Theor Appl Genet (1997) 94: 1072–1077

P. Annicchiarico

Additive main effects and multiplicative interaction (AMMI) analysis of genotype-location interaction in variety trials repeated over years

Received: 12 December 1996 / Accepted: 28 February 1997

Abstract Genotype-location (GL) interaction effects are of special interest for breeding programmes to identify adaptation targets, adaptive traits and test sites. These effects, generally having relatively low repeatability between years, should be studied on a multi-year basis in annual crops. Their assessment by additive main effects and multiplicative interaction (AMMI) analysis is currently defined for this situation. Two procedures based on cross validations are proposed for testing the GL-interaction principal component axes, exploiting the utilities of the computer programme MATMODEL. The use of Gollob's F test, F_{GH2} test, F_{R} test and the heuristic criterion based on the signal-to-noise ratio is also envisaged. The consistency of results provided by the testing procedures was verified on four data sets of different cereal crops. Gollob's test tended to be the most liberal, while the F_{GH2} test appeared somewhat more liberal than the F_{R} test. The signal-to-noise ratio gave results consistent with the F_{R} test considered at a $P \leqslant 0.01$ level of significance. These criteria disagreed in two data sets with the conclusions provided by the two cross-validation procedures which, in turn, also disagreed in one data set. Preference could be given to different testing procedures depending on the number of test years, locations and genotypes.

Key words Adaptation · AMMI model · Cross validations · Genotype-environment interaction · Predictive accuracy

Introduction

Genotypes of annual crops evaluated for grain yield on a multi-locational, multi-year basis frequently show genotype-environment (GE) interactions that complicate the selection and/or recommendation of materials. The study of genotype-location (GL) interaction is of special interest for breeding programmes since it may lead to the subdivision of a target region into different sub-regions, relatively uniform with respect to genotypic responses, in order to exploit specific adaptation effects; alternatively, it may provide guidelines for the choice of crucial test sites in a wide adaptation prospect (Annicchiarico 1992; De Lacy et al. 1994). Adaptive traits can concurrently be identified (Annicchiarico and Perenzin 1994). Coping with genotype-year (GY) and genotype-location-year (GLY) interaction effects is possible only by selection for yield stability across environments defined as location-year combinations.

Various statistical methods have been developed for the analysis of GE interaction. Amongst them, Additive main effects and multiplicative interaction (AMMI) analysis (Gauch 1992) is particularly effective for depicting adaptive responses (Crossa 1990; Annicchiarico 1997). After fitting the genotype and environment main effects in the model, a crucial step in the analysis is the determination of the amount of pattern, namely the portion of GE-interaction variation representing real responses to genotypes and environments, and of noise, i.e. the random variation affecting GE-interaction effects. Ideally, only pattern is included in the selected AMMI model by retaining in its multiplicative term the statistically significant GE-interaction principal-component (PC) axes. An early F test devised by Gollob (1968) for the assessment of PC axes proved too liberal both on theoretical grounds and following simulation results (Mandel 1971; Cornelius 1993). Other F tests have been developed that allow a better control of Type-I error rates (Cornelius et al. 1992; Cornelius 1993; Piepho 1995). An alternative criterion is based on a cross-validation procedure in which the observations available for each GE combination are randomly split into modelling and validation data

---

Communicated by H. C. Becker

P. Annicchiarico
Istituto Sperimentale per le Colture Foraggere, viale Piacenza 29, I-20075 Lodi, Italy

(Gauch 1992 p. 134). The AMMI models, including zero (AMMI 0) to all possible N (AMMI N) PC axes, are compared in terms of predictive accuracy. The assessment in terms of predictive accuracy, where predictive stands for the ability of a model to explain variation in data not used in constructing the model itself, can be considered in principle more objective than other testing criteria (Gauch and Zobel 1988; Gauch 1992 p. 120). The extensive computation required for the cross-validation procedure can be performed by the computer programme MATMODEL (Gauch 1993 a). After the assessment, the selected AMMI model is usually re-parametrized using information from all available data. One last criterion of great simplicity, which provides a guideline rather than a real statistical test of predictive accuracy, is based on the estimation of the signal-to-noise ratio contained in the interaction matrix (Gauch 1992 p. 147). The selected model is the one which allows discarding the portion of GE-interaction variation closest to the expected noise estimated from the experimental error variance.

