1-1-2006

On the design of early generation variety trials with correlated data

Brian R. Cullis
University of Wollongong, bcullis@uow.edu.au

A B. Smith
Wagga Wagga Agricultural Institute, alismith@uow.edu.au

N. E. Coombes
Wagga Wagga Agricultural Institute

On the design of early generation variety trials with correlated data

Abstract

This article considers the design of early generation variety trials with a prespecified spatial correlation structure and introduces a new class of partially replicated designs called p-rep designs in which the plots of standard varieties are replaced by additional plots of test lines. We show how efficient p-rep designs can be readily generated using the modified Reactive TABU search algorithm. The expected and realized genetic gain of p-rep and grid plot designs is compared in a simulation study.

Keywords

early, generation, design, variety, data, trials, correlated

Disciplines

Physical Sciences and Mathematics

Publication Details

Cullis, B. R., Smith, A. B.. & Coombes, N. E.. (2006). On the design of early generation variety trials with correlated data. Journal of Agricultural, Biological, and Environmental Statistics, 11 (4), 381-393.

On the Design of Early Generation Variety Trials With Correlated Data

B. R. CULLIS, A. B. SMITH, and N. E. COOMBES

This article considers the design of early generation variety trials with a prespecified spatial correlation structure and introduces a new class of partially replicated designs called p-rep designs in which the plots of standard varieties are replaced by additional plots of test lines. We show how efficient p-rep designs can be readily generated using the modified Reactive TABU search algorithm. The expected and realized genetic gain of p-rep and grid plot designs is compared in a simulation study.

Key Words: Genetic gain; Grid plot design; REML; Spatial correlation.

1. INTRODUCTION

Early generation variety trials (EGVTs) are an integral part of all plant improvement programs. These trials present the first opportunity for breeders to undertake selection for key quantitative traits such as grain yield. In an excellent review article Kempton (1984) suggested that the usual paradigms which underpin experimental design such as estimation of treatment comparisons with minimum error and provision of a valid estimate of that error, may not be relevant to EGVTs. The aim of EGVTs is to maximize the genetic gain from selection of a subset of superior breeding lines (hereafter referred to as test lines). Selection is usually undertaken with respect to a range of traits. In this article we focus on the key trait of grain yield.

In EGVTs there may be insufficient seed to replicate all test lines and so the most widely adopted designs are so-called grid-plot designs. These designs are formed by interposing a (regular) grid of plots containing a standard (or several standard) variety(ies) among plots of unreplicated test lines. Historically, in grid-plot designs local control of heterogeneity was achieved by subtraction of a "fertility index" based on the yields of the standard varieties. Cullis, Lill, Fisher, and Read (1989) proposed a spatial analysis approach for EGVTs which is readily extended to two dimensions or modeling extraneous variation (Gilmour, Cullis, and Verbyla 1997).

B. R. Cullis is Research Leader, Biometrics, NSWDPI Wagga Wagga Agricultural Institute, NSW 2650, Australia (E-mail: *brian.cullis@dpi.nsw.gov.au*). A. B. Smith is Senior Research Scientist, NSWDPI Wagga Wagga Agricultural Institute, NSW 2650, Australia. N. E. Coombes is Senior Biometrician, NSWDPI Wagga Wagga Agricultural Institute, NSW 2650, Australia.

Journal of Agricultural, Biological, and Environmental Statistics, Volume 11, Number 4, Pages 381–393
DOI: 10.1198/108571106X154443

Adoption of spatial methods for the analysis of variety trials has led to interest in the design of experiments when the data are correlated (see, e.g., Martin 1996). Limited theoretical results are available for replicated designs with smaller numbers of treatments, though these are not applicable to EGVTs. This has led to an algorithmic approach for the construction of efficient designs for these trials (Coombes 2002; Chan 1999; Martin and Eccleston 1997) that requires numerical optimization of an objective function. Most approaches use the A-optimality criterion as the objective function. Chan (1999); Martin and Eccleston (1997) used simulated annealing (Kirkpatrick, Gelatt, and Vecchi 1983) to obtain optimal or near optimal designs for a range of correlation models. Coombes (2002) used a modified TABU (Glover 1989, 1990) search algorithm and found it often produced better designs than simulated annealing, but required more processing time. Chauhan (2000) and Chan (1999) considered the design of unreplicated trials with a prespecified spatial correlation structure. They suggested that the overall A-optimality which minimizes the average variance of all elementary contrasts between all treatments (i.e., test lines and standard varieties) may not be the most relevant criterion for EGVTs.

This article considers the design of early generation variety trials with a prespecified correlation structure that reflects the spatial analysis model of Gilmour et al. (1997). In developing these designs we consider a new class of designs in which the plots of standard varieties are replaced by additional plots of test lines, whenever resources allow. It has long been recognized that the use of systematically located check plots is not as efficient as the use of an appropriate incomplete block design (Atiqullah and Cox 1962). This principle can be applied to the design of EGVTs, in that the use of additional plots of test lines in place of plots of standard varieties should result in a greater response to selection.

The structure of the article is as follows. Section 2 describes methods of analysis for data from early generation field trials. Section 3 discusses issues associated with the generation of trial designs, including the choice of optimality criteria. Section 4 introduces the new designs. The efficiency of these designs compared with grid-plot designs is examined via a simulation study, the methodology and results of which are presented and discussed in Section 5. Finally, some concluding remarks are given in Section 6.

