CARIBBEAN FOOD CROPS SOCIETY

(C F C S)

XIV th Meeting

Quatorzième Congrès

de la

SOCIETE INTERCARAIBE POUR LES PLANTES ALIMENTAIRES

Guadeloupe — *27 - 29 Juin*

Martinique — *30 Juin - 2 Juillet 1977*

Sponsored by

Organise par

L'INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (I.N.R.A.)

with the aids of

Avec les aides

de la

DELEGATION GENERALE A LA RECHERCHE SCIENTIFIQUE ET TECHNIQUE

(D.G.R.S.T.)

and of the

et des

CONSEILS GENERAUX

CHAMBRES D'AGRICULTURE

DE LA GUADELOUPE ET DE LA MARTINIQUE

with the technical assistance of the following organisms

avec le concours technique des organisations suivantes

ORSTOM - IRFA - IRAT - CTGREF - DDA -

And the participation of Institutions of 15 Caribbean territories

Et la participation des Institutions de 15 pays de la Caraîbe

SOUS le PATRONNAGE de MM. LES PREFETS de la GUADELOUPE et de la MARTINIQUE

Hôtel Arawak
Gosier - Guadeloupe

Hôtel Méridien
Trois Ilets - Martinique

Nouv. agron. Antilles-Guyane, 3, 3/4, 513-522, 1977.

BEYOND THE RANDOMIZED COMPLETE BLOCK : NEW STATISTICAL CONCEPTS FOR THE DESIGN AND ANALYSIS OF FIELD EXPERIMENTS

A.C. BREWER(°)

INTRODUCTION

By far the most common statistical technique employed in the conduct of field experiments in the region is that associated with the randomized complete block design or occasionally modifications of this which retain balanced block and treatment structures which are comparatively simple to statistically analyse and interpret. On occasions these designs and methods of analysis prove to be entierly satisfactory and produce clear and precise results for the experimenter. However, all too frequently the results of these methods are in varying degress found to be disappointing, giving little or no resultant information on the subject under investigation. For example, Brewer (2) in a survey of coefficients of variation in a series of maize experiments, found the median value of this sample to be 21.48 % which can hardly be described as satisfactory. In order to introduce the new concepts it will be instructive to consider some of the causes of such poor results.

One such cause is that the experiment is either poorly managed or that one or more unfortunate occurrences have happened during the conduct of the trial such as waterlogging, praedial larceny etc. Although astute experimenters or biometricians can occasionally retrieve something from the disaster, no statistical technique can of course insure against such a happening. Another is that the intrinsic variability of the material used in the study (land or plants) is higher than expected, giving less precise estimates than one would have hoped for. This means that a larger trial than the one actually laid down would have been necessary and thus it is always beneficial to know the level of variability one is dealing with when an experiment is being planned. It will be seen that the application of the new methods can sometimes be of assistance here.

A third cause is where the blocking used has been ineffective in its purpose of accounting for a substantial amount of the variability in the field. Symptomatic of this is the presence of a small mean square for block effects in the analysis of variance of the data. Now some experimenters are more skilful in their allocation of blocks to the experimental areas than others, but I expect all at some time or another have had the experience of the pattern of variation in the field not following the pattern expected when the blocks were allocated. This brings us to the question of whether there are other methods of explaining and controlling variability in the field which are not so dependent upon the configuration of plots in the experiment decided upon beforehand. The methods to be described here are being developed from what is considered to be one of the fundamental properties of field variation.

Models of Spatial Variation

The customary method of forming blocks in a field experiment is to place them on integrated areas of land, possibly elongated in one direction of greateer variability. This is in recognition of the generally accepted principle that areas of land in close prowimity to each other are more similar in nature than those which are further apart. Rather than split the area into discreet units which are assumed within themselves to be homogeneous, models of spatial variation attempt to apply this general property over the whole area.

(°) Seconded on technical assistance by the United Kingdom Ministry of Overseas Development.

There are in general three rules upon which such models can be thought to apply. The lowest and most fundamental is that between points in a field. Conceptually each point in an experimental plot has a response potential and the response of the plot, apart from the treatment effect, is made up of the sum of the responses of each of its individual points. Response potentials of points closer together are more similar than those farther away and neghbouring plots will tend to be similar in their response because their respective points are close to each other. Matern (1970) has described some models of this sort and they will generally be referred to as Matern's models.

The intermediate scale is that between individual plants, an example of a model of this nature being that considered by Mead (1967). However, models of this nature are normally used for a different purpose, that of investigating competition effects, resulting in adjacent plants being less similar in their response then in general.

