The Usefulness of Optimal Design for Generating Blocked and Split-Plot Response Surface Experiments

Peter Goos

Universiteit Antwerpen

Abstract

This article provides an overview of the recent literature on the design of blocked and split-plot experiments with quantitative experimental variables. A detailed literature study introduces the ongoing debate between an optimal design approach to constructing blocked and split-plot designs and approaches where the equivalence of ordinary least squares and generalized least squares estimates are envisaged. Examples where the competing approaches lead to totally different designs are given, as well as examples in which the proposed designs satisfy both camps.

Keywords: equivalence of OLS and GLS, ordinary least squares, generalized least squares, orthogonal blocking, D-optimal design

1 Introduction

Blocked experiments are typically used when not all the runs of an experiment can be conducted under homogeneous circumstances. In such experiments, the runs are partitioned in blocks so that the within-block variability is considerably smaller than the between-block variability. Split-plot experiments are used for a different reason, that is because the levels of some of the experimental variables are not independently reset for every experimental run as this is often impractical, costly or time-consuming. As a consequence, the runs of a split-plot experiment are also partitioned in groups, and within each group the levels of the so-called hard-to-change variables are constant. Split-plot experiments can thus be interpreted as special cases of blocked experiments. Despite the similarity between the two types of experiments, the terminology utilized in the context of blocked and split-plot experiments is different: the groups of observations in a blocked experiment are called blocks, whereas those in split-plot experiments are named whole plots. The hard-to-change variables in a split-plot experiment are often called whole-plot variables, whereas the remaining (easier-to-change) variables are referred to as sub-plot variables.

The design and analysis of blocked and split-plot experiments has a long, initially agricultural, history. Originally, the focus was on experiments involving qualitative experimental factors. The focus in this article is on continuous factors, although it should be stressed that the optimal design approach can easily handle categorical factors as well.

1.1 Blocked experiments

The problem of arranging response surface designs (sometimes referred to as regression designs) in blocks did not receive attention until Box and Hunter (1957) established conditions for orthogonal blocking in the case of one blocking variable, the effects of which are modelled as fixed effects. Conditions for the case of one blocking variable the effects of which are treated as random were discussed by Khuri (1992) and those for the case of more than one crossed blocking variable, the effects of which are treated as random and/or fixed were given by Goos and Donev (2005a). The special case of orthogonally blocked mixture experiments has received attention by Nigam (1976) and John (1984).

For orthogonally blocked designs, the same estimates for the effects of the experimental variables are obtained when ignoring the blocks, when modelling them as fixed and when modelling them as random (see, for example, Khuri (1992) and Goos and Vandebroek (2003b) for the case of one blocking factor). Another way to put this result is to say that the intra-block estimator is equivalent to the combined inter- and intra-block estimator. The former estimator uses only within-block information and is obtained by treating the block effects as fixed, whereas the latter is obtained by assuming the block effects are random and also uses between-block information. The combined inter- and intra-block estimator is thus the generalized least squares estimator and the corresponding analysis is called mixed model analysis. In this article, we will focus on the situation in which the block effects are assumed random because the case of fixed block effects can be seen as a limiting case of that in which the block effects are treated as random (see, for example, Gilmour and Trinca (2000), Goos and Vandebroek (2001a) and Grasshoff and Schwabe (2003)).

The optimal design of blocked experiments first received attention by Atkinson and Donev (1989) and Cook and Nachtsheim (1989), who concentrated on experiments with one blocking factor that generates fixed block effects. Their work was later extended to problems involving random block effects by Goos and Vandebroek (2001a) and to cases with more than one blocking factor by Goos and Donev (2005a). The problem of designing blocked mixture experiments has been discussed by Donev (1989) and Goos and Donev (2006a, 2006b).

1.2 Split-plot Experiments

The problem of designing split-plot experiments in industry first received attention in Anbari and Lucas (1994), who discussed the arrangement of factorial designs in a split-plot format, Letsinger, Myers and Lentner (1996), who investigated the efficiency of various second-order designs when run as a split-plot experiment, and Draper and John (1998), who discuss modifications of central composite designs and Box-Behnken designs to be run in a split-plot format. A sequential strategy for designing multi-stratum designs, special cases of which are split-plot designs, is presented by Trinca and Gilmour (2001). The optimal design of first- and second-order split-plot experiments later received attention by Goos and Vandebroek (2001b, 2003a, 2004). The optimal design of split-plot experiments for spherical design regions receives special attention in Mee (2005). Standard experimental designs for split-plot mixture-process variable designs are proposed by Kowalski, Cornell and Vining (2002). Optimal designs for this type of experiment are reported in Goos and Donev (2005b).

A recent line of articles focuses on the arrangement of standard response surface designs like central composite and Box-Behnken designs in a split-plot format so that ordinary least squares estimation provides the same estimates as generalized least squares estimation (see Vining, Kowalski and Montgomery (2005) and Parker, Kowalski and Vining (2005a, 2005b)). A key feature of some of these so-called equivalent-estimation split-plot designs is that they allow a model-independent estimation of the variance components needed for statistical inference.