2. MIXED MODEL FOR EARLY GENERATION VARIETY TRIALS

We consider the general framework proposed by Gilmour et al. (1997) and assume that the trial has a two-dimensional layout indexed by field rows $(1, \ldots, r)$ and columns $(1, \ldots, c)$ so that the total number of observations is $n = rc$. Our model for the $n \times 1$ vector of yields, $\mathbf{y}$ (assumed ordered as rows within columns) is given by

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}_v\mathbf{u}_v + \mathbf{Z}_b\mathbf{u}_b + \mathbf{e} \\ &= \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}_s\mathbf{u}_s + \mathbf{Z}_t\mathbf{u}_t + \mathbf{Z}_b\mathbf{u}_b + \mathbf{e}, \end{aligned} \tag{2.1}$$

where $\boldsymbol{\tau}$ is the $p \times 1$ vector of fixed effects with associated $n \times p$ design matrix $\mathbf{X}$ (assumed to have full column rank), $\mathbf{u}_v$ is the $v \times 1$ vector of random variety effects with associated

$n \times v$ design matrix $\mathbf{Z}_v$, $\mathbf{u}_b$ is the $b \times 1$ vector of random block effects with associated $n \times b$ design matrix $\mathbf{Z}_b$, and $\mathbf{e}$ is the vector of residual effects. As discussed previously EGVTs usually comprise both test lines and standard varieties so let $\mathbf{u}_v = (\mathbf{u}_s', \mathbf{u}_t')'$, where $\mathbf{u}_s$ and $\mathbf{u}_t$ are vectors of standard and test variety effects, respectively (with length s and t such that $v = s+t$). Similarly let $\mathbf{Z}_v = [\mathbf{Z}_s\ \mathbf{Z}_t]$, where $\mathbf{Z}_s$ and $\mathbf{Z}_t$ are the design matrices for standard and test varieties (with dimension $n \times s$ and $n \times t$). The vector of block effects corresponds to effects associated with experimental design or extraneous sources of variation. Note that the assumption of random (rather than fixed) variety effects is consistent with the aim of EGVTs, namely selection (also see Smith, Cullis, and Thompson 2005).

We assume that the joint distribution of $(\mathbf{u}_v',\ \mathbf{u}_b',\ \mathbf{e}')'$ is Gaussian with zero mean and variance matrix

$$\mathbf{V} = \sigma^2 \begin{bmatrix} \mathbf{G}_v(\boldsymbol{\gamma}_v) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_b(\boldsymbol{\gamma}_b) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}(\boldsymbol{\phi}) \end{bmatrix},$$

where $\boldsymbol{\gamma}_v$, $\boldsymbol{\gamma}_b$, and $\boldsymbol{\phi}$ are vectors of unknown variance parameters and σ^2 is the (unknown) scale parameter. The matrix $\mathbf{G}_v$ is often a scaled identity matrix, that is, $\mathbf{G}_v = \gamma_v \mathbf{I}_v$ or possibly $\mathbf{G}_v = \gamma_v \mathbf{A}$ where $\mathbf{A}$ is a known relationship matrix. At present pedigrees are not generally used in routine analyses of EGVTs.

The matrix $\mathbf{G}_b$ is typically a direct sum of scaled identity matrices, each component corresponding to different terms within $\mathbf{u}_b$.

Here we assume that $\mathbf{R} = \mathbf{R}_c(\boldsymbol{\phi}_c) \otimes \mathbf{R}_r(\boldsymbol{\phi}_r)$, where $\mathbf{R}_r$ and $\mathbf{R}_c$ are correlation matrices for the row and column processes, respectively. A commonly used correlation model is an autoregressive process of order one. A nugget effect may also be included but for simplicity we do not consider this here.

Estimation of the mixed model in (2.1) consists of two linked processes, namely the estimation of fixed and random effects and the estimation of variance parameters. The BLUEs of fixed and BLUPs of random effects are obtained as solutions to the mixed model equations, which, for the model in (2.1) are given by

$$\begin{bmatrix} \mathbf{Z}_v'\mathbf{R}^{-1}\mathbf{Z}_v + \mathbf{G}_v^{-1} & \cdot\cdot & \cdot\cdot \\ \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_v & \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \cdot\cdot \\ \mathbf{Z}_b'\mathbf{R}^{-1}\mathbf{Z}_v & \mathbf{Z}_b'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_b'\mathbf{R}^{-1}\mathbf{Z}_b + \mathbf{G}_b^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{u}}_v \\ \hat{\boldsymbol{\tau}} \\ \tilde{\mathbf{u}}_b \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_v'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_b'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix},$$

where the symbol $\cdot\cdot$ is used to indicate the symmetry of the matrix. Variance parameters are estimated using residual maximum likelihood (Patterson and Thompson 1971), resulting in E-BLUEs and E-BLUPs. Thus, henceforth we use the notation $\hat{\boldsymbol{\tau}}$ to represent the E-BLUE of $\boldsymbol{\tau}$ and, for example, $\tilde{\mathbf{u}}_v$ to represent the E-BLUP of $\mathbf{u}_v$.

