



# Genotype–Environment Interaction, Megaenvironments and Two-Table Coupling Methods for Sugarcane Yield Studies in Venezuela

Ramón Rea<sup>1</sup> · Orlando De Sousa-Vieira<sup>2</sup> · Alida Díaz<sup>2</sup> · Miguel Ramón<sup>3</sup> · Rosaura Briceño<sup>2</sup> · José George<sup>2</sup> · Milagros Niño<sup>2</sup> · Leandro Balzano-Nogueira<sup>1</sup>

Received: 29 May 2015 / Accepted: 21 September 2015 / Published online: 20 October 2015  
© Society for Sugar Research & Promotion 2015

**Abstract** The selection of sugarcane cultivars adapted to different environments becomes difficult when there is genotype–environment interaction (GEI). The data were analyzed from twenty sugarcane genotypes evaluated in eight locations over two crop cycles to identify megaenvironments (ME), through GEI methods for higher cane yield measured in tons of cane per hectare (TCH) and percentage of sucrose (Pol% cane) using biplot multivariate GEI models. The best genotypes and the environment were determined with better mean yield by the two-table coupling coinertia method. Additive main effects and multiplicative interaction (AMMI) analyses revealed significant GEI with respect to both variables. The AMMI stability value exposed high genotypes stability for Pol% cane, but for TCH just G4, G8, G1, G20 and G17 have stability in all environments. The site-type regression SREG-GGE biplot showed two ME for TCH and one for Pol% cane. Although both yield variables showed mean negative correlation, through coinertia analysis it was possible to determine that G15, G17 and G13 were the best genotypes for both variables in all environments, besides “Los Tamarindos” was the best environment, with both variables correlated positively, and G11, G13, G12 could be considered the better

genotypes. This work revealed the necessity of using coinertia as a complementary analysis to AMMI and GGE, which needs to be applied to determine the genotypes and environments that favor multiple yield variables, in order to increase the productivity.

**Keywords** AMMI analysis · SREG-GGE biplot · Coinertia analysis · Pol% cane · *Saccharum* spp.

## Introduction

In Venezuela, sugarcane is cultivated under different soil types, fertility levels and humidity. The unequal performance of genotypes in different environments (genotype × environment interaction or GEI) in yield assays is a challenge for breeders. It has been shown that GEI reduces selection progress and complicates the identification of superior cultivars of sugarcane in regional trials (Rea et al. 2011; Rea and De Sousa-Vieira 2002). When GEI is present, one of the options is to use stability analysis to identify cultivars with higher and more stable yields. Therefore, several statistical methods have been proposed and used to study adaptation and stability of varieties in different locations (Roostaei et al. 2014; Dolédec and Chessel 1994; Dray et al. 2003b). Some multivariate models such as additive main effects and multiplicative interaction (AMMI), site regression model (SREG) and coinertia analysis have been used for the interpretation of GEI. The AMMI and SREG models are similar, but in SREG the linear terms of genotypes (G) and environments (E) are not considered individually, instead they are added to the multiplicative term of genotype × environment (GE) interaction (Crossa et al. 2002; Silveira et al. 2013; Yan 2011).

AMMI models are capable of measuring the weight of the environments, the genotypes and their interactions

✉ Leandro Balzano-Nogueira  
leobalzano@gmail.com

<sup>1</sup> Fundación Instituto de Estudios Avanzados, Carretera Nacional Hoyo de la Puerta, Valle de Sartenejas, Baruta, Edo. Miranda, Apartado 17606, Caracas 1015, Venezuela

<sup>2</sup> Instituto Nacional de Investigaciones Agrícolas-Yaracuy, Carretera vía aeropuerto, Flores Boraure, Km.3, al frente de Multifrut, San Felipe, Estado Yaracuy, Venezuela

<sup>3</sup> Instituto Nacional de Investigaciones Agrícolas-Portuguesa, Carretera Nacional Acarigua a Barquisimeto Km. 5., Araure, Estado Portuguesa, Venezuela

throughout a value that measures how stable a genotype is in all environments in terms of yield. In contrast, the SREG emphasizes the behavior of genotypes through a regression upon locations, which is important for studying the possible existence of different megaenvironments in a region (Yan et al. 2007). In fact, SREG models include  $G + GE$  or  $E + GE$  in the bilinear terms and provide a graphical analysis of easy interpretation called GGE biplot (Yan and Tinker 2006) that has been used in many cultivar  $\times$  environment interaction studies (Akcura et al. 2011; Jalata 2011; Mortazavian et al. 2014; Rodríguez et al. 2012; Ramburan and Zhou 2011; Rao et al. 2011; Rea et al. 2011). The biplot representation displays the grouping of genotypes and environments with similar response patterns and permits to identify the most representative and most discriminating environments (Yan 2011). GGE biplot is developed from the first two principal components (PC) of the SREG model. Genotypes close to each other in the biplot indicate similar response patterns across environments. Meanwhile, nearby environments exhibit an acute angle between them, indicating similar environmental conduct. The lack of association between environments is given by a  $90^\circ$  angle between vectors and negative association by angles greater than  $90^\circ$ – $180^\circ$  (Ibáñez et al. 2006). Genotypes located on the vertices of the polygon performed either the best or the poorest in one or more environments (Yan and Tinker 2006). Perpendicular lines drawn on each side of the polygon form groups of locations or genotypes with similar behavior. In turn, coinertia analysis is a multivariate method that identifies trends or co-relationships in multiple datasets that contain the same samples and simultaneously finds ordinations (dimension reduction diagrams) from the datasets that are most similar. It does this by finding successive axes from the two datasets with maximum covariance (Culhane et al. 2003; Dray et al. 2003a). Separate analyses find axes maximizing inertia in each hyperspace. These axes are projected in a subspace, and each individual is represented by an arrow, where the beginning of the arrows is the position of the variety described by one data matrix and the arrowhead is the position of the variety described by the other matrix (Dolédec and Chessel 1994). The aim of this study was to identify megaenvironments (ME) and to determine optimal genotypes and environments for higher sugarcane yield measured by tons of cane per hectare (TCH) and percentage of sucrose (Pol% cane).

## Materials and Methods

### Material Selection

To carry out this work, data from the regional (outfield) trials of sugarcane in the last stage of selection breeding

program at the National Institute for Agricultural Research (INIA) were used. These assays were performed using a randomized complete block design with three replicates and experimental units of  $45 \text{ m}^2$  (3 rows of 10 m with 1.5 m spacing). These assays were conducted for 2 years with successive cuts (plant and first ratoon).