1.3 Pros and cons of orthogonally blocked and equivalent-estimation designs

The main advantage of orthogonally blocked and equivalent-estimation designs is that knowledge or estimation of the model's variance components are not required for estimating the factor effects. This is because ordinary and generalized least squares provide the same estimates. This equivalence is especially useful for researchers who have no access to software that allows mixed model estimation. This advantage is, however, only minor as, despite the equivalence of ordinary and generalized least squares, estimation of the variance components remains necessary for statistical inference.

For the estimation of the variance components, there are several options and the options available depend on the experimental design. The now widely accepted restricted (or residual) maximum likelihood estimation (REML) is recommended for split-plot designs by Letsinger, Myers and Lentner (1996) as it yields good results for various relative magnitudes of the variance components even for smaller and second-order designs. For blocked experiments, restricted maximum likelihood is recommended by Gilmour and Trinca (2000), who also discuss an alternative approach for calculating estimates for the variance components. A substantial advantage of REML is that it is applicable to every possible blocked or split-plot experiment. A disadvantage is that the resulting estimates depend on the specified model. This led Kowalski, Cornell and Vining (2002) and Vining, Kowalski and Montgomery (2005) to suggesting a replication-based approach which yields model-independent estimates of the variance components. Evidently, this approach requires experimental designs with sufficient replication making that a substantial part of the experimental design is tailored to make the replication-based estimation of the variance components possible.

One of the goals of this article is to show that this is done at the expense of an inefficient estimation of the factor effects, which is the primary concern of the researcher. Another goal of the article is to illustrate that, even if no replication-based estimation of the variance components is envisaged, constructing orthogonally blocked and equivalent-estimation designs may lead to substantial losses in efficiency. The resulting imprecision of the estimates of course has a negative impact on the power for detecting active effects and leads to wider prediction intervals.

The remainder of this article is organized as follows. First, the linear statistical model traditionally used for analyzing data from blocked or split-plot experiments is introduced, as well as a number of design optimality criteria for evaluating alternative experimental designs. In Sections and 4 and 6, it is shown that D-optimal blocked and split-plot designs in some cases possess the property that ordinary least squares and generalized least squares are equivalent. Situations in which this result does not hold are investigated in Sections 5 and 7. It is shown that some of the orthogonally blocked and equivalent-estimation experimental designs presented in the literature perform poorly in terms of the estimation-based and prediction-based design optimality criteria introduced in Section 3. Finally, Section 8 shows that it may be beneficial to consider heterogeneous sizes of the blocks or the whole plots in an experiment. A discussion of the computational results concludes the paper.

2 Statistical model and estimation

The statistical model corresponding to a blocked or split-plot experiment can be written as

$$\boldsymbol{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\delta}, \tag{1}$$

where $\boldsymbol{Y}$ represents an n-dimensional vector of responses, $\mathbf{X}$ is the $n \times p$ model matrix, $\boldsymbol{\beta}$ is the p-dimensional vector of factor effects (possibly including an intercept) and $\boldsymbol{\delta}$ is a vector of random effects with zero mean and covariance matrix $\mathbf{V}$.

The nature of the covariance matrix $\mathbf{V}$ depends on the type of experimental data that is modelled. For the type of experiments considered here, the matrix $\mathbf{V}$ has the form

$$\mathbf{V} = \sum_{i=1}^{b} \sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i' + \sigma_\varepsilon^2 \mathbf{I}_n, \tag{2}$$

respectively. For all split-plot designs described in this article, the parameter b equals one, the number of whole plots is denoted by b_1 and $\mathbf{Z}_1$ is an $n \times b_1$ matrix of zeroes and ones, with (j,k)th element equal to one if the jth experimental run belongs to the ith whole plot, and zero otherwise. For blocked experiments, b represents the number of blocking factors and each of these has b_i different levels. For that type of experiment, the $\mathbf{Z}_i$ matrices are $n \times b_i$ matrices of zeroes and ones, with (j,k)th element equal to one if the jth experimental run is conducted at the kth level of the ith blocking variable, and zero otherwise. The parameters σ_i^2, $i = 1, \ldots, b$, are referred to as block-effect variances in the case the model is used for data from a blocked experiment. For split-plot experiments, where we assume b to be one, σ_1^2 is named the whole-plot error variance.

The researcher's primary interest when conducting a blocked or split-plot experiment is in estimating the factor effects contained within the vector $\boldsymbol{\beta}$. As the responses generated by blocked and split-plot experiments are assumed to be correlated, the best linear unbiased estimator of the factor effects is the generalized least squares estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{Y},$$

whose covariance matrix is given by

$$\mathrm{var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{3}$$

3 Design evaluation criteria

The optimal designs discussed in this article are so-called D-optimal designs, which means that they are constructed such that they minimize the determinant of the variance-covariance matrix (3). The relative D-efficiency of a design with model matrix $\mathbf{X}_1$ with respect to a design with model matrix $\mathbf{X}_2$ is expressed as

$$\left\{\frac{|(\mathbf{X}_2'\mathbf{V}_2^{-1}\mathbf{X}_2)^{-1}|}{|(\mathbf{X}_1'\mathbf{V}_1^{-1}\mathbf{X}_1)^{-1}|}\right\}^{1/p} = \left\{\frac{|\mathbf{X}_1'\mathbf{V}_1^{-1}\mathbf{X}_1|}{|\mathbf{X}_2'\mathbf{V}_2^{-1}\mathbf{X}_2|}\right\}^{1/p},$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ are the covariance matrices corresponding to $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively, and p is the dimension of $\boldsymbol{\beta}$. In this article, we initally focus on comparing experimental designs with the same covariance structure, so that $\mathbf{V}_1$ and $\mathbf{V}_2$ will usually be identical. In Section 8, however, we do consider design options with different covariance matrices $\mathbf{V}_1$ and $\mathbf{V}_2$.

