

University of Wollongong

Research Online

Faculty of Engineering and Information
Sciences - Papers: Part A

Faculty of Engineering and Information
Sciences

1-1-1989

Efficiency of neighbour analysis for replicated variety trials in Australia

Brian R. Cullis

NSW Agriculture and Fisheries Research Institute, bcullis@uow.edu.au

A. C. Gleeson

NSW Agriculture and Fisheries Research Centre

Follow this and additional works at: <https://ro.uow.edu.au/eispapers>



Part of the [Engineering Commons](#), and the [Science and Technology Studies Commons](#)

Recommended Citation

Cullis, Brian R. and Gleeson, A. C., "Efficiency of neighbour analysis for replicated variety trials in Australia" (1989). *Faculty of Engineering and Information Sciences - Papers: Part A*. 812.

<https://ro.uow.edu.au/eispapers/812>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Efficiency of neighbour analysis for replicated variety trials in Australia

Abstract

Use of a one-dimensional neighbour method of analysis in 1019 variety trials of a range of crops conducted by plant breeders in four states of Australia in 1985–87 resulted in an average reduction of 42% in variances of varietal yield differences compared with conventional randomized complete block analysis. Of these trials, 219 were designed as square, rectangular or generalized lattices and the average reduction in variances of varietal yields with incomplete block analysis and recovery of interblock information was 33%. The results emphasized that plots should be wide enough to avoid interplot competition, and that neighbour analysis is of most benefit in trials with short plots or when the field layout has many plots in a row.

Keywords

replicated, neighbour, field, efficiency, australia, analysis, trials

Disciplines

Engineering | Science and Technology Studies

Publication Details

Cullis, B. R. & Gleeson, A. C. (1989). Efficiency of neighbour analysis for replicated variety trials in Australia. *The Journal of Agricultural Science*, 113 (2), 233-239.

Efficiency of neighbour analysis for replicated variety trials in Australia

B. R. CULLIS AND A. C. GLEESON*

NSW Agriculture and Fisheries Research Institute, Wagga Wagga, NSW 2650, Australia

** NSW Agriculture and Fisheries Research Centre, Tamworth, NSW 2340, Australia*

(Revised MS received 30 March 1989)

SUMMARY

Use of a one-dimensional neighbour method of analysis in 1019 variety trials of a range of crops conducted by plant breeders in four states of Australia in 1985–87 resulted in an average reduction of 42% in variances of varietal yield differences compared with conventional randomized complete block analysis. Of these trials, 219 were designed as square, rectangular or generalized lattices and the average reduction in variances of varietal yields with incomplete block analysis and recovery of interblock information was 33%. The results emphasized that plots should be wide enough to avoid interplot competition, and that neighbour analysis is of most benefit in trials with short plots or when the field layout has many plots in a row.

INTRODUCTION

Since the introduction of randomized complete block (RCB) designs by Fisher (1925), there have been three main statistical approaches to improving the efficiency of analyses of variety trials. These approaches remove the effects of within-site heterogeneity from variety and variance estimates and are (i) the use of incomplete block designs, (ii) Papadakis' covariate adjustment and (iii) neighbour or spatial analysis in which an attempt is made to model the error structure.

Within-site heterogeneity is of particular concern in trials with a large number of varieties. This prompted Yates (1936) to introduce lattice designs for such trials. However, it was not until Patterson & Williams (1976) extended Yates' method of construction to remove restrictions on numbers of varieties and generate generalized lattice designs that widespread use was made of incomplete block designs in variety trials.

In a study of 244 cereal variety trials in the UK since 1975, Patterson & Hunter (1983) estimated an average reduction of 30% in the variances of varietal yields from using generalized lattices instead of RCBs; they also suggested that there is potential for further improvement using neighbour methods.

Papadakis (1937) proposed adjusting yields of field experiments by covariance analysis with respect to treatment-corrected yields of adjacent plots. Kempton & Howes (1981), in a study of 118 wheat variety trials, estimated that the iterated Papadakis procedure

(Bartlett 1978) resulted in an 8.7% increase in precision compared with incomplete block analysis. Wilkinson *et al.* (1983) subsequently showed, however, that the iterated Papadakis procedure had a substantial positive bias in the treatment F-ratio, resulting in more significant differences being declared than actually exist.

