Australian Journal of Crop Science AJCS
AJCS 7(11):1772-1779 (2013) ISSN:1835-2707

AMMI and sequential path analyses of soybean [*Glycine max* (L.) Merrill] experimental lines in a breeding program in the Mexican tropics

Guillermo Ascencio-Luciano[1], Nicolas Maldonado-Moreno[1], Jesus Gerardo Garcia-Olivares[2] and Homar Rene Gill-Langarica[2*]

[1]Las Huastecas Experimental Station, INIFAP, Km 55 Highway Tampico-Mante, Villa Cuauhtemoc, Tamaulipas, Mexico
[2]Center for Genomic Biotechnology, Instituto Politecnico Nacional (The National Polytechnic Institute), Reynosa, Tamaulipas, Mexico

***Corresponding author: hgill@ipn.mx**

Abstract

The selection of soybean genetic material for crop production is complicated by genotype × environment interactions (G×E) during the development of breeding materials. The Additive Main Effects and Multiplicative Interactions (AMMI) model is the most commonly used model for modeling G×E and analyzing field data in plant breeding experiments. The objective of this study was to evaluate the significance and magnitude of the effect of G×E on soybean grain yield using a biplot based on AMMI analysis. The study also aimed to determine the relationship between total yield and various yield components for soybean using a sequential path analysis. For this purpose, data on grain yields and agronomic traits for 12 soybean genotypes over seven growing seasons (2004-2010) were collected. The first two components of the AMMI model explained 61.5% of the G×E. According to the biplot analysis, no single plant material had the highest yield across all environmental growing conditions. The genotype H98-1521 had the highest yields overall, in addition to a high average yield and less variable yields across different environments. The sequential path analysis demonstrated that soybean yield is significantly related to the following yield components: Weight of 100 seeds (100SEW), plant height at first pod (FPOH), and plant height at R7 (PLHR7).

Keywords: Genotype by environment interaction, soybean yield in Mexico.
Abbreviations: AMMI- Additive Main Effects and Multiplicative Interaction Analysis, ANOVA- Analysis of variance, E- Environment, FPOH- Plant height at first pod, G- Genotypes, G×E- Genotype by environment interaction, INIFAP- National Institute for Agricultural, Animal and Forestry Research, PCA- Principal component analysis, PLHR2- Plant height at R2, PLHR7- Plant height at R7, POPL- Number of pods per plant, YIELD-Yield, SNICS- National Service of Inspection and Certification of Seeds, SQ- Seed quality, VIF- Variance inflation factor, 100SEW- Weight of 100 seeds, R2 and R7- Reproductive Stages, where R2- Full Bloom, R7- Beginning Maturity.

Introduction

In Mexico, soybean [(*Glycine max* (L.) Merr.] is grown without irrigation in states with a subtropical climate, such as Campeche, Chiapas, San Luis Potosi, Sonora, Veracruz, and Tamaulipas. Soybean is in high demand by the food industry for both human and animal consumption. Domestic production is not sufficient to meet national demand; approximately 3.5 million tons/year are imported, amounting to 4.5% of the worldwide market (SoyStats, 2012). The steady increase in demand for soybean in Mexico presents a challenge to plant breeders, who must undertake the breeding and selection of soybean genotypes with high production potential and a range of adaptations to the agroclimatic conditions of the Mexican tropics. The soybean breeding program for the tropical region of Mexico is located at the Las Huastecas experimental station of the Instituto Nacional de Investigaciones Forestales, Agricolas, y Pecuarias (INIFAP) (National Institute for Agricultural, Animal and Forestry Research) in southern Tamaulipas. This breeding program began in the 1980s with an evaluation of plant materials imported from Brazil and Taiwan. Various breeding experiments have been implemented, including the first hybridizations of these plant materials. A soybean breeding program is currently underway that is aimed at obtaining genotypes that have the potential to produce a yield greater than 3 tons ha^{-1} and that are adapted to various production regions (particularly the states of Chiapas, Michoacan, Veracruz, and San Luis Potosi). The resulting materials will belong to maturity group IX, a group of pest-resistant (primarily resistant to leaf-eating pests) and disease-resistant cultivars that will have high protein and fat contents (Pathan et al., 2013). It is essential to analyze the materials produced in the breeding program for yield stability, total yield, and individual yield components because soybean production is known to be strongly influenced by environmental conditions. Variation in yield can occur in a given environment from year to year (E), among genotypes in a single year (G), and among genotypes and