We also compare the alternative design options in terms of A-, G- and V-optimality. The A-optimality criterion seeks to minimize the average variance of the parameter estimates and thus to minimize the trace of the covariance matrix (3). A G-optimal design minimizes the maximum prediction variance,

$$\max_{\mathbf{x}\in\chi} \ \mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{f}(\mathbf{x})$$

over the region of interest χ, while a V-optimal design minimizes the average prediction variance over that region:

$$\operatorname*{avg}_{\mathbf{x}\in\chi} \ \mathbf{f}'(\mathbf{x})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{f}(\mathbf{x}).$$

In these expressions, $\mathbf{f}(\mathbf{x})$ is a p-dimensional vector containing the model expansion of the point $\mathbf{x}$ at which a prediction is made. Relative A-, G- and V-efficiencies of two designs with model matrices $\mathbf{X}_1$ and $\mathbf{X}_2$, and corresponding covariance matrices $\mathbf{V}_1$ and $\mathbf{V}_2$, are then computed as

$$\frac{\text{trace}\,(\mathbf{X}_2'\mathbf{V}_2^{-1}\mathbf{X}_2)^{-1}}{\text{trace}\,(\mathbf{X}_1'\mathbf{V}_1^{-1}\mathbf{X}_1)^{-1}},$$

$$\frac{\max_{\mathbf{x}\in\chi}\ \mathbf{f}'(\mathbf{x})(\mathbf{X}_2'\mathbf{V}_2^{-1}\mathbf{X}_2)^{-1}\mathbf{f}(\mathbf{x})}{\max_{\mathbf{x}\in\chi}\ \mathbf{f}'(\mathbf{x})(\mathbf{X}_1'\mathbf{V}_1^{-1}\mathbf{X}_1)^{-1}\mathbf{f}(\mathbf{x})},$$

and

$$\frac{\operatorname{avg}_{\mathbf{x}\in\chi}\ \mathbf{f}'(\mathbf{x})(\mathbf{X}_2'\mathbf{V}_2^{-1}\mathbf{X}_2)^{-1}\mathbf{f}(\mathbf{x})}{\operatorname{avg}_{\mathbf{x}\in\chi}\ \mathbf{f}'(\mathbf{x})(\mathbf{X}_1'\mathbf{V}_1^{-1}\mathbf{X}_1)^{-1}\mathbf{f}(\mathbf{x})},$$

respectively.

The relative D-, A-, G- and V-efficiencies are defined in such a way that values larger than one indicate an improvement over the benchmark design. As the relative efficiencies depend on the relative magnitudes of the variance components σ_ε^2 and σ_i^2, $i = 1, \ldots, b$, through the covariance matrices $\mathbf{V}_1$ and $\mathbf{V}_2$, we use several relative magnitudes to evaluate the different design options in this paper.

4 D-optimal blocked experiments that are orthogonally blocked

It is not at all unusual that D-optimal designs turn out to be orthogonally blocked. For example, a D-optimal and orthogonally blocked design with two blocks of four observations for estimating a main-effects model in three factors x_1, x_2 and x_3 is obtained by arranging the runs of a 2^3 factorial design in two blocks of size four using the three-factor interaction contrast. More generally, it is not difficult to see that any two-level factorial or fractional factorial design arranged in orthogonal blocks is D-optimal for estimating a model containing the main-effects and any subset of unconfounded interaction effects (that is, unconfounded with block effects, main effects and other interaction effects). A proof of the D-optimality of these designs is given in Goos and Vandebroek (2001a).

That D-optimal designs are orthogonally blocked does not just happen in special cases like these but it also occurs in more complex situatons. Goos and Donev (2005a), for example, find a D-optimal orthogonally blocked design for an 18-run experiment involving two blocking factors, one of which acts at three levels and the other of which acts at two levels, on a square design region $\chi = [-1, 1]^2$ when the variance component σ_1^2 is sufficiently large. The model under investigation was a full second-order model in two continuous variables. This design is displayed in Figure 1. It is easy to verify that the

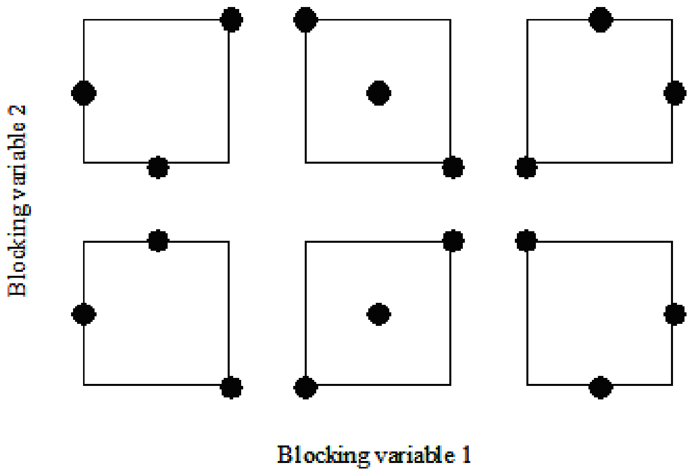


Figure 1: D-optimal, orthogonally blocked design for estimating a full second-order model involving two blocking factors.

