

# A comparison of experimental designs for selection in breeding trials with nested treatment structure

H. P. Piepho · E. R. Williams

Received: 16 March 2006 / Accepted: 16 August 2006 / Published online: 7 October 2006  
© Springer-Verlag 2006

**Abstract** Plant breeders frequently evaluate large numbers of entries in field trials for selection. Generally, the tested entries are related by pedigree. The simplest case is a nested treatment structure, where entries fall into groups or families such that entries within groups are more closely related than between groups. We found that some plant breeders prefer to plant close relatives next to each other in the field. This contrasts with common experimental designs such as the  $\alpha$ -design, where entries are fully randomized. A third design option is to randomize in such a way that entries of the same group are separated as much as possible. The present paper compares these design options by simulation. Another important consideration is the type of model used for analysis. Most of the common experimental designs were optimized assuming that the model used for analysis has fixed treatment effects. With many entries that are related by pedigree, analysis based on a model with random treatment effects becomes a competitive alternative. In simulations, we therefore study the properties of best linear unbiased predictions (BLUP) of genetic effects based on a nested treatment structure under these design options

for a range of genetic parameters. It is concluded that BLUP provides efficient estimates of genetic effects and that resolvable incomplete block designs such as the  $\alpha$ -design with restricted or unrestricted randomization can be recommended.

## Introduction

Plant breeders frequently conduct field trials where the objective is to select the best entries. After the entries to be compared have been identified, the breeder will seek the best design and method of analysis. Both will depend on the genetic architecture of the population under selection. Entries often have a nested treatment structure arising from genetic relationships. Here, we consider two-stage nested structures, where entries fall into groups (Lush 1947; Falconer and Mackay 1996; Bueno and Gilmour 2003). For example, entries may correspond to F3-derived F4 families. We here consider the case where selection is based on the entire set of entries in the experiment, as opposed to selection within groups of entries. The main issues are how to accommodate the nested treatment structure in the choice of experimental design, whether to model the treatments as fixed or random effects in the analysis, and how these issues relate to each other.

Most plant breeding trials involve a large number of entries, so some kind of incomplete blocking is usually worthwhile, e.g., a resolvable incomplete block design such as an  $\alpha$ -design (Williams et al. 2002). The optimality of these  $\alpha$ -designs has been established assuming fixed treatment effects. However, it is not clear that  $\alpha$ -designs are best when the analysis will treat the treatments as correlated random effects. Bueno and

---

Communicated by M. Cooper.

---

H. P. Piepho (✉)  
Fachgebiet Bioinformatik, Institut für Pflanzenbau  
und Grünland, Universität Hohenheim,  
70593 Stuttgart, Germany  
e-mail: piepho@uni-hohenheim.de

E. R. Williams  
Statistical Consulting Unit, The Graduate School,  
c/- MSI John Dedman Building,  
The Australian National University,  
Canberra, ACT 0200, Australia

Gilmour (2003) considered the design problem, when treatment effects are regarded as random. They compared various genetic correlation structures arising from a pedigree, including the nested structure considered here, and concluded that the optimal design is not independent of the genetic correlation structure and the heritability.

Bueno and Gilmour (2003) used the mean variance of a difference between treatments as an optimality criterion on the grounds that the treatment ranking can be obtained from these comparisons. The genetic variances and the heritability were assumed known, whereas in practice, variance components need to be estimated, and the resulting estimators for random effects, known as empirical BLUP (best linear unbiased prediction), do not perform as efficiently as BLUP with known variance components. It is therefore useful to study the small-sample behaviour of different models and estimation methods by simulation. In doing so, one may directly relate the ranking of effect estimates compared to the ranking of true effects. This is more readily interpretable for a plant breeding context than the average variance of a difference, though both criteria are closely related.

In light of Bueno and Gilmour (2003), it is not obvious whether a resolvable incomplete block design such as an  $\alpha$ -design is optimal or at least competitive under all analysis scenarios, including BLUP of treatment effects for a nested pedigree, though it can be hoped to be close to the optimum. If genetic variance and covariance parameters are known beforehand, one can find optimal designs for a specific pedigree and analysis method (Bueno and Gilmour 2003). This approach may not be within reach for routine applications, though, where genetic variance parameters are usually unknown before the trial, time available for planning is limited, and breeders prefer to rely on robust and easy-to-use procedures. Efficient incomplete block designs are conveniently constructed using software such as CycDesign (Whitaker et al. 2006), and so it is reasonable to prefer such designs in practice. In the present paper, we compare the performance of  $\alpha$ -designs, the most common type of resolvable incomplete block design for large numbers of treatments, to some simple commonly used alternatives.