environments (years) (G×E) (Pacheco et al., 2009). Assuming favorable breeding of soybean plants, yield will depend on the range of genetic variation, the influence of polygenic traits and other genetic factors that affect yield, and interaction with the environment (Rowntree et al., 2013). Selection based only on yield is ineffective; therefore, varieties must be selected based on a more diversified group of correlated traits (Sedghi and Amanpour, 2010). Therefore, information regarding correlated traits and an estimation of the direct and indirect effects of these traits on yield would increase the success of this soybean breeding program (Malik et al., 2007). Using the relationships among agro-morphological traits to increase grain yield is most effective during the selection of progenitors and progeny. It is necessary to evaluate the genetic potential of plant materials using methodologies that combine stability with productivity, as estimated by yield and yield components (Oz et al., 2009). For economically important crops such as maize (Alejos et al., 2006), soybean (Silva and Duarte, 2006), cotton (González et al., 2007), rice (Morais et al., 2008), chickpea, fava, and beans (Fikere et al., 2008; Fikere et al., 2009; Padilla et al., 2008; Pereira et al., 2009 and Tamkoc et al., 2009), different methodologies have been used to analyze the effects of E and G×E on yield and to compare yield stability among experimental lines and varieties. The most advantageous and widely used methodology used to evaluate the stability of experimental lines in the final experimental phase is the AMMI model (Additive Main Effects and Multiplicative Interaction Analysis), which includes both multiplicative interactions and additive main effects (Gauch, 2006; Gauch et al 2008). To evaluate the effect of different yield components on total yield, Wright (1921) implemented a path analysis methodology that separates the estimated correlations of traits from their direct and indirect effects on a single response variable (Dewey and Lu, 1959). This method is frequently used for crop breeding because it separates correlations between variables into causal effects, thus identifying chains of causality among variables. Previous studies (Whittaker et al., 2009; Gaikwad et al., 2007) have shown the importance of this type of analysis for developing selection strategies. A conventional pathway analysis could show a multicollinearity effect, particularly among correlated variables (Hair et al., 1995). Interpreting the true effect of each variable can be problematic because the effects are mixed because of collinearity. In an attempt to address this problem, Samonte et al. (2005) adopted the sequential path analysis method to determine the relationships between rice (*Oryza sativa* L.) yield and yield components by using an analysis of predictor variables and ordering the pathways into first-, second-, and third-order pathways. The sequential path model has an advantage over the conventional path model because it measures the true effects of the predictor variables and allows adjustments for different datasets (Mohammadi et al., 2003). Path and sequential path analysis coefficients have been used in soybean to identify selection criteria for total yield and yield components, such as agronomical traits and fat and protein contents (Board et al., 1997; Shukla et al., 1998; Ball et al., 2001; Malik et al., 2006a). These models have not been used in conjunction with the AMMI model for soybean, although they have been used together for other crops such as wheat and potato (Li et al., 2006; Hassanpanah and Azimi, 2010). Accordingly, the objectives of this study were to apply the AMMI model to calculate the effect of G, E, and G×E on soybean yields in southern Tamaulipas and to use path analysis and sequential path analysis to identify the relationships between total yield and agro-morphological traits.

Results and discussion

Variance and AMMI analysis

Grain yield (kg ha^{-1}) for the twelve soybean genotypes at seven different environments (years) was analyzed at the study site at the Las Huastecas experimental station (Table 2). This analysis demonstrated that grain yield was significantly affected ($p<0.01$) by E, G, and G×E. The large variation in soybean yield associated with E indicated that the values for E (agricultural year) generally had large differences from the mean, thereby accounting for the majority of the variation in yield. In the model, only 4.78% of the total sum of squares was attributed to the effect of G (Table 3). Grain yield attributed to G (average yield across environments) varied between 2,426.6 kg ha^{-1} (Huasteca 200) and 2,868.4 kg ha^{-1} (H98-1052) (Table 2). G×E had a significant impact on only 2.12% of the variation in grain yield in the model. The variety Huasteca-300 was the genotype associated with the highest yield in all four environments, whereas each of the experimental lines H98-1052, H98-1521, and H88-1880 was associated with high yield in only one environment. Huasteca-300 exhibited the highest yield (3,708.3 kg ha^{-1}) in 2004 (this year was associated with high yields trends overall) and the highest yield (2,851.3 kg ha^{-1}) in 2010 (a year exhibiting low yield trends overall) (Table 2). Figure 2 is AMMI biplot where genotypes and environments are depicted as points on a plane. The abscissa represents the main effects, and the ordinate represents the first multiplicative axis term (PC1). The horizontal line represents the interaction score of zero, and the vertical lines represent the grand mean yield. Displacement along the vertical axis indicates interaction differences between genotypes and between environments. Displacement along the horizontal axis indicates differences in the main effects of genotype and the environment. The solid line connecting the environment markers represents the year-to-year variation within an individual location. The genotypes with PC1 scores close to zero indicate general adaptation, whereas the higher scores indicate more specific adaptation to environments with PC1 scores of the same sign (Ebdon and Gauch, 2002a). Genotypes with a higher PC1 score, such as H98-1365, H98-1075, and Huas-100, were better adapted to the 2009 year, with a higher PC1 score of the same sign (Fig. 2). This combination results in a larger positive interaction. In contrast, genotype Huas-300 was adapted to years 2004, 2005, and 2007 with higher negative PC1 scores. The relative magnitude and direction of genotypes along the abscissa and ordinate axis in the biplot are important to understand the response pattern of the genotypes across environments. The best genotype should combine high yield and stable performance across a range of production environments. For example, the high-yielding (averaged over environments) genotypes H98-1052, H98-1076, H98-1521, Huas-300, and H88-1880 can be considered the most favorable based on their stability. H88-1880 combined a low absolute PC1 score and a high yield and may be the best genotype overall, with relatively less variable yield across environments. The years 2007, 2008, 2009, and 2010 had a relatively lower variation in the interaction (PC1) score, whereas 2004, 2005, and 2006 had the highest variation (Fig. 2). Therefore, the relative ranking of genotypes was stable between 2007 and 2010 compared to the ranking between 2004 and 2006. The years 2004, 2005, and 2006 combined larger main effects with lower interaction effects, making the study area a less predictable location for soybean variety evaluation. Huasteca 200 and 100 were the first plant materials developed by the breeding program at Las Huastecas and

Table 1. List of materials used in the study.

Maturity group	Genotype	Genetic Status	Progenitor ♀		Progenitor ♂
IX	Huas-100*	V	Santa Rosa	X	Jupiter
IX	Huas-200	V	F81-5344	X	Santa Rosa
IX	Huas-300	V	H82-1930	X	H80-2535
IX	Huas-400	V	Dois Marcos 301**		
IX	H88-1880	AEL	Santa Rosa	X	H80-2535
IX	H98-1052	AEL	H88-1880	X	H88-3868
IX	H98-1521	AEL	BR-15	X	H88-1880
IX	H98-1021	AEL	Padre X Santa Rosa	X	Santa Rosa
IX	H98-1075	AEL	H88-1880	X	H88-3868
IX	H98-1076	AEL	H88-1880	X	H88-3868
IX	H98-1092	AEL	H88-1880	X	H91-0235
IX	H98-1365	AEL	H88-3964	X	H88-0445

*Huas = Huasteca; **Selection of individual plants; V = Variety; (AEL) = Advanced experimental line.