overall projection of this design obtained by ignoring the blocks is a duplicated three-level factorial design for two variables. Also, it is not hard to verify that the average levels of all the regressors x_1, x_2, x_1^2, x_2^2 and x_1x_2 in the full second-order model are equal at each level of each blocking factor. This is the condition for having an orthogonally blocked design. The average levels of the regressors amount to 0, 0, 2/3, 2/3 and 0, respectively.

It is remarkable that the design points in Figure 1, that is the points of the duplicated three-level factorial design, are D-optimal for estimating the full second-order model in the presence of the two blocking factors as these points are not the most informative ones on that model in the absence of blocking factors (when $\sigma_1^2 = \sigma_2^2 = 0$), in terms of the D-optimality criterion. The D-optimal set of points in the absence of blocking factors is the set consisting of three replicates of a 2^2 factorial design, a duplicated center point and each of the midpoints of the edges of the experimental region. This design can, however, not be arranged in orthogonal blocks: a D-optimal blocked arrangement of that set of points is displayed in Figure 2. The resulting design is a D-optimal blocked design for values of σ_1^2 and σ_2^2 close to zero.

This example demonstrates the trade-off that is performed by the D-optimality criterion between orthogonality and optimality. When the variance components associated with the blocking factors are small (in which case blocking is not that important), the D-optimality criterion selects the most informative points irrespective of whether the points can be arranged in an orthogonally blocked design or not. When, on the contrary, the magnitudes of the variance components are higher, it is important to select different (and less informative) sets of design points that allow the construction of an orthogonally blocked design. Otherwise, too much information on the factor effects is absorbed by the block effects.

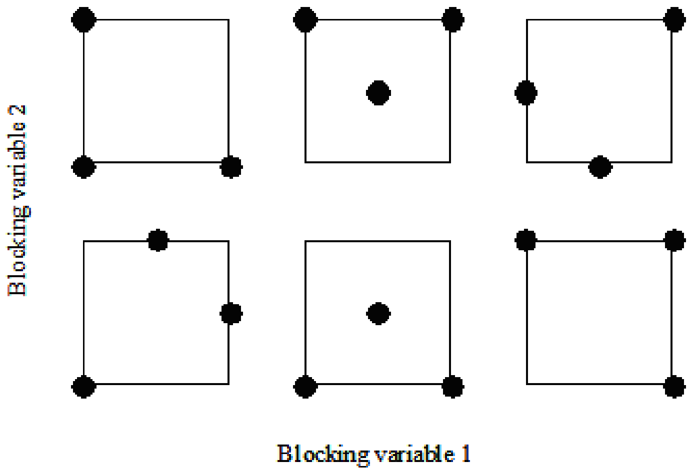


Figure 2: D-optimal, non-orthogonally blocked design for estimating a full second-order model involving two blocking factors.

5 The inefficiency of some orthogonally blocked designs

The previous section has demonstrated that the D-optimality criterion in some cases leads to the selection of a design that is orthogonally blocked. In most practical situations however, a D-optimal design will not be orthogonally blocked. In some of these situations, it is simply impossible to construct an orthogonally blocked design. The most important reason for this to happen is when the number of observations in each of the blocks of the experiment is too small compared to the number of model parameters. However, in some cases where the D-optimal designs are not orthogonally blocked, it is possible to construct orthogonally blocked designs using different methods.

For example, the design with two blocks of four observations in Figure 3a is an orthogonally blocked design for estimating a second-order Scheffé model in three mixture variables:

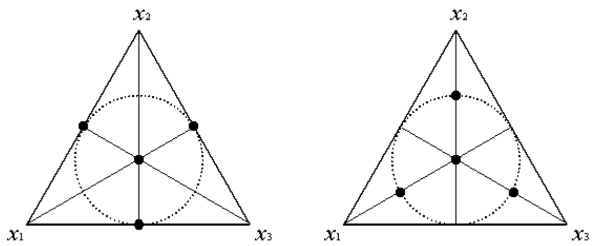
$$E(Y) = \sum_{i=1}^{3} \beta_i x_i + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3.$$

The design in Figure 3a was obtained by projecting a 2^3 factorial design arranged in two orthogonal blocks onto the mixture design region which is defined by the equations

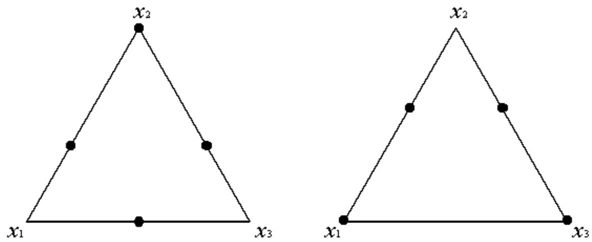
$$0 \leq x_i \leq 1, \quad i = 1, 2, 3,$$

and

$$\sum_{i=1}^{3} x_i = 1.$$



(a) Orthogonally blocked design



(b) $\mathcal{D}$-optimal design

Figure 3: Graphical representation of two three-variable mixture designs with two blocks of size four for estimating a second-order Scheffé polynomial.