AMMI analysis and relative testing procedures for PC axes have generally been applied to GE-interaction data matrices, environments being either location-year combinations or locations in trials from only one year. No clear definition of these procedures has been provided for multi-locational variety trials repeated over years, which are very frequent in crops such as grain cereals. Indeed, the relatively low repeatability of GL adaptive patterns between different years (Léon and Becker 1988; Annicchiarico 1997) entrains a need for these crops to study GL-interaction effects on a multi-year basis, in order to retain in the AMMI model only those effects that are consistent across years. One objective of the present study was the definition of procedures for the predictive assessment of GL-interaction PC axes based on cross validations for this situation, exploiting utilities of the programme MATMODEL. The use of four other tests, namely Gollob's *F* test, F_{GH2} test, F_{R} test and the criterion based on the signal-to-noise ratio, was also envisaged. The consistency of results provided by the different testing procedures was verified on four multi-year data sets of different cereal crops.

Materials and methods

Experimental data

The grain yield of four data sets was considered. The sets are those referenced as bread wheat, data set 1 of durum wheat, maize FAO class 700, and oat, in the study of Annicchiarico (1997), where some information on the trials can be found together with reference to other reports providing further details. For the maize data set, having three or four replicates per environment, analyses were currently performed on data of three replicates. All sets included data from three cropping seasons. The number of genotypes and locations was, respectively, 11 and 11 for maize, 18 and 31 for bread wheat, 9 and 6 for durum wheat, and 10 and 7 for oat.

AMMI model

The analysis of variance (ANOVA) model for the individual yield value Y_{ijk} of the levels i of genotype (G), j of location (L), and k of year (Y) factors, omitting the block factor, is:

$$Y_{ijk} = m + G_i + L_j + Y_k + (GL)_{ij} + (GY)_{ik} + (LY)_{jk} + (GLY)_{ijk} + e_{ijk},$$

where m is the grand mean and e_{ijk} is the experiment error. The GLY interaction is the appropriate error term for testing the GL interaction under the common assumption of the year factor as random (Cochran and Cox 1957 p. 565). Variance components were estimated for ANOVA effects from the expectations of mean squares according to a completely random model (Wricke and Weber 1986 p. 104), using PROC VARCOMP of the SAS software, in order to assess the relative importance of various sources of variation in the different data sets. The genotype main effect was tested as described by Cochran and Cox (1957 p. 567).

The $(GL)_{ij}$ effect in the ANOVA was partitioned according to different AMMI models as:

$$(GL)_{ij} = \sum l_n u_{ni} v_{nj} + r_{ij},$$

where $\sum$ is the sum of the $n = 1, 2, \ldots, N$ PC axes included in the model, l_n is the singular value for the PC axis n, u_{ni} and v_{nj} are the genotype and the location eigenvectors, respectively, for the PC axis n, and r_{ij} is the residual GL interaction. Eigenvectors are scaled as unit vectors.

Cross validations

In the procedure devised by Gauch (1992, 1993 a), targetted to analysing GE-interaction data matrices, individual plot values are randomly split into modelling and validation data and AMMI models are compared in terms of root mean square prediction difference (RMSPD) between model expected values and validation observations. The larger the inconsistency between observations submitted to same genotype and environment effects (i.e. the pooled experiment error), the greater the noise relative to the pattern. Since results of different data splittings may be inconsistent, especially in the presence of relatively few validation observations, it is recommended to average the results of at least ten independent cross validations to obtain a reliable assessment for each AMMI model (Crossa et al. 1990). In the computer programme MATMODEL, different splittings of single observations of GE combinations can be obtained by choosing method R of replicate selection (Gauch 1993 a p. 24). The observations are chosen regardless of the possible presence of a block factor in the experiments. Among alternative options also available, method V allows the selection of the same replicate(s) throughout for all GE combinations (Gauch 1993 a p. 25). This method is not used frequently, as it rarely enables one to perform the recommended number of validation runs.