### Locations and Genotypes

The trials locations in Venezuela were: Quebrada Arriba (QA) and Montaña Verde (MV) located in Lara state; Las Majaguas (LM), Ivonne (Iv) and Castillera (Ca) in Portuguesa state; Santa Lucía (SL) and FUNDACAÑA (FC) in the Yaracuy state, and Los Tamarindos (LT) in Aragua state (Table 1). The main characteristics of soils and rainfall of these areas are also given in Table 1. The evaluated experimental material consisted of the following clones: V91-1 (G1), V91-2 (G2), V91-6 (G3), V91-8 (G4), V91-15 (G5), V98-62 (G6), V98-86 (G7), V98-120 (G8), V99-117 (G9), V99-190 (G10), V99-203 (G11), V99-208 (G12), V99-213 (G13), V99-217 (G14), V99-236 (G15), V99-245 (G16), V00-50 (G17) and three reference clones: B80-408 (G18), C323-68 (G19) and CP74-2005 (G20).

### Analyzed Yield Variables

The analyzed variables were agronomic performance in tons of cane per hectare (TCH), determined by weighing the effective area of each experimental unit at harvest time and the estimate of industrial yield (sugar) through Pol% cane, determined by laboratory tests on samples composed of 10 stalks per experimental unit.

### Statistical Analyses

To evaluate yield (TCH and Pol% cane) and to consider general and specific adaptability and possible groupings of environments, AMMI and GGE biplot methodologies were performed (Crossa et al. 2002; Yan 2011). AMMI models

**Table 1** Main characteristics of soil and precipitation/year of the locations tested

Location	Soil (texture)	Precipitation (mm/year)	pH
Quebrada Arriba (QA)	Clay loam	1101	7.7
Santa Lucía (SL)	Silty clay loam	700	8.0
Montaña Verde (MV)	Loam	1048	7.3
Las Majaguas (LM)	Clay loam	1500	7.0
Ivonne (Iv)	Clay loam	1500	7.0
Castillera (Ca)	Clay loam	1500	7.0
FUNDACAÑA (FC)	Silt loam	1111	8.1
Los TamarindoT (LT)	Sandy loam	1051	6.7

were performed using R package “agricolae” (De Mendiburu 2015). SREG-GGE biplot analyses were performed using R package “GGEBiplotGUI” (Frutos et al. 2014). To identify genotypes and environments with positive relationship between both yield characteristics simultaneously

**Table 2** Mean squares analysis of variance for cane yield (TCH) and Pol% cane

Source	df	TCH <sup>a</sup>	Explained variation (%)	Pol% cane	Explained variation (%)
Treatments	159	4340.54**		1819.38**	
Environment (E)	7	48,801.08**	49.50	83.89**	32.27
Genotype (G)	19	11,902.64**	32.77	47.21**	49.29
GEI	133	920.22*	17.73	2.52*	18.45
Error	800	839.65		1.69	

df degrees of freedom

\*, \*\* Significant at 5 and 1 % probability, respectively

<sup>a</sup> Tons of cane per hectare

(TCH and Pol% cane), coinertia analysis (Dolédec and Chessel 1994; Dray et al. 2003b) was performed using R package “ade4” (Dray and Dufour 2007; Chessel et al. 2004; Dray et al. 2007). For the selection of genotypes, it was used two  $n \times p$  matrices, one for each yield variable, where the genotypes were classified as individuals and the environments were the variables. On the other hand, for best environment determination, the above matrices were transposed so the environments were the individuals and the genotypes the variables. Then, each matrix was used for a principal components analysis (PCA), and then, each pair of matrices was constrained into a subspace of maximal coinertia and maximal correlation between both yield variables. To determine the best suited genotypes for the best environment taking into account both yield variables, a covariance analysis was performed, applying the formula:

$$(G_{ij\text{TCH}} - \bar{G}_i\text{TCH})(G_{ij\text{Pol\%}} - \bar{G}_i\text{Pol\%})$$

where  $G_{ij\text{TCH}}$  is the value of TCH of a genotype  $i$  in an environment  $j$ .  $\bar{G}_i\text{TCH}$  is the mean value of TCH of a genotype  $i$  in all environments.  $G_{ij\text{Pol\%}}$  is the mean value of Pol% cane of a genotype  $i$  in an environment  $j$ .  $\bar{G}_i\text{Pol\%}$  is

**Table 3** Average yield in tons of cane per hectare (TCH) of twenty sugarcane genotypes at eight locations in two cycles of harvest

Genotypes/ locations	Quebrada Arriba (QA)	Santa Lucia (SL)	Fundacaña (FC)	Montaña Verde (MV)	Las Majaguas (LM)	Ivonne (Iv)	Castillera (Ca)	Los Tamarindos (LT)	Mean Genotypes	ASV	Rank of ASV
V91-1 (G1)	127.69	92.25	117.22	129.80	88.86	96.53	69.70	145.46	108.44	0.99	3
V91-2 (G2)	126.06	75.44	116.96	122.42	102.61	78.15	84.30	108.37	101.79	2.88	12
V91-6 (G3)	149.21	57.73	110.76	137.87	102.48	90.26	97.69	122.91	108.61	3.47	14
V91-8 (G4)	125.15	86.67	109.62	126.67	100.51	87.79	96.18	132.16	108.09	0.21	1
V91-15 (G5)	110.14	96.59	108.68	147.11	110.98	100.49	111.65	190.48	122.01	6.28	20
V98-62 (G6)	165.67	111.52	140.55	178.31	146.32	104.48	108.75	142.00	137.20	3.59	16
V98-86 (G7)	113.77	86.86	94.21	118.25	106.52	112.03	94.37	130.14	107.02	2.20	9
V98-120 (G8)	163.78	116.46	128.04	179.91	138.16	116.80	126.67	166.54	142.04	0.77	2
V99-117 (G9)	122.67	98.88	135.79	129.16	112.45	84.17	113.74	143.55	117.55	1.99	8
V99-190 (G10)	177.54	121.25	122.38	147.87	110.33	103.20	115.40	159.65	132.20	4.06	18
V99-203 (G11)	160.97	121.61	110.81	140.70	117.57	107.85	103.54	153.75	127.10	3.56	15
V99-208 (G12)	186.65	105.56	119.51	144.38	148.22	115.64	124.49	173.03	139.68	4.35	19
V99-213 (G13)	136.12	103.95	140.53	168.03	146.31	118.29	124.72	163.95	137.74	2.78	11
V99-217 (G14)	126.41	88.19	92.66	137.97	98.05	106.99	94.93	138.53	110.47	1.88	7
V99-236 (G15)	171.29	97.83	143.54	156.35	122.55	107.48	121.00	159.37	134.93	2.30	10
V99-245 (G16)	89.02	60.80	103.44	132.84	64.93	79.29	78.05	109.88	89.78	3.60	17
V00-50 (G17)	147.69	110.04	133.07	161.04	148.79	120.92	133.98	166.84	140.30	1.28	5
B80-408 (G18)	119.70	75.92	137.18	130.66	115.04	105.95	100.46	140.85	115.72	3.17	13
C323-68 (G19)	155.80	89.23	132.01	156.89	124.13	111.99	103.87	152.33	128.28	1.55	6
CP74-2005 (G20)	111.53	80.69	98.85	112.40	93.31	88.69	93.95	122.35	100.22	1.12	4
Mean locations	139.34	93.87	119.79	142.93	114.91	101.85	104.87	146.11	120.46		
MSD $\alpha = 0.05$	30.68	28.83	32.32	22.23	30.53	49.79	40.74	20.08			

Rank of ASV: Ranking orders of the 20 genotypes

ASV AMMI stability values, MSD minimal significant difference

the mean value of Pol% cane of a genotype *i* in all environments.