More details about this approach for constructing orthogonally blocked design can be found in Prescott (2000). A D-optimal design for the same model and blocking structure is given in Figure 3b. The exact levels of the variables in the eight runs of the designs are displayed in Table 1.

The small mixture experiment illustrates that considering only orthogonally blocked designs as an option entails a substantial cost in terms of statistical efficiency. The relative D-efficiency of the D-optimal design with respect to the orthogonally blocked design equals 3.55 when $\sigma_1^2 = 0$, 3.38 when $\sigma_1^2 = \sigma_\varepsilon^2$ and 3.34 when $\sigma_1^2 = 10\sigma_\varepsilon^2$. Treating the block effects as fixed would result in a relative D-efficiency of 3.33. These numbers show that the orthogonally blocked design would have to be replicated more than three times to achieve the same efficiency as the D-optimal design. In terms of A-efficiency, the difference between the two designs is even larger. For example, when $\sigma_1^2 = \sigma_\varepsilon^2$, the relative A-efficiency of the D-optimal design with respect to the orthognally blocked design is 6.80. When $\sigma_1^2 = \sigma_\varepsilon^2$, the relative G- and V-efficiency of the D-optimal design with respect to the orthogonally blocked one amount to 8.77 and 3.79, respectively. This means that the maximum prediction variance obtained by using the D-optimal design is almost nine times smaller than by using the orthogonally blocked design. On average, the prediction variances obtained using the D-optimal design are 3.79 times smaller. These numbers, although already enormous, are even larger for smaller relative magnitudes of σ_1^2.

Table 1: Coordinates of the points of the orthogonally blocked and D-optimal designs in Figure 3.

	Orthogonal			$\mathcal{D}$-optimal		
Block	x_1	x_2	x_3	x_1	x_2	x_3
1	2/3	1/6	1/6	0.6	0.4	0
1	1/6	2/3	1/6	0.5	0	0.5
1	1/6	1/6	2/3	0	1	0
1	1/3	1/3	1/3	0	0.4	0.6
2	1/2	1/2	0	0.4	0.6	0
2	1/2	0	1/2	0	0.6	0.4
2	0	1/2	1/2	1	0	0
2	1/3	1/3	1/3	0	0	1

Fortunately, not all orthogonally blocked designs in the literature are that inefficient. However, other examples can be found where the orthogonally blocked designs are substantially less attractive in terms of efficiency than, for example, D-optimal ones. A few examples are provided in Goos and Donev (2006a).

Finally, it should be stressed that D-optimal designs that are not perfectly orthogonally blocked, are in many cases close to being orthogonally blocked. For example, Goos (2002) (page 123) reports a D-optimal design for a pastry dough mixing experiment involving three variables. The design is not orthogonally blocked, but it has an efficiency factor of 96.10%. The efficiency factor serves as a measure for the orthogonality of a blocked experiment and equals 100% for designs that are fully orthogonally blocked (see Trinca and Gilmour (2000) for a definition of the efficiency factor). The D-optimal design in Figure 3b has an efficiency factor of 90.99%. These large efficiency factors signify that the estimates of the factor effects will be influenced by the blocked nature of the experimental design only to a small extent, and that researchers should not be overly worried about the non-orthogonality of the D-optimal designs.

6 D-optimal split-plot experiments for which OLS and GLS are equivalent

For specific first-order design problems, there are several situations with even whole plot sizes for which D-optimal split-plot designs possess the property that the ordinary least squares estimator and the generalized least squares estimator are equivalent. Goos and Vandebroek (2003a) proved that this was the case for the minimum aberration designs constructed in Huang, Chen and Voelkel (1998) and Bingham and Sitter (1999b, 1999a, 2001). The minimum aberration split-plot designs in Bingham, Schoen and Sitter (2004) are also

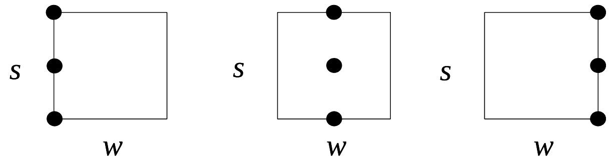


Figure 4: D-optimal second-order split-plot design with three whole plots of size three for estimating a full second-order model in one whole-plot variable w and one sub-plot variable s.

D-optimal and possess the equivalence property. The D-optimality and the equivalence property hold for models including main-effects and every subset of unconfounded interaction effects.

Constructing D-optimal second-order split-plot designs for which ordinary and generalized least squares estimation produce the same estimates for $\boldsymbol{\beta}$ is also possible. Perhaps the simplest design in that class is displayed in Figure 4. The design is D-optimal for estimating a full second-order model in one whole-plot variable w and one sub-plot variable s when three whole plots of three observations are available. The design in Figure 4 is called a crossed design because the same combinations of levels of the whole-plot variables are used in every whole plot. That ordinary and generalized least squares are equivalent for crossed designs was shown in Letsinger, Myers and Lentner (1996).