Absorption of the equations for $\boldsymbol{\tau}$ and $\mathbf{u}_b$ leads to a reduced set of equations for $\mathbf{u}_v$ given by

$$\left(\mathbf{Z}_v'\mathbf{S}\mathbf{Z}_v + \mathbf{G}_v^{-1}\right)\tilde{\mathbf{u}}_v = \mathbf{Z}_v'\mathbf{S}\mathbf{y}, \tag{2.2}$$

where $\mathbf{S}$ is given by

$$\mathbf{S} = \mathbf{R}^{-1} - \mathbf{R}^{-1}[\mathbf{X}\ \mathbf{Z}_b] \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \cdot\cdot \\ \mathbf{Z}_b'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_b'\mathbf{R}^{-1}\mathbf{Z}_b + \mathbf{G}_b^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}' \\ \mathbf{Z}_b' \end{bmatrix} \mathbf{R}^{-1}.$$

The coefficient matrix in (2.2) plays a key role in design efficiency since the scaled inverse of this matrix is the matrix of (asymptotic) prediction error variances for the variety effects. That is, var $\left(\tilde{\mathbf{u}}_v - \mathbf{u}_v\right) \simeq \sigma^2 \left(\mathbf{Z}_v'\mathbf{S}\mathbf{Z}_v + \mathbf{G}_v^{-1}\right)^{-1}$. This is the random effects analogue of the variance matrix for fixed variety effects. This link will be explored in greater detail in Section 3.1.

3. DESIGN GENERATION

3.1 OPTIMALITY MEASURES

A number of different optimality criteria have been proposed for EGVTs. Several authors have suggested using contrasts between standard varieties and test lines, among standard varieties only, between standard varieties and test lines or between the average of standard varieties and test lines. Motivated by the aim of EGVTs we seek an optimality measure which will maximize the expected genetic gain (EGG). Kempton (1984) also used EGG as the measure for comparing the effectiveness of different methods of analysis of EGVTs. For simplicity we assume that $\mathbf{G}_v = \gamma_v \mathbf{I}_v$. In the case of EGVTs, selection is made on the basis of the E-BLUPs of the test line effects. Thus, expected genetic gain depends upon the distribution of these E-BLUPs. The bias and precision of predictions of random effects when the variance parameters are unknown and must be estimated was discussed by several authors including Kackar and Harville (1981). In the literature on experimental design for correlated data there appears to be no consideration of the effect on design construction of using estimated variance parameters. Thus, authors compute optimality measures assuming that variance parameters are known. We proceed in the same manner but use an optimality measure based on random rather than fixed variety effects. Consequently we assume that the distribution of the E-BLUPs of the test line effects can be approximated by the distribution of the BLUPs so that

$$\tilde{\mathbf{u}}_t \sim \mathcal{N}(\mathbf{0}, \sigma^2(\gamma_v \mathbf{I}_t - \mathbf{C}^{tt})), \tag{3.1}$$

where $\mathbf{C}^{tt}$ is the (scaled) prediction error variance matrix for the test line effects and is given by the partition of $\left(\mathbf{Z}_v'\mathbf{S}\mathbf{Z}_v + \mathbf{G}_v^{-1}\right)^{-1}$ corresponding to $\mathbf{u}_t$. We could then calculate EGG as the mean of the top $m\%$ of values from the distribution in (3.1). One of the aims of the simulation study in Section 5.2.5 is to show that EGG computed in this way, that is, based on the approximation to the distribution of the E-BLUPs as given in (3.1) correlates well with realized genetic gain (RGG).

A computationally simpler method of calculation for EGG is obtained by using a further approximation, namely, to replace the full prediction error variance matrix in (3.1) with a scaled identity matrix. A good approximation can be obtained using the concept of "effective error variance" (see, e.g., Cochran and Cox 1957). Thus, we approximate $\mathbf{C}^{tt}$ by $A_{tt}\mathbf{I}_t/2$ where $\sigma^2 A_{tt}$ is the average pairwise prediction error variance of test line effects, that is,

$$A_{tt} = \frac{2}{t-1}\left(\operatorname{tr}\left(\mathbf{C}^{tt}\right) - \frac{1}{t}\mathbf{1}_t'\mathbf{C}^{tt}\mathbf{1}_t\right). \tag{3.2}$$

Then we can calculate EGG as

$$\begin{aligned} \text{EGG} &= i\sqrt{\sigma^2(\gamma_v - A_{tt}/2)} \\ &= i\sigma_v h_g, \end{aligned} \tag{3.3}$$

where i is the selection intensity corresponding to m (i.e., the mean of the top $m\%$ of order statistics from a standard normal distribution of size t) and h_g is the square root of a generalized measure of heritability with $h_g^2 = 1 - A_{tt}/(2\gamma_v)$. Note that the equation for EGG in (3.3) is analogous to the standard quantitative genetics formula (see, e.g., Falconer and Mackay 1996). The difference is that we propose the use of the generalized measure of heritability rather than the standard measure that is calculated as the ratio of genetic variance to total (genetic plus error) variance. In the simplest case of balanced data and a model with no fixed effects other than an overall mean and no random effects other than those associated with variety and residual error (both with simple scaled identity variance matrices) the generalized and standard heritability measures are identical. In all other cases they will differ and importantly, the standard heritability measure will not relate to response to selection.

Note that we have found that values of EGG based on the full distribution of (3.1) and the approximate method of (3.3) are very similar (see, e.g., Section 5.2.5). Given the simplicity of calculation of (3.3) and the analogy with the standard formula we therefore choose to use (3.3).

Thus it is clear that, for a given selection intensity and genetic variance, EGG is maximized if the average of the prediction error variances of all elementary contrasts between test lines is minimized. Thus we choose A_{tt} as our design optimality measure.

