National Institute for Applied Statistics Research Australia

University of Wollongong, Australia

Working Paper

13-19

The Design of Early Stage Plant Breeding Trials using Genetic Relatedness

Brian R. Cullis, Alison B. Smith, Nicole A. Cocks and David G. Butler

The design of early stage plant breeding trials using genetic relatedness

Brian R. Cullis, Alison B. Smith, Nicole A. Cocks and David G. Butler
Centre for Bioinformatics and Biometrics,
National Institute for Applied Statistics Research Australia,
University of Wollongong, Australia

Abstract

The use of appropriate statistical methods has a key role in improving the accuracy of selection decisions in a plant breeding program. This is particularly important in the early stages of testing in which selections are based on data from a limited number of field trials that include large numbers of breeding lines with minimal replication. The method of analysis currently recommended for early stage trials in Australia involves a linear mixed model that includes genetic relatedness via ancestral information; non-genetic effects that reflect the experimental design and a residual model that accommodates spatial dependence. Such analyses have been widely accepted as they have been found to produce accurate predictions of both additive and total genetic effects, the latter providing the basis for selection decisions. In this paper we present the results of a case study of 34 early stage trials to demonstrate this type of analysis and to re-inforce the importance of including information on genetic relatedness. In addition to the application of a superior method of analysis, it is also critical to ensure the use of sound experimental designs. Recently, model-based designs have become popular in Australian plant breeding programs. Within this paradigm, the design search would ideally be based on a linear mixed model that matches, as closely as possible, the model used for analysis. Therefore, in this paper, we propose the use of models for design generation that include information on genetic relatedness and also include bespoke non-genetic and residual models based on historic data for individual breeding programs. At present the most commonly used design generation model omits genetic relatedness information and uses non-genetic and residual models that are supplied as default models in the associated software packages. The major reasons for this are that pre-existing software is unacceptably slow for designs incorporating genetic relatedness and the accuracy gains resulting from the use of genetic relatedness have not been quantified. Both of these issues are addressed in the current paper. An updating scheme for calculating the optimality criterion in the design

1

search is presented and is shown to afford prodigious computational savings. An in silico study that compares three types of design function across a range of ancillary treatments shows the gains in accuracy for the prediction of total genetic effects (and thence selection) achieved from model-based designs using genetic relatedness and bespoke non-genetic and residual models.

Keywords: linear mixed model; model-based design; genetic relatedness

1   Introduction

Plant breeding is focused on the objective of genetic improvement, producing new varieties with increased productivity and quality. Most plant breeding programs follow a method of breeding referred to as a pedigree selection method. This method implies that programs are structured around the grouping of breeding lines which have been derived as progeny of a fixed number of crosses between elite parents. Different crosses are made each year and the cohort of breeding lines then undergo selection through preliminary and advanced stages of testing.

Traits of interest for selection in the preliminary stage include disease and herbicide tolerance, phenology type and functional grain quality. Selection intensity is high, reflecting the relatively high heritability of these traits or the ability to use marker assisted selection techniques for simply inherited traits.

Typically, selection in the advanced stages occurs in a sequential manner. These stages are referred to as the S1, S2, S3 and S4 stages. They key selection trait in the advanced stages is grain yield. Yield data for each stage is generated from a series of field trials sown at several locations. Trials typically include the set of breeding lines of interest and also some check varieties, which will collectively be termed entries. The focus of this paper are S1 and S2 trials and we refer to these as the early stage trials. The number of entries which are evaluated at the S1 and S2 stages often exceeds 1,000, while at the S3 and S4 stages the number of entries tested is generally less than 100. At the S1 and S2 stages resource and seed limitations reduce the numbers of locations and plots that can be used for each breeding line. Replication within a location is often less than two and the number of locations is usually less than four. For the S3 and S4 stages entries are evaluated at up to ten locations, with between two to three replicates per location. Given the relatively low heritability for the trait of grain yield, and the presence of variety (entry) by environment interaction, it is critical to adopt efficient experimental designs and appropriate methods of analysis.

In Australia, the preferred approach for the analysis of multi-environment trial (MET) data-sets is the factor analytic linear mixed model of Smith et al. (2001). Their original approach considered modelling of the non-genetic effects within each environment using the spatial approach advocated by Gilmour et al. (1997) and modelling the variety by environment effects using a factor analytic model. Their approach did not include genetic relatedness for the variety effects for each environment. Oakey et al. (2006) addressed this for the analysis of a single trial and later Oakey et al. (2007) incorporated genetic relatedness through the use of ancestral information for the models advocated by Smith et al. (2001). More recently Smith & Cullis (2018) developed factor analytic selection tools to assist with seletion decisions from a factor analytic linear mixed model analysis of MET data-sets. These methods

are in widespread use for the analysis of MET data-sets in the advanced stages of selection in Australia.

There is an extensive literature on the design of field trials for plant breeding programs. Designs used for these trials fall into two broad categories: classical or optimal model-based designs. The former include complete and incomplete block designs, row-column designs or α-designs (see Bailey, 2008; John & Williams, 1995, 1998; Patterson & Williams, 1976, for example). The principle of model-based design is to search the design space for a design function (Bailey, 2008) which is near optimal under a pre-specified model. The model used for the design of field trials for plant breeding programs is usually a linear mixed model, which is consistent with the linear mixed model used for the analysis. Early work on model-based design for plant breeding selection trials focussed on methods to find optimal designs for spatially dependent data (Martin, 1986; Martin et al., 2006; Martin & Eccleston, 1992; Chan, 1999). More recently, model-based designs have been considered for more general linear mixed models which include spatial dependence and blocking factors (Butler et al., 2008; Coombes, 2002; Williams & John, 2006).

Cullis et al. (2006) introduced p-rep designs to address the problems associated with the design of S1 and S2 stage trials where the number of plots for each breeding line in a trial is less than two. They showed that p-rep designs improved the accuracy of selection of breeding lines compared to so-called grid-plot designs (Kempton, 1982). The p-rep designs are in widespread use in most plant breeding programs in Australia. These designs can be produced efficiently using statistical software such as the DiGGeR package (Coombes, 2009). Williams et al. (2011) developed augmented p-rep designs by combining p-rep and augmented check plot designs.

In order to align model-based design with current methods for the analysis of S1, S2 and S3 stage field trials, recent work has focused on the inclusion of genetic relatedness through the use of ancestral information. Bueno Filho & Gilmour (2003) consider the construction of non-resolvable incomplete block designs when the treatments effects are correlated. Their study considered three choices of genetic relatedness in a simplistic setting of six treatments in blocks of size four. Within the limited scope of their study they concluded that for most situations it would be reasonable to use a design which is optimal for unrelated treatment effects.

Piepho & Williams (2006) investigated three types of design for a simple genetic structure for the set of treatment effects. On the basis of a simulation experiment, they concluded that the assumption of correlated treatment effects was superior to assuming fixed treatment effects. They also recommended that an unrestricted α-design with a simple family-based treatment structure could be used for the design of field trials in plant breeding programs.

Butler et al. (2014) presented a more general approach for the design of field trials in a plant breeding program. Their approach was based on finding designs which were (near) optimal under a linear mixed model which partitioned the total genetic variance into additive and non-additive effects for in-bred crops or simpler models for non in-bred crops which only included additive effects. Models for plot effects (Bailey, 2008) were either classical (using fixed or random block effects), or based on spatial dependence or a mixture of both. Butler et al. (2014) illustrated their methods using two examples from S1 and S2 stage trials from a canola and sorghum breeding program respectively. Their designs were generated using the OD-V1 statistical software package (Butler, 2013).

There remain at least two issues which hinder the adoption of the Butler et al. (2014) designs for selection experiments in plant breeding programs. Firstly, there is a lack of empirical evidence regarding the improvement in accuracy relative to other classical or model-based design which are currently in use. Secondly, the computational load for the design search is prohibitive for trials with a larger number of entries. The aim of this paper is to therefore address these issues. The issue of computational burden is addressed by development of an updating algorithm for evaluation of the optimality criteria, which is an extension of the algorithm described by Martin & Eccleston (1992) and Chan (1999), to allow for correlated treatment effects. Secondly the empirical advantage of model-based designs with correlated effects is compared to two other commonly used designs using an in silico study, based on a large set of field trials from a plant breeding program based in Australia.

The paper is arranged as follows. In section 2 we present a detailed analysis of a case study based on a set of early stage trials grown in 2018 by the four publicly funded pulse breeding programs in Australia. In section 3 we present an approach to model-based design, including derivation of the updating formula for correlated treatment effects which addresses the computational burden associated with calculating the optimality criteria for each interchange in the design search. In this section we also demonstrate how to obtain a near optimal design using the R package OD-V2 (Butler & Cullis, 2018) and provide some results on the reduction in computing time using the updating formula. We conclude in section 4 with an in-silico experiment designed to assess the performance of model-based designs using genetic relatedness and so-called bespoke non-genetic models.

2 Case study

Given that the model used for model-based design is "usually chosen to be as close as possible to that expected for the analysis" (Butler et al., 2014), a case study involving the analysis of 34 early stage trials was conducted. The trials comprised S1 and S2 stage trials from four Australian public pulse breeding programs in 2018. The case study was used to highlight the importance of including information on genetic relatedness in the analysis and to summarise models used for non-genetic effects.

2.1 Genetic material, experimental designs and phenotyping

Comprehensive ancestral information was provided for all four programs and this was used to form numerator relationship matrices (Meuwissen & Luo, 1992) that are summarised in table 1. The median generation number was 8, 5, 10 and 6 for chickpeas, fababeans, field peas and lentils respectively. The median inbreeding coefficient for entries in all programs is consistent with the high level of inbreeding achieved prior to inclusion in yield evaluation trials. The average genetic relatedness (Dunner et al., 1998) is also high, reflecting the strong familial structure of the breeding lines in each of the programs. It is important to note however, that there is substantial heterogeneity of the genetic relatedness within each program. Those entries with low genetic relatedness were predominantly older commercial varieties which have not been used as parents in recent crosses, while those entries with high genetic relatedness were either recently released commercial varieties which have been used

as parents in many crosses or breeding lines from (full sib) families with a large representation in trials.

	Inbreeding			Relatedness			Entries in	
Crop	Min	Mean	Max	Min	Mean	Max	Pedigree	Data
Chickpea	0.50	0.809	0.998	0.091	0.532	0.915	1307	872
Fababean	0.50	0.934	0.999	0.000	0.305	0.490	712	564
Fieldpea	0.75	0.900	0.998	0.000	0.244	0.389	1717	1209
Lentil	0.75	0.881	0.997	0.000	0.278	0.421	2459	2005

Table 1: Summary of numerator relationship matrices for each crop in the case study: information includes the minimum, mean and maximum inbreeding coefficient and genetic relatedness. Final columns give the total number of entries in the pedigree and the number of entries with phenotypic data.

