# Evaluation of Multienvironment Trials of Peanut Cultivars

F. Casanoves,* J. Baldessari, and M. Balzarini

## ABSTRACT

Multienvironment yield trials (MET) for advanced peanut lines are conducted each year at the EEA-Manfredi Peanut Breeding Program, the main INTA program for developing new peanut (*Arachis hypogaea* L.) cultivars for cultivation in the Argentinean crop area. The main objective of this work was the simultaneous analysis of several multienvironment yield tests first to identify superior cultivars for the peanut crop area in Argentina, and second to investigate if different megaenvironments exist. The simultaneous evaluation of several years of MET provides information that allows researchers to better guide breeding strategies. We analyze a 6-yr series of grain yield data from MET, involving 18 genotypes and five test locations using six by-year analyses of complete yield data sets and an Additive Main Effect and Multiplicative Interaction (AMMI) mixed model analysis combining all 6 yr of MET. AMMI models in a mixed model framework were used for exploring genotype–environment (GE) interaction since the lists of genotypes annually tested in multienvironment trials vary from year to year since new genotypes are introduced every year and others are withdrawn. The results allowed us to identify mf484 and mf505 as superior cultivars and confirm the existence of a unique megaenvironment for identifying high yield cultivars in the peanut crop area of Argentina. The mixed model approach of MET data was successfully implemented to analyze highly unbalanced GE data sets.

THE GERMPLASM EVALUATION is a crucial activity in plant breeding (Stroup, 2000). MET are commonly conducted annually to obtain information that supports recommendations of superior cultivars for cultivation. MET are used to evaluate several genotypes in multiple environments (locations and/or years), and they are essential because of the presence of GE interaction, i.e., differential genotypic responses in different environments. The main objective in the evaluation of a series of MET is to identify superior cultivars for a target region and to determine if this region can be subdivided into different megaenvironments to better guide breeding strategies (Kang, 2002). Important concepts such as ecological regions, ecotypes, megaenvironments, specific adaptations, and stability originated from the GE interaction analysis (Yan and Hunt, 2002). The identification of megaenvironments is associated with the exploration of the annually repeatable GE interaction patterns. For a particular megaenvironment, genotypes are evaluated on the basis of mean yield and stability of yields across environments.

The EEA-Manfredi, INTA, Argentina, conducts MET in a Peanut Breeding Program (PBP-INTA) in different environments (locations and years). The selected locations of MET are representative of the environmental characteristics of the northern, central and the southern zones of the Argentinean peanut crop area (Córdoba province) that extends from 32° to 33°50′ S latitude and from 63° to 64°35′ W longitude. This region produces 95% of the Argentinean peanuts since it shows high homogeneity of climate and soil characteristics appropriated for the cropping. The genotypes evaluated during all 6 yr of MET vary from year to year because new genotypes are introduced every year and others are withdrawn. Therefore, the MET databases through the years are incomplete, i.e., all of the genotypes are not present in all of the combinations of locations and years.

The unbalanced data sets obtained from compiling several years of MET data was handled by modeling the response under the mixed model theory (Balzarini, 2000). This modeling strategy allows inferring the genotype performance across environments even when the GE tables are incomplete. Additionally, the stability and interaction measures can be obtained as certain mixed model parameters (Piepho, 1998; Balzarini, 2002).

In this study, the genotype effects were considered fixed because they refer to genotype sets in advanced selection stages. The environmental effects were considered random, except for an annual MET analysis where differences among environments refer to differences among locations and they could be associated with predictable differences among locations (Allard and Bradshaw, 1964). For the combined MET analysis (through 6 yr), the environments were defined as combinations of location and year effects, and treated as random. For both by-year and all-year MET analyses, the GE interaction was modeled by an AMMI model, but with the fixed or the mixed approach to AMMI according to the balancedness of the data sets. The evaluation of several years of multienvironment yield trials provides useful information to help researchers to better guide breeding strategies.

Up until now, studies have not been done to investigate better cultivars and the possible existence of different megaenvironments using yield data from several MET. The main objective of this work was the simultaneous analysis of several multienvironment yield tests to identify superior cultivars for the peanut crop area

---

F. Casanoves, Centro Agronómico Tropical de Investigación y Enseñanza, 7170 Turrialba, Costa Rica; J. Baldessari, EEA-Manfredi, Instituto Nacional de Tecnología Agropecuaria, Manfredi, Córdoba, Argentina; and M. Balzarini, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, cc 509, (5000) Córdoba, Argentina. Received 16 Oct. 2003. *Corresponding author (casanoves@catie.ac.cr)

**Table 1. Participating locations in multienvironment trials of the INTA-EEA-Manfredi Peanut Breeding Program of Argentina from 1996–1997 to 2001–2002.**

| Locations (sites) | Soil type | Altitude (meters above sea level) | Latitude south | Longitude west |
|---|---|---|---|---|
| Location 1 (General Deheza and General Cabrera) | loam-sandy loam | 325 | 32°44′ | 63°45′ |
| Location 2 (Manfredi) | silty loam | 300 | 31°50′ | 63°44′ |
| Location 3 (Santa Eufemia) | finest sandy loam | 250 | 33°12′ | 63°17′ |
| Location 4 (El Sur, San Ambrosio, Las Vertientes, Reducción, Las Acequias) | finest sandy loam | 330 | 32°48′ | 64°25′ |
| Location 5 (Gigena) | sandy loam | 450 | 32°44′ | 64°17′ |

in Argentine, and secondarily to determine if different megaenvironments exist.

## MATERIALS AND METHODS

### Database

The information sources used include the MET of PBP-INTA conducted from year 1996–1997 to 2001–2002, which correspond to the first 6 yr of multienvironmental evaluation in the program. The trials were conducted in 10 sites, some of which because of their proximity and similarity in soil and weather conditions were regarded as the same location for the purposes of the analysis (Table 1). Altogether 18 genotypes, four short cycle and 14 long cycle, were evaluated (Table 2). The set of evaluated genotypes was the same for all the locations within a year, but the genotype list in the study period varied from year to year. The MET of the PBP-INTA for advanced lines began in 1996 with 11 genotypes without structural changes (genotypes and locations) in the first 2 yr. During the next 4 yr, seven new genotypes were incorporated, six genotypes were dropped and new locations were added. Florman, mf484, mf485, mf487, and mf489 were the only cultivars evaluated during all 6 yr. The genotype × location × year database is highly unbalanced. Table 2 contains a list of the evaluated genotypes. In all of the sites, the trials were laid out as a randomized complete block design, with four replications. The plots were made of two 10-m-long furrows with 70 cm between furrows. Recommended seeding rates (15 seeds/m$^2$) and cultural practices were used at all locations. Each plot was manually harvested. The analyzed yield values correspond to kilograms of peanuts per plot at constant moisture (80 g kg$^{-1}$). In Table 3, the main events