A problem with the simple design in Figure 4 is that it does not allow the estimation of the variance σ_1^2 associated with the settings of the hard-to-change variables. This is because there are only three whole-plot degrees of freedom in the design, and these are used for estimating the intercept, the main effect of the whole-plot variable w and its quadratic effect. As a result, no degrees of freedom are left for estimating the whole-plot error variance σ_1^2 and a proper split-plot analysis of the data from such a design cannot be done with the classical mixed model analysis. A Bayesian analysis is then required to take into account the split-plot nature of the design (see Gilmour and Goos (2005)). A simple way to solve this problem and to obtain a more practical experiment is to increase the number of whole plots of the design. The D-optimal designs with four, five or six whole plots of size three are also crossed. The D-optimal design with six whole plots consists of two replicates of the design in Figure 4. A D-optimal design with four whole plots replicates the whole plot at $w = -1$ or the one at $w = 1$. The D-optimal design with five whole plots duplicates both the whole plot at $w = -1$ and the one at $w = 1$.

A number of other situations with one whole-plot variable for which the D-optimal design is an equivalent estimation design are identified by Parker, Kowalski and Vining (2005a). The design in Table 2 shows that it is possible to find D-optimal equivalent-estimation designs even when there is more than one whole-plot factor. It is a design with six whole plots of four observations for estimating a full second-order model in two whole-plot factors w_1 and w_2 and two sub-plot variables s_1 and s_2. The design was constructed using the

Table 2: D-optimal split-plot design with two whole-plot variables and two sub-plot variables for which ordinary and generalized least squares are equivalent.

wp	w_1	w_2	s_1	s_2	wp	w_1	w_2	s_1	s_2
1	-1	-1	0	-1	4	-1	1	-1	1
	-1	-1	-1	0		-1	1	-1	-1
	-1	-1	1	1		-1	1	1	1
	-1	-1	-1	1		-1	1	1	-1
2	0	-1	0	1	5	1	0	-1	0
	0	-1	-1	-1		1	0	0	-1
	0	-1	1	-1		1	0	1	1
	0	-1	1	0		1	0	-1	1
3	1	-1	1	-1	6	1	1	1	-1
	1	-1	-1	1		1	1	1	1
	1	-1	0	0		1	1	-1	-1
	1	-1	-1	-1		1	1	-1	1

algorithm of Goos and Vandebroek (2003a). This design suffers from the same problem as the on in Figure 4 in that it has no degrees of freedom for estimating the whole-plot error variance. For practical experiments, it is recommended to construct a design with more than six whole plots.

7 The inefficiency of some equivalent-estimation split-plot designs

The construction of split-plot designs for which ordinary and generalized least squares are equivalent has received much attention lately. Unfortunately, a lot of the equivalent-estimation designs presented in the literature are not very efficient. The inefficiency of these designs is to a large extent due to the fact that they have large numbers of replicated design points. The replications are included in the designs to enable the construction of pure error estimates of the variance components. Equivalent-estimation designs are described in detail in Vining, Kowalski and Montgomery (2005) and Parker, Kowalski and Vining (2005a, 2005b).

A Box-Behnken design, involving one hard-to-change variable w and two easy-to-change variables s_1 and s_2, arranged in six whole plots of four observations so that ordinary least squares and generalized least squares are equivalent is given in Vining, Kowalski and Montgomery (2005). The design is displayed in the left panel of Table 3. A striking feature of the design is that three of its whole plots contain replicates of the center point only. The purpose of this is to construct estimates for the residual and whole-plot error variances that do not depend on a possibly incorrect model specification.

Table 3: Equivalent-estimation split-plot Box-Behnken design and D-optimal design with six whole plots of size four for estimating a full second-order model in one whole-plot variable w and two sub-plot variables s_1 and s_2.

Box-Behnken				D-optimal			
wp	w	s_1	s_2	wp	w	s_1	s_2
1	-1	-1	0	1	1	0	1
	-1	1	0		1	1	0
	-1	0	-1		1	0	-1
	-1	0	1		1	-1	0
2	1	-1	0	2	0	0	0
	1	1	0		0	-1	-1
	1	0	-1		0	-1	1
	1	0	1		0	1	1
3	0	-1	-1	3	-1	1	0
	0	1	1		-1	0	1
	0	-1	-1		-1	-1	0
	0	1	1		-1	0	-1
4	0	0	0	4	-1	-1	0
	0	0	0		-1	0	-1
	0	0	0		-1	0	1
	0	0	0		-1	1	0
5	0	0	0	5	0	1	1
	0	0	0		0	0	0
	0	0	0		0	-1	-1
	0	0	0		0	1	-1
6	0	0	0	6	1	-1	0
	0	0	0		1	0	-1
	0	0	0		1	0	1
	0	0	0		1	1	0

Using the design construction algorithm of Goos and Vandebroek (2003a), a D-optimal design with six whole plots of four observations was constructed as an alternative to the equivalent-estimation Box-Behnken design. A spherical design region was assumed and the points of the Box-Behnken design were used as candidate points in the algorithmic construction. Using the points of a Box-Behnken design as candidates in a design construction algorithm is recommended by Mee (2005) for spherical design regions. The resulting D-optimal design is displayed in the right panel of Table 3. The design only has two center point replicates, which leads to a substantial increase in efficiency: the relative D-efficiency of the D-optimal design with respect to the equivalent-estimation design is 1.65 when $\sigma_1^2 = \sigma_\varepsilon^2$.