Table 2 presents a summary of key features of the 34 trials. All trials were laid out in rectangular arrays of plots, indexed by field columns and field rows. Plot dimensions were long and thin, with columns longer than rows. The trait of interest here is grain yield (t/ha). The mean yield varied substantially between trials, with some trials being severely affected by drought. Generally stage 1 trials had more entries than stage 2 trials but fewer plots per entry. Partially replicated designs (Cullis et al., 2006) were used for 10 out of the 12 stage 1 trials and 4 out of the 22 stage 2 trials. These so-called *p*-rep designs were resolvable (or near resolvable) with respect to the entries with more than one plot. The remaining trials were designed as resolvable two or three replicate block designs. In most cases there were also additional plots of commercially important varieties. The resolvable blocks were aligned either with columns (referred to as column replicate blocks: CRep), aligned with rows (referred to as row replicate blocks: RRep) or in both directions (trials 12 and 13). In all cases the resolvable column or row replicate blocks spanned multiple columns or rows respectively. Most designs were constructed by staff within each breeding program using either DiGGeR (Coombes, 2009), with default parameter settings, or Agrobase (Mulitze, 1990). Pedigree information was not used in the design process for these trials. Designs for trials 1, 11, 14, 15, 22, 27 and 29 were constructed using the methods to be described in the current paper using OD-V2 (Butler & Cullis, 2018).

2.2   Statistical models for analysis

The approach used for the analysis is similar to that proposed by Oakey et al. (2006). They considered two linear mixed models with differing assumptions regarding the distribution of the (random) genetic effects. They referred to these models as the *standard* and *pedigree* models. In both cases the non-genetic effects are modelled according to the approach devised by Gilmour et al. (1997), who allowed for three possible sources of variation, namely global, extraneous and local. The linear mixed models considered in this paper also include, by default, random terms which respect the design construction process. These included terms for resolvable or near resolvable column or row replicate blocks for all designs (either one or both as required) and terms for columns and rows when the designs were constructed

| | | | | | | | Entry replication | | | | | | Mean | NAs |
Crop	Trial	Stage	Plots	Cols	Rows	Entries	1	2	3-4	≥ 5	CRep	RRep	yield	yield
Chickpea	1	S1	720	10	72	359	0	358	1	0		36	1.861	4
Chickpea	2	S1	360	10	36	180	0	180	0	0		18	2.611	1
Chickpea	3	S1	468	18	26	389	332	52	4	1	9		3.133	6
Chickpea	4	S1	408	12	34	336	274	57	5	0	6		1.558	1
Chickpea	5	S2	288	8	36	144	0	144	0	0		18	1.565	0
Chickpea	6	S2	252	9	28	126	0	126	0	0		14	2.624	0
Chickpea	7	S2	288	6	48	145	10	134	0	0	3		1.637	11
Chickpea	8	S2	288	6	48	145	4	140	0	0	3		1.077	5
Chickpea	9	S2	288	6	48	144	0	144	0	0		24	1.195	7
Chickpea	10	S2	288	6	48	144	0	144	0	0		24	1.976	1
Fababean	11	S1	240	12	20	159	88	67	2	2	6		1.607	1
Fababean	12	S2	144	12	12	45	0	0	42	3	6	6	1.442	0
Fababean	13	S2	144	12	12	45	0	0	42	3	6	6	1.014	0
Fababean	14	S1	360	12	30	261	178	79	0	4	6		0.935	0
Fababean	15	S1	360	12	30	255	166	85	0	4	6		1.214	1
Fababean	16	S2	144	6	24	68	0	64	4	0	3		4.831	0
Fababean	17	S2	240	12	20	109	0	105	0	4	6		1.147	1
Fababean	18	S2	168	12	14	80	0	76	4	0		7	3.201	5
Fababean	19	S2	240	12	20	109	0	105	0	4	6		1.200	1
Fababean	20	S2	168	12	14	80	0	76	4	0	6		1.545	0
Fababean	21	S2	168	12	14	80	0	76	4	0		7	3.036	0
Fieldpea	22	S1	960	12	80	797	637	159	0	1	6		2.336	2
Fieldpea	23	S2	840	12	70	409	1	407	0	1	6		0.819	1
Fieldpea	24	S2	840	12	70	411	4	405	0	1	6		2.506	5
Fieldpea	25	S2	840	12	70	410	2	407	0	1	6		1.847	0
Fieldpea	26	S2	840	12	70	420	0	420	0	0	6		1.159	2
Lentil	27	S1	960	12	80	788	616	172	0	0		40	1.479	0
Lentil	28	S1	168	12	14	128	93	34	0	1		7	0.871	2
Lentil	29	S2	600	12	50	499	398	99	1	0	6		0.219	1
Lentil	30	S2	600	12	50	498	396	100	1	0	6		1.682	1
Lentil	31	S2	600	12	50	499	398	101	0	0	6		0.981	0
Lentil	32	S2	600	12	50	498	397	100	1	0	6		1.142	1
Lentil	33	S1	636	12	53	521	406	115	0	0	6		2.287	5
Lentil	34	S1	120	12	10	88	58	29	1	0			1.643	2

Table 2: Summary of trials in the case study: information includes the number of plots, columns, rows and entries; the number of entries with specified levels of replication (1, 2, 3-4, ≥5); the number of columns per column replicate and the number of rows per row replicate where applicable. Final columns give the mean yield (t/ha) and the number of plots with missing yield values.

using model-based approaches. Low-order polynomials in columns and/or rows were not considered. The residual vector was modelled as a separable first order auto-regressive process.

Following Oakey et al. (2006), the vector of genetic effects is partitioned into additive and non-additive effects given by

$$\boldsymbol{u_g} = \boldsymbol{u_a} + \boldsymbol{u_e} \tag{1}$$

where it is assumed that

$$
\begin{bmatrix} \boldsymbol{u_a} \\ \boldsymbol{u_e} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \sigma_a^2 \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_e^2 \boldsymbol{I_v} \end{bmatrix} \right)
$$

where $\boldsymbol{A}$ is the $v \times v$ numerator relationship matrix and v is the number of entries in the pedigree. The *pedigree* model estimates σ_a^2 and σ_e^2 by fitting both terms, while the *standard*

model excludes the additive effects, ignoring genetic relatedness. All analyses were conducted using Version 4 of the ASReml-R package (Butler et al., 2018).

2.3   Results of analysis

To allow for an unambiguous comparison of the fit of the two genetic models, all terms and variance models not associated with the genetic effects were chosen while fitting the *pedigree* model and thence retained for the fit of the *standard* model.

Table 3 presents a summary of the residual maximum likelihood (REML) estimates of the variance parameters associated with the non-genetic terms in the random model and the variance model for the residual. Random terms associated with resolvable blocks were always included in the model, while random terms which were non-resolvable, namely columns and rows, were only included as required. The major source of non-genetic variation in the random model terms is due to the (long) columns, being fitted on 33/34 occasions. This result supports the findings in Oakey et al. (2006), who also fitted terms associated with columns on most occasions (11/14 trials). This has important implications for choosing terms to be included in the random model for construction of near optimal model-based designs (see section 3.5.1).

The REML estimates of the row and column autoregressive parameters are moderate to high for all programs with the estimates for the row dimension being greater than that for the column dimension (which is consistent with the shape of the plots). The REML estimates of these parameters vary between crops but are somewhat lower than those reported by Oakey et al. (2006), since they fitted a measurement error component.

		Random terms											
		Resolvable				Non-resolvable				Residual			
		CRep		RRep		Col		Row		Col AR		Row AR		Var
Crop	Trials	Med	Ni	Med	Ni	Med	Ni	Med	Ni	Med	Nl	Med	Nl	Med
Chickpea	10	55	4	26	6	89	10	23	1	0.053	4	0.374	10	285
Fababean	11	307	9	101	4	183	11	172	3	0.171	6	0.482	10	568
Fieldpea	5	0	5		0	47	5		0	0.192	5	0.499	5	347
Lentil	8	4	5	71	2	33	7	6	1	0.244	6	0.440	6	441

Table 3: Summary of median of the REML estimates of the non-genetic variance parameters for each crop in the case study ($\times 10^4$ for all parameters except autoregressive correlations). The column labelled Ni for the random terms is the number of trials which included the term. The column labelled Nl for the residual autoregressive correlations (Col AR and Row AR) is the number of trials where the ratio of the REML estimate to its asymptotic standard error exceeded 1.5

Table 4 presents a summary of the REML estimates of the genetic variance parameters and the model-based reliability of the prediction of the total genetic effects. Additive genetic variance is the dominant source of genetic variance in all programs, though the relativity does vary between programs. These findings are consistent with those of Oakey et al. (2006) who reported a median percentage additive genetic variance of 84 for the 14 stage 3 wheat trials. The reliability of the predicted total genetic effects is typical of these trials.

Figure 1 presents a scatter plot of the log base 10 of the REML likelihood ratio test to test the hypothesis of zero additive genetic variance against the log base 10 of the number of entries with data for 33 trials (one trial was removed due to a near zero estimate of genetic variance). The hypothesis of zero additive genetic variance was strongly rejected for the majority of the trials, with the exceptions associated with trials with low values of genetic variance or small numbers of entries. These results are again consistent with those of Oakey et al. (2006).

Crop	Additive	Non-Additive	Total	%Additive	Reliability
Chickpea	165	15	306	97.6	0.606
Fababean	133	45	455	84.3	0.575
Fieldpea	110	45	279	75.1	0.593
Lentil	269	45	648	90.4	0.662

Table 4: Summary of analysis results for the case study: median of the REML estimates of the genetic variance parameters ($\times 10^4$), the percent additive genetic variance relative to the total genetic variance and the reliability of the prediction of the total genetic effect for entries with a single replicate.

3 Model-based design

Bailey (2008) defines a comparative experiment as an experiment in which we are interested only in contrasts between treatments. Early stage trials are comparative experiments which typically have a simple treatment structure. The aim is selection of the subset of breeding lines which are superior and hence will progress to the next stage of testing. An experimental design has three key elements: the plot structure, the treatment structure and the so-called design function. The design function is a function, T which allocates treatments to plots. Bailey (2008) defines the treatment structure as meaningful ways of dividing up the set of treatments ($\mathcal{T}$), the plot structure as meaningful ways of dividing up the set of plots (Ω), ignoring treatments, and the design function as a function from Ω to $\mathcal{T}$. In classical design the function $\mathcal{T}$ is chosen to satisfy certain combinatorial properties, whereas in model-based design the function $\mathcal{T}$ is chosen to result in a design which is optimal or near optimal for a pre-specified model. The model considered in this paper is a linear mixed model.

