Genotype-environment interactions, megaenvironments and winner genotypes and environments for soybean grain yield in Argentina

Julia Lúquez[1*], M Capurro[2] & Luis Erazzú[3]

[1]* Unidad Integrada Balcarce Ruta 226,km 73,5, Balcarce, Argentina, [2] Unidad Integrada Balcarce, CONICET, Ruta 226, km 73,5, Balcarce, Argentina, [3] INTA Famaillá, Tucumán, Argentina, *jluquez@balcarce.inta.gov.ar*

In Argentina, soybean [(*Glycine max* L.) Mer.] can be planted in a wide area. Currently, multienvironment trials (MET's) for yield performance for cultivars of different Maturity Groups are conducted in three different agroecological regions: North, North Pampean and South Pampean. Analysis and interpretation of MET's data related with genotype x environment interactions (GE) and selection of best genotypes have been made with analysis of variance and mean comparisons where GE interactions were not exploited. The objectives of this study were to identify megaenvironments and winner genotypes and environments using GGE biplot based on the site and genotype regression (SREG and GREG) models to exploit MET's data sets from soybean regional trials. The GGE biplots display graphically the relationship among test environments, genotypes and GE interactions. Grain yield data of 19 soybean cultivars of Maturity Group long IV from three seasons (2005, 2007 and 2008) across 27 environments in the three agroecological regions in Argentina were analyzed. The GGE biplots based on the SREG model showed that yield grain performance of soybean cultivars was determined by environments and GE interactions. Practically, three megaenvironments were determined (27 environments were grouped here) suggesting useful cultivar specific adaptations (Principal Components 1 and 2 explained 70.1% of variation). The GGE biplots based on the GREG model showed that two megaenvironments concentrated all cultivars on 13 locations with the best grain yields (Principal Components 1 and 2 explained 87.8% of variation) in contrast with the test locations typically used. The utilization of this information, could lead to improve soybean cultivars evaluation in Argentina.

KEYWORDS: *Glycine max*, GGE biplot, megaenvironments, GE interactions, grain yield

La soja [(*Glycine max* L.) Mer.] puede sembrarse en una amplia zona del país en Argentina. Los ensayos comparativos de rendimiento, normalmente se realizan en 3 zonas agroecológicas: Norte, Pampeana Norte y Pampeana Sur. El análisis y la interpretación de los datos que surgen de estos ensayos en relación con las interacciones genotipo x ambiente (IGA) y la selección de los mejores genotipos se realiza analizando las varianzas y comparando medias, y así, las IGA no son explotadas debidamente. El objetivo de este estudio fue identificar mega-ambientes (MA) y genotipos y ambientes superiores utilizando modelos de regresión por sitios (SREG) y por genotipos (GREG) que expresan la respuesta en función de G+GA y A+GA respectivamente y la exhiben a través de sus gráficos ("biplot"). Se utilizaron los rendimientos de grano de 19 cultivares de soja del Grupo de Madurez IV largo de las campañas agrícolas 2005, 2007 y 2008 provenientes de 27 ambientes de las 3 zonas agroecológicas del país. El GGA biplot basado en SREG mostró que el rendimiento de grano de los cultivares estuvo determinado por los ambientes y las IGA (los Componentes Principales 1 y 2 explicaron el 70,1% de la variación). Se delinearon prácticamente 3 MA, donde se agruparon los 27 ambientes probados, indicando valiosas adaptaciones específicas de los cultivares. El GGA biplot basado en GREG mostró, en tanto, que 2 MA concentraron a todos los cultivares evaluados en 13 ambientes con los mejores rendimientos de grano, en contraste con las típicas localidades de prueba usadas (los Componentes Principales 1 y 2 explicaron el 87,8% de la variación). El uso de esta aproximación ofrece un punto de vista que podría mejorar el uso de los recursos asignados para realizar ensayos de evaluación de cultivares de soja en Argentina.

PALABRAS CLAVE: *Glycine max,* GGE biplot, mega-ambientes, interacciones genotipo x ambiente, rendimiento de grano.

