A SIMPLE ALTERNATIVE TO THE STANDARD STATISTICAL MODEL FOR THE ANALYSIS OF FIELD EXPERIMENTS WITH LATIN SQUARE DESIGNS

C. Philip Cox

Jeff B. Meeker

Recommended Citation

A SIMPLE ALTERNATIVE TO THE STANDARD STATISTICAL MODEL
FOR THE ANALYSIS OF FIELD EXPERIMENTS
WITH LATIN SQUARE DESIGNS

C. Philip Cox, Professor Emeritus
Iowa State University, Department of Statistics,
312 Snedecor Hall, Ames, IA 50011-1210

Jeff B. Meeker
Bristol-Myers Squibb Co., P. O. Box 5100, Wallingford, CT 06492-7660

ABSTRACT

Latin Square (LS) designs have long been advocated for field crop experiments on the grounds that '... soil fertility and other variations in two directions are controlled.' As counter-evidence, the published standard analyses of eight LS experiments showed that in only two did the sum of squares for both between-rows and between-columns account for appreciable background variability.

Regarding the background concomitant variability as a continuous surface to which treatment effects are additive, it is suggested that a contributory shortcoming of the standard model is that it admits only a restricted class of surfaces because parameters for warp, or row x column interaction, components are excluded.

It is shown that, at the loss of some orthogonality between background and treatment effects, the deficiency can be remedied by fitting a more general polynomial surface. The principle is exemplified using a backward selection multiple regression procedure to analyze LS data in Cochran and Cox (1957). The procedure gave a considerable reduction in the coefficient of variation, from 12.9 to 6.3%, and permitted more sensible inferences than those (null) from the standard analysis.

A note on medieval cultivation practices and experimental design is appended.

1. Introduction

The Latin Square has achieved a time honored eminence as an experimental design with a wide variety of application areas as a result of its balance and the associated statistical efficiency. In a standard exposition, Cochran and Cox (1957), it is stated "The effect of the double grouping is to eliminate from the errors all differences among rows and equally all differences among columns, ..., The experimental material should be arranged so that the differences among rows and columns represent major sources of variation." This could be realistic if the row and column components were each discrete, individually uniform and non-interacting entities giving additive effects as the background for the observation of additive treatment effects. The authors were at least implicitly aware that statistical models should be veridical and, more cautiously, continue: "In field experiments the plots are usually laid out in a square formation so that soil fertility and other variations in two directions are controlled". Relatedly, of the background response in such contexts and in the absence of discontinuities Cox (1958) - remarked that if this is taken as a continuous function of distances along rows and columns, the mathematical (statistical) model could be regarded as representing a response surface passing through one set of parallel curves for the rows and another set of parallel curves for the columns. The curves were those definable in terms of the linear, quadratic ... components of the inter-row and inter-column contrasts.

In an initial discussion of linear model relevance two aspects were also noted in Cox (1958), one relating to the adequacy with which the background response was represented, the second being concerned with the validity of the model-implicit assumptions, additivity in particular, for the behavior of the treatment responses; a third aspect, of course, concerns the assumptions that the residual terms are additive with some specified distribution.

It was further suggested that, in some applications, the usual Latin Square model could be implicitly over-parameterized in the sense that some of the inter-row and inter-column contrasts might not estimate realistic non-zero parameters of the background changes and might more efficiently be incorporated to improve the residual variance estimate. The examinations presented below of some published Latin Square analyses also suggest that only rarely does the design succeed in removing significant background or concomitant variability attributable to both the rows and the columns. Conversely, another possible source of non-veridicality in some cases is that 'warp' or row x column cross-product or interaction components do not appear in the structural part of the conventional model leading - if their effects are not zero - to parametric contamination of the residual variance estimate and hence to possibly biased and inefficient estimations.

When an observed response, or some transformation of it, is described in terms of parametric and aleatory components, the term 'model' is widely used for the former, the parametric structure assumed for the expected value of the response. Many considerations of model adequacy, however, require discrimination between variability sources reasonably identifiable as 'structural' and those to be designated as 'unexplained' or residual variability. Because, as exemplified in multiple regression determinations, the aleatory components are often conditional on the structural, and conversely, it is suggested that the term 'statistical (as distinct from mathematical) model' should be extended to include the expression of the particular distributional attributes being posited for the residual variability. In principle, it should then be possible to assess model adequacy by comparing data generated by a computer programmed by the model with the actual observations under study. Here it will be shown that a multiple regression linear model can have advantages over the regular model for analyses of field-plot experiments. It will be apparent that the procedure is extensible to other areal layouts although only latin square designs are considered in detail.