3.1 Linear mixed model

In the following we present a linear mixed model for $\boldsymbol{y}$, the n-vector of data, which is suitable for the model-based design (and analysis) of a comparative experiment. Nelder (1977) introduced the concept of two aspects of a random effect. He remarked that one kind of random term in a linear model is a component of error, while the other kind of random term represents those effects of interest. The latter type of random term will therefore be in the set of treatment factors, while the former type of random term will be in the set of plot factors. Applying this broad principle, we consider a linear mixed model with four
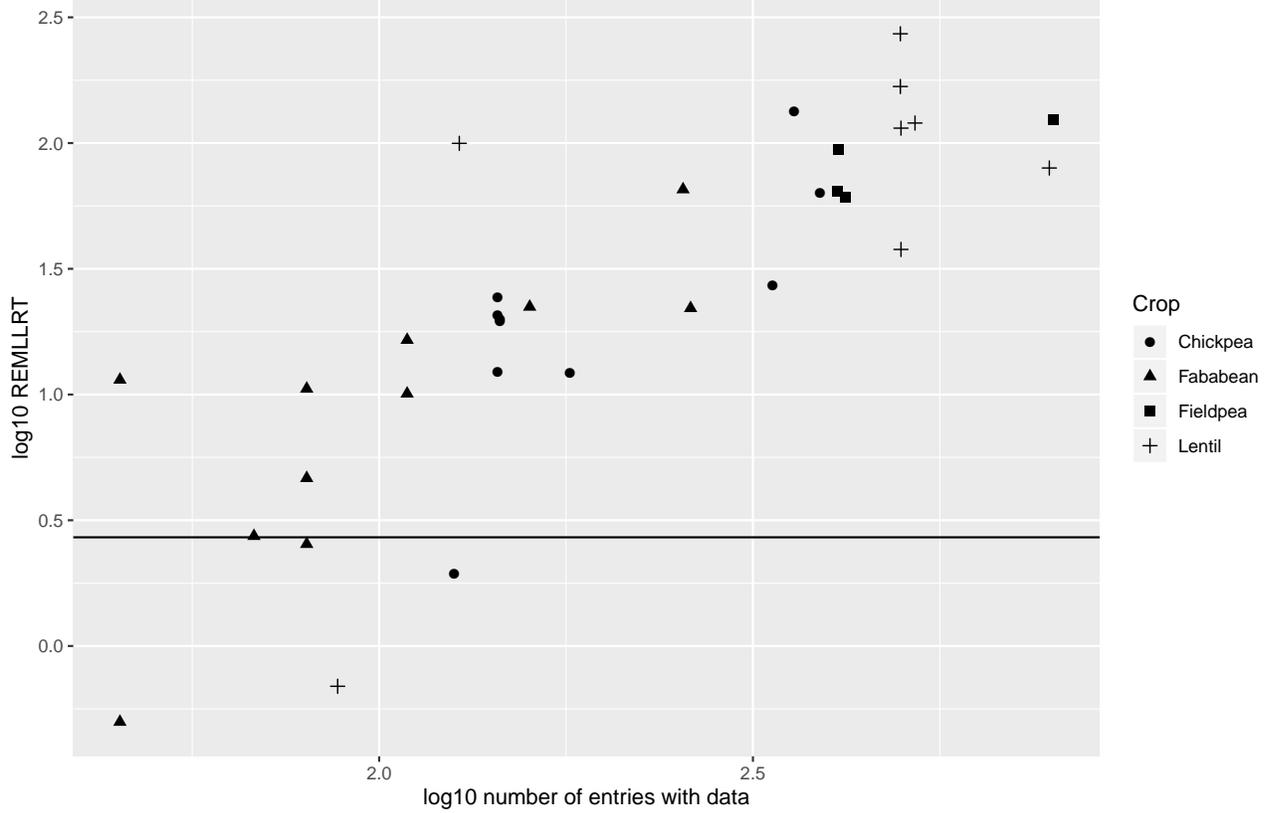
Figure 1: Scatter plot of the log base 10 of REML likelihood ratio statistic for the test that $\sigma_a^2 = 0$, against log base 10 of the number of entries in the trial. One trial removed due to near zero genetic variance. The solid horizontal line is located at the nominal 5% critical value for the statistic.

components given by

$$y = X_o\tau_o + X_p\tau_p + Z_ou_o + Z_pu_p + e \tag{2}$$

where τ_o and τ_p are vectors of fixed effects with associated design matrices X_o and X_p with c_{x_o} and c_{x_p} columns respectively; u_o and u_p are vectors of random effects with associated design matrices Z_o and Z_p with c_{z_o} and c_{z_p} columns respectively and e is the vector of residuals. The subscripts o and p identify so-called objective and peripheral fixed and random effects respectively. Equation (2) can be written succinctly as

$$\begin{aligned} y &= W_o\beta_o + W_p\beta_p + e \\ &= W\beta + e \end{aligned} \tag{3}$$

where $W_o = [X_o\ Z_o]$, $W_p = [X_p\ Z_p]$, $\beta_o = (\tau_o^\top, u_o^\top)^\top$, $\beta_p = (\tau_p^\top, u_p^\top)^\top$, $W = [W_o\ W_p]$ and $\beta = (\beta_o^\top, \beta_p^\top)^\top$.

The random effects and residuals in (2) are assumed to follow a Normal distribution such that:

$$\begin{bmatrix} u_o \\ u_p \\ e \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} G_o & 0 & 0 \\ 0 & G_p & 0 \\ 0 & 0 & R \end{bmatrix} \right)$$

9

where $\boldsymbol{G_o}, \boldsymbol{G_p}$ and $\boldsymbol{R}$ are positive definite matrices assumed to be functions of vectors of variance parameters $\boldsymbol{\sigma_{g_o}}, \boldsymbol{\sigma_{g_p}}$ and $\boldsymbol{\sigma_r}$ respectively. Model-based design requires values for these parameters so in the following they are regarded as known.

3.2 Predictions of interest

The aim is to find an optimal or near optimal design with respect to a d-vector of estimable functions $\boldsymbol{\pi} = \boldsymbol{D}\boldsymbol{\beta_o}$ where $\boldsymbol{D}$ is a known matrix with c_{w_o} columns and $c_{w_o} = c_{x_o} + c_{z_o}$. The vector of estimable functions, $\boldsymbol{\pi}$, involves only objective effects, but may involve fixed, random or both fixed and random effects. Gilmour et al. (2004) provide a computationally efficient algorithm for forming predictions from the linear mixed model specified in (3). For brevity in the following we use the terminology of Gilmour et al. (2004) and refer to $\boldsymbol{\pi}$ as the vector of predictions, which we assume are estimable. Gilmour et al. (2004) provide a simple test of estimability of predictions which fit naturally into their prediction algorithm. Briefly, given $\boldsymbol{D}$, the vector of predictions and associated prediction error variance/covariance matrix are formed by recursive absorption from an extended set of mixed model equations (see Robinson, 1991, for example). Full details can be found in Gilmour et al. (2004). The mixed model equations (MMEs) for (3) are given by

$$\boldsymbol{C}\tilde{\boldsymbol{\beta}} = \boldsymbol{W}^\top \boldsymbol{R}^{-1} \boldsymbol{y} \tag{4}$$

where the coefficient matrix in (4) is

$$\begin{bmatrix} \boldsymbol{W}_o^\top \boldsymbol{R}^{-1} \boldsymbol{W}_o + \boldsymbol{G}_o^* & \boldsymbol{W}_o^\top \boldsymbol{R}^{-1} \boldsymbol{W}_p \\ \boldsymbol{W}_p^\top \boldsymbol{R}^{-1} \boldsymbol{W}_o & \boldsymbol{W}_p^\top \boldsymbol{R}^{-1} \boldsymbol{W}_p + \boldsymbol{G}_p^* \end{bmatrix} \tag{5}$$

with

$$\boldsymbol{G}_o^* = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_o^{-1} \end{bmatrix} \quad \text{and} \quad \boldsymbol{G}_p^* = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_p^{-1} \end{bmatrix}$$

It follows that the reduced set of MMEs for $\tilde{\boldsymbol{\beta}}_o$ are given by

$$\boldsymbol{C}_{oo}\tilde{\boldsymbol{\beta}}_o = \boldsymbol{W}_o^\top \boldsymbol{P}_p \boldsymbol{y} \tag{6}$$

where $\boldsymbol{C}_{oo} = \boldsymbol{W}_o^\top \boldsymbol{P}_p \boldsymbol{W}_o + \boldsymbol{G}_o^*, \boldsymbol{P}_p = \boldsymbol{R}^{-1} - \boldsymbol{R}^{-1}\boldsymbol{W}_p(\boldsymbol{W}_p^\top \boldsymbol{R}^{-1}\boldsymbol{W}_p + \boldsymbol{G}_p^*)^-\boldsymbol{W}_p^\top \boldsymbol{R}^{-1}$ and $(\boldsymbol{W}_p^\top \boldsymbol{R}^{-1}\boldsymbol{W}_p + \boldsymbol{G}_p^*)^-$ is any particular generalised inverse of $\boldsymbol{W}_p^\top \boldsymbol{R}^{-1}\boldsymbol{W}_p + \boldsymbol{G}_p^*$. It can be shown that

$$\boldsymbol{P}_p = \boldsymbol{V}_p^{-1} - \boldsymbol{V}_p^{-1}\boldsymbol{X}_p(\boldsymbol{X}_p^\top \boldsymbol{V}_p^{-1}\boldsymbol{X}_p)^-\boldsymbol{X}_p^\top \boldsymbol{V}_p^{-1}$$

where $\boldsymbol{V}_p = \boldsymbol{Z}_p\boldsymbol{G}_p\boldsymbol{Z}_p^\top + \boldsymbol{R}$ and $(\boldsymbol{X}_p^\top \boldsymbol{V}_p^{-1}\boldsymbol{X}_p)^-$ is any particular generalised inverse of $\boldsymbol{X}_p^\top \boldsymbol{V}_p^{-1}\boldsymbol{X}_p$. The matrix $\boldsymbol{P}_p$ has rank $n - \operatorname{rank}(\boldsymbol{X}_p)$ and is unique, and it is the Moore-Penrose inverse of $\boldsymbol{T} = \boldsymbol{M}_p\boldsymbol{V}_p\boldsymbol{M}_p$, where $\boldsymbol{M}_p = \boldsymbol{I}_n - \boldsymbol{X}_p(\boldsymbol{X}_p^\top \boldsymbol{X}_p)^-\boldsymbol{X}_p^\top$. That is $\boldsymbol{T} = \boldsymbol{P}_p^+$.

For known $\boldsymbol{\sigma_{g_o}}, \boldsymbol{\sigma_{g_p}}$ and $\boldsymbol{\sigma_r}$

$$\boldsymbol{D}(\boldsymbol{\beta_o} - \tilde{\boldsymbol{\beta}}_o) \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Lambda}) \tag{7}$$

where $\boldsymbol{\Lambda} = \boldsymbol{D}\boldsymbol{C}_{oo}^-\boldsymbol{D}^\top$ and $\boldsymbol{C}_{oo}^-$ a particular generalised inverse of the coefficient matrix of (6).

3.3   Optimal design criteria

In the context of early stage trials, or more generally for comparative experiments, the most widely used and useful optimality criteria is the $\mathcal{A}$-optimality criteria (Martin, 1986). Bueno Filho & Gilmour (2007) developed a Bayesian design criterion for selection experiments in plant breeding based on a utility function that minimizes the risk of an incorrect selection. They show that this is in fact the $\mathcal{A}$-optimality criteria based on the PEV matrix for the vector of random entry effects. Bueno Filho & Gilmour (2003) and Cullis et al. (2006) use this criteria for generating optimal or near-optimal designs for plant breeding designs when the treatments are correlated and for so-called p-rep designs used in early stage trials. In this case it can be shown that

$$\mathcal{A} = \sum_i \sum_{j<i} \mathrm{pev}\,(\tilde{\pi}_i - \tilde{\pi}_j)\,/d/(d-1)$$

where $d \leq v$ is the number of varieties for which the $\mathcal{A}$-value is computed and pev () refers to the prediction error variance of its scalar (or matrix) argument. A convenient form for computing $\mathcal{A}$ is given by

$$\mathcal{A} = \frac{2}{d-1}(\mathrm{tr}\,(\boldsymbol{\Lambda}) - \mathbf{1}_d^\top \boldsymbol{\Lambda} \mathbf{1}_d/d)$$

We seek a design which minimises $\mathcal{A}$ over all valid design functions of the design.

3.4   Design search and updating formulae

The design search process presented in Butler et al. (2014) and implemented in OD-V2 (Butler & Cullis, 2018) can be summarised as follows: given an initial design, the $\mathcal{A}$-value is optimized under the supervision of a search strategy (for example a tabu search Glover, 1989) for (pairwise) interchanges of the rows of $\boldsymbol{W_o}$.