**Table 2. Genotypes evaluated in multienvironment trials in the EEA-Manfredi Peanut Breeding Program, INTA, from 1996–1997 to 2001–2002.**

| Genotype | Cycle† | Ancestry |
|---|---|---|
| manf393 | S | Robut 33-1/NC Ac 2698 |
| mf447 | S | Florman/Manfredi Virginia 5 |
| mf478 | S | MGS 9/NC Ac 2232 |
| mf480 | S | CS 9/ICGS 5 |
| Florman | L | Selection of Florunner |
| mf457 | L | Florman/Tachimasari |
| mf484 | L | Florman/Marc 1‡ |
| mf485 | L | Florman/Marc 1‡ |
| mf486 | L | Florman/Marc 1‡ |
| mf487 | L | Florman/Marc 1 |
| mf489 | L | Florman/Marc 1 |
| mf496 | L | Florman/(Mf321/RCM 1451) |
| mf499 | L | Florman§2/Colorado Irradiado |
| mf505 | L | Florman/F435-2-3-B-2-1-b4-B-3-b3-1-B |
| mf506 | L | Florman/F435-2-3-B-2-1-b4-B-3-b3-1-B |
| mf508 | L | Florman/F435-2-3-B-2-1-b4-B-3-b3-1-B |
| mf510 | L | Florman/F435-2-3-B-2-1-b4-B-3-b3-1-B |
| Tegua | L | Selection of Florunner |

† S = short cycle, L = long cycle.
‡ Their pedigree translates to the same plant F$_1$.
§ Crossing,/indicates parental separation.

that characterized the years involved in this evaluation of MET data are shown.

### Statistical Analysis of MET

Considering the locations as the basic megaenvironment units and that the genotype × year and location × year data tables are highly unbalanced, genotype × location (GL) interaction was explored via individual yearly analyses (complete data tables). In this first modeling attempt, we run an analysis of variance model which included fixed effects for location (L), genotypes (G), and genotype × location interaction (GL). The adjusted model was:

$$Y_{ijk} = \mu + L_j + B(L)_{k(j)} + G_i + GL_{(ij)} + \varepsilon_{ijk} \quad [1]$$

where $Y_{ijk}$ is the yield of Genotype $i$, in the Location $j$, Block $k$; $\mu$ is the overall mean; $L_j$ is the effect of Location $j$ with $j = 1,…,s$; $B(L)_{k(j)}$ is the effect of Block $k$ within Location $j$ with $k = 1,...,n$; $G_i$ is the effect of Genotype $i$ with $i = 1,.. g$; $GL_{(ij)}$ is the interaction effect between Genotype $i$ with Location $j$ and $\varepsilon_{ijk}$ is the random error term associated with observation $Y_{ijk}$. Two models were developed, one model assumed homogenous residual variances and the other allowed heterogeneous residual variances across locations. The latter model was evaluated because the trials conducted in different locations could have different precision (residual variance).

For each year, the comparison of genotype performance across environments (broad inference) was based on the least squares means. Since a fixed effects model was considered, the standard errors of mean differences only depended on the residual variance. The narrow inferences (environment-specific inference) about a genotype performance were obtained from genotype means in each environment. Symmetrical biplots were used (Gabriel, 1971) to analyze GL interaction patterns associated with the AMMI model (Gauch, 1988) and with the site regression model (SREG) (Cornelius et al., 1996). The biplot algorithm is based on the single value decomposition of a residual matrix. The AMMI model biplots are constructed from the residual matrix of the additive model, i.e., Eq. [1] without the GL interaction effect. The SREG model biplots, known as GGE biplot (Yan et al., 2000), are con-

**Table 3. Main characteristics of the agricultural years 1996–1997 to 2001–2002 in the EEA-Manfredi Peanut Breeding Program, INTA.**

| Year | Events |
|---|---|
| 1996–1997 | Low precipitation. |
| 1997–1998 | Late harvest (mainly in Manfredi). Sclerotinia [caused by *Sclerotinia sclerotiorum* (Lib.) de Bary and *Sclerotinia minor* Jagger] infection in General Deheza. |
| 1998–1999 | Low precipitation during the crop cycle and high during the harvest. Early frost. |
| 1999–2000 | Low precipitation in the reproductive stage. Late harvest. |
| 2000–2001 | Low precipitation (except in Manfredi). Late harvest. |
| 2001–2002 | Low precipitation in the late stages in Sta. Eufemia and Cabrera. |

**Table 4. Fitting criteria for the fixed effects multienvironment trials model with and without heterogeneous residual variance across the locations**

| Year | Homogeneous residual variance (HoRV) | | | Heterogeneous residual variance (HeRV) | | | Best model |
|------|------|------|------|------|------|------|------|
| | AIC | BIC | −2 res Log(likelihood) | AIC | BIC | −2 res Log(likelihood) | |
| 1996–1997 | −96.3 | −97.5 | 190.6 | −84.1 | −88.5 | 162.3 | HeRV |
| 1997–1998 | −70.0 | −71.2 | 138.0 | −68.8 | −73.2 | 131.7 | HeRV |
| 1998–1999 | −129.9 | −131.4 | 257.7 | −132.4 | −141.0 | 254.8 | HoRV |
| 1999–2000 | −109.7 | −111.2 | 217.5 | −104.0 | −110.6 | 200.1 | HeRV |
| 2000–2001 | −60.3 | −61.7 | 118.6 | −46.6 | −53.1 | 85.2 | HeRV |
| 2001–2002 | −113.4 | −114.9 | 224.8 | −115.2 | −121.8 | 225.5 | HoRV |

structed from the residual matrix corresponding to an adjustment Eq. [1], which omits G and GL.

Scatter plots were generated for each year using mean yield (*x* axis) and stability measurement (*y* axis) for each genotype. Four stability statistics were calculated: (i) the CV (Francis and Kannenberg, 1978) because it is the statistic traditionally used in the PBP-INTA, (ii) the stability variances of Shukla (1972), (iii) the first principal component (PC1) of the AMMI model analysis, and (iv) the first and/or second principal component (PC1 and PC2) of the SREG model.