INTRODUCTION

Multiple environment trials are conducted annually throughout the world by seed companies and breeding institutions. The aim is to identify superior cultivars for a region. Another goal is to develop a comprehensive understanding of the region and to determine if part of the cropping area can be treated as a mega-environment (ME). Investigation about ME is a prerequisite for meaningful cultivar evaluation and recommendation (Yan & Hunt, 2002). CIMMYT defined ME as "a broad, not necessarily contiguous area, occurring in more than one country and frequently transcontinental, defined by similar biotic and abiotic stresses, cropping system requirements, consumer preferences, and, for convenience, by volume of production" (Braun *et al.*, 1996). Gauch & Zobel (1996; 1997) pointing out that the need for growing different cultivars in different regions is due to the existence of genotype x environment interaction (GE), defined a ME as a portion of a crop species growing region with a homogeneous environment that causes some genotypes to perform similarly. Yan et *al.*, in 2000, presented a "which wins where" methodology for identify ME.

In Argentina, soybean [(*Glycine max* L.) Mer.] can be cultivated in a extensive area, from 40° to 23° latitude. This allows a great diversity of sowing dates, and for this, crop rotations. In 2008-2009, 18 millions ha were planted, with an average yield of 3000 kg/ha. Soybean cropping area that is taken into account in this study is divided in three agroecological regions: North (I), North Pampean (II) and South Pampean (III). Each region is divided into a number of subregions. This division is based on geographical, climatic and even administrative factors rather than on Multiple Environmental Trials data (MET´s). Usually the data analyses have been made with statistical tools such as analysis of variance and mean comparisons, so GE interactions were not exploited. The objective of this study was to take advantage of MET's data sets for grain yield from soybean sowing regions, to identify ME and winner genotypes and best environments using GGE biplot based on the site and genotype regression (SREG and GREG) models. The GGE biplots displays graphically the relationship among test environments, genotypes and GE interactions (Yan *et al.*, 2000).

MATERIALS AND METHODS

Cultivars
Nineteen cultivars of Maturity Group long IV were tested. At each environment a randomized complete block design with three replications was used. Each plot was of seven rows of 5,5 metres at 21 cm apart. Harvest area per plot was 2 rows of 5,5 metres. Average grain yield (kg/ha) are in Table 1.

Environments
Regions and subregions of soybean sowing area are in Fig.1. Twenty seven environments were chosen, derived from different locations and different sowing dates in some locations (Table 2). All environments were tested in 2005, 2007 and 2008 seasons. For GGE methodology, average of grain yield for those three years was considered.

Table 1. *Grain yield (kg/ha) of cultivars of Maturity Group long IV tested*
Tabla 1. *Rendimiento de grano (kg/ha) de los cultivares de soja del GM IV largo   evaluados*