2. Examples of Latin Square analyses

Some published analyses were examined for evidence on the above indication that the rows and/or columns sets in latin square models for crop experiments may contain redundant parameters. The results are given in Table 1.

The examples in Table 1 are merely those most conveniently available and it is not suggested that they constitute a random sample. Taking them as a haphazard sample, however, they do not support the commonly suggested view that two-way background variability is removed; it can be seen that both row and column mean squares achieved 5% significance in only two, Fisher (1947) and Rayner (1967), of the eight cases. Both these examples appear to use data obtained under the auspices of the Department of Statistics at Rothamsted Experimental Station in which context the remark, Fisher (1946), is interesting -- "Such a double elimination may be especially fruitful if the blocks of strips coincide with some physical feature of the field such as the ploughman's "lands" which often produce a characteristic periodicity in fertility due to variations in depth of soil, drainage and such factors." It could therefore be that, instead of holding for designs laid out arbitrarily or with primary concern for cultivation convenience, the successful removal of Cartesian two-way

background variability requires the exploitation of prior knowledge about the particularities of the experimental plots. Such knowledge might well be expected at Rothamsted where experimental areas have long been closely studied thus facilitating the recognition of associations between the topography of field systems, cultivation practices, environmental influences such as wind–sheltering, and crop yields.

It is accordingly apparent that, as results of both geological history and cultivation practices, field physiography will importantly influence crop production – commonly because of induced changes in the percolation and retention rates of soil moisture. Neither such general factors nor those which can be more crop specific, such as soil density and pH, will necessarily follow a simple Cartesian structure so that the adequacy of the conventional Latin Square model must, even as a first approximation, be suspect in this context. I noted in Cox (1974) that it was surprising that the issue had not been further pursued by Fisher in view of his remark, Fisher (1946) – '...; 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.'

3. Model specification

It will be immediately clear from the foregoing that the specification of anything approaching a fully veridical model will be formidably complicated. One approach, for example, would involve parametrically nonlinear functions admitting plot location effects, treatment effects, their possible interactions, effects due to differential crop maturation rates together with differential, possibly non–Gaussian and non–additive aleatory effects. Less ambitiously here we have tried to preserve the heuristic utility of the conventional model while admitting, towards improved adequacy, the simplest change – the introduction of cross–product terms in the row and column dimensions. The use of polynomial functions, per se, is not new; for example, comments by R. A. Fisher, on research in 1912-25, and A. Hald, on research in 1948, are reported in Cox (1950). Some agricultural – as distinct from experimental design – precedents for this approach have also been described by, as examples, Eden and Maskell (1928), Wiebe (1935) for wheat, Hutchinson and Panse (1935) for cotton and Cox et al.(1958) for a ryegrass–clover ley. The comment by Fisher (loc. cit.)

'The principal difficulty ... was that fitted polynomials in two dimensions might easily absorb so many as 20 or 30 degrees of freedom without removing a corresponding proportion of the residual sum of squares'

was made before the concept of model adequacy was introduced in Cox (1958) and is a reminder that over–parameterization was once regarded as an unavoidable nuisance. It is shown by the example in Section 5 here that the nuisance can be much reduced by the use of stepwise regression methods.

The contour studies on soil moisture and crude protein percentage in Cox et al. (1958) suggest beginning with representation of a differential element of the background effect as

$$\phi(x,z)dxdz = \sum_r \sum_s \alpha_{rs} x^r z^s \qquad r, s = 0,1,\ldots,$$

at the point (x,z) with respect to Cartesian axes parallel to the rows and columns respectively and, in principle, the definite double integral of this could be used for the background contribution over the (i,j)–th plot. Hence, on the assumptions that:

  i) the background changes over each plot are sufficiently symmetrical
with respect to the plot centers,

  ii) there are no discontinuities over the whole experimental area,

  iii) the experimental layout is made up of equally spaced plots of the
same size and shape;

the mean value theorem permits the yield, y_{ij}, of the plot in the ith row
and jth column, to be expressed as

$$y_{ij} = \sum_r \sum_s \beta_{rs} x_i^r z_j^s + \theta_{ij} + \epsilon_{ij}, \quad r, s = 0,1,... \qquad (1)$$

where θ_{ij} is the effect of the treatment on the i,jth plot of which the
coordinates of the center are (x_i, z_j) and ϵ_{ij} is the residual aleatory
element.