All three programs “agricolae,” “GGEbiplotGUI” and “ade4” were developed as Comprehensive R Archive Network (CRAN) in R software.

## Results and Discussion

### Interrelationship Between Genotypes and Locations

The average sugarcane yield (TCH) and Pol% cane were significantly affected by environmental and genotypic effects ( $p \leq 0.01$ ). Environmental effects accounted for 49.50 and 32.77 %, while genotypic effects explained 32.27 and 49.29 % of the total (E + G + GEI) variation (Table 2). This response to environmental and genotypic effects coincides with those found by Rea et al. (2011). Average TCH yields of genotypes ranged from 89.78 for V99-245 to 142.04 for V98-120, while TCH yield per location ranged from 93.87 for Santa Lucía to 146.11 for Los Tamarindos (Table 3). The average sugar content expressed in Pol% cane

ranged from 12.29 % for V91-15 to 15.51 % for V99-245, while Pol% cane per location ranged from 12.54 % in Montaña Verde to 14.63 % in Santa Lucía (Table 4).

Also, the AMMI stability values (ASVs) are reported, noticing that for the case of TCH (Table 3), there are genotypes with high values of ASV, mainly G5 and G12, which means more unstability; in other words, the environments tend to affect more strongly the sugarcane growth rate. In the case of yield measured in Pol% cane (Table 4), all genotypes are relatively low (stable) in all studied environments which means the sucrose content is not so affected by environment like TCH.

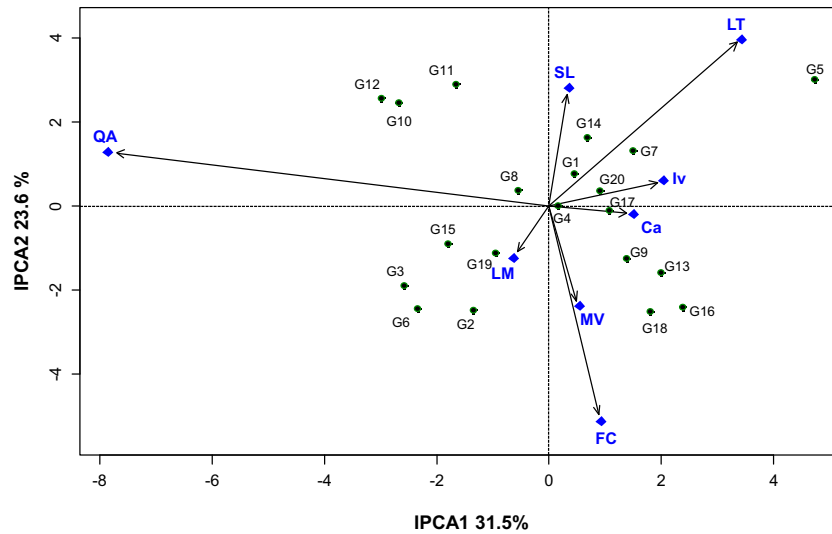
Biplots of AMMI models for both yield variables were generated using genotypic and environmental scores of the first two AMMI components (Rea et al. 2011). AMMI stability values of the genotypes in all environments are graphically represented for both yield variables (Figs. 1, 2). The statistically stable genotypes are represented by points near the origin in the AMMI2 biplot, with values near zero for the two axes of interaction (IPCA1 and IPCA2). Distribution of genotype points in the AMMI2 biplot for TCH (Fig. 1) revealed that the genotypes G4, G8, G1, G20

**Table 4** Average yield in Pol% cane of twenty sugarcane genotypes at eight locations in two cycles of harvest

Genotypes/ locations	Quebrada Arriba (QA)	Santa Lucia (SL)	Fundacaña (FC)	Montaña Verde (MV)	Las Majaguas (LM)	Ivonne (Iv)	Castillera (Ca)	Los Tamarindos (LT)	Mean Genotypes	ASV	Rank of ASV
V91-1 (G1)	13.59	15.82	14.81	13.47	13.14	14.72	13.99	14.92	14.31	0.54	9
V91-2 (G2)	11.12	13.54	12.04	11.35	12.11	11.91	10.80	14.77	12.20	0.51	8
V91-6 (G3)	11.96	14.05	13.37	11.98	13.45	13.10	12.52	14.16	13.07	0.42	5
V91-8 (G4)	12.34	13.32	11.73	11.04	11.73	13.79	13.16	14.29	12.67	0.56	10
V91-15 (G5)	11.09	14.21	12.01	10.82	12.64	12.75	11.96	12.82	12.29	0.22	3
V98-62 (G6)	13.87	14.34	14.15	13.06	12.36	13.26	12.57	14.04	13.45	0.62	11
V98-86 (G7)	14.15	15.66	13.75	11.64	13.96	14.73	13.99	15.85	14.22	0.97	19
V98-120 (G8)	12.88	14.08	13.93	13.00	12.43	13.64	13.76	13.50	13.40	0.77	16
V99-117 (G9)	11.85	13.09	12.66	11.11	12.64	13.02	12.38	13.19	12.49	0.15	1
V99-190 (G10)	11.92	14.48	13.35	11.76	11.38	13.12	13.30	14.35	12.96	0.48	7
V99-203 (G11)	12.88	14.53	13.30	11.38	12.02	14.43	12.93	15.81	13.41	0.72	14
V99-208 (G12)	12.75	14.60	13.62	14.79	13.09	14.37	13.73	15.33	14.03	0.93	18
V99-213 (G13)	13.07	14.93	14.42	13.33	13.18	13.72	14.13	15.80	14.07	0.20	2
V99-217 (G14)	12.45	13.86	12.33	10.65	13.10	12.92	12.64	14.29	12.78	0.72	13
V99-236 (G15)	14.23	16.13	15.51	12.92	13.94	15.75	14.95	15.82	14.91	0.44	6
V99-245 (G16)	14.34	17.14	16.34	14.10	15.91	15.27	15.51	17.72	15.79	0.62	12
V00-50 (G17)	13.73	14.48	14.92	13.03	14.33	14.34	14.01	15.40	14.28	0.32	4
B80-408 (G18)	13.59	15.44	14.10	14.22	13.37	15.46	14.74	14.91	14.48	0.78	17
C323-68 (G19)	11.60	14.33	14.35	13.54	13.56	14.92	13.87	15.16	13.92	0.74	15
CP74-2005 (G20)	13.48	14.55	15.94	13.72	16.35	15.55	14.94	16.99	15.19	1.24	20
Mean locations	12.84	14.63	13.83	12.54	13.23	14.04	13.9	14.50	13.69		
MSD $\alpha = 0.05$	1.68	1.60	1.39	1.74	1.47	1.48	1.35	1.37			

