Additive Main Effects and Multiplicative Interaction Analysis of Two International Maize Cultivar Trials

J. Crossa,* H. G. Gauch, Jr., and R. W. Zobel

ABSTRACT

The methodology used by the International Maize and Wheat Improvement Center (CIMMYT) to develop and improve its maize (*Zea mays* L.) germplasm involves evaluation of families or experimental varieties in extensive international testing trials. The genotype-environment interaction is produced by differential genotypic responses to varied environmental conditions. Its effect is to limit the accuracy of yield estimates and complicate the identification of specific genotypes for specific environments. The objective of this study was to use the Additive Main effects and Multiplicative Interaction (AMMI) method, with additive effects for genotypes and environments and multiplicative terms for genotype-environment interaction, for analyzing data from two international maize cultivar trials. Results from the first trial were: (i) predictive assessment selected AMMI with one principal component axis, (ii) AMMI increased the precision of yield estimates equivalent to increasing the number of replications by a factor of 4.30, (iii) AMMI provided much insight into genotype-environment interactions, and (iv) AMMI selected a different highest-yielding genotype than did treatment means in 72% of the environments. Results for the second trial were that predictive assessment selects the AMMI with none of the principal component axes, which increased precision equivalent to increase the number of replications by a factor of 2.59.

THE MAIZE PROGRAM of CIMMYT can be described as a multistage process with a continuous flow of germplasm. Gene pools are improved and their best fractions are further advanced to form populations (Pandey et al., 1986). The superior families from each germplasm pool or population are combined to form high-yielding experimental varieties. The major objective of the CIMMYT Maize Program is to develop widely adapted, high-yielding, stable genotypes with resistance to environmental stresses, including drought, heat, insects, and diseases.

Multilocation yield trials play a major role in this system of maize germplasm development. Extensive international trials of families and experimental varieties are carried out over a wide range of environments. Superior families and varieties are selected for two main purposes: (i) use within the maize breeding program of (ii) distribution to national programs for eventual use by farmers. A vital goal in breeding and agronomic research is to provide reliable guidance for selecting the best genotypes for planting in future years and at new sites, i.e., to predict yield as precisely as possible based on limited experimental data.

Given these objectives, statistical analyses may feature any one of three markedly different objectives, namely, (i) between-trial predictive success, (ii) within-trial predictive success, and (iii) within-trial postdictive success (Gauch, 1988). If between-trial prediction is the objective, trial data, perhaps with concomitant environmental data, would be used to construct a model, and inferences are made for other sites and years not included in the yield trial. The focus of this paper, however, will be the two within-trial objectives.

In within-trial postdiction, a statistical model is constructed for a data set and success is measured in terms of the model's ability to fit this same data set. To evaluate within-trial prediction, Gauch (1988) and Gauch and Zobel (1988) proposed that the data from within a yield trial be split into modeling data and validation data. They used the AMMI model, along with such a data splitting method, to analyze New York soybean [*Glycine max* (L.) Merr.] trial data. Success was measured in terms of the ability of the model fitted to the modeling data to predict the validation observations. They concluded that for those data, the number of multiplicative interaction terms should be one. Additional terms worsened predictive value rather than improving it. They also found that AMMI analysis with two replications was as precise as treatment (i.e., genotype × environment) means based on five replications.

In cooperation with national programs in developing countries, CIMMYT carries out a large number of multilocation yield trials with a great diversity of

J. Crossa, Biometrics and Statistics Unit, CIMMYT, Apartado Postal 6-641, 06600 Mexico D.F., Mexico; H.G. Gauch, Jr., Dep. of Agronomy, Cornell Univ.; and R.W. Zobel, USDA-ARS and Dep. of Agronomy and Plant Breeding, Cornell Univ., Ithaca, NY 14853. Received 28 June 1989. *Corresponding author.

Published in Crop Sci. 30:493-500 (1990).

germplasm and over a wide range of environmental conditions. The objective of this study was to examine the results of AMMI analysis applied to two CIMMYT maize cultivar trials conducted in 1984 and 1985.

MATERIALS AND METHODS

Data from two multilocation maize cultivar trials were analyzed in this study. Trial 1 had 17 genotypes (i.e., cultivars) (Table 1) grown in 36 environments in 1984. Trial 2 had nine genotypes (Table 2) evaluated in 38 environments in 1985. Both trials used a randomized complete-block design with four replications in each environment. Grain yield (kg ha^{-1}) was calculated using average shelling percentage of 80% and adjusted to 15% moisture. Details of the CIMMYT maize breeding program and field procedures are described by Vasal et al. (1982) and in publications by CIMMYT (1982, 1984). In this study the term *genotype* is used as a generic term for cultivar.

The reason for presenting results of two international trials is that they differed not only in the germplasm included, but in the test environments. Trial 1 was conducted under three different environmental conditions: tropical, subtropical, and temperate, whereas Trial 2 was only in subtropical environments.

Table 1. Genotypes included in Trial 1 grown in 36 environments in 1984.