4. Analysis

  If, apart from the parametric coefficients, the form of the background
surface is known, a standard linear model program can be used for the
analysis. As an alternative for small squares, a simple multiple analysis
of covariance can be carried out. Orthogonal polynomial coefficients can
be used for the purely row and purely column terms and products of the
coefficients will provide the cross-product position covariates. The
purely inter-row and inter-column sums of squares (SS) can be isolated
first but the SS for the inter-treatment and the background cross-product
quantities will not be orthogonal. Accordingly the ANOVA SS for
treatments is that accounted for by the treatment parameters after the SS
accounted for by all the background parameters have been calculated; it
can conveniently be designated in the ANOVA as 'treatments adjusted' – for
the background parameters – following Cochran and Cox (1957).

  As different from the standard analysis based on the tacitly known row,
column, treatment, additive model, the alternative involved here
essentially raises the question of model adequacy first discussed in the
one dimensional case, by Cox (1958). In some cases previous experience
may be available to assist in model determination; a Bayesian approach for
this has been investigated by Meeker (1975). When, however, the
determination is to be made solely on the basis of a single experiment
under analysis, stepwise regression procedures are indicated. If, for an
n x n latin square, a background polynomial surface of degree p, with no
non-zero coefficients, is fitted, there will be n^2 – (n–1) –
(1/2)(p+1)(p+2) degrees of freedom (df) for the estimate s^2 of $\sigma^2 = V(\epsilon_{ij})$
as compared with the (n–1)(n–2)df provided by the standard model. A cubic
surface could be fitted for a 5x5 square, for example, leaving 11df as
compared with 12df for the residual variance estimate. The 'warp-model'
would be expected to be superior in this case therefore, if the
expectations of the quartic row and column terms, but not those of lower
order cross-product terms, are equal to σ^2.

  Lastly here, a loss in simplicity by comparison with the usual model
is to be noted – that estimates of the parameters are no longer
independent. Accordingly it will be necessary to select the appropriate
elements from the inverse of the model specification matrix (X'X) for
determinations of, for example, confidence intervals for inter-treatment
contrasts.

5. Example

As part of an extensive investigation of millet cultivation reported by Li, Meng and Liu (1936), a Latin Square experiment was carried out to investigate the effects of different inter-row spacings on the grain yield. The data, which have long been used to illustrate the standard latin square analysis in the text Snedecor and Cochran (1989), are

reproduced in Table 2 wherein A,B,...,E denote inter-row spacings from A = 2" to E = 10" in steps of 2".

If a general surface background model can be specified independently of the particular data, inter-row and inter-column contrasts up to degree (n-1) for an n x n square can be isolated in the analysis of variance either individually or in two sets as in the usual analysis. The joint sum of squares for treatments and the $x^r z^s$ cross-product terms can then be subdivided to obtain the sum of squares for treatments eliminating the cross-product terms by covariance analysis as noted earlier. Thus, for the term in xz, the covariate w_{ij} for the plot in the ith row and jth column can be generated from the simple vector product

$$[w_{ij}] = [-2 \ -1 \ 0 \ 1 \ 2]'[-2 \ -1 \ 0 \ 1 \ 2].$$

Covariates for higher order cross-product terms can be similarly generated and a relatively simple forward model building analysis of covariance can be carried out to the degree required. When, however, the degree must be determined from the particular data, as in this example, a backward stepwise deletion procedure is recommended.

A convenient SAS program for this allowing the between-treatment sum of squares to be isolated at each stage, as is required, has been written by author J.B.M. With this, no coefficient was admitted into the general polynomial background function without achieving an F-ratio exceeding 3.0; for the data here this criterion gave the same model as that given on the basis of individual significance tests using a uniform exceedance probability level α, of less than 0.10. The ANOVA for the general polynomial model so obtained and the standard model are given in Tables 3 and 4 respectively.