Each interchange requires evaluation of the $\mathcal{A}$-value, which is based on the prediction error variance matrix $\boldsymbol{\Lambda}$. This requires expensive matrix calculations and hence an exhaustive search of the design space for moderate to large problems is implausible. Updating formulae for $\boldsymbol{C}_{oo}^-$ can be used to significantly reduce the computational burden.

Martin & Eccleston (1992) developed an updating method for finding optimal designs under a linear model with correlated errors. Their algorithm is extended here, but allowing the more general setting of correlated objective effects, within the framework of a linear mixed model. The interchange of two rows of $\boldsymbol{W_o}$ is equivalent to first removing two rows from Ω, followed by adding the two units back, but in reverse order. Martin & Eccleston (1997) suggest using a four-step approach which involves adding or removing one unit to obtain the new $\boldsymbol{C}_{oo}^-$. Chan (1999) presented a two-step approach which she claims to be simpler and easier to implement. We begin by developing a two-step approach, similar to that proposed by Chan (1999), but a four-step approach is also presented as this has proven to be competitive to the two-step approach in terms of computational load. The four-step approach is also more suitable for use in the context of early stage trials, where the presence of so-called singleton treatments (that is, those treatments which occur on only one plot) can cause computational problems as discussed by Coombes (2002).

Let the current design contain $n = r + s$ units, which are referred to as plots in the following. Consider the retention of r plots by removing s plots, and both $\boldsymbol{W_o}$ and $\boldsymbol{P_p}$ are partitioned conformably with the removal of s plots as follows

$$\boldsymbol{W_o} = \left[\begin{array}{c} \boldsymbol{W_{or}} \\ \boldsymbol{W_{os}} \end{array} \right] \quad \text{and} \quad \boldsymbol{P_p} = \left[\begin{array}{cc} \boldsymbol{P_{p;rr}} & \boldsymbol{P_{p;rs}} \\ \boldsymbol{P_{p;sr}} & \boldsymbol{P_{p;ss}} \end{array} \right]$$

where for example $\boldsymbol{W_{or}}$ is the design matrix for the subset of the design corresponding to the r retained plots so has r rows and c_{w_o} columns. The matrix $\boldsymbol{T}$ is also partitioned conformably with $\boldsymbol{P_p}$. We note that in many cases $\boldsymbol{X_p}$ is null in which case $\boldsymbol{P_p} = \boldsymbol{V_p}^{-1}$ and hence $\boldsymbol{T} = \boldsymbol{V_p}$. It follows that the coefficient matrix of the reduced MMEs for the subset of the design corresponding to the r retained plots is

$$\boldsymbol{C_{oo}^{(r)}} = \boldsymbol{W_{or}^\top} \boldsymbol{T_{rr}^\sim} \boldsymbol{W_{or}} + \boldsymbol{G_o^*}$$

where $\boldsymbol{T_{rr}^\sim}$ is any particular generalised inverse of $\boldsymbol{T_{rr}}$.

Using a similar argument to Martin & Eccleston (1992) and results on the inverse of partitioned matrices (see for example Searle, 1982) it can be shown that

$$\boldsymbol{C_{oo}} - \boldsymbol{C_{oo}^{(r)}} = \boldsymbol{F_{rs}} \boldsymbol{P_{p;ss}^-} \boldsymbol{F_{rs}^\top} \tag{8}$$

where $\boldsymbol{F_{rs}} = \boldsymbol{W_{or}^\top} \boldsymbol{P_{p;rs}} + \boldsymbol{W_{os}^\top} \boldsymbol{P_{p;ss}}$. Hence an updating formula for $\boldsymbol{C_{oo}^{(r)-}}$ can now be derived using (8) as follows. Using Corollary 18.2.15 in Harville (1997) (p.432)

$$\boldsymbol{C_{oo}^{(r)-}} = \boldsymbol{C_{oo}^-} + \boldsymbol{C_{oo}^-} \boldsymbol{H_{rs}} (\boldsymbol{P_{p;ss}} - \boldsymbol{H_{rs}^\top} \boldsymbol{C_{oo}^-} \boldsymbol{H_{rs}})^- \boldsymbol{H_{rs}^\top} \boldsymbol{C_{oo}^-} \tag{9}$$

where $\boldsymbol{H_{rs}} = \boldsymbol{F_{rs}} \boldsymbol{P_{p;ss}^-} \boldsymbol{P_{p;ss}}$. If $\boldsymbol{P_{p;ss}}$ is non-singular then $\boldsymbol{H_{rs}} = \boldsymbol{F_{rs}}$. Furthermore, if $(\boldsymbol{P_{p;ss}} - \boldsymbol{H_{rs}^\top} \boldsymbol{C_{oo}^-} \boldsymbol{H_{rs}})$ is non-singular then (9) becomes

$$\boldsymbol{C_{oo}^{(r)-}} = \boldsymbol{C_{oo}^-} + \boldsymbol{C_{oo}^-} \boldsymbol{H_{rs}} (\boldsymbol{P_{p;ss}} - \boldsymbol{H_{rs}^\top} \boldsymbol{C_{oo}^-} \boldsymbol{H_{rs}})^{-1} \boldsymbol{H_{rs}^\top} \boldsymbol{C_{oo}^-} \tag{10}$$

Next we consider adding s units back to the design, but with the appropriate permutation applied to the rows of $\boldsymbol{W_{os}}$. In the simple case of $s = 2$, then provided that the interchange is legal, then the two rows of $\boldsymbol{W_{os}}$ are interchanged. The new design matrix for the full set of $n = r + s$ units is therefore given by

$$\boldsymbol{W_o^*} = \left[\begin{array}{c} \boldsymbol{W_{or}} \\ \boldsymbol{W_{os}^*} \end{array} \right]$$

where $\boldsymbol{W_{os}^*}$ is the permuted design matrix associated with the s units.

Using a similar approach to the deletion of r units, we consider the coefficient matrix of the reduced MMEs for the full, new design with $r + s$ units which is given by

$$\boldsymbol{C_{oo}^{(r+s)}} = \boldsymbol{W_o^{*\top}} \boldsymbol{P_p} \boldsymbol{W_o^*} + \boldsymbol{G_o^*}$$

It can be shown that

$$\boldsymbol{C_{oo}^{(r+s)}} - \boldsymbol{C_{oo}^{(r)}} = \boldsymbol{F_{rs}^*} \boldsymbol{P_{p;ss}^-} \boldsymbol{F_{rs}^{*\top}} \tag{11}$$

where $\boldsymbol{F}^*_{rs} = \boldsymbol{W}^\top_{or}\boldsymbol{P}_{p;rs} + \boldsymbol{W}^{*\top}_{os}\boldsymbol{P}_{p;ss}$. Hence an updating formula for $\boldsymbol{C}^{(r+s)-}_{oo}$ is

$$\boldsymbol{C}^{(r+s)-}_{oo} = \boldsymbol{C}^{(r)-}_{oo} - \boldsymbol{C}^{(r)-}_{oo}\boldsymbol{H}^*_{rs}(\boldsymbol{P}_{p;ss} + \boldsymbol{H}^{*\top}_{rs}\boldsymbol{C}^{(r)-}_{oo}\boldsymbol{H}^*_{rs})^-\boldsymbol{H}^{*\top}_{rs}\boldsymbol{C}^{(r)-}_{oo} \tag{12}$$

where $\boldsymbol{H}^*_{rs} = \boldsymbol{F}^*_{rs}\boldsymbol{P}^-_{p;ss}\boldsymbol{P}_{p;ss}$. If $\boldsymbol{P}_{p;ss}$ is non-singular then $\boldsymbol{H}^*_{rs} = \boldsymbol{F}^*_{rs}$. Furthermore, if $(\boldsymbol{P}_{p;ss} + \boldsymbol{H}^{*\top}_{rs}\boldsymbol{C}^{(r)-}_{oo}\boldsymbol{H}^*_{rs})$ is non-singular then (12) becomes

$$\boldsymbol{C}^{(r+s)-}_{oo} = \boldsymbol{C}^{(r)-}_{oo} - \boldsymbol{C}^{(r)-}_{oo}\boldsymbol{H}^*_{rs}(\boldsymbol{P}_{p;ss} + \boldsymbol{H}^{*\top}_{rs}\boldsymbol{C}^{(r)-}_{oo}\boldsymbol{H}^*_{rs})^{-1}\boldsymbol{H}^{*\top}_{rs}\boldsymbol{C}^{(r)-}_{oo} \tag{13}$$

Using the two updating formulae, one for the removal and one for the addition allows us to compute the new $\mathcal{A}$-value from $\boldsymbol{\Lambda}^{(r+s)} = \boldsymbol{D}\boldsymbol{C}^{(r+s)-}_{oo}\boldsymbol{D}^\top$.

Setting $s = 1$ leads to the four-step updating scheme proposed by Martin & Eccleston (1992). The four-step approach can have computational advantages and additionally it can be implemented to handle designs in which all of the objective effects are fixed effects and the design contains singletons. Using the updating formulae in (10) and (13), as an example, when $s = 1$ we have:

$$\begin{aligned}
\boldsymbol{C}^{(r)-}_{oo} &= \boldsymbol{C}^-_{oo} + \boldsymbol{C}^-_{oo}\boldsymbol{h}_{rs}(p_{p;ss} - \boldsymbol{h}^\top_{rs}\boldsymbol{C}^-_{oo}\boldsymbol{h}_{rs})^{-1}\boldsymbol{h}^\top_{rs}\boldsymbol{C}^-_{oo} \tag{14} \\
\boldsymbol{C}^{(r+s)-}_{oo} &= \boldsymbol{C}^{(r)-}_{oo} - \boldsymbol{C}^{(r)-}_{oo}\boldsymbol{h}^*_{rs}(p_{p;ss} + \boldsymbol{h}^{*\top}_{rs}\boldsymbol{C}^{(r)-}_{oo}\boldsymbol{h}^*_{rs})^{-1}\boldsymbol{h}^{*\top}_{rs}\boldsymbol{C}^{(r)-}_{oo} \tag{15}
\end{aligned}$$

for removal and addition of a unit respectively where

$$\boldsymbol{W}_o = \left[\begin{array}{c} \boldsymbol{W}_{or} \\ \boldsymbol{w}^\top_{os} \end{array}\right], \quad \boldsymbol{W}^*_o = \left[\begin{array}{c} \boldsymbol{W}_{or} \\ \boldsymbol{w}^{*\top}_{os} \end{array}\right] \quad \text{and} \quad \boldsymbol{P}_p = \left[\begin{array}{cc} \boldsymbol{P}_{p;rr} & \boldsymbol{p}_{p;rs} \\ \boldsymbol{p}^\top_{p;sr} & p_{p;ss} \end{array}\right]$$

and

$$\begin{aligned}
\boldsymbol{h}_{rs} &= \boldsymbol{f}_{rs} = \boldsymbol{W}^\top_{or}\boldsymbol{p}_{p;rs} + \boldsymbol{w}_{os}p_{p;ss} \\
\boldsymbol{h}^*_{rs} &= \boldsymbol{f}^*_{rs} = \boldsymbol{W}^\top_{or}\boldsymbol{p}_{p;rs} + \boldsymbol{w}^*_{os}p_{p;ss}
\end{aligned}$$

for $p_{p;ss} - \boldsymbol{h}^\top_{rs}\boldsymbol{C}^-_{oo}\boldsymbol{h}_{rs} > 0$.

