Theor Appl Genet (1985) 71:314–319

An adequate design for regression analysis of yield trials

L. Gusmão
Department of Genetics, Estação Agronómica Nacional, P-2780 Oeiras, Portugal

Accepted April 28, 1985
Communicated by J. MacKey

Summary. Based on theoretical demonstrations and illustrated with a numerical example from triticale yield trials in Portugal, the Completely Randomized Design is proposed as the one suited for Regression Analysis. When trials are designed in Complete Randomized Blocks the regression of plot production on block mean instead of the regression of cultivar mean on the overall mean of the trial is proposed as the correct procedure for regression analysis. These proposed procedures, in addition to providing a better agreement with the assumptions for regression and the philosophy of the method, induce narrower confidence intervals and attenuation of the hyperbolic effect. The increase in precision is brought about by both a decrease in the t Student values by an increased number of degrees of freedom, and by a decrease in standard error by a non proportional increase of residual variance and non proportional increase of the sum of squares of the assumed independent variable. The new procedures seem to be promising for a better understanding of the mechanism of specific instability.

Key words: Regression analysis – Yield trials – Triticale – Instability

Introduction

Regression analysis, as a tool to assess genotypes for their adaptation to a range of environments, was developed by Finlay and Wilkinson (1963) twenty-five years after the proposal of this same technique by Yates and Cochran (1938), as a means for further analysis of genotype-environment interaction in groups of experiments.

Further refinements to the method (e.g. Eberhart and Russel 1966; Perkins and Jinks 1968) have been particularly related to the aim of analysis, ignoring the design of trials conforming to the regression model.

In spite of criticisms (Baker 1969; Witcombe and Whittington 1971; Freeman and Perkins 1971; Hardwick and Wood 1972; Shukla 1972; Freeman 1973; Hill 1975; Wright 1976; Fischer and Maurer 1978; Brennan and Bith 1979; etc.) – often emerging from the particular work conditions and the handling of the results – the method proved to be quite effective in many situations.

Although there is no independence of the predictor variable, there has been no evidence for not accepting the linearity of the underlying regression and the normal distribution of the errors. Further, constance of errors has not been proved to be rejected in the regions of the regression to be met in practice.

As far as an ultimate screening for cultivars (which have already been selected for a particular area) is concerned, the regression technique allows for a better understanding of its possible relative performance within the range of environmental indexes which were tested.

As a matter of fact, in this particular situation, cultivars to be tested in a regional programme to forecast its relative performance in field conditions should have been previously scrutinized for adequacy with climatic rythms, and resistance to prevalent diseases in that particular region, where they are to be licenced, so that specific instability (in the sense defined by Joppa et al. 1971) would not be expected.

The method turns out much more efficient the wider the range of the putative yielding ability of the environment, which can be approximated in terms of local edaphic conditions and technical inputs available (and its relations with climatic limitations).

As Cooper (1970) stressed, certain environmental limitations can be avoided by technical measures, but in some situations the only solution may be the selection of adapted genetic material. On the other hand, low-yielding environments coexist very often together with high-yielding ones in the same target region for the trials. In this situation it is important to assess, for the same region on a whole, which cultivars are more suited in each case.

Although in most crops substantial progress has been made in developing cultivars with both an improved stability of performance and a high yielding ability, seldom is this

desideratum achieved whenever the environmental productivity levels to be found in practice greatly vary.

A dynamic method for evaluating relative cultivar behaviour can then be most useful, providing uniform geoclimatic situations are available. Linear regression of the mean yield of individual cultivars on mean yield of all cultivars (and variations within this same principle) has widely proved to be such a method because of its predictive value under the described conditions, although always applied on data resulting from trials designed without taking into account the implied assumptions for regression.

Complete Randomized Blocks (CRB) has been the most frequently used design, even if Split-plots (cf. e.g. Becker et al. 1982) and single or multiple Lattice designs (cf. e.g. Hill et al. 1983; Gama and Hallauer 1980; Ghaderi et al. 1980) have been used in the context previously referred to.