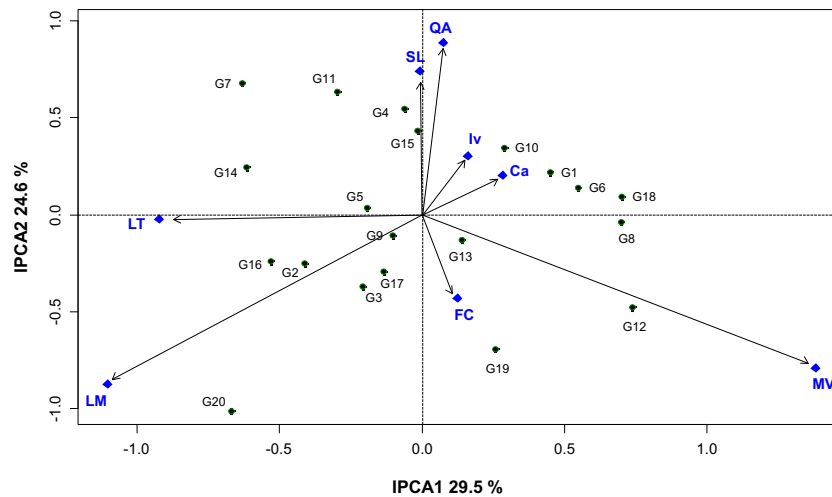
Rank of ASV: Ranking orders of the 20 genotypes

ASV AMMI stability values, MSD minimal significant difference



**Fig. 1** AMMI biplot for TCH yield of the first two GE interaction principal components axes of 20 genotypes and 8 environments. The biplot reveals stability of yield just for G4, G8, G1, G20 and G17; the remaining 15 genotypes scattered away from the origin in the biplot, indicating that the genotypes were more sensitive to environmental interactive forces. Principal components explained 55.1 % of GEI interaction. G1 V91-1; G2 V91-2; G3 V91-6; G4 V91-8; G5 V91-15;

G6 V98-62; G7 V98-86; G8 V98-120; G9 V99-117; G10 V99-190; G11 V99-203; G12 V99-208; G13 V99-213; G14 V99-217; G15 V99-236; G16 V99-245; G17 V00-50; G18 B80-408; G19 C323-68; G20 CP74-2005. QA Quebrada Arriba, MV Montaña Verde, LM Las Majaguas, Iv Ivonne, Ca Castellera, SL Santa Lucía, FC FUNDACAÑA, LT Los Tamarindos

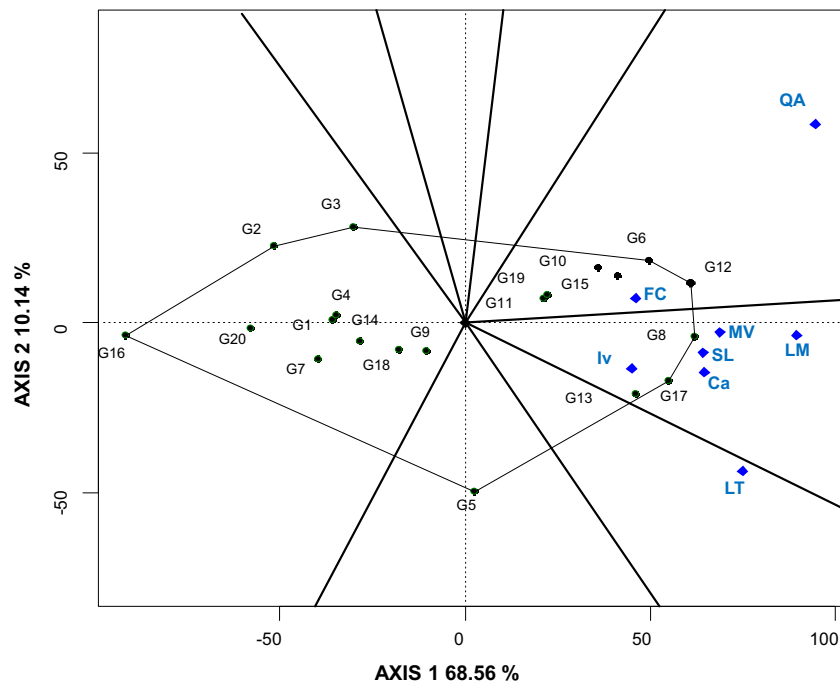


**Fig. 2** AMMI biplot for Pol% cane yield of the first two GE interaction principal components axes of 20 genotypes and 8 environments. It can be observed in the biplot the low values of the both axes, indicating that all genotypes are represented close to the origin; this reveals stability for all genotypes, indicating that the environment does not affect majorly this yield variable. This was also corroborated by the relatively low value of the explained variation percentage in the analysis of variance (Table 2). Principal

components explained 54.1 % of GEI interaction. G1 V91-1; G2 V91-2; G3 V91-6; G4 V91-8; G5 V91-15; G6 V98-62; G7 V98-86; G8 V98-120; G9 V99-117; G10 V99-190; G11 V99-203; G12 V99-208; G13 V99-213; G14 V99-217; G15 V99-236; G16 V99-245; G17 V00-50; G18 B80-408; G19 C323-68; G20 CP74-2005. QA Quebrada Arriba, MV Montaña Verde, LM Las Majaguas, Iv Ivonne, Ca Castellera, SL Santa Lucía, FC FUNDACAÑA, LT Los Tamarindos

and G17 scattered close to the origin, indicating minimal influence of these genotypes with environments. The remaining 15 genotypes scattered away from the origin in the biplot, indicating that the genotypes were more sensitive to environmental interactive forces. Meanwhile,

distribution of genotype points in the AMMI2 biplot for Pol% cane (Fig. 2) revealed that all genotypes are close to the origin (note the scales of both axis), indicating that all are stable in terms of sucrose production, but not the environment.