| Cultivar number	Cultivars	Released by	Grain yield(kg/ha)
1	ADM 4800	DonMario Semillas	3932
2	DM 4870	DonMario Semillas	4041
3	A 4910 RG	Nidera Semillas	3744
4	A 4613 RG	Nidera Semillas	4116
5	Tj 2049	La Tijereta- Seminium	3970
6	FN-485	Ferias del Norte	4074
7	A 4725 RG	Nidera Semillas	3762
8	DM 4600	DonMario Semillas	3824
9	Areco 4550	Areco	3897
10	Natalia 49	Relmó	3661
11	ACA 480 RG	Asociación de Cooperativas Argentinas	3632
12	ADM 5048	DonMario Semillas	3920
13	SPS 4500	SPS	3893
14	AS 4801	ASP	4107
15	DM 4970	DonMario Semillas	4369
16	NA 4990 RG	Nidera Semillas	4070
17	ALM 4930	DonMario Semillas	4265
18	ACA 4-60 RG	Asociación de Cooperativas Argentinas	3600
19	NA 4553 RG	Nidera Semillas	3830
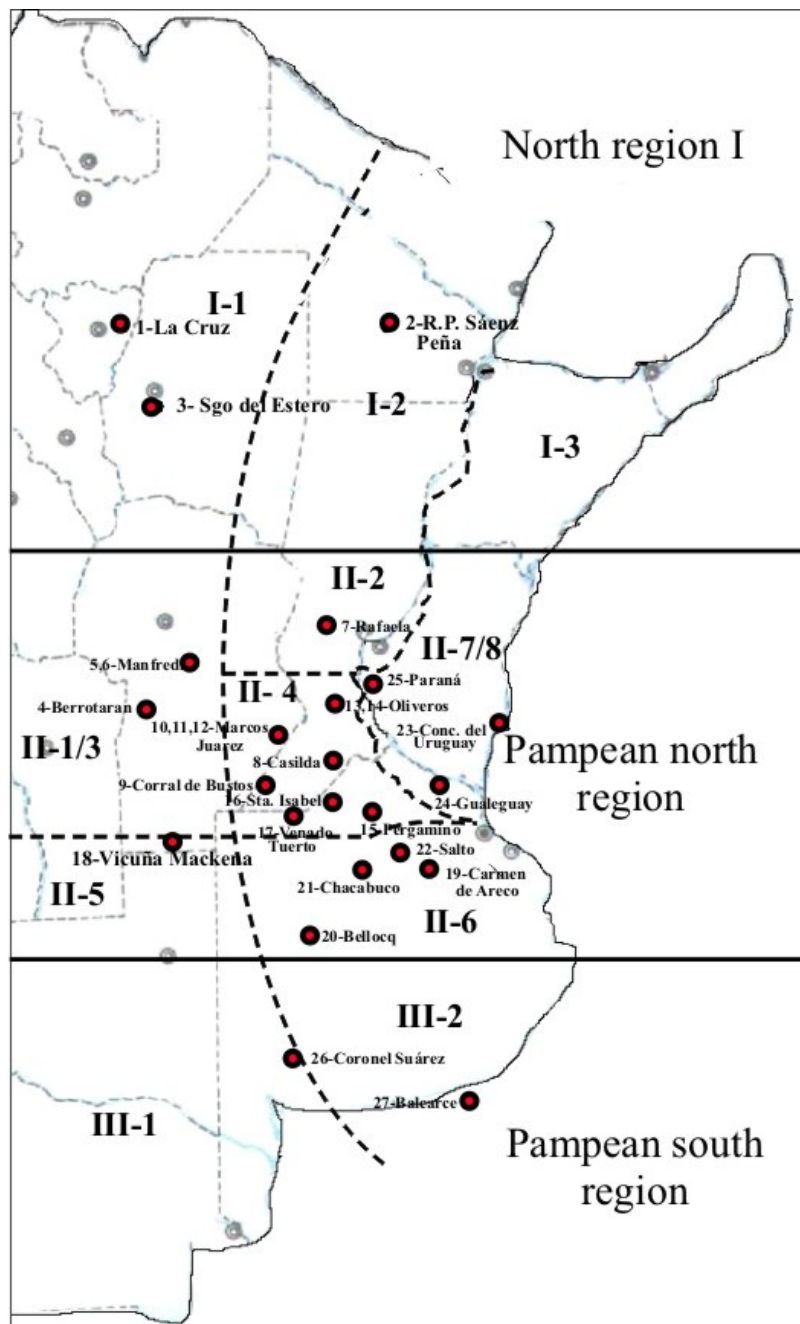
Figure1. *Regions and subregions of soybean sowing in Argentina*
Figura 1. *Regiones y subregiones de siembra de soja en Argentina*

Data analysis
Yield performance data were extracted from the information supplied annually by National Network of Evaluation of Soybean Cultivars (Red Nacional de Evaluación de Cultivares de Soja, RECSO) in Argentina.
Analysis of variance was conducted to investigate the relative contribution to total variability of cultivars (gen), year (year), environments (loc) and cultivar x environment interaction effects. To explore variability for grain yield and to consider the presence of ME, the GGE methodology proposed by Yan et al. (2000) was used.

Generated GGE biplot models were based on site regression (SREG) and genotype regression (GREG) using Info-Gen software (Balzarini & Di Rienzo, 2004). SREG removes environment effect and use the information of genotype (G) plus GE interaction. The objective is to identify mega-environments and winner genotypes for each of them. GREG removes genotype effect and response in function of environment (E) and GE interaction. The objective is to identify environmental contribution to GE interaction and the winner environments in a mega-environment for a particular group of genotypes.

