AMMI adjustment for statistical analysis of an international wheat yield trial

J. Crossa [1]*, P. N. Fox [1], W. H. Pfeiffer [1], S. Rajaram [1] and H. G. Gauch, Jr. [2]

[1] International Maize and Wheat Improvement Center (CIMMYT) Apdo. Postal 6-641, 06600 Mexico D.F., Mexico
[2] Department of Agronomy, Cornell University, Ithaca, NY 14853, USA

Summary. Multilocation trials are important for the CIMMYT Bread Wheat Program in producing high-yielding, adapted lines for a wide range of environments. This study investigated procedures for improving predictive success of a yield trial, grouping environments and genotypes into homogeneous subsets, and determining the yield stability of 18 CIMMYT bread wheats evaluated at 25 locations. Additive Main effects and Multiplicative Interaction (AMMI) analysis gave more precise estimates of genotypic yields within locations than means across replicates. This precision facilitated formation by cluster analysis of more cohesive groups of genotypes and locations for biological interpretation of interactions than occurred with unadjusted means. Locations were clustered into two subsets for which genotypes with positive interactions manifested in high, stable yields were identified. The analyses highlighted superior selections with both broad and specific adaptation.

Key words: Wheat – Multilocation trials – Predictive accuracy – Cluster analysis – Stability

Introduction

The Bread Wheat Breeding Program of the International Maize and Wheat Improvement Center (CIMMYT) aims to develop widely adapted, high-yielding, stable germ plasm with adequate industrial quality and resistance to combinations of disease and abiotic stresses (Rajaram et al. 1984). Cooperative multilocation trials, managed by national programs, are fundamental in developing and identifying such germplasm.

Interpretation of performance of a number of genotypes evaluated in a broad range of environments is always affected by genotype × environment interaction (*GE*). Differential genotypic responses to variable environmental conditions, especially associated with changes in ranks of genotypes, limit accurate yield estimates and identification of superior, stable genotypes.

Regression and other methods for partitioning the *GE* interaction sum of squares are discussed by Gauch (1988).

In multivariate analysis, the response of a genotype in *E* different environments may be conceptualized as a pattern in *E* dimensional space, with the coordinate of an individual spatial axis being the yield of the genotype in one environment.

Multivariate methods can facilitate interpretation of multilocation genotype trials. Ordination techniques such as principal components analysis, principal coordinates analysis, and factor analysis often present a large percentage of the original *E* dimensional variation in a few dimensions, and thus may simplify interpretation of *GE* structure. Often, complex relationships among locations or among genotypes can be adequately represented in a scattergram (Freeman and Dowker 1973; Westcott 1987). Alternatively, cluster analysis has been used to group locations that discriminate among genotypes in a similar manner or to summarize patterns of genotypic performance across environments (Abou-El-Fittouh et al. 1969). The combination of ordination and cluster analysis is termed pattern analysis (Byth et al. 1976; Shorter et al. 1977).

The Additive Main effects and Multiplicative Interaction (AMMI) procedure with prediction assessment was proposed for analysis of two-way tables (Gauch 1988; Gauch and Zobel 1988). The method integrates additive main effects and multiplicative components, extracting first the additive main effects and then using

* To whom correspondence should be addressed

principal components analysis to investigate the *GE*. With the biplot facility from AMMI analysis, both genotypes and locations occur on the same scattergram, and inferences about specific genotype × location combinations can be made. In plant breeding, the method has focused upon accuracy of estimates of genotypic yields in across-location trials. This has been referred to as yield prediction assessment and functions by splitting the data into modelling and validation data, with the values expected by the model compared with the validation data (Gauch and Zobel 1988).

In this study, complementary statistical procedures were applied to a hexaploid wheat (*Triticum aestivum* L. em Thell.) yield trial to: (1) assess the success of genotypic yield prediction, (2) form homogeneous subgroups of locations and genotypes, and (3) identify stable, high-yielding genotypes.

Materials and methods

Grain yield data were analyzed from the eighth Elite Selection Wheat Yield Trial (ESWYT) in which 18 bread wheat genotypes (Table 1) were tested under rain-fed and irrigated conditions. A total of 25 locations spanning 15 countries was used (Table 2). A randomized complete block design with three replicates was employed, and the plot size recommended to cooperators was six 3-m rows sown 30 cm apart, with the central four rows harvested. Variable local checks, which were not always identified, and the tetraploid durum wheat check were excluded.

Statistical analysis

AMMI analysis. For genotypic yields in across-location trials, prediction assessment was conducted using the AMMI method (Bradu and Gabriel 1978; Gauch 1988; Gauch and Zobel 1988). The AMMI model is

$$Y_{ij} = u + g_i + l_j + \sum_{k=1}^{n} \lambda_k \alpha_{ik} \gamma_{jk} + E_{ij} ,$$

where Y_{ij} is the yield of the ith genotype in the jth environment; u is the grand mean; g_i and l_j are the genotype and location deviations from the grand mean, respectively; λ_k is the eigenvalue of the principal component analysis axis k; α_{ik} and γ_{jk} are the genotype and location principal component scores for axis k; n is the number of principal components retained in the model; and E_{ij} is the error term.

Data were split into two subgroups: modelling and validation data. For each combination of genotype and location, two randomly chosen replicates were used to estimate parameters of the AMMI model, and the other replicate was used to validate the model. Seven models were fitted to the data. The first was the AMMI0 model, which estimated the additive main effects (i.e., genotypes and locations) without considering interaction; the second, AMMI1, combined the main effects from AMMI0 with interaction effects estimated from the first interaction principal component axis (PCA1). The third model, AMMI2, considered main effects plus two interaction principal components. AMMI3 – AMMI5 included cumulatively one more interaction principal component axis. The seventh model (DATA) with 17 PCA axes was the full model, which completely specified the data matrix and equalled the average of the two replicates selected at random for modelling.

Table 1. Genotypes of the eight ESWYT

Number	Cross or name
1	PRL/VEE#6
2	CN079*2/PRL
3	BUC/GLEN
4	KAUZ#1
5	KAUZ#2
6	RRV/WW15/3/BJ/2*ON//BON/4/NAC
7	GENARO 81
8	AGA/3*YR
9	SERI 82
10	BUC/PVN
11	KEA/TOW
12	GLENNSON 81
13	MON/CROW
14	BJY/COC
15	HAHN*2/PRL
16	R37/GHL121//KAL/BB/3/KLT
17	VEE#5/PVN
18	HD2206/HORK

Table 2. Locations from the eighth ESWYT

Code	Country	Location	Latitude (°N)	Elevation (masl)
AK	Algeria	El Khroub	36	640
AL	Algeria	Setif	36	1,023
BJ	Bangladesh	Joydebpur	24	8
CA	Cyprus	Athalassa	35	142
EG	Egypt	El Gemmeiza	31	8
ES	Egypt	Sakha	31	6
EB	Egypt	Beni-Suef	29	28
IL	India	Ludhiana	31	247
ID	India	Delhi	29	228
JM	Jordan	Madaba	36	785
KN	Kenya	Njoro	0	2,165
MG	Mexico	Guanajuato	21	1,765
MS	Mexico	Sonora	27	38
MM	Mexico	Michoacán	20	1,517
NB	Nepal	Bhairahwa	27	105
PI	Pakistan	Islamabad	34	683
PA	Pakistan	Ayub	32	213
SR	Saudi Arabia	Riyadh	24	600
SG	Sudan	Gezira	14	411
SE	Spain	Encinar	38	20
SJ	Spain	Jerez	37	180
SC	Spain	Cordoba	38	110
SS	Spain	Sevilla	38	20
TB	Tunisia	Beja	37	150
TC	Thailand	Chiang Mai	18	820

Concepts of data division or cross-validation are presented for multiple regression by Allen (1971) and Snee (1977), and for principal components analysis by Wold (1978) and Krzanowski (1983).