## Combined MET analysis

A model with environment, genotype, and genotype × environment effects was used. Environmental effects were defined as the combination of the location and years. All the effects used in the model were considered random, except for genotype. The equation for the response of Genotype *i* in Block *k* in Environment *j* is:

$$Y_{ijk} = \mu + G_i + E_j + GE_{ij} + B(E)_{k(j)} + \varepsilon_{ijk} \quad [2]$$

where $Y_{ijk}$ is the response of Genotype *i*, in Environment *j* and Block *k*, $\mu$ is the overall mean, $G_i$ is the fixed effect of Genotype *i*, $E_j$ is the random effect of the Environment *j*, $GE_{ij}$ is the random effect of the interaction of Genotype *i* and Environment *j*, $B(E)_{k(j)}$ is the effect of Block *k* within the Environment *j*, $\varepsilon_{ijk}$ is the random error term associated with Observation $Y_{ijk}$.

**Table 5. Variability of yields between genotypes (G), locations (L), block between site [B(L)], and G × S interaction for PBP-INTA multienvironment trials, for the 1996–1997 to 2001–2002 years.**

| Year | Source | df† | SS‡ | *P* value | % (G+L+GL) |
|------|--------|-----|------|-----------|------------|
| 1996–1997 | L | 2 | 77.83 | <0.0001 | 85.8 |
| | G | 10 | 6.91 | 0.0038 | 7.6 |
| | G × L | 20 | 6.03 | 0.2340 | 6.6 |
| | B(L) | 9 | 9.33 | <0.0001 | |
| 1997–1998 | L | 2 | 112.96 | <0.0001 | 74.9 |
| | G | 10 | 6.35 | <0.0001 | 4.2 |
| | G × L | 20 | 31.45 | <0.0001 | 20.8 |
| | B(L) | 9 | 6.29 | <0.0001 | |
| 1998–1999 | L | 4 | 173.78 | <0.0001 | 90.0 |
| | G | 11 | 6.29 | <0.0001 | 3.2 |
| | G × L | 44 | 12.83 | 0.0026 | 6.7 |
| | B(L) | 14 | 7.23 | 0.0001 | |
| 1999–2000 | L | 3 | 269.32 | <0.0001 | 86.7 |
| | G | 11 | 21.71 | <0.0001 | 7.0 |
| | G × L | 33 | 19.66 | <0.0001 | 6.3 |
| | B(L) | 12 | 6.52 | 0.0001 | |
| 2000–2001 | L | 3 | 191.50 | <0.0001 | 96.0 |
| | G | 11 | 2.10 | 0.0043 | 1.0 |
| | G × L | 33 | 5.96 | 0.0001 | 3.0 |
| | B(L) | 12 | 5.60 | <0.0001 | |
| 2001–2002 | L | 3 | 169.76 | <0.0001 | 81.5 |
| | G | 11 | 18.14 | <0.0001 | 8.7 |
| | G × L | 33 | 20.28 | <0.0001 | 9.8 |
| | B(L) | 12 | 8.10 | <0.0001 | |

† df = degrees of freedom.
‡ SS = sum of squares.

The overall genotype performance (broad inference) was based on the genotype means, but the standard errors of the mean differences depended on the mixed model used for the variance of genotype means that included variance components associated with the GE interaction. Thus, the statistical comparison of the genotype performance, both yield differences and differences in yield stability were considered. The variance components were estimated via restricted maximum likelihood (REML) using SAS (SAS Institute, 1997) PROC MIXED.

Different models were developed for the variance and covariance matrix of the random interaction terms for a given environment: (i) traditional mixed model, i.e., Eq. [2] assuming homoscedastic variances for the random GE terms and (ii) mixed stability variance model, i.e., model analogous to Eq. [2] considering heteroscedastic variances across genotypes for the GE interaction terms, and 3) AMMI mixed model. The AMMI mixed model is expressed as

$$y_{ijk} = \mu + G_i + E_j + B(E)_{k(j)} + \sum_{m=1}^{M} \lambda_m \gamma_{im} \alpha_{mj} + E_{ijk},$$

where $y_{ijk}$ is the response of Genotype *i*, in Environment *j* and Block *k*, $\mu$ is the overall mean, $G_i$ is the fixed effect of Genotype *i*, $E_j$ is the random effect of the Environment *j*, and $\sum_{m=1}^{M} \lambda_m \gamma_{im} \alpha_{mj}$ models the GE interaction as the sum of *M* multiplicative components that explain the GE interaction in *M* orthogonal directions [in each direction, fixed genotype scores ($\gamma_{im}$) and random environment ($\alpha_{mj}$) intervene]; $\lambda_m$ is a weight factor associated with the *m*th direction; and $\varepsilon_{ijk}$ represents the portion of the $GE_{ij}$ interaction not explained by the model plus the random error term associated with $y_{ijk}$.