The largest scale is that between plots of the experiment in which the responses of individual plots are regressed upon their neighbours. Besag (1) lays a comprehensive theoretical foundation for one class of such models, whilst other are available. Whilst it is recognised that such models can only be an approximation in the field experiment situation because the response of a plot is the aggregation of such models, whilst other types are available. Whilst it is recognised that such models can only be an approximation in the field experiment situation because the response of a plot is the aggregation of the response over an integrated area of land, they do however constitute the easiest tools to use on an empirical set of data and can in some situations form a fair approximation to Matern's models.

It will be seen in the following three sections how consideration and investigation of these models can have an impact on the conduct of field experiments through the stages of planning, design and analysis. They may be particularly valuable in the Caribbean region where experimental land is often less homogeneous than in other areas.

Planning of Experiments

There are a number of questions to be decided upon during the planning stages of a field trial butthis paper is concerned with the particular problem of the field layout and the statistical design. Questions on statistical design will be covered in the following section but once a design has been proposed, within the general framework of the design the number, size, shape and arrangement of the plots has to be determined. It is then necessary to draw up a detailed field plan incorporating the properties decided upon in this determination.

In its present state, this process is almost entirely subjective in nature, the efficiency of the various decisions being dependent entirely on the experimenter's experience and expertise in "reading" the likely performance of the experimental land and material that he is dealing with. In practice he will base his judgement on what has happened before on the same or similar land and on the past performance of the crop. It would be a great help to him if more precise information of this sort were available and it is in this respect that an empirical investigation of spatial variability in field trials can be of great use.

As an example let us suppose that the experimenter has 4 treatments which he is to arrange in blocks of 4 in a randomized complete block design. From fitting spacial variation models to past experiments on similar land, and by considering the crop now under study, a particular model having a component of variability due to the environment, genetic variability of the planting material, and other random variability can be assumed to hold in the present case. Figure 1 shows some possible configurations of plots within blocks and table 1 shows the expected coefficients of variation which would result from these and other configurations from a particular model of this type. The actual units of measurement are not important for the purposes of the example.

The model used in the example has no direction where variability is greater than in others, but it is evident from table 1 that even in this entirely symmetric case the size, shape and configuration of plots within blocks all have an effect on the resulting precision to be expected. Cases 4 and 6 produce the lowest C of V.'s which supports the proposition that on well balanced land, the best arrangement for block experiments is to have elongated plots laid side by side to form compact blocks Also it can be seen that larger size plots tend to be associated with lower C. of V.'S although, by comparing cases 2 and 3, we can ascertain that this does not always hold Case 3 has the same arrangement as case 2 but the lengths have been doubled (and consequently the areas quadrupled) yet it results in a higher C. of V. This effect has resulted from the fact that at this particular point of the scale the greater homogeneity within 2x2 blocks compared to 4x4 blocks has outweighed the greater accuracy of the larger plots in estimating the response of the crop. Although it is certainly not assumed that this particular model applies to any known situation, it should nevertheless serve as a warning against the impression that larger plots always imply greater accuracy.

Also included in Table 1 are estimates of how many replicates are needed to detect treatment differences of 10%, 15% and 20% of the overall mean response. Thus models of spatial variability can also assist in determining the size of the experiment necessary to achieve the objectives, which has hitherto been a very difficult question to answer. Alternatively one might be able to say whether it is worthwhile conducting the trial with the present resources.

The example illustrates the usefulness of the approach in the case of a simple design under well behaved conditions. When the situation is more complex due to the presence of directional variation when slopes are encountered, for instance. systemmatic variation due for instance to bedding, or irregular variation due to observable factors such as drainage or soil nutrients the task facing the experimenter of deciding upon the optimum field layout can sometimes be almost insuperable. However, with sufficient care and prior information, spatial models of greater complexity which incorporate the expected effects can be formed and investigated in like manner.

The Experimental Design

For reasons of ease of computation and interpretation the use of balanced statistical designs such as randomized complete blocks are the standard techniques used. However, such designs are not without disadvantages in respect of the constraints which they necessarily impose upon the field plan. All blocks contain the same number of plots and therefore are of the same size. It is left to the experimenter to find homogeneous areas of land of equal size on which to place these blocks and quite frequently a considerable amount of compromise is necessary. This results in blocks being formed which are not as homogeneous as one would hope and consequently has an undesirable effect on the precision of the experiment. Professor Pearce, in a series of seminars during a visit to the region, propounded views that with the ever increasing presence of computers which can easily perform the more complex calculations involved with unbalanced designs given a suitable computer program, the advantages of balalanced designs have been diminished. Thus it is possible for the experimenter to choose his blocks, not neceesarily all of the same size, more according to his expectation of how the land will behave. When using unbalanced block designs, techniques for the optimum allocation of treatments to blocks (Jones (3) and for the analysis of data (Pearce et al. (7) can be used.