One common alternative is to plant close relatives together in blocks, which is common practice. This type of field layout is akin to a split-plot design with groups of relatives allocated to main plots although groups may differ in size. Another alternative investigated here was suggested by Williams and John (1999) in a tree breeding context. The genetic context is a little different with trees grouped in families and families

nested in provenances, and selection is generally focussed on families within provenances. For this nested structure, Williams and John (1999) proposed a restricted randomization such that the number of different groups (families) per incomplete block is maximized. Software packages such as CycDesign (Whitaker et al. 2006) conveniently generate such designs, even when group sizes are not equal. This restricted randomization maximizes the power for detecting between group differences and so is considered here for the nested pedigree case.

The present paper uses simulation to compare these two design alternatives to an unrestricted randomization for  $\alpha$ -designs. Our assumption throughout is that entries are a random two-stage sample from a hypothetical population of entries, with families sampled in the first stage and entries within families sampled in the second stage. This justifies the assumption of random entry effects. The fact that entries are a random sample does not necessarily imply, however, that best linear unbiased prediction (BLUP) is the best method of analysis. For example, when variance components are poorly estimated, it may be worthwhile to take entries as fixed and use best linear unbiased estimation (BLUE). We therefore compare BLUE and BLUP in simulations. Different designs and estimation methods are evaluated using the rank correlation of estimated and true treatment effects.

No claim is made here that any of the studied designs is optimal under a random treatment model. While optimal designs could be devised for such scenarios (Bueno and Gilmour 2003), routine application is difficult due to unknown variance components and the need to consider several traits simultaneously. Our purpose is therefore to identify general-purpose designs for nested pedigrees that perform well over a wider range of genetic correlation structures.

## Materials and methods

### Simulation of treatment effects

Throughout this paper, the treatment effect is defined as the sum of a group (family) effect and an entry within group effect. This gives rise to a genetic variance of treatments of  $V_t = V_g + V_w$ , a genetic covariance between treatments in the same family of  $V_g$ , in different families of zero. In simulations, treatment effects were generated according to the nested structure  $\text{group/entry} = \text{group} + \text{group} \cdot \text{entry}$  based on this simple variance component model, where groups and entries were coded as `group` and `entry`, respectively

(Piepho et al. 2003). The random group effect implies a genetic correlation  $\rho_g = V_g/V_t$  among treatment effects in the same group.

#### Simulation of block effects

Block effects were simulated as random with variance  $V_b$  and the residual variance was set to one, inducing an environmental correlation  $\rho_e = V_b/(1 + V_b)$  between plots in the same block. In a given simulated data set, block effects and treatment (entry + group) effects were generated as described previously. The treatments were then randomized to plots according to each of the experimental designs, using the same set of treatment and block effects. Simulations were performed using the SAS System.

#### Variance parameters

We define the ratio  $R = V_t/(V_t + V_e)$  where  $V_t$  is the variance of treatment effects defined above and  $V_e = V_b + 1$  is the variance of block plus residual error effects. It should be stressed that  $R$  is not identical to a heritability, because the denominator involves the block variance. We prefer  $R$  to the heritability for the purpose of this paper due to the symmetry in the definition of  $V_t$  and  $V_e$ . In simulations, each of the parameters  $\rho_e$ ,  $\rho_g$ , and  $R$  took values 0.2, 0.5, and 0.8. This covered a relatively wide range of settings. Note that each of the three design parameters can take on values between 0 and 1. The choice of values for  $\rho_g$  was deliberate, while for  $\rho_e$  and  $R$ , these values reflected estimates for parameters found in analyses of a wide range of trials by German plant breeders (results not shown). Each combination of the values for  $\rho_e$ ,  $\rho_g$ , and  $R$  was assessed with 1,000 simulations.

#### Experimental designs

In simulations, we considered designs with incomplete blocks of sizes three, five, eight, and ten. The designs were (1) the  $\alpha$ -design with unrestricted randomization, (2) the  $\alpha$ -design with randomization restricted so that the number of groups per block is maximized, (3) a split-plot design with groups randomized to main-plots (equivalent to incomplete blocks). We simulated groups such that the size was constant, equalling the block size. In addition to the simulation of equal group sizes, we also simulated experiments with unequal group sizes. These had two large groups comprising half the treatments with the remaining treatments in groups, which were the same size as the block size. The simulation of unequal groups did not cover the split-

plot design. The number of treatments was 120 in all simulated designs. The simulated designs had two, three or four complete replicates. CycDesignN (Whitaker et al. 2006) was used to generate  $\alpha$ -designs with and without restricted randomization for groups. When randomization was unrestricted, the design was generated using a single treatment factor and group/entry labels were randomly allocated to the treatment numbers for each simulation run. For restricted randomization, the design was generated using group and entry factors and group/entry labels were randomly allocated to the design group and entry numbers for each simulation run.