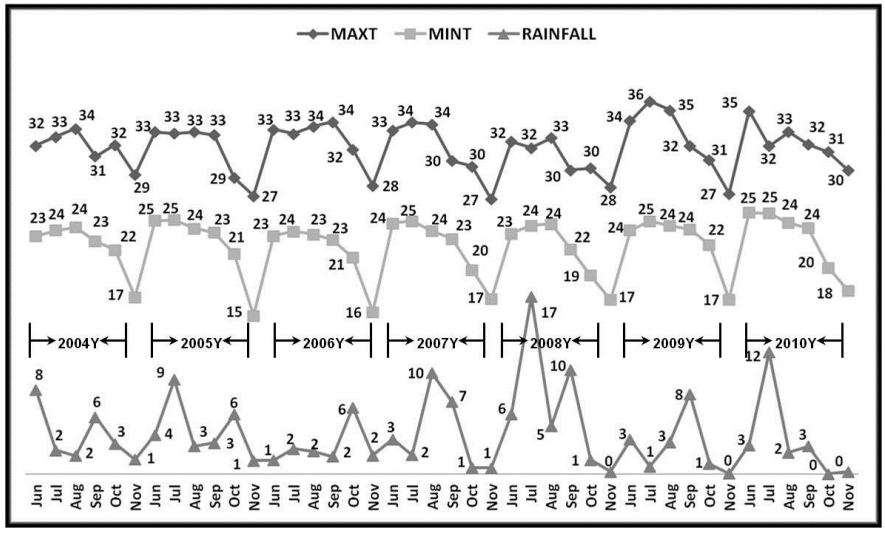

Fig 1. Monthly maximum (MAXT) and minimum (MINT) temperatures (°C) and monthly average rainfall (mm^3) during the experimental years.

the first progenitors in the development of new hybrids adapted to tropical climate and latitude; they belong to the maturity group IX. The line H88-1880 was a well-adapted genotype with the same progenitor (variety Santa Rosa) as Huasteca 100 and 200. The line H88-1880 has been certified by the Servicio Nacional de Inspección y Certificación de Semillas (SNICS) (National Service of Inspection and Certification of Seeds) and will soon be grown in southern Tamaulipas (Guillermo Ascencio Luciano, personal communication) (Table 1).

In Figure 3 cross-validated the interaction pattern of the 12 soybean genotypes with 7 environments. The distances from the origin (0, 0) are indicative of the amount of the interaction that was exhibited by genotypes over environments or environments over genotypes (Voltas *et al*., 2002). The genotypes Huasteca-300, Huasteca 400, H98-1092, H98-1365, H98-1052, and H98-1021 exhibited strong interactions (either positive or negative). The environments (years) 2004 and 2005 showed only a weak interaction. The additive behavior for 2004 shows that in this particular year, genotype yield in general was highly correlated with average genotypic traits across all environments. Genotypes that are connected and located on opposite sides of the axes are linked together by a polygon in the biplot graph. The lines connecting the data points for genotype and environment (year) form sectors perpendicular to the sides of the polygon (Hernández and Crossa, 2000). The genotypes near the apex of each sector, i.e., close to the origin (0, 0), are the best genotypes for the environments within that sector. The AMMI biplot (Fig. 3) shows five sectors, four of which include environments. This figure shows that three different environments (years 2005, 2008, and 2010) were grouped together in a single sector, which indicates that the genotypes exhibited similar development during these years. The year 2004 is closer to the origin, and thus, there was a weaker environment-genotype interaction in that year. This result suggests that environmental conditions during 2004 were not stable enough to produce high yields in these genotypes (Fig. 1).

Sequential path analysis

Based on the VIF and tolerance values, the plant height at R7 (PLHR7), the first pod height (FPOH), and the weight of 100 seeds (100SEW) were considered to be first-order traits. To identify the first-order variables corresponding to the traits above, the procedure was repeated using the variables PLHR7, FPOH, and 100SEW as dependent variables. These three variables were considered to be second-order variables for soybean yield. The direct effects on yield traits were calculated using the procedure described by Williams et al. (1990). A partial determination coefficient (analogous to R^2 in a linear regression analysis) was calculated from the path coefficient for all predictor variables. The path analysis was used to estimate the direct effects of the variables (agronomi

Table 2. Average grain yields (kg ha^{-1}) of 12 soybean genotypes in different environments (years) at the Las Huastecas experimental station in southern Tamaulipas, Mexico.

	Environment							
Genotypes*	2004	2005	2006	2007	2008	2009	2010	Average
H98-1052	3308.3	3279.1	2855.7	2925.7	2768.7	2504.8	2436.1	2868.4a
H98-1076	3274.9	3316.0	2610.0	3022.1	2591.5	2363.9	2510.6	2812.7b
H98-1521	3223.5	3365.1	1902.2	3294.8	2957.2	2387.4	2338.7	2781.3c
Huas-300	3708.3	3440.5	1396.1	2810.7	3107.1	2108.5	2851.3	2774.6c
H88-1880	3428.8	3390.3	1766.3	2871.0	2659.2	2527.4	2507.6	2735.8c
H98-1021	3234.5	3275.4	2215.5	2399.1	2768.8	2444.1	2253.1	2655.8d
H98-1092	3543.8	3161.9	1373.5	3132.1	2476.1	2405.1	2259.1	2621.6e
H98-1075	3408.6	2784.7	2023.5	2822.1	2484.6	2437.8	2384.6	2620.8e
Huas-100	3080.4	3100.6	1732.7	3090.6	2442.7	2384.0	2282.6	2587.7e
Huas-400	3107.0	3136.2	1091.2	3056.7	2508.4	2103.4	2125.2	2446.9e
H98-1365	3124.4	2703.4	2060.1	2402.4	2384.1	2461.6	1930.4	2438.1e
Huas-200	2697.0	2899.3	1693.0	2569.6	2501.2	2258.5	2367.4	2426.6e
Average	3261.6a	3154.4a	1893.3e	2866.4b	2637.5c	2365.5d	2353.9d	2647.5