Cross-validation, used in AMMI to measure the success of the models, computed the sum of squared differences (SSD) between the models' fitted values and the validation data. The root mean square predictive difference (RMS PD) was the square root of SSD divided by the number of validation observations. Smaller values of RMS PD indicated more accurate prediction.
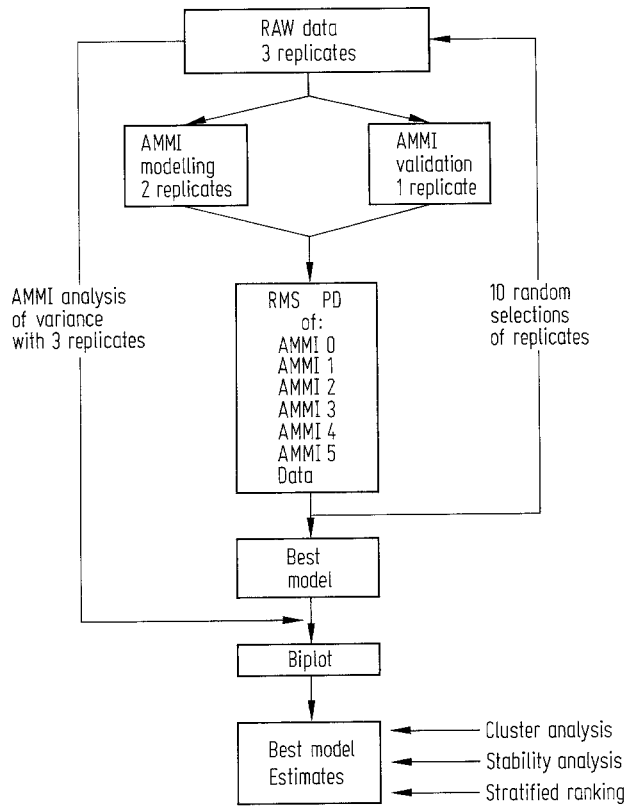
Fig. 1. Sequence of statistical analyses

This procedure was repeated using ten different random selections of replicates from the genotype × location matrix, and the average RMS PDs over the ten randomizations were calculated for each model.

On the basis of RMS PD, the best model was identified. Finally, AMMI estimates for all genotype by location combinations were generated using this model with three replicates (Fig. 1).

AMMI analysis constituted a preliminary investigation, providing predictively accurate yield estimates for input into cluster and stability analysis.

Average rank difference. Genotypes at each location were ranked on the basis of their yields. For the ith genotype in the jth location, the absolute value of the rank difference (RD) between any two predictive models, e.g. x and y, was defined as

$$\text{RD}_{ij} = (|R_x - R_y|)_{ij} ,$$

where R_x and R_y were the ranks of models x and y, respectively.

Then, the average rank difference (ARD) between models x and y for all genotypes and locations was:

$$\text{ARD}_{xy} = \left[\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \text{RD}_{ij} \right] \bigg/ n_i n_j ,$$

$$\text{ARD}_{xy} = \left[\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (|R_x - R_y|)_{ij} \right] \bigg/ n_i n_j ,$$

where n_i and n_j were the numbers of genotypes and locations, respectively.

Cluster analysis. Locations and genotypes were clustered using the HACLUS program from the University of Queensland, with a dissimilarity matrix of squared Euclidean distance and an in-

cremental sum of squares fusion strategy (Burr 1970). The first two entities to fuse minimize the increment in the within-group sum of squares. This process repeats for the second and subsequent fusions so that, at any level of the hierarchy, the within-group sum of squares is a minimum. Location main effects were removed prior to clustering locations by standardizing the data, so that each location had a mean of zero and a phenotypic standard deviation of one (Fox and Rosielle 1982). This caused location clusters to be largely determined by the relative performance of genotypes within locations and was judged the appropriate transformation for plant breeding purposes.

Stratified ranking. The stratified ranking technique of Fox et al. (1990) was applied to unadjusted means and AMMI estimates. The procedure consisted of scoring the number of locations in which each genotype ranked in the top, middle, and bottom third of trial entries.

Stability analysis. A spatial model proposed by Westcott (1987) and applied to international maize trials (Crossa 1988; Crossa et al. 1988 a, b; Crossa et al. 1989) was used.

The spatial method was used for assessing genotypic stability within the two major groups of locations obtained from cluster analysis. Locations were ranked in descending order of mean yields. For characterizing genotypic stability, principal coordinates analysis was performed in cycles. The first cycle (cycle 1) included the lowest yielding location, the second cycle (cycle 2) included the two lowest yielding locations, and so on, terminating with the final analysis based upon all locations in the group. For each cycle, a three-dimensional scattergram of principal coordinate axes represented the relationships among the genotypes.

Wescott's (1987) measure of similarity between two genotypes X and Y, in a given location i is

$$S_i(XY) = [H_i - (X_i + Y_i)/2]/H_i - L_i ,$$

where H_i is the highest yielding genotype in location i, L_i is the lowest yielding genotype in location i, and X_i and Y_i are the performances of genotypes X and Y in location i, respectively. For more than one location, the similarity between genotypes X and Y is the average of $S_i(XY)$ across locations. This measure of similarity between any pair of genotypes indicates their average proximity to H_i. The analysis determines a point at the center of the scattergram with a maximum value for S. High-yielding genotypes are close to H_i, have small values for S, and are represented by points further from the center of the scattergram (Crossa 1988). In contrast, low-yielding genotypes are far from H_i, have large values for S, and are represented by points near the center of the scattergram.

Genotypes that consistently show above average performance throughout the cycles occupy the periphery of the scattergrams. These outliers are the most stable genotypes, even though they may occupy different positions in three-dimensional space. Genotypes with below average performance are represented by points that tend towards the center.

Results

AMMI analysis

All models fitted approached the validation data more closely than the mean of the two modelling replicates (Table 3). The AMMI1 model showed the least deviation from validation data (RMS PD of 818 kg/ha), thus providing the best estimates.