There has been substantial statistical development of neighbour or spatial analysis since Wilkinson *et al.* (1983) first proposed 'trend plus error' models for field experiments. Gleeson & Cullis (1987) present a unified approach to one-dimensional neighbour analysis based on modelling the trend as an autoregressive-integrated-moving average (ARIMA) process and estimating the parameters using the method (Patterson & Thompson 1971) of residual maximum likelihood (REML). The class of ARIMA models for trend encompass other one-dimensional neighbour models proposed by Green *et al.* (1985), Besag & Kempton (1986) and Williams (1986).

Cullis *et al.* (1989) have derived a diagnostic test of whether a particular ARIMA model adequately describes the trend in a field experiment. In a large simulation study based on uniformity data, Lill *et al.* (1988) showed that substantial gains in the accuracy of treatment estimation are achieved using neighbour analysis instead of incomplete block analysis. Their procedure also had low levels of bias in the treatment F-ratio (between -1.4 and 2.4% depending on the trend model fitted and the relative magnitude of trend and measurement error).

Table 1. Summary of design size and field layouts of variety trials from three states in Australia

State	Treatments		Replicates		Rows		Plots per row	
	Range	Mean	Range	Mean	Range	Mean	Range	Mean
South Australia	38-120	74	3-6	4.8	12-15	13	13-45	27
Western Australia	12-120	54	2-6	3.2	3-24	12	4-40	14
New South Wales	9-125	31	2-5	3.2	2-15	4	10-74	25

With the improved accuracy of treatment estimation and low levels of bias in variance estimation demonstrated in simulation studies, the question arises as to how efficient neighbour methods are with data from actual experiments. The primary aim of this paper is to examine the efficiency of neighbour methods as discussed by Gleeson & Cullis (1987). As a basis for this study, we used the results of 1019 variety trials conducted by plant breeders in four states of Australia in 1985-87 on wheat, barley, oats, triticale, lupin, chickpea and rape. A secondary aim is to compare the efficiency of the neighbour analysis with the efficiency of incomplete block analysis with recovery of interblock information.

DESCRIPTION OF TRIALS

The trials used in this study were selected from mid to late-stage variety trials in plant breeding programmes across Australia. Any trial with more than ten missing values was rejected. The trials had 9-125 varieties in 2-6 replicates. Each trial was laid out as a rectangular array of several rows of plots, with the same number of plots in each row. Plots were generally long and thin, and plots within rows were adjacent on their longer side. For such trials, correlation among plots is likely to be dominant in the direction along rows and in such cases one-dimensional neighbour analysis is usually appropriate.

The 76 trials from South Australia (SA) were from a wheat breeding programme based at Roseworthy Agricultural College. The trials were all designed as RCBs and generally had more replication and more varieties than trials from other states (Table 1). The mean yield was low, reflecting the poor soils and low average annual rainfall in SA. The range of mean yields and RCB coefficients of variation (c.v. %) are listed in Table 2. Two plot widths were used in these trials: plots were 0.83 m wide in 31 trials and 1.25 m wide in 45 trials (Table 3). All plots were sown 5 m long.

The 74 trials from Queensland (Qld.) are from a wheat breeding programme based at the Queensland Wheat Research Institute, Toowoomba. All trials were designed as 7×7 triple lattices. The trial layout was either a single row of 147 plots or a rectangular

array of seven rows of 21 plots in three contiguous replicates.

In Western Australia (WA), the 526 trials represent five different crops (Table 2). All trials were designed as RCBs. For the lupin trials, the sown plot was 5×1 m, whereas for the other four crops the sown plot was 5×1.25 m. Of the 526 trials, 516 were laid out in 12 rows. The predominant use of a 12-row layout resulted in the average number of plots per row being lower than in other states (Table 1).

In New South Wales (NSW), the 343 trials covered five crops. Trials were designed as either RCBs or incomplete block (IB) designs; the IB designs were square, rectangular or generalized lattices (Patterson & Williams 1976). Plot sizes were generally constant within each breeding programme. The 113 NSW wheat trials were from four separate breeding programmes, hence the various plot sizes. There was a substantial range in the RCB c.v. and trial mean yields for all crops (Table 2).

CALCULATION OF EFFICIENCY

The efficiency (Yates 1939) of an IB (or neighbour) analysis is calculated as the average variance of pairwise varietal differences from an RCB analysis divided by the average variance of pairwise varietal differences from an IB (or neighbour) analysis. This is the measure of efficiency used by Patterson & Hunter (1983).