The simulation of unequal groups did not cover the split-plot design because this would have necessitated generating block effects with variance depending on block size, thus making block structures different among designs. Different variance values would then have been assumed for the split-plot and  $\alpha$ -designs, thus making the comparison difficult. Alternatively, plot values could have been generated based on a spatial model (Gilmour et al. 1997; Qiao et al. 2000), but this would have involved many choices among alternative models. It should be stressed that in our simulations' field, trends were completely aligned with the split-plot blocks, which represents a worst-case scenario for the split-plot design.

#### Analysis of simulated data

All designs were analysed using the block model `rep/blk`, where `rep` codes complete replicates, while `blk` codes incomplete blocks/main plots within replicates. Taking main plots as random was a necessity for the split-plot design. Block effects in  $\alpha$ -designs were taken as random for recovery of information. Treatment effects of entries, corresponding to the sum of group effect and entry-within-group effect, were modelled as `group + group•entry`. The full model reads

`group + group•entry + rep + rep•blk`.

Treatment effects, given by `group + group•entry`, were estimated via both best linear unbiased estimation (BLUE) by taking treatment effects fixed and via best linear unbiased prediction (BLUP) by taking treatment effects as random. In order to evaluate different designs and methods of estimation, it was assumed that selection is to be based on combined group/entry effects. The objective therefore is to identify estimators which best agree with the ranking of true treatment effects (the sum of the `group` and `group•entry` effects). Thus, Spearman's rank correlation was

computed between true and estimated treatment effects in each simulation run. Rank correlations were averaged across 1,000 simulation runs per setting. Our simulations indicate that the standard error of a difference in the mean rank correlation for different estimators was usually in the order of 0.001. Thus, we reported mean rank correlation up to three decimal places.

## Results

The simulation results are fairly clear-cut, so we only present a limited number of tables. The full set of tables is available from the first author upon request. Results for equal group sizes and two replicates are shown in Tables 1 and 2, while those for unequal group sizes are given in Table 3. In addition, we performed simulations for three and four replicates using group size 5. Results for three replicates are reported in Table 4. The tables show how the rank correlation among true and estimated values depends on genetic correla-

tion ( $\rho_g$ ), environmental correlation ( $\rho_e$ ), and repeatability ( $R$ ). The results show that the split-plot design is always outperformed by the  $\alpha$ -design and restricted randomization for the  $\alpha$ -design leads to little, if any, improvement in terms of rank correlation among estimated and true treatment effects. Also, for the  $\alpha$ -designs, BLUP of treatment effects is better than BLUE throughout. The difference between designs diminishes most notably with increasing repeatability  $R$ . The performance of BLUP and BLUE improves consistently and markedly with increasing environmental correlation ( $\rho_e$ ). Interestingly, for the split-plot design, BLUE outperforms BLUP in most cases.

## Discussion

Plant breeders exercise various selection strategies with pedigreed lines. For example, one may select solely on the performance of individual entries. This is a poor strategy when heritability is low. Information from relatives may be exploited, e.g., by selection

**Table 1** Rank correlation among estimated and true genetic values for 12 groups of size 10 (=block size). Block effects were taken as random for all analyses. Each simulated design had two

replicates. 1,000 simulations were used for each setting. Best rank correlation for a setting is boldfaced