Table 3. The average RMS PD (kg/ha) for seven models based on ten randomizations

Model	RMS PD
AMMI0	829
AMMI1	818
AMMI2	839
AMMI3	850
AMMI4	855
AMMI5	862
DATA	885

Table 4. Estimated differences in RMS PD and ARD among several predictive models and their rank correlations over 18 genotypes and 25 locations

Models	RMS PD	ARD	Rank correlation
AMMI0–AMMI1	11	2.88	0.92
AMMI1–AMMI2	21	2.20	0.94
AMMI2–AMMI3	11	1.45	0.95
AMMI3–AMMI4	5	1.51	0.96
AMMI4–AMMI5	7	1.44	0.96

Table 5. AMMI analysis of variance of the eighth ESWYT

Source of variation	df	$SS \times 10^6$	$MS \times 10^3$
Replicates within L	50	96	1,927
Treatment	449	7,551	16,819 **
Model	81	7,262	89,656 **
Location (L)	24	7,100	295,871 **
Genotype (G)	17	54	3,160 **
Interaction PCA 1	40	108	2,687 **
Residual	368	290	787 **
Error	850	372	438

** Significant at 0.01 probability level

Changes in RMS PD and in ARD with successive addition to the number of principal components axes modelled, from zero to five, was quantified (Table 4). Although the largest change in RMS PD did not occur between AMMI0 and AMMI1, the greatest effect on ranking was observed here, paralleled by the fact that AMMI0 and AMMI1 exhibited the lowest Kendall's Tau rank correlation (Knight 1966). Changes in the ARD did not seem related to changes in RMS PD. The stratified ranking procedure presented below showed distinctly different genotypic rankings between location groups using AMMI1 estimates, whereas for the AMMI0 model, which does not accommodate interaction and rank change, each location would order genotypes identically. Thus, moving from the AMMI0 to the AMMI1 model has important implications for selecting genotypes. Overall, successive addition of axes, after the first which accounts for the largest fraction of GE variation, showed a diminishing effect on ranking. Increasing rank correla-

tions between consecutive models, from AMMI1 onwards, reflected this trend.

A significant feature of multivariate models, including AMMI analysis, is that they account for a large proportion of pattern in the first few dimensions, with subsequent dimensions accounting for a diminishing percentage of pattern and an increasing percentage of noise (Gauch 1982, 1988). Often, arbitrary decisions are taken with respect to the number of dimensions considered. However, the RMS PD provided an objective basis for the decision. Our a priori expectations were that higher interaction PCA axes would contribute to genotypic yield prediction, but to a negligible degree above a certain number of axes. However, the poorer fit to validation data of models involving higher PCA axes (AMMI2 to AMMI5, Table 3) suggested that they involve noise, and not highly complex interaction among genotypes and locations. Noise, or nonsystematic components, of the genotype × location interaction reduced the accuracy of estimates and were thus considered residual variation, even though the axes involved may have been statistically significant using an F-test.

In the AMMI analysis of variance using three replicates (Table 5), the treatment sum of squares (SS) was partitioned into two components: 96% due to the model or pattern (including main effects for genotypes and locations and the first interaction PCA) with 81 df, and 4% residual or random variation (noise) with 368 df. The model-validation procedure (Table 3) indicated that the 4% of the treatment SS allocated to the residual term was not of predictive value.

The root mean square residual (or adjustment) for R replicates, L locations, and G genotypes was Residual $SS/RLG = 464$ kg/ha or 10% of the grand mean (Table 5). This is a measure of the departure of AMMI1 estimates from the means across replicates.

The AMMI1 estimates from two replicates were more accurate than the unadjusted means of those replicates. For three replicates, the prediction accuracy could not be measured empirically since no validation data remained. However, it was assumed that AMMI1 estimates based on three replicates were more accurate than those based on two replicates. The effect of the AMMI estimates on genotypic means was assessed.

Within locations, AMMI1 frequently ranked genotypes differently than unadjusted means (Table 6), and in 18 of the 25 cases AMMI1 estimation changed the top-yielding entry. For unadjusted means, 13 genotypes were top yielders in at least one location. With AMMI1 estimates there were only three top yielders, Kauz#2, Seri 82, and Hahn*2/Prl, suggesting that noise in unadjusted means elevated some genotypes to first position.

Studies in maize (Crossa et al. 1990) and soybean yield trials (Gauch and Zobel 1989) also showed that AMMI1 estimates and unadjusted means led to a different

Table 6. AMMI1 estimates (above), unadjusted means (below), and their rank differences (RD_{ij}) (in parenthesis) for grain yield (tons/ha) of 18 genotypes grown in 25 locations