Considerable differences between the two analyses and the corresponding inferences are indicated as a result of the appreciable reduction in the residual mean square; the coefficients of variation are 6.3% and 12.9% for the general polynomial and the standard models respectively. Notably also, neither the between-all-treatments mean square nor any of its individual, one degree of freedom, contrasts achieved 5% significance in the standard analysis whereas the between-all-treatments mean square in Table 4 was highly significant (P < 0.005). Further, partitioning the sum of squares gave values of 4313.08, 1825.28, 959.77 and 130.47 for the individual, Type 1 mean squares, for the linear, quadratic, cubic and quartic components of the response to spacing so that both the linear and quadratic components of the regression of yield on spacing achieved significance using the general polynomial model.

Estimates of mean yields are given in Table 5 and illustrated in Figure 1.

By comparison with the general polynomial model it appears that the usual model gave appreciable underestimates of the yields for the 6 and 8-inch spacings and an overestimate for the yield at the 10-inch spacing. Relatedly the usual model suggests that the yield decreased linearly with increased spacing from a maximum at the 2-inch spacing; the regression coefficient _is_ significantly different from zero if 65, from the quadratic, cubic and quartic components, is taken as a denominator mean square. The significance achieved for the linear and quadratic components

is particularly supportive of the general polynomial model because, together, they indicate that the maximum yield was achieved at the 6 inch spacing in agreement with the comment by the experimenters, Li et al. (1936) that this was the spacing commonly used for this crop – perhaps as a consequence of empirical sequential optimizations. Also supportive of the general surface model is the pragmatic and somewhat prescient comment (attributable to Cochran) on the standard analysis – 'The results of this experiment are probably disappointing. In trying to discover the best width of spacing, an investigator hopes to obtain a curved regression, with reduced yields at the narrowest and widest spacings, so that the range of spacings straddles the optimum. As it is, assuming the linear regression real, all that we have learned is that the best spacing may lie below 2 in.' – Snedecor and Cochran (1989).

The residuals from the two fitted models are given in Tables 6 and 7, those for the general polynomial model being based on the estimate

$$\text{background polynomial} = 29.69 + 85.93x - 18.43x^2$$
$$+203.29z - 87.69z^2 + 10.65z^3$$
$$-16.29xz + 4.62x^2z$$

based on axes giving $x = 1, ..., 5$ for plots along the top row and $z = 1, ..., 5$ for those down the left column and quasi-estimates 47.77, 40.86, 55.09, 41.47 and 0 for treatments A, ..., E respectively, obtained from the SAS General Linear Models Procedure.

The residuals in Table 7 are naturally in general smaller than those in Table 6 exceptions being the residuals in column 2. From Table 6 something of a pattern emerges in that over-estimates are found in the SE and NW corners of the square with under-estimates in the NW and SE corners; the pattern does not appear in the residuals in Table 7.

Finally, from an agronomic view it is of interest to estimate the mean yields per individual plant since these can be expected to change because of the relaxation of competition as spacing increases. The values obtained from the two models are compared in Figure 2. The difference is seen to be that, whereas on the basis of the row-column model the increase in yield per plant due to reduction in competition continues undiminishingly up to at least the 10-inch spacing, the background polynomial model suggests that the yield per plant asymptotically approaches a maximum at slightly more than the 10-inch spacing. This latter seems the more intuitively convincing behavior and it is noteworthy that the response curve itself could belong to either of the families of curves to which such relationships 'tend to conform qualitatively ...' illustrated and discussed in Seber and Wild (1989).

6. Discussion

Since the original presentation of this material, considerably more sophisticated and computationally intensive procedures have been advanced for treating the persistent and ubiquitous problem of identifying background effects, signals, and noise. The spatial, nearest neighbor and smoothing approaches, Wilkinson et al. (1983), Green (1984), Green et al. (1985) are highly relevant to the agronomic context. Green (1984) noted that '... the various methods proposed are based on differing statistical principles, so that comparison between methods has been difficult.' One such difficulty may be the assessment of just what is tacitly contributed to (or imposed on) the situation via assumptions such as isotropic or dominant direction inter-plot correlation. In this respect a 'good' linear model has pragmatic appeal because it can be regarded as a first order Taylor series approximation to a more fully veridical nonlinear model towards which sequential improvements can be envisaged. Relatedly the linear model approach gropes heuristically towards causes of phenomena