It is interesting to explore the link between our optimality measure and the standard A-optimality measure used in the fixed effects setting. With the assumption of fixed variety effects the analogous mixed model to (2.1) could be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}_v\boldsymbol{\tau}_v + \mathbf{Z}_b\mathbf{u}_b + \mathbf{e},$$

where $\boldsymbol{\tau}_v$ is the $v \times 1$ vector of fixed variety effects. The A-optimality measure, namely the average pairwise variance for fixed test line effects, can be written as

$$A_{tt}^* = \frac{2}{t-1}\left(\text{tr}\left(\mathbf{C}^{*tt}\right) - \frac{1}{t}\mathbf{1}_t'\mathbf{C}^{*tt}\mathbf{1}_t\right),$$

where $\mathbf{C}^{*tt}$ is the partition of $\left(\mathbf{Z}_v'\mathbf{S}\mathbf{Z}_v\right)^-$ corresponding to the test line effects. Thus, for large values of γ_v, the values of A_{tt} and A_{tt}^* will be very similar, and, in the limit, identical.

3.2 SEARCH ALGORITHM

We seek to optimize A_{tt} for the model in (2.1) with prespecified design matrices for effects not associated with varieties (e.g., fixed covariate effects and random block effects) and prespecified variance parameters for the random and residual effects. This requires a search algorithm. The fundamental concept is the permutation of rows of $\mathbf{Z}_v$ to achieve a

design with minimum value of A_{tt}. Formally we define a design as a permutation vector $\mathbf{p}$ of length n representing the allocation of varieties to plots through the ordering of the rows of $\mathbf{Z}_v$. We then define the $n \times n$ matrix $\mathbf{P}$ as the row permutation of $\mathbf{I}_n$ corresponding to $\mathbf{p}$. Given a design matrix $\mathbf{Z}_v$ and permutation matrix $\mathbf{P}$ a new design is obtained as $\mathbf{Z}_v^* = \mathbf{P}\mathbf{Z}_v$. We also define a perturbation function $g()$ that operates on $\mathbf{p}$ to produce a new permutation $\mathbf{p}^*$, that is, $\mathbf{p}^* = g(\mathbf{p})$. In the simplest case, and in our algorithm, $g()$ is the function performing a two plot interchange subject to constraints such as resolvability (i.e., restricting interchanges to swaps within resolvable blocks).

The algorithm minimizes the objective function for either fixed or random variety effects (that is using A_{tt}^* or A_{tt}) by repeated application of $g()$ to $\mathbf{p}$ supervised by an optimisation strategy. The strategy we use is the Reactive TABU Search (RTS) [as adapted by Coombes (2002) for use in experimental design].

4. A NEW CLASS OF DESIGNS

The proposed designs involve the use of replicated plots for a percentage, p, of the test lines. They are henceforth referred to as "p-rep" designs. The original motivation was as an alternative to grid-plot designs so that the total trial size (number of plots) is maintained but grid plots are replaced by replicated plots of test lines. Typically the ratio of grid plots to test plots in EGVTs is 1:4 so we base our designs on $p = 25\%$ and use two plots of each replicated line. Other values of p and levels of replication may be used and the basic design is easily modified to suit specific requirements. For example, standard varieties may be included with higher levels of replication than test lines. The choice of test lines to be replicated may be made completely at random or may be influenced by the availability of seed or interest to the breeder.

As an example we consider a field trial in which 120 lines are to be tested. One possible grid-plot design based on a grid frequency of 1:4 involves 15 plots of each of two standard varieties making a total of 150 plots. We assume these are laid out in a rectangular array of 5 columns by 30 rows. A p-rep design for this scenario involves the use of two replicates of each of 30 test lines (so that $p = 25\%$) and single plots of the remaining 90 lines. Example randomizations for this setting are shown in Figure 1. These randomizations were obtained using the optimization strategy described in Section 3.2 with three different scenarios for the variety effects, namely fixed effects (so the standard A-optimality criteria, A_{tt}^*, was used) and random effects with values of $\gamma_{v_{\text{des}}}$ of 0.5 and 0.1. Note that the subscript in $\gamma_{v_{\text{des}}}$ denotes that this is the value of the genetic variance ratio used to construct the design. The nongenetic part of the model used to generate the design was the same in all cases, namely a model with random row and column effects (each with a variance component ratio of 0.1) and an AR1$\times$AR1 spatial process with row and column autocorrelation parameters of 0.6 and 0.4, respectively. The designs were constructed to be resolvable by restricting the search to ensure that all 30 replicated lines occurred once in the top half of each design (i.e., rows 1 to 15) and then again in the bottom half (rows 16 to 30). It is clear that there are marked differences between designs generated using the standard optimality measure compared with the measure based on expected genetic gain, particularly when $\gamma_{v_{\text{des}}}$ is small.