Table 2. *Environments (locations and planting dates) tested*
Tabla 2. *Ambientes (localidades y fechas de siembra) evaluados*

Environments	Region	Subregion
1. La Cruz	I	1
2. Pres.R. Sáenz Peña	I	2
3. Sgo. del Estero	I	1
4. Berrotarán	II	1/3
5. Manfredi (p.d. 1)	II	1
6. Manfredi (p.d. 2)	II	1
7. Rafaela	II	2
8. Casilda	II	2
9. Corral de Bustos	II	4
10. M. Juárez (p.d. 1)	II	4
11. M. Juárez (p.d. 2)	II	4
12. M. Juárez (p.d. 3)	II	4
13. Oliveros (p.d. 1)	II	2
14. Oliveros (p.d. 2)	II	2
15. Pergamino	II	6
16. Sta. Isabel	II	4
17. V. Tuerto	II	4
18. V. Mackenna	II	5
19. C. de Areco	II	6
20. Bellocq	II	6
21. Chacabuco	II	6
22. Salto	II	6
23. C. del Uruguay	II	8
24. Gualeguay	II	8
25. Paraná	II	8
26. C. Suárez	III	1
27. Balcarce	III	2

Ref.: p.d.: planting date

GGE biplot models were constructed with the Principal Components (PC 1 and PC 2) through singular value decomposition (SVD) of centered data by environments effects for SREG and genotypes for GREG.

RESULTS AND DISCUSSION

Analysis of variance
Results showed significant differences (P<0.0001) among the levels of all sources of variation included in the analysis. The environments and GE interactions accounted for the largest sum of squares (31%), (year x loc: 42%, year x gen: 0.29%, loc x gen: 4.29%, year x loc x gen: 4.5%), and genotypes accounted 3.47%. These results are common in MET´s. Similar results were found by Setimela et *al.*(2007) in a study of evaluation of 35 open pollinated maize varieties in 59 locations in Africa.

GGE biplot based on SREG
The graphic generated (polygon) is divided in five sectors and can be observed in Fig.2. Cultivars situated in the extremes of the polygon represent those with extreme performance: the best or the worst grain yield in some environments. Cultivars 10, 8 and 12 (clock

wise) limit one sector of the polygon and perform better than cultivars 3, 7, 2, 5 and 1 in all the environments contained into this sector; cultivars 15 and 17 perform better than cultivars 6, 4, 16, 14, 13 and 9 in all environments contained into the sector of the polygon limited by them; cultivar 19 has the best grain yield at environment 2, while cultivars 18 and 11 don't perform good in any environment. All the environments into each sector of the polygon belong to one ME. So, GE interactions were more important than genotype component, because 3 ME were determined where genotypes with highest grain yield were identified. This fact indicates the possibility to exploit cultivar environment specific adaptations (PC 1 and PC 2 explained 70.1% of variation). Dardanelli et *al.* (2006) found 2 ó 3 ME for protein and oil + protein content in soybean based in MET´s data in Argentina using GGE biplot methodology.

GGE biplot based on GREG
The graphic generated has four sectors and can be observed in Fig.3. Environments situated in the extremes of the polygon (7, 25, 20, 22, 16, 2 and 12) represent those with extreme performance: the best or the worst grain yield for some cultivars. The sector of the polygon limited by the environments 25, 7 and 12 and that contains the environments 23,14, 24, 3, 5, 17 and 26, does not contains cultivars, suggesting that these are the worst environments for all cultivars evaluated and that contribute to GE interactions. The same situation occurs in the sector of the polygon with environment 2 in the extreme and environments 4, 11 and 13 inside it. Those two sectors of the polygon limited by environments 20 and 16, followed by environments 27,1,15, 22, 21,19, 14,10, 6, 8 and 9 (clock wise) contain all the cultivars evaluated, suggesting that they are the best environments for them ("winner environments"). So, the GGE biplot based on the GREG model showed that two ME concentrated all cultivars on 13 environments with the best grain yields (Principal Components 1 and 2 explained 87.8% of variation).