Location	Genotype																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	7.8	8.2	8.0	7.7	7.5	7.9	7.7	7.7	7.9	7.7	8.3	8.0	8.1	8.2	8.4	7.8	7.9	7.9
	7.8	8.3	7.7	6.1	8.1	8.4	8.1	8.8	8.6	6.7	7.9	7.5	8.6	7.1	8.5	7.7	9.1	7.9
	(1)	(4)	(8)	(1)	(9)	(2)	(7)	(12)	(5.5)	(1)	(8)	(8)	(1.5)	(12)	(4)	(1)	(9)	(0)
2	5.6	5.9	5.7	5.4	5.3	5.7	5.5	5.5	5.7	5.5	6.0	5.8	5.8	5.9	6.2	5.6	5.6	5.6
	5.2	5.4	5.5	5.5	5.3	5.3	5.5	6.2	6.5	5.0	6.2	5.5	5.9	6.0	6.7	5.5	5.0	5.8
	(3)	(11)	(1.5)	(7)	(5)	(7)	(6.5)	(11)	(7)	(2)	(2)	(5)	(1)	(1)	(0)	(0)	(7)	(4)
3	2.2	2.3	2.0	2.3	2.3	2.3	2.3	1.9	2.4	2.0	2.2	2.3	2.0	2.3	2.3	2.1	2.1	2.2
	2.2	2.1	2.1	2.0	1.9	2.6	2.2	1.8	2.3	2.4	2.1	2.1	1.7	2.0	2.2	2.6	2.3	2.7
	(2)	(4)	(4)	(9)	(13)	(5)	(1)	(1)	(4)	(11)	(2)	(2)	(2)	(11)	(6)	(11)	(7)	(11)
4	1.4	1.6	1.3	1.5	1.4	1.5	1.5	1.2	1.6	1.3	1.6	1.5	1.4	1.6	1.7	1.4	1.4	1.4
	1.5	1.6	1.0	1.2	1.4	1.6	1.1	1.1	1.8	1.9	1.5	1.6	1.5	2.0	1.4	1.2	1.6	1.4
	(3)	(2)	(2)	(6)	(1)	(1)	(7)	(1)	(1)	(15)	(3)	(2)	(6)	(1)	(10)	(1)	(4)	(2)
5	5.9	6.1	5.9	5.9	5.9	6.0	5.9	5.7	6.1	5.8	6.1	6.0	5.9	6.1	6.3	5.9	5.9	5.9
	6.0	7.2	6.5	6.7	6.6	4.5	6.0	5.5	5.2	5.0	6.1	5.6	5.5	5.8	6.9	6.1	6.6	5.3
	(2)	(1)	(8)	(10)	(12)	(11)	(2)	(4)	(11)	(0)	(3)	(6)	(5)	(8)	(1)	(7)	(4)	(5)
6	3.1	3.0	2.5	3.6	3.8	3.3	3.4	2.7	3.5	3.0	2.7	3.1	2.5	3.1	2.9	3.0	3.0	3.1
	2.7	3.2	2.7	3.1	4.0	3.2	3.5	2.5	3.8	3.0	3.8	3.2	2.4	2.8	2.0	3.3	2.4	3.1
	(7)	(1)	(4)	(9)	(0)	(2)	(1)	(0)	(0)	(9)	(14)	(1)	(0.5)	(3)	(4)	(5)	(4.5)	(2)
7	3.4	3.2	2.7	3.8	4.0	3.5	3.6	2.9	3.7	3.2	2.9	3.3	2.7	3.3	3.1	3.2	3.2	3.3
	3.5	2.8	2.8	3.9	3.9	4.9	2.1	2.5	4.3	3.5	2.6	3.3	3.4	3.3	2.5	3.9	2.9	2.9
	(0)	(4)	(5)	(2.5)	(3.5)	(4)	(14)	(1)	(1)	(6)	(1)	(2)	(9)	(1)	(3)	(8)	(0.5)	(3.5)
8	10.8	10.4	9.7	11.6	12.0	11.0	11.3	10.1	11.3	10.7	9.8	10.7	9.7	10.5	10.0	10.7	10.6	10.7
	10.6	9.6	9.1	11.8	12.5	10.9	10.9	9.9	11.2	9.4	8.6	11.3	10.8	10.6	11.3	10.7	10.8	11.3
	(7)	(2)	(0)	(0)	(0)	(2)	(5)	(0)	(2)	(7)	(2)	(3)	(8)	(0)	(12)	(1)	(2)	(3)
9	7.4	7.4	7.0	7.5	7.7	7.5	7.6	7.0	7.6	7.3	7.2	7.4	7.0	7.4	7.4	7.3	7.3	7.4
	6.7	7.7	6.3	8.3	7.9	7.6	7.5	7.5	7.5	7.3	7.4	7.1	7.2	7.5	7.8	6.9	7.4	7.1
	(8)	(4)	(0)	(1)	(1)	(0)	(4)	(9)	(6)	(2)	(4)	(9)	(4)	(1)	(8)	(3)	(2)	(4)
10	5.8	6.0	5.7	5.9	5.9	5.9	5.9	5.6	6.0	5.7	5.9	5.9	5.7	6.0	6.1	5.8	5.8	5.8
	6.0	6.0	5.4	6.3	5.9	6.1	6.0	5.9	5.0	6.5	6.5	6.1	5.3	6.6	6.2	5.4	5.2	5.6
	(2)	(5)	(2)	(4)	(2)	(1)	(1)	(7)	(15)	(14.5)	(4.5)	(1)	(1)	(1)	(4)	(1)	(4)	(2)
11	6.5	7.0	6.9	6.3	6.0	6.6	6.4	6.6	6.6	6.4	7.2	6.8	7.0	7.0	7.3	6.6	6.6	6.6
	6.9	7.1	7.5	6.9	5.8	7.2	5.7	6.7	6.5	6.4	6.3	7.3	6.6	7.9	7.5	5.9	6.1	6.2
	(6.5)	(3)	(3.5)	(9.5)	(1)	(4)	(2)	(4)	(1)	(3)	(11)	(3)	(6)	(4)	(1.5)	(4)	(7)	(3)
12	5.9	5.7	5.2	6.5	6.8	6.1	6.3	5.4	6.3	5.8	5.3	5.9	5.1	5.8	5.4	5.8	5.7	5.9
	5.6	6.1	4.7	6.4	6.6	6.7	6.8	4.8	5.5	6.0	5.5	6.5	5.1	5.6	5.5	4.6	5.8	6.8
	(5)	(6)	(1)	(4)	(3)	(2)	(1.5)	(1)	(8.5)	(1)	(3.5)	(3)	(2)	(2)	(0)	(8)	(2)	(5.5)
13	2.7	3.1	3.1	2.4	2.1	2.8	2.5	2.7	2.7	2.6	3.3	2.9	3.2	3.1	35	2.7	2.8	2.7
	3.5	2.9	3.2	2.2	1.8	2.7	2.1	2.2	2.7	2.9	3.1	2.9	2.9	2.8	3.7	2.7	3.6	3.1
	(11)	(3)	(2)	(2)	(0)	(3)	(1)	(3)	(2)	(7)	(3)	(2)	(7)	(6)	(0)	(3)	(6)	(5)
14	8.0	8.2	7.9	8.0	8.0	8.1	8.0	7.8	8.2	7.9	8.1	8.1	7.9	8.2	8.3	7.9	8.0	8.0
	7.7	7.8	7.7	8.2	8.1	7.3	8.4	7.8	8.4	7.3	8.2	8.6	7.9	8.4	8.8	7.8	7.6	8.4
	(3)	(8.5)	(2)	(0)	(4)	(10)	(4)	(6.5)	(1)	(1)	(2)	(4)	(5)	(3)	(0)	(1)	(5)	(5)
15	5.5	5.3	4.7	5.9	6.1	5.6	5.8	5.0	5.8	5.3	4.9	5.4	4.8	5.3	5.1	5.3	5.3	5.4
	5.6	5.2	5.1	5.7	6.7	5.3	5.6	4.5	5.8	5.5	4.9	5.4	4.9	4.8	5.4	5.5	5.4	5.1
	(1)	(1)	(5)	(1)	(0)	(6)	(0)	(3)	(1)	(3)	(1)	(1)	(1)	(8)	(5)	(5)	(2)	(6)
16	4.7	4.6	4.2	5.1	5.2	4.8	5.0	4.3	5.0	4.6	4.3	4.7	4.2	4.7	4.5	4.6	4.6	4.7
	4.0	5.2	3.6	4.9	5.0	4.7	5.3	4.3	4.9	4.5	4.2	4.7	4.1	5.4	4.3	5.1	4.3	5.3
	(10)	(6)	(0)	(6)	(5)	(4.5)	(1.5)	(3)	(4)	(2)	(0)	(3.5)	(1)	(8)	(1)	(7)	(2)	(5.5)
17	3.8	3.9	3.5	4.0	3.9	3.9	3.5	3.5	4.0	3.7	3.8	3.9	3.6	3.9	3.9	3.7	3.7	3.8
	4.0	2.9	3.8	4.6	3.9	4.3	3.9	1.5	4.0	3.2	4.6	4.1	4.4	3.7	3.9	4.3	3.5	3.9
	(3)	(10)	(4)	(2)	(8)	(1)	(4)	(0)	(7)	(1)	(11)	(3)	(13)	(8)	(6)	(9)	(2)	(0)
18	4.7	4.4	3.8	5.3	5.6	4.8	5.1	4.1	5.1	4.5	4.0	4.6	3.8	4.5	4.2	4.5	4.5	4.6
	4.1	4.9	4.6	5.5	5.3	4.8	6.0	4.8	4.8	3.9	3.7	4.8	3.0	4.8	4.2	4.0	4.8	4.3
	(8)	(9)	(7)	(0)	(2)	(0)	(2)	(8)	(2)	(7)	(1)	(1)	(1)	(4)	(1)	(5)	(1)	(5)
19	1.8	1.8	1.4	2.1	2.2	1.9	2.0	1.4	2.1	1.7	1.6	1.8	1.4	1.8	1.8	1.7	1.7	1.8
	2.0	2.3	1.4	2.1	1.8	2.1	1.9	1.1	2.0	1.9	1.9	1.5	1.1	1.7	1.3	2.1	1.7	2.1
	(2)	(7)	(3)	(1)	(10)	(3)	(5.5)	(2)	(3)	(6)	(5.5)	(8)	(0)	(6)	(5)	(8)	(0)	(6)
20	3.8	3.9	3.6	4.0	4.0	4.0	4.0	3.6	4.1	3.7	3.8	3.9	3.6	3.9	4.0	3.8	3.8	3.8
	3.6	3.8	3.5	3.4	3.9	4.3	4.4	4.7	4.2	4.1	3.6	3.9	3.9	4.1	3.4	3.4	3.5	3.7
	(2)	(2)	(3)	(13)	(5)	(1)	(4)	(17)	(3)	(9)	(0)	(1)	(7)	(2)	(12.5)	(3.5)	(3)	(0)

Table 6 (continued)