Data from each trial were analysed at least twice, once as an RCB (all IB designs were resolvable), and then by one-dimensional neighbour analysis within rows as described by Gleeson & Cullis (1987), hereafter referred to as GC analysis. For the GC analysis, successive trend models of higher order are fitted until an error model diagnostic proposed by Cullis *et al.* (1989) indicates that the model is adequate. This test can be applied to examine whether the simplest trend model, equivalent to fitting the row mean as the trend in each row, is adequate or whether the next simplest model, which assumes the first differenced trend is independently, normally distributed, should be tried, and so on. Other neighbour analyses, for example those of Williams (1986) and Besag & Kempton (1986), rather than testing the

Table 2. Summary of design, plot size, c.v. % and yield (kg/ha) of variety trials from four states of Australia

State	Crop	Number of trials		Plot size (m)	C.v. %		Yield	
		Incomplete block	Randomized complete block		Range	Mean	Range	Mean
South Australia	Wheat	0	76	5 × 0.8-1.25	7.0-39.4	16.0	172-1821	831
Queensland	Wheat	74	0	5 × 1.25	3.9-33.4	10.4	1123-5423	3010
Western Australia	Wheat	0	176	5 × 1.25	2.3-44.4	7.1	250-2973	1309
	Barley	0	54	5 × 1.25	2.4-44.5	13.0	510-3636	2102
	Oats	0	54	5 × 1.25	3.3-15.0	7.2	360-3506	1722
	Triticale	0	14	5 × 1.25	3.8-12.9	6.8	688-3104	1745
New South Wales	Lupin	0	228	5 × 1.0	9.9-76.4	25.3	113-2067	903
	Wheat	59	54	17-20 × 1.0-1.8	4.0-91.2	10.7	232-8382	3591
	Barley	65	38	10-20 × 1.8	3.5-40.1	10.6	755-6025	3255
	Oats	26	27	20 × 1.8	4.5-19.7	10.4	1107-5476	3677
	Chickpea	15	0	6 × 1.8	5.7-44.6	13.5	335-3447	2176
	Rape	0	59	18 × 1.8	4.1-30.9	14.0	311-3836	2005

adequacy of the model, assume that the 'independent first differenced trend' model is always adequate. To make the results of this efficiency study relevant to other neighbour analyses, the independent first differenced trend model was taken to be the simplest trend model, so that the simpler row-mean trend model was not accepted even when the diagnostic indicated it was adequate. Because of this, the reported mean efficiencies of GC analysis in this paper may have been underestimated.

The 239 IB trials (165 in NSW and 74 in Qld.) were analysed by the appropriate IB analysis with recovery of interblock information using the program REML (Robinson & Digby 1987).

RESULTS AND DISCUSSION

The distributions of the GC efficiencies were positively skew for each state (Fig. 1), as were the distributions of the IB efficiencies for trials in NSW and Qld. GC efficiency increased with increasing site variability, as measured by the RCB c.v. For most of the trials the independent first differenced trend model was adequate.

South Australia

The arithmetic mean GC efficiency was 2.01, with a median of 1.90. An interesting feature of the SA results is the difference between wide and narrow plots (Table 3). The mean GC efficiency was greater for wide (1.25 m) than for narrow (0.83 m) plots. For wide plots, the mean GC efficiency increased with increasing number of plots per row, whereas for narrow plots there was no apparent gain in GC efficiency with more plots per row. These results suggested that interplot competition can occur in very narrow plots: the negative correlation between neighbouring plots due to interplot competition could ameliorate the positive correlation due to trend, and a weaker (positive) correlation would lead to a lower efficiency for neighbour analysis.

For plots wide enough to avoid interplot competition, the increased GC efficiency resulting from more plots per row was also intuitively reasonable. More plots per row results in a greater area covered by the row and the likelihood of greater variability along the row; the greater the number of plots per row the better the estimate of the between-plot correlations within a row.

Queensland

The arithmetic mean GC efficiency was 1.94, with a median of 1.61, whilst for IB analysis the mean efficiency was 1.63, with a median of 1.44. This corresponds to an average reduction of 51% in average pairwise variance of varietal yield differences