$R$	$\rho_g$	$\rho_e$	BLUE			BLUP			
			Split-plot	$\alpha$ -Design/randomization		Split-plot	$\alpha$ -Design/randomization		
				Unrestricted	Restricted		Unrestricted	Restricted	
0.2	0.2	0.2	0.554	0.577	0.579	0.457	0.581	<b>0.592</b>	
		0.5	0.562	0.658	0.655	0.469	<b>0.673</b>	0.672	
		0.8	0.565	0.805	0.803	0.562	<b>0.815</b>	0.814	
	0.5	0.2	0.538	0.563	0.564	0.470	0.649	<b>0.655</b>	
		0.5	0.553	0.649	0.649	0.411	0.723	<b>0.730</b>	
		0.8	0.567	0.798	0.797	0.461	<b>0.835</b>	0.834	
	0.8	0.2	0.538	0.555	0.557	0.572	<b>0.768</b>	<b>0.768</b>	
		0.5	0.548	0.638	0.638	0.438	0.809	<b>0.812</b>	
		0.8	0.559	0.788	0.788	0.368	0.873	<b>0.876</b>	
	0.5	0.2	0.2	0.797	0.814	0.811	0.784	<b>0.821</b>	0.820
			0.5	0.800	0.863	0.864	0.781	0.868	<b>0.869</b>
			0.8	0.807	0.935	0.935	0.813	<b>0.937</b>	<b>0.937</b>
0.5		0.2	0.795	0.807	0.806	0.798	0.838	<b>0.839</b>	
		0.5	0.795	0.859	0.860	0.771	0.878	<b>0.880</b>	
		0.8	0.801	0.934	0.933	0.778	<b>0.939</b>	0.938	
0.8		0.2	0.801	0.808	0.807	0.778	0.838	<b>0.839</b>	
		0.5	0.801	0.859	0.859	0.778	0.878	<b>0.879</b>	
		0.8	0.801	0.933	0.933	0.778	<b>0.938</b>	<b>0.938</b>	
0.8		0.2	0.2	0.934	0.939	0.939	0.933	<b>0.940</b>	<b>0.940</b>
			0.5	0.935	0.958	0.959	0.934	<b>0.959</b>	<b>0.959</b>
			0.8	0.937	0.982	0.982	0.939	<b>0.982</b>	<b>0.982</b>
	0.5	0.2	0.931	0.937	0.936	0.932	<b>0.941</b>	0.940	
		0.5	0.932	0.956	0.956	0.931	<b>0.958</b>	<b>0.958</b>	
		0.8	0.935	0.981	0.981	0.934	<b>0.981</b>	<b>0.981</b>	
	0.8	0.2	0.925	0.931	0.931	0.934	0.945	<b>0.946</b>	
		0.5	0.926	0.952	0.952	0.927	0.959	<b>0.960</b>	
		0.8	0.928	0.978	0.979	0.926	<b>0.980</b>	<b>0.980</b>	

**Table 2** Rank correlation among estimated and true genetic values for 24 groups of size 5 (=block size). Block effects were taken as random for all analyses. Each simulated design had two

replicates. 1,000 simulations were used for each setting. Best rank correlation for a setting is boldfaced

R	$\rho_g$	$\rho_e$	BLUE			BLUP		
			Split-plot	$\alpha$ -Design/randomization		Split-plot	$\alpha$ -Design/randomization	
				Unrestricted	Restricted		Unrestricted	Restricted
0.2	0.2	0.2	0.552	<b>0.565</b>	0.562	0.484	0.556	0.559
		0.5	0.555	0.627	0.630	0.485	0.639	<b>0.645</b>
		0.8	0.562	0.774	0.774	0.583	<b>0.791</b>	<b>0.791</b>
	0.5	0.2	0.549	0.558	0.559	0.509	0.621	<b>0.622</b>
		0.5	0.555	0.625	0.624	0.467	<b>0.692</b>	0.690
		0.8	0.562	0.770	0.772	0.498	<b>0.814</b>	<b>0.814</b>
	0.8	0.2	0.541	0.556	0.559	0.590	0.722	<b>0.725</b>
		0.5	0.550	0.619	0.621	0.471	0.774	<b>0.776</b>
		0.8	0.558	0.763	0.764	0.438	<b>0.858</b>	<b>0.858</b>
0.5	0.2	0.2	0.800	0.806	0.805	0.794	<b>0.812</b>	0.811
		0.5	0.801	0.850	0.849	0.798	<b>0.857</b>	0.856
		0.8	0.803	0.924	0.923	0.821	<b>0.927</b>	<b>0.927</b>
	0.5	0.2	0.796	0.803	0.804	0.800	<b>0.830</b>	<b>0.830</b>
		0.5	0.800	0.846	0.847	0.791	0.867	<b>0.868</b>
		0.8	0.802	0.922	0.923	0.799	<b>0.930</b>	<b>0.930</b>
	0.8	0.2	0.802	0.806	0.803	0.799	<b>0.832</b>	<b>0.832</b>
		0.5	0.802	0.848	0.847	0.799	<b>0.869</b>	<b>0.869</b>
		0.8	0.802	0.921	0.922	0.799	0.929	<b>0.930</b>
0.8	0.2	0.2	0.934	0.937	0.937	0.934	<b>0.938</b>	<b>0.938</b>
		0.5	0.934	0.953	0.953	0.934	<b>0.954</b>	<b>0.954</b>
		0.8	0.934	0.977	0.977	0.938	<b>0.978</b>	<b>0.978</b>
	0.5	0.2	0.932	0.935	0.935	0.933	<b>0.939</b>	<b>0.939</b>
		0.5	0.933	0.951	0.952	0.932	<b>0.954</b>	<b>0.954</b>
		0.8	0.934	0.977	0.977	0.934	<b>0.978</b>	<b>0.978</b>
	0.8	0.2	0.929	0.933	0.932	0.937	<b>0.946</b>	0.945
		0.5	0.932	0.950	0.950	0.934	<b>0.959</b>	<b>0.959</b>
		0.8	0.932	0.976	0.976	0.931	<b>0.979</b>	0.978