Location	Genotype																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
21	2.7	2.3	1.6	3.4	3.9	2.8	3.2	2.0	3.1	2.5	1.7	2.6	1.6	2.4	1.9	2.5	2.5	2.6
	2.9	2.6	1.4	3.1	3.4	2.6	3.2	2.0	3.6	3.0	2.4	1.6	1.1	3.0	1.0	3.5	2.4	2.4
	(2)	(4)	(1)	(3)	(2)	(5)	(1)	(0)	(3)	(2)	(3)	(7)	(1)	(6)	(3)	(8)	(0)	(5)
22	2.7	2.6	2.0	3.1	3.3	2.8	2.9	2.2	3.0	2.5	2.2	2.7	2.1	2.6	2.4	2.5	2.5	2.6
	3.2	2.1	2.6	3.0	2.8	2.3	3.6	2.2	3.0	3.1	2.0	3.1	2.2	2.3	2.0	2.7	2.2	2.4
	(4)	(6)	(9)	(3)	(6)	(7)	(3)	(1)	(3)	(10)	(2)	(3)	(4)	(2)	(4)	(3)	(2)	(2)
23	1.0	1.3	1.0	1.0	1.0	1.2	1.1	0.9	1.2	0.9	1.2	1.2	1.1	1.3	1.4	1.0	1.1	1.1
	0.8	1.3	0.7	0.7	1.0	0.8	1.0	1.0	1.3	1.3	1.4	1.2	1.1	1.1	1.2	1.5	1.5	1.0
	(3)	(4)	(4)	(5)	(2)	(8)	(0)	(5)	(0)	(13)	(1)	(2)	(2)	(6)	(6)	(14)	(7)	(2)
24	3.9	4.0	3.6	4.1	4.2	4.1	4.1	3.6	4.2	3.8	3.8	4.0	3.6	4.0	4.0	3.9	3.9	3.9
	4.1	4.3	3.9	4.6	4.0	4.1	4.5	3.5	3.9	3.8	3.7	3.9	3.9	3.9	3.7	4.0	3.5	3.5
	(6)	(5)	(9)	(2)	(6)	(0)	(2)	(0)	(10)	(1)	(1)	(1)	(5)	(4)	(6)	(7)	(4)	(7)
25	5.7	5.0	4.2	6.6	7.2	5.8	6.3	4.8	6.2	5.5	4.3	5.4	4.1	5.1	4.5	5.4	5.3	5.5
	6.7	4.9	4.4	6.5	7.5	5.4	5.7	5.6	6.4	5.5	4.0	5.2	3.6	4.7	5.2	4.5	6.0	5.2
	(4)	(0)	(1)	(1)	(0)	(4)	(3)	(7)	(0)	(0)	(1)	(2)	(0)	(2)	(3)	(5)	(6)	(3)
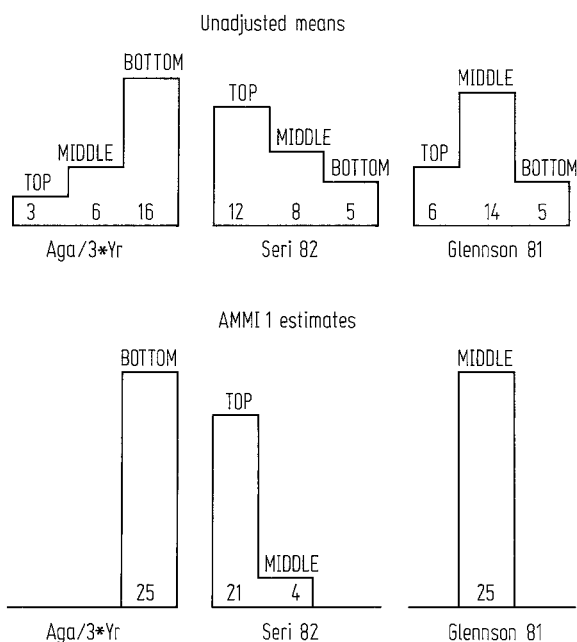
Fig. 2. Stratified ranking analyses of three selected genotypes for unadjusted means and AMMI1 estimates over all locations, giving number of locations for which a genotype occurred in the top, middle, and bottom third of entries

top-ranking genotype in over half the locations. AMMI1 estimation was recommended, since ranking discrepancies between AMMI1 estimates and unadjusted means were attributable to random statistical variation.

Figure 2 presents three contrasting genotypes and shows that AMMI1 estimation had a profound effect, producing sharper, stratified ranking patterns. Aga/3 * Yr was generally in the bottom third using unadjusted means, but also occurred three times in the top and six

times in the middle. With AMMI1 estimates, it was always in the bottom third. Seri 82 was often in the top, but eight times in the middle and five times in the bottom, for unadjusted means. The pattern for Seri 82 using AMMI1 estimates was 21, four, and zero occurrences in the top, middle, and bottom, respectively. Glennson 81 was frequently in the middle in the unadjusted analysis, but was also in the top third six times and in the bottom, five times. AMMI1 estimates always placed it in the middle. It appeared that, without AMMI estimation, noise in the data partially blurred adaptation patterns of genotypes to the extent that a relatively well-adapted genotype fell by chance in the middle and bottom several times. Conversely, a relatively poorly adapted line sometimes occurred at the top for the same reason.

The distribution of differences between unadjusted genotype × location mean yields and their corresponding AMMI1 estimates was plotted for locations and genotypes. It revealed that differences followed a relatively even pattern throughout the data matrix, without large changes focused on a few genotypes or locations. This even pattern seemed consistent with the idea of a general noise reduction in the data matrix, while a centering of changes may have suggested an oversimplification by the model in not accounting for atypical genotypes or locations.