Genotype code number	Genotype name†
1	Antalya 8233
2	Antalya 8233 (1)
3	Chuquisaca 8233
4	Rampur 8233 (1)
5	Tlaltizapan 8233
6	Capinopolis 8245
7	Islamabad 8245
8	Islamabad 8245 (1)
9	Sids 8245
10	Sida 8245 (1)
11	Taltizapan 8245
12	Across 8245
13	Pirsabak 8248
14	Pirsabak 8248 (1)
15	Tlaltizapan 8248
16	Across 7845
17	Across 7748

† Genotypes are identified by the population number from which the genotype was derived and by the year (XXYY, where XX is the year and YY is the population number) and the site where selection was performed. When the name of a genotype is followed by a number in parentheses, the selection was made by breeders in the national programs. The "*Across*" genotypes were selected as progenies that performed best across sites.

Table 2. Genotypes included in Trial 2 grown in 38 environments in 1985.

Genotype code number	Genotype name†
1	Ferkessedougou 8223 (1)
2	Across 8126
3	Across 8130
4	San Jeronimo 8232 (1)
5	Ferkessdougou 8235 (1)
6	Across 8235
7	Ikenne 8149 (1)
8	Across 7726
9	Pirsabak 7930 (1)

† Genotypes are identified by the population number from which the genotype was derived and by the year (XXYY, where XX is the year and YY is the population number) and site where selection was performed. When the name of a genotype is followed by a number in parentheses, the selection was made by breeders in the national programs. The "*Across*" genotypes were selected as progenies that performed best across sites.

Statistical Analysis

The AMMI analysis, performed using MATMODEL (Gauch, 1987), first fits additive effects for genotypes (G) and environments (E) by the usual additive analysis of variance procedure, and then fits multiplicative effects for genotype-environment (GE) interaction by principal components analysis (PCA). The AMMI model is

$$Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^{n} \lambda_k \alpha_{ik} \gamma_{jk} + R_{ij}$$

where Y_{ij} is the yield of the ith genotype in the jth environment; g_i is the mean of the ith genotype minus the grand mean; e_j is the mean of the jth environment minus the grand mean; λ_k is the square root of the eigenvalue of the PCA axis k; α_{ik} and γ_{jk} are the principal component scores for PCA axis k of the ith genotype and the jth environment, respectively, and R_{ij} is the residual. Environment and genotype PCA scores are expressed as unit vector times the square root of λ_k (i.e., environment PCA score = $\lambda_k^{0.5}\ \gamma_{jk}$; genotype PCA score = $\lambda_k^{0.5}\ \alpha_{ik}$) (Zobel et al., 1988).

The GE interaction sum of squares (SS) is subdivided into PCA axes where axis k is regarded as having $g+e-1-2k$ df, where g and e are the number of genotypes and environments, respectively, since this is the increase in the number of mathematically independent model parameters that results from incorporation of the kth PCA axis (Gollob, 1968). The model including one or more PCA axes is nonlinear in its parameters, so the allocation of df must be regarded as an approximation. A different method of allocating the df, not used here, has been suggested by Mandel (1971) and was used by Cornelius (1978) in a postdictive method for choosing a model for the analysis of an unreplicated yield trial of 49 maize cultivars grown at four plant densities.

Postdictive success was measured by approximate F-tests at the 0.05 probability level by comparing each principal component's mean square with the pooled within-environment error mean square. Those PCA axes that were not significant were pooled into a residual term.

Predictive assessment was carried out by the cross-validation procedure described by Gauch and Zobel (1988). The data were split into two subgroups, model data and validation data. For each treatment (i.e., genotype and environment combination), two replicates were selected at random to be modelled by AMMI, and the other two were reserved as validation observations. Thus, Trial 1 had $17 \times 36 \times 2 = 1224$ validation observations, and Trial 2 had $9 \times 38 \times 2 = 684$ validation observations.

The general principles of data splitting or cross-validation are set forth in the related literature on multiple regression (Laird and Cady, 1969; Allen, 1971; Anderson et al., 1972; Snee, 1977; Berk, 1984). Cross-validation methods have been used to select the optimal number of axes to retain in principal components analysis (Wold, 1978; Krzanowski, 1983). Using these methods, each observation is predicted by a model using the remainder of the data, and simulation of prediction is performed without requiring any new data. It can be argued that a single observation does not contribute much to the multivariate structure, so neither does its deletion remove such. On the other hand, it is expected that the deletion of one-half of a portion of the replications of a yield trial does affect the multivariate structure of the data. This suggests that a large number of random partitions of the trial data is not needed.

Several models were fitted to the model prediction data: additive model (AMMI0), which estimates the additive main effects (i.e., genotypes and environments) and retains none of the PCA axes; AMMI1, which combines the additive main effects from AMMI0 with the GE interaction effect

estimated from the first principal component axis (PCA 1) (this model retains one interaction PCA and relegates the rest to the residual); AMMI2 and so on, up to the full AMMI model (AMMIp, where p = min[g − 1, e − 1]) with all PCA axes. The predictive values from the full model are equal to the cell (i.e., genotype × environment) means, namely, the average of the two replications selected at random for modeling; thus, the full model will be characterized as the CELL MEANS model.