In the case considered here, the amount of noise for the assessment of GL-interaction PC axes depends instead on the degree of inconsistency for GL-interaction effects between observations of different years (i.e. the GLY interaction). An application of this concept implies the performance of cross validations on data of GL combinations of different years previously averaged across experiment replicates. This can be accomplished by the programme MATMODEL inputing genotype-location data matrices of different years as if they were genotype-environment data matrices of different experimental replicates. Two major approaches could then be envisaged for the assessment. The first (termed hereafter as Procedure 1) is using observations of some years for modelling and of remaining years for validation, choosing method V of replicate selection in the programme. A similar approach was used by Weber and Westermann (1994), who verified the consistency between predicted and validation data of different AMMI models by simple correlation. Two of the three years of data were currently used for modelling and the remaining year for validation, averaging the results of the three possible data splittings carried out, holding a different year for validation each time. The second approach (Procedure 2) is performing data splittings on single observations of GL combinations, i.e. regardless of the year factor, choosing method R of replicate selection in MATMODEL. This approach allows a considerable increase in the number of validation runs. However, the actual noise in the model (the GLY interaction) is completely confounded in this case with the year main effect as well as with the location-year (LY) and GY interactions. This is expected to underestimate the contribution to prediction of the GL-interaction PC axes, as modelling only accounts for genotype, location and GL-interaction effects. Therefore, the predictive assessment should be carried out after the removal of the year, LY-interaction and GY-interaction effects from the genotype-location-year data matrix submitted to cross validation. The need for the removal of each of these effects was empirically verified by performing the predictive assessment on original data, as well as on data adjusted for one, two, or all of the three effects. For adjustment, year levels were standardized to the same mean yield, LY-interaction effects were consistently eliminated across location-year combinations, and GY-interaction effects were consistently eliminated across genotype-year combinations. Another method for performing Procedure 2, which possibly offers a simpler alternative to the above mentioned adjustment of original data for the three ANOVA effects, is the execution of cross validations on data matrices of GL-interaction effects computed separately for each year and inputed in the programme as described for matrices of the original data. Also genotype and location main effects are set to zero in this case; but this is not expected to influence the assessment, as the tested hypothesis remains the same (i.e. whether modelling of GL-interaction effects in some data can be useful for predicting these effects in other data). The GL-interaction effects can easily be computed for each year by PROC GLM of the SAS software as residuals of a two-way ANOVA including the genotype and location factors. The identity of results provided by the two methods for performing Procedure 2 based, respectively, on adjustment of original data and on GL-interaction effects was preliminarily verified for each data set. Cross-validation results based on the splitting of single observations of GL combinations were always averages of 50 validation runs, holding two observations for modelling and one for validation.

Other testing criteria

Details of the different testing procedures were reported by Gollob (1968) for the F test he proposed, Cornelius et al. (1992) and Cornelius (1993) for the F_{GH2} test, Cornelius et al. (1992) and Piepho (1995) for the F_R test, and Gauch (1992 p. 147) for the assessment based on the signal-to-noise ratio. Cornelius (1993) also described the F_{GH1} test, as well as simulation tests based on the F_{GH1} and F_{GH2} statistics, which were not presently considered since they are expected to provide results very similar to the F_{GH2} test while requiring more computation. The u_1 and u_2 values in the F_{GH2} test were computed using the approximations given by Cornelius (1980).

Common applications of these tests, referring to the analysis of GE-interaction matrices, retained the pooled experiment error as the error term for computing F ratios or estimating the amount of noise in the interaction data matrix. For the current assessment of GL-interaction PC axes, such an error term has to be replaced by the GLY interaction, already indicated as the error for testing the entire GL interaction. When needed, the sum of squares and the mean squares of PC axes on a cell mean basis can be obtained by dividing their values by the number of years.