Although the statistical efficiency of unbalanced designs are not as high for balanced designs in general, a skilful use of them in some situations can undoubtedly dramatically increase the resulting precision of the results. It is evident that

.../...

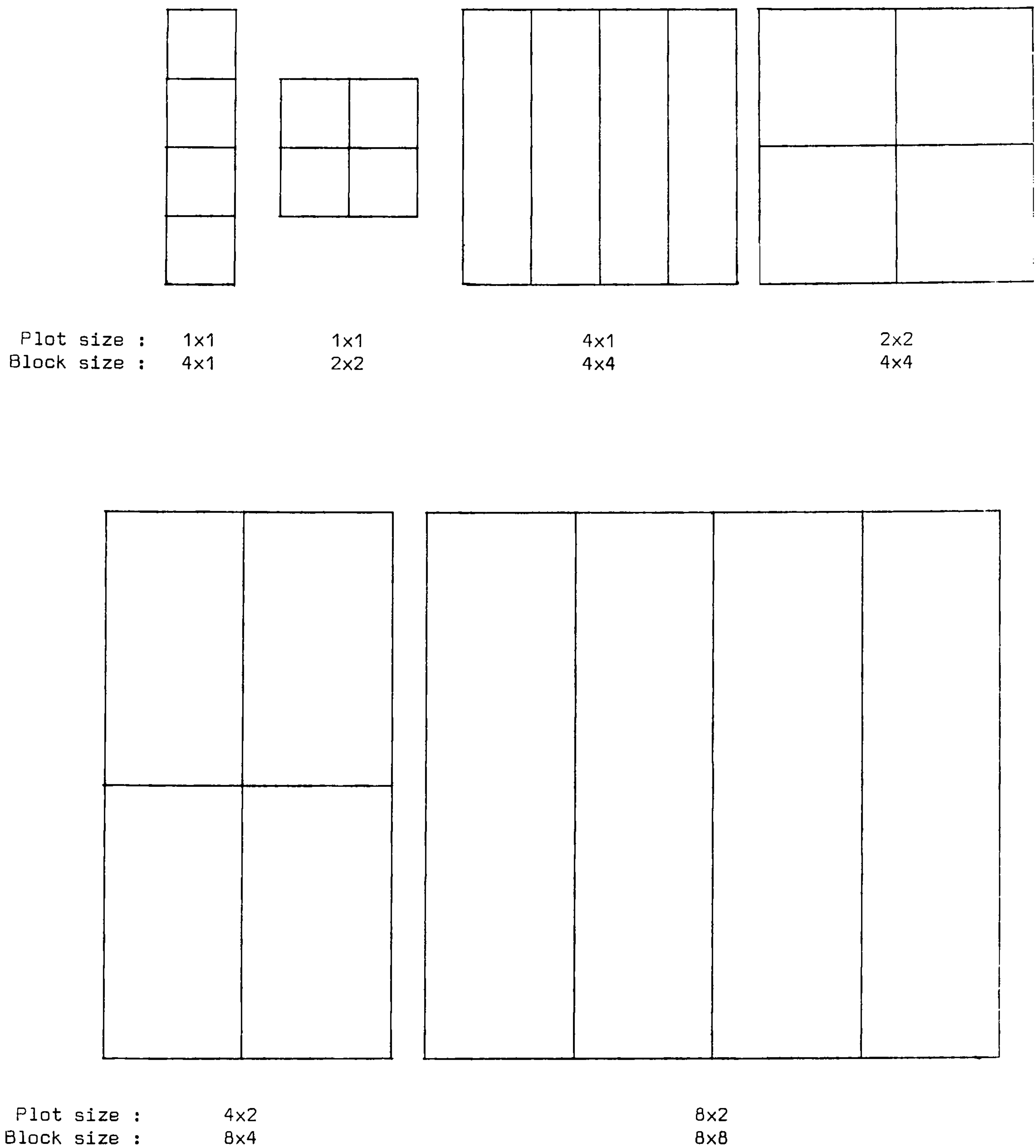

Fig. 1 - Some possible shapes and arrangements of plots within blocks of 4 plots.

Table 1 - Expected coefficients of variation and sizes of experiment necessary for various plot and block configurations under a particular spatial variability model.