Predicted values for each model were compared with the validation data by computing the sum of squared differences. For example, the model predicted value of Genotype 1 in Environment 1 was compared with the corresponding validation observations. The differences between the predicted value and validation observations were squared and summed over all genotypes and environments. This sum of squared differences was divided by the number of validation observations and the square root was taken to give the root mean square predictive difference (RMS PD). A small value of RMS PD indicates good predictive success. This procedure was repeated 10 times, using different randomizations, and the results averaged. Trial 1 and 2 data were also randomly split 15, 30, 50, 70, and 100 different times; the calculated average RMS PD for each model was similar to the one obtained for 10 random splittings. Therefore, using 10 different random splittings of the data appears to be adequate for cross-validation.

In the process of comparing model prediction with validation data, errors are present both in the model's predictions and in the validation data when estimating the true mean (Gauch and Zobel, 1988). A useful criterion to measure the prediction error is the mean square error (MSE) of the estimate (Cochran, 1963). Therefore,

$$\text{MSE(model} - \text{validation)} = \text{MSE(model)} + \text{VAR(validation data)} \quad [1]$$

$$\text{where MSE(model)} = \text{Var(model)} + (\text{Bias})^2 \quad [2]$$

See Appendix for development of Eq. [1] and [2].

The Var(validation data) is the mean square error of a single replication, while MSE(model) is based on r_m replicates (where r_m = number of replicates used for modeling). According to Gauch and Zobel (1988), the theoretical question of basing prediction on the mean of r_l replicates to equal the performance of the best predictive model based on r_m can be approximated by r_l = Var(validation data)/MSE(model). Because the best predictive model is based on r_m replicates, the approximate gain factor can be defined as $GF = r_l/r_m$.

For both trial data sets, careful examination of the residuals from the recommended models indicated that they were approximately normally distributed.

RESULTS

Trial 1

Additive main effects and multiplicative interaction analysis showed that environments, genotypes, and GE interaction were highly significant ($P < 0.01$) and accounted for 75, 11, and 14% of the treatment combinations SS, respectively (Table 3). The criterion of postdictive success for AMMI using all the data (all four replications) and *F*-tests at the 0.05 probability level recommended including the first five interaction PCA axes in the model (Table 3).

However, predictive assessment, measured by RMS PD, selected AMMI1 with the first interaction PCA axis as most predictively accurate in each of the 10 different random partitions of the data. The AMMI1 model had the lowest average RMS PD (906.46 kg ha^{-1}) (Table 4). This model has 101 df (35 for environments plus 16 for genotypes plus 50 for interaction PCA axis 1) and is 2.8 times as parsimonious (few df) as AMMI5, which was selected for postdictive success (i.e., AMMI5 contains 2.8 times as many df as AMMI1).

Table 3 indicates that the AMMI1 model captures 93% of the treatment combinations SS, and the model-validation procedure (Table 4) indicates that the 7% of the treatment combinations SS remaining in the residual is of no predictive value.

A final model may be constructed applying AMMI1 to all the data (all four replications). In its first interaction PCA axis, this model recovers more than half of the GE interaction SS (54.6%) in only 8.9% of the interaction df (Table 3). The higher interaction PCA axes are judged by predictive assessment to be just noise for the purpose of yield prediction, and thus may be pooled with the residual.

An interesting question concerns the number of replicates within an environment needed to equal the precision obtained by using the AMMI1 model. The quantity of greatest interest is the MSE(model) [1] and it can be estimated as the difference between MSE(model − validation) and Var(validation data). The MSE(model − validation) is the square of the RMS PD for AMMI1 (906.46 kg ha^{-1}, Table 4). The Var(validation data) is simply the pooled error mean square (ignoring replications within environments) from the AMMI analysis of variance, which may be expressed in terms of its square root as 858.01 kg ha^{-1} (Table 3). Hence for AMMI1, MSE(model) is esti-

Table 3. Additive main effects and multiplicative interaction analysis of variance for grain yield (kg ha^{-1}) for Trial 1 including the first five interaction principal component analysis (PCA) axes.

Source of variation	df	Sum of squares	Mean squares
		× 10^{-3}†	
Treatment combinations	611	6 310 687	10 329**
Genotype (G)	16	687 747	42 984**
Environment (E)	35	4 717 825	134 795**
GE	560	905 299	1 617**
Interaction PCA 1	50	493 869	9 877**
Residual + PCA 2 to PCA 5	510	411 430	807**
Interaction PCA 2	48	80 290	1 673**
Interaction PCA 3	46	57 139	1 242**
Interaction PCA 4	44	53 437	1 214**
Interaction PCA 5	42	46 202	1 100*
Residual	330	174 360	528
Pooled error	1 836	1 351 633	736

*,** Significant at the 0.05 and 0.01 probability levels, respectively.
† Multiply the reported value by this to get the actual number.

Table 4. Average root mean square predictive difference (RMS PD) for seven models constructed based on grain yield data (kg ha^{-1}) from Trial 1.

Model	df	RMS PD
AMMI0	51	983.84
AMMI1	101	906.46
AMMI2	149	937.82
AMMI3	195	963.22
AMMI4	239	986.27
AMMI5	281	996.80
CELL MEANS†	611	1054.77

† Full model based on all genotype × environment combinations.

mated as $[(906.46)^2 - (858.01)^2]$, which expressed in terms of its square root is 292.41 kg ha^{-1}. This represents an error of only 6.2% relative to the grand yield mean for the trial.