*H = experimental line; Huas = variety developed and grown in southern Tamaulipas. Average values (means) with the same letter are not significantly different ($p \leq 0.05$) based on a Duncan mean comparison analysis. Underlined values are the highest values for each environment.

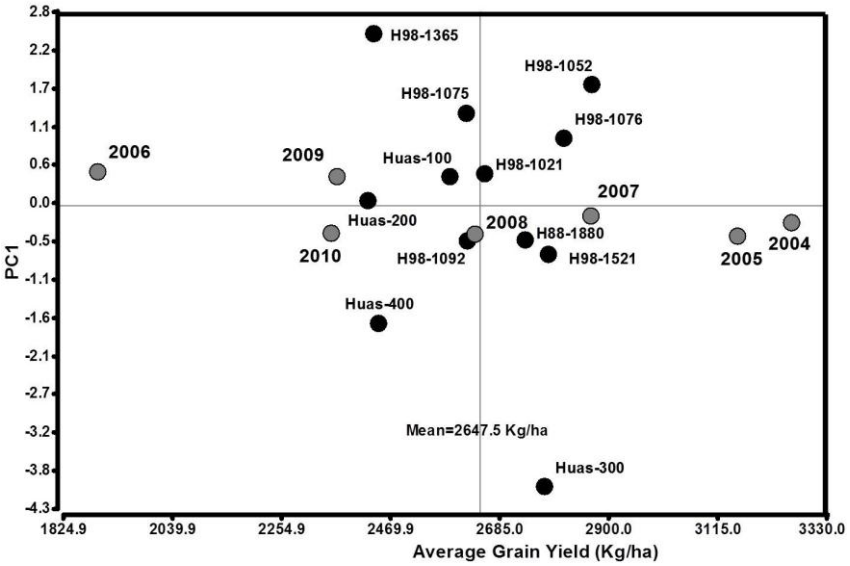

Fig 2. Graphical analysis of PC1 for 12 soybean genotypes at seven growing sites in southern Tamaulipas, Mexico. Principal Component 1 (PC1) was obtained using mean yield values. H = experimental line; Huas = variety developed and grown in southern Tamaulipas.

cal traits) with YIELD as the response variable. The multicollinearity analysis showed moderate collinearity for all traits, even those that showed strong effects such as PLHR7 (VIF = 1.23) and FPOH (VIF = 1.05) (Table 4). The sequential path analysis using YIELD as the dependent variable and the remaining variables as the independent variables identified PLHR7, FPOH, and 100SEW as first-order variables that explained 26% of the variation in soybean grain yield (Table 5 and Fig. 4). The sequential path analysis of the second- and first-order variables demonstrated that 43% of the total variation in PLHR7 was explained by the following three traits: PLHR2, SQ, and POPL. This analysis also showed that 18% of the total variation in 100SEW was explained by PLHR2 and POPL (Table 5). Not all variables had significant direct effects on soybean grain yield. Of the first-order variables that had direct effects on grain yield, PLHR7 and 100SEW had direct positive effects and FPOH had a direct negative effect on yield. The results of this study demonstrate the differential behavior of genotypes during different tests. Differences in yield were explained by significant differences ($p \leq 0.01$) in E, G, and G×E. The differential response of soybean genotypes to their environments could reflect the influence of varying climate conditions during the seven years of the experiment (Fig. 2). Our results corroborate findings for other economically important crops such as chickpea (*Cicer arietinum*), fava beans (*Vicia faba*), and pinto beans (*Phaseolus vulgaris* L.) (Fikere et al., 2008; Fikere et al., 2009; García et al., 2012; Pereira et al., 2009; Tamkoc et al., 2009). In the present study, the PC1 graphs indicate that the experimental lines performed better than the established soybean varieties and exhibited high yield and reliability in favorable environments; these results were consistent with those of Karasu et al. (2009). Some reports for other crops have used both AMMI models and path coefficient analyses to evaluate the genotype response to the environment and identify correlations among yield components. These studies demonstrated that these models are good statistical tools that aid the selection of soybean genotypes and the study of their performance (Tai et al., 1994; Li et al., 2006; Tai and Tarn, 2003). In this study, the use of an AMMI

Table 3. Variance analysis of the yields of 12 soybean genotypes in different environments (years) at the Las Huastecas experimental station in southern Tamaulipas, Mexico.

Variation source	d.f.	Mean Square
Repetition	2	0.22389236 NS
Environment	6	8.50910805 **
Genotype	11	0.48193936 **
Interaction (GxE)	66	0.21348101 **
AMMI 1	16	0.3685000**
AMMI 2	14	0.1170100 *
AMMI 3	12	0.08173 NS
Model	85	0.83565601 **
Error	166	0.10076134
CV (%)	12 %	

d.f. = degrees of freedom; CV = coefficient of variation; NS = not significant ($p \leq 0.05$); **significant ($p \leq 0.05$).