Case number	Plot size	Block size	C. of V.	No. of reps. required to detect differences of:		
				10 %	15 %	20 %
1	1x1	4x1	18.7 %	28	13	7
2	1x1	2x2	15.6 %	20	9	5
3	2x2	4x4	17.2 %	24	11	6
4	4x1	4x4	13.8 %	16	7	4
5	4x2	4x8	16.0 %	21	10	6
6	8x2	8x8	12.4 %	13	6	3

when this promising fresh approach is used, the greatest advantages will accrue when the experimenter has a clear picture of the expected performance of the land he is dealing with. Since models of spatial variation attempt to characterize the performance of a field, they can be used to great effect in this situation. With the investigation of a representative model, the experimenter will be able to see for himself the effects of blocks of different sizes, shapes and arrangements on the precision of the experiment and will thus be in a better position to judge the relative merits of different formations. Indeed, this can be seen to be similar to and interactive with the application of the models in the last section, and thus can be viewed together as helping to form a useful field plan - design strategy for field experiments.

An alternative approach to experimental design is to at least partially disregard blocking techniques and work more directly on the assumption that neighbouring plots are more similar to each other than those farther away. We shall see in the next section how alternative methods of analysis of the data have developed using this property, whilst here some designs which "balance" the treatments with the neighbouring plot effect will be briefly described.

One possibility is the familiar method of the chequer-board pattern of planting a standard treatment as neighbours to all experimental treatments (Fig. 2). All contrasts of experimental treatments with the standard treatment are balanced with respect to the neighbouring plot effect. When the object of the experiment is to investigate these contrasts in particular, rather than contrasts between experimental treatments, this can be a useful design. One has, however, necessarily to make the assumption that competition effects between the standard and other treatments are negligible.

We could alternatively consider designs such as the 4x4 and 5x5 latin squares shown in Fig. 3. It can be seen that each treatment neighbours every other treatment exactly once in each direction in the case of the 4x4 square, and twice both lattidudinally and longitudinally in the case of the 5x5 square. Thus every contrast of two treatments is balanced with respect to the neighbouring plot effect.

Although such designs hold well-balanced properties when applying neighbouring plot techniques to them, to achieve this their possible arrangements in the field are even more restrictive than for balanced block designs. A further severe restriction on the latin squares is that only a small subset of the possible randomizations retain the properties described above. To justify their use, more work will need to be done on an investigation of their relative advantages over other designs. It is possible that less restrictive partially balanced designs (c.f. partially balanced incomplete blocks) can be generated.