Using the homocedastic factor analytic (FA1) structure (Jennrich and Schluchter, 1986) to model the variance and covariance matrix of the GE interaction terms within environment, the estimated covariance parameters can be used as genotype scores (genotypic sensitivity) to explain the GE interaction. The general factor analytic (FA) model is **QQ'** + **D**, where **Q** is a scores matrix of **q**s and **D** is a diagonal matrix with the possibility of nonnegative parameters on the diagonal. If matrix **D** is omitted, that is to say **D** = **0**, then the model is denoted as FA0. If matrix **D** has all its diagonal elements equal (**D** = $\sigma^2$**I**), then the model is denoted as FA1. In this work structures FA1, of order 1, 2, and 3 corresponding to AMMI models with $M = 1$, $M = 2$ and $M = 3$ multiplicative terms, respectively, were used. The selection of the most suitable parameterization to model the GE interaction was made through maximum likelihood ratio tests, the Akaike information criterion (AIC) and the Schwarz Bayesian criteria (BIC) (Littell et al., 1996). The graphical biplots associated with a AMMI mixed model (AMMI biplot) were obtained by graphing the covariance (standardized) parameters associated with each genotype in each multiplicative term versus the empirical best linear unbiased estimator (BLUP) of the environmental
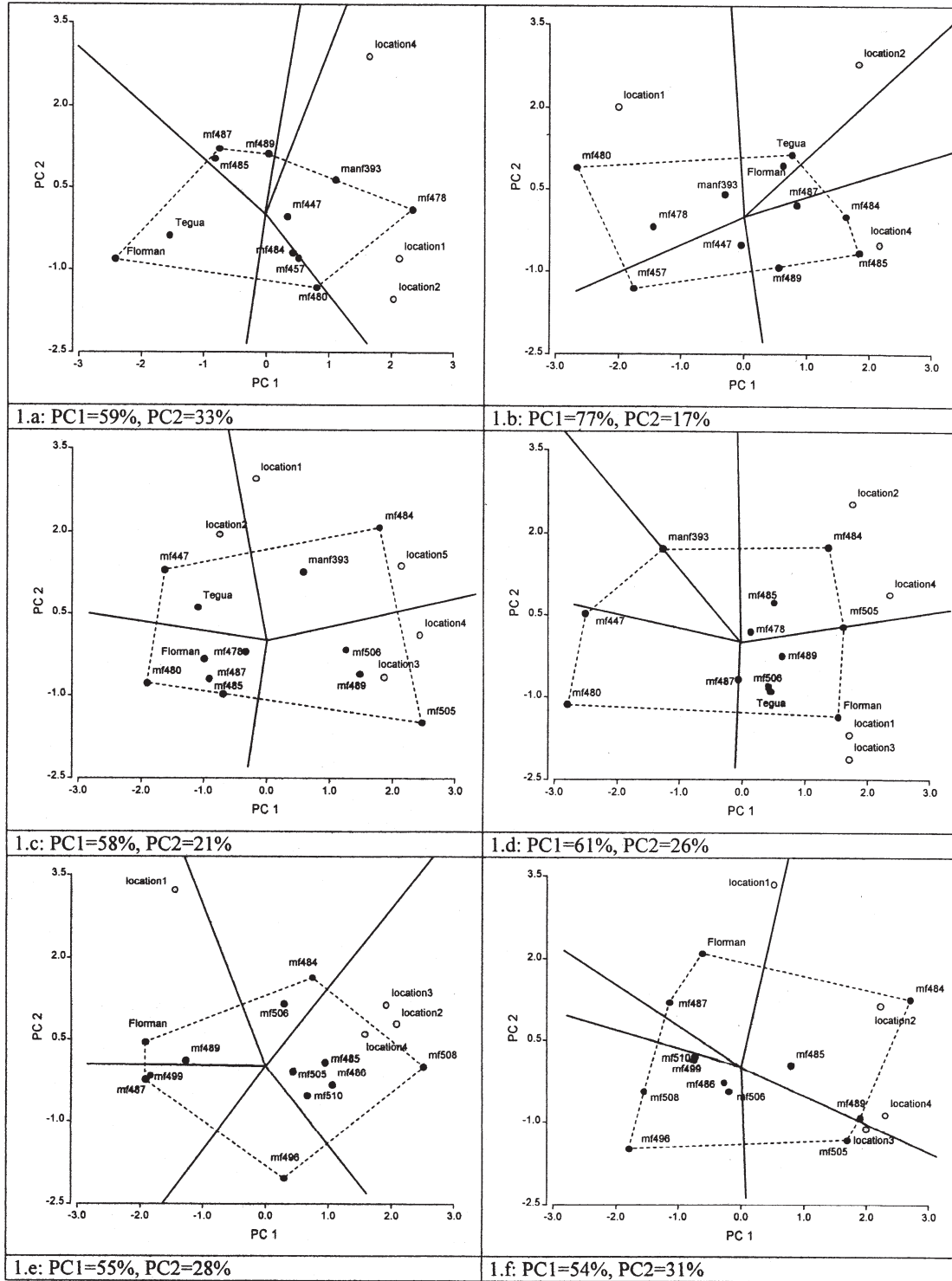
**Fig. 1. GGE biplots based on 6 yr of PBP-INTA multienvironment trials. The dark points represent genotypes and the light points represent locations.**

effect on the same term using the statistical software Info-Gen (Info-Gen, 2003).

## RESULTS AND DISCUSSION

### Analysis of MET by Year

The model with heterogeneous residual variances across locations fit better than the model with homoge-

nous variances in 4 of the 6 yr of MET data analyzed (Table 4). The Akaike information criteria (AIC) and Schwarz (BIC) (in which a greater value implies a better fit) coincided with the likelihood ratio test in identifying the years with different precision from the trials conducted in different locations.

Table 5 contains the analysis of variance for each year of MET data processed with the models with better fits

as it is suggested in Table 4. In all the years analyzed, except 1996–1997, the GL interaction was statistically significant ($P < 0.01$). The PBP-INTA MET began in 1996, the year in which the residual variance was highest. This low accuracy of the tests could mask the presence of GL interaction. The relative magnitude of the GL interaction with respect to the variability explained by (G+L+GL) for each year is given in Table 5. The variation caused by the GL interaction was smaller than the variation among genotypes in 5 of the 6 yr of the MET analyzed. The variation between locations was always the most important one, explaining between 74.9 and a 96.0% of the total variation, which justifies the selection of the biplots based on the SREG model for MET analysis (Yan et al., 2000).

The first two principal components (PC1 and PC2) obtained by singular value decomposition of the centered data (SREG model) explained more than 85% of the total variability caused by (G+GL) for all the years except in 1998–1999 and 2000–2001 where these percentages were 79 and 83%, respectively. Figure 1 represents the GGE biplots obtained for each year of MET.

In the year 1996–1997 (Fig. 1.a), there was a high correlation of the PC1 ($r = 0.98$, $P < 0.0001$) and low correlation of the PC2 ($r = 0.15$, $P = 0.641$) with yields. The extreme genotypes on the PC1 axis, mf478 (positive PC1) and Florman (negative PC1), do not show a crossover interaction (COI); these were genotypes of higher and lower rankings, respectively. The GL interaction was not significant and all the locations behaved as a unique megaenvironment where some advantages were demonstrated for the short-cycle cultivars over those of long cycle. This is possibly due to the low precipitation registered during this year. The mf487 and mf489 genotypes (positive PC2) and the mf480 genotype (negative PC2) were the only ones that showed a differential response, COI across environments.

In the year 1997–1998 (Fig. 1.b), the yields of the long-cycle check cultivars were relatively high and homogenous across environments, showing advantages with respect to the other genotypes in Location 2 where the harvest was late. The mf485 and mf484 genotypes showed COI, with a relatively superior performance in Location 4 and with a lower ranking in Location 1. The mf480 and mf457 genotypes showed the opposite behavior with relatively high performances in Location 1, where the crop was subjected to stress by a *Sclerotinia minor* Jaegger infection. In this year, the yield did not correlate with the PC1 ($r = 0.062$; $P = 0.857$). Therefore, the genotypes with greater projections of the PC1 (mf485, mf484, mf457 and mf480) showed COI, whereas the Tegua, Florman, and mf480 genotypes, with greater positive projections of the PC2, yielded relatively better than the rest during this year.