The approximate number of replicates needed for the CELL MEANS model to equal the performance of AMMI1 is $r_i = (858.01/292.41)^2 = 8.61$. Because AMMI1 based on two replicates is as precise as CELL MEANS based on 8.61 replicates, the estimated statistical GF is 4.30.

Given these results, it can be argued that for Trial 1 the AMMI1 fitted values are more precise than CELL MEANS in estimating yield. It is of interest to compare, for each environment, the genotype ranked first in AMMI1, with the one ranked first in the CELL MEANS. Results showed that the use of AMMI1 leads to a different selection in 72% of the environments (i.e., in 26 environments AMMI1 and CELL MEANS picked different winners).

Biplot Display

In the biplot (Fig. 1), showing main effects means on the abscissa and PCA 1 values as the ordinates, genotypes (or environments) that appear almost on a perpendicular line have similar means and those that fall almost on a horizontal line have similar interaction patterns. Genotypes (or environments) with large PCA 1 scores (either positive or negative) have high interactions, whereas genotypes (or environments) with PCA 1 scores near zero have small interactions.

As pointed out by Zobel et al. (1988), the AMMI1 expected yield for any genotype and environment combination can be calculated from Fig. 1. The additive AMMI0 part of the AMMI model is simply the genotype mean plus the environment mean minus the grand mean. The interaction part is simply the genotype PCA score times the environment PCA score. These two parts are added to produce the expected value of the AMMI1 model. Genotypes and environments with PCA 1 scores of the same sign produce positive interactions effects, whereas combination of PCA 1 scores of opposite signs have negative specific interactions.

Three groupings of genotypes are evident from Fig. 1:

Group 1 includes Genotypes 13, 14, 15, and 17. They are all selections derived from Population 48 ('Compuesto de Hungria'). They show a similar mean yield response (below the grand mean) and a similar, large negative interaction. For these genotypes, the AMMI1 model predicts genotype yields that are close to those of the AMMI0 model in environments with PCA scores near zero, larger yields than the AMMI0 model in environments with negative environment PCA scores, and smaller yields than the AMMI model in environments with positive environment PCA scores.

Group 2 consists of Genotypes 1, 2, 3, 4, and 5. These are selections out of Population 33 ('Amarillo Subtropical') in 1982. Their mean yields are similar, but their interactions with environment differ. The interaction PCA 1 score for Genotypes 1, 2, and 4 are positive and large; for Genotypes 3 and 5 they are positive and smaller.

Group 3 consists of Genotypes 6 to 12 and 16. They are genotypes formed from Population 45 ('Amarillo Bajio'). They show high mean yield response. Genotypes 6 and 12 have the smallest interactions, and are therefore the most stable genotypes. The yield stability of a group of CIMMYT's subtropical maize populations, measured by the performance of genotypes de-