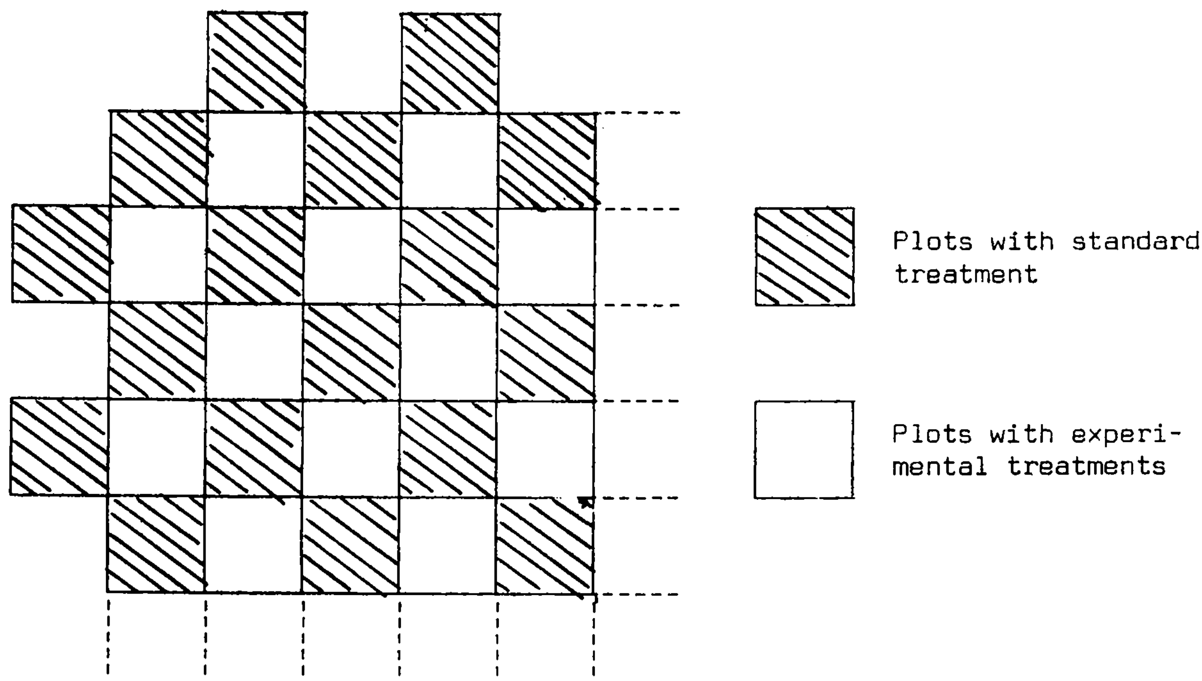

Fig. 2 - The check-board design

A	B	C	D
C	A	D	B
B	D	A	C
D	C	B	A

A	B	C	D	E
C	A	E	B	D
B	D	A	E	C
E	C	D	A	B
D	E	B	C	A

Fig. 3 - Latin squares balanced for neighbouring plot effects.

The Analysis of Experiments

In a recent paper Pearce (8) resurrected a method of data analysis due originally to Papadakis (6) of adjusting the performance of a plot with that of its neighbours. He reported that in a varied sample of experiments with perennial crops the methods were in many cases more successful in removing variation than the conventional block analysis. The method is simply to treat the experiment as a completely randomized design and to remove treatment effects from the plot responses. The residual for each plot is then adjusted for the performance of its neighbours by means of a regression analysis so that the resulting residuals can be viewed as a deficit or a surplus in the response of the plot compared with the average response of those around it. The error sum of squares is then calculated as the sum of squares of these new residuals.

When the fundamental property of neighbouring plots being closely related in terms of response is in evidence, it is not surprising that this method gives good results, and the results obtained from it so far show that it does warrant greater attention in the future. What it lacks is as sound a theoretical base as there is for the more usual forms of analysis of variance, and only through a study of models of spatial variation between plots such as those described by Besag (1) can the formalized and possibly modified. When such schemes as this are accepted as useful and well based, then designs such as those shown in figs. 2 and 3 will be in greater demand.

Block experiments and analyses will still retain certain advantages, one of which is that the blocks tehemselves can be treated as administrative units, making it unnecessary for example, that the entire experiment should be planted or harvested on the same day. In the analysis of data there is no reason why the advantages of both methods cannot be retained and the experiment analysed using a hybrid scheme allowing for both block and neighbouring plot effects.

Conclusion

In the widespread use of balanced block designs, experimenters are often constraining themselves in statistical straight-jackets resulting in a loss of accuracy in the results. By considering more closely the nature of field variation and alternative methods of design and analysis to deal with it, it is possible to greatly improve the methods now used. What is needed is for experimenters to realize that the practical application of statistical method is as in need of further research as the subject they are investigating in their field trials, and consequently will be prepared to venture a little further into the realms of unconventional statistical approaches - with the aid of a biometrician.

REFERENCES

BESAG (J.). - Spatial interaction and the statistical analysis analysis of lattice systems. J.R. Statis. Soc. B, 36, 192-236, 1974.

BREWER (A.C.). - A pilot survey on coefficients of variations in maize experiments and its use in predicting the efficiency of simple experimental designs. Proc. 12th Ann. Meeting Carib. Food Crops Soc. (In press).

JONES (B.). - An algorithm for deriving optimal block designs. Technometrics, 18, 451-458, 1976.

MATERN (B.). - Performance of various designs of field experiments when applied in random fields. Proc. 3rd Conf. Advisory Group of Forest Statisticians, 119-128, 1970.

MEAD (R.). - A mathematical model for the estimation of interplant competion. Biometrics, 23, 189-205, 1967.

PAPADAKIS (J.). - Bull. 23, Sci. Inst. d'Amelioration des plantes a Salonique (Greece). 1937.

PEARCE (S.C.), CALINSKI (T.) and MARSHALL (T.F.) de C. - The basic contrasts of an experimental design with special reference to the analysis of data. Biometrician, 61, 449-460, 1974.

PEARCE (S.C.) and MOORE (C.S.). - Reduction of experimental error in perennial crops, using adjustment by neighbouring plots. Expl. Agric. 12, 267-272, 1976.

-!-!-!-!-!-!-!-!-!-!-!-!-

SUMMARY

Field experiments in the region often give results of disappointingly low precision. The paper ascerts that this is partially due to the widespread and sometimes indiscriminate use of balanced block designs. As an alternative to this models of spatial variation are introduced and their application to the planning, design and analysis of experiments is discussed. It is shown how they can help to determine the size, shape, number and arrangement of plots in the field ; and a new strategy is proposed of using the models in conjunction with unbalanced block designs. An alternative method of analysis involving the adjustment of the response of the plot according to the response of its neighbours is described, and some experimental designs thought to be suitable for this form of analysis are introduced.