The year 1997–1998 was the only year of MET data where the GL interaction explained a greater percentage of (G+L+GL) than G. Although this result could suggest the possible existence of different megaenvironments in the PBP-INTA target region, it is important to note that the observed pattern was not consistent through the years. Besides that, short- and long-cycle

cultivars were evaluated, and the harvests in all of the MET, mainly in Location 2, were delayed during this year.

In the year 1998–1999 (Fig. 1.c), the correlation of the PC1 with yield was 0.89 ($P < 0.0001$). The genotypes with greater PC1 values were mf505 and mf484 (long cycle), both with the highest ranking in most of the locations, except Location 2. The short-cycle cultivars, mf447 and mf480 (negative PC1), showed relatively poor performances except in Location 2. The precipitation was low during the cultivation cycle and high during the harvest, which could explain the differential performance of the short-cycle cultivars. The differences in the genotype response across environments were less than the differences in the mean genotype response, thus the data did not suggest the presence of megaenvironments. Location 4 was more favorable than Location 2 in this year for genotype mf505. Nevertheless, in the following year (Fig. 1.d), this genotype showed advantages relative to the other genotypes in Location 4 and 2. The remaining genotypes are grouped into a unique megaenvironment where the short cycle mf447, mf480, and manf393 genotypes had the worst performance.

In the year 2000–2001 (Fig. 1.e), the contribution of the GL interaction was approximately equal to the variability among genotypes and the yield correlated significantly ($P < 0.05$) with PC1 and PC2, which suggested that the projections of PC2 did not necessarily indicate COI. The extreme genotypes in their projections on the PC1, mf508, Florman, and mf487, showed greater COI. The Florman and mf487 rankings were relatively higher in Location 1 than in the rest, which showed a behavior opposite to mf508. The mf484 and mf496 genotypes with opposite PC2 values responded proportionally to the differences between high- and low-ranking locations, respectively.

In the year 2001–2002 (Fig. 1.f), the variability between genotypes was much greater than the GL interaction and the yields correlated significantly with PC1 ($r = 0.95$, $P < 0.0001$). The mf484, mf505 and mf489 genotypes had the best performance in Locations 2, 3, and 4, whereas Florman and mf487 showed an advantage in Location 1. The genotypes mf508 and mf496 were the worst performers. The separation of Location 1 with respect to the rest, where Florman and mf487 showed relatively high yields, could be due to factors such as scarce rainfall during the grain filling and fruit loss due to the high precipitation registered at harvest in that location. In Location 3, the precipitation was also low, but the differences between the yields of the genotypes were small since a low yield in all of the genotypes was registered.

Figure 2 represents a comparison of the performances of mf480 (short-cycle genotype) and mf489 genotypes (long-cycle) across three locations using the GGE biplot for year 1997–1998, where the GL contribution was greater than that of G. The comparison was made by connecting the point that represents the mf480 genotype with the point that represents mf489 genotype with a straight line and drawing a perpendicular line that passes through the origin (Yan and Hunt, 2002). The perpen-
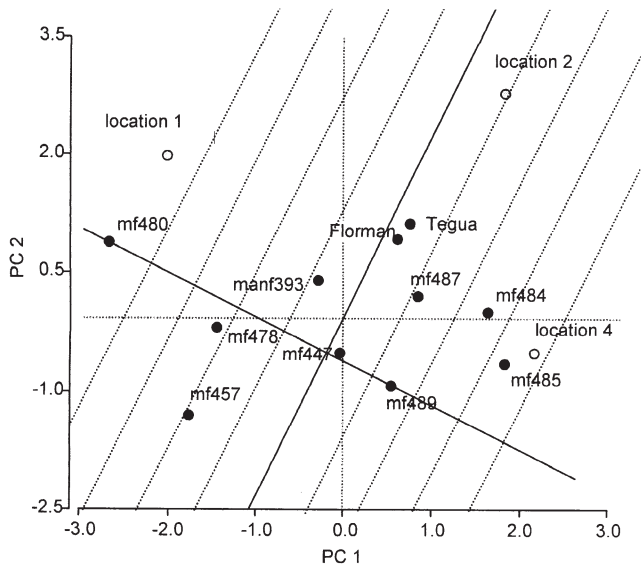
**Fig. 2. Comparison of short and long cycle genotype performance in all of the intervening multienvironment trials locations during the 1997-1998 year. The dark points represent genotypes and the light points represent locations.**

**Table 7. Variance components for genotypes (G), environments (E) and G × E interaction for PBP-INTA multienvironment trials. Joint analysis of the 1996–1997 to 2001–2002 years.**

| Variance component | Estimate | Percent |
|---|---|---|
| Environment | 1.1813 | 90.5 |
| Genotype | 0.0137 | 1.0 |
| G × E | 0.1100 | 8.5 |
| Genotype × location | 0.029 | 23.9 |
| Genotype × year | 0.026 | 21.5 |
| Genotype × year × location | 0.066 | 54.6 |
| Residual | 0.1489 | |

result in Fig. 2, suggest that the peanut region in which cultivars were tested does not show repeatable megaenvironments for breeding purposes. From the results obtained with the analysis of MET data conducted in 1997–1998 where the COI was significant because of differences in natural infections across locations, experiments specially designed to create different environmental conditions appear as appropriated for breeding with respect to pathogen resistance.

Table 6 contains for each one of the analyzed years of MET data, the mean yield for the genotypes, the AMMI model PC1, and the coefficient of variation (CV) as a stability statistics. Although the simultaneous interpretation of the CV and the yield mean constitute a common strategy of analysis implemented in this breeding program, our results show that use of the CV is associated with recommendations for low performance genotypes. For example, in the year 1996–1997, the mf480 genotype has a low CV, and nevertheless, it is a major contributor to the GL interaction. In Location 1, this genotype had a ranking of 11 whereas in Location 4, it ranked third. Something similar happened for the 1997–1998 data where although the ranking of mf480

dicular line separates the locations into two groups, which shows that mf480 genotype had greater relative yields than mf489 genotype in Location 1. Meanwhile, for the mf489 genotype the favorable location was Location 4 followed by Location 2, although yields at Location 2 were barely above average. If genotypes mf480 (short cycle) and mf485, mf484, mf487, and Florman (long-cycle) are compared this way, one concludes that Locations 4 and 2 are favored by the long-cycle cultivars.