of which correlation procedures are essentially descriptive; a distinction analogous to that between investigations on genotypes and descriptions of phenotypes. A connection between the two approaches was, in fact, obtained in Cox (1962) where, with time instead of space, it was shown that correlations between observations on individuals (cows) at different times, could be 'predicted' from, or 'explained in terms of', their individual linear regressions. The polynomial background fitting described here is itself, of course, a smoothing procedure and other types of smoothing in the hands of versatile practitioners can also doubtless provide good description; both information and infants can, however, be suffocated by over-zealous smoothing.

The objective of the preceding exposition is to suggest that there are promising possibilities for the use of multiple regression and model adequacy testing procedures in analyses of data from compact yield experiments. The three main differences from the standard procedure are:

1) As initially noted by Hald (in Cox, 1950) the background polynomial model obviates a basic assumption of the standard model that the background response is constant over each plot.

2) As shown in Table 1, redundant-row and/or column parameters commonly occur in the standard model. The ensuing inefficiency may then partially explain the relatively rare use of Latin Square designs in field experiments; Cox (1950), for example, reported that, of 6317 experiments at North Carolina Experiment Station during 1942-48, only 91 (1.4%) used Latin Square designs.

3) The standard model excludes some 'warp'-parameters which can explain important components of background variability and reduce bias in the estimates of treatment effects and residual variance.

The possible advantages are, however, not free and further research is required on, as examples, the procedure used to select the background polynomial parameters, the validity - and possible relaxation – of assumptions re homogeneity of variance, position x treatment interactions and, especially, to provide confidence and prediction intervals in particular contexts. This latter relates to questions of: whether and in what contexts, parameters are fixed or random?, what quantities are estimable? and on the discrimination versus estimation properties of designs discussed in Cox (1961), subsumed in the remark - 'A prerequisite for the formal analysis of the experimental data is a relevant conceptual "frame of reference" ... based on some explicit elementary or primitive assumptions' (Wilk and Kempthorne, 1957), that which F. Yates had described as 'The inductive basis'.

Another important area for research is the provision of experimental designs specially adapted to the basis of a general polynomial background function. These might, by relaxation of the, each treatment equally often in each row and column condition, and by appropriate modification of randomization procedures, lead to improved adequacy testings and estimations. These might be particularly advantageous if a background function was demonstrably a durable, for example edaphic, attribute of the actual locale used for an experiment.

Finally here we would suggest the addition of the parenthesized to the admonition, Hunter (1987) that –

'It is seldom wise to assume a postulated (and, a priori, dubiously adequate) first order model to represent an unknown function exactly'.

Appendix A

An interesting discussion of cultivation practices in relation to fertility is available in White (1962), "The earliest plough was essentially an enlarged digging stick dragged by a pair of oxen. ... . Its conical or triangular shape does not normally turn over the soil and, it leaves a wedge of undisturbed earth between each furrow. Thus cross-ploughing is necessary, with the result that, in regions where the scratch-plough is used, fields tend to be squarish in shape, ...". Because of generally heavier soils a new plough, the heavy plough, was developed for Northern Europe. White continues: ". the heavy plough handled the clods with such violence that there was no need for cross-ploughing. ... . The new plough, by eliminating cross-ploughing, tended to change the shape of fields ... from squarish to long and narrow with a slightly rounded vertical cross-section for each strip field ... . These strips were normally ploughed clockwise, with the sod turning over and inward to the right. As a result, with the passing of the years, each strip became a long low ridge, assuring a crop on the crest even in the wettest years, and in the intervening long depression, or furrow, in the driest seasons." And, because it took eight rather than two oxen to pull, peasant cooperatives were formed and ... all the lands of a village had to be reorganized into vast, fenceless 'open fields' ploughed in long narrow strips." The above thesis is attributed by White to one Marc Bloch.