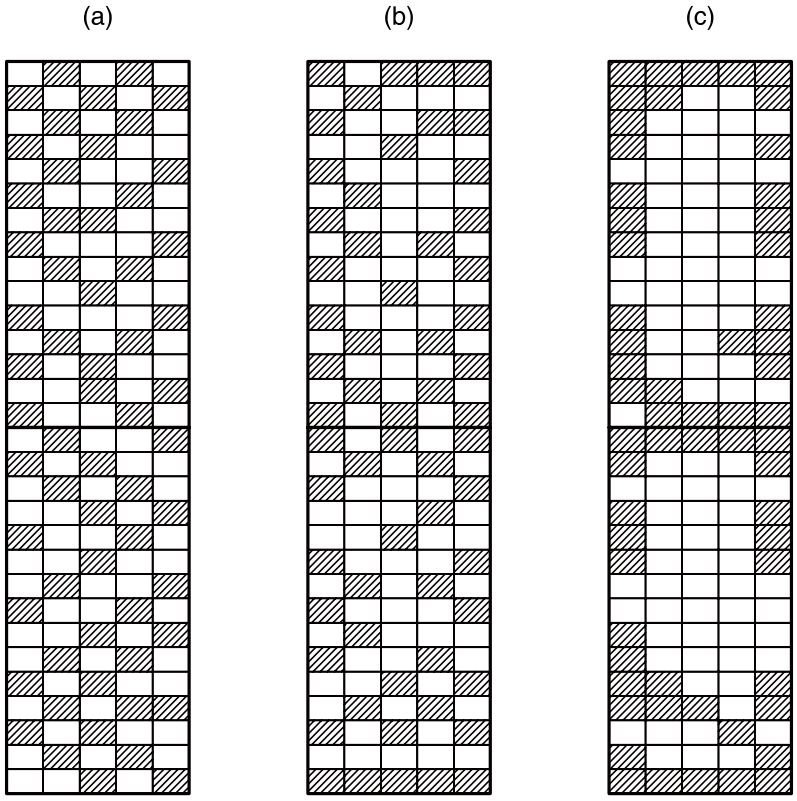

Figure 1. p-rep designs for 120 test lines with $p = 25\%$ replicated in a trial laid out as 5 columns by 30 rows. Designs are based on (a) fixed variety effects (b) random variety effects with $\gamma_{v_{\text{des}}} = 0.5$ and (c) random variety effects with $\gamma_{v_{\text{des}}} = 0.1$. Plots with replicated lines are shaded. Resolvable blocks comprise 5 columns by 15 rows.

We reiterate that extensions to the basic design are easily obtained. As a real example we cite the recent case of an EGVT that required a design for 193 test lines and 5 check varieties with a layout of 40 rows by 10 columns. The generated design involved two replicates for 150 of the test lines, single plots for the remaining 43 lines, 12 plots each for two of the check varieties and 11 plots each for the remaining three check varieties (see Coombes 2002 for an optimal randomization).

5. RELATIVE EFFICIENCY OF *P*-REP DESIGNS

5.1 METHODOLOGY

The performance of the new designs was compared against grid-plot designs using a simulation study. We assessed six designs for the case of 120 lines being tested in a field trial with a total of 150 plots (laid out as 5 columns by 30 rows). The designs comprised the three *p*-rep designs described in Section 4 together with three grid-plot designs based

on the same scenarios for variety effects (i.e., fixed variety effects or random variety effects with $\gamma_{v_{\text{des}}} = 0.5,\ 0.1$). The grid-plot designs involved 2 standard varieties with 15 plots of each. All designs were generated using the model for nongenetic effects as given in Section 4.

Data for each design were generated according to 12 different models that were chosen as being typical of data models found in EGVTs. The models comprised the factorial combinations of three values for $\gamma_{v_{\text{dat}}}$, namely, 0.1, 0.5, and 1 and 4 models for nongenetic effects labeled (a) $\gamma_{r_{\text{dat}}} = \gamma_{c_{\text{dat}}} = 0.1, \phi_r = 0.6, \phi_c = 0.4$; (b) $\gamma_{r_{\text{dat}}} = \gamma_{c_{\text{dat}}} = 0, \phi_r = 0.6, \phi_c = 0.4$; (c) $\gamma_{r_{\text{dat}}} = \gamma_{c_{\text{dat}}} = 0.1, \phi_r = 0.4, \phi_c = 0.6$; (d) $\gamma_{r_{\text{dat}}} = \gamma_{c_{\text{dat}}} = 0, \phi_r = 0.4, \phi_c = 0.6$. Note that the subscript in $\gamma_{\cdot_{\text{dat}}}$ denotes that this is the value of the genetic variance ratio used to generate the data (as distinct from constructing the design). The scale parameter, σ^2, was set at a value of 1.0. Note that the nongenetic data model "a" corresponds to the model used to generate the designs. Simulated data were analysed using models that matched the data generation model. For example, data generated using model "a" (for any of the six designs) were analyzed using a mixed model with random row and column effects and an AR1 $\times$ AR1 spatial process for the residual effects. We note that for the p-rep designs we would, in practice, also include random effects for resolvable blocks since this reflects the randomization employed in the design. However, since resolvable blocks are not part of the design model but are accommodated using a restriction of the search algorithm we have chosen to exclude them from the data (and thence analysis) error models. This does not compromise the generality of the results. All analyses were conducted using the samm (Butler, Cullis, Gilmour, and Gogel 2003) suite of functions. A total of 1,000 simulations was conducted for each of the 72 scenarios (6 designs by 12 data models). Note that for an individual simulation only three sets of genetic effects were generated (corresponding to the three values of $\gamma_{v_{\text{dat}}}$). These were then applied or subsetted across all designs and nongenetic data models in order to improve the accuracy of the associated comparisons.

