EXPERIENCE IN USE INCOMPLETE BLOCK DESIGNS IN POLAND

Wiesław Pilarczyk
Department of Mathematical and Statistical Methods, Agricultural University of Poznań and
The Research Centre of Cultivar Testing, Słupia Wielka, Poland

1. Introduction

Incomplete block designs are now widely used in plant breeding and variety testing around the world. At the beginning of their use, practically just after the Second World War, mainly lattice designs (square, rectangular and cubic) were used. These designs were resolvable and blocks were of the same size. They suffer from one very serious disadvantage that they exist only for some very limited numbers of treatments. In Poland they were commonly used in plant breeding (mainly in s sugar beet and cereals) in the years 1950-1980 despite of problems with their statistical analysis in pre-computer times. In 1981 the wide class of so called α-designs (Patterson and Williams, 1976) was introduced to practice in official variety testing in Poland by The Research Centre of Cultivar Testing. These designs exist for every number of treatment for different block sizes, but in some cases blocks of two sizes must be used. The aim of this paper is to report and conclude on the efficiency of these designs in official variety testing in Poland.

2. Designs currently used

When resolvable designs were introduced in 1981, it was approximately at the same time when the number of replicates was reduced from six to four. It was expected that introducing new (potentially more efficient) designs could compensate, at least partly, the loss caused by reduction of number of replicates. This expectation was justified by hope that smaller blocks could better reflect the field variability than complete blocks and lattice designs with very limited numbers of treatment possible to be compared and with fixed sizes of incomplete blocks. The arrays prepared by Williams [1975] were used for generating α-designs. In fact new designs are 1-resolvable block designs, what means that every complete replicate (superblock) consists of s incomplete blocks of k or (k-1) plots each. New designs existed for every number of treatments and were much more flexible from the point of view of block sizes. Having no experience in their use, new designs were applied with block sizes (number of plots within every block) approximately equal to square root of the number of treatments. It followed from the suggestion given in a paper by Patterson and others [1978]. Usually the designs with blocks of 4-7 plots were used. After collecting extensive experimental data, and after their analysis using so-called post-blocking method, Pilarczyk [1987] shoved that – at least in experiments on cereals – the most effective were designs with blocks consisting of 8-12 plots. In experiments on cereals in Poland, plots of 15m^2 (1.5m by 10m) are in use till now. So since then a new designs with bigger blocks were introduced and are used till now. The typical 1-resolvable incomplete block design has a form presented in Figure 1. It is a design for V=24 treatments, R=3 replicates (superblocks) with s=3 incomplete blocks of k=8 plots within each superblock. The efficiency factor of this design is E=0.8824.

Replicate (superblock) 1

14	15	8	5	22	23	1	12	3	18	11	13	6	7	2	17	19	21	24	20	4	6	9	10

Replicate (superblock) 2

13	1	19	9	4	23	17	22	21	5	2	7	24	11	8	15	14	3	6	10	18	12	20	16

Replicate (superblock) 3

3	24	15	20	14	7	9	23	1	14	12	8	17	6	16	11	5	21	22	18	13	2	10	19

Fig.1 Example of 1-resolvable incomplete block design with three replicates

All experimental stations belonging to official variety testing in Poland are supplied with catalogues of 1-resolvable designs for number of treatments between 16 and 100 and for different numbers of replicates and blocks sizes. The stations are obliged to make a proper randomisation before application of the design in a field. Randomisation consists of three steps:

a) randomisation of superblocks,
b) randomisation of incomplete blocks within superblocks,
c) randomisation of plots within incomplete blocks.

3. The data

All the conclusion presented here are based on experimental data consisting of 174 trials conducted (year of harvesting) in the years 1998 and 1999 on cereals. Among them there are 72 experiments on winter wheat (40 and 32 trials respectively), 63 trials on spring wheat (26 and 37) and 39 trials on rye (22 and 17). From the year 2000 the number of trials was seriously decreased, so the presented set of experiments is the last so extensive one in Poland. In all experiments there were four replicates and the plots of $15m^2$ (1.5m by 10m) for harvesting were applied.

The different parameters of analysed trials are collected in Table 1.