The two-step approach is straightforward, however it is useful to illustrate the four-step approach in more detail. As a simple example of the four-step approach consider interchange of two units in the design, denoted by (ω_a, ω_b). The rows of the objective design matrix $\boldsymbol{W}_o$ are indexed using Ω such that $\boldsymbol{W}_o = \boldsymbol{W}_o[\Omega,]$ and let Ω_{-a} be the set of plots excluding ω_a and Ω_{-b} be the set of plots excluding ω_b, then the four-step approach consists of the following four steps:

1. Remove ω_a: $\boldsymbol{W}_{or} = \boldsymbol{W}_o[\Omega_{-a},]$ and $\boldsymbol{w}^\top_{os} = \boldsymbol{W}_o[\omega_a,]$ then use (14);

2. Add ω_b: $\boldsymbol{W}_{or} = \boldsymbol{W}_o[\Omega_{-a},]$ and $\boldsymbol{w}^{*\top}_{os} = \boldsymbol{W}_o[\omega_b,]$ then use (15);

3. Remove ω_b: $\boldsymbol{W}_{or} = \boldsymbol{W}^*_o[\Omega_{-b},]$ and $\boldsymbol{w}^\top_{os} = \boldsymbol{W}^*_o[\omega_b,]$ then use (14);

4. Add ω_a: $\boldsymbol{W}_{or} = \boldsymbol{W}^*_o[\Omega_{-b},]$ and $\boldsymbol{w}^{*\top}_{os} = \boldsymbol{W}_o[\omega_a,]$ then use (15)

noting that if a singleton is included in the set (ω_a, ω_b) then the steps must be arranged so that the singleton is not removed first.

13

3.5 Implementation in OD-V2

3.5.1 Example

The implementation of model-based design for early stage trials is shown here via an example based on one of the breeding programs introduced in section 2. This example also forms the basis of the simulation study presented in section 4. A design is required for testing 256 S1 stage entries from the field peas program with the aim of making selections for progression to S2 stage testing. The trial will also include four check varieties, which is standard practice in most S1 stage field trials. The predictions of interest are the total genetic effects for each of the 260 entries. In the linear mixed model the total genetic effects will be partitioned into additive and non-additive effects through the use of the numerator relationship matrix for the 260 entries. Summaries of this matrix included a mean inbreeding coefficient of 0.941, while the minimum, mean and maximum genetic relatedness were 0.130, 0.346 and 0.532. One of the check varieties had the highest mean genetic relatedness, as it had been widely used as a parent of the test lines.

A p-rep design will be constructed, with 50% of the breeding lines having two plots each. (Note that percentage replication will be examined in section 4.) The check varieties will also each have two plots so the full design involves 392 plots that will be arranged as 12 columns by 28 rows. The linear mixed model for the design search will include terms for non-genetic effects as determined from historic analyses of relevant trials. In the current example this relates to a comprehensive set of S1 and S2 stage trials from the field peas breeding program for the years 2013 to 2017 inclusively. This led to the inclusion of random effects for a two-level blocking factor CRep (corresponding to columns 1-6 and 7-12); random Column and Row effects and finally a separable first order autoregressive process for the residuals. The design search requires values of the associated variance parameters and these will be taken as the median values from the historic analyses. This approach leads to bespoke designs for individual breeding programs and selection decisions.

The design for the example can be constructed in OD-V2 (Butler & Cullis, 2018) using the following call:

```
example.od <- od(fixed=~ 1,
      random=~ ric(Entry, Ainv) + CRep + Column + Row,
      residual = ~ar1(Column):ar1(Row),
      permute=~ ric(Entry, Ainv),
      G.param = sv, R.param = sv, maxit=10, search="tabu", data=data.df)
```

In this call the data-frame data.df has 392 rows and corresponds to an initial (user-supplied) design. The genetic effects are specified in the term ric(Entry, Ainv) and relate to the total genetic effects ($\boldsymbol{u_g}$). The variance function ric() indicates that the variance model for the effects in the factor argument (that is, Entry) is given by $\sigma_a^2 \boldsymbol{A} + \sigma_e^2 \boldsymbol{I}_{260}$ where $\boldsymbol{A}$ is the numerator relationship matrix for the 260 entries. Note that in the form given here, the inverse of the numerator relationship matrix, ie. $\boldsymbol{A}^{-1}$ is supplied rather than $\boldsymbol{A}$ as it has been stored in sparse form (see Butler, 2013). This specification for the total genetic effects will be termed the direct formulation and is a new feature of OD-V2 (Butler & Cullis, 2018) that affords computational savings over the alternative, indirect formulation. The latter

involves the inclusion of two terms in the model, associated with the additive effects ($\boldsymbol{u_a}$) and non-additive effects ($\boldsymbol{u_e}$), so that the resultant vector of objective effects ($\boldsymbol{\beta_o}$) has twice as many elements as the direct formulation. The other terms in the `random` model formula are associated with random `CRep,` `Column` and `Row` effects respectively. Lastly the `residual` model formula specifies that the variance model for the residuals is one associated with a separable first order autoregressive process. The `permute` model formula specifies that `Entry` is the factor to be permuted and the use of the direct formulation means that the $\mathcal{A}$-value will be computed for total genetic effects. The (bespoke) variance parameters are supplied using the `G.param` and `R.param` arguments (also see the Appendix).

3.5.2 Timings for updating schemes

Table 5 presents the elapsed time (seconds) for a full tabu search and 1000 interchanges for OD-V2 using the direct formulation for total genetic effects; 1000 interchanges for OD-V2 using the indirect formulation for total genetic effects and 1000 interchanges for OD-V1 using the indirect formulation for total genetic effects. The timings were conducted on 8 trials used in the case study (trials numbered 3, 5, 15, 17, 22, 23, 27 and 31 in table 7). The trials included an S1 and S2 stage trial for each of the four crops. The results illustrate the substantial reduction in computational load by using the updating formula developed in section 3.4 and approximately agree with the total number of FLOPs involved, that is, $O(c_{w_o}^3)$ without an updating scheme compared with $O(c_{w_o})$ with the updating scheme.

		tabu	V2 Direct		V2 Indirect	V1 Indirect
Trial	Entries	interchanges	full tabu	1000	1000	1000
22	797	460320	1887	4.101	22.731	25000
27	788	460320	1573	3.418	20.892	21250
31	499	179700	262	1.458	10.976	9050
23	409	352380	457	1.297	10.837	6500
3	389	109278	116	1.063	8.695	4080
15	255	64620	27	0.432	0.717	150
5	144	41328	29	0.703	1.118	230
17	109	28680	10	0.350	0.120	50

Table 5: Timings for designs for 8 trials (trial numbers correspond to those in table 2): elapsed time (seconds) for a full tabu loop (number of interchanges given in column 3) and for 1000 interchanges for OD-V2 using the direct formulation for total genetic effects; 1000 interchanges for OD-V2 using the indirect formulation for total genetic effects and 1000 interchanges for OD-V1 using the indirect formulation for total genetic effects. Table is ordered on the number of entries.

4 Accuracy of designs using genetic relatedness

The performance of model-based designs constructed using genetic relatedness and bespoke non-genetic models was assessed using an in silico experiment based on the example presented in section 3.5.1. Of particular interest were comparisons with model-based designs

using bespoke non-genetic models but no information on genetic relatedness and with more classical approaches to design which use neither bespoke non-genetic models nor genetic relatedness. Thus there were three key treatments under investigation, corresponding to three design functions (DF), labelled as DF++, DF−+ and DF−− where the two final characters indicate the use/nonuse of genetic relatedness and bespoke non-genetic models in the design construction. Further details are provided later.

The simulation experiment was also designed to cover the range of genetic variance parameters and levels of replication in current use for S1 and S2 stage trials in the pulse breeding programs described in section 2. This led to the construction of a total of 63 ancillary treatments to consider, comprising the factorial combinations of the three factors described in table 6. These factors are labelled `Prep` (the percentage of breeding lines which have two plots per trial), `Padd` (the proportion of additive genetic variance expressed as a ratio of the total genetic variance) and `Rsq` (the baseline reliability of the data for the design with level 0 of the Prep factor).

Factor	Description	Levels
`Prep`	Percentage of breeding lines replicated	$p = 0, 5, 10, 15, 25, 50, 100$
`Padd`	Proportion of additive genetic variance	$k = 0.5, 0.7, 0.9$
`Rsq`	Baseline reliability	$r_0^2 = 1/3, 1/2, 2/3$
`T`	`Prep`∧`Padd`∧`Rsq`	$63 = 7 \times 3 \times 3$

Table 6: Summary of the ancillary treatment factors and their levels used in the simulation experiment.

Table 7 presents summaries of the layouts for the seven basic field trial configurations arising from the different replication levels used in the simulation experiment. Every trial was assumed to be laid out in a contiguous rectangular array with 12 columns and between 22 and 28 rows. The blocking factor `CRep` was constructed to span two sets of contiguous columns in each case.

Prep	Plots	Cols	Rows	Tworep
0	264	12	22	4
5	276	12	23	16
10	288	12	24	28
15	300	12	25	40
25	324	12	27	64
50	392	12	28	132
100	520	12	27	260

Table 7: Summary of the field trial layouts used in the simulation experiment. The final column (Tworep) provides the number of entries, including the four check varieties, which had two plots per trial.

The linear mixed model chosen to generate the non-genetic effects for the data-sets used in this study included terms, variance models and variance parameters as determined from the

historic analyses discussed in section 3.5.1. The linear mixed model also partitioned the total genetic effects into additive and non-additive effects. Given the set of variance parameters for the non-genetic effects, the two genetic variance parameters in the data generation model were then chosen to be consistent with the levels of the two ancillary treatment factors `Padd` and `Rsq`. Table 8 presents the variance parameters for all of the random and residual variance models in the data generation linear mixed model for each level of `Padd` and `Rsq`.

		Random terms			Residual			Genetic	
Padd	Rsq	CRep	Column	Row	Var	Col AR	Row AR	Additive	Non-Additive
0.5	1/3	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.00380	0.00738
0.5	1/2	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.01100	0.02130
0.5	2/3	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.02810	0.05450
0.7	1/3	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.00315	0.00262
0.7	1/2	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.01160	0.00969
0.7	2/3	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.03400	0.02830
0.9	1/3	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.00195	0.00042
0.9	1/2	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.00980	0.00211
0.9	2/3	0.0336	0.0183	0.00749	0.0682	0.214	0.392	0.03520	0.00759

Table 8: Variance parameters for the data generation linear mixed models in the simulation experiment. All parameters are variance components except for the residual autoregressive correlations (Col AR and Row AR).