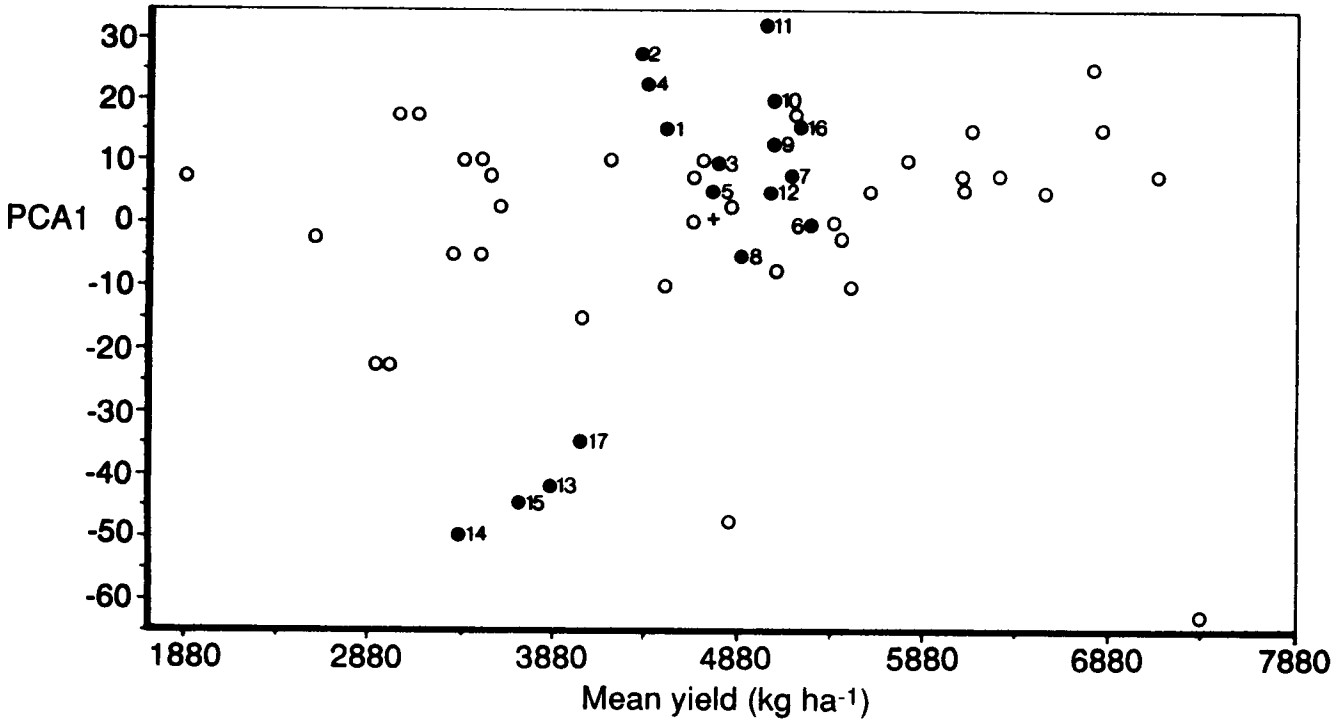

Fig. 1. Biplot of the yield means and the first principal component axis scores of 17 maize genotypes (●) and 36 environments (○) from Trial 1. The grand yield mean of the experiment is represented by "+".

rived from them show that Population 45 produces many genotypes that give stable yields under both favorable and unfavorable environmental conditions (Crossa et al., 1989).

Populations 33, 45, and 48 are all adapted to the subtropics, and all have short or intermediate maturity and yellow grain color. However, their different genetic compositions cause different interaction patterns and yield responses in genotypes derived from them. Population 33 was derived from subtropical, intermediate maturity, yellow flint germplasm. It has medium to short plant height and has been selected for ear rot resistance. Population 45, with a broad germplasm base, was derived from crosses among lowland tropical maize types from Mexico and the Caribbean Islands and from Corn Belt dents from the USA. Population 48 was formed using temperate germplasm from the central USA Corn Belt and southern Europe.

The environments show much variability in both main effects and interactions (Fig. 1). Two sites in Greece had the greatest negative PCA 1 scores. Nea Zoe and Thessaloniki, with PCA scores of −64.72 and −48.56, respectively, were quite different in their main effects. Both sites are located at about 40 ° N lat. At the other extreme of the PCA axis, the two locations with the greatest positive PCA scores are Iboperenda in Bolivia (with PCA score 24.88), located at 20 ° S lat., and Duc Trong Farm in Vietnam (with PCA score 18.70), located at 20 to 25 ° N lat. Interaction PCA axis 1 arranges the environments in a sequence from subtropical (positive PCA) to temperate locations (negative PCA).

Several cycles of selection of Population 48 have been conducted in Turkey under temperate environmental conditions similar to those in Greece. The biplot diagram shows that the two sites in Greece have a positive interaction with temperate genotypes derived from Population 48 (Group 1). Iboperenda and Duc Trong Farm, on the other hand, favor subtropically adapted germplasm like Genotypes 2 and 11 (at the top of Fig. 1) derived from Populations 33 and 45, respectively.

Hence, the AMMI interaction PCA 1 axis has a consistent and agronomically important interpretation in terms of a subtropical to temperate gradient in both environments and adaptation of genotypes.

Trial 2

Additive main effects and multiplicative interaction analysis of variance partitioned the treatment SS into additive genotype and environment effects, and nonadditive GE interaction effects. These sources were all significant at the 0.01 probability level and accounted for 2, 91, and 7% of the treatment combinations SS, respectively (Table 5). Note that the environment effect dominated the analysis, and although the GE interaction SS was about 3.8 times larger than the genotype SS, it was still much smaller than the GE interaction SS found in Trial 1 (only 7% of the treatment combinations SS instead of 14%).

Postdictive assessment by *F*-tests at the 0.05 probability level recommended AMMI4, that is, inclusion in the model of four interaction PCA axes (Table 5).

Table 5. Additive main effects and multiplicative interaction analysis of variance for grain yield (kg ha^{-1}) for Trial 2 including the first five interaction principal component analysis (PCA) axes.

Source of variation	df	Sum of squares	Mean squares
		× 10^{-3}†	
Treatment combinations	341	3 531 530	10 356**
Environment (E)	37	3 216 475	86 932**
Genotype (G)	8	65 049	8 131**
GE	296	250 006	845**
Interaction PCA 1	44	73 182	1 663**
Interaction PCA 2	42	44 785	1 066**
Interaction PCA 3	40	37 714	943**
Interaction PCA 4	38	29 552	778*
Interaction PCA 5	36	23 866	663
Residual	96	40 906	426
Pooled error	1026	564 284	550

*,** Significant at the 0.05 and 0.01 probability levels, respectively.
† Multiply the reported value by this to get the actual number.

Table 6. Average root mean square predicitive difference (RMS PD) for seven models constructed based on grain yield data (kg ha^{-1}) from Trial 2.

Model	df	RMS PD
AMMI0	45	810.15
AMMI1	89	834.84
AMMI2	131	854.64
AMMI3	171	867.58
AMMI4	209	875.80
AMMI5	245	885.40
CELL MEANS†	341	897.25

† Full model based on all genotype × environment combinations.

