des cultivars. Plusieurs méthodes statistiques ont été développées pour l'analyse de l'interaction GE. Le biplot GGE découlant du modèle SREG est une méthodologie récente basée sur la courbe formée des deux composantes principales (PC1 et PC2). Le GGE représente l'effet principal G plus les effets de l'interaction GE. Donc, le biplot GGE fournit un outil graphique adéquat pour une analyse visuelle des données MET. Les objectifs de cette recherche étaient d'évaluer le comportement des cultivars à travers les environnements afin d'identifier les cultivars à forts rendements à adaptations larges ou spécifiques en utilisant l'analyse biplot GGE. Quatorze cultivars de canne à sucre ont été évalués sur trois sites en vierge, première et deuxième repousses dans la zone centrale où la canne est cultivée au Guatemala, totalisant neuf environnements (trois sites x trois récoltes). Le rendement en tonne de canne à l'hectare (TCH) était recueilli. Les deux premières composantes principales (PC1 et PC2) étaient hautement significatives ($P < 0.01$) et contribuaient à hauteur de 73% de l'interaction GGE. L'analyse biplot GGE a permis la sélection du cultivar PR75-2002 en deuxième position pour le rendement en canne avec une large aire d'adaptation large (stable). Deux groupes d'environnements (mega-environnements) étaient définis; le premier avec sept environnements et le second avec deux environnements. Les cultivars gagnants avec les rendements les plus élevés étaient CG00-120 et CG00-092 pour chacun de ces deux groupes (adaptation large et spécifique) respectivement. L'effet principal G et l'interaction GE peuvent être exploités en sélectionnant les cultivars avec les plus forts rendements pour chacun de ces mega-environnements.

**ANALISIS GGA BILOT UTILIZADO PARA EVALUAR EL RENDIMIENTO
DE CAÑA EN CULTIVARES DE CAÑA DE AZÚCAR
(*Saccharum spp.*) A TRAVES DE SITIOS Y CICLOS DEL CULTIVO**

Por

J.L. QUEME, H. OROZCO y M. MELGAR

*Centro Guatemalteco de Investigación y Capacitación de la Caña de Azúcar (CENGICAÑA)
Estación Experimental, Finca Camantulul, Km 92.5 Santa Lucía Cotzumalguapa,
Guatemala, Centro América
jlqueme@hotmail.com*

**PALABRAS CLAVE: GGA biplot, Interacción
Genotipo × Ambiente (GA), Componentes Principales (CP).**

Resumen

ENSAYOS de rendimiento en multi-ambientes (EMA) son una serie de experimentos en los cuales un grupo de genotipos (G) son evaluados en diferentes ambientes (A). La presencia de la interacción genotipo × ambiente (GA) observada en EMA complica la selección y / o recomendación de cultivares. Varias metodologías estadísticas se han desarrollado para el análisis de la interacción GA. GGA biplot del modelo de Regresión de sitios (SREG) es una metodología reciente basada en una gráfica formada con los dos primeros componentes principales (CP1 y CP2). GGA representa el efecto principal de G más los efectos de la interacción GA. Así, GGA biplot proporciona una herramienta gráfica adecuada para el análisis visual de datos provenientes de los EMAs. Los objetivos de esta investigación fueron evaluar las respuestas de cultivares de caña de azúcar a través de ambientes con el fin de identificar cultivares de alto rendimiento con una adaptación amplia y específica mediante la aplicación análisis GGA biplot. Catorce cultivares de caña de azúcar fueron evaluados en tres sitios en plantía, primera y segunda soca a través de la zona media de producción de caña de azúcar de Guatemala, formando nueve ambientes (tres localidades × tres ciclos del cultivo). Se registraron datos de toneladas de caña por hectárea (TCH). Los dos primeros componentes principales (CP1 y CP2) fueron altamente significativos ($P < 0.01$) y explicaron el 73 % del GGA. GGA biplot permitió seleccionar el cultivar PR75-2002 como segundo lugar en el promedio en rendimiento de caña con adaptación amplia (estable). Dos grupos de ambiente (mega-ambientes) fueron definidos, el primero compuesto por siete ambientes y el segundo por dos ambientes. Los cultivares ganadores con mayor rendimiento de caña fueron: CG00-120 y CG00-092 para cada uno de los grupos, respectivamente (adaptación específica). El efecto principal G y los efectos de la interacción GA pueden ser efectivamente explotadas mediante la selección de cultivares de mayor rendimiento para cada mega-ambiente.