based on family (group) means, also known as family selection (Falconer and Mackay 1996). A further strategy, often used in animal breeding, is within-family selection, which selects individuals based on their deviation from the family mean. A compromise between family selection and within-family selection is to combine both individual means and family means via a selection index. It turns out that the optimal index suggested by Lush (1947) is equivalent to BLUP of individual treatment effects (group + group•entry) based on the nested structure considered in this paper. Thus, selection using BLUP provides an optimal strategy in the sense of Lush (1947). By contrast, selection based on BLUE of individual treatment effects does not exploit information from relatives and is therefore equivalent to individual selection.

In some breeding applications, there may be a nested pedigree, and yet the breeder decides to use the field trial only for selecting within groups (families, crosses, etc.), because selection of groups is based on other considerations, but will be done in parallel with the field trials. In this case, it will be worthwhile to

have a separate trial for each group, because this maximizes the accuracy of within-group comparisons. Clearly, in designing a field trial for evaluation of entries, the guiding principle should be to test in a single trial all entries that need to be directly compared for making a selection decision. If selection is across several groups, then all groups should be tested in a single trial. If selection is to be done only within groups, but not across groups, then it is preferable to perform a separate trial for each group.

It should be emphasized that our conclusions are strictly valid only for the family and experiment sizes we investigated. Simulations conducted for other settings not shown in this paper did yield the same tendency in favour of unrestricted randomization. Also, our simulations assume absence of competition effects, which may not be a realistic scenario when plots are small.

This paper considered different options in designing experiments with nested treatment structure. The simplest option is to ignore nesting altogether and use a standard design such as the  $\alpha$ -design. Alternatively,

**Table 3** Rank correlation among estimated and true genetic values for block size 5. Two groups (families) were of size 30, while the remaining groups were of size 5. Block effects were taken as random for all analyses. Each simulated design had two replicates. 1,000 simulations were used for each setting. Best rank correlation for a setting is boldfaced

R	$\rho_g$	$\rho_e$	BLUE		BLUP	
			Randomization		Randomization	
			Unrestricted	Restricted	Unrestricted	Restricted
0.2	0.2	0.2	0.560	0.561	0.576	<b>0.577</b>
		0.5	0.626	0.624	<b>0.652</b>	0.651
		0.8	0.768	0.767	<b>0.790</b>	0.789
	0.5	0.2	0.550	0.551	0.633	<b>0.635</b>
		0.5	0.612	0.614	0.691	<b>0.697</b>
		0.8	0.761	0.760	0.807	<b>0.809</b>
	0.8	0.2	0.531	0.529	0.721	<b>0.727</b>
		0.5	0.597	0.598	0.769	<b>0.772</b>
		0.8	0.743	0.738	<b>0.842</b>	<b>0.842</b>
0.5	0.2	0.2	0.801	0.802	<b>0.810</b>	<b>0.810</b>
		0.5	0.847	0.847	<b>0.856</b>	<b>0.856</b>
		0.8	0.921	0.921	<b>0.925</b>	<b>0.925</b>
	0.5	0.2	0.793	0.793	0.823	<b>0.824</b>
		0.5	0.836	0.835	0.859	<b>0.860</b>
		0.8	0.915	0.917	0.924	<b>0.925</b>
	0.8	0.2	0.793	0.792	<b>0.823</b>	0.822
		0.5	0.838	0.837	<b>0.862</b>	0.861
		0.8	0.916	0.916	<b>0.925</b>	<b>0.925</b>
0.8	0.2	0.2	0.935	0.935	<b>0.936</b>	<b>0.936</b>
		0.5	0.952	0.952	<b>0.953</b>	<b>0.953</b>
		0.8	<b>0.977</b>	<b>0.977</b>	<b>0.977</b>	<b>0.977</b>
	0.5	0.2	0.930	0.930	<b>0.935</b>	<b>0.935</b>
		0.5	0.948	0.949	0.950	<b>0.952</b>
		0.8	0.975	0.975	<b>0.976</b>	<b>0.976</b>
	0.8	0.2	0.919	0.918	<b>0.933</b>	<b>0.933</b>
		0.5	0.937	0.937	<b>0.947</b>	<b>0.947</b>
		0.8	0.970	0.970	0.972	<b>0.973</b>

randomization may be restricted such that either the power with within-group comparisons is maximized (split-plot design), or the power for between-group comparisons is maximized, as in the design suggested by Williams and John (1999). These authors argued that variation between groups is often perceived to be greater than that within groups, and that in such cases it may be desirable to ensure that groups occur as equally as possible within each block. Our simulations indicate that neither of these two restrictions on the randomization provide any advantage in the settings studied.