Predictive success, as measured by the average RMS PD for 10 random partitions, recommended the additive AMMI0, which may be considered a subcase of AMMI, namely, zero interaction PCA axes (Table 6). The predictive model with 45 df (8 for genotypes + 37 for environments) is 4.6 times as parsimonious as the postdictive model with four PCA axes. The interaction is small, only 7% of the treatment combinations SS, and does not contribute to predictive success.

The Var(validation data) calculated in the AMMI analysis of variance, ignoring replications within environments, is 550 000 (Table 5) or expressed in terms of its square root is 741.61 kg ha^{-1}. The mean squared difference between AMMI0 predictions and validation observations, MSE(model − validation) is $(810.15)^2$ (Table 6). However, most of this variance comes from the validation observations, and only a small part from the AMMI0 model itself. Recalling that the variance of validation observations is 550 000, and removing this from RMS PD, it turns out that the estimated root mean square error of AMMI0 is 326.13 kg ha^{-1}. Therefore, the approximate number of replicates needed in order for the CELL MEANS model to equal the performance of the AMMI0 model with only two replicates is $r_1 = (741.61/326.13)^2 = 5.17$, giving a theoretical gain factor of 2.59. The AMMI0 model with two replicates is predictively superior to the CELL MEANS model based on five replicates. The approximate gain factor of 2.59 implies that if CELL MEANS had been used, approximately 2.59 times as many replicates would have been required to detect differences of the same magnitude as in AMMI0.

DISCUSSION

The AMMI model used previously in a New York soybean trial (Zobel et al., 1988; Gauch and Zobel, 1988) was applied on a larger scale to two of CIMMYT's international maize trials.

Results of AMMI analysis of Trial 1 reflect the expected GE interaction for temperate vs. tropical environments and demonstrate the value of this method (within-trial prediction success) when GE interaction is associated with environment specific factors. However, when most factors causing GE interaction are associated with unpredictable weather, such as variable rainfall and temperature, environmental data should be used to construct a model, so that inferences can be made for other sites and other years not included in the yield trial. This last objective is termed between-trial prediction success or agrotechnology transfer (Gauch, 1988).

Trial 1, involving a wider range of environments (tropical, subtropical, and temperate) than trial 2, had a larger theoretical gain in precision (GF of Trial 1 is 4.30, whereas GF for Trial 2 is 2.59) and a more complex best predictive model than Trial 2. The AMMI analysis applied to a soybean yield trial resulted in a GF of 2.5 (Gauch and Zobel, 1988) and AMMI with one multiplicative factor (AMMI1) as the best predictive model. The AMMI analysis of Trial 2 showed AMMI0 to be the best predictive model with the same GF as the soybean trial data.

Additive main effects and multiplicative interaction analysis has also been used in CIMMYT's international wheat (*Triticum aestivum* L. em Thell) yield trials, and results (not presented here) showed AMMI with either two, three, or four PCA axes to be the best predictive model.

Trial 2 exhibits an outcome expected from theory. Sometimes AMMI's role will be to diagnose a particular submodel that provides a more appropriate analysis (Zobel et al., 1988). By using the AMMI model for an initial analysis, it is possible to determine the submodel that is best for a particular data set.

These results indicate that the number of terms to be included in an AMMI model cannot be specified a priori, without first trying AMMI predictive assessment with the estimation of RMS PD for each model. In general, factors like type of crop, diversity of the germplasm, and range of environmental conditions will affect the degree of complexity of the best predictive model.

The theoretical gain in precision achieved by using AMMI predictive assessment should be regarded as an approximation since bias present in both the MSE(model) and MSE(model − validation) cannot be estimated and the number of replicates needed for the CELL MEANS model to equal the performance of the AMMI model is calculated based on the error mean square when all data are used.

In Trial 2, the gain in precision by using AMMI analysis was not impressive. The AMMI0 with two replicates was as good as the CELL MEANS model with $r_1 = 5.17$ replicates. The predicted value of AMMI0 model in the absence of interaction is:

$$\text{AMMI0} = \bar{Y}_{..} + \hat{g}_i + \hat{e}_j,$$

where $\hat{g}_i = \bar{Y}_{i.} - \bar{Y}_{..}$ and $\hat{e}_j = \bar{Y}_{.j} - \bar{Y}_{..}$ are the mean deviations of the ith genotype and the jth environment, respectively, and $\bar{Y}_{..}$ is the grand mean. Assuming homogeneity of variances the terms $\bar{Y}_{..}$, $\hat{g}_i$, and $\hat{e}_j$ are uncorrelated. Therefore, for a random model, the variance of AMMI0 is

$$\text{Var(AMMI0)} = \text{Var}(\bar{Y}_{..}) + \text{Var}(\hat{g}_i) + \text{Var}(\hat{e}_j)$$

where $\text{Var}(\bar{Y}_{..}) = \sigma^2/r_m ge$, $\text{Var}(\hat{g}_i) = \sigma^2 (g - 1)/r_m ge$, $\text{Var}(\hat{e}_j) = \sigma^2(e - 1)/r_m ge$, and $r_m = 2$, $g = 9$, and $e = 38$.

$$\begin{aligned}\text{Thus, Var(AMMI0)} &= \sigma^2 (e + g - 1)/r_m eg \\ &= \sigma^2(38 + 9 - 1)/(2)(38)(9) \\ &= \sigma^2/14.87.\end{aligned}$$

This implies that in the absence of interaction, AMMI0 with two replicates is expected to be as good as the CELL MEANS model based on 14 replicates within a location. It can be argued that the failure of the AMMI0 model to do that well ($r_1 = 5.17$) may be due, in part, to the real but unpredictable GE interaction.