Table 1

Parameters of trials on cereals conducted in Poland in 1998-1999

Species/year	Number of varieties	Block sizes applied	Number of cases
winter wheat/1998	48	8	7
	38	8	16
	32	8	17
winter wheat/1999	42	6	16
	39	8	10
	27	7	6
spring wheat /1998	36	6	5
	28	7	20
	28	5	1
spring wheat/1999	29	8	16
	17	6	21
winter rye/1998	25	5	1
	22	8	3
	20	4	18
winter rye/99	21	7	13
	20	4	4

3. Efficiency of the designs

For all trials considered here their efficiency was calculated for two cases. For so-called intra-block analysis (without recovery of inter-block information) and for inter-block analysis (with recovery of this information). The intra-block efficiency E_i was calculated according to the formula

$$E_i = MS_e(V)^*E/MS_e(k)$$

where $MS_e(V)$ stands for mean square for error in randomised complete block analysis of trial result (ignoring subdivision of complete superblocks into incomplete blocks), E – for the mean harmonic efficiency factor of the design and $MS_e(k)$ – for mean square for error for incomplete block analysis. Similarly, the inter-block efficiency E_b was calculated using the formula given by Patterson and Hunter [1983] of the form

$$E_b = \gamma \left\{E + \frac{(1-E)(s-1)}{\gamma(V-1)-(V-s)}\right\}$$

where $\gamma = MS_e(V)/MS_e(k)$, and s denotes the number of incomplete blocks within superblock. The averages of these efficiencies for all considered here species and years are given in Table 2. For comparison, the results reported by Pilarczyk [1990] for similar series of experiment conducted in 1983, 1985 and 1987 on winter wheat are also given. It

Table 2

Average efficiencies of α-designs in some series of experiments on cereals

Species/year	Number of trials	Number of varieties	Prevailing block size	E_i	E_b	E_i^*
Winter wheat/1998	40	see Table 1	8	1.121	1.190	1.165
Winter wheat/1999	32	see Table 1	7	1.111	1.186	1.149
--------------------	-----	---------	-------	------	------	------
Spring wheat/1998	26	see Table 1	7	1.299	1.368	1.331
Spring wheat/1999	37	see Table 1	7	1.206	1.110	1.164
--------------------	-----	---------	-------	------	------	------
Winter rye/1998	22	see Table 1	4	0.972	1.138	1.069
Winter rye/1999	17	see Table 1	7	1.010	1.234	1.093
Winter wheat/1983	59	33	5	1.24	-	1.51
Winter wheat/1985	71	38	6	1.10	-	1.25
Winter wheat/1987	68	43	7	1.07	-	1.23

is worth of noting, that in the years 1983 and 1985 the plots of $25m^2$ were used.

Additionally, the value of E_i^* was also calculated for every series. It is again the mean value of intra-block efficiencies but with replacement of all values smaller than 1 by 1 (efficiency of randomised complete block design). It is justified by the fact that often, in the case of low efficiency of incomplete block analysis of particular trial, the analysis is performed ignoring incomplete blocks, what results in analysis as randomised complete block data.

4. Extent of spatial variability

There are many papers in which the spatial dependence between results in variety trials is assumed and investigated, Brewer and Mead [1986], Grondona and Cressie [1991], Kristensen and Ersbøl [1992], Watson [2000], Zimmerman and Harville [1991]. Very often the spherical function is assumed to model semivariograms in variety trials. Its features are illustrated in Figure 2. In particular, the nugget effect (c_0) is a measure of non-spatial variation, the range (a) is the distant beyond which the variance takes its maximum values equal to sill (c_0+c). The partial sill (c) is a measure of strength of spatial dependence. To check if there was spatial dependence in analysed trial data, the following procedure was applied:

1^0 The semivariances were calculated basing on residuals both from complete block analysis and from incomplete block analysis. Let e_{ij} means residual for i-th successive plot (i=1,2,...V) within j-th superblock (j=1,2,...,r). The semivariance at lag h, named S(h), was calculated as

$$S(h)=\frac{1}{2N(h)}\sum_{j=1}^{r}\sum_{i=1}^{V-h}\left(e_{ij}-e_{i+h,j}\right)^2 \quad ,$$

where N(h) means number of cases that differences ($e_{ij} - e_{i+h,j}$) exist. It is equivalent with calculation of semivariances within particular superblocks and then averaging them. The semivariances were calculated up to lag equal half of the number of treatments.