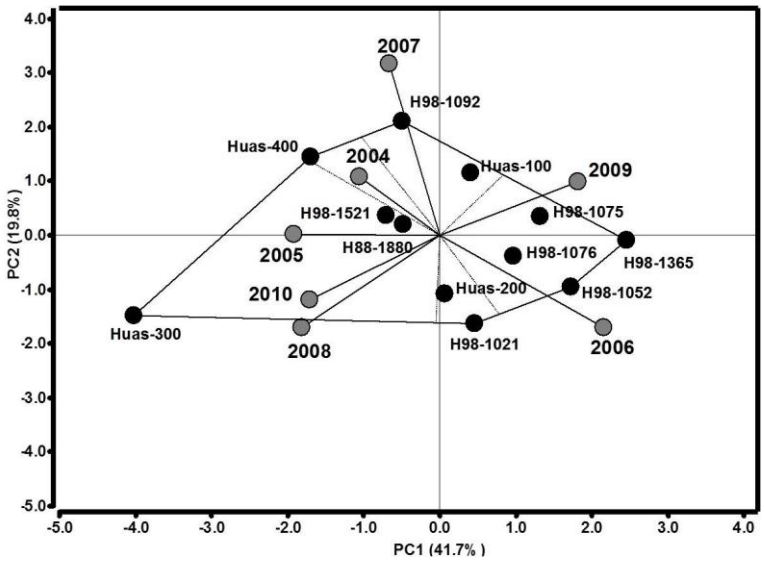

Fig 3. Graphic analysis of PC1 and PC2 for 12 soybean genotypes at 7 growing sites in southern Tamaulipas, Mexico. H = experimental line; Huas = variety developed and grown in southern Tamaulipas.

model combined with a sequential path analysis permitted the identification of genotypes with relatively high stability and medium to high yields (H88-1880 and H98-1521) (based on PC1 and PC2). The sequential path analysis also indicated that the yield components PLHR7, FPOH, and 100SEW had the highest contribution to soybean crop yields.

Materials and methods

Plant materials

Twelve genotypes were evaluated within the framework of the breeding program in the Mexican tropics implemented by INIFAP. Eight of the 12 soybean genotypes are experimental lines (H88-1880, H98-1021, H98-1028, H98-1052, H98-1075, H98-1076, H98-1092, H98-1365, and H98-1521), and four cultivars are commercially available and commonly used in the region (Huasteca 100, Huasteca 200, Huasteca 300, and Huasteca 400). The genotypes H88 and H98 are advanced experimental lines that are characterized by medium size and determinate growth and were developed from various individual crosses. The Huasteca 100 variety was developed from a Santa Rosa x Jupiter hybridization. Both progenitors are adapted to tropical regions and therefore are classified into maturity group IX. The varieties Santa Rosa and Jupiter were brought from Brazil and the United States, respectively (Maldonado, 1994). The Huasteca 200 variety, which is tall-sized with semi-determinate growth, was grown for the first time at the Las Huastecas experimental station (Maldonado and Ascencio, 2010b) and is a product of a F81-5344 x Santa Rosa cross. The progenitor F81-5344 is a variety developed by Dr. Kuell Hinson of the U.S. Department of Agriculture in Gainesville, FL. The characteristics of this variety include low sensitivity to the photoperiods of tropical latitudes, full size with semi-determinate growth, strong attachment of the seed to the pod, resistance to leaf disease, and a long vegetative cycle. The variety Huasteca 300 was developed from a cross between H82-1930 and H80-2535. Variety H82-1930 is the product of a Jupiter x F76-9835 cross. The H80-2535 variety is the product of a Jupiter x Iowa cross. The variety Huasteca 400 was developed at the Las Huastecas experimental station using material from the Brazilian genetic variety Dois Marcos 301(DM301), which was introduced in Mexico in 1998. Plants developed from these material exhibit valuable agronomic traits such as plant height, strong attachment of the seed to the pod, and resistance to plant disease, lodging, and threshing (Table 1). The study was conducted at the Las Huastecas experimental station, which is located at 15 m a.s.l. in Estacion Cuauhtemoc, Altamira, Tamaulipas, Mexico at 18° 50' N, 96° 10' W. The experimental plots were established in the spring-summer production cycle during the rainy season between 2004 and 2010 (Fig. 1) in a field with slightly alkaline vertisol soil. The climate in this region is classified as warm sub-humid according to Garcia (1987). The mean annual temperature is 25.4 °C, and the maximum and minimum temperatures are 42.5 °C and 7.0 °C, respectively. The mean annual rainfall is 800 mm. By the end of May, the

Table 4. Direct effects of yield predictor variables and measures of collinearity (Model 1) for 12 soybean genotypes grown in southern Tamaulipas.

Predictor variable	Direct effect	Tolerance	VIF	p-value
PLHR2	-0.06	1.01	1.00	NS
PLHR7	0.57	0.81	1.23	***
FPOH	-0.36	0.96	1.05	**
100SEW	0.30	0.96	1.04	**
SQ	0.11	0.99	1.01	NS
POPL	-0.02	1.00	1.00	NS

NS = not significant ($p \leq 0.05$); **significant ($p \leq 0.05$); *** highly significant ($p \leq 0.05$).