In practical plant breeding, one is primarily interested in obtaining good point estimates of treatment effects for making good selection decisions. It is not immediately obvious that maximization of power for between-group comparisons will maximize the response to selection for nested treatment structures such as hierarchical pedigrees. In fact, intuitively it seems equally reasonable to expect that maximization of

power for within-group comparisons is advantageous when differences between groups are relatively large, because large differences between groups should be easier to detect than small differences within groups (Bueno and Gilmour 2003). This reasoning leads to a split-plot design as an obvious choice, and some breeders seem to think along these lines when designing their trials—they prefer to plant close relatives next to each other in the field, as they perceive this design to facilitate visual assessment of diseases and other traits. While a split-plot type design may have some practical merit for the breeder, our simulations have shown that this design performs relatively poorly compared to  $\alpha$ -designs, as far as the ranking of entries is concerned. This suggests that planting close relatives near to each other in the field is not usually an efficient strategy, when selection is to be exercised across groups. There may be a need to use a split-plot design, however, when strong competition is expected among entries from different groups, for example, when groups differ markedly in plant height. This will be particularly useful when plot size is too small to avoid neighbour effects on plants that will be used for measuring traits of interest.

It must be stressed that group sizes will usually be heterogeneous, so use of a split-plot design would not be straightforward mainly at the analysis stage. In particular, variance within main plots is expected to increase with the number of subplots. The unequal sizes of main plots could be accommodated, e.g., by modelling the variance within main plots as a log-linear regression on the number of plots per main plot (Littell et al. 1996, Sect. 8.4). Analysing a split-plot design by spatial methods is another option that could be contemplated, but there is a danger that group effects are removed by the spatial error component, thus leading to biased estimates of treatment effects.

Simulations did not reveal substantial gains, if any, from restricted randomization as suggested by Williams and John (1999). The danger that group and block effects are confounded, causing an inflated block variance estimate, seems low when the number of entries is large. The REML method bases variance estimation on error contrasts, thus making variance component estimates independent of fixed effects (Patterson and Thompson 1971). Therefore, treatment effects will not cause a bias in the block variance estimate, when treatment effects are taken as fixed. The situation is less clear with random treatment effects. With groups of equal size, the worst case scenario would occur when blocks are completely confounded with groups, i.e., when (as a result of randomization) complete groups would be allocated to blocks. This allocation is exactly

**Table 4** Rank correlation among estimated and true genetic values for group size 5 (=block size) and 24 groups (families). Block effects were taken as random for all analyses. Each

simulated design had three replicates. 1,000 simulations were used for each setting. Best rank correlation for a setting is boldfaced

R	$\rho_g$	$\rho_e$	BLUE			BLUP		
			Split-plot	$\alpha$ -Design/randomization		Split-plot	$\alpha$ -design/randomization	
				Unrestricted	Restricted		Unrestricted	Restricted
0.2	0.2	0.2	0.630	0.647	0.648	0.600	0.652	<b>0.653</b>
		0.5	0.634	0.713	0.713	0.615	<b>0.722</b>	<b>0.722</b>
		0.8	0.639	0.841	0.843	0.678	0.846	<b>0.849</b>
	0.5	0.2	0.626	0.642	0.642	0.622	0.695	<b>0.698</b>
		0.5	0.634	0.713	0.715	0.587	0.758	<b>0.763</b>
		0.8	0.633	0.840	0.837	0.607	<b>0.861</b>	0.860
	0.8	0.2	0.625	0.640	0.640	0.699	<b>0.781</b>	0.781
		0.5	0.623	0.708	0.707	0.599	<b>0.824</b>	0.823
		0.8	0.635	0.835	0.836	0.581	0.891	<b>0.892</b>
0.5	0.2	0.2	0.851	0.860	0.860	0.849	<b>0.863</b>	<b>0.863</b>
		0.5	0.852	0.896	0.896	0.854	0.898	<b>0.899</b>
		0.8	0.854	0.950	0.951	0.871	<b>0.951</b>	<b>0.951</b>
	0.5	0.2	0.850	0.858	0.859	0.854	0.873	<b>0.875</b>
		0.5	0.851	0.894	0.895	0.848	0.904	<b>0.905</b>
		0.8	0.852	0.949	0.950	0.855	<b>0.952</b>	<b>0.952</b>
	0.8	0.2	0.852	0.858	0.857	0.855	0.872	<b>0.873</b>
		0.5	0.852	0.894	0.894	0.855	<b>0.904</b>	<b>0.904</b>
		0.8	0.852	0.950	0.949	0.855	<b>0.952</b>	<b>0.952</b>
0.8	0.2	0.2	0.954	0.957	0.956	0.954	<b>0.957</b>	<b>0.957</b>
		0.5	0.954	0.969	0.969	0.955	<b>0.969</b>	<b>0.969</b>
		0.8	0.955	0.986	0.986	0.957	<b>0.986</b>	<b>0.986</b>
	0.5	0.2	0.953	0.956	0.956	0.953	<b>0.958</b>	<b>0.958</b>
		0.5	0.954	0.969	0.968	0.953	<b>0.970</b>	0.969
		0.8	0.954	0.986	0.986	0.954	<b>0.986</b>	<b>0.986</b>
	0.8	0.2	0.951	0.954	0.953	0.955	<b>0.961</b>	0.960
		0.5	0.952	0.966	0.967	0.953	<b>0.970</b>	<b>0.970</b>
		0.8	0.952	0.985	0.985	0.951	<b>0.986</b>	0.985