AMMI analysis provides a graphical representation (or biplot, Fig. 3) to summarize information on main effects and interactions (PCA 1) of both genotypes and environments simultaneously (Kempton 1984). In the current study, displacement along the abscissa reflected differences in main effects, whereas displacement along the ordinate exhibited differences in interaction effects. For example, sites TC and MG differed in main effect but not in interaction, whereas TC and KN had similar main
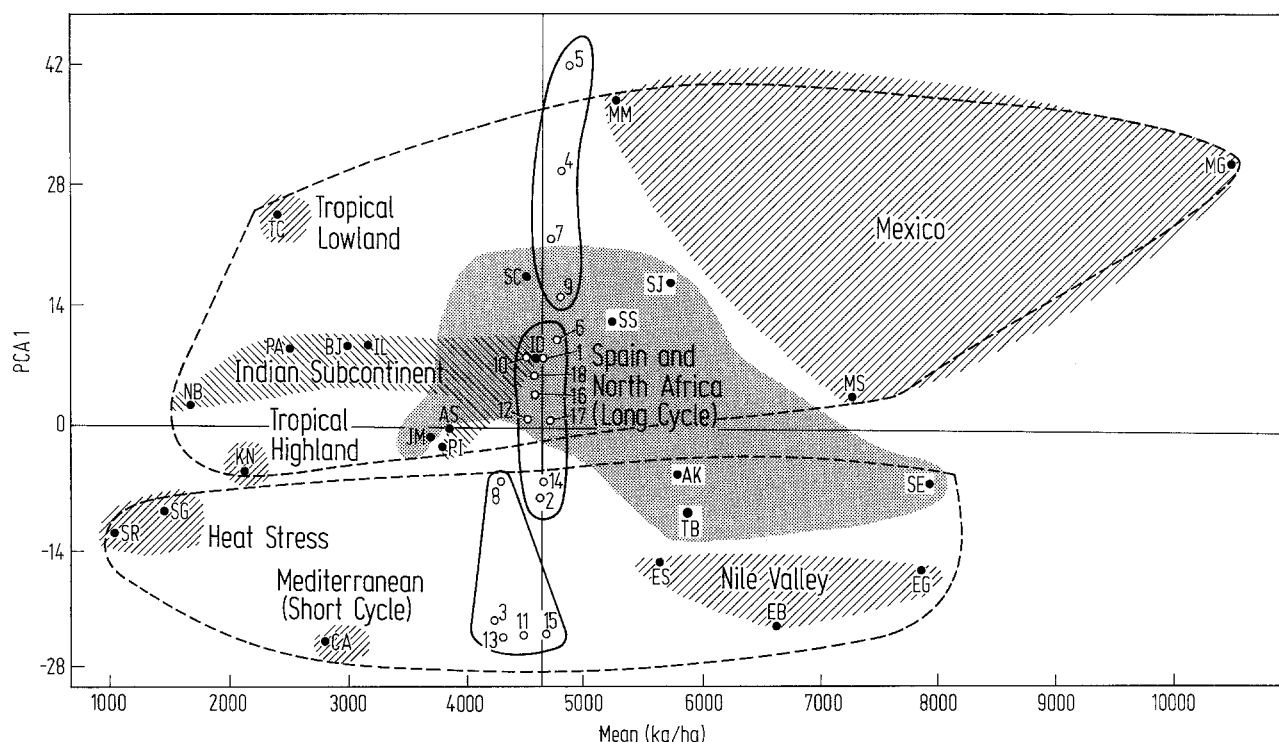
Fig. 3. Biplot of the unadjusted means (kg/ha) and the first PCA axis for interaction of 18 genotypes (o) and 25 locations (●). Subgroups of locations from clustering are shown by *dashed* (– –) lines and genotypic subgroups are delineated by *unbroken* lines (——). The shaded regions correspond to CIMMYT breeders' division of locations. The vertical line represents the grand mean of the experiment. The horizontal line is PCA 1 = 0

effects but different interaction with genotypes. Sites TC and MS differed in both main effects and interaction.

The AMMI1-estimated yield for any genotype in any location may be calculated from the biplot (Fig. 3), as described by Zobel et al. (1988). The biplot captured 96% of the treatment SS. Genotypes with PCA 1 scores near zero had little interaction across environments (and, likewise, locations with PCA 1 scores near zero had little interaction across genotypes and low discrimination among genotypes). For such genotypes, overall mean rankings were fairly reliable, whereas for genotypes with large PCA 1 scores (either positive or negative) overall rankings were less reliable. Genotype and location combinations with PCA 1 scores of the same sign produced positive specific interaction effects, whereas combinations of opposite sign had negative specific interactions.

Cluster analysis

Locations were clustered using the AMMI1-estimated values of 18 genotypes as attributes and, conversely, the genotypes were grouped using the AMMI1-estimated values of 25 locations as attributes. Figure 3 shows: (1) the biplot with unadjusted main effects of genotypes and locations on the abscissa and the AMMI interaction PCA 1 on the ordinate, and (2) the superimposed group-

ing of genotypes (unbroken lines) and locations (dashed lines) from cluster analysis using AMMI1 estimates. Also, superimposed on the biplot (Fig. 3) is the more intuitive grouping of locations (shaded regions) by CIMMYT breeders.

The largest location subgroup (Group 1) comprised 16 locations: 12 with positive first principal components plus Njoro in Kenya (KN), Madaba in Jordan (JM), Setif in Algeria (AS), and Islamabad in Pakistan (PI) with small negative values for PCA 1. Group 2 included the nine locations with the largest negative first PCA scores.

In location Group 1, Bhairahwa in Nepal (NB), Ayub and Islamabad in Pakistan (PA and PI), Joydebpur in Bangladesh (BJ), and Delhi and Ludhiana in India (ID and IL) represent the Indian subcontinent, where wheat is grown in rotation with rice and yields are frequently limited by terminal heat stress after a short growing cycle. The separation of PI from the other subcontinent locations was probably due to its high stripe-rust incidence.

Also in Group 1, Chiang Mai in Thailand (TC) is a location in the lowland tropics with high relative humidity and heat stress, while Njoro in the tropical highlands of Kenya experiences high diurnal temperature fluctuations. Three of the four sites in Spain – Córdoba, Jerez, and Sevilla (SC, SJ, and SS) – were tightly grouped. The separation into Group 2 of Encinar (SE) in Spain can be
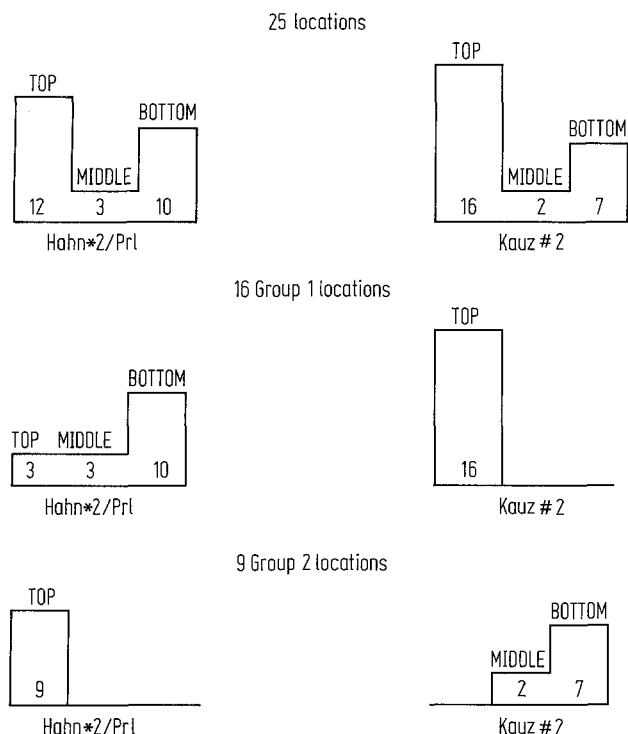
Fig. 4. Stratified ranking analyses of two selected genotypes using AMMI1 estimates over all locations and for Groups 1 and 2, giving number of locations for which a genotype occurred in the top, middle, and bottom third of entries

explained by higher rainfall and associated attacks of powdery mildew and septoria tritici blotch. Higher disease pressure differentiated the germplasm differently from the other three locations, as seen from the first principal component axis, but was not severe enough to prevent the location mean yield from reflecting increased rainfall.

Separation between the two Algerian locations, Setif (AS) and El Khroub (AK), might have been due to differences in growing cycle duration and precipitation (AK − 597 mm, AS − 354 mm). At Setif, average mature plant height was 77 cm. Such short stature usually indicates severe environmental stress, probably drought. The higher altitude of setif also produces cold stress.

Three Mexican locations that differed in main effects and interaction also joined Group 1. Guanajuato, Sonora, and Michoacán (MG, MS, and MM) are irrigated, high-yielding environments with a shorter growing cycle than occurs in Spain and North Africa. The difference in PCA 1 between Michoacán and Guanajuato versus Sonora may be explained by elevation and latitude. MM and MG [1,517 and 1,765 m above sea level (masl)] at 20° and 21°N, respectively, are cooler than MS, which is an irrigated desert environment with an elevation of 38 masl at 27°N.