Data-sets were generated according to the linear mixed model given by

$$\boldsymbol{y}_{ij}^{[k]} = \boldsymbol{X}_i \boldsymbol{\tau}_i + \boldsymbol{Z}_{ij} \boldsymbol{u}_{\boldsymbol{a},i}^{[k]} + \boldsymbol{Z}_{ij} \boldsymbol{u}_{\boldsymbol{e},i}^{[k]} + \boldsymbol{\eta}_i^{[k]} \tag{16}$$

where $i = 1, \ldots, 63$ indexes the ancillary treatments, $j = 1, \ldots, 3$ indexes the design functions and $k = 1, \ldots, N$ indexes the simulations ($N = 4000$). For each k $(= 1, \ldots, N)$, $\boldsymbol{u}_{\boldsymbol{a},i}^{[k]}$ and $\boldsymbol{u}_{\boldsymbol{e},i}^{[k]}$ are sampled from a normal distribution with mean zero and variance matrices given by $\sigma_{a,i}^2 \boldsymbol{A}$ and $\sigma_{e,i}^2 \boldsymbol{I}_{260}$ respectively. Note that the genetic variance components, although indexed by the level of `T` are constant on the levels of `Padd` and `Rsq`, with only nine unique combinations as listed in table 8. The vector $\boldsymbol{\eta}_i^{[k]}, i = 1, \ldots, 63; k = 1, \ldots, N$ represents the plot effects, that is the sum of all the random non-genetic effects and residuals. These are obtained via sampling from a normal distribution with mean zero and variance matrix given by

$$\sigma_{cr}^2 \boldsymbol{Z}_{\boldsymbol{cr},i} \boldsymbol{Z}_{\boldsymbol{cr},i}^\top + \sigma_c^2 \boldsymbol{Z}_{\boldsymbol{c},i} \boldsymbol{Z}_{\boldsymbol{c},i}^\top + \sigma_r^2 \boldsymbol{Z}_{\boldsymbol{r},i} \boldsymbol{Z}_{\boldsymbol{r},i}^\top + \sigma^2 \boldsymbol{\Sigma}_{\boldsymbol{c}i}(\rho_c) \otimes \boldsymbol{\Sigma}_{\boldsymbol{r}i}(\rho_r)$$

where $\boldsymbol{\Sigma}_{\boldsymbol{c}i}$ and $\boldsymbol{\Sigma}_{\boldsymbol{r}i}$ are scaled variance matrices of order c_i and r_i for first order autoregressive processes where c_i and r_i are the numbers of columns and rows in the trial layout. These matrices are functions of autocorrelation parameters, ρ_c and ρ_r, for the column and row dimensions respectively. The design matrices for the fixed and non-genetic random effects are given by $\boldsymbol{X}_i, \boldsymbol{Z}_{\boldsymbol{cr},i}, \boldsymbol{Z}_{\boldsymbol{c},i}$ and $\boldsymbol{Z}_{\boldsymbol{r},i}$ where the latter correspond to the terms in the random model formula, namely `CRep, Column` and `Row` respectively. Values for all the variance parameters are set out in table 8. There is only one fixed effect associated with an overall mean. Lastly, we consider the design matrix for the genetic effects. A separate design matrix, $\boldsymbol{Z}_{ij}, i = 1, \ldots, 63; j = 1, \ldots, 3$ is formed for each level of the design function factor as described below

DF$_{++}$ a design function generated from (16)

DF$_{-+}$ a design function generated from the plot effects component of (16) but with genetic effects assumed to be fixed effects

DF$_{--}$ a design function generated from the model of Williams et al. (2011) with genetic effects assumed to be fixed effects

Hence there was a total of 63 different designs for DF$_{++}$ but only seven designs for DF$_{-+}$ and DF$_{--}$, one for each level of `Prep`. When forming the design function for the DF$_{++}$ design, for levels $5, 10, 15, 25, 50$ of `Prep` we used the method of Huang et al. (2013) to subset those breeding lines to be tested with two plots in the trial.

All design functions were obtained using **OD-V2** (Butler & Cullis, 2018), by varying the model formulae for the fixed, random and residual components. The **OD-V2** call for DF$_{++}$ designs was as given in section 3.5.1. Details and scripts for all designs are available in the supplementary materials.

The correlations between the true and predicted total genetic effects for all entries were calculated for each ancillary treatment and design function, (a total of 189 combinations), where the predicted genetic effects were the empirical BLUPs obtained from the fitting of the linear mixed model that corresponded to the data generation model.

4.1   Simulation study results

The estimation of the variance parameters is considered first. Summaries of the MSEP, simulation variance and bias (expressed as a % of the true value) for all variance parameters in the model are available in the supplementary materials. Only the results for the additive and total genetic variance parameters are presented here. Figures 2 and 3 present plots of the percentage bias for the additive genetic variance and total genetic variance against the levels of `Prep`. Significant bias occurs for some combinations of k and r_0^2. There is consistent, negative bias for $k = 0.9$ for the additive genetic variance, for all levels of `Prep` and `Rsq`. This negative bias is associated with a significant positive bias for the non-additive genetic variance for these combinations, which then results in bias for the total genetic variance, especially for $k = 0.9$ and $r_0^2 = 1/3$. Bias for the total genetic variance is acceptable for the four `Padd` and `Rsq` combinations of $(k = 0.5, r_0^2 = 1/2)$, $(k = 0.5, r_0^2 = 2/3)$, $(k = 0.7, r_0^2 = 1/2)$ and $(k = 0.7, r_0^2 = 2/3)$ for all levels of `Prep`. The remaining combinations of `Padd` and `Rsq` all exhibit unacceptable bias for total genetic variance, particularly for small values of `Prep`. Bias for the additive genetic variance is acceptable for all levels of `Prep` except when $k = 0.9$. This latter bias is largely a result of the significant positive bias in estimation of the non-additive genetic variance, when it is such a small component of the total genetic variance and is so close to zero (relative to the residual variance).

Table 9 presents the mean squared error of prediction (MSEP-CV) for the REML estimate of the additive genetic variance expresssed as a coefficient of variation (relative to the true value) for each level of replication of the test lines, the proportion of additive genetic variance as a ratio to the total genetic variance and for baseline heritabilities of $r_0^2 = 1/3$ and $r_0^2 = 1/2$. The results for $r_0^2 = 2/3$ are similar and are omitted. The MSEP-CV for the DF$_{++}$ design function is consistently lower than the MSEP-CV for the other two design functions. Of
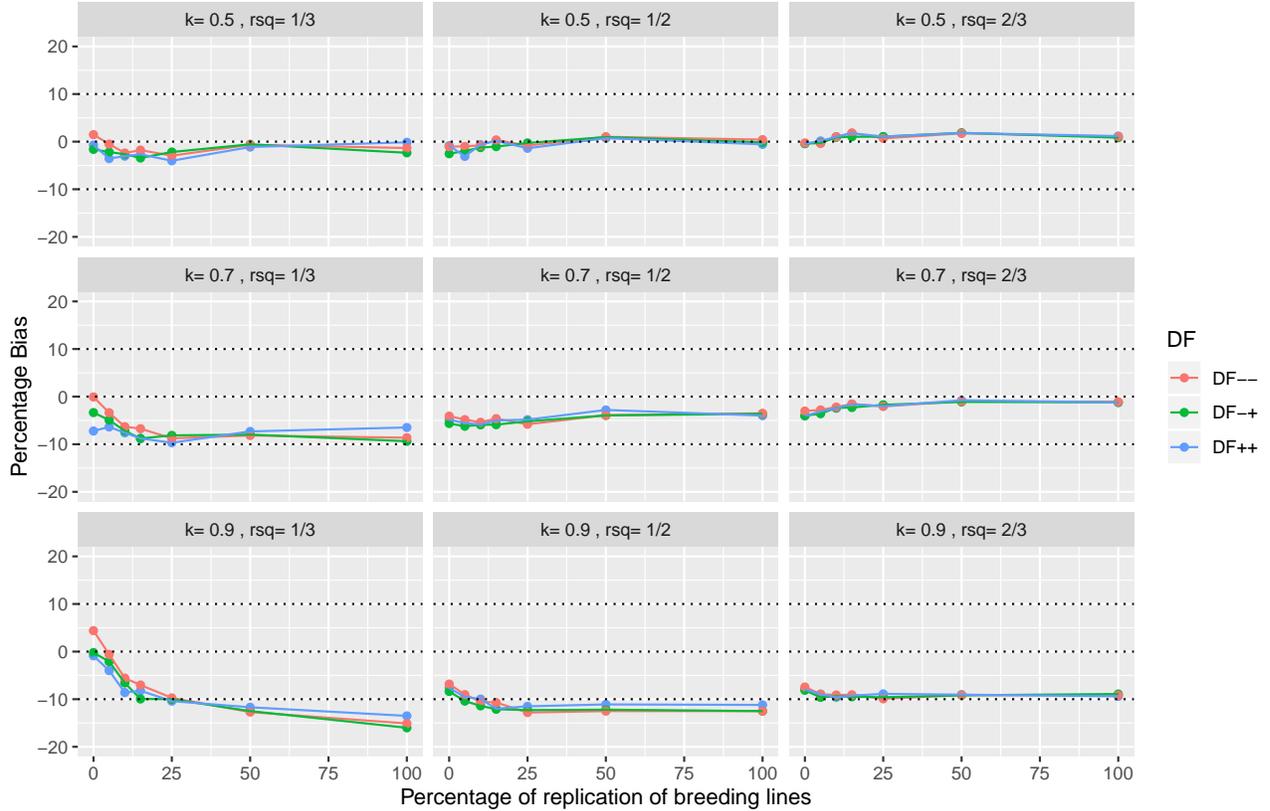
Figure 2: Plot of the percentage bias for the additive genetic variance versus percentage of replication of test lines. Each panel represents the combination of baseline heritability and proportion of additive genetic variance as a ratio to the total genetic variance. Superimposed dotted lines indicate $+/- 10\%$.

particular interest, is the ability of the DF_{++} design function to achieve comparable levels of MSEP-CV achieved by DF_{-+} and DF_{--} for much lower values of `Prep`. For example, the MSEP-CV for the design functions DF_{--} and DF_{-+} for a `Prep` of 25% is approximately the same as the MSEP-CV for the DF_{++} design function for a `Prep` of only 10%.

The correlations between the true and predicted entry effects are summarised in Figures 4 and 5 for $r_0^2 = 1/3$ and $r_0^2 = 1/2$, respectively. Each figure presents panel line plots of the mean simulation based correlations (in blue) and model-based reliability (in black) versus the levels of `Prep` for each of the three design functions. The results for $r_0^2 = 2/3$ are similar, with the DF_{++} design function superior to the other design functions for all levels of `Prep` and `Padd`. The model-based reliability was derived from the $\mathcal{A}$-values obtained from either the final $\mathcal{A}$-value for the DF_{++} design function, or the initial $\mathcal{A}$-values for the DF_{-+} and DF_{--} design functions evaluated under the (true) data model of (16), using the approximation of Cullis et al. (2006).