A collective analysis of the six GGE biplots, plus the

**Table 6. Yield means and stability measures for each genotype during the 6 yr (1996–1997 to 2001–2002).**

| Genotype | 1996–1997 Year Mean† | CV | PC1 | Genotype | 1997–1998 Year Mean† | CV | PC1 | Genotype | 1998–1999 Year Mean† | CV | PC1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| mf478 | 2.24 a | 45.70 | 0.21 | Florman | 3.18 a | 33.48 | 0.17 | mf484 | 2.96 a | 25.51 | 1.28 |
| manf393 | 2.03 ab | 54.64 | 1.03 | Tegua | 3.16 ab | 34.42 | −0.10 | mf505 | 2.84 ab | 38.17 | 3.23 |
| mf447 | 1.94 ab | 53.31 | −0.19 | mf480 | 2.97 abc | 15.88 | 3.14 | mf489 | 2.82 ab | 39.95 | 1.44 |
| mf457 | 1.88 abc | 41.57 | −1.21 | mf487 | 2.90 abcd | 46.98 | −0.69 | manf393 | 2.79 abc | 29.07 | 0.29 |
| mf484 | 1.87 abc | 42.72 | −1.09 | mf484 | 2.86 cdef | 59.20 | −1.60 | mf506 | 2.78 abc | 35.46 | 1.41 |
| mf480 | 1.85 abc | 32.20 | −1.89 | mf457 | 2.69 cdef | 41.39 | 1.87 | mf478 | 2.67 bcd | 41.35 | −0.41 |
| mf489 | 1.85 bc | 65.18 | 1.73 | manf393 | 2.68 def | 37.82 | −0.11 | mf447 | 2.61 bcde | 36.25 | −1.97 |
| mf487 | 1.76 bcd | 70.35 | 1.74 | mf478 | 2.68 def | 27.71 | 1.27 | Tegua | 2.58 bcde | 46.52 | −1.96 |
| mf485 | 1.66 bcd | 70.71 | 1.65 | mf485 | 2.65 def | 73.83 | −2.25 | Florman | 2.53 cde | 48.03 | −1.36 |
| Tegua | 1.51 cd | 53.57 | −0.62 | mf447 | 2.57 ef | 49.55 | −0.46 | mf487 | 2.49 de | 42.23 | −0.81 |
| Florman | 1.36 d | 51.84 | −1.37 | mf489 | 2.49 f | 63.71 | −1.24 | mf485 | 2.44 e | 32.94 | 0.16 |
| | | | | | | | | mf480 | 2.39 e | 41.42 | −1.29 |
| LSD = 0.398 | | | | LSD = 0.297 | | | | LSD = 0.252 | | | |

| Genotype | 1999–2000 Year Mean† | CV | PC1 | Genotype | 2000–2001 Year Mean† | CV | PC1 | Genotype | 2001–2002 Year Mean† | CV | PC1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| mf505 | 3.39 a | 44.25 | −0.98 | mf484 | 2.00 a | 59.22 | 0.13 | mf484 | 3.53 a | 42.35 | 0.22 |
| mf484 | 3.34 ab | 45.36 | −2.74 | mf508 | 1.94 ab | 65.76 | −2.22 | mf489 | 3.17 b | 39.64 | −1.60 |
| Florman | 3.30 abc | 46.04 | 1.01 | mf506 | 1.93 ab | 60.36 | 0.39 | mf505 | 3.13 b | 36.90 | −2.13 |
| mf489 | 3.17 abc | 37.08 | 0.61 | mf485 | 1.89 abc | 69.43 | −1.31 | mf485 | 3.04 bc | 44.23 | −0.56 |
| mf485 | 3.15 abc | 51.43 | −1.40 | mf486 | 1.83 abcd | 59.52 | −1.47 | Florman | 2.93 bcd | 40.01 | 2.57 |
| Tegua | 3.09 bc | 45.99 | 1.00 | mf505 | 1.82 abcde | 61.96 | −0.20 | mf506 | 2.81 cd | 37.93 | −0.62 |
| mf506 | 3.09 bc | 51.23 | 0.61 | mf510 | 1.77 bcde | 58.84 | 0.18 | mf486 | 2.77 cd | 31.61 | 0.21 |
| mf487 | 3.04 c | 54.52 | 0.80 | Florman | 1.77 bcde | 71.35 | 2.02 | mf487 | 2.76 cd | 31.65 | 1.86 |
| mf478 | 3.03 c | 39.21 | 0.00 | mf489 | 1.76 bcde | 63.96 | 1.01 | mf499 | 2.72 de | 45.29 | 0.11 |
| manf393 | 2.71 d | 49.24 | −1.78 | mf499 | 1.71 cde | 68.33 | 1.37 | mf510 | 2.69 de | 35.15 | 1.00 |
| mf447 | 2.42 e | 51.60 | 0.17 | mf487 | 1.69 de | 68.25 | 1.91 | mf508 | 2.48 ef | 40.98 | 0.23 |
| mf480 | 2.29 e | 47.67 | 2.70 | mf496 | 1.64 e | 68.52 | −1.82 | mf496 | 2.35 f | 50.32 | −1.29 |
| LSD = 0.273 | | | | LSD = 0.188 | | | | LSD = 0.281 | | | |

† Different letters indicate differences in the mean yields across the locations (LSD, $P < 0.05$).

**Table 8. Fitting criteria for 3 mixed models adjusted to the PBP-INTA multienvironment trials data. Joint analysis of the 1996–1997 to 2001–2002 years.**

| Fitting criteria | Homogeneous variance for G × E terms | Heterogeneous variance for G × E terms | Mixed AMMI [2] for G × E terms |
|---|---|---|---|
| AIC | −802.3 | −799.2 | −765.1 |
| BIC | −804.3 | −810.0 | −787.3 |
| −2 res log(likelihood) | 1596.1 | 1560.4 | 1452.2 |
| No. of covariance G × A parameters | 1 | 18 | 36 |
| P value (no genotype effects) | 0.0027 | 0.0017 | <0.0001 |

changed from 11 to 1 across locations, this genotype showed the lowest CV. On the contrary, the use of the PC1 from the AMMI model as a stability measure allowed for better identification of the cultivars with COI. Genotype mf484 was among the superior ones in 5 of the 6 yr and mf505 genotype in 3 of the 4 yr of MET data.

## Combined Analysis of MET

The environments (combination of years and locations) constituted a source of important variation (90.5% of the total variation). The high variations due to environmental differences is expected in MET conducted through several years (Yan and Kang, 2003). The GE interaction was considerably higher than the variability attributable to G, and more than 50% of GE interaction variation is due to the $G \times L \times Y$ interaction (Table 7). The GL interaction only represents 23.9% of the GE interaction.