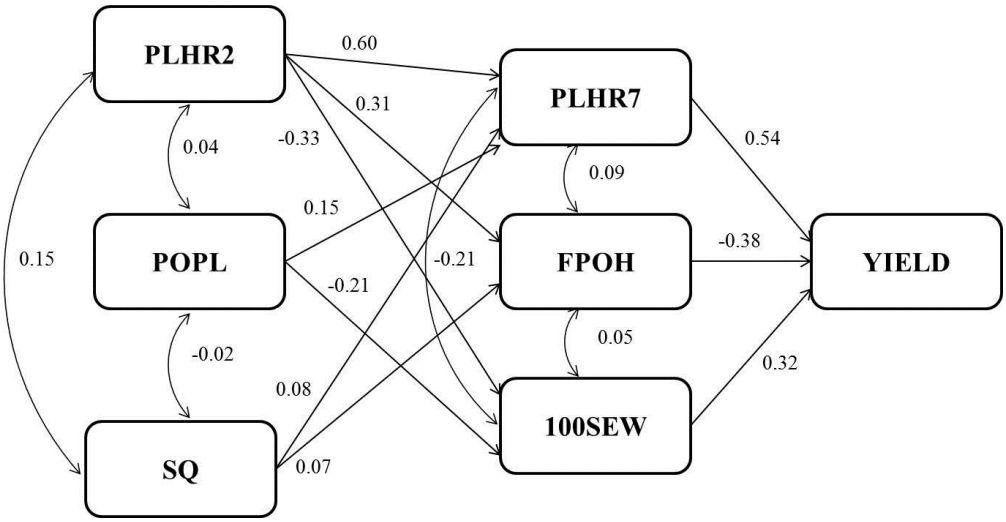

Fig 4. Sequential path analysis showing the inter-relationships of traits contributing to grain yield for 12 genotypes of soybean grown in southern Tamaulipas. FPOH- Plant height at first pod, PLHR2- Plant height at R2, PLHR7- Plant height at R7, POPL- Number of pods per plant, YIELD-Yield, SQ- Seed quality, VIF- Variance inflation factor, 100SEW- Weight of 100 seeds, R2 and R7- Reproductive Stages, where R2- Full Bloom, R7- Beginning Maturity

Table 5. Direct effects of the first- and second-order predictor variables on grain yield and collinearity (Model 2) for 12 genotypes of soybean grown in southern Tamaulipas.

Response variable	Predictor Variable	Adjusted R^2	Direct Effect	Tolerance	VIF
YIELD	PLHR7	0.26	0.54	0.82	1.22
	FPOH		-0.38	0.95	1.05
	100SEW		0.32	0.96	1.05
PLHR7	PLHR2	0.43	0.60	0.62	1.62
	CAL		0.08	0.98	1.01
	POPL		0.15	0.96	1.04
FPOH	PLHR2	0.08	0.31	0.91	1.10
	POPL		-0.09	0.99	1.00
100SEW	PLHR2	0.18	-0.33	0.87	1.14
	POPL		-0.21	0.94	1.06

FPOH- Plant height at first pod, PLHR2- Plant height at R2, PLHR7- Plant height at R7, POPL- Number of pods per plant, YIELD-Yield, SQ- Seed quality, VIF- Variance inflation factor, 100SEW- Weight of 100 seeds, R2 and R7- Reproductive Stages, where R2- Full Bloom, R7- Beginning Maturity.

fields were weeded, and the soil was plowed to 30 cm. This was followed by raking and row digging in separate sessions. The experimental design used random blocks with four repetitions, arranged in rows to facilitate cultivation practices. The distance between rows was 75 cm, and the plant density was 300,000 plants per ha. For the purposes of this study, the following data were collected for agronomic, vegetative, and reproductive traits: YIELD (the grain yield), plant height at R2 (PLHR2), plant height at R7 (PLHR7), plant height at first pod (FPOH), weight of 100 seeds (100SEW), seed quality (SQ), and number of pods per plant (POPL).

Variance and AMMI analysis

The grain yield (YIELD) was first calculated for each genotype to determine the variation in yield caused by differences in G, E, and G×E. The analysis of G×E was performed using the AMMI model. The analysis of variance allowed the determination of the relative contributions of G, E, and G×E to total variance. SAS 6.03 (SAS Institute, 1998) software was used to run the AMMI model using an algorithm by Hernandez and Crossa (2000) and the biplot model (Burgueño et al. 2002). The two analyses were used to create biplots in which

"GE" refers to G×E. The AMMI model is represented by the following equation:

$$Y_{ij} = \mu + G_i + a_j + \sum_{i=1}^{k} \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij}$$

where Y_{ij} is the yield of the *i*-th genotype in the *j*-th environment and the additive parameters are as follows: μ = mean, G_i = effect of the *i*-th genotype, a_j = effect of the *j*-th environment, $\boldsymbol{\lambda_k}$ $(\boldsymbol{\lambda 1} \trianglerighteq \boldsymbol{\lambda 2} \trianglerighteq \ldots \trianglerighteq \boldsymbol{\lambda t})$ are scaling constants (singular values) that allow the imposition of orthonormality constraints on the singular vectors for genotypes, $\boldsymbol{\alpha ik} = \alpha_{1k}, \ldots, \alpha_{gk})$ and environments, $\boldsymbol{\gamma_{jk}} = (\gamma_{jk}, \ldots, \gamma_{ek})$, such that $\sum i\alpha 2ik = \sum j2\lambda_k = 1$ and E_{ij} = experimental error. The results of the AMMI model analysis were interpreted on the basis of two AMMI graphs: (a) the graph that showed the main and first multiplicative axis term (PC1) of both genotypes and environments, and (b) the biplot that used scores of environments and genotypes PC1 against scores of environments and genotypes of the second multiplicative axis term (PC2). Both AMMI biplots were constructed using the software InfoStat V2011p (Di Rienzo et al., 2008).