It is immediately clear that there is a consistent and important increase in correlation for DF_{++} designs relative to both DF_{-+} and DF_{--} designs. The advantage is larger for higher values of k and lower values of r_0^2. The DF_{++} designs are superior for all levels of `Prep`, though the advantage diminishes for designs with `Prep` less than 10%. It is also noteworthy
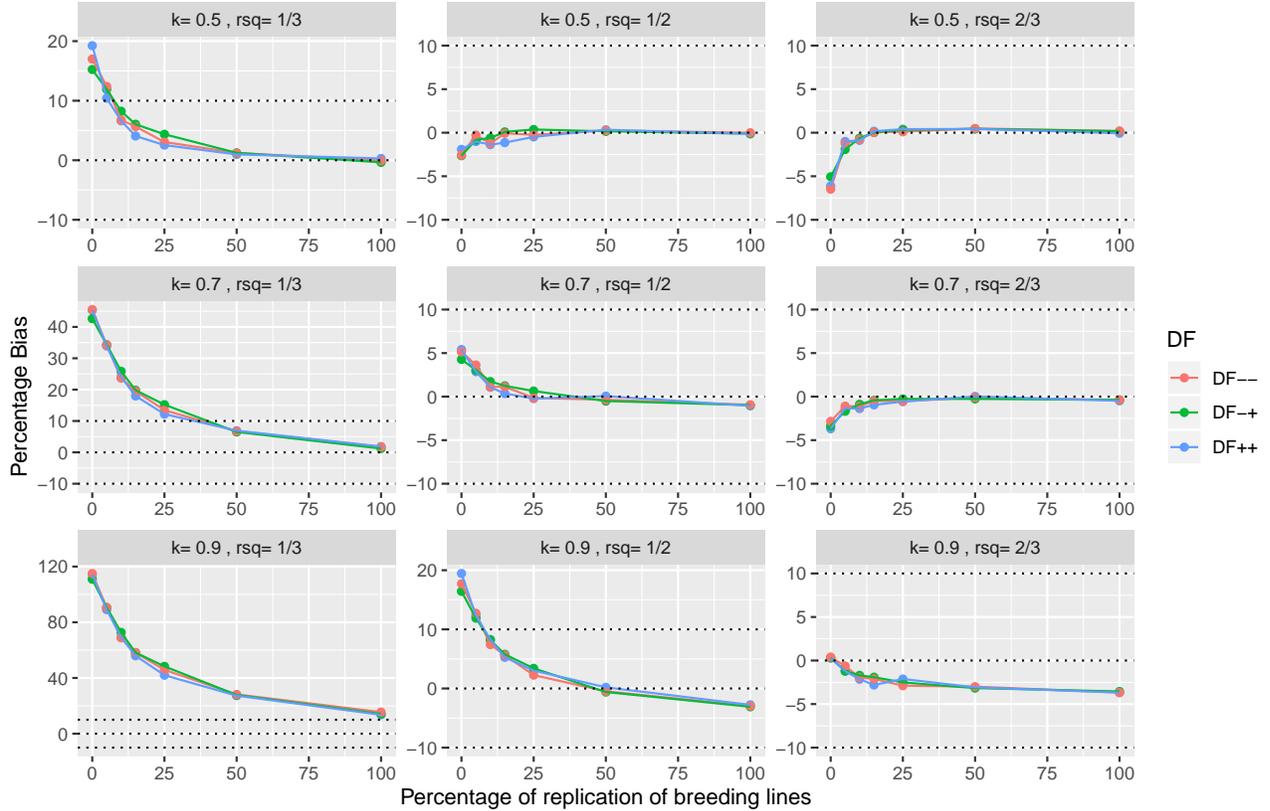
Figure 3: Plot of the percentage bias for the total genetic variance versus percentage of replication of test lines. Each panel represents the combination of baseline heritability and proportion of additive genetic variance as a ratio to the total genetic variance. Superimposed dotted lines indicate $+/- 10\%$.

that the advantage of DF$_{++}$ designs for values of `Prep` greater than 25% is stable, suggesting there is little impact of using the subset selection method of Huang et al. (2013) in terms of improving the accuracy of the design.

The plots also highlight the interaction between the increase in simulation based correlation and model-based reliability due to increasing the level of replication and the value of k. This suggests that the information coming from the data plays an important role in improving the accuracy of the prediction of total genetic effects when the non-additive component is present (and can be reliably estimated). The impact of variance parameter estimation on accuracy is clear. All simulation based correlations are significantly lower than model-based reliabilities, but particularly so for designs with values of `Prep` less than 10%. This is consistent with the MSEP-CV shown in table 9.

One final remark is that across all levels of k, r_0^2 and `Prep` there is no measurable difference between design functions DF$_{-+}$ and DF$_{--}$. This may be a result of the choice of variance parameters for the residual process which gave rise to only modest levels of spatial dependence arising from the inclusion of spatial dependence.

Padd	Prep	$r_0^2 = 1/3$			$r_0^2 = 1/2$		
		DF++	DF−+	DF−−	DF++	DF−+	DF−−
0.5	0	88.5	93.4	96.3	68.6	70.6	71.9
0.5	5	83.6	90.0	91.2	65.4	69.1	70.0
0.5	10	84.0	86.9	87.7	67.3	68.9	68.9
0.5	15	81.8	87.7	86.1	65.0	68.8	68.6
0.5	25	78.7	85.7	84.0	65.2	68.6	67.3
0.5	50	75.8	80.2	78.7	62.9	65.7	65.0
0.5	100	71.5	71.7	73.5	60.8	60.8	61.9
0.7	0	85.5	94.1	97.0	59.9	62.1	62.8
0.7	5	83.3	90.1	90.6	58.1	60.3	61.1
0.7	10	80.4	84.7	85.7	56.5	58.8	58.5
0.7	15	77.4	85.1	83.9	55.7	58.1	58.0
0.7	25	76.1	81.6	80.2	54.9	57.3	55.7
0.7	50	68.7	74.5	73.2	52.0	53.7	53.6
0.7	100	62.9	64.6	66.5	49.7	50.0	50.4
0.9	0	105.0	112.8	117.3	56.9	59.8	60.4
0.9	5	98.2	108.3	107.2	55.3	57.3	57.6
0.9	10	90.4	99.2	100.8	53.3	54.4	54.5
0.9	15	89.3	98.4	97.6	51.0	54.0	53.8
0.9	25	84.4	92.7	91.3	49.4	52.0	50.6
0.9	50	75.9	83.8	81.7	45.6	47.4	46.9
0.9	100	66.7	70.5	72.9	40.4	41.7	42.2

Table 9: Mean squared error of prediction for the REML estimate of the additive genetic variance expresssed as a coefficient of variation (relative to the true value) for each design function, level of replication of the test lines, the proportion of additive genetic variance as a ratio to the total genetic variance and for baseline heritabilities of 1/3 and 1/2.
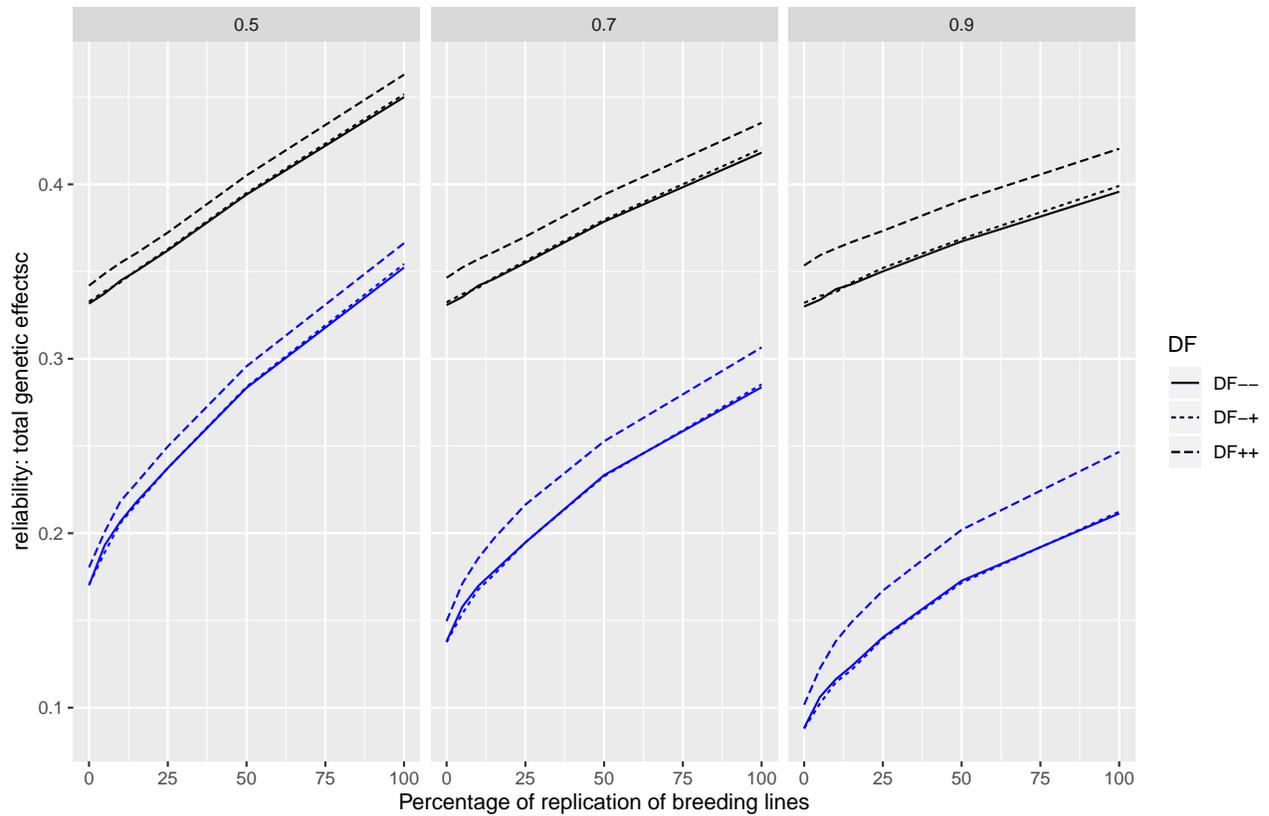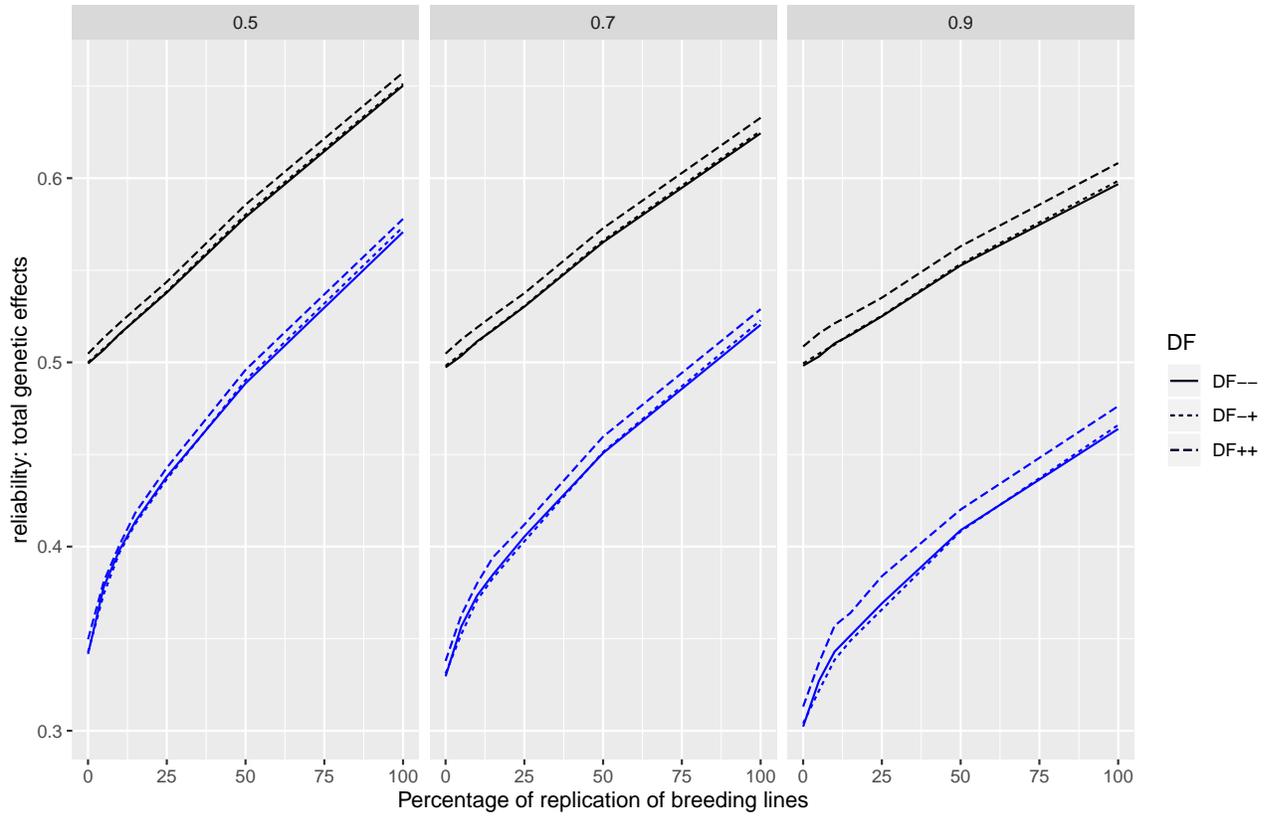
Figure 4: Line plot of the mean simulation based correlation (in blue) and model-based reliability (in black) versus the percentage replication of breeding lines for baseline heritability of 1/3. Each panel represents a different proportion of additive genetic variance as a ratio to the total genetic variance.