Data from cultivar yield trials are used to estimate genotype yield and to make selections. More precise yield estimates will increase the probability of making successful selections. Gauch and Zobel (1989) have used order statistics to quantify the relationship between precision in yield estimates and selection. They found in the soybean trial that: (i) AMMI analysis significantly improved the probability of successful selection and (ii) the CELL MEANS model rankings and AMMI1 rankings picked different winners in slightly over half of the environments. In Trial 1, the CELL MEANS model and AMMI1 model rank genotypes differently in 72% of the environments.

The most difficult question is which genotype selection to prefer when CELL MEANS and AMMI predictions disagree. Our recommendation is to prefer AMMI rankings when the discrepancies are moderate and readily attributed to random statistical variation. When the discrepancy is large, however, the principal inference is not that the AMMI model is best or that the CELL MEANS model is best, but rather that this particular treatment (i.e., genotype and environmental combination) mean estimation is unreliable. If possible, more data should be sought, or perhaps the field data records should be inspected carefully for the source or probable cause of unusually large deviations.

Based on the results from soybean and maize trials, it can be suggested that in plant breeding the approximate gain in precision achieved with AMMI provides a tool for selecting better genotypes and therefore achieving higher realized progress from selection. Agronomic predictive assessment with AMMI can be used for statistical analysis of on-farm experiments.

However, more research is needed to determine the general usefulness of the AMMI model for analyzing yield trial experiments in agronomy and plant breeding. In particular, further investigation is required to: (i) quantify the probability of successful selection of a genotype when using AMMI predictive values, compared with the probability of selection based on the predictive value of Finlay-Wilkinson (1963) regression and on treatment means and (ii) compare AMMI

analysis with other traditional multivariate techniques, such as cluster analysis, principal coordinate analysis, and factor analysis.

ACKNOWLEDGMENTS

The authors would like to thank the numerous cooperators in national maize research programs that carried out the variety trials. The authors also appreciate the contributions of Dr. Paul L. Cornelius who offered valuable comments that significantly improved the manuscript.

APPENDIX

Let

y_{ijk} = the observed value of the ith genotype in the kth replicate at the jth environment,
$E(y_{ijk}) = \mu_{ij}$,
$\hat{\mu}_{ij}$ = model estimate of μ_{ij},
$\tilde{y}_{ijk}$ = validation observation,
$E(\hat{\mu}_{ij}) = m_{ij}$,

where $i = 1,2,\ldots g$ = number of genotypes,
$j = 1,2,\ldots e$ = number of environments, and
$k = 1,2,\ldots r_m$ = number of replicates used for modeling.

Define:

$$\text{MSE(model} - \text{validation)} = \text{MSE}(\hat{\mu}_{ij} - \tilde{y}_{ijk}) = \frac{1}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m} E(\hat{\mu}_{ij} - \tilde{y}_{ijk})^2 = E\left[\frac{1}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m}(\hat{\mu}_{ij} - \tilde{y}_{ijk})^2\right].$$

Note that the quantity in brackets is (RMS PD)^2. Further define:

$$\text{MSE(model)} = \text{MSE}(\hat{\mu}_{ij}) = \frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e} E(\hat{\mu}_{ij} - \mu_{ij})^2 = E\left[\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - \mu_{ij})^2\right],$$

and

$$\text{BIAS} = E(\hat{\mu}_{ij}) - \mu_{ij} = m_{ij} - \mu_{ij}.$$

Then,

$$\text{MSE}(\hat{\mu}_{ij} - \tilde{y}_{ijk}) = E\left[\frac{1}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m}(\hat{\mu}_{ij} - \tilde{y}_{ijk})^2\right]$$

$$= E\left\{\frac{1}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m}[(\hat{\mu}_{ij} - \mu_{ij}) - (\tilde{y}_{ijk} - \mu_{ij})]^2\right\}$$

$$= E\left\{\frac{1}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m}[(\hat{\mu}_{ij} - \mu_{ij})^2 - 2(\hat{\mu}_{ij} - \mu_{ij})(\tilde{y}_{ijk} - \mu_{ij}) + (\tilde{y}_{ijk} - \mu_{ij})^2\right\}$$

$$= E\left[\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - \mu_{ij})^2 - \frac{2}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m}(\hat{\mu}_{ij} - \mu_{ij})(\tilde{y}_{ijk} - \mu_{ij}) + \frac{1}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m}(\tilde{y}_{ijk} - \mu_{ij})^2\right].$$