CONCLUSION

GGE biplot showed that three ME concentrated the 27 environments tested in this study. Grain yield was determined principally by environments and GE interactions offering opportunity to exploit cultivar specific adaptations. Thirteen out of 26 environments belonging 2 ME, concentrated all cultivars with the best grain yields. The existence of these ME displayed a different grouping of environments respect to traditional grouping, showing together, locations of different agroecological regions and viceversa. So, GGE biplot offered for us a new concept about exploitation of soybean MET's data related with importance of GE interactions, test environments and selection of best cultivars in Argentina.
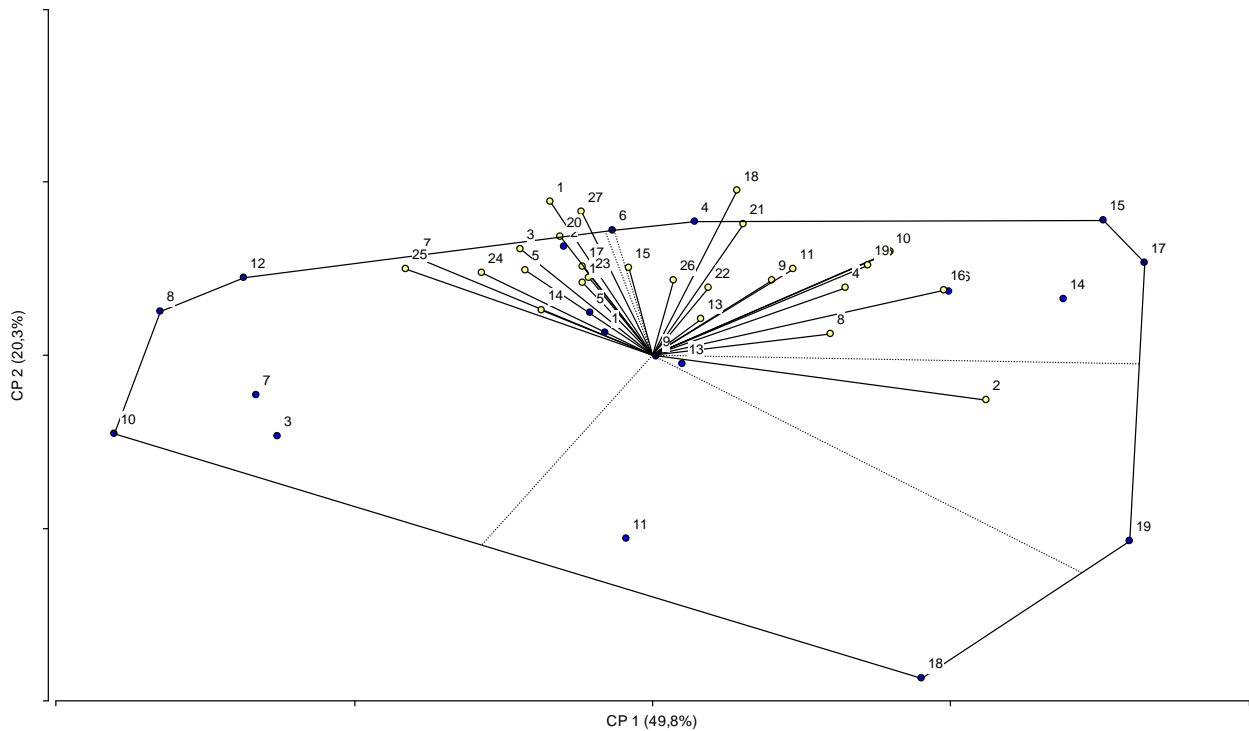
Figure 2. *GGE biplot based on SREG. Principal Components 1 and 2 explained 70.1% of variation. Environments are marked as vectors. Cultivars are inside and in vertices of the polygon.*
Figura 2. *Gráfico "biplot" Genotipo + Genotipo x Ambiente basado en el modelo de regresión sobre sitios. Los Componentes Principales 1 y 2 explicaron el 70,1% de la variación. Los ambientes son los vectores y los cultivares están en los vértices y dentro del polígono*
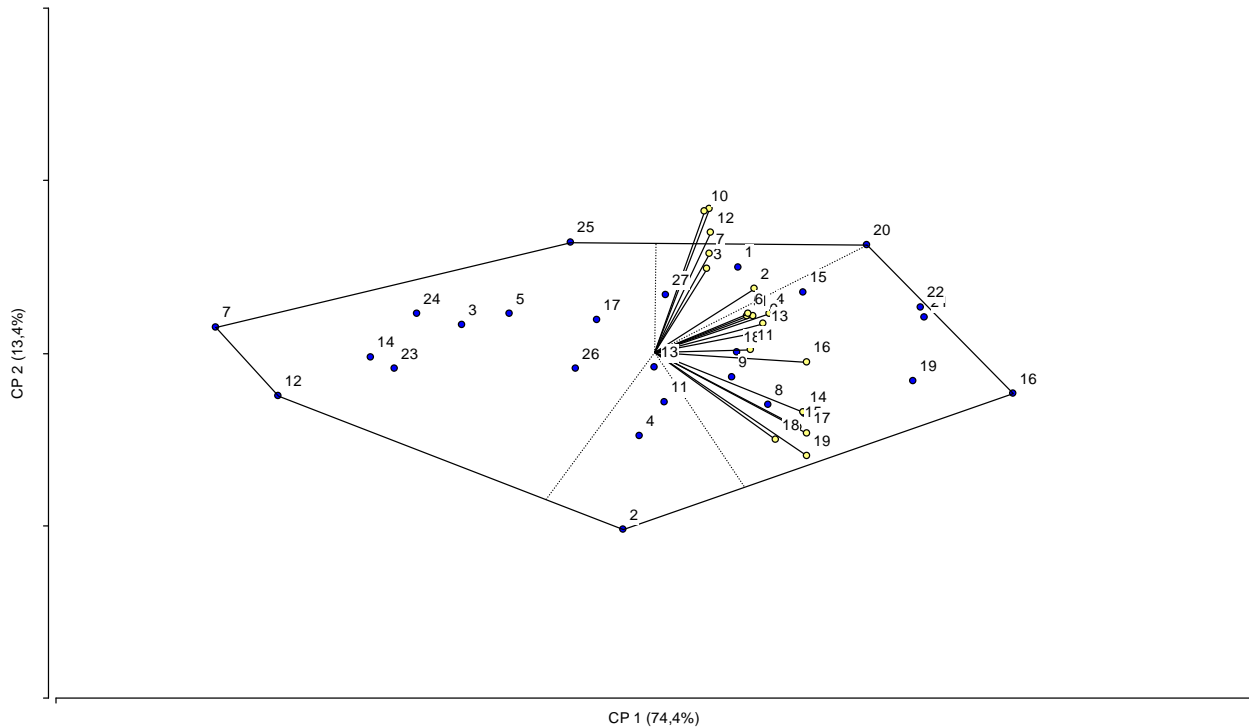