that of a split-plot design. Here, the error term for the comparison of groups would involve the block variance, while within-group comparisons would be based entirely on within-block comparisons. In this sense, our split-plot scenario considered in simulations corresponds to the extreme case envisioned by Williams and John (1999). The relatively poor performance of the split-plot design confirms the assertion by Williams and John (1999) that it is preferable to avoid blocks that are entirely made up of only a single group. The chances of obtaining a randomization that is equal or close to that of a split-plot design are usually so small, however, that the performance of unrestricted versus restricted randomization of an  $\alpha$ -design is very similar. Nevertheless, one may prefer the restricted randomization as an insurance against a particularly unfavourable randomization.

In this paper, we have focussed on  $\alpha$ -designs, which are a special class of resolvable incomplete block designs. The main advantage of  $\alpha$ -designs is their speed of construction for larger treatment numbers and most of the time they will be optimal or near-optimal when

treatment effects are fixed. For smaller treatment numbers, in particular, it is possible to improve on the  $\alpha$ -designs, albeit only slightly in most cases. For example, some of the square and rectangular lattices (known to be optimal for their parameter specification) are not  $\alpha$ -designs (Patterson and Williams 1976). Cyc-DesigN gives users the option to construct designs using the alpha option or the more general (and slower) unrestricted option. As a general recommendation, one would use the unrestricted option for small treatment numbers and alpha for bigger designs. In a plant breeding context, it is often preferable to concentrate on  $\alpha$ -designs, because the number of treatments is large.

We here have concentrated on designs with one-way blocking structures and found  $\alpha$ -designs to perform well compared to alternative designs. It is stressed, however, that for field layouts involving a rectangular grid of plots, row-column designs can be more efficient than one-dimensional blocking such as with  $\alpha$ -designs or more generally resolvable incomplete block designs. The advantages of two-dimensional over one-dimen-

sional blocking can be considerable, especially when plots are not long and thin.

This paper has considered BLUP and BLUE as alternative methods of estimation when entries can be regarded as a random sample. Use of BLUE implies falling back to a fixed effects assumption. It should be stressed that the random assumption is the stronger one. Switching to fixed effects when the assumption of a random sample is tenable, can, in fact, be regarded as a relaxation of statistical assumptions. From a Bayesian perspective, use of BLUE implies use of a noninformative prior for genetic effects. Clearly, both BLUE and BLUP are valid methods of analysis, when entries can be regarded as random. In particular, randomness of treatment effects does not invalidate BLUE. To further support this point, consider the example of incomplete blocks. Proper randomization justifies the random assumption for blocks, and yet, one may prefer to take blocks as fixed during analysis, when inter-block information is low or when weights for optimal combination of intra-block and inter-block information cannot be accurately estimated. In fact, the intra-block analysis is a limiting case of the mixed-model analysis (i.e., with recovery of inter-block information) when the block variance component is large. Hence, algebraically, the BLUP of blocks tends towards the BLUE as the block variance increases.

Our results show that BLUP of treatment effects almost invariably outperforms BLUE, often markedly so. The most critical issue is the estimation of genetic variance components. If all treatment variance components can be estimated on “enough” degrees of freedom, BLUP is the method of choice for estimating treatment effects. This finding is in agreement with a large number of empirical studies in animal and plant breeding (Panter and Allan 1995; Piepho and Möhring 2006). When genetic variances cannot be reliably estimated for a given trial, it may be worthwhile to use variance component data from similar past trials. The advantage of BLUP does not seem to have gained much attention, however, for the particular type of breeding scenario considered in this paper (but see Cervantes-Martinez et al. 2001 for an application), although a nested random model is often used for estimation of genetic parameters in plant breeding trials (Yang 2002; Soh et al. 2003; Persson and Andersson 2004; Wu et al. 2006).