Results

The results of variance component estimation are reported in Table 1 for the four data sets. The GL interaction was statistically significant ($P \leqslant 0.001$) for all crops except oat, where it reached a significance of only $P \leqslant 0.12$. Its error term, the GLY interaction, was always significant ($P \leqslant 0.001$). The estimated variance of the GL interaction was higher than the genotype variance in durum wheat and maize, higher than the GY-interaction variance in all crops except oat, and lower than the GLY-interaction variance in all crops except durum wheat. Therefore, GL-interaction effects appeared of limited importance in oat, important in bread wheat and maize, and of special importance in durum wheat where they showed limited inconsistency between years and were associated with low variation for the genotype main effect.

Results of cross validations based on splitting of single observations of GL combinations are given in Table 2 for the original data and for data adjusted for the effects of year, LY interaction and GY interaction. As already mentioned, the test results based on an adjustment of original data for the three effects were identical to those based on the GL-interaction effects of individual years for cross-validation Procedure 2. Obviously, the removal of the three sources of variation

Table 1 Degrees of freedom (*df*) and estimate of variance components for yield (t ha^{-1} × 10^3) of four sets of crop variety trials (G = genotype factor; L = location; Y = year)

Source of variation	Bread wheat		Durum wheat		Maize		Oat	
	df	Variance	*df*	Variance	*df*	Variance	*df*	Variance
G	17	117 ***	8	19 NS	10	82 *	9	57 *
L	30	1313 ***	5	401 NS	10	1559 **	6	402 *
Y	2	150 *	2	0 NS	2	626 **	2	0 NS
G × L	510	43 ***	40	151 ***	100	196 ***	54	38 NS
G × Y	34	24 ***	16	23 *	20	53 *	18	62 **
L × Y	60	1457 ***	10	1340 ***	20	1240 ***	12	439 ***
G × L × Y	1020	179 ***	80	116 ***	200	211 ***	108	311 ***

NS, *, **, *** = not significant and significant at $P \leqslant 0.05$, $P \leqslant 0.01$ and $P \leqslant 0.001$, respectively

Table 2 Root mean square prediction difference (t ha^{-1}) of AMMI models including zero (AMMI 0) to three (AMMI 3) genotype-location interaction principal component axes for cross validations based on original data and splitting of complete years (CY), and original data or data adjusted for the effects of year, location-year interaction and genotype-year interaction and splitting of single observations (SO), for four sets of crop variety trials

Data set	AMMI model	Original[ab] data, CY	Original[a] data, SO	Adjusted[ac] data, SO
Bread wheat	0	1.645 X	1.230 X	0.538 X
	1	1.648	1.341	0.538
	2	1.651	1.418	0.542
	3	1.657	1.477	0.551
Durum wheat	0	1.455	1.181 X	0.494
	1	1.445	1.299	0.416
	2	1.434 X	1.373	0.382 X
	3	1.436	1.421	0.393
Maize	0	1.877 X	1.555 X	0.879
	1	1.897	1.707	0.871
	2	1.890	1.807	0.854 X
	3	1.896	1.868	0.873
Oat	0	1.041 X	0.915 X	0.597 X
	1	1.078	1.006	0.609
	2	1.089	1.054	0.628
	3	1.095	1.086	0.637

[a] Data of genotype-location combinations previously averaged across experimental replicates; model results averaged over three cross validations for CY and 50 cross validations for SO splittings; X = model with best predictive accuracy

[b] Procedure 1 in the text

[c] Equivalent to the splitting of genotype-location interaction effects computed for individual years; Procedure 2 in the text

from the original data decreased the RMSPD of all AMMI models as a result of the decreased variation between observations submitted to the same genotype and location effects. However, the decrease tended to be lower for models of greater complexity, leading to largely inconsistent results between original and adjusted data as regards model comparison. The model including no GL-interaction PC axes (AMMI 0) was distinctly the most accurate for all crops according to original data. The assessment based on adjusted data showed the superiority of AMMI 2 for durum wheat and maize, the distinct but less marked advantage of AMMI 0 for oat, and the substantial equality of AMMI 0 and AMMI 1 (only 0.0003 t ha^{-1} of difference in RMSPD) for bread wheat. Inconsistencies also took place between the assessment based on data adjusted for the three effects and those based on data adjusted for only some of the effects (data not reported), supporting on the whole a need for the removal of all of the three effects from the genotype-location-year data matrix in cross-validation Procedure 2. The adjustment for LY-interaction effects had a marked influence on the results of cross validations, as expected from the higher variance of this effect relative to those of year and GY interaction (Table 1).