2^0 The spherical function was used of the form

$$\gamma(h,\theta)=\begin{cases} 0 & h=0 \\ c_0+c\left(\dfrac{3h}{2a}-\dfrac{h^3}{2a^3}\right) & 0<h\le a \\ c_0+c & a<h \end{cases}$$

to model the underlying semivariograms. The weighted least squares method as described by Watson [2001] was used for fitting parameters.

Spatial dependence was assumed if semivariances were satisfactorily fitted by spherical function curve, i.e. the iterative procedure converged and the estimates were feasible, what means that values of c_0 and c were positive and range parameter a>1.

3^0 The number of cases was counted for which the spatial dependence was detected after randomised complete block (RCB) analysis of variance and after incomplete block (IB) analysis.

The described procedure was applied to the results of all trials on winter wheat conducted in 1998 and 1999. These trials were chosen because of high number of treatments compared, so also the number of tested lags equal to half of it was relatively high.

The results are collected in Table 3. There is also an information given about the goodness of fit measure as a percent of weighted sum of squares explained by the fitted curve in all the trials in which the spatial dependence was detected. As one can see, the spatial dependence was detected in about 50% of all trials.

All presented data concern estimation of spherical function made using residuals after RCB analysis of variance as there were only two trials (in a year 1998) in which the spatial dependence was detected also when residuals from IB analysis of variance were used. In these two trials the goodness of fit was respectively 87.3% and 26.9% for the first trial and 95.4% and 20.2% for the second one, for the residuals from RCB and IB analysis. So in these two trials incomplete blocks seriously reduced spatial variability but part of it still was unexplained. The range ‘a’ was reduced from 12 to 5 in the first trial and from 15 to 5 for the second one.

Table 3

Summary characterisation of spatial dependence in wheat trials

Year	Number of varieties V	Number of trials		Number of trials with goodness of fit		Range a	
		all	with spatial dependence detected	≤50%	>50%	Min	Max
1998	48	7	5	2	3	10	32
	38	16	8	4	4	7	42
	32	17	6(2)	2	4	6	30
total		40	19(2)	8	11	6	42
1999	42	16	9	4	5	6	35
	39	10	7	4	3	8	48
	27	6	1	0	1	14	14
Total		32	17	8	9	6	48

Note that (2) means that there were two trials with spatial dependence detected after incomplete block analysis of data.

5. Conclusions

The performed analysis of large data set from trials on cereals conducted for variety testing purposes in years 1998 and 1999 allows to conclude that:

a) advantage of incomplete blocks with blocks of about 7-8 plots over complete ones was again confirmed in trials on cereals with higher than 20 number of varieties;
b) after complete block analysis the spatial dependence was detected in approximately half of all analysed trials on winter wheat;
c) incomplete blocks shoved their effectiveness in removing spatial dependence in variety testing field trials;
d) there were only two (out of 72) trials where some trace spatial dependence was detected after incomplete block analysis;
e) there was no difference observed between extent of spatial dependence between years.

Literature

Brewer A.C., Mead R., 1986, Continuous second order models of spatial variation with application to the efficiency of field crop experiments, Journal of the Royal Statistical Society, Series A 149, pp.314-337.

Grondona M.O., Cressie N., 1991, Using spatial considerations in the analysis of experiments, Technometrics 33, 381-392.

Kristensen K., Ersbøl A.K., 1992, The use of geostatistical methods in planning variety trials, Biuletyn Oceny Odmian (Cultivar Testing Bulletin) 24-25, pp. 139-157.

Patterson H.D., Hunter E.A., 1983, The efficiency of incomplete block designs in National List and Recommended List cereal Variety trials. J. Agric. Sci. Camb. 103, pp. 427-433.

Patterson H.D., Williams E.R., 1976, A new class of resolvable incomplete block designs. Biometrika 63, pp. 83-92.

Pilarczyk W., 1990, Precision on field trials in incomplete block designs for several species, Cultivar Testing Bulletin XVII-XVIII, pp.161-169.

Watson S., 2000, Spatial dependence and block designs in spaced plant herbage trials, Journal of Agric. Sci. Cambridge, 134, pp.245-258.

Williams E.R., 1975, A new class of resolvable block designs. Ph. D. Thesis, University of Edinburgh.

Zimmerman D.L., Harville D.A., 1991, A random field approach to the analysis of field-plot experiments and other spatial experiments, Biometrics 47, pp.223-239.