A remarkable feature of the D-optimal design is that ordinary and generalized least squares produce identical estimates for eight of the ten parameters in the full second-order model. Only the main effects of the two sub-plot variables are estimated differently. The inequivalence is caused by the fact that the center point is required in the design for the model's estimability and it makes the design in the sub-plot variables s_1 and s_2 unbalanced.

The D-optimal design is not just better in terms than the equivalent-estimation design in terms of D-efficiency. The relative A-efficiency of the D-optimal design with respect to the Box-Behnken design is 1.24, whereas the relative G- and V-efficiency amount to 1.49 and 1.75, respectively. These relative efficiencies were obtained for $\sigma_1^2 = \sigma_\varepsilon^2$.

The comparison between the D-optimal split-plot design in the right panel of Table 3 is of course unfair in a way. The reason is that the twelve replicates of the center points are included in the design for estimating the variance components. A full second-order model could be estimated with just one of the three center-point whole plots. This design, which would still be an equivalent-estimation design, is however still substantially less efficient than a D-optimal design with four whole plots of four observations in terms of D-, A- and V-efficiency. The relative D-, A- and V-efficiencies when $\sigma_1^2 = \sigma_\varepsilon^2$ equal 1.21, 1.27 and 1.15, respectively. In terms of G-efficiency, the D-optimal design is only marginally better than the four whole-plot equivalent-estimation design. The D-optimal design is shown in Table 4.

8 Designs with heterogeneous block and plot sizes

Traditionally, researchers in experimental design have focused on experiments with equal block or whole-plot sizes. In some experimental situations, there is however no hard practical constraint that forces a researcher to restrict his/her attention to designs with homogeneous block sizes or whole-plot sizes. For example, an alternative to the design in Table 2 is given in Table 5. The design was constructed using the algorithm of Goos and Vandebroek (2004), which allows the researcher to find the D-optimal number of whole plots and the D-optimal numbers of observations in each whole plot in addition to the

Table 4: D-optimal design with four whole plots of size four for estimating a full second-order model in one whole-plot variable w and two sub-plot variables s_1 and s_2.

wp	w	s_1	s_2
1	1	1	0
	1	0	-1
	1	0	1
	1	-1	0
2	0	0	0
	0	1	1
	0	-1	-1
	0	-1	1
3	-1	1	0
	-1	-1	0
	-1	0	-1
	-1	0	1
4	0	1	1
	0	1	-1
	0	0	0
	0	-1	-1

D-optimal design points. The design in Table 5 was obtained by imposing an upper limit of six on the number of whole plots. It has two whole plots of three observations, two of four observations and two of five observations. The relative D-efficiency of the design in Table 5 with respect to the one in Table 2 equals 1.03 when $\sigma_1^2 = \sigma^2$, which means that allowing heterogeneous whole-plot sizes leads to more efficient designs. This would even be more so when more than six whole plots were allowed. Actually, more than six whole plots are needed for being able to estimate the whole-plot error variance.

Finally, it should be mentioned that the split-plot design in Table 5 turns out to be an equivalent-estimation design even though it has heterogeneous whole-plot sizes. This demonstrates the fact that the D-optimality criterion does lead to equivalent-estimation or orthogonally blocked designs if it doesn't imply too substantial a loss in efficiency.

9 Discussion

The focus in this article was on orthogonally blocked and equivalent-estimation split-plot designs, and on D-optimal designs. It was shown that in some cases D-optimal blocked and split-plot designs happen to be fully orthogonally blocked designs or equivalent-estimation split-plot designs. In many cases however, D-optimal design are not perfectly orthogonally blocked or they lead to different ordinary and generalized least squares estimates for some of the model parameters. This may make the D-optimal design less attractive and inspires

Table 5: D-optimal split-plot design with six whole-plots of heterogeneous whole-plot sizes for estimating a full second-order model with two whole-plot variables and two sub-plot variables.

wp	w_1	w_2	s_1	s_2
1	-1	-1	1	-1
	-1	-1	-1	-1
	-1	-1	1	1
	-1	-1	-1	1
2	1	-1	-1	0
	1	-1	1	1
	1	-1	-1	1
	1	-1	0	-1
	1	-1	1	-1
3	1	0	-1	-1
	1	0	1	0
	1	0	0	1
4	-1	1	-1	-1
	-1	1	-1	1
	-1	1	1	-1
	-1	1	1	1
5	0	1	-1	0
	0	1	1	1
	0	1	0	-1
6	1	1	1	1
	1	1	-1	-1
	1	1	1	-1
	1	1	0	0
	1	1	-1	1

some researchers to put the optimal design approach aside. However, several examples in this paper demonstrated that some of the orthogonally blocked or equivalent-estimation designs proposed in the literature are very inefficient. The inefficiency is not just in terms of the computationally convenient D-optimality criterion but also in terms of the A-, G- and V-optimality criteria. It was also shown that it may be beneficial to consider experimental designs with heterogeneous block or whole plots sizes. Although this might be counterintuitive, more efficient designs can often be found by leaving the well-paved path of balanced designs.