Table 8 includes the statistical values that were used to select between the best mixed models: (i) homogenous variance model for GE, (ii) heterogeneous across genotypes variance model for GE, and (iii) AMMI mixed model for GE. The best model for the joint analysis across the 6 yr of MET data was AMMI mixed model with two multiplicative terms (AMMI [2]).

Figure 3, obtained from the AMMI mixed model, was used to explore the GE interaction in the joint analysis across 6 yr. The mf480 and Florman genotypes made high contributions to the GE interaction; the first had important COI (Environments 1 and 4). Florman responded proportionally to the differences between Environment 1 and 4 environments in the majority of the years. The dispersion of the environmental scores, although they demonstrate a high interaction between years and locations, also show that Location 1 was one of the greater contributors to the GE interaction, especially in the 1997–1998 year because the mf480 and mf457 genotypes yielded better in this location than in others.

The GE interaction pattern could be difficult to repeat since groups are not visible between the environmental scores from the same location. The GL interaction is small in relation to the GLY interaction. Therefore, the GE interaction is not expected to be repeatable. The results suggest that this set of locations should be considered a unique megaenvironment for breeding objectives.

Figure 4 is a dispersion graph of the product of the genotype scores in the PC1 and PC2 of the AMMI model (2) versus the yield means for each genotype. It shows that Florman, mf480 and mf457 genotypes had lower stability. The mf484 and mf505 genotypes had the best performance across all environments.
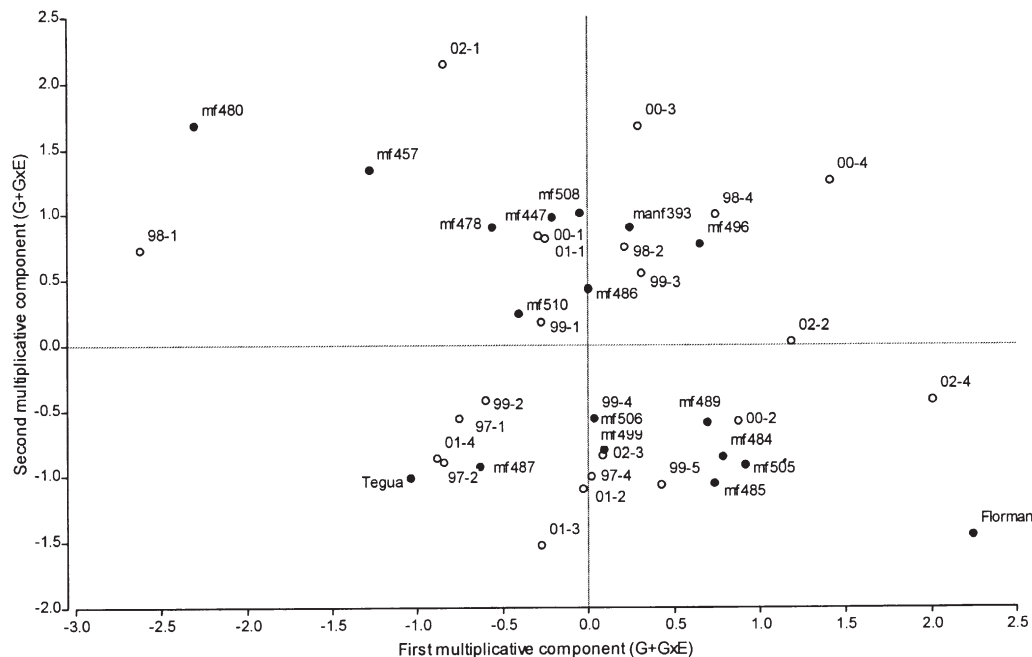


**Fig. 3. Biplot based on 6-yr multienvironment trials data for PBP-INTA. Dark points represent genotypes and light points represent environments coded according to the harvest year-location.**
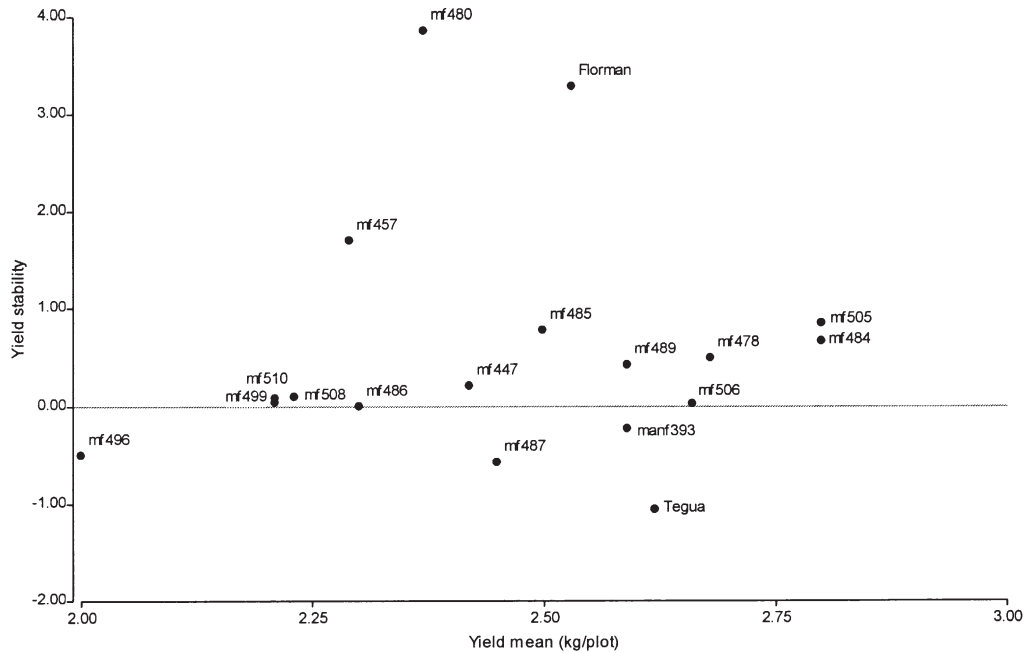
**Fig. 4. Yield stability versus yield mean.**

Yield means across all of the environments and the standard errors to evaluate mean differences, obtained from the AMMI[2] mixed model are shown in Table 9. These errors depend on the amount of information available for each genotype through the collection of analyzed MET and on the contribution of the genotype to the GE interaction (Balzarini, 2000). In addition, the table contains the coefficients of variation (CV) across environments and the stability variances (SV) (Shukla, 1972) obtained as covariance parameters of model 2. The stability variances for genotypes mf480, mf457, mf484, Florman, mf447, and manf393 were significantly different from zero ($P < 0.05$).