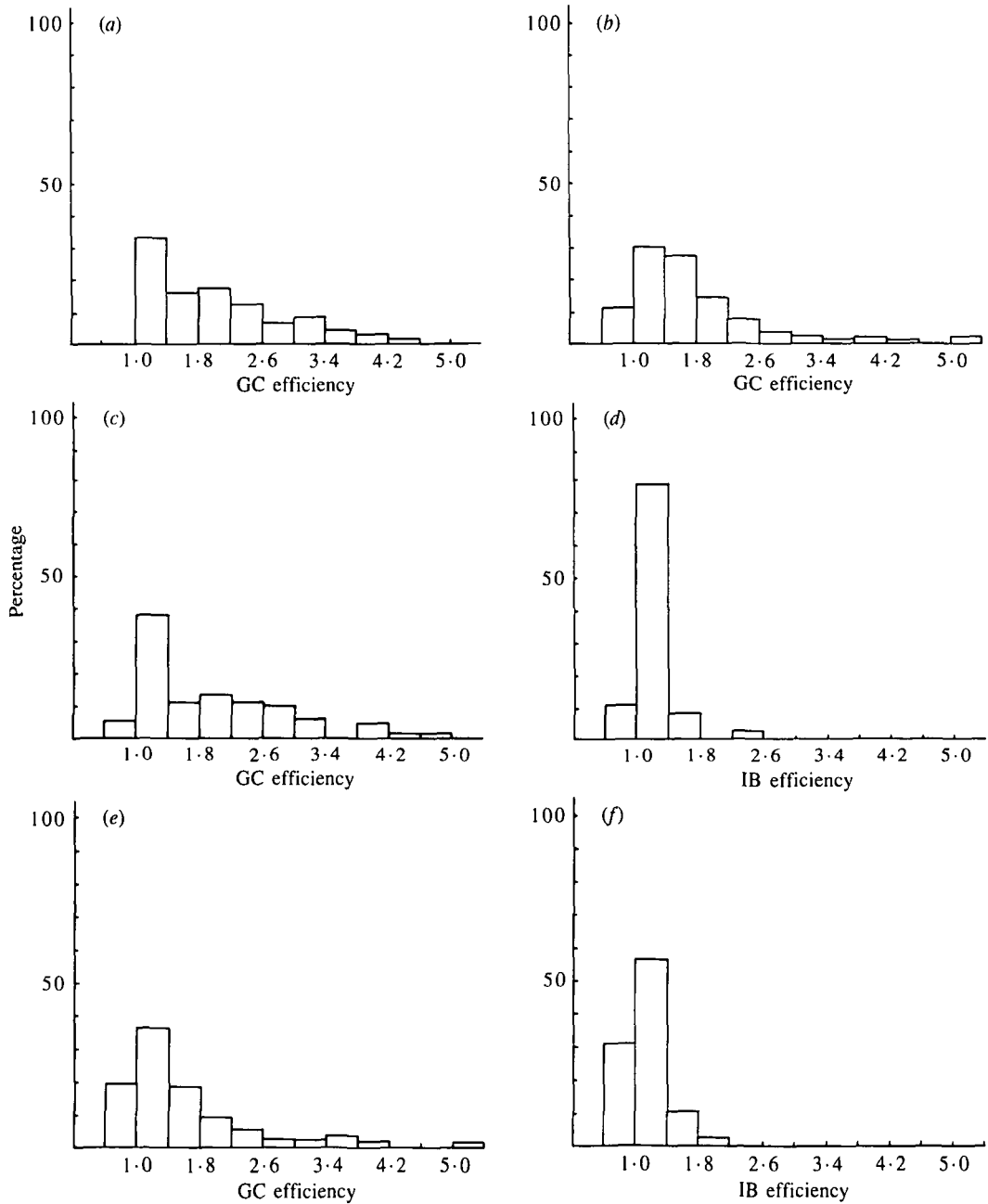


Fig. 1. Histograms of the Gleeson & Cullis (GC) and incomplete block (IB) efficiencies in (a) South Australia, (b) Western Australia, (c) and (d) Queensland and (e) and (f) New South Wales.

when using GC analysis, compared with a 39% reduction for IB analysis. No attempt was made to examine the effect of trial layout on the efficiencies because each site had only one of the layouts, hence layout effects were confounded with site effects.

Western Australia

There was no discernible effect of crop type on the mean GC efficiency. The arithmetic mean efficiency of GC analysis was 1.74, with a median of 1.48. The WA

Table 3. Mean efficiency of Gleeson & Cullis analysis for South Australia trials classified by plot width and number of plots per row. Number of trials shown in parentheses

Plot width (m)	Number of plots per row				Margin
	< 10	11-20	21-30	> 30	
1.25	—	1.93 (15)	2.03 (7)	2.46 (23)	2.22 (45)
0.83	—	1.77 (13)	1.67 (14)	1.65 (4)	1.71 (31)
	—	1.86 (28)	1.79 (21)	2.34 (27)	2.01 (76)

Table 4. Mean efficiency of Gleeson & Cullis (GC) analysis for Western Australia trials classified by number of plots per row. Number of trials shown in parentheses

Number of plots per row	Mean GC efficiency
< 10	1.53 (223)
11-20	1.83 (211)
21-30	1.92 (75)
> 30	2.37 (17)
	1.74 (526)

results indicated increased efficiency from a larger number of plots per row (Table 4), in the same way as the SA results did for wide plots.