In the present study we intend to define the most suitable design in order to increase the accuracy and precision of the regression method.

Comparison of regression lines for the same cultivar based either on trial average or block average

In the model proposed by Perkins and Jinks (1968 a, b) which predicts the performance y_{ij} of the i^{th} line in the j^{th} environment ($j = l, \ldots, s$), or even in the model extended by Freeman and Perkins (1971) to the r replicates of y_{ijk} ($k = l, \ldots, r$), the genotype-environment interaction of the i^{th} line in the j^{th} environment, g_{ij}, consists of a linear portion $\beta_i \varepsilon_j$ and a deviation δ_{ij} from regression (being ε_j the additive environmental contribution of the j^{th} environment).

Homogeneity within blocks and heterogeneity between blocks are implied in CRB designs, so, seeing that the environmental index is no more than a quantitative grading of the environment, each block should be regarded as a different environment.

Thus, assuming that the average of all cultivars for environmental index is x_j and e_{ij} is the observed deviation of y_{ij} from the regression line, instead of being

$$y_{ij} = \beta_{io} + \beta_{il} x_j + e_{ij} \tag{1}$$

for the regression line, we should have

$$y_{ijk} = \beta'_{io} + \beta'_{il} x_{jk} + e_{ijk}, \tag{2}$$

where e_{ijk} is the observed deviation of y_{ijk} from the regression line.

The least square line parameter estimators, will therefore have, respectively, the following forms,

$$\hat{\beta}_{il} = \frac{s \sum_j x_j y_{ij} - \left(\sum_j x_j\right)\left(\sum_j y_{ij}\right)}{s \sum_j x_j^2 - \left(\sum_j x_j\right)^2}$$

and

$$\hat{\beta}_{io} = \bar{y}_i - \hat{\beta}_{il} \bar{x}$$

in the first case, and

$$\hat{\beta}'_{il} = \frac{s\,r \sum_j \sum_k x_{jk} y_{ijk} - \left(\sum_j \sum_k x_{jk}\right)\left(\sum_j \sum_k y_{ijk}\right)}{s\,r \sum_j \sum_k x_{jk}^2 - \left(\sum_j \sum_k x_{jk}\right)^2}$$

and

$$\hat{\beta}'_{io} = \bar{y}_i - \hat{\beta}'_{il} \bar{x}$$

in the second case.

So, lines (1) and (2) have a common point "$\bar{x}, \bar{y}_i$" and, due to the relations

$$(1/r) \sum_k x_{jk} = x_j$$

and

$$(1/r) \sum_k y_{ijk} = y_{ij}$$

which give,

$$\hat{\beta}_{il} = \frac{s \sum_j \left(\sum_k x_{jk}\right)\left(\sum_k y_{ijk}\right) - \left(\sum_j \sum_k x_{jk}\right)\left(\sum_j \sum_k y_{ijk}\right)}{s \sum_j \left(\sum_k x_{jk}\right)^2 - \left(\sum_j \sum_k x_{jk}\right)^2}$$

we can see that $\hat{\beta}_{il}$ only equals $\hat{\beta}'_{il}$ (and so do the respective lines) when the mean yields of each block of the same trial are equal, which doesn't conform to the design assumption.

Adequate design for regression and alternative procedures for Complete Randomized Blocks

The logic of the restriction, which was imposed for the hypothesis of identity between regression coefficients, relies upon the intrinsic philosophy informing the probabilistic model of the regression, in which the independent variable is assumed as fixed, and the dependent variable is characterized by random variation.

Consequently, a conceptually more adequate model for regression analysis of cultivar mean yield on the mean yield of all cultivars, should be based on Completely Randomized designs (CR), where homogeneity in the whole trial is implied and safeguarded by the implicit method of randomization.

In most of the CRB designs, as applied in field yield trials, the significance of the variance between blocks is seldom revealed and so the overall mean yield in the trials can be assumed as the real block yield mean. In this situation, a more accurate procedure would estimate the linear relationship by means of the regression of cultivar plot yield on the overall mean yield of the trial (so, simulating the CR design).

Such a procedure will lead, as previously shown, to estimators identical to those attained with the standard

method, but keeping the total amount of degrees of freedom (d. f.).

The high determination coefficients arising from the standard procedure (r_i^2) can be shown to be biased by overestimation, in comparison with this same parameter estimate ($r_i'^2$) as evaluated by the CR simulation procedure we just mentioned.

As a matter of fact, being

$$r_i^2 = \frac{\left[\sum_j x_j y_{ij} - \left(\sum_j x_j\right)\left(\sum_j y_{ij}\right)/s\right]^2}{\left[\sum_j x_j^2 - \left(\sum_j x_j\right)^2/s\right]\left[\sum_j y_{ij}^2 - \left(\sum_j y_{ij}\right)^2/s\right]}$$

and

$$r_i'^2 = \frac{\left[\sum_j x_j y_{ij} - \left(\sum_j x_j\right)\left(\sum_j y_{ij}\right)/s\right]^2}{\left[\sum_j x_j^2 - \left(\sum_j x_j\right)^2/s\right]\left[\left(\sum_j \sum_k y_{ijk}^2\right)/r - \left(\sum_j y_{ij}\right)^2/s\right]}$$

we can see that the difference between these two expressions is in the denominator of the first term of the second expression and, as for r_i^2,

$$\sum_j y_{ij}^2 = (1/r) \sum_j (1/r) \left(\sum_k y_{ijk}\right)^2$$

and for $r_i'^2$,

$$\left(\sum_j \sum_k y_{ijk}^2\right)/r = (1/r) \sum_j \sum_k y_{ijk}^2 ,$$

being

$$y_{ijk} \geqq 0 ,$$

we will have

$$(1/r) \left(\sum_k y_{ijk}\right)^2 \leqq \sum_k y_{ijk}^2 ,$$

and so

$$r_i'^2 \leqq r_i^2 .$$

The equality can only arise on the stocasticly impossible hypothesis of equality of the values y_{ijk} for all k (due to the random character of this variable), and the inequality grows with the number of replicates in each trial and the number of trials, which shows that the standard procedure will seldom produce unbiased results.

On the other hand, and again in the theoretical assumption of equality among block mean yields, assuming constant variance and expectation zero for the residuals (respectively e_{ij} and e_{ijk} in each case) there will be an identity of variances of the parameter estimators, because:

Making

$$V(e_{ij}) = \sigma_i^2$$

and

$$V(e_{ijk}) = \sigma_i'^2 ,$$

respectively for the regression, based either on standard procedure or CR simulation procedure, we will have

$$\sigma_i^2 = \sigma_i'^2/r ,$$

as e_{ij} is the average of the r values e_{ijk} in each trial.

Variances of the parameter estimators in each procedure will so result, in the first case:

$$V(\hat{\beta}_{io}) = \frac{(\sigma_i'^2/r) \sum_j x_j^2}{s \sum_j (x_j - \bar{x})^2}$$

for interception, and

$$V(\hat{\beta}_{il}) = \frac{\sigma_i'^2/r}{\sum_j (x_j - \bar{x})^2}$$

for slope, and in the second case:

$$V(\hat{\beta}'_{io}) = \frac{\sigma_i'^2 \sum_j \sum_k x_{jk}^2}{s\, r \sum_j \sum_k (x_{jk} - \bar{x})^2}$$

and

$$V(\hat{\beta}'_{il}) = \frac{\sigma_i'^2}{\sum_j \sum_k (x_{jk} - \bar{x})^2} ,$$

respectively for those same parameter estimators.

Simplification of the variance expression in the second case, based on the assumed relations,

$$\sum_k x_{jk}^2 = r\, x_j^2$$

and

$$\sum_k (x_{jk} - \bar{x})^2 = r\, (x_j - \bar{x})^2 ,$$

will result in

$$V(\hat{\beta}'_{io}) = \frac{(\sigma_i'^2/r) \sum_j x_j^2}{s \sum_j (x_j - \bar{x})^2}$$

and

$$V(\hat{\beta}'_{il}) = \frac{\sigma_i'^2/r}{\sum_j (x_j - \bar{x})^2} ,$$

which are the same as for the standard procedure.

Concerning the inference of the mean response, and yet based on the assumptions we have being made, the estimated standard error for mean response at an x* value of x can be expressed, in each case, by

$$\text{S.E.} = \sqrt{(\sigma_i'^2/r) \left[(1/s) + \frac{(x^* - \bar{x})^2}{\sum_j (x_j - \bar{x})^2}\right]} ,$$

and

$$\text{S.E.}' = \sqrt{\sigma_i'^2 \left[(1/s\, r) + \frac{(x^* - \bar{x})^2}{\sum_j \sum_k (x_{jk} - \bar{x})^2}\right]} ,$$

which are equal, under the condition of equality between blocks of the same trial.

Consequently, the increased precision induced by CR design vs. CRB design may be approximated as an exclusive function of the t probabilistic values, to be found in the Student's distribution, based on $sr-2$ d.f. for CR design and $s-2$ d.f. for CRB design, within each level of significance.

In practical terms, we may anticipate that regression based on CRB design by the standard procedure can only result in a reasonable approximation when: a) the block means in each trial are very similar; b) the number of trials is high; c) the number of replicates and the level of significance are low.

In an ultimate situation, where only one replicate exists, CRB design will be the same as CR design, which leads to what we call the correct procedure for regression analysis on CRB designs. However, in this situation, the constancy and independence of errors could not be so clearly depicted from scatter plot diagrams based on observations.

In Table 1 the increase of t values can be seen and, as a result, the decrease of precision by enlarged confidence intervals (once standard errors are the same in both cases), when regression is based in standard procedure instead of CR simulation procedure, in several hypothetical situations of 5 replicate trials of equal block means.

In Fig. 1 we can visualise what has been said concerning increase of accuracy and precision.

Table 1. Increase of tabulated t values for inference, following standard procedure vs. CR simulation procedure, in trials with 5 replicates assuming equal block means at different number of trials and different levels of significance (S.L.)

No. of trials	d.f.		100(1−α/2)% S.L.	
	(s−2)	(sr−2)	95	99
5	3	23	54%	108%
10	8	48	15%	25%
15	13	73	8%	14%

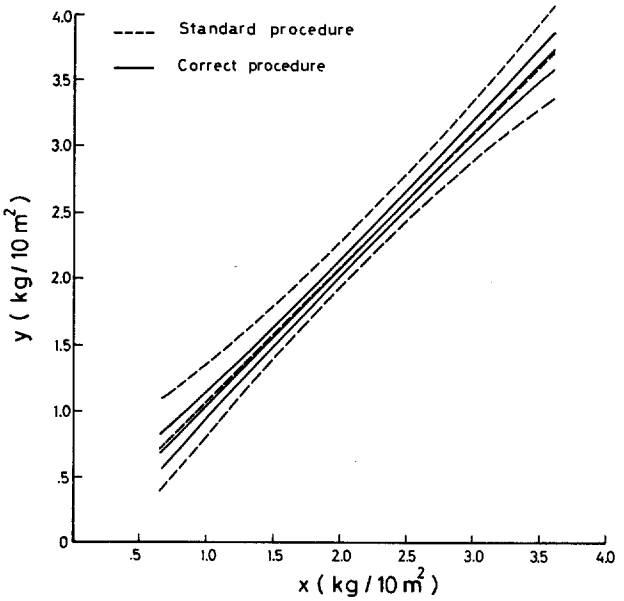

Fig. 1. Comparison of the regression lines and respective limit lines for 95% confidence intervals for strain 75IT8, when evaluated both by standard procedure and correct procedure

Numerical example

Data for illustration purposes were taken from triticale yield trials in Portugal (Cidraes and Gusmão 1982; Gusmão and Cidraes 1982).

Only triticale strains common to all trials were considered and the different analysis procedures were applied as if they constituted the only ones in the trial.

A previous analysis (through regression using what we have named correct procedure) showed a progressive increase in the determination coefficients, when a greater number of sites, beginning at a climatically homogeneous area was successively put into analysis, except when sites of the northern part of the country (known to be geoclimatically different) were included.

This fact confirms, on one hand, that only geoclimatically homogeneous regions should be considered for regression purposes (otherwise inducing to unacceptable levels the determination coefficients for linear regression) and, on the other hand, that for adaptation purposes, the whole area located south of the "Montejunto-Estrela" mountaineous system (representing more than half of the whole country) is much more homogeneous than traditionally thought. The main differences rely mainly in pedological features and technical facilities.

Consequently, only 25 out of 31 trials, located at 11 sites of this area, during the years 76/77 to 78/79, were considered for the present study.

The trials under analysis were designed in CRB with 5 replicates, and three distinct procedures were put into comparison, namely: standard procedure (regression of y_{ij} values on x_j values); correct procedure (regression of y_{ijk} values on x_{jk} values); and CR simulation procedure (regression of y_{ijk} values on x_j values).

In Table 2 estimates of intercept, slope and determination coefficients, as evaluated by these procedures, are presented. Table 3 contains the residual variations and the standard errors at three different points of the predictor variable.

Discussion and conclusions

As actual differences between blocks exist in some of the trials, the CR simulation procedure cannot produce so high correlation coefficients as if trials had been designed as CR. Nevertheless, an identity for this procedure with standard procedure in what concerns

Table 2. Intercept (I), slope (S1), and determination coefficient (r^2), as evaluated by the three procedures (y_{ij} on x_j, y_{ijk} on x_{jk}, and y_{ijk} on x_j), expressed in kg/10 m²

Strains	y_{ij} on x_j			y_{ijk} on x_{jk}			y_{ijk} on x_j		
	I	S1	r^2	I	S1	r^2	I	S1	r^2
72IT7	−0.177	1.04	0.94	−0.179	1.04	0.86	−0.177	1.04	0.83
72IT13	−0.188	1.04	0.92	−0.189	1.04	0.85	−0.188	1.04	0.82
74IT8	0.0772	0.935	0.95	0.0912	0.928	0.78	0.0772	0.935	0.76
75IT6	0.0965	0.946	0.96	0.0941	0.947	0.83	0.0965	0.946	0.80
75IT7	0.0330	0.986	0.88	0.0470	0.980	0.79	0.0330	0.986	0.77
75IT8	0.0427	1.01	0.95	0.0193	1.02	0.84	0.0427	1.01	0.79
75IT12	−0.128	1.05	0.95	−0.117	1.05	0.87	−0.128	1.05	0.84
75IT13	0.0311	1.02	0.94	0.0394	1.01	0.84	0.0311	1.02	0.82
75IT14	0.0798	1.02	0.92	0.0488	1.03	0.81	0.0798	1.02	0.76
75IT16	0.133	0.963	0.91	0.145	0.957	0.81	0.133	0.963	0.79

Table 3. Residual variance ($S^2_{y/x}$) and standard error (S.E.) for $x^* = \bar{x}$ (x_1^*), $x^* = \bar{x} \pm 1$ kg/10 m² (x_2^*), and $x^* = \bar{x} \pm 1.5$ kg/10 m² (x_3^*), expressed in kg/10 m²