With respect to experimental design in relation to ploughing practice, the suggestion is that, after the scratch-plough was superseded by the heavy plough, randomized block designs would be more appropriate than latin square designs (cf. Table 1). White, however, notes that Bloch's thesis has been authoritatively challenged and gives counter indications of narrow strips with scratch-ploughing and flat strips with heavy ploughing noting that strips were often related to inheritance divisions. Nevertheless he further remarks "Scholars like ... Bloch had a keen sense of fact and they observed the averages." Whatever the merits of Bloch's thesis, its basis changed with history and various enclosure acts until, according to Harris (1973), "By 1850, certainly by 1870, open fields had disappeared from all but a few rare villages." The photograph, Figure 3, adapted from Morgan (1975), shows residuals of peasants' strips at Stoke Newington, Oxfordshire, England.

Summary

Data are presented to suggest that the regular row, column, treatment additive structural model used to analyze Latin Square experiments is not necessarily veridical for field experiments wherein non-Cartesian edaphic, background variability components can occur to contaminate the residual mean square. Using only a simple criterion for term selection it is shown, and exemplified, that representation of the background by a more general polynomial can, by separating out row x column interaction terms, substantially reduce bias in the residual mean square. A note on medieval cultivation practices and experimental design is appended.

Acknowledgements

References

Cochran, W. G. and Cox, G. M. (1957, 2nd Edition). Experimental Designs. John Wiley, New York.

Cox, C. P. (1974). An alternative model for the analysis of latin square experiments. Department of Statistics Seminar. Iowa State University

Cox, C. P. (1962). The relation between covariance and individual curvature analyses of experiments with background trends. <u>Biometrics</u> <u>18</u>, 12-21.

Cox, C. P. (1961). A practical application of a theoretically inefficient design. <u>Biometrics</u> <u>17</u>, 646.

Cox, C. P. (1958). The analysis of latin square designs with individual curvatures in one direction. <u>J. R. Statist. Soc. B</u> <u>20</u> 193-204.

Cox, C. P., Hosking, Z. D. and Line, C. (1958). Within-field changes in herbage composition and soil moisture. <u>J. Brit. Grassland Soc.</u> <u>13</u> 187.

Cox, G. M. (1950). A survey of types of experimental designs. Discussion. <u>Biometrics</u> <u>6</u> 317-318.

Eden, T. and Maskell, E. J. (1928). The influence of soil heterogeneity on the growth and yield of successive crops. <u>J. Agric. Sci.</u> <u>18</u> 163.

Fisher, R. A. (1946, 10th Edition). Statistical Methods for Research Workers. Oliver and Boyd, London.

Fisher, R. A. (1947, 4th Edition). The Design of Experiments. Oliver and Boyd, London.

Goulden, C. H. (1939). Methods of Statistical Analysis. John Wiley, New York.

Green, P. J. (1984). Linear models for field trials, smoothing and cross-validation. Tech. Report 2779. Math Res. Center. University of Wisconsin.

Green, P. J., Jennison, C. and Seheult, A. H. (1985). Analysis of field experiments by least squares smoothing. <u>J. R. Stat. Soc. B</u> <u>47</u>, 299-315.

Harris, A. In Derby H. C. (Ed). (1973). A New Historical Geography of England. Cambridge Univ. Press.

Hunter, J. S. (1987). Are some latin squares better than others? Ch. 9 In Mallows C. L. (Ed) Design, Data and Analysis. John Wiley, New York.

Hutchinson, J. B. and Panse, V. G. (1935). Studies in the technique of field experiments. <u>Indian J. Agric. Sci.</u> <u>5</u> 523.

Li, H. W., Meng, J. and Liu, T. N. (1936). <u>J. Amer. Soc. Agron.</u> <u>28</u> 1.

Meeker, J. B. (1975). The use of model information indicating more than one possible model. Ph.D. thesis. Department of Statistics, Iowa State University.

Morgan, G. (1975). Life in a Medieval Village. Cambridge University Press. London.

Rayner, A. A. (1967). Biometry for Agriculture Students. University of Natal Press, Pietermaritzburg.

Seber, G. A. F. and Wild, C. J. (1989). Nonlinear Regression. John Wiley, New York.

Snedecor, G. W. and Cochran, W. G. (1989, 8th Edition). Statistical Methods. Iowa State University, Ames, Iowa.

Steel, R. G. D. and Torrie, J. H. (1960). Principles and Procedures of
    Statistics. McGraw-Hill, New York.