In the present paper, we have evaluated designs, which have been optimized under a fixed treatment effects model. At the same time, simulations revealed that BLUP of treatment effects is preferable to BLUE. This raises the question whether the design can be further improved under a random treatment effects

model. The results of Bueno and Gilmour (2003) suggest that there is likely to be some scope for improvement. Those authors concluded that two designs which are both optimal for unrelated treatments are not necessarily equally good for some treatment relatedness. For the special case of a nested treatment structure (half-sibs or full-sibs) and low-to-medium heritability, the optimal design is in the class of optimal designs for unrelated treatments. Thus, it is quite possible that an optimal design for BLUP with a nested treatment structure can be found among the class of resolvable incomplete block designs.

Pedigrees giving rise to a two-stage nested treatment design are but one possible scenario in plant breeding, albeit a rather common one. Alternative designs include various crossed designs, including diallels and factorial crosses. Typically, many different crosses are made, and the resulting pedigree may be quite complex overall. In principle, one could attempt to find optimal designs for each specific setting, though the effort would be quite considerable. Also, genetic variance components are typically not accurately known, and they differ among traits, so finding a universally optimal design that capitalizes on genetic correlation structure is difficult. It seems desirable to have general-purpose designs that can be expected to yield good results over a range of different settings. Our simulations suggest that the  $\alpha$ -design is fairly robust to different pedigree structures and analysis methods (BLUP vs. BLUE), when the pedigree has a simple two-stage structure. Also, unequal group sizes can be accommodated with no difficulty because grouping is ignored at the design stage. Thus, we recommend this design for nested treatment structures in plant breeding. Additional simulations with other pedigree structures, including crossed designs, would be useful.

**Acknowledgments** The first author was supported within the breeding and informatics (BRAIN) project of the genome analysis of the plant biological system (GABI) initiative ([www.brain.de](http://www.brain.de)). We thank all breeders of GABI-BRAIN who have provided data and information on their field trial designs. Three referees are thanked for exceptionally constructive and detailed comments.

## References

- Bueno JSD, Gilmour SG (2003) Planning incomplete block experiments when treatments are genetically related. *Biometrics* 59:375–381
- Cervantes-Martinez CT, Frey KJ, White PJ, Wesenberg DM, Holland JB (2001) Selection for greater  $\beta$ -glucan content in oat grain. *Crop Sci* 41:1085–1091
- Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. 4th edn. Prentice Hall, Harlow



- Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneous variation in the analysis of field experiments. *J Agric Biol Environ Stat* 2:269–293
- Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996) SAS System for Mixed Models. SAS Institute, Cary
- Lush JL (1947) Family merit and individual merit as bases for selection. *Am. Nat.* 81:241–261 and 362–379
- Panter DM, Allen FL (1995) Using best linear unbiased predictions to enhance breeding for yield in soybean. 2. Selection of superior crosses from a limited number of yield trials. *Crop Sci* 35:405–410
- Patterson HD, Williams ER (1976) A new class of resolvable incomplete block designs. *Biometrika* 63:83–92
- Patterson HD, Thompson (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika* 58:545–554
- Persson BT, Andersson B (2004) Accuracy of single- and multi-trait REML evaluation of data including non-random missing records. *Silvae Genetica* 53:135–139
- Piepho HP, Büchse A, Emrich K (2003) A hitchhiker's guide to the mixed model analysis of randomized experiments. *J Agron Crop Sci* 189:310–322
- Piepho HP, Möhring J (2006) Selection in cultivar trials—is it ignorable? *Crop Sci* 146:193–202
- Qiao CG, Basford KE, DeLacy IH, Cooper M (2000) Evaluation of experimental designs and spatial analyses in wheat breeding trials. *Theor Appl Genet* 100:9–16
- Soh AC, Gan HH, Wong G, Hor TY, Tan CC (2003) Estimates of within family genetic variability for clonal selection in oil palm. *Euphytica* 133:147–163
- Whitaker D, Williams ER, John JA (2006) CycDesignN: a package for computer generation of experimental designs. <http://www.cycdesign.co.nz>
- Williams ER, John JA (1999) Construction of resolvable designs with nested treatment structure. *Biom J* 41:341–349
- Williams ER, Matheson AC, Harwood CE (2002) Experimental design and analysis for tree improvement, 2nd edn. CSIRO, Canberra
- Wu J, Jenkins JN, McCarty JC, Wu D (2006) Variance component estimation using the additive, dominance, and additive x additive model when genotypes vary across environments. *Crop Sci* 46:174–179
- Yang RC (2002) Likelihood-based analysis of genotype-environment interactions. *Crop Sci* 42:1434–1440