The main differences in yield stability across environments were observed in mf480, mf457, mf484, and Florman genotypes according to stability variance, which was consistent with the results obtained from the mixed AMMI biplot. The coefficient of variation fails to identify unstable genotypes that showed relatively high interaction with environments, e.g., mf480, mf478, manf393, and mf447, which are all short-cycle cultivars. Taking into account variability across environments, the cultivar with the best mean response was mf484.

## CONCLUSIONS

The by-year analysis of MET data showed that the variability due to the GL interaction in the PBP-INTA is relatively small in relation to the variability among genotypes. Although the GGE biplots obtained for each year of MET data allowed us to identify superior genotypes in some locations, the locations that were favorable in a particular year for some genotypes were not favorable through out the years. In 5 of the 6 yr of MET, the GGE PC1 biplot significantly correlated with yield means demonstrating proportional genotype responses across the locations; on occasion, it implied rank changes in genotype order. Using additional information of registered climatic contingencies in each year could explain the interaction. The results obtained from mixed AMMI models using data from 6 yr confirmed the random nature and the relatively low magnitude of the GE interaction in the PBP-INTA. Both strategies to evaluate MET data, the by-year analyses and the joint analysis indicate the existence of a unique megaenvironment for breeding purposes in the peanut crop area of Argentina. The results suggest that it is not necessary to partition the region into subregions to make cultivar recommendations. It also suggests that instead of increasing the number of locations where MET are conducted, it would be better to redirect the resources available toward implementing more efficient experimental designs.

**Table 9. Genotypic performance across 23 environments.**

| Genotype | Yield mean (kg plot$^{-1}$)† | SE‡ | SV§ | CV¶ |
|---|---|---|---|---|
| mf484 | 2.80a | 0.27 | 0.14 | 45.45 |
| mf505 | 2.66abc | 0.27 | 0.07 | 44.67 |
| mf478 | 2.61abde | 0.21 | 0.09 | 36.46 |
| mf489 | 2.59bcd | 0.26 | 0.07 | 47.37 |
| mf506 | 2.56bcdf | 0.25 | 0.01 | 43.98 |
| Florman | 2.53bcdef | 0.26 | 0.14 | 50.28 |
| manf393 | 2.53bcde | 0.22 | 0.11 | 38.14 |
| Tegua | 2.51bcdef | 0.25 | 0.07 | 46.71 |
| mf485 | 2.50defg | 0.27 | 0.10 | 52.64 |
| mf487 | 2.45bcdefg | 0.25 | 0.07 | 49.06 |
| mf486 | 2.43bcdef | 0.23 | 0.02 | 45.39 |
| mf510 | 2.38bcdef | 0.23 | 0.05 | 46.88 |
| mf499 | 2.36egh | 0.26 | 0.01 | 55.73 |
| mf447 | 2.35fh | 0.21 | 0.12 | 42.22 |
| mf508 | 2.33defgh | 0.23 | 0.09 | 50.05 |
| mf457 | 2.29cfgh | 0.22 | 0.35 | 42.38 |
| mf480 | 2.26cfgh | 0.21 | 0.47 | 36.64 |
| mf496 | 2.09h | 0.24 | 0.05 | 56.79 |

† Yield means across environments, different letters indicate statistically significant differences ($p < 0.05$).
‡ Mean difference standard errors (SE) used to mixed model.
§ Stability variances (SV) for each genotype obtained from a mixed model.
¶ Coefficient of variation (CV) across environments.

## REFERENCES

Allard, R.W., and A.D. Bradshaw. 1964. Implications of genotype-environmental interactions in applied plant breeding. Crop Sci. 4:503–508.

Balzarini, M. 2000. Biometrical models for predicting future performance in plant breeding. Ph.D. Diss., Louisiana State University, Baton Rouge, LA, USA.

Balzarini, M. 2002. Applications of mixed models in plant breeding. p. 353–363. *In* M.S. Kang (ed.) Quantitative genetics, genomics, and plant breeding. CABI Publishing, New York.

Cornelius, P.L., J. Crossa, and M.S. Seyedsadr. 1996. Statistical test and estimates of multivariate models for genotype-by-environment interaction. p. 199–234. *In* M.S. Kang and H.G. Gauch (ed.) Genotype-by-environment interaction. CRC Press, Boca Raton, FL

Francis, T.R., and L.W. Kannenberg. 1978. Yield stability studies in short season maize: I. Descriptive method for grouping genotypes. Can. J. Plant Sci. 58:1029–1034.

Gabriel, K.R. 1971. Biplot display of multivariate matrices with application to principal components analysis. Biometrika 58:453–467.

Gauch, H.G., Jr. 1988. Model selection and validation for yield trials with interaction. Biometrics 44:705–715.

Info-Gen. 2003. Info-Gen Versión 1. FCA, Universidad Nacional de Córdoba.

Jennrich, R.L., and M.D. Schluchter. 1986. Unbalanced repeated measures models with structured covariance matrices. Biometrics 42:805–820.

Kang, M.S. 2002. Genotype-environment interaction: Progress and prospects. p. 221–243. *In* M.S. Kang (ed.) Quantitative genetics, genomics, and plant breeding. CABI Publishing, New York.

Littell, R.C. G.A Milliken, W.W. Stroup, and R.D. Wolfinger. 1996. SAS® System for mixed models. SAS Institute Inc., Cary, NC.

Piepho, H.P. 1998. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. Theor. Appl. Genet. 97:195–201.

SAS Institute. 1997. SAS/STAT software: Changes and enhancements through release 6.12. SAS Inst., Cary, NC.

Shukla, G.K. 1972. Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29:237–245.

Stroup, W.W. 2000. Mixed models issues in the analysis of multilocation trials. p. 135–142. *In* The XXth International Biometrics Conference, Vol. II. University of California, Berkeley.

Yan, W., L.A. Hunt, Q. Sheng, and Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on GGE Biplot. Crop Sci. 40:597–605.

Yan, W., and L.A. Hunt. 2002. Biplot analysis of multi-environment trial data. p. 289–303. *In* M.S. Kang (ed.) Quantitative genetics, genomics, and plant breeding. CABI Publishing, New York.

Yan, W., and M.S. Kang. 2003. GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL.