Theor Appl Genet (1986) 72, 98-104

Inadequacy of blocking in cultivar yield trials*

L. Gusmão

Genetics Department, Estação Agronómica Nacional, P-2780 Oeiras, Portugal

Received July 20, 1985; Accepted November 25, 1985
Communicated by J. MacKey

Summary. Based on the literature, theoretical considerations and a numerical example on triticale, Complete Randomized Blocks design is shown to be inadequate for cultivar yield trial purposes. Assumptions required for validity and convenience are shown not to be verified throughout most of the published experiments as well as in the present numerical example. It has been referred to the difficulty in forecasting homogeneity within blocks together with heterogeneity between blocks. This is difficult to achieve even in well-known experimental fields, let alone local fields chosen at random, which leads to unacceptably low correlation between plots from block to block in each trial. Heteroscedasticity, as supported by different regression coefficients in Joint Regression Analysis, does not allow for ANOVA, unless the overall variation of soil fertility level is reduced to an amount comparable with that expected for the unknown errors. In this instance, the loss of degrees of freedom in the two-way ANOVA is known not to be compensated for by block effect deduction. The need to generalize trial results calls attention to the emphasis that should be given to cultivar performance pattern within the area they are to be released. Thus, we advocate the need for precise point evaluations in order to give accurate estimation of that pattern. Genotype-environment interaction, within situations where ecological diversity does not include stress mechanisms of different natures, can be reasonably described through its genotype-fertility level component, since specific instability, related to climatic features, is supposed to be strongly reduced by the screening process of both cultivar production and introduction. Sensitivity of the regression techniques (even through robust methods) requires a broad range of trial fertility levels and, besides an adequate number of degrees of freedom and detection of eventual remaining specific instabilities, demands an experienced evaluation of particular ecological situations; however, randomization is not required except within trials, which should be designed as completely randomized. To carry on trials beyond one year is not an *a priori* demand and should only be considered when very abnormal seasonal conditions occur.

Key words: Yield trials – Complete Randomized Blocks – Triticale – Joint Regression Analysis – Genotype-environment interaction

Introduction

In assessing cultivars for yield, we are not usually concerned with the narrow limits within the experimental conditions, but rather with a wider situation for which individual trials are supposed to be representative.

On the other hand, as Student (1923) claimed "the object of testing varieties of cereals is to find out which will pay the farmer best"; this we can easily extend to all possible crops.

The "static statistically-based models" of analysis (Rose 1975) can only respond in a global sense as they make use of average effects. It appears convenient to replace these models in favour of more dynamic ones capable of providing more specific answers.

It is since Tuckey (1977), that a growing interest has been turned to modern methods of robust analysis as a way to overcome the need for previous assumptions which seldom could be proved. Nevertheless, as Hoaglin et al. (1983) advocate, whenever Gaussian behaviour of errors applies, classic statistical methods are still the most convenient.

Through many "uniformity trials" (or the "blank experiments" of Fairfield Smith (1938)), errors in field trials have been proved to be in accordance with normal distribution,

* This research was partially supported by the Calouste Gulbenkian Foundation, Lisboa

providing an adequate randomization had been adopted. Furthermore, as Student (1908) pointed out, "the deviation from normality must be very extreme to lead to serious error", and as Kempthorne (1952) said, "tests of significance of treatment effects, obtained by the analysis of variance, when randomization is used, are reliable, regardless of the distribution involved".

Multiple regression, though useful in enlightning the physiological bases of yield, fails to give an objective answer, namely in what concerns cultivar relative performance in the diversity of potential environments of a crop. As a matter of fact "even where many climatic and other measurements are reliable their meaningful integration in biological terms is not readily reproduced in mathematical terms" (Goodchild and Boyd 1975).

The description of such a pattern of relative performance seems to be possible, not in terms of the different environment variables but in terms of their integration in a quantitative environmental index. Yates and Cochran (1938) showed that the "association between varietal differences and general fertility" could be "further investigated by calculating the regression of the yields of the separate varieties on the mean of all varieties".

Even then those authors stated, in that same paper, that "the average difference between varieties, is of considerable importance even when this response varies from place to place or from year to year. For unless we both know the causes of this variation and can predict the future incidence of these causes we shall be unable to make allowance for it, and can only base future practice on the average effects".

Probably because of this last statement and due to the recognized authority of these scientists, further development of the technique, known today as "Joint Regression Analysis" (JRA), was first resumed 25 years later (Finlay and Wilkinson 1963). It is not yet sufficiently widespread as a tool for assessing cultivars in a regional scheme since the conditions for its applicability have not been established.

Nevertheless, the increasing indication of the linear behaviour of yield, under diverse environmental conditions, as tested by JRA, is enough to conclude that, very often, a large component of interaction "cultivar-fertility level" can not allow for a correct evaluation of errors by elimination of block effects, unless: (a) the joint regression lines for the different cultivars are coincident or parallel (this rarely can be expected), or (b) the block effect is only the reflection of the within block errors (a bias towards error evaluation is introduced by the incomplete randomization; no advantage would be brought that could surmount the loss of degrees of freedom).

By means of a simple mathematical approach I have already demonstrated the inadequacy of blocking for JRA in cultivar yield trials as it leads to loss of precision (Gusmão 1985).

Though I do not dispute the accuracy and precision of Complete Randomized Blocks design (CRB) whenever very strong correlation between treatments exists, I stress its bias and lack of precision in most of the situations concerning yield trials.

Theoretical considerations

Blocking in general

The idea of blocking as a means to minimize experimental errors and, at the same time, to provide a correct estimate of those errors by valid statistical methods, was shown by Student (1908) as applied to a series of duplicate experiments. Assuming a normal distribution of errors, he discussed the data of Cushny and Peebles on sleeping drugs and stated: "the low value of the S.D. is probably due to the different drugs reacting similarly on the same patient, so that there is a correlation between the results."

The underlying theory of reducing the block effect can, thus, be expressed in the following sentence: "the art of designing all experiments is even more in arranging matters so that r_{AB} is as large as possible, than in reducing σ_A^2 and σ_B^2" (Student 1923).

To illustrate the problems which arise when comparing several varieties grown together on a "chess--board" Student (1923) used an experiment from Beaven.

Stressing the bias of errors for systematic arrangement of treatments, in his classic "Statistical Methods for Research Workers", Fisher (1925) launches the method of design and analysis of CRB.

The main scope was to provide a simple "layout which fulfils the condition of supplying a valid estimate of error and at the same time possesses the advantage of eliminating a substantial fraction of the soil heterogeneity" (Fisher and Wishart 1930). As in the expression of Haydock and Sandland (1975), blocking is, consequently, "an attempt to concentrate the environmental differences in a removable main effect".

As far as the null hypothesis is concerned, namely that the different samples are drawn from the same population, the equality of the variances is a necessary part (Fisher 1942), which does not imply the à priori knowledge of such variances. However, the assumption of no interaction between main effects is essential in order to validate the test of the main effects against residual variation.

In summary, apart from the fact that block means cannot be evaluated without error, three points must be considered in order to both take advantage from blocking and allow for correct hypothesis test: (i) homogeneity within blocks, (ii) heterogeneity between blocks and (iii) no interaction between effects to be tested.

Blocking in cultivar yield trials

Regarding the above mentioned first two points, which concern efficiency of blocking, some considerations about the causes of variation in the yield of a crop are pertinent. Two kinds of such causes are recognized by Student (appendix to Mercer and Hall 1911): those which are random, and those which "occur with more regularity increasing from point to point or having centers from which they spread outwards".

Though these two causes may be considered globally in their frequencies as Gaussian, as far as complete randomization is put forward (see Wood and Stratton 1910; Fairfield Smith 1938), each one deserves more careful attention.

The first ones, random as they are, do not concern our present purpose and may, obviously, be reduced by refinement of the measurements and care in performing the experiment. The second ones, which refer to variation in soil fertility, are those we intend to discuss more carefully.

Errors of this kind demand criterious randomization in order to be evaluated, and the way they have been described by Student (appendix to Mercer and Hall 1911) clearly alerted Fisher (1942) who stated: "the peculiarity of agricultural field experiments lies in the fact, verified in all careful uniformity trials, that the area of ground chosen for the experimental plots may be assumed to be markedly heterogeneous, in that its fertility varies in a systematic and often a complicated manner from point to point."

Homogeneity within blocks is assumed to be secured by correlation between adjacent plots (frequently in relation with plot size) as it has been often stressed since Mercer and Hall (1911). However, if no major objections can be made when it relates to duplicate experiments of small plots, the same is not so straightforward when several treatments are to be tested, even if particular care has been devoted to perform blocking.

On the other hand, "the practical universallity of yield heterogeneity" (Harris 1920) was never refuted, but "the systematic variation, point to point or having centers from which they spread outwards" can be objected to as, for instance, by Fairfield Smith (1938). Only in 5 of the 39 blank experiments revised by Fairfield Smith was there an indication that the soil was more variable in one direction than in another. In his own experiment, "the soil heterogeneity of the area was found to be patchy and equally variable in all directions", though "the distribution of yield from unit plots was symmetrical and differed from normal distribution only very slightly in direction of a flat-topped curve".

Finally, but not less important, is that it is "impossible to forecast accurately the relative efficiencies of different arrangements for a field of whose heterogeneity little or nothing is known" (Smith 1938).

The analysis of these two points suggests that hardly any advantage can be brought out by blocking in cultivar yield trials. On the other hand, the third point at issue, related to the interaction between effects to be tested, is the most sensitive, as it concerns the validity of both tests and inferences.

Genotype environment interactions were recognized long ago so, to take into account such interactions in the analysis of groups of experiments, special techniques have been developed.

The complexity of environmental factors and of their interactions call the experimenter's attention to more sizeable targets. As put by Allard and Bradshaw (1964) "to determine what may be profitable in terms of present day technology we therefore have to turn to analysis of genotype-environment interactions as we actually find them".

It was noticed earlier that "no correlation can be drawn regarding the suitability of the varieties for particular soil types" (Gregg and Sharrock 1955), and though some authors, as for instance Sandison and Bartlett (1958), stated that "it is possible to examine differences between varieties related to soil type, fertility level . . .", no evidence has been given as to the influence of soil type as different from its reflection in the fertility level. We also believe that small climatic variations, which do not strongly interfere with the hormonal processes of yield (Gusmão and Cidraes 1979), can be considered in that same perspective.

Interaction of "genotype-soil fertility level" within trials has been ignored, and not accounted for, in the design of single experiments. However, even before CRB, Fisher and Mackenzie (1923) recognized that "if important differences exist in the manurial response of varieties a great complication is introduced into both variety and manurial tests", thus concluding, "only if such differences are non-existent or quite unimportant can variety tests conducted with a single manurial treatment give conclusive evidence as to the relative value of different varieties".

But response to manure implies response to fertility, and the yield, according to this last feature, is, as noticed by Bradshaw (1974), of an inevitable nature.

Phenotypic changes produced by environment include variations in yield with a general pattern specific for each genotype and appear linear when assessed through the JRA (at least, when sub-optimal and super-optimal environments do not coexist, as stressed by Knight (1970)).

Though we may say, as Hardwick and Wood (1972), that "There is variation between genotypes in more than one dimension", we may also say that, as far as fairly uniform climatic conditions are to be found, a different response to soil fertility is, in most of the cases, the most important and, in practical terms, the only one allowed for consideration (when an experiment in a single place is concerned).

In fact, references abound in the literature, which show the high frequency with which the genotypes exhibit different patterns of response to the environmental index (as assessed by the general soil fertility). We must, then, generalize to "fertility levels" the expression "manurial treatment" in the following statement of Fisher and Mackenzie (1923): "The yields of different varieties under different manurial treatment are better fitted by a product formula than by a sum formula".

The main purpose in producing new cultivars is to attain yield superiority in all situations which would allow one to ignore the interaction effect, on its statistical perspective in the two-way ANOVA. The most probable, in this type of trials, is type *2* interaction, as described by Haldane (1946). Type *1a* interaction is quite improbable, nevertheless this type together with type *1b* and *2* are the only ones to be expected whenever fertility level in its interaction with genotype overcomes all other possible components of the environment.

In fact, in this case a CRB for a cultivar yield trial can be more adequately described in terms of a factorial design without replications. This does not allow for any degree of freedom for error, or, for what turns out to be the same, the degrees of freedom other than those for the main effects can not be ascribed for error but rather, in a larger or smaller extent, as interaction "cultivar-block".

According to Dagnelie (1980), CRB for yield trials can be expressed in a mixed theoretical model with blocks being random. This allows the main factor (fixed) to be contrasted against interaction, as far as its variances are considered to be the same.

However, blocks represent different variations within a certain general environment to be tested, and so, in order to be considered as random they must not be allotted in the systematic way they generally are (according to presumed fertility gradient, or simply in a preconceived, more or less, compact display).

Even if blocks had been randomly assigned within a possible block population (which would originate a true mixed theoretical model), the lack of homogeneity within blocks, if not sufficiently compensated by heterogeneity between blocks, would produce a residual error which would mask the interaction effect.

As Fisher (1942) stressed, "For our test of significance to be valid the differences in fertility between plots chosen as parallels must be truly representative of the differences between plots with different treatments".

Finally, we would like to point out that, in the search for interactions too much attention has been payed to sites and years, but, as written by Hanson (1964), "Years are consecutive elements in a time series which could be cyclic, and sites are chosen for convenience or to represent a range of different types of environments rather than as random choices". Thus, groups of experiments must be assumed in a fixed effect model, for which, there is general agreement that "the main effects should always be tested against the residual variation" (Freeman 1973).

Lack of homoscedasticity

The major component of the error, in yield trials, is ascribable to heterogeneity of soil fertility. Yates and Cochran (1939) refer to "a somewhat higher residual variation of 'Trebi' from its own regression than of the other varieties from their regression". So, they probably suspected that some correlation existed between regression coefficient and individual residual variance. It is the existence of such a correlation that led Freeman (1973) to say: "It is better to test the significance of the β_i for a particular genotype by comparing the appropriate sum of squares against the deviation for regression for that genotype rather than against the pooled deviation term."

The frequent association between high $S_{\hat{d}}$ and high yield, referred by Joppa et al. (1971) suggests that resistance mechanisms may frequently have a detrimental effect on yield ability. On the other hand, specific genotype-environment interaction, in the way described by those same authors, is not necessarily correlated with a reduction of regression coefficient but, most probably, with a reduction of the coordinate at the origin. As a matter of fact, if the same specific interaction occurs in all trials, a proportional reduction at all levels of the regression line is to be expected thus, the regression coefficient is not affected. Specific interaction can then be interpreted in terms of a reduction of the mean fertility of the environment.

In this way, we have to admit a lack of homoscedasticity whenever yield trials are concerned. This situation, though legitimate for testing the null hypothesis, will invalidate even the Completely Randomized Design (CR) in what concerns the test of differences after discredit of the null hypothesis unless the overall variation is restrained within narrow limits.

Numerical example

When for one trial, fertility level is the only relevant environmental component for which interaction with genotypes is to be considered, strong correlation between cultivars is to be expected whenever regression coefficients of genotypes are similar.

In such a situation, the non-existence of block-to-block correlation within trials can only mean the lack of homogeneity within blocks and of heterogeneity between blocks.

To illustrate the aforesaid we, use data from Cidraes and Gusmão (1980), namely for the cohort of ten triticale genotypes considered in Gusmão (1985), having regression coefficients of 0.928 to 1.05.

Figure 1 shows box-and-whisker plots of the correlation coeficient comparisons concerning each of the 25 CRB five replicate trials.

The median of within trials correlation varies from a maximum of 0.534 to a minimum of −0.131, when the overall correlation for the 25 trials are distributed in the following five-number summary.

45			
M23	0.814		d_F
F 11.5	0.781	0.835	0.054
1	0.718	0.867	

These results discredit the assumptions in any trial, in spite of the general upward tendency revealed in Fig. 1. This can only testify the care put in setting the trials in the field, otherwise no such tendency would have been shown.

Lack of homoscedasticity is not so easy to show. Causes for this are, namely: the greater proportion of errors in rela-

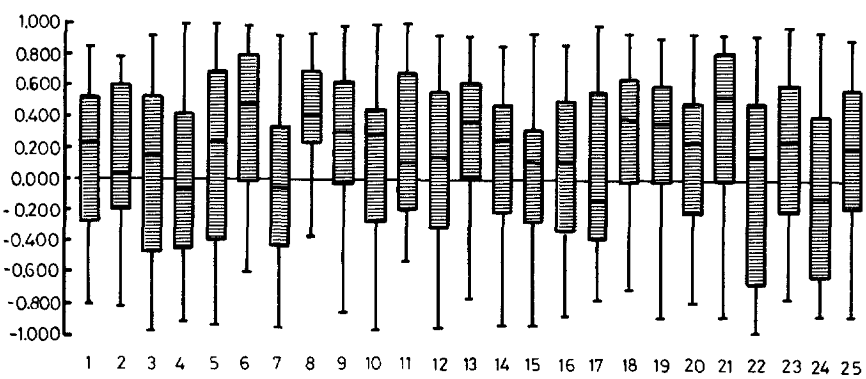

Fig. 1. Box-and-whisker plots for the 45 correlation coefficients between genotype pairs in each trial

Table 1. Fourth-spread values of plot yield in each trial (kg/10 m²)

Triticale										Rye
72IT7	72IT13	74IT8	75IT6	75IT7	75IT8	75IT12	75IT13	75IT14	75IT16	'Mirandela'
0.270	0.170	0.090	0.300	0.310	0.090	0.170	0.430	0.370	0.350	0.340
0.060	0.130	0.050	0.490	0.450	0.250	0.500	0.340	0.220	0.630	0.150
0.080	0.190	0.650	0.260	0.140	0.220	0.660	0.120	0.800	0.140	0.140
0.370	0.150	0.220	0.070	0.010	0.580	0.310	0.170	0.620	0.310	0.210
0.340	0.480	0.180	0.630	0.780	0.190	0.480	0.360	1.250	0.400	0.340
0.520	0.680	0.910	0.300	0.100	0.840	0.320	0.470	0.720	0.560	0.560
0.800	0.360	0.640	0.500	0.450	0.970	0.210	0.090	0.170	0.250	0.220
0.050	0.230	0.310	0.500	0.290	0.220	0.120	0.140	0.290	0.260	0.100
0.300	0.100	0.870	0.850	0.810	0.180	0.500	0.680	0.520	0.380	0.250
0.840	0.130	0.830	1.150	0.770	0.750	0.920	0.840	0.540	0.310	0.270
0.320	0.640	0.040	0.490	0.220	0.380	0.140	0.220	0.550	0.550	0.530
0.380	0.440	0.090	0.690	0.640	0.100	0.030	0.190	0.080	0.200	0.230
0.300	0.120	0.200	0.180	0.100	0.310	0.050	0.190	0.110	0.200	0.130
0.470	0.270	0.360	0.600	0.490	0.250	0.630	0.620	0.430	0.280	0.160
0.260	0.170	0.570	0.110	0.410	0.420	0.450	0.250	0.410	0.120	0.200
0.170	0.490	0.270	0.210	0.430	0.870	0.280	0.670	0.560	0.790	0.040
0.190	0.250	0.900	0.060	0.310	0.160	0.030	0.270	0.200	0.690	0.210
0.220	0.250	0.150	0.210	0.240	0.120	0.010	0.170	0.180	0.300	0.350
0.320	0.040	0.170	0.390	0.150	0.160	0.260	0.050	0.100	0.180	0.110
0.350	0.990	0.230	0.860	0.570	0.690	0.210	0.650	0.350	1.110	0.540
0.620	0.170	0.530	0.350	0.280	0.510	0.020	0.250	0.310	0.340	0.480
0.140	0.070	0.110	0.050	0.240	0.180	0.520	0.220	0.120	0.470	0.080
0.150	0.200	0.240	0.280	0.120	0.280	0.030	0.100	0.020	0.180	0.200
0.390	0.250	1.160	0.750	0.420	0.790	0.150	1.020	0.590	0.640	0.190
0.420	0.390	0.370	0.250	0.460	0.550	0.130	0.230	1.310	0.320	0.220

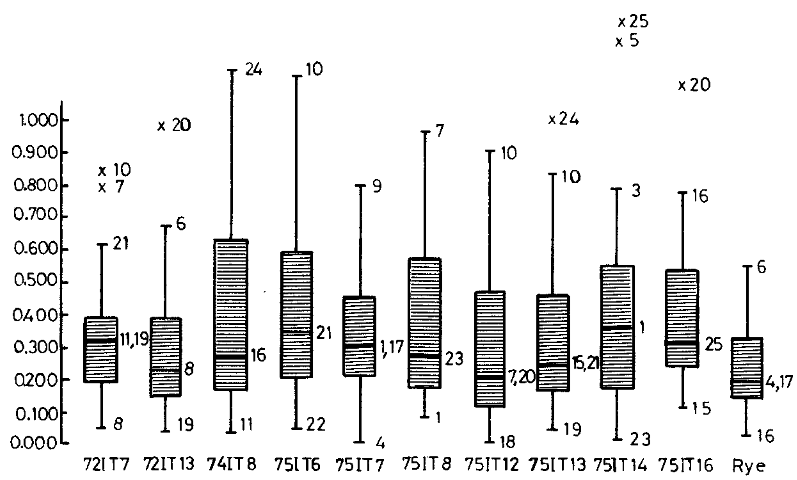

Fig. 2. Box-and-whisker plots of fourth-spread values of plot yield in each of the 25 trials (expressed in kg/10 m²)

tion to differences due to different regression coefficients, the relatively narrow limits of soil fertility levels, the paucity of degrees of freedom and the lack of complete randomization.

For the analysis of individual variation, and in order to include a noticeable different regression coefficient, the results on the rye cultivar 'Mirandela' (used as the standard in the trials) were aggregated to the so-far considered cohort. Individual variation in each trial was evaluated through the fourth-spread as presented in Table 1 and Fig. 2.

Extreme variability of spread within the group of identical regression behaviour (227 kg/ha to 377 kg/ha), as measured by the broadened median (corresponding to a 40% trimmed mean) does not allow for testing heteroscedasticity. However, the estimation of that same parameter for the rye cultivar produces a smaller value (213 kg/ha) corresponding to the smaller regression coefficient.

Discussion and conclusions

The analysis of the assumptions for CRB in cultivar yield trials, through a numerical example on triticale cultivars, supports the theoretical considerations which degrade this design for the present purpose.

After discredit of the null hypothesis, the test of differences against pooled SSE, even in completely randomized design, may induce errors whenever very different regression coefficients and great variation of fertility levels are concerned. In this situation, type I errors are likely to occur for comparison of high coefficient regression cultivars, and type II errors in the opposite case. Cultivar yield trials should, consequently, be planned according to the following *consideranda*. Using the words of Horner and Frey (1957), "small grain varieties are most often recommended on a regional or large area basis", which implies their lack of specialization for restricted environments. Very often seasonal climatic differences within the same place surmount local differences within the same year and normally the selection process and/or preliminary evaluation lasts much longer than the ultimate yield tests can endure. Large specific instability, related to climatic features is then strongly reduced by the screening process of cultivar production. It is not surprising that the main interaction component for each cultivar yield trial can be ascribed to the local yield potential.

JRA has been shown to give excellent descriptions of relative cultivar behaviour in such situations where ecological amplitude involves stress mechanisms of identical nature. As pointed out by Hill (1975), referring to JRA, "despite its imperfection, it does have the twin merit of simplicity and biological relevance". Individual trial responses have, therefore, the only merit to contribute to the description of a pattern of behaviour, through which a more correct evaluation of errors can be achieved.

On the other hand, specific instability has been described in terms of residuals in the JRA; but residuals comprise a fraction of true instability (as the result of a set of unknown independent variables) and a fraction arising from heterogeneity within each trial (which can be thought as a function of the regression). Thus, when in search for the relative pattern of behaviour, particular care should be devoted to reduction of overall variability within trials. This can be attained through a reduction of the area, namely by a lower number of repetitions, in spite of its negative reflection on the causal variable precision. As concluded by Rasmusson and Glass (1976) "the need for sampling of several environments (years and locations) in a testing program should not be taken for granted, but should be carefully evaluated". However, this should not be taken exclusively under the theoretical point of view of El-Enein et al. (1977); "to evaluate the performance of new lines, test locations should probably be divided to include both high and low-yielding sites". As a matter of fact, this only comtemplates the precise evaluation of the regression parameters estimators, and does not conciliate this objective with the need for searching of eventual specific interactions.

As comparisons are made through a well defined pattern, randomization of sites is no longer needed, and environmental yield potential limits can be induced through technical means, as far as they do not interfere with main climatic features. However, as pointed out by Eberhart and Russell (1966) "the varieties must be grown in an adequate number of environments covering the full range of possible environmental conditions".

Finally, and in spite of the remark of Immer et al. (1934), that relative performance of different cultivars is not the same in different years, there is, most frequently, no advantage to be expected in carrying on the trials beyond one year, unless very extreme climatic conditions occur.

Therefore, we partly support the conclusions of Teich (1983) in that "Increasing the number of sites and reducing the number of test years is a possible means of increasing the rate of gain", but, in opposition to him, we think that, in most of the circumstances, this is the recommended procedure. Taking into account the number of degrees of freedom, a minimum of three repetitions in each CR trial will also improve the chances of explaining occasional specific interactions.

Acknowledgements. I am very grateful to Professor M. Mota, Department of Genetics, Estação Agronómica Nacional, for all his encouragement and helpful criticism, and 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 Professor Cândido Pinto Ricardo, Department of Physiology, Estação Agronómica Nacional, for his kind help in the revision of the manuscript.

References

Allard RW, Bradshaw AD (1964) Implications of genotype-environmental interactions in applied plant breeding. Crop Sci 4:503–508

Bradshaw AD (1974) Environment and phenotypic plasticity. Brookhaven Symp Biol 25:75–94

Cidraes AG, Gusmão L (1982) Adaptação de triticales em Portugal. 1. Comportamento das linhas e variedades em ensaio no triénio 1976/1977 a 1978/1979, Melhoramento 27:161–201

Dagnelie P (1973) Théorie et methods statistique, vol 2. Press Agronomiques, Gembloux

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

El-Enein RA Abo, Morsi LR, Gomaa AA (1977) Yield stability parameters for barley cultivars in Egypt. Agric Res Rev 55: 189–206

Farfield Smith H (1938) An empirical law describing heterogeneity in the yields of agricultural crops. J Agric Sci 28: 1–23

Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant-breeding programme. Aust J Agric Res 14: 724–754

Fisher RA (1925) Statistical methods for research workers, 8th edn. Oliver and Boyd, Edinburgh London

Fisher RA, Mackenzie WA (1923) Studies in crop variation. 2. The manurial response of different potato varieties. J Agric Sci 13:311–320

Fisher RA, Wishart J (1930) The arrangement of field experiments and the statistical reduction of the results. Imp Bureau Soil Sci Tech Comm 10

Freeman GH (1973) Statistical methods for the analysis of genotype-environment interactions. Heredity 31:339–354
Goodchild NA, Boyd WJR (1975) Regional and temporal variations in wheat yield in Western Australia and their implications in plant breeding. Aust J Agric Res 26: 209–217
Gregg AS, Sharrock JN (1955) NAAS/NIAB trials of spring wheat varieties 1949–1951. J Nat Inst Agric Bot 7:277–288
Gusmão L (1985) The adequate design for regression analysis of yield trials. Theor Appl Genet 71:314–319
Gusmão L, Cidraes AG (1982) Adaptação de triticales em Portugal. 2. Análise do comportamento quantitativo em relação às espécies parentais. Melhoramento 27:203–223
Haldane JBS (1946) The interaction of nature and nurture. Ann Eugen 13:197–205
Hanson WD (1964) Genotype-environment interaction concepts for field experimentation. Biometrics 20:540–552
Hardwick RC, Wood JT (1972) Regression methods for studying genotype-environment interactions. Heredity 28: 209–222
Harris JA (1920) Practical universality of field heterogeneity as a factor influencing plot yields. J Agric Res 19:279–314
Haydock KP, Sandland RL (1975) Introduction: Design and analysis in field experiments. In: Bofinger VJ, Wheeler JL (eds) Developments in field experiment design and analysis. Commonwealth Agricultural Bureaux, England, pp 1–10
Hill J (1975) Genotype-environment interactions – a challenge for plant breeding. J Agric Sci 85:477–493
Hoaglin DC, Mosteller F, Tuckey JW (1983) Introduction to more refined estimators. In: Hoaglin DC, Mosteller F, Tuckey JW (eds) Understanding robust and exploratory data analysis. Wiley and Sons, New York Chichester Brisbone Toronto Singapore, pp 283–296
Horner TW, Frey KJ (1957) Methods for determining natural areas for oat varietal recommendations. Agron J 49: 313–315
Immer FR, Hayes HK, Powers L (1934) Statistical determination of barley varietal adaptation. J Am Soc Agron 26: 403–419
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
Kempthorn O (1952) The design and analysis of experiments. Wiley and Sons, New York; Chapman and Hall, London
Knight R (1970) The measurement and interpretation of genotype-environment interactions. Euphytica 19:225–235
Mercer WB, Hall AD (1911) The experimental error of field trials. J Agric Sci 4:107–132
Rasmusson DC, Glass RL (1976) Estimates of genetic and environmental variability in barley. Crop Sci 7:185–188
Rose CW (1975) The role of modelling and field experiments in understanding complex systems. In: Bofinger VJ, Wheeler JL (eds) Developments in field experiment design and analysis. Commonwealth Agricultural Bureaux, England, pp 129–154
Sandison A, Bartlett BO (1958) Comparison of varieties for yield. J Nat Inst Agric Bot 8:351–357
Student (1908) The probable error of a mean. Biometrika 6:1–25
Student (1923) On testing varieties of cereals. Biometrika 15: 271–293
Tuckey JW (1977) Exploratory data analysis. Addison-Wesley, Reading Menlo Park London Amsterdam Don Mills Sydney
Teich AH (1983) Genotype-environment interaction variances in yield trials of winter wheat. Cereal Res Commun 11: 15–20
Wood TB, Stratton JH (1910) The interpretation of experimental results. J Agric Sci 3:417–440
Yates F, Cochran WG (1938) The analysis of groups of experiments. J Agric Sci 28:556–580