Sequential path analysis

A path analysis (Wright, 1921) was used for the correlation analysis of plant traits and to evaluate the relative contribution of each yield component to grain yield. Grain yield (YIELD) was the response variable, and the yield components were the independent variables. A sequential path analysis was used to predict the relationships between grain yield and the yield components and to eliminate variables with very low contributions to the model. First a sequential multiple regression analysis was used to organize the predictor variables into sequential order by genotype based on their contribution to the total variation in yield and minimal collinearity (Fig. 4). The sequential path model consisted of three component paths, each component with its respective predictor and response variables. The level of multicollinearity determined in each component path was measured according to the value of tolerance and the variance inflation factor (VIF) (Hair et al., 1995) (Table 4-5). The tolerance value ($1-R^2_i$) is the amount of variation in a selected independent variable that is not explained by the other independent variables. Here, R^2_i is the determination coefficient for the prediction of the *i*-th variable by the predictor variables. The VIF is determined by the magnitude of the effect of the other independent variables on the variance of the selected independent variable [$VIF = 1 / (1-R^2_i)$]. If the response variables are not redundant, the VIF value approximates 1. A low tolerance value (< 0.1) or a high VIF (> 10) indicates high multicollinearity of the variables (Hair et al., 1995). In such cases, the data were subsequently analyzed using the software InfoStat V2011p (Di Rienzo et al., 2008).

Conclusion

G×E is a common phenomenon during field testing of experimental crop lines and associated varieties. The occurrence of G×E increases the difficulty of selecting the best plant genetic materials. This study demonstrated for the first time the application of the AMMI model in field experiments for a soybean breeding program in Mexico. The model was effective for identifying G×E patterns and explaining grain yield data from multi-environment (multi-year) tests. The AMMI model provided an analysis of the relative magnitude and significance of G×E effects and the mutual effects of G and E. Based on seven years of evaluation and testing, this study demonstrated that G×E is an important source of variation in soybean yield in southern Tamaulipas, Mexico. The biplots generated by the AMMI model were used to effectively visualize the response pattern of the genotypes and the environments at the study site. The use of sequential path analysis as a predictive tool for the relationships between total yield and the six components of soybean yield showed that plant height at R7 (p = 0.54), plant height at first pod (p = -0.38), and the weight of 100 seeds (p = 0.32) had the largest direct effects on individual plant yields in the seven years of the study. These yield components could be effective selection criteria for the improvement of soybean yield in the soybean breeding program in southern Tamaulipas, Mexico.

Acknowledgments

The authors wish to thank the Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional (IPN) (20131450 grant) and the Instituto Nacional de Investigaciones Forestales, Agricolas, y Pecuarias (22281931997 grant).

References

Alejos G, Monasterio P, Rea R (2006) Análisis de la interacción genotipo-ambiente para rendimiento de maíz en la región maicera del estado Yaracuy, Venezuela. Agron Tropical 56(3): 369-384.

Ball R, McNew R, Vories E, Keisling T, Purcell L (2001) Path analyses of population density effects on short-season soybean yield. Agron J. 93: 187-195.

Board J, Kang M, Harville B (1997) Path analyses identify indirect selection criteria for yield of late-planted soybean. Crop Sci. 37: 879-884.

Burgueño J, Crossa J Vargas M (2002) SAS Programs for graphing GE and GGE biplots. *En*: Biometrics and Statistics Unit, CIMMYT, http://www.cimmyt.org. Verified 8 October 2012.

Crossa J (1990) Statistical analyses of multilocation trials. Adv. Agron. 44: 55-85.

Dewey DR, Lu KH (1959) A correlation and path coefficient analysis of components of crested wheatgrass seed production. Agron. J. 51: 515-518.

Di Rienzo JA, Casanoves F, Balzarini MG, González L, Tablada M, Robledo CW (2008) InfoStat, versión 2008, Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina.

Ebdon JS, Gauch HG (2002a) Additive main effect and multiplicative interaction analysis of national turfgrass performance trials: I. Interpretation of genotype x environment interaction. Crop Sci. 42: 489-496.

Fikere M, Fikere E, Tadele T, Legesse T (2009) Parametric stability analyses in field pea (*Pisum sativum L.*) under south Ethiopian condition. Agr Sci. 5: 146-151.

Fikere M, Tadesse T, Letta T (2008) Genotype-environment interactions and stability parameters for grain yield of faba bean (*Vacia faba* L.) genotypes grown in south eastern Ethiopia. Biotechnol. 7: 2847-2853.

Gaikwad S, Bangar N, Chavan B (2007) Correlation and path coefficient analysis in soybean. J Maharashtra Agr U. 32: 276-277.

García E (1987) Modificaciones al sistema de clasificación climática de Köppen (para adaptarlo a las condiciones de la República Mexicana). Cuarta edicion, Universidad Nacional Autónoma de México. México, D.F. 130 p.

García JG, López E, Cumpían J, Cantú MA, Zavala F, Mayek N (2012) Grain yield charcoal tot resistance stability in common beans under terminal drought conditions. J Phytopatol. 160: 98-105.

Gauch HG (2006) Statistical Analysis of Yield Trials by AMMI and GGE. Crop Sci. 46: 1488-1500.

Gauch HG (1988) Model selection and validation for yield trials with interaction. Biometrics. 44: 705-715.

Gauch HG, Piepho HP, Annicchiarico P (2008) Statistical analysis of yield trials by AMMI and GGE: further considerations. Crop Sci. 48: 866-889.

González T, Monteverde E, Marín C, Madriz IPM (2007) Comparación de tres métodos para estimar estabilidad del rendimiento en nueve variedades de algodón. Interciencia. 32(5): 344-348.

Hair Jr JF, Anderson RE, Tatham BWC (1995) Multivariate data analysis with readings, Fourth ed. Prentice Hall. International edition. Englewood. NJ.

Hassanpanah D, Azimi J (2010) Yield stability analysis of potato cultivars in spring cultivation and after barley harvest cultivation. American-Eurasian J Agr Environ Sci. 9 (2): 140-144.

Hernandez MV, Crossa J (2000) The AMMI analysis and graphing the biplot. CIMMYT, INT. Mexico. [Online]: Available at http://www.cimmyt.org/biometrics/biplots.exe; verified 8 November 2012.