**Fig. 3** Site regression GGE biplot based on symmetrical scaling for the which-won-where pattern for TCH yield. In *black* are shown the genotypes and in *blue* are shown the studied environments. Biplot based on a “Tester-centered (G + GE)” table, without any scaling and dual metric preserving. The perpendicular lines to each side of the polygon divide the biplot into several sectors to allow visualization. Principal components explained 78.7 % of TCH variation. Note the

formation of two megaenvironments. *G1* V91-1; *G2* V91-2; *G3* V91-6; *G4* V91-8; *G5* V91-15; *G6* V98-62; *G7* V98-86; *G8* V98-120; *G9* V99-117; *G10* V99-190; *G11* V99-203; *G12* V99-208; *G13* V99-213; *G14* V99-217; *G15* V99-236; *G16* V99-245; *G17* V00-50; *G18* B80-408; *G19* C323-68; *G20* CP74-2005. *QA* Quebrada Arriba, *MV* Montaña Verde, *LM* Las Majaguas, *Iv* Ivonne, *Ca* Castellera, *SL* Santa Lucía, *FC* FUNDACAÑA, *LT* Los Tamarindos (color figure online)

### Megaenvironment and Genotype Representativeness Determination

As reported previously (Yan et al. 2007), the best methodology for megaenvironment and genotype representativeness determination is GGE biplots. Figures 3 and 4 show the “which-won-where” SREG biplots of TCH and Pol% cane, respectively, in which it can be observed what genotype performed better in what location for the yield variable measured. The angle between the vectors of two environments is related to the correlation coefficient between them (Kempton 1984; Yan 2002). The distance between two environments (locations) measured by the cosine of the angle between the vectors indicates their similarity or dissimilarity in discriminating the genotypes (Yan and Tinker 2006). For the case of TCH (Fig. 3), the eight locations can be grouped into two megaenvironments, one conformed by QA and FC, in which G11, G19, G10, G15, and in major extent G6 and G12 are the best genotypes. The other megaenvironment is formed by MV, LM, SL, Iv, Ca and LT, in which G8, G17 and G13 are the best genotypes. Due to the establishment of two megaenvironments and based on the length of the QA and LT vectors, these would be the ideal environments for selecting and producing genotypes

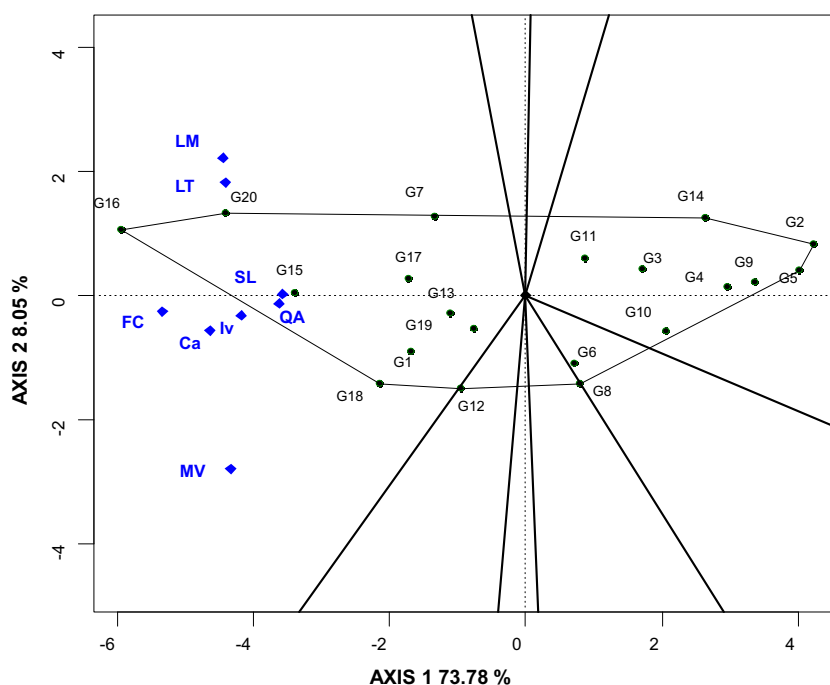
adapted specifically to both megaenvironments. LM is the most representative environment, and G8 and G12 are the best genotypes in terms of TCH yield.

In the case of Pol% cane (Fig. 4), just one megaenvironment formation occurs. Based on the length of the vectors, it can be said that FC is the most representative of all locations tested and G18, G15, G20 and G16 are the best adapted genotypes to this environment.

### Selection of Best Genotypes in Terms of Both Yield Variables

As can be noticed, the best genotypes for TCH (Fig. 3) turned out to be the worst for Pol% cane (Fig. 4), situation that results challenging for the selection of genotypes with positive relationship between both yield variables. To select the best genotypes, a coinertia analysis was performed (Dolédec and Chessel 1994; Dray et al. 2003b).

Coinertia analysis is a multivariate method often neglected, because it is a bit difficult to interpret, but is a very powerful methodology based on a different strategy of data analysis in which, instead of creating a single table with all variables, they are grouped according to a particular characteristic and are treated as two



**Fig. 4** Site regression GGE biplot based on symmetrical scaling for the which-won-where pattern for Pol% cane yield. In black are shown the genotypes and in blue are shown the studied environments. Biplot based on a “Tester-centered (G + GE)” table, without any scaling and dual metric preserving. The perpendicular lines to each side of the polygon divide the biplot into several sectors to allow visualization. Principal components explained 81.83 % of Pol% cane variation. Note that despite the environment variability, there is no