Figure 5: Line plot of the mean simulation based correlation (in blue) and model-based reliability (in black) versus the percentage replication of breeding lines for baseline heritability of 1/2. Each panel represents a different proportion of additive genetic variance as a ratio to the total genetic variance.

A OD timing syntax

Below are the calls used to obtain the timings in table 5. The first sequence of calls provides an estimate of the elapsed time to complete a full tabu search of the design space and 1000 interchanges using OD-V2 and the direct formulation for the total genetic effects. The function `od.init` sets or displays various options that affect the behaviour of OD-V2. The first call to OD returns a template from which variance parameters for each of the random and residual terms can be specified. These are assigned in the subsequent two lines of the code and these are supplied in the call to OD-V2 via the `R.param` and `G.param` arguments. The `permute` argument specifies that the column labelled `Entry`, which contains the entry code, is to be permuted and the optimality criteria is for the total genetic effects for those entries with data. Use of the `time` argument provides an estimate of the time for a single tabu loop based on a set of random interchanges. If `time> 0`, the objective criterion is evaluated the specified number of times prior to the optimisation step. The elapsed time is reported and used to estimate the execution time for a scavenging tabu loop.

```
sv <- od(fixed=~ 1,
         random=~ ric(Entry, Ainv) + RRep  + Column,
         residual = ~ar1(Column):ar1(Row),
         permute=~ ric(Entry, Ainv),
         start.values = TRUE, data=dat)
sv <- sv$vparameters.table
sv$Value <- c(sigma.vmg,sigma.ideg, sigma.rrep, sigma.col,
              sigma,colphi,rowphi)
od.options(time=1e3)
temp.od <- od(fixed=~ 1,
              random=~ ric(Entry, Ainv) + RRep  + Column,
              residual = ~ar1(Column):ar1(Row),
              permute=~ ric(Entry, Ainv),
              R.param = sv, G.param = sv,
              search = 'tabu',maxit=1,data=dat)
```

The next sequence of calls estimates the elapsed time for 1000 random interchanges using the indirect model, but in this call to OD-V2 the optimality criteria is based on the additive genetic effects. OD-V2 does not allow the user to specify the calculation of the optimality criteria for the total genetic effects using the indirect formulation for total effects. The pipe operator (|) is used to associate the permute factor with other terms in the model that should be permuted in parallel. The performance of the indirect formulation for the total genetic effects was found to be substantially inferior to the direct formulation, and hence it has been deprecated. Therefore the elapsed times using these calls will not be exactly equivalent to the use of the indirect formulation within the updating scheme, but trials suggest that they are adequate to ascertain the approximate penalty associated with use of the indirect formulation, which is the only form available in OD-V1 for total genetic effects.

```
sv <- od(fixed=~ 1,
         random=~ vm(Entry, Ainv) + ide(Entry) +
```

```
            RRep  + Column,
            residual = ~ar1(Column):ar1(Row),
            permute=~ vm(Entry, Ainv)|ide(Entry),
            start.values = TRUE, data=dat)
sv <- sv$vparameters.table
sv$Value <- c(sigma.vmg,sigma.ideg, sigma.rrep, sigma.col,
              sigma,colphi,rowphi)
od.options(time=1e3)
temp.od <- od(fixed=~ 1, random=~ vm(Entry, Ainv) + ide(Entry) +
              RRep + Column,
              residual = ~ar1(Column):ar1(Row),
              permute=~ vm(Entry, Ainv) | ide(Entry),
              R.param = sv, G.param = sv,
              search = 'tabu',maxit=1,data=dat)
```

The final sequence of calls using **OD-V1** provide the elapsed times for the indirect formulation
for the total genetic effects. There is no option to evaluate timings within the call to **OD-V1**
and so crude estimates of elapsed time were estimated from a sequence of two calls to isolate
the time for evaluating the optimality criteria for 1000 random interchanges.

```
sv <- od.init('Column'=sigma.col,
              'RRep'=sigma.rrep,'ped(Entry)'=sigma.vmg,
              'ide(Entry)'=sigma.ideg,
              'Column:Row|Column' = colphi,
              'Column:Row|Row' = rowphi, 'rcov' = sigma)
date()
temp.od <- od(fixed=~ 1,
              random=~ ped(Entry) + ide(Entry) + RRep  + Column,
              rcov = ~ar1(Column):ar1(Row), ginv = list(Entry=Ainv),
              geneticEffects="total", permute=~ ped(Entry)|ide(Entry),
              Rstart=sv, Gstart=sv,
              search = 'random',maxit=1,data=dat)
date()
temp.od <- od(fixed=~ 1,
              random=~ ped(Entry) + ide(Entry) + RRep  + Column,
              rcov = ~ar1(Column):ar1(Row), ginv = list(Entry=Ainv),
              geneticEffects="total", permute=~ ped(Entry)|ide(Entry),
              Rstart=sv, Gstart=sv,
              search = 'random',maxit=1001,data=dat)
date()
```

References

BAILEY, R. A. (2008). *Design of comparative experiments.* Cambridge University Press.

BUENO FILHO, J. S. D. S. & GILMOUR, S. G. (2007). Block designs for random treatment effects. *Journal of Statistical Planning and Inference* **137**, 1446–1451.

BUENO FILHO, J. S. S. & GILMOUR, S. G. (2003). Planning incomplete block experiments when treatments are genetically related. *Biometrics* **59**, 375–381.

BUTLER, D. & CULLIS, B. (2018). Optimal Design under the Linear Mixed Model: https://mmade.org/optimaldesign/. Technical report, National Institute for Applied Statistics Research Australia, University of Wollongong.

BUTLER, D., ECCLESTON, J., & CULLIS, B. (2008). On an approximate optimality criterion for the design of field experiments under spatial dependence. *Australian and New Zealand Journal of Statistics* **50**.

BUTLER, D. G. (2013). *On The Optimal Design of Experiments Under the Linear Mixed Model.* PhD thesis, University of Queensland.

BUTLER, D. G., CULLIS, B. R., GILMOUR, A. R., & THOMPSON, R. (2018). ASReml version 4. Technical report, University of Wollongong.

BUTLER, D. G., SMITH, A. B., & CULLIS, B. R. (2014). On the Design of Field Experiments with Correlated Treatment Effects. *Journal of Agricultural, Biological, and Environmental Statistics* **19**, 539–555.

CHAN, B. (1999). *The design of field experiments when the data are spatially correlated.* PhD thesis, The University of Queensland.

COOMBES, N. E. (2002). *The reactive tabu search for efficient correlated experimental designs.* PhD thesis, Liverpool John Moores University, Liverpool, U.K.

COOMBES, N. E. (2009). DiGGeR, a Spatial Design Program. Biometric bulletin, NSW DPI.

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**, 381–393.

DUNNER, S., CHECA, M. L., GUTIERREZ, J. P., MARTIN, J. P., & CANON, J. (1998). Genetic analysis and management in small populations: The asturcon pony as an example. *Genetics Selection Evolution* **30**, 397–405.

GILMOUR, A., CULLIS, B., WELHAM, S., & GOGEL, B. (2004). An efficient computing strategy for prediction in mixed linear models. *Computational Statistics & Data Analysis* **44**, 571–586.

GILMOUR, A. R., CULLIS, B. R., VERBYLA, A. P., & 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.

GLOVER, F. (1989). Tabu search–Part 1. *Journal on Computing* **1**, 190–206.

HARVILLE, D. A. (1997). *Matrix algebra from a statisticians perspective.* Springer-Verlag, New York.

HUANG, B. E., CLIFFORD, D., & CAVANAGH, C. (2013). Selecting subsets of genotyped experimental populations for phenotyping to maximise genetic diversity. *Theoretical and Applied Genetics* **126**, 379–388.

JOHN, J. A. & WILLIAMS, E. R. (1995). *Experimental designs.* Wiley and Sons.

JOHN, J. A. & WILLIAMS, E. R. (1998). t-Latinized designs. *Australian and New Zealand Journal of Statistics* **40**, 111–118.

KEMPTON, R. A. (1982). The design and analysis of unreplicated field trials. *Vortrage fur Pflanzenzuchtung* **7**, 219–242.

MARTIN, R. & ECCLESTON, J. (1997). Construction of optimal and near-optimal designs for dependent observations using simulated annealing. Technical report, Dept. Prob. Statist, Unversity of Sheffield.

MARTIN, R. J. (1986). On the design of experiments under spatial correlation. *Biometrika* **73**, 247–277.

MARTIN, R. J., CHAUHAN, N., ECCLESTON, J. A., & CHAN, B. S. P. (2006). Efficient experimental designs when most treatments are unreplicated. *Linear Algebra and Its Applications* **417**, 163–182.

MARTIN, R. J. & ECCLESTON, J. A. (1992). Recursive formulae for constructing block designs with dependent errors. *Biometrika* **79**, 426–430.

MEUWISSEN, T. H. E. & LUO, Z. (1992). Computing inbreeding coefficients in large populations. *Genetics, Selection and Evolution* **24**, 305–315.

MULITZE, D. (1990). AGROBASE/4: A microcomputer database management and analysis system for plant breeding and agronomy. *Agronomy Journal* **82**, 1016—-1021.

NELDER, J. A. (1977). A reformulation of linear models. *Journal of the Royal Statistical Society, Series A* **140**, 48–76.

OAKEY, H., VERBYLA, A., CULLIS, B., WEI, X., & PITCHFORD, W. (2007). Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. *Theoretical and Applied Genetics* **114**.

OAKEY, H., VERBYLA, A. P., CULLIS, B. R., PITCHFORD, W. S., & KUCHEL, H. (2006). Joint modelling of additive and non-additive genetic line effects in single field trials. *Theoretical and Applied Genetics* **113**, 809–819.

PATTERSON, H. D. & WILLIAMS, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika* **63**, 83–92.

Piepho, H.-P. & Williams, E. R. (2006). A comparison of experimental designs for selection in breeding trials with nested treatment structure. *Theoretical Applied Genetics.* **113**, 1505–1513.

Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science* **6**, 15–51.

Searle, S. R. (1982). *Matrix algebra useful for Statistics.* Wiley Inter Science.

Smith, A., Cullis, B. R., & Thompson, R. (2001). Analysing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* **57**, 1138–1147.

Smith, A. B. & Cullis, B. R. (2018). Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. *Euphytica* **214**, 1–19.

Williams, E., Piepho, H. P., & Whitaker, D. (2011). Augmented p-rep designs. *Biometrical Journal* **53**, 19–27.

Williams, E. R. & John, J. A. (2006). Row-column factorial designs for use in agricultural field trials. *Applied Statistics* **62**, 103–108.