Karasu A, Oz M, Göksoy AT, Turan ZM (2009) Genotype by environment interactions, stability, and heritability of seed yield and certain agronomical traits in soybean [*Glycine max* (L.) Merr.]. Afr J Biotechnol. 8(4): 580-590.

Li W, Yan ZH, Wei YM, Lan XJ, Zheng YL (2006) Evaluation of genotype × environment Interactions in Chinese spring wheat by the AMMI model, correlation and path analysis. J Agr Crop Sci. 192: 221-227.

Maldonado NM (1994) Huasteca 100 y Huasteca 200 nuevas variedades de soya para el sur de Tamaulipas. CIRNE-INIFAP. Altamira, Tamaulipas., México. Foll tec Num. 9. 20 p.

Maldonado MN, Ascencio LG (2010b) Huasteca 200, variedad de soya de baja sensibilidad al fotoperiodo corto para el trópico de México. Rev Mex Cien Agr. 1(5): 707-714.

Malik M, Ashraf M, Qureshi A, Ghafoor A (2007) Assessment of genetic variability, correlation and path analysis for yield and its components in soybean. Pak J Bot. 39: 405-413.

Malik M, Qureshi A, Ashraf M, Ghafoor A (2006a) Genetic variability of the main yield related characters in soybean. Int J Agric Biol. 8(6): 815-819.

Mohammadi SA, Prasanna BM, Singh NN (2003) Sequential path model for determining interrelationships among grain yield and related characters in maize. Crop Sci. 43: 1690-1697.

Morais LK, Da Silva RM, Chiorato AF, Azzini LE, Villela OV, Gallo PB, Sakai M, Basto CR, Malavolta VM (2008) Adaptabilidade e estabilidade fenotípica de genótipos de arroz irrigado para o estado de São Paulo. Rev Bras Biocienc. 14(1): 9-16.

Oz M, Karasu A, Goksoy A, Turan Z (2009) Interrelationships of agronomical characteristics in soybean (*Glycine max)* grown in different environments. Int J Agric Biol. 11: 85-88.

Pacheco RM, Duarte JB, Souza PIM, De Silva SA, Da Nunes Jr J (2009) Key locations for soybean genotype assessment in central Brazil. Pesq Agropec Bras. 44: 478-486.

Padilla VI, Valenzuela VRI, Armenta CCM, Salinas P, Sánchez SE (2008) Comportamiento agronómico de genotipos de garbanzo en siembra tardía en el valle del Mayo, Sonora, México. Rev Fit Mex. 3: 43-49.

Pathan SM, Vuong T, Clark K, Lee JD, J. Shannon G, Craig AR, Ellersieck MR, Burton JW, Cregan PB, Hyten DL, Nguyen HT, Sleper DA (2013) Genetic mapping and confirmation of quantitative trait loci for seed protein and oil contents and seed weight in soybean. Crop Sci. 53: 765-774.

Pereira HS, Melo LC, Del Peloso MJ, De Faria MJC, Da Costa LGC, Cabrera DL (2009) Comparison of methods for phenotypic adaptability and stability analysis in common bean. Pesq Agropec. 44: 374-383.

Rowntree SC, Suhre JJ, Weidenbenner NH, Wilson EW, Davis VM, Naeve SL, Casteel SN, Diers BW, Esker PD, Specht JE, Conley SP (2013) Genetic gain × management interactions in soybean: I. Planting date. Crop Sci. 53:1128-1138.

Samonte S, Wilson L, McClung A, Medley J (2005) Targeting cultivars onto rice growing environments using AMMI and SREG GGE Biplot analysis. Crop Sci. 45: 2414-2424.

SAS Institute (1998) SAS/STATR User Guide. Release 6.03 Edition.

Sedghi M, Amanpour BB (2010) Sequential path model for grain yield in soybean. Not Sci Biol. 2(3): 104-109.

Shukla S, Singh K, Pushpendra (1998) Correlation and path coefficient analysis of yield and its components in soybean (*Glycine max* L. Merrill). Soybean genet news. 25: 67-70.

SoyStats (2012) A Reference guide to important soybean facts and figures http://www.soystats.com/2012/Default-frames.htm. Verified 12 November 2012.

Silva WC and Duarte JB (2006) Métodos estadísticos para estudo de adaptabilidade e estabilidade fenotípica em soja. Pesq Agropec Bras. 41 (1): 23-30.

Tai GCC, Levy W, Coleman K (1994) Path analysis of genotype interactions of potatoes exposed to increasing warm climate constraints. Euphytica. 75: 49-61.

Tai GCC, Tarn TR (2003) Evaluation of varieties: genotypes, environments and genotype-environment interactions. Acta Hort. 619: 31-34.

Tamkoc A, Ustum A, Altinok S, Acikgoz E (2009) Biomass and seed stability pea genotypes. Agr Environ. 7: 140-146.

Voltas J, Van EF, Igartua E, García del Moral LF, Molina-Cano JL, Romagosa I (2002). Genotype by environment interaction and adaptation in barley breeding: Basic concepts and methods of analysis. In: Slafer GA, Molina-Cano JL, Savin R, Araus JL, Romagosa I (eds) Barley Science: Recent Advances from Molecular Biology to Agronomy of Yield and Quality. The Harworth Press Inc., New York, pp. 205-241.

Whittaker A, Vazzana C, Vecchio V, Benedettelli S (2009) Evaluation of direct and indirect effects of flavonoids, mineral elements and dry weight on antiradical scavenging activity in leaf material of field-grown Trifolium pratense cultivars using Path Analysis. Field Crop Res. 113(1): 1-11.

Williams WA, Jones MB, Demment MW (1990) A Concise Table for Path Analysis Statistics. Agron J. 82: 1022-1024.

Wright S (1921) Correlation and causation. J Agr Res. 20: 557-585.