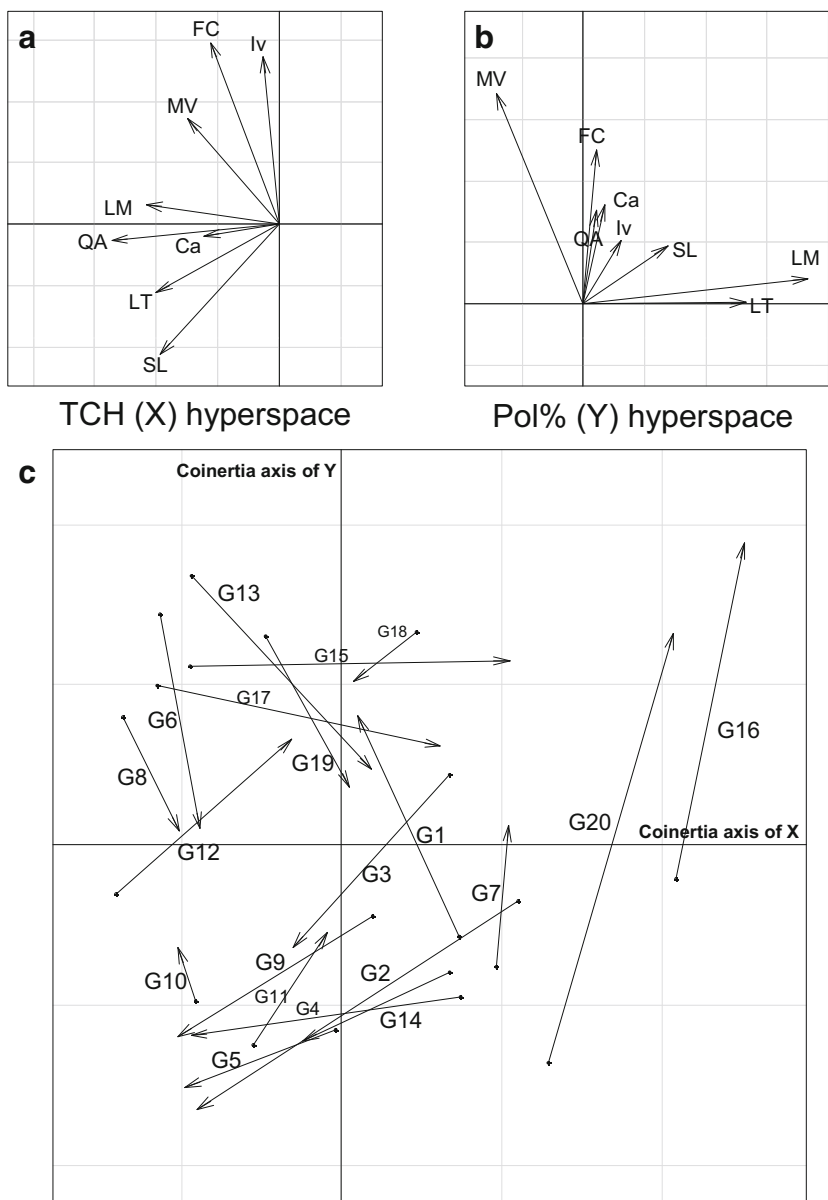
megaenvironments formation. *G1* V91-1; *G2* V91-2; *G3* V91-6; *G4* V91-8; *G5* V91-15; *G6* V98-62; *G7* V98-86; *G8* V98-120; *G9* V99-117; *G10* V99-190; *G11* V99-203; *G12* V99-208; *G13* V99-213; *G14* V99-217; *G15* V99-236; *G16* V99-245; *G17* V00-50; *G18* B80-408; *G19* C323-68; *G20* CP74-2005. *QA* Quebrada Arriba, *MV* Montaña Verde, *LM* Las Majaguas, *Iv* Ivonne, *Ca* Castellera, *SL* Santa Lucía, *FC* FUNDACANA, *LT* Los Tamarindos (color figure online)

hyperdimensions, which then are coupled so that the amount of information gathered by the axes of coinertia is relatively large, making it very efficient.

Coinertia analysis is very flexible and allows many possibilities for coupling, besides is suitable for quantitative and/or qualitative or fuzzy variables. Moreover, various weighting of sites and various transformations and/or centering of species data are available for this method. Hence, more biological considerations can be taken into account in the statistical procedures. Moreover, the principle of this method is very general and can be easily extended to the case of distance matrices (Dray et al. 2003b). To apply coinertia, two data matrices were generated, one for each yield variable, in order to perform a principal components analysis (PCA) for both data matrices. Separate analysis of each data table permits to create two hyperspaces, one for each yield variable. In each hyperspace, it can be determined an axis, which is the vector direction maximizing the projected variability or inertia. Both axis can be isolated and plotted in a multi-dimensional subspace so that the covariance between the two new sets of projected scores is maximal. This maximal covariance means a maximal correlation between both yield variables (Fig. 5). Coinertia analysis explained 80.09

and 12.48 % of the observed inertia in the TCH hyperspace ( $X$  matrix in  $x$  axis) and the Pol% cane hyperspace ( $Y$  matrix in  $y$  axis), respectively. The beginning of the arrows is the position of the genotypes described by the TCH data matrix, and the arrowhead is the position of the genotypes described by the Pol% cane data matrix. Despite the high quality of representation, measured by the high quantity of information gathered by the two axes, the maximal correlation between the behavior of both yield variables is very low; this is explained by the low value of the  $R_v$  Escoufier similarity coefficient 0.07817213, corroborating that the best genotypes for a variable result in the worst for the other one. Nevertheless, it can yet be inferred that the best genotypes with both yield variables positively correlated must be represented as an arrow toward the right from the upper left to the upper right quadrant. In this sense, the best genotypes for all environments are *G15*, *G17* and *G13* in that order. These results demonstrate not only the best adapted genotypes to all environments, but also the large capability of the proposed strategy for data analysis in this paper, which succeeds in finding the greatest differences between genotypes minimizing the differences between the environments in terms of both yield variables.

**Fig. 5** Coinertia analysis of the genotypes combining TCH and Pol% cane yield variables. **a**, **b** Scatterplots represent the coefficients of the combinations of the variables (environments) for each data matrices to define the coinertia axes. Separate analyses find axes maximizing inertia in each hyperspace. These axes of maximum inertia are projected in **(c)** scatterplot on which the genotypes are also projected. The beginning of the arrows is the position of the genotypes described by the TCH data matrix, and the arrowhead is the position of the genotypes described by the Pol% cane data matrix. The analysis explained 80.095 % in the TCH hyperspace and 12.478 % in the Pol% cane hyperspace of the observed inertia with a Rv Escoufier similarity coefficient of 0.07817213. *G1* V91-1; *G2* V91-2; *G3* V91-6; *G4* V91-8; *G5* V91-15; *G6* V98-62; *G7* V98-86; *G8* V98-120; *G9* V99-117; *G10* V99-190; *G11* V99-203; *G12* V99-208; *G13* V99-213; *G14* V99-217; *G15* V99-236; *G16* V99-245; *G17* V00-50; *G18* B80-408; *G19* C323-68; *G20* CP74-2005. *QA* Quebrada Arriba, *MV* Montaña Verde, *LM* Las Majaguas, *Iv* Ivonne, *Ca* Castellera, *SL* Santa Lucía, *FC* FUNDACAÑA, *LT* Los Tamarindos