The predictive assessment based on original data and splitting of complete years (Procedure 1) distinctly disagreed with the results of Procedure 2 for maize, in which AMMI 0 ranked first for accuracy followed by AMMI 2 and AMMI 1 (Table 2). Minor inconsistencies took place also for durum wheat and oat as concerns the relative differences in RMSPD value between models. The results for bread wheat were almost identical, the advantage of AMMI 0 over AMMI 1 remaining substantially negligible.

The results provided by the other testing criteria are shown in Table 3. The conclusions of these tests were in complete accordance only for the smallest data set, that

Table 3 Proportion of genotype-location (GL)-interaction sum of squares (SS), and assessment of GL-interaction principal component (PC) axes in AMMI models by four testing criteria, for four sets of crop variety trials

Data set	PC	GL SS (%)	Gollob's[ab] F test	F_{GH2}[ab] test	F_R[ac] test	Signal-to-noise[ad] ratio
Bread wheat	1	22	**	**	**	X
	2	16	**	**	**	X
	3	10	**	NS	NS	–
	4	9	**	NS	NS	–
	5	8	*	NS	NS	–
	6	6	NS	NS	NS	–
Durum wheat	1	63	**	**	**	X
	2	27	**	**	**	X
	3	5	NS	NS	NS	–
Maize	1	37	**	**	**	X
	2	26	**	**	*	–
	3	13	NS	NS	NS	–
Oat	1	55	**	*	NS	–
	2	20	NS	NS	NS	–

[a] Error term for PC axes is genotype-location-year interaction

[b] NS, *, ** = PC axis not significant and significant at $P \leqslant 0.05$ and $P \leqslant 0.01$, respectively

[c] NS, *, ** = residual GL-interaction SS after fitting higher ranking PC axes not significant and significant at $P \leqslant 0.05$ and $P \leqslant 0.01$, respectively

[d] X = PC axes included in the model

of durum wheat. Gollob's F test proved far more liberal than the other tests for the largest data set, that of bread wheat, where it declared four PC axes significant ($P \leqslant 0.01$) rather than only two as did the other tests. The F_{GH2} test tended to be more liberal than the F_R test for maize and oat. The test based on the signal-to-noise ratio provided results in consistent agreement with those of the F_R test considered at $P \leqslant 0.01$. These criteria disagreed with the conclusions of cross validation in two data sets, those of bread wheat and maize (Tables 2 and 3). For bread wheat, however, the predictive accuracy of the model including two PC axes was just slightly lower than that of AMMI 1.

Discussion

The four data sets, markedly different in the extent of GL interaction relative to genotype and GLY-interaction effects, are examples of different situations possibly occurring in multi-environment variety trials and may therefore provide information of interest for a wide range of AMMI analysis users.