In Group 2, three locations from the Nile Valley of Egypt − El Gemmeiza, Sakha, and Beni-Suef (EG, ES,

and EB) − occurred with Riyadh in Saudi Arabia (SR) and Gezira in Sudan (SG) with a negative PCA1. All are irrigated, but heat stress may have differentiated SR and SG from the environment found in the Nile Valley in Egypt.

Clustering of genotypes (Fig. 3) showed three major subgroups. The first included four closely related, high-yielding genotypes with high positive values for PCA1. Two were Kauz sister lines (genotypes 4 and 5) and two, Veery sister lines, Genaro 81 and Seri 82 (genotypes 7 and 9). Kauz was developed from the cross Jup/Bjy//Ures (Ures being another Veery, closely related to Genaro 81) and is phenotypically very similar to Genaro 81.

The second subgroup consisted of nine genotypes (1, 2, 6, 10, 12, 14, 16, 17, 18) with intermediate positive and negative first PCA scores.

Five genotypes with negative values for the first PCA comprised the third subgroup of later-maturing genotypes (3, 8, 11, 13, 15). Genotype 8, the earliest-maturing entry in the nursery, was an exception and appeared a distant member of the third group. It was the only two-gene dwarf in the trial and showed the highest susceptibility of any genotype to stripe rust.

The general lateness of subgroup 3 genotypes was consistent with their negative interaction with Indian subcontinent locations, where earliness is sought. The apparent anomaly of the positive interaction of these later entries with locations subject to severe environmental stress, e.g. heat stress at SG and SR, may be explained by lateness providing some check against bolting. The latest genotype matured in 110 days at SG.

In general, the main effects axis of the biplot diagram reflected the overall quality for locations and general breeding status for genotypes, whereas the interaction axis discriminated early (positive PCA scores) to late (negative PCA scores) maturing genotypes and, correspondingly, the lengths of growing season of locations.

The advantage of superimposing clustering on the biplot was that subgroups were delineated clearly. The genotypic subgroup with the highest positive PCA 1 values (4, 5, 7, 9) and some genotypes from the subgroup with intermediate positive values of PCA 1 (1, 6, 10, 16, 18) had positive interactions with most Group 1 locations, except KN, JM, AS, and PI.

Nine Group 2 locations with the lowest negative PCA 1 scores showed positive, specific interactions with genotypes 2, 3, 8, 11, 13, 14, and 15.

Figure 4 presents stratified ranking patterns for Kauz#2 and Hahn*2/Prl using AMMI1 estimates, first for the 25 locations and then individually for location Groups 1 and 2. They showed a similar overall pattern, but Kauz#2 was well adapted to location Group 1 and relatively poorly adapted to Group 2. Hahn*2/Prl showed the opposite trend. These results confirm the interaction patterns inferred from Fig. 3.
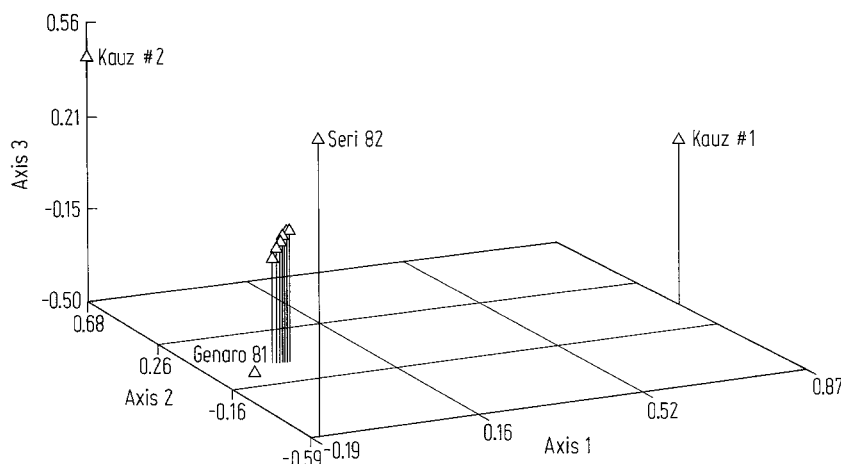
Fig. 5. Plot of the first three principal axes from a principal coordinate analysis of a set of 18 wheat genotypes in all Group 1 locations

Cluster analysis performed on unadjusted mean values formed slightly different groups of locations and genotypes: (1) Madaba in Jordan (JM) was located in environment group 2; (2) genotypes 7 and 9 joined the subgroup with intermediate positive and negative first PCA scores; and (3) genotype 8 was a single-member group. Using AMMI1 estimates as input to cluster analysis, groups of genotypes and locations were more cohesive, leading to clearer explanations of response patterns. Clusters based on AMMI1 estimates and unadjusted means were different because, as mentioned, AMMI1 estimated values differed from unadjusted data by a root mean square residual of 464 kg/ha.

While concurrence among the AMMI biplot, cluster analysis, and CIMMYT breeders' location grouping was imperfect, the fact that three schemes could be superimposed is significant. Figure 3 provided a powerful tool for agronomic and biological interpretation of such a large data set.

Stability analysis

For location Group 1, Westcott's (1987) spatial technique using AMMI1 estimates indicated that the Kauz lines (genotypes 4 and 5) and Seri 82 (genotype 9) were outliers in all cycles and thus represented the most stable genotypes. They were the highest-yielding genotypes in most Group 1 locations, with Kauz#2 first in all but one. Genaro 81 also showed a stable pattern but, being closer to the center of the diagram in most cycles, did not equal the Kauz lines and Seri 82.

Rather than include a large number of scattergrams (one for each cycle in each location group), only one three-dimensional plot associated with the analysis for all locations in Group 1 is presented in Fig. 5. Regardless of the direction of their separation from the center, the Kauz lines and Seri 82 are outliers. Genaro 81 was closer to the center of the plot.

Results of the stability analysis across nine Group 2 locations using the spatial model showed genotypes 2, 11, 14, and 15 to be outliers in all diagrams and, thus, the most stable. Hahn*2/Prl (genotype 15) was the highest yielder in all nine environments.

Discussion

The primary aim of multilocation trials in plant breeding is to estimate genotypic yields. One obstacle in this estimation is noise and error in field data.

Empirically, the noise in the AMMI analysis is quantified by the residual SS after adjusting for the best predictive model, whereas the error is estimated by the differences among individual experimental units (replicates) with the same treatment (genotype × location combination). In Table 5, the residual term could be thought of as corresponding to noise, which is extractable by AMMI analysis from the genotype × environment SS and, in the same vein, the error term could be described as that error variation which is extractable by the randomized complete block design. This does not imply, however, that these or any other sources of variation are free of error and noise.

From a practical viewpoint, a precise distinction between noise and error is difficult. Nonuniformity in irrigation, application of agrochemicals, insect and animal damage, and disease inoculum load contribute to undesired variation, as do heterogeneous soils and imprecision in field operations. Statistically, the degree to which this can be accounted for in the replicate, incomplete block, error, residual, or other term in an analysis of variance depends on the appropriateness of (1) the experimental design, and (2) the statistical analysis.

Three basic types of error control strategies exist for increasing the accuracy of genotypic yield estimates: (1) partitioning error variation, (2) partitioning genotypic

variation, and (3) partitioning genotype × location variation. The first uses incomplete block designs in the form of lattices (Yates 1936) or generalized lattices (Patterson et al. 1978), which remove interblock variation from experimental error. The second approach includes spatial methods (neighbor analysis) to adjust for soil fertility trends (Wilkinson et al. 1983; Besag and Kempton 1986). Both strategies relate to experimental design. In contrast, the AMMI model removes residual or noise variation from genotype × location interaction.