**Selection of Best Environment for Both Yield Variables**

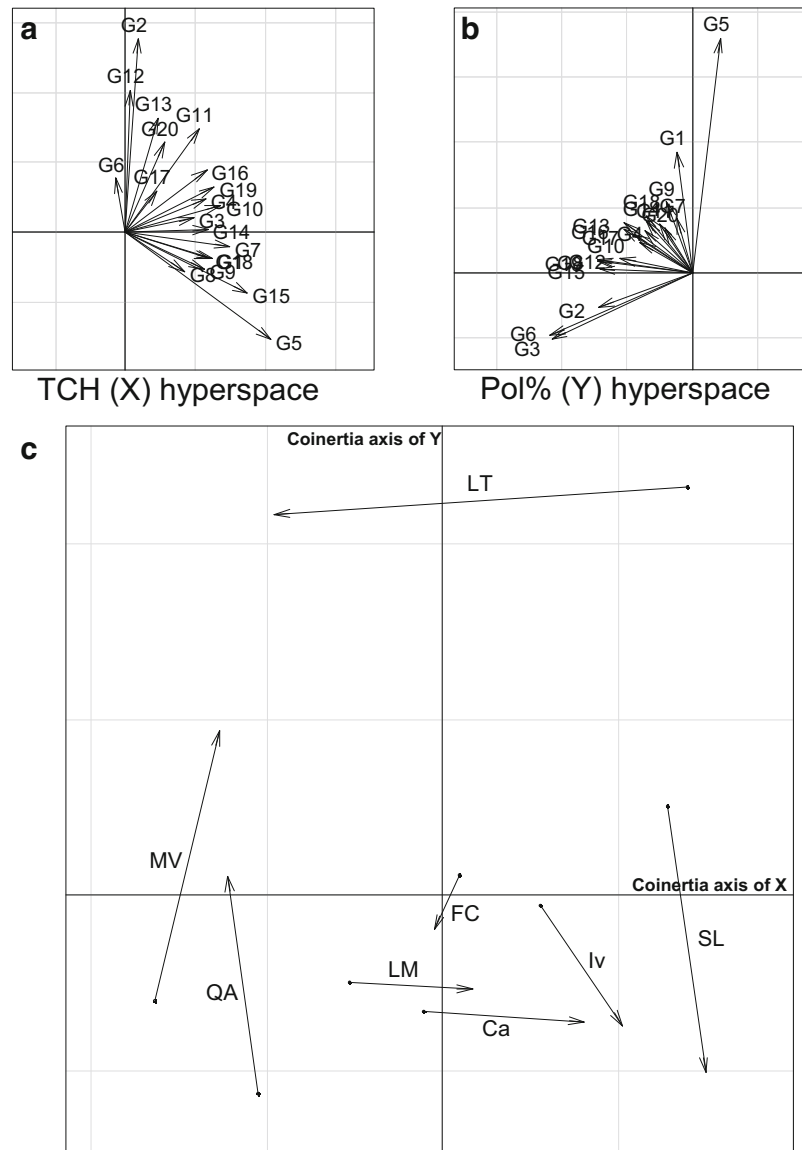
Another strategy of determination of the best suited genotype was to determine the environment with major positive correlation between both TCH and Pol% cane variables and then determining the genotypes most adapted to that environment.

In this case, another coinertia analysis was performed but with the transposed matrices of the two yield variables so the hyperspaces were constructed by the genotypes information to generate a subspace analysis describing the environments behavior (Fig. 6). Coinertia analysis explained 81.482 % in the TCH hyperspace and 10.617 % in the Pol% cane hyperspace, respectively. Also in this

case, the maximal correlation between the behaviors of both yield variables is low, Rv Escoufier similarity coefficient of 0.2203166, but taking into account the arrows interpretation explained above, the better environment must be represented as an arrow toward left from the upper right to the upper left quadrant. In this sense, the best environment, with both yield variables correlated positively, is Los Tamarindos (LT).

Once it is determined that Los Tamarindos (LT) was the best environment for growth (TCH) and production (Pol% cane) of sugarcane, it was decided to establish the best suited genotypes for that particular environment. A covariance analysis per genotype between the yield variables was performed. It can be seen that in LT all genotypes showed positive numbers, which means that in this





**Fig. 6** Coinertia analysis of the environments combining TCH and Pol% cane yield variables. **a, b** Scatterplots represent the coefficients of the combinations of the variables (genotypes) for each data matrices to define the coinertia axes. Separate analyses find axes maximizing inertia in each hyperspace. These axes of maximum inertia are projected in **(c)** scatterplot on which the environments are also projected. The beginning of the arrows is the position of the environments described by the TCH data matrix, and the arrowhead is the position of the environments described by the Pol% cane data

matrix. The analysis explained 81.482 % in the TCH hyperspace and 10.617 % in the Pol% cane hyperspace of the observed inertia with a Rv Escoufier similarity coefficient of 0.2203166. *G1* V91-1; *G2* V91-2; *G3* V91-6; *G4* V91-8; *G5* V91-15; *G6* V98-62; *G7* V98-86; *G8* V98-120; *G9* V99-117; *G10* V99-190; *G11* V99-203; *G12* V99-208; *G13* V99-213; *G14* V99-217; *G15* V99-236; *G16* V99-245; *G17* V00-50; *G18* B80-408; *G19* C323-68; *G20* CP74-2005. *QA* Quebrada Arriba, *MV* Montaña Verde, *LM* Las Majaguas, *Iv* Ivonne, *Ca* Castellera, *SL* Santa Lucía, *FC* FUNDACAÑA, *LT* Los Tamarindos

environment all genotypes exhibited results above the mean of all environments to both, TCH and Pol% cane. The more positive the result, the more efficient the genotype for yield in TCH and Pol% cane in this environment, so *G11*, *G13* and *G12* could be considered the better genotypes for successive selection processes throughout adaptation and stability (Table 5).

As demonstrated, this data analysis strategy is robust and permits to determine the best genotype for a particular environment or vice versa, depending of the breeding program needs. Nevertheless, for studies concerning national production, the genotypes selection process must be considering the best adapted for the megaenvironments determined.

**Table 5** Covariance analysis between both yield variables to determine best adapted genotype to “Los Tamarindos (LT)” environment

Genotype	LT
V99-203 (G11)	63.9600
V99-213 (G13)	45.2821
V99-208 (G12)	43.1818
V99-217 (G14)	42.3762
CP74-2005 (G20)	39.8317
V91-8 (G4)	38.8670
V99-245 (G16)	38.7655
V99-190 (G10)	38.2206
V98-86 (G7)	37.7743
V91-15 (G5)	36.4576
C323-68 (G19)	29.9106
V00-50 (G17)	29.7290
V91-1 (G1)	22.6755
V99-236 (G15)	22.3355
V99-117 (G9)	18.1341
V91-2 (G2)	16.8809
V91-6 (G3)	15.5293
B80-408 (G18)	10.8373
V98-62 (G6)	2.8020
V98-120 (G8)	2.3883

Certainly, the environment (when GEI is present) is one of the most influential factors affecting genotype selection. Although a hurdle, it should not be seen as a problem but as a challenge, grouping the set of variables as any other. It is just necessary to apply the adequate tools to solve and to understand the situation. Here, we demonstrate that besides AMMI models and GGE biplots, it is very necessary to add another statistical methodology like two-table coupling methods (coinertia) that expose a lot more information useful for breeders in this field of genotype  $\times$  environment interaction studies. AMMI and GGE models permit to determine stability characteristics, megaenvironments and determining the best adapted genotype to some environment, but if coinertia analysis is added, more thorough conclusions can be obtained, specially because it is necessary to select genotypes with both yield variables positively correlated; in other words, it is useless to obtain genotypes with large quantities of TCH if the sucrose quantities are low or genotypes with high sucrose content but with a low growth rate. No doubt the use of the coinertia along with AMMI and GGE methodologies could enhance breeding programmes to obtain better productivity levels, always maintaining growth rates, to fulfill required needs.