5.2 RESULTS AND DISCUSSION

5.2.1 Frequency of Zero Values for Estimated Genetic Variance

Table 1 gives the percentage of analyses in which the genetic variance ratio was estimated as zero. As expected, this percentage is higher for smaller data values of $\gamma_{v_{\text{dat}}}$. The percentages for the lowest true value, that is, $\gamma_{v_{\text{dat}}} = 0.1$, are significantly ($p < 0.001$) higher for data generated from the grid-plot designs (ranging from 16.5 to 18.9% for averages across nongenetic models) compared with the p-rep designs (ranging from 13.4 to 14.9%). This has important consequences since an estimated genetic variance of zero implies that the predicted genetic effects for all lines are zero. Thus, the data cannot be used to make selections. In terms of the p-rep designs this only appears to be an issue for lowest data value of $\gamma_{v_{\text{dat}}} = 0.1$. Table 1(a) reveals a pattern among the nongenetic models for this value, with the frequency of zero estimates of genetic variance being higher for data models that include random row and column effects compared with those that do not (i.e., models "a" versus "b" and "c" versus "d"). This is particularly so for the design based on $\gamma_{v_{\text{des}}} = 0.1$. This is not surprising given the appearance of such designs (see, e.g., Figure

Table 1. Percentage of Analyses With Genetic Variance Estimated as Zero (that is, $\hat{\gamma}_{v_{dat}} = 0$).

	Grid-plot designs			*p-rep designs*		
Nongenetic	*Random effects*		*Fixed*	*Random effects*		*Fixed*
data model	$\gamma_{v_{des}} = 0.1$	$\gamma_{v_{des}} = 0.5$	*effects*	$\gamma_{v_{des}} = 0.1$	$\gamma_{v_{des}} = 0.5$	*effects*
	(a) Genetic data model: $\gamma_{v_{dat}} = 0.1$. Standard errors of means for individual data models: min = 1.02; max = 1.33; mean = 1.14					
a	17.6	16.8	16.8	15.4	14.8	14.0
b	17.5	14.3	16.9	12.4	12.6	12.5
c	22.9	18.7	18.2	18.5	13.8	15.4
d	17.5	16.2	15.7	13.3	12.8	11.8
average	18.9	16.5	16.9	14.9	13.5	13.4
	(b) Genetic data model: $\gamma_{v_{dat}} = 0.5$. Standard errors of means for individual data models: min = 0.10; max = 0.62; mean = 0.36					
a	3.7	4.0	3.3	0.3	0.8	0.4
b	2.3	2.5	1.7	0.1	0.1	0.7
c	3.1	3.6	3.3	0.6	0.2	0.4
d	2.6	3.5	2.5	0.3	0.2	0.1
average	2.9	3.4	2.7	0.3	0.3	0.4
	(c) Genetic data model: $\gamma_{v_{dat}} = 1.0$. Standard errors of means for individual data models: min = 0; max = 0.30; mean = 0.12					
a	0.4	0.7	0.8	0	0.1	0
b	0.5	0.4	0.4	0	0	0
c	0.4	0.6	0.8	0	0	0
d	0.4	0.9	0.1	0	0	0
average	0.4	0.7	0.5	0	0	0

1(c)) which are characterized by replicated varieties appearing in edge rows and columns so that the varietal contrasts of replicated versus unreplicated lines are confounded with row and column effects.

5.2.2 Realized Genetic Gain

Table 2 gives the realized genetic gain as a percentage of the true genetic gain for selection of the top 20% of lines. The realized gain was calculated for each simulation as the mean of the largest 24 test line E-BLUPs. The true genetic gain was calculated in the same manner but using the true test line effects as generated for that simulation. As discussed in Section 5.2.1, when the genetic variance is estimated as zero the realized genetic gain is also zero. The figures in Table 2 are means over those simulations where the estimated genetic variance was positive. The relative genetic gains for the p-rep designs (averaged over nongenetic models) are significantly higher ($p < 0.0001$) than the corresponding grid-plot values. In terms of individual data models there is a similar pattern as observed

Table 2. Realized Genetic Gain (RGG) Expressed as Percentage of True Genetic Gain for Selection of 20% of Test Lines. Means over simulations where genetic variance estimated as positive. Standard errors of means for individual data models: min = 0.44; max = 0.49; mean = 0.46.

	Grid-plot designs			*p-rep designs*		
Nongenetic	*Random effects*		*Fixed*	*Random effects*		*Fixed*
data model	$\gamma_{v_{des}} = 0.1$	$\gamma_{v_{des}} = 0.5$	*effects*	$\gamma_{v_{des}} = 0.1$	$\gamma_{v_{des}} = 0.5$	*effects*
	(a) Genetic data model: $\gamma_{v_{dat}} = 0.1$.					
a	43.9	44.7	44.0	46.9	47.3	47.0
b	44.2	43.7	43.9	48.5	47.8	48.0
c	43.8	43.7	43.5	47.4	46.4	46.7
d	44.1	44.1	43.7	48.2	47.8	47.4
average	44.0	44.1	43.8	47.7	47.4	47.3
	(b) Genetic data model: $\gamma_{v_{dat}} = 0.5$					
a	69.7	69.7	70.0	72.1	72.6	72.4
b	71.1	71.3	71.3	73.8	73.4	73.2
c	68.8	69.2	69.2	71.9	72.1	71.9
d	70.4	70.6	70.5	72.9	73.1	73.1
average	70.0	70.2	70.3	72.7	72.8	72.7
	(c) Genetic data model: $\gamma_{v_{dat}} = 1.0$					
a	78.8	78.9	78.7	80.9	81.4	80.9
b	80.4	80.5	80.6	82.1	82.6	82.9
c	78.8	79.2	78.9	80.8	81.2	81.1
d	79.5	80.0	80.0	82.1	82.3	82.3
average	79.4	79.6	79.6	81.5	81.9	81.8

in Table 1, namely that the genetic gain is lower for models that include random row and column effects compared with those that do not.