Because these three error control strategies are applied to different and orthogonal sources of variation (error, genotypes, and interaction), they can be used independently or simultaneously. Despite the fact that practical complementarity of these error control strategies requires investigation, it can be postulated that, for appropriately planned trials, more accurate yield estimates would be obtained by integrating the three approaches. Effective methods for adjusting treatment means, demonstrated empirically to improve predictive success, may become increasingly important given a trend towards fewer replicates and more locations (Bradley et al. 1988).

The sequence of statistical procedures used in this study features the third strategy, AMMI analysis, as a first step. This improved the predictive accuracy of the eighth ESWYT trial by separating pattern from nonsystematic variation.

A significant feature of AMMI analysis is that adjustment is carried out using information from other locations to refine the estimates within a given location, whereas the other strategies pertain to an intralocation adjustment, i.e., one location has no influence upon the adjustment at another. Gauch (1990) suggests that this is one of the reasons why AMMI estimators may be more predictively accurate than treatment means. AMMI models consider all RGL observations (R replicates, G genotypes, and L locations), whereas treatment means focus exclusively on R replicates and disregard the other $R\,(GL-1)$ observations. The second reason relates to the Stein effect, which indicates that a small sacrifice in bias, in this case using AMMI estimators instead of unbiased treatment means, can produce a gain in accuracy. Gauch (1990) provides detailed theoretical bases for both these factors, which relate to predictive accuracy of AMMI estimators.

From a biological standpoint, one outlying location dominated by a stress (i.e., toxic levels of soluble aluminum in the soil) among several other locations without such a problem could be overridden by a relatively coordinated message or pattern coming from the other locations. The majority of associated and biologically meaningful genotype × location interactions would be relegated to the residual term of the AMMI analysis. However, even distribution of AMMI adjustments through the eighth ESWYT data matrix, without large

changes centered on a minority of genotypes or locations, allayed fear of such oversimplification. Nonetheless, risk of oversimplification should be recognized. With regard to the question of the range over which AMMI will function, it is recommended that an examination of the distribution of changes due to AMMI estimation should be routinely conducted.

The CIMMYT Wheat Breeding Program has classified major agroclimatic areas of wheat production mainly based on climatic conditions, foliar disease complexes, and soil type (Rajaram et al. 1984). Even though based on one year's data, the biplot and cluster analyses appeared reasonably congruent with the breeders' grouping.

Integration of the analyses was useful for interrelating the different objectives of multilocation trials: (1) to assess the success of yield prediction; (2) to group sites for breeding; and (3) to interpret genotype × environment interaction. Results suggested that definition of subgroups of locations for which specifically improved lines could be developed is feasible, with Kauz#2 and Hahn*2/Prl demonstrating superiority for specific subgroups of locations. Seri 82, however, was the most broadly adapted genotype in the trial, and the extent to which a breeder should capitalize on specific interactions between genotypes and subsets of locations, at the expense of aiming for broader adaptation, will depend upon the degree to which the demonstrated relationships among locations are repeatable across years.

References

Abou-El-Fittouh HA, Rawlings JO, Miller PA (1969) Classification of environments to control genotype by environment interactions with an application to cotton. Crop Sci 9:135–140

Allen DM (1971) Mean square error of prediction as a criterion for selecting variables. Technometrics 13:469–475

Besag J, Kempton R (1986) Statistical analysis of field experiments using neighboring plots. Biometrics 42:231–251

Bradley JP, Knittle KH, Troyer AF (1988) Statistical methods in seed corn production selection. J Prod Agric 1:34–38

Bradu D, Gabriel KR (1978) The biplot as a diagnostic tool for models of two-way tables. Technometrics 20:47–68

Burr EJ (1970) Cluster sorting with mixed character types. II. Fusion strategies. Aust Comput J 2:98–103

Byth DE, Eisemann RL, De Lacy IH (1976) Two-way pattern analysis of a large data set to evaluate genotypic adaptation. Heredity 37:215–230

Crossa J (1988) A comparison of results obtained with two methods for assessing yield stability. Theor Appl Genet 75:460–467

Crossa J, Westcott B, Gonzalez C (1988a) The yield stability of maize genotypes across international environments: Full season tropical maize. Exp Agric 24:253–263

Crossa J, Westcott B, Gonzalez C (1988 b) Analyzing yield stability of maize genotypes using a spatial model. Theor Appl Genet 75:863–868

Crossa J, Westcott B, Gonzalez C (1989) The yield stability of CIMMYT's maize germplasm. Euphytica 40:245–251

Crossa J, Gauch HG, Zobel RW (1990) Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. Crop Sci 30:493–500

Fox PN, Rosielle AA (1982) Reducing the influence of environmental main-effects on pattern analysis of plant breding environments. Euphytica 31:645–656

Fox PN, Skovmand B, Thompson BK, Braun H-J, Cormier R (1990) Yield and adaptation of hexaploid spring triticale. Euphytica 47:57–64

Freeman GH, Dowker BD (1973) The analysis of variation between and within genotypes and environments. Heredity 30:97–109

Gauch HG (1982 a) Noise reduction by eigenvector ordinations. Ecology 63:1643–1649

Gauch HG (1988) Model selection and validation for yield trials with interaction. Biometrics 88:705–715

Gauch HG (1990) Full and reduced models for yield trials. Theor Appl Genet (in press)

Gauch HG, Zobel RW (1988) Predictive and postdictive success of statistical analyses of yield trials. Theor Appl Genet 76:1–10

Gauch HG, Zobel RW (1989) Accuracy and selection success in yield trial analyses. Theor Appl Genet 77:473–481

Kempton RA (1984) The use of the bi-plots in interpreting variety by environment interactions. J Agric Sci 103:123–135

Knight WE (1966) A computer method for calculating Kendall's Tau with ungrouped data. J Am Stat Assoc 61:436–439

Krzanowski WJ (1983) Cross-validatory choice in principal components analysis – some sampling results. J Stat Comp Simul 18:299–314

Patterson HD, Williams ER, Hunter EA (1978) Block designs for variety trials. J Agric Sci 90:395–400

Rajaram S, Skovmand B, Curtis BC (1984) Philosophy and methodology of an international wheat breeding program. In: Gustafson JP (ed) Gene manipulation in plant improvement. 16th Stadler Genet Symp, 1984, University of Missouri-Columbia. Plenum Press, New York London. pp 33–60

Shorter R, Byth DE, Mungomery VE (1977) Genotype × environment interactions and environmental adaptation. II. Assessment of environmental contributions. Aust J Agric Res 28:223–235

Snee RD (1977) Validation of regression models: methods and examples. Technometrics 19:415–428

Westcott B (1987) A method of assessing the yield stability of crop genotypes. J Agric Sci 108:267–274

Wilkinson GN, Eckert SR, Hancock TW, Mayo O (1983) Nearest neighbor (NN) analysis of field experiments (with discussion). J Stat Soc, Ser B 45:152–212

Wold S (1978) Cross-validatory estimation of the number of components in factor and principal components models. Technometrics 20:397–405

Yates F (1936) A new method of arranging variety trials involving a large number of varieties. J Agric Sci 26:424–455

Zobel RW, Wright MJ, Gauch HJ (1988) Statistical analysis of a yield trial. Agron J 80:388–393