**Acknowledgments** The authors would like to thank the Venezuela's National Institute for Agricultural Research (INIA) and the Venezuelan Endowment for Science, Technology and Innovation (FONACIT) for the financing of regional trials.

## References

- Akcura, Mevlüt, Seyfi Taner, and Yuksel Kaya. 2011. Evaluation of bread wheat genotypes under irrigated multi-environment conditions using GGE biplot analyses. *Agriculture* 98(1): 35–40.
- Chessel, Daniel, Anne B. Dufour, and J. Jean Thioulouse. 2004. The ade4 package-I-one-table methods. *R News* 4: 5–10.
- Crossa, José, Paul L. Cornelius, and Weikai Yan. 2002. Biplots of linear-bilinear models for studying crossover genotype  $\times$  environment interaction. *Crop Science* 42: 619–633.
- Culhane, Aedín, Guy Perrière, and Desmond G. Higgins. 2003. Cross-platform comparison and visualisation of gene expression data using co-inertia analysis. *BMC Bioinformatics* 4(1): 59.
- De Mendiburu, Felipe. 2015. *Agricolae: Statistical procedures for agricultural research*. R package version 1.2-2. <http://CRAN.R-project.org/package=agricolae>.
- Dolédec, Sylvain, and Daniel Chessel. 1994. Co-inertia analysis: An alternative method for studying species-environment relationships. *Freshwater Biology* 31: 277–294.
- Dray, Stéphane, and Anne B. Dufour. 2007. The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software* 22(4): 1–20.
- Dray, Stéphane, Anne B. Dufour, and Daniel Chessel. 2007. The ade4 package-II: Two-table and K-table methods. *R News* 7(2): 47–52.
- Dray, Stéphane, Daniel Chessel, and Jean Thioulouse. 2003a. Procrustean co-inertia analysis for the linking of ecological tables. *Ecoscience* 10: 110–119.
- Dray, Stéphane, Daniel Chessel, and Jean Thioulouse. 2003b. Co-inertia analysis and the linking of ecological data tables. *Ecology* 84(11): 3078–3089.
- Frutos, Elisa, María P. Galindo, and Victor Leiva. 2014. An interactive biplot implementation in R for modeling genotype-by-environment interaction. *Stochastic Environmental Research and Risk Assessment* 28(7): 1629–1641.
- Ibáñez, Maximiliano, Mario Cavanagh, Natalia Bonamico, and Miguel Di Renzo. 2006. Análisis gráfico mediante biplot del comportamiento de híbridos de maíz. *Revista de Investigaciones Agropecuarias* 5(3): 83–93.
- Jalata, Zerihun. 2011. GGE-biplot analysis of multi-environment yield trials of barley (*Hordeum vulgare*. L.) genotypes in Southeastern Ethiopia Highlands. *International Journal of Plant Breeding and Genetics* 5(1): 59–75.
- Kempton, Robert. 1984. The use of biplots in interpreting variety by environment interactions. *The Journal of Agricultural Science* 103: 123–135.
- Mortazavian, Seyed, H. Nikkhal, F. Hassani, M. Sharif-al-Hosseini, M. Taheri, and M. Mahlooji. 2014. GGE biplot and AMMI analysis of yield performance of barley genotypes across different environments in Iran. *Journal of Agricultural Science and Technology* 16: 609–622.
- Ramburan, Sanesh, and Marvellous Zhou. 2011. Investigating sugarcane genotype  $\times$  environment interactions under rainfed conditions in South Africa using variance components and biplot analysis. *Proceedings of the South African Sugar Technology Association* 84: 245–362.
- Rao, Srinivasa, Sanjana Reddy, Abhishek Rathore, Belum V. Reddy, and Sanjeev Panwar. 2011. Application GGE biplot and AMMI model to evaluate sweet sorghum (*Sorghum bicolor*) hybrids for genotype  $\times$  environment interaction and seasonal adaptation. *Indian Journal of Agricultural Science* 81(5): 438–444.
- Rea, Ramón, and Orlando De Sousa-Vieira. 2002. Genotype  $\times$  environment interaction in sugarcane yield trials in the central-western region of Venezuela. *Interiencia* 27(11): 620–624.
- Rea, Ramón, Orlando De Sousa-Vieira, Miguel Ramón, Gleenys Alejos, Alida Díaz, and Rosaura Briceño. 2011. AMMI analysis

- and its application to sugarcane regional trials in Venezuela. *Sugar Tech* 13(2): 108–113.
- Rodríguez, Reynaldo, Yaquelin Puchades, Norge Bernal, Héctor J. Suárez, and Héctor García. 2012. Métodos estadísticos en el estudio de la interacción genotipo–ambiente en caña de azúcar. *Ciencia en su PC* 1: 47–60.
- Roostaei, Mozaffar, Reza Mohammadi, and Ahmed Amri. 2014. Rank correlation among different statistical models in ranking of winter wheat genotypes. *The Crop Journal* 2: 154–163.
- Silveira, Luís, Volmir Kist, Thiago O. Paula, Márcio H. Barbosa, Luiz A. Peternelli, and Edelclaiton Daros. 2013. AMMI analysis to evaluate the adaptability and phenotypic stability of sugarcane. *Scientia Agricola* 70(1): 27–32.
- Yan, Weikai, and Nicholas A. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science* 86: 623–664.
- Yan, Weikai, Manjit S. Kang, Baoluo Ma, Sheila Woods, and Paul L. Cornelius. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science* 47: 641–653.
- Yan, Weikai. 2002. Singular value partition for biplot analysis of multi-environment trial data. *Agronomy Journal* 94: 990–996.
- Yan, Weikai. 2011. GGE Biplot vs. AMMI graphs for genotype-by-environment data analysis. *Journal of the Indian Society of Agricultural Statistics* 65(5): 181–193.