White, Lynn, Jr. (1962). Medieval Technology and Social Change.
    Clarendon Press, Oxford.

Wiebe, G. A. (1935). Variation and correlation in grain yield among 1500
    wheat nursery plots. J. Agric. Res. 50 331.

Wilk, M. B. and Kempthorne, O. (1957). Non-additivities in a latin square
    design. J. Amer. Stat. Assoc. 52, 218.

Wilkinson, G. N., Eckert, S. R., Hancock, T. W. and Mayo, O. (1983).
    Nearest neighbor (NN) analysis of field experiments (with discussion).
    J. R. Stat. Soc. B 45 151-211.

Table 1
Examples of latin square analyses

Source		rows	columns	treatments	Residual ms;df
			F-ratios		
Goulden (1939)		4.99	1.50	20.78	2.34;12
5x5, wheat	P[a]	0.013	0.263		
Fisher (1946)		8.14	1.35	uniformity	130.3;16
5x5, mangolds	P	9E-4	0.295	trial	
Fisher (1947)		7.10	3.20	162.53	1527;20
6x6, potatoes	P	6E-4	0.028		
Cochran & Cox (1957)		2.39	4.00	10.38	26.2;12
5x5, potatoes[b]	P	0.109	0.027		
Steel and Torrie (1960)		1.44	5.04	58.47	0.45;6
4x4, wheat	P	0.321	0.044		
Jowett (1967)		6.42	2.03	2.56	58.7;20
6x6, sorghum	P	0.001	0.118		
Rayner (1967)		34.54	13.38	19.07	16.84;12
5x5, grain	P	1.7E-6	2E-4		
Snedecor & Cochran (1989)		3.22	1.45	0.98	1056;12
5x5, millet	P[a]	0.052	0.277		

[a]Exceedance probability for the above F-ratio

[b]Individual plot yields not given

Table 2
yield of millet per plot (g)

row	1	2	column 3	4	5
1	B:257	E:230	A:279	C:287	D:202
2	D:245	A:283	E:245	B:280	C:260
3	E:182	B:252	C:280	D:246	A:250
4	A:203	C:204	D:227	E:193	B:259
5	C:231	D:271	B:266	A:334	E:338

The required subtotals are:

	1	2	3	4	5
row	1255	1313	1210	1086	1440
column	1118	1240	1297	1340	1309
treatment	1349	1314	1262	1191	1188
	A	B	C	D	E

Table 3
ANOVA of millet yield data, general polynomial model

Source of variation	df	ms
x	1	4646.48
x^2	1	1462.86
z	1	408.98
z^2	1	4657.73
z^3	1	8166.42
xz	1	5387.56
x^2z	1	1333.03
treatments (adjusted)	4	1807.15
residual	13	252.29
total	24	1523.81

Table 4
ANOVA of millet yield data, standard model

Source of variation	df	ms
rows (z)	4	3400.3
columns (x)	4	1536.5
treatments	4	1039.1
⎡linear	1	3960 ⎤
⎢quadratic, cubic,	3	65 ⎢
⎣quartic		⎦
residual	12	1055.6
total	24	1523.8

Table 5
mean yield estimates (g)

model	spacing (in) 2	4	6	8	10
general polynomial	262.89	255.98	270.21	256.59	215.12
standard	269.8	262.8	252.4	238.2	237.6

Table 6
residuals from the standard row-column model

row	column 1	2	3	4	5
1	23.92:B	− 2.28:E	3.12:A	19.92:C	−44.68:D
2	24.92:D	6.92:A	−10.28:E	− 9.08:B	−12.48:C
3	−16.88:E	3.52:B	30.52:C	2.12:D	−19.28:A
4	− 3.28:A	− 9.28:C	16.52:D	−25.48:E	21.52:B
5	−28.68:C	1.12:D	−39.88:B	12.52:A	54.92:E

Table 7
residuals from the general polynomial model

row	column 1	2	3	4	5
1	4.36	− 9.98	− 9.33	18.39	1.69
2	−11.34	− 5.42	− 3.05	2.08	− 2.78
3	11.62	17.40	16.96	− 8.49	− 6.72
4	13.02	−14.23	1.37	−12.29	− 8.38
5	−18.34	16.78	−15.46	8.45	13.70