Figure 3. *GGE biplot based on GREG.Principal Components 1 and 2 explained 87.8% of variation. Cultivars are marked as vectors. Environments are inside and in vertices of the polygon.*
Figura 3. *Gráfico "biplot" Genotipo + Genotipo x Ambiente basado en el modelo de regresión sobre genotipos. Los Componentes Principales 1 y 2 explicaron el 87,8% de la variación. Los cultivares son los vectores y los ambientes están en los ángulos y dentro del polígono*

REFERENCES

Balzarini, M. & J. Di Rienzo. 2004. Info-Gen. Programa estadístico para datos de Mejoramiento y Genética de Poblaciones. http://www.info-gen.com.ar

Braun, H. J., S. Rajaram & M. Van Ginel. 1996. CIMMYT´s approach to breeding for wide adaptation. Euphytica 92: 175-183.

Dardanelli, J. L., M. Balzarini, M.J.Martínez, M. Cuniberti, S. Resnik, S. Ramunda, R. Herrero & H. Baigorri. **2006**. Soybean Maturity Groups, Environments and their Interaction Define Mega-environments for Seed Composition in Argentina. Crop Science 46:1939-1947.

Gauch, H. G. & R. W. Zobel. 1996. AMMI analysis of yield trials. In: Genotype –by-environment interaction. Kangrand MS & Gauch Ed. CRC Press,. Boca Raton, Fl. pp 1-40.

Gauch, H. G. & R. W. Zobel. 1997. Identifying mega-environments and targeting genotypes. Crop Science. 37: 311-326.

Setimela, P.S., B. Vivek, M. Bazinger, J. Crossa & F.Maideni. 2007. Evaluation of early to medium maturing open pollinated maize varieties in SADC region using GGE biplot based on the SREG model. Field Crop Research 103: 161-169.

Yan, W. & L. Hunt. 2002. Biplot analysis of diallel data. Crop Science 42: 21-30.

Yan, W., L. A. Hunt, Q. Sheng & Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Science 40: 597-605.