Strains	y_{ij} on x_j				y_{ijk} on x_{jk}				y_{ijk} on x_j			
	$S^2_{y/x}$	S.E.			$S^2_{y/x}$	S.E.			$S^2_{y/x}$	S.E.		
		x_1^*	x_2^*	x_3^*		x_1^*	x_2^*	x_3^*		x_1^*	x_2^*	x_3^*
72IT7	0.0556	0.0976	0.147	0.192	0.144	0.0339	0.0507	0.0658	0.180	0.0751	0.113	0.148
72IT13	0.0749	0.113	0.171	0.222	0.150	0.0346	0.0517	0.0672	0.186	0.0764	0.115	0.150
74IT8	0.0400	0.0828	0.125	0.162	0.196	0.0396	0.0591	0.0768	0.214	0.0819	0.123	0.161
75IT6	0.0291	0.0710	0.106	0.139	0.151	0.0348	0.0519	0.0674	0.182	0.0756	0.114	0.148
75IT7	0.110	0.137	0.207	0.269	0.208	0.0408	0.0609	0.0791	0.229	0.0848	0.128	0.166
75IT8	0.0437	0.0865	0.130	0.170	0.162	0.0360	0.0537	0.0698	0.213	0.0817	0.123	0.160
75IT12	0.0539	0.0961	0.145	0.189	0.135	0.0329	0.0491	0.0637	0.162	0.0713	0.107	0.140
75IT13	0.0564	0.0983	0.148	0.193	0.159	0.0357	0.0532	0.0692	0.186	0.0764	0.115	0.150
75IT14	0.0783	0.116	0.174	0.227	0.200	0.0400	0.0597	0.0776	0.257	0.0898	0.135	0.176
75IT16	0.0782	0.116	0.174	0.227	0.177	0.0376	0.0562	0.0730	0.198	0.0788	0.119	0.155

the estimates of intercept and slope (as was demonstrated in the last chapter) can be seen in Table 2.

In this same Table, the bias of the standard procedure, as checked by the correct procedure, is obvious. The differences, in many cases, may not be very relevant, but the main disadvantage of applying the standard procedure in the regression analysis is the loss of d.f. (as was shown in the last chapter) and its projection first in the *t* values, and then in the confidence limits.

In Table 3, where the advantage of the correct procedure over the standard procedure is evident (and, in most cases, even that of the CR simulation procedure), two conclusions can be drawn.

First, the difference of the residual variance from standard to correct or even CR simulation procedure is not proportional to the number of replicates. This fact could be assigned to random variation, but its repeatability, seen in Table 3, suggests a stronger mechanism related to the asymmetry within each trial of the distribution of the y_{ijk} values in relation with the regression line.

Further analysis of this phenomenon and the search for possible relationships with the characteristics of the particular environment where it is observed, can be a valuable means of elucidating what has been called specific instability according with the side (respectively underneath or above) of the regression line where they tend to concentrate in a particular environment.

Comparing the standard error for correct and CR simulation procedure, the distortion brought about by the simulation induces a greater residual variance to this last procedure, not to be expected if the correct design (CR) was used. Furthermore, as the sum of squares of the deviations for x values (S^2_x), in view of

the Schwarz inequality, is always higher in the correct procedure, this one benefits of a lower standard error, as it can be seen in Table 3.

Higher values of S_x^2 in the correct procedure, together with the equality of the number of d.f. in both these two above-mentioned procedures lead us to the second relevant conclusion, which is the comparatively lower hyperbolic effect on the confidence limit line for the correct procedure, also apparent in Table 3. This effect and general increase of precision can be visualized in Fig. 1.

In view of theoretical demonstrations, illustrated with a numerical example, we can summarize:

a) a Completely Randomized design is the most adequate design for Regression Analysis;
b) whenever a set of trials performed in Complete Randomized Blocks is to be analysed through Regression, the analysis should be based on block results instead of trial results (i.e. y_{ijk} on x_{jk} instead of y_{ij} on x_j);
c) in addition to having a greater accuracy, the proposed procedures, a) and b), induce a higher precision for inference purposes;
d) lack of assumption of independence of errors, whenever suspected by data analysis, can be used for elucidating possible specific instability according to the pattern of distribution of the y_{ijk} values in relation to the regression line;
e) although the linearity of the regression line can be strongly supported by this and many other works with several species, the choice of locations within a homogeneous region for performing the trials should not only emphasize extreme situations but try to cover the greatest diversity of environments in order to allow for the detection of specific instability.