Since the validation observations $\tilde{y}_{ijk}$ are uncorrelated with the model estimates, $\hat{\mu}_{ij}$, the cross-products term has expectation zero. Therefore,

$$\text{MSE}(\hat{\mu}_{ij} - \tilde{y}_{ijk}) = E\left[\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - \mu_{ij})^2\right] + E\left[\frac{1}{\text{ger}_m}\sum_{i=1}^{g}\sum_{j=1}^{e}\sum_{k=1}^{r_m}(\tilde{y}_{ijk} - \mu_{ij})^2\right] = \text{MSE}(\hat{\mu}_{ij}) + \text{Var}(\tilde{y}_{ijk}) = \text{MSE(model)} + \text{Var(validation)}.$$

Now, if the amount of bias is $m_{ij} - \mu_{ij}$, then

$$\text{MSE(model)} = \text{MSE}(\hat{\mu}_{ij}) = E\left[\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - \mu_{ij})^2\right] =$$

$$= E\left\{\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}[(\hat{\mu}_{ij} - m_{ij}) + (m_{ij} - \mu_{ij})]^2\right\} =$$

$$= E\left\{\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}[(\hat{\mu}_{ij} - m_{ij})^2 + 2(\hat{\mu}_{ij} - m_{ij})(m_{ij} - \mu_{ij}) + (m_{ij} - \mu_{ij})^2]\right\}$$

$$= E\left[\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - m_{ij})^2 + \frac{2}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - m_{ij})(m_{ij} - \mu_{ij}) + \frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(m_{ij} - \mu_{ij})^2\right].$$

Taking expectations gives

$$\text{MSE}(\mu_{ij}) = E\left[\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - m_{ij})^2\right] + E\left[\frac{2}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(\hat{\mu}_{ij} - m_{ij})(m_{ij} - \mu_{ij})\right] + E\left[\frac{1}{ge}\sum_{i=1}^{g}\sum_{j=1}^{e}(m_{ij} - \mu_{ij})^2\right] = \text{Var}(\hat{\mu}_{ij}) + (\text{BIAS})^2$$

since the cross-product terms $E[2(\hat{\mu}_{ij} - m_{ij})(m_{ij} - \mu_{ij})] = 2(m_{ij} - \mu_{ij})E(\hat{\mu}_{ij} - m_{ij}) = 0$, since m_{ij} and μ_{ij} are fixed parameters and $E(\hat{\mu}_{ij} - m_{ij}) = 0$.

REFERENCES

Allen, D.M. 1971. Mean square error of prediction as a criterion for selecting variables. Technometrics 13:469–475.

Anderson, R.L., D.M. Allen, and F.B. Cady. 1972. Selection of predictor variables in linear multiple regression. p. 3–17. T.A. Bancroft (ed.) *In* Statistical Papers in Honor of George W. Snedecor. The Iowa State Univ. Press, Ames.

Berk, K.N. 1984. Validating regression procedures with new data. Technometrics 26:331–338.

CIMMYT. 1982. CIMMYT's maize program: An overview. CIMMYT, El Batan, Mexico.

CIMMYT. 1984. CIMMYT report on maize improvement 1980–1981. CIMMYT, El Batan, Mexico.

Cochran, W.G. 1963. Sampling techniques. 2nd ed. John Wiley & Sons, Inc., John Wiley & Sons, New York.
Cornelius, P.L. 1978. Empirical models for the analysis of unreplicated lattice-split-plot cultivar trials. Crop Sci. 18:627–633.
Crossa, J., B. Westcott, and C. Gonzalez. 1989. The yield stability of CIMMYT's maize germplasm. Euphytica 40:245–251.
Finlay, K.W., and G.N. Wilkinson. 1963. An analysis of adaptation in a plant breeding programme. Aust. J. Agric. Res. 14:742–754.
Gauch, H.G. 1987. MATMODEL. Microcomputer Power, Ithaca, NY.
Gauch, H.G. 1988. Model selection and validation for yield trials with interaction. Biometrics 44:705–715.
Gauch, H.G., and R.W. Zobel. 1988. Predictive and postdictive success of statistical analyses of yield trials. Theor. Appl. Genet. 76:1–10.
Gauch, H.G., and R.W. Zobel. 1989. Accuracy and selection success in yield trial analyses. Theor. Appl. Genet. 77:473–481.
Gollob, H.F. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33:73–115.
Krzanowski, W.J. 1983. Cross-validatory choice in principal components analysis—Some sampling results. J. Statist. Computation and Simulation. 18:299–314.
Laird, R.J., and F.B. Cady. 1969. Combined analysis of yield data from fertilizer experiments. Agron. J. 69:829–834.
Mandel, J. 1971. A new analysis of variance model for non-additive data. Technometrics 13:1–18.
Pandey, S., A.O. Diallo, T.M.T. Islam, and J. Deutsch. 1986. Progress from selection in eight tropical maize populations using international testing. Crop Sci. 26:879–884.
Snee, R.D. 1977. Validation of regression models: Methods and examples. Technometrics 19:415–428.
Vasal, S.K., A. Ortega, and S. Pandey. 1982. CIMMYT's maize germplasm management improvement, and utilization program. CIMMYT, El Batan, Mexico.
Wold, S. 1978. Cross-validatory estimation of the number of components in factor and principal components models. Technometrics 20:397–405.
Zobel, R.W., M.J. Wright, and H.G. Gauch, Jr. 1988. Statistical analysis of a yield trial. Agron. J. 80:388–393.