New South Wales

The arithmetic mean efficiencies of GC and IB analyses were 1.62 and 1.46, respectively, corresponding to an average 38% reduction in average pairwise variance from GC analysis, compared with a 32% reduction for IB analysis. The median efficiencies of GC and IB analyses were 1.32 and 1.26, respectively.

An increase in GC efficiency with increasing plots per row was again evident for most crops. Table 5 presents the mean GC efficiency classified by crops \times number of plots per row in each trial; the wheat results were subdivided into narrow- and wide-plot trials. The results from the narrower plots used in 40 of the wheat trials again suggested that interplot competition reduces the efficiency of GC analysis.

Another feature of Table 5 is the low mean GC efficiencies for rape. If we assume the same level of trend at sites where the rape trials were conducted as at the sites where the other crops were grown, it appears that the experimental error associated with rape is larger than that of other crops. This may be because rape is an outcrossing species with poor seedling growth, which may increase variability both within and between plots of the same variety.

CONCLUSIONS

The arithmetic mean efficiency of GC analysis over the four states was 1.73 ($n = 1019$), corresponding to a 42% reduction in average pairwise variance. In the Qld. and NSW IB designs, the average efficiency of IB analysis was 1.50 ($n = 219$), a 33% reduction in average pairwise variance. This result is similar to the value of 1.43 ($n = 244$) reported by Patterson &

Table 5. Mean efficiency of Gleeson & Cullis analysis for New South Wales trials classified by crop and number of plots per row. Number of trials shown in parentheses

Crop	Plot size (m)	Number of plots per row				Margin
		< 10	11-20	21-30	> 30	
Rape	18 \times 1.8	1.06 (1)	1.27 (52)	1.25 (6)	—	1.26 (59)
Chickpea	6 \times 1.8	—	—	2.04 (15)	—	2.04 (15)
Wheat:						
narrow plots	17 \times 1.0	—	1.03 (6)	1.69 (30)	1.62 (4)	1.58 (40)
wide plots	20 \times 1.8	—	1.65 (31)	1.66 (20)	1.87 (22)	1.72 (73)
Oats	20 \times 1.8	0.92 (1)	1.46 (4)	1.51 (18)	1.66 (30)	1.58 (53)
Barley	10-20 \times 1.8	1.13 (5)	1.52 (16)	1.65 (72)	1.87 (10)	1.63 (103)
		1.09 (7)	1.41 (109)	1.67 (161)	1.92 (66)	1.62 (343)

Hunter (1983). The mean GC efficiency for the 219 IB designs was 1.77 (or a 44% reduction in average pairwise variance), so that an average further reduction of 11% in average pairwise variance can be expected from using GC rather than IB analysis.

The consistently lower gains in GC efficiency when using narrow plots may be due to interplot competition. The types of variety trial considered in this paper were aimed at evaluating the potential of a variety as a monoculture crop, so that interplot competition in these trials acted as an extraneous source of variation. A sensible precaution for such trials is to have the plots wide enough or sufficiently well spaced to avoid interplot competition.

In NSW, where plots were longest, the mean efficiency of GC (and IB) analysis was least. The mean efficiencies of GC analysis for the three states with shorter plots (Qld., SA and WA) and the chickpea breeding programme in NSW were all larger than that for the other NSW plant breeding programmes. Such an effect is not surprising because long plots cover a larger slice of the mosaic of field variation, and this should lead to a reduction in the correlation between adjacent plots. There is therefore a greater need for, and benefit from, neighbour analysis in trials with short plots.

There are some examples (e.g. Williams & Lockett 1988) for which row+column analysis has proved more efficient than one-dimensional neighbour analysis such as the GC analysis used in this study. For variety trials with a large number of rows, such as several of the WA and SA trials in this study, a two-dimensional analysis may be more efficient. B. R. Cullis & A. C. Gleeson (unpublished) have recently extended the GC analysis used in this paper to two

dimensions, with appropriate diagnostic checks of model adequacy. A comparison of the efficiency of the two-dimensional spatial analysis with the conventional row+column analysis of such trials will be of interest for investigating the difference in efficiency between spatial and randomization-based analysis in two dimensions.