Acknowledgements. I am grateful to Professor M. Mota, Department of Genetics, "Estação Agronómica Nacional", for his encouragement and helpful criticism, to Professor A. StAubyn, "Instituto Superior de Agronomia", Technical University of Lisbon, for his teaching, computer facilities at the Informatic Center, and revision of the theoretical demonstrations; and to Dr. L. Costa-Rodrigues, Department of Genetics, "Estação Agronómica Nacional", for his kind help in the revision of the manuscript.

References

Baker RJ (1969) Genotype-environment interactions in yield of wheat. Can J Plant Sci 49:743–751

Becker HC, Geiger HH, Morgenstern K (1982) Performance and phenotypic stability of different hybrid types in winter rye. Crop Sci 22:340–344

Brennan PS, Byth DE (1979) Genotype-environmental interactions for wheat yields and selection for widely adapted wheat genotypes. Aust J Agric Res 30:221–232

Cidraes AG, Gusmão L (1982) Adaptação de triticales em Portugal. 1. Comportamento das linhas e variedades em ensaio no triénio 1976/77 a 1978/79. Melhoramento 27: 161–201

Cooper JP (1970) Environmental physiology. In: Frankel O, Bennet E (eds) Genetic resources in plants. Blackwell Sci Publ, Oxford Edinburgh, pp 131–142

Eberhart SA, Russel WA (1966) Stability parameters for comparing varieties. Crop Sci 6:36–40

Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant breeding programme. Aust J Agric Res 14: 742–754

Fischer RA, Maurer R (1978) Drought resistance in spring wheat cultivars. 1. Grain yield responses. Aust J Agric Res 29:897–912

Freeman GH, Perkins JM (1971) Environmental and genotype-environmental components of variability. 8. Relations between genotypes grown in different environments and measures of these environments. Heredity 27:15–23

Gama EG, Hallauer AR (1980) Stability of hybrids produced from selected and unselected lines of maize. Crop Sci 20: 623–626

Ghaderi A, Everson EH, Gress CE (1980) Classification of environments and genotypes in wheat. Crop Sci 20: 707–710

Gusmão L, Cidraes AG (1982) Adaptação de triticales em Portugal. 2. Análise do comportamento quantitativo em relação às espécies parentais. Melhoramento 27:203–223

Hill J (1975) Genotype-environment interaction a challenge for plant breeding. J Agric Sci 85:477–493

Hill RR Jr, Baylor JE (1983) Genotype x environment interaction analysis for yield in alfalfa. Crop Sci 23:811–815

Joppa LR, Lebsock KL, Busch RH (1971) Yield stability of selected spring wheat cultivars (*Triticum aestivum* L. em Thell) in the uniform regional nurseries, 1959 to 1968. Crop Sci 11:238–241

Perkins JM, Jinks JL (1968 a) Environmental and genotype-environmental components of variability. 3. Multiple lines and crosses. Heredity 23:339–356

Perkins JM, Jinks JL (1968 b) Environmental and genotype-environmental components of variability. 4. Non-linear interactions for multiple inbred lines. Heredity 23: 525–535

Shukla GK (1972) Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29:237–245

Teich AH (1983) Genotype-environment interaction variances in yield trials of winter wheat. Cereal Res Commun 11: 15–20

Witcombe JR, Whittington WJ (1971) A study of the genotype seeds of *Brassica napus*. Heredity 26:397–411

Wright AJ (1976) Bias in the estimation of regression coefficients in the analysis of genotype-environmental interaction. Heredity 37:299–303

Yates F, Cochran WG (1938) The analysis of groups of experiments. J Agric Sci 28:556–580