Table 3 gives the realized genetic gains averaged over all simulations, that is, including those simulations in which the genetic variance was estimated as zero. These figures therefore reflect a combination of the information contained in Tables 1 and 2. Thus, the superiority of *p*-rep over grid-plot designs shown in Table 3 is even greater than that of Table 2.

5.2.3 SELECTION OF REPLICATED LINES

It is important to investigate whether there is any bias towards selection of replicated lines in the *p*-rep designs. Thus in each simulation we recorded the number of replicated lines in the top 20% of lines as selected on the basis of the E-BLUPs. The number of replicated lines in the corresponding true top 20% was also recorded. These values were transformed to the logit scale then averaged across simulations. To reduce bias and avoid problems of infinite values we added 0.5 to both the numerator and denominator in the logit

Table 3. Realized Genetic Gain Expressed as Percentage of True Genetic Gain for Selection of 20% of Test Lines. Means over all simulations including those where genetic variance estimated as zero. Figures are averages over the four nongenetic data models.

	Grid-plot designs			*p-rep designs*		
Data	*Random effects*		*Fixed*	*Random effects*		*Fixed*
model $\gamma_{v_{dat}}$	$\gamma_{v_{des}}$ *=0.1*	$\gamma_{v_{des}}$ *=0.5*	*effects*	$\gamma_{v_{des}}$ *= 0.1*	$\gamma_{v_{des}}$ *= 0.5*	*effects*
0.1	35.7	36.8	36.4	40.6	41.0	40.9
0.5	68.0	67.8	68.4	72.4	72.6	72.4
1.0	79.1	79.1	79.1	81.5	81.9	81.8

transformation. Table 4 gives the results (back-transformed to percentages) for selection based on the E-BLUPs. These values must be compared with the value of 20% which was the analogous figure for selection based on the true test line effects. Table 4 shows that for the smaller data model values of γ_v there is a slight bias towards selection of replicated lines.

5.2.4 Matching Designs to Data

There is some evidence to suggest that, for the p-rep designs, realized genetic gain for a given true value of $\gamma_{v_{\text{dat}}}$ is highest for a design based on a similar value of $\gamma_{v_{\text{des}}}$ (see Table 2). However, one would be reluctant to recommend use of designs based on very low values of $\gamma_{v_{\text{des}}}$ (e.g., $\gamma_{v_{\text{des}}} = 0.1$) since they have some undesirable features. In particular the frequency of zero estimates of genetic variance is high (Table 1). When this is factored into realized genetic gain then for true $\gamma_{v_{\text{dat}}} = 0.1$ this design performs worse than the other two designs (Table 3). Also the designs have a very unusual appearance with replicated plots grouped around the edges of the resolvable blocks (Figure 1(c)). As a general recommendation it appears that designs based on the "moderate" value of $\gamma_{v_{\text{des}}} = 0.5$ provide a reasonable solution for true values of $\gamma_{v_{\text{dat}}}$ in the range of 0.1 to 1.0. These designs perform relatively well both in terms of the frequency of zero estimates of genetic variance and realized genetic gain so that they have the highest values in Table 3.

Table 4. Percentage of Replicated Lines in Observed Top 20% of Lines. Percentage of replicated lines in true top 20% is 20%.

Data	*Random effects*		*Fixed*
model $\gamma_{v_{dat}}$	$\gamma_{v_{des}} = 0.1$	$\gamma_{v_{des}} = 0.5$	*effects*
0.1	22.6	22.9	22.9
0.5	21.8	21.9	21.7
1.0	21.7	21.0	21.0

5.2.5 Performance of Optimality Measure

It is instructive to compare the basis of the design optimality measure, that is, EGG calculated using equation (3.3), with the realized genetic gain. Correlations were calculated between the RGG values from Table 2 and the corresponding EGG values calculated using Equation (3.3) for the p-rep designs only and separately for the three data model values of $\gamma_{v_{\text{dat}}}$ (so correlations were based on $n = 12$ data points). All correlations were high ($r = 0.802,\ 0.932$, and 0.940 for $\gamma_{v_{\text{dat}}} = 0.1,\ 0.5$, and 1.0, respectively) with the strongest association for the larger data values of $\gamma_{v_{\text{dat}}}$. Additionally we considered the relationship between the two measures of EGG, namely, based on the full (approximate) distribution of the predicted test line effects (Equation (3.1)) and based on A_{tt} (Equation (3.3)). The correlations (for the p-rep designs only) were very high ($r = 0.998,\ 0.997$, and 0.995 for $\gamma_{v_{\text{dat}}} = 0.1,\ 0.5$, and 1.0, respectively). This suggests that there is little loss in using A_{tt} compared with the full distribution. Finally, the strength of association observed here between EGG (calculated using A_{tt}) and realized genetic gain gives confidence in the use of our design optimality criteria, since it suggests that designs chosen as superior on the basis of A_{tt} are likely to also be superior in terms of realized genetic gain.