The results of this study suggest that, given an adequate plot width, GC analysis efficiency increases as the number of plots per row in the trial increases. This is intuitively sensible, because the increased area covered by a row of more plots is likely to be more variable, and the larger number of plots will enable more accurate estimation of that variability.

We believe that, with (i) the demonstrated increase in accuracy of treatment estimation, (ii) the low levels of bias in variance parameters and treatment F-ratios (Lill *et al.* 1988), (iii) the development of a test of whether a neighbour analysis is needed and, if so, whether the field neighbour model is adequate (Cullis *et al.* 1989) and now (iv) the results of this paper showing greater efficiency than IB analysis, neighbour analysis can now be confidently recommended for small-plot field experiments such as variety trials.

We gratefully acknowledge the skilled and diligent technical assistance of R. Prangnell in this work. We also acknowledge R. Gilmour (WA), G. Hollamby (SA), R. Eisemann (Qld.) and the plant breeders of NSW Agriculture and Fisheries for their assistance in providing the data for this study. We thank R. Gilmour (WA) for his helpful suggestions on an earlier draft. The Commonwealth Wheat Industry Research Council is gratefully acknowledged for their financial assistance.

REFERENCES

- BARTLETT, M. S. (1978). Nearest neighbour models in the analysis of field experiments. *Journal of the Royal Statistical Society, B* **40**, 147-174.
- BESAG, J. & KEMPTON, R. (1986). Statistical analysis of field experiments using neighbouring plots. *Biometrics*, **42**, 231-251.
- CULLIS, B. R., MCGILCHRIST, C. A. & GLEESON, A. C. (1989). Error model diagnostics in the general linear model with applications to the analysis of repeated measures and field experiments. *Journal of the Royal Statistical Society, B* (in press).
- FISHER, R. A. (1925). *Statistical Methods for Research Workers*. (1st edn) Edinburgh: Oliver and Boyd.
- GLEESON, A. C. & CULLIS, B. R. (1987). Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. *Biometrics*, **43**, 277-288.
- GREEN, P. J., JENNISON, C. & SEHEULT, A. H. (1985). Analysis of field experiments by least squares smoothing. *Journal of the Royal Statistical Society, B* **47**, 299-315.
- KEMPTON, R. A. & HOWES, C. W. (1981). The use of neighbouring plot values in the analysis of variety trials. *Applied Statistics*, **30**, 59-70.
- LILL, W. J., GLEESON, A. C. & CULLIS, B. R. (1988). Relative accuracy of a neighbour method for field trials. *Journal of Agricultural Science, Cambridge*, **111**, 339-346.
- PAPADAKIS, J. S. (1937). Méthode statistique pour des expériences sur champ. *Bulletin de l'Institut d'Amélioration des Plantes à Salonique*, No 23.
- PATTERSON, H. D. & HUNTER, E. A. (1983). The efficiency of incomplete block designs in National List and Recommended List cereal variety trials. *Journal of Agricultural Science, Cambridge*, **101**, 427-433.
- PATTERSON, H. D. & THOMPSON, R. (1971). Recovery of interblock information when block sizes are unequal. *Biometrika*, **58**, 545-554.
- PATTERSON, H. D. & WILLIAMS, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika*, **63**, 83-92.

- ROBINSON, D. & DIGBY, P. G. N. (1987). REML – *Residual Maximum Likelihood Program*. Edinburgh: Scottish Agricultural Statistics Service.
- WILKINSON, G. N., ECKERT, S. R., HANCOCK, T. W. & MAYO, O. (1983). Nearest neighbour (NN) analysis of field experiments (with discussion). *Journal of the Royal Statistical Society, B* **45**, 152–212.
- WILLIAMS, E. R. (1986). A neighbour model for field experiments. *Biometrika*, **73**, 279–287.
- WILLIAMS, E. R. & LUCKETT, D. J. (1988). The use of uniformity data in the design and analysis of cotton and barley variety trials. *Australian Journal of Agricultural Research*, **39**, 339–350.
- YATES, F. (1936). A new method of arranging variety trials involving a large number of varieties. *Journal of Agricultural Science*, **26**, 424–455.
- YATES, F. (1939). The recovery of inter-block information in variety trials arranged in three-dimensional lattices. *Annals of Eugenics*, **9**, 135–156.