Figure 1

Millet: Yield/plot related to inter-row spacing

inter-row spacing (in.)
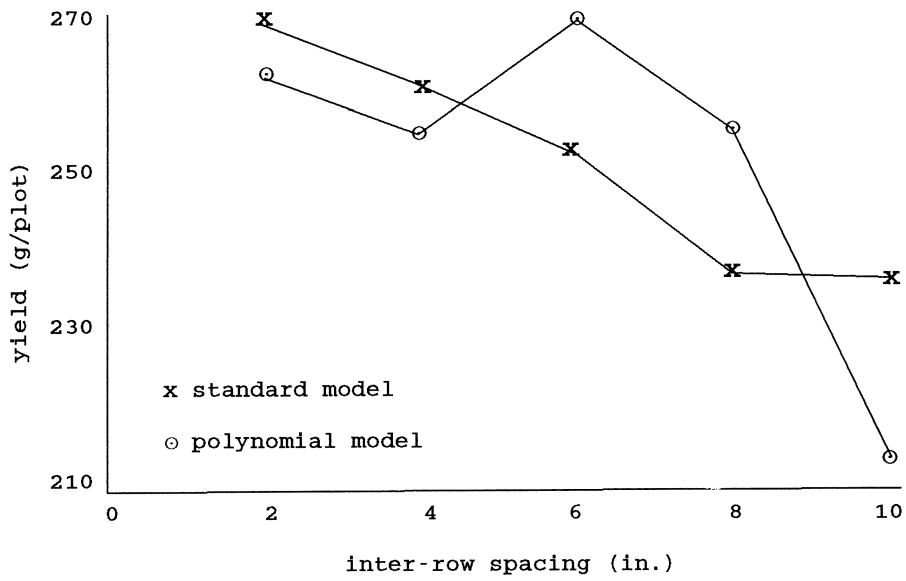
Figure 2

Millet: Yield/plant related to inter-row spacing
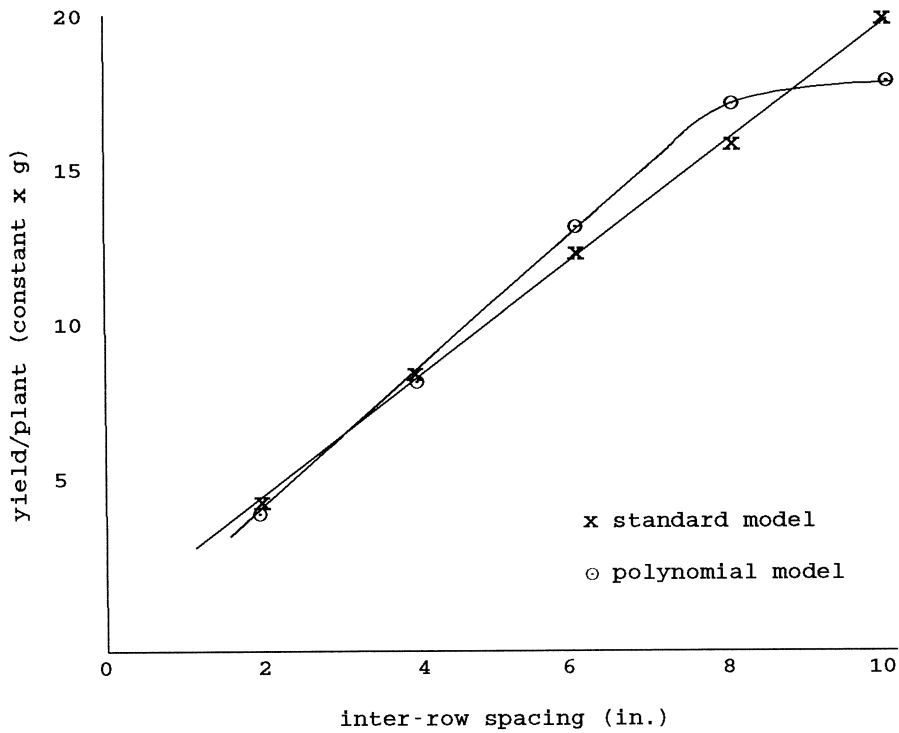

inter-row spacing (in.)

Figure 3

Peasants' strips in England
(adapted from Morgan, 1975)