With reference to cross validations for the testing of GE-interaction PC axes, Piepho (1994) showed that the assessment based on the splitting of single observations of GE combinations, regardless of the block factor for experiments laid out as a randomized complete block design, is biased towards the selection of simpler AMMI models, as a result of the inflated noise created by the confounding of the pooled experiment error with the block main effect. Piepho suggested, therefore, to split complete blocks within environments. In a similar way, the confounding of the year, LY-interaction and GY-interaction effects with the appropriate noise, i.e. the GLY interaction, in the original data submitted to a splitting of single observations of GL combinations produced a trend towards the selection of simpler AMMI models. One solution in this case could be the splitting of complete years (Procedure 1) which has, however, the drawback of permitting only a limited number of validation runs when trials are repeated over just a few years. Averaging the results of many different data splittings is recommended, especially when as few as about 200–300 validation observations are available (Gauch 1993 a p. 48). In this respect, the results of Procedure 1 cannot be considered reliable for relatively small genotype-location data matrices such as the current ones for durum wheat, maize and oat, which provided only 121–54 validation observations. The unreliability of Procedure 1 for the maize data set is also suggested by the rather peculiar shape of the RMSPD response in increasing the complexity of AMMI models, characterized by two minima (Table 2). Only one minimum is typically expected, corresponding to one maximum of statistical efficiency for the prediction known as Ockham's hill (Gauch 1993 b). The diagnosis of absence of pattern in the GL interaction of this data set also contrasts somewhat with the high level of ANOVA significance of the interaction ($P \leqslant 0.001$), although in principle a significant interaction may well contain only noise from a prediction standpoint (Crossa et al. 1990). The performance of a number of validation runs by different splittings of single observations of GL combinations, without inflating the appropriate error term for the assessment, is made possible by the current adjustment of original data for the effects of year and its first-order interactions with locations and genotypes (Procedure 2). However, this solution also has a major inconvenience represented by the fact that the adjusted data, or the GL-interaction effects of single years which can alternatively be used for the assessment, are not stochastically independent. This can introduce a bias in the test results difficult to quantify but expected to increase when decreasing the size of the GL-interaction data matrix. Indeed, Procedures 1 and 2 provided almost identical results for the only data set, that of bread wheat, where both could reasonably be considered fully reliable. It should be noted that the application of a third possible procedure for predictive assessment, based on the original data and the splitting of complete years within locations rather than throughout locations, which is analogous to the splitting of complete blocks within environments for the assessment of GE interactions, would allow an increase in the number of validation runs but would leave the LY and GY interactions completely confounded with the appropriate model noise.

The substantial equality in predictive accuracy between AMMI 0 and AMMI 1 for bread wheat makes difficult the model selection in this case. The choice of AMMI 1 is supported by the fact that the predictive assessment of PC axes based on the cross-validation criterion tends to be conservative because it refers to the modelling of a subset of the available data (two of three genotype-location observations), which is expected to be somewhat less accurate than modelling of all data performed ultimately after model choice (Cornelius 1993). Moreover, the choice of AMMI 1 is more consistent with the ANOVA results, which declared the entire GL interaction effect highly significant ($P \leqslant 0.001$).

The present results for the remaining test criteria agree well with the expectation that Gollob's F test has increased liberality by increasing the size of the interaction matrix submitted to principal components analysis (Cornelius 1993), suggesting to possibly adopt this test only for small GL-interaction data matrices. The tendency of the F_R test to be somewhat more conservative than the F_{GH2} test is in accordance with the simulation results of Piepho (1995). A reason here contributing to such a trend, representing a possible disadvantage of the F_R criterion, lies in the equivalence of the null hypothesis for the testing of AMMI models including no GL- (or GE)-interaction PC axes with that for ANOVA testing of GL (or GE) interaction. No first PC axis could ever be declared significant in the absence of significant interaction, even if it accounted for all of the interaction variation. No such restriction applies to the F_{GH2} criterion, as here shown by the fact that the first PC axis in the oat data set, accounting alone for most of the GL-interaction variation (Table 3), was declared significant ($P \leqslant 0.05$) by the test despite the lack of significance of the entire interaction in the ANOVA. However, the F_R test has the advantage, compared with the F_{GH2} test, of being more robust to departures from normality and homoscedasticity (Piepho 1995) and simpler to compute.

The results of cross validations, reportedly in close agreement with those of the criterion based on the signal-to-noise ratio (Gauch 1992) and of the F_{GH2} test (Cornelius 1993), presently differed for at least one of four data sets, the bread wheat one, which is also the data set for which the predictive assessment could be regarded as most reliable. This may emphasize the interest of the proposed cross-validation procedures for the assessment of GL-interaction PC axes. However, the mentioned drawbacks of these procedures should also be kept in mind for their application. Cross validations can be preferred to other testing criteria for relatively large GL-interaction data matrices, adopting Procedure 1 when several test years are available. Their results for PC testing could be interpreted with some liberality for the evaluation of modelling based on the complete data set when a high proportion (30–50%) of observations of GL combinations were used for validation in the assessment. Missing data for GL combinations of some years, quite frequent in large data sets, can be handled for modelling in the MATMODEL programme by means of an Expectation-Maximization algorithm (Gauch and Zobel 1990; Gauch 1993 a). Other testing criteria appear preferable for smaller data sets, for which the F_{GH2} and F_R tests at $P \leqslant 0.01$ and the estimate of predictive accuracy provided by the signal-to-noise ratio criterion may often lead to similar conclusions for model choice based on the current results.