6. CONCLUDING REMARKS

In this article we have described a new class of designs, so-called p-rep designs, that provide an alternative to grid-plot designs for EGVTs. The basic premise is to replicate a percentage, p, of test lines and use single plots of the remainder. A simulation study showed that, for a fixed trial size, p-rep designs resulted in higher genetic gains than grid-plot designs. The use of these designs should therefore have a positive impact on industry. As examples, we consider the Australian wheat and barley industries. The average price paid to farmers for these crops over the last five years is \$187/ha for wheat and \$154/ha for barley. Based on these figures the average benefit of p-rep designs over grid-plot designs observed in this study was \$1.63/ha for wheat and \$1.34/ha for barley.

We have proposed a new optimality measure for EGVTs, namely the average pairwise prediction error variance of test line effects (A_{tt}). This reflects the aim which is to maximize the response to selection (or genetic gain). The simulation study showed that this optimality measure performed well in the sense that designs identified as superior in terms of A_{tt} were, in general, also superior in terms of realised genetic gain. The assumption implicit with A_{tt} is that variety effects are random. This is in contrast to standard approaches to design in which treatment effects are regarded as fixed. In terms of design generation the use of random variety effects requires specification of a value for the genetic variance component ratio ($\gamma_{v_{\text{des}}}$). The simulation study showed that the use of a value of $\gamma_{v_{\text{des}}} = 0.5$ provided designs that performed well over the range of data values of $\gamma_{v_{\text{dat}}}$ under consideration (and typical of EGVTs).

An important issue is that EGVTs are often conducted and analyzed as series of trials, known as multi-environment trials. The proposed designs are ideally suited to this setting since there is potential to balance test line replication across trials. For example, consider

the trial used in this study in which 120 test lines were assessed using 150 plots. If four such trials were grown as a series, then replication of $p = 25\%$ (i.e., 30 lines) in each would allow a total of five plots for each test line. The issue of how best to allocate test lines to plots across trials is complex and is the subject of current research.

ACKNOWLEDGMENTS

This work has received financial support from the Grains Research and Development Corporation of Australia. NEC was supported by a Grains Research and Development Corporation scholarship. We thank referees for useful comments that have improved the article.

[Received August 2005. Revised May 2006.]

REFERENCES

Atiqullah, M., and Cox, D. (1962), "The Use of Control Observations as an Alternative to Incomplete Block Designs," *Biometrika*, 49, 464–471.

Butler, D. G., Cullis, B. R., Gilmour, A. R., and Gogel, B. J. (2003), "samm Reference Manual," Training series, No QE02001, QLD Department of Primary Industries and Fisheries, Brisbane, QLD.

Chan, B. (1999), "The Design of Field Experiments When the Data are Spatially Correlated," PhD thesis, The University of Qld, Brisbane.

Chauhan, N. (2000), "Efficient and Optimal Designs for Correlated Observations," PhD thesis, University of Sheffield, U.K.

Cochran, W. G., and Cox, G. M. (1957), *Experimental Designs* (2nd ed.), New York: Wiley.

Coombes, N. (2002), "The Reactive Tabu Search for Efficient Correlated Experimental Designs." PhD thesis, Liverpool John Moores University, Liverpool, U.K.

Cullis, B., Lill, W., Fisher, J., and Read, B. (1989), "A New Procedure for the Analysis of Early Generation Variety Trials," *Applied Statistics*, 38, 361–375.

Falconer, D. S., and Mackay, T. (1996), *Introduction to Quantitative Genetics* (4th ed.), Essex, UK: Longman Scientific and Technical.

Gilmour, A. R., Cullis, B. R., and Verbyla, A. P. (1997), "Accounting For Natural And Extraneous Variation In The Analysis Of Field Experiments," *Journal of Agricultural, Biological, and Environmental Statistics*, 2, 269–273.

Glover, F. (1989), "Tabu Search—Part I," *ORSA Journal of Computing*, 1, 190–206.

——— (1990), "Tabu Search—Part II," *ORSA Journal of Computing*, 2, 4–32.

Kackar, R., and Harville, D. (1981), "Unbiasedness of Two-Stage Estimation and Predictions for Mixed Linear Models," *Communications in Statistics: Theory and Methods*, A10, 1249–1261.

Kempton, R. A. (1984), "The Design and Analysis of Unreplicated Field Trials," *Vortrage für Pflanzenzuchtung*, 7, 219–242.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983), "Optimization by Simulated Annealing," *Science*, 220, 671–680.

Martin, R. (1996), "Spatial Experimental Design," in *Handbook of Statistics 13: Design and Analysis of Experiments*, eds. S. Ghosh and C. R. Rao, Amsterdam: Elsevier Science, pp. 477–514.

Martin, R. J., and Eccleston, J. (1997), "Construction of Optimal and Near-Optimal Designs for Dependent Observations Using Simulated Annealing," Research report 479/97, Dept. of Probability and Statistics, University of Sheffield.

Patterson, H. D., and Thompson, R. (1971), "Recovery of Interblock Information When Block Sizes are Unequal," *Biometrika*, 31, 100–109.

Smith, A. B., Cullis, B. R., and Thompson, R. (2005), "The Analysis of Crop Cultivar Breeding and Evaluation Trials: An Overview of Current Mixed Model Approaches," *Journal of Agricultural Science*, 143, 449–462.