The significant GL-interaction PC axes in the present data sets were agronomically interpretable, as reported in earlier studies which also pointed out the possible implications of the results on the breeding of each crop for the target area (Annicchiarico and Perenzin 1994; Annicchiarico et al. 1995; Annicchiarico and Mariani 1996; Annicchiarico 1997).

Acknowledgements I thank Dr. H. G. Gauch for making available the programme MATMODEL and for his useful suggestions on an earlier draft of the paper. I am also grateful to an unknown referee for his insightful comments. The Italian networks of bread wheat, maize and oat variety trials and the Agrimed research project "Durum wheat: breeding for higher and more stable yield in semi-arid Mediterranean zones", coordinated by the Experimental Institute for Cereal Crops, provided the experimental data.

References

Annicchiarico P (1992) Cultivar adaptation and recommendation from alfalfa trials in northern Italy. J Genet Breed 46 : 269–278

Annicchiarico P (1997) Joint regression vs AMMI analysis of genotype-environment interactions for cereals in Italy. Euphytica 94 : 53–62

Annicchiarico P, Mariani G (1996) Prediction of adaptability and yield stability of durum wheat genotypes from yield response in normal and artificially drought-stressed conditions. Field Crops Res 46 : 71–80

Annicchiarico P, Perenzin M (1994) Adaptation patterns and definition of macro-environments for selection and recommendation of common-wheat genotypes in Italy. Plant Breed 113 : 197–205

Annicchiarico P, Bertolini M, Mazzinelli G (1995) Analysis of genotype-environment interactions for maize hybrids in Italy. J Genet Breed 49 : 61–68

Cochran WG, Cox GM (1957) Experimental designs, 2nd edn. J Wiley and Sons, Chichester, UK

Cornelius PL (1980) Functions approximating Mandel's tables for the means and standard deviations of the first three roots of a Wishart matrix. Technometrics 22 : 613–616

Cornelius PL (1993) Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. Crop Sci 33 : 1186–1193

Cornelius PL, Sayedsadr M, Crossa J (1992) Using the shifted multiplicative model to search for "separability" in crop cultivars trials. Theor Appl Genet 84 : 161–172

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

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

De Lacy IH, Fox PN, Corbett JD, Crossa J, Rajaram S, Fischer RA, Van Ginkel M (1994) Long-term association of locations for testing spring wheat. Euphytica 72 : 95–106

Gauch HG (1992) Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier, Amsterdam

Gauch HG (1993 a) MATMODEL Version 2.0. Cornell University, Ithaca

Gauch HG (1993 b) Prediction, parsimony and noise. Am Scientist 81 : 468–478

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 (1990) Inputing missing yield trial data. Theor Appl Genet 79 : 753–761

Gollob HF (1968) A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33 : 73–115

Léon J, Becker HC (1988) Repeatability of some statistical measures of phenotypic stability — correlations between single-year results and multi-year results. Plant Breed 100 : 137–142

Mandel J (1971) A new analysis of variance model for non-additive data. Technometrics 13 : 1–8

Piepho HP (1994) Best Linear Unbiased Prediction (BLUP) for regional yield trials: a comparison to additive main effects and multiplicative interaction (AMMI) analysis. Theor Appl Genet 89 : 647–654

Piepho HP (1995) Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trials. Theor Appl Genet 90 : 438–443

Weber WE, Westermann T (1994) Prediction of yield for specific locations in German winter-wheat trials. Plant Breed 113 : 99–105

Wricke G, Weber WE (1986) Quantitative genetics and selection in plant breeding. De Gruyter, Berlin New York