References

Anbari, F. T. and Lucas, J. M. (1994). Super-efficient designs: How to run your experiment for higher efficiency and lower cost, *ASQC Technical Conference Transactions*, pp. 852–863.

Atkinson, A. C. and Donev, A. N. (1989). The construction of exact D-optimum experimental designs with application to blocking response surface designs, *Biometrika* **76**: 515–526.

Bingham, D. R., Schoen, E. D. and Sitter, R. R. (2004). Designing fractional factorial split-plot experiments with few whole-plot factors, *Journal of the Royal Statistical Society, Ser. C* **53**: 325–339.

Bingham, D. R. and Sitter, R. R. (1999a). Some theoretical results for fractional factorial split-plot designs, *Annals of Statistics* **27**: 1240–1255.

Bingham, D. R. and Sitter, R. R. (2001). Design issues in fractional factorial split-plot designs, *Journal of Quality Technology* **33**: 2–15.

Bingham, D. and Sitter, R. R. (1999b). Minimum-aberration two-level fractional factorial split-plot designs, *Technometrics* **41**: 62–70.

Box, G. E. P. and Hunter, J. S. (1957). Multi-factor experimental designs for exploring response surfaces, *Annals of Mathematical Statistics* **28**: 195–241.

Cook, R. D. and Nachtsheim, C. J. (1989). Computer-aided blocking of factorial and response-surface designs, *Technometrics* **31**: 339–346.

Donev, A. N. (1989). Design of experiments with both mixture and qualitative factors, *Journal of the Royal Statistical Society, Ser. B* **51**: 297–302.

Draper, N. R. and John, J. A. (1998). Response surface designs where levels of some factors are difficult to change, *Australian and New-Zealand Journal of Statistics* **40**: 487–495.

Gilmour, S. G. and Goos, P. (2005). Bayesian analysis of data from multi-stratum response surface designs. Unpublished manuscript.

Gilmour, S. G. and Trinca, L. A. (2000). Some practical advice on polynomial regression analysis from blocked response surface designs, *Communications in Statistics: Theory and Methods* **29**: 2157–2180.

Goos, P. (2002). *The Optimal Design of Blocked and Split-plot Experiments*, New York: Springer.

Goos, P. and Donev, A. N. (2005a). Blocking response surface designs. To appear in *Computational Statistics and Data Analysis*.

Goos, P. and Donev, A. N. (2005b). Tailor-made split-plot mixture-process variable designs. Unpublished manuscript.

Goos, P. and Donev, A. N. (2006a). The D-optimal design of blocked experiments with mixture components. To appear in *Journal of Quality Technology*.

Goos, P. and Donev, A. N. (2006b). D-optimal minimum support mixture designs in blocks. To appear in *Metrika*.

Goos, P. and Vandebroek, M. (2001a). D-optimal response surface designs in the presence of random block effects, *Computational Statistics and Data Analysis* **37**: 433–453.

Goos, P. and Vandebroek, M. (2001b). Optimal split-plot designs, *Journal of Quality Technology* **33**: 436–450.

Goos, P. and Vandebroek, M. (2003a). D-optimal split-plot designs with given numbers and sizes of whole plots, *Technometrics* **45**: 235–245.

Goos, P. and Vandebroek, M. (2003b). Estimating the intercept in an orthogonally blocked experiment, *Communications in Statistics: Theory and Methods* **33**: 873–890.

Goos, P. and Vandebroek, M. (2004). Outperforming completely randomized designs, *Journal of Quality Technology* **36**: 12–26.

Grasshoff, U. and Schwabe, R. (2003). On the analysis of paired observations, *Statistics & Probability Letters* **65**: 7–12.

Huang, P., Chen, D. and Voelkel, J. (1998). Minimum-aberration two-level split-plot designs, *Technometrics* **40**: 314–326.

John, P. W. M. (1984). Experiments with mixtures involving process variables, *Technical Report 8*, Center for Statistical Sciences, University of Texas, Austin, Texas.

Khuri, A. I. (1992). Response surface models with random block effects, *Technometrics* **34**: 26–37.

Kowalski, S. M., Cornell, J. A. and Vining, G. G. (2002). Split-plot designs and estimation methods for mixture experiments with process variables, *Technometrics* **44**: 72–79.

Letsinger, J. D., Myers, R. H. and Lentner, M. (1996). Response surface methods for bi-randomization structures, *Journal of Quality Technology* **28**: 381–397.

Mee, R. (2005). Optimal three-level designs for response surfaces in spherical experimental regions. Unpublished manuscript.

Nigam, A. K. (1976). Corrections to blocking conditions for mixture experiments, *Annals of Statistics* **47**: 1294–1295.

Parker, P. A., Kowalski, S. M. and Vining, G. G. (2005a). Construction of balanced equivalent estimation second-order split-plot designs. Unpublished manuscript.

Parker, P. A., Kowalski, S. M. and Vining, G. G. (2005b). Unbalanced and minimal point equivalent estimation second-order split-plot designs. Unpublished manuscript.

Prescott, P. (2000). Projection designs for mixture experiments in orthogonal blocks, *Communications in Statistics: Theory and Methods* **29**: 2229–2253.

Trinca, L. A. and Gilmour, S. G. (2000). An algorithm for arranging response surface designs in small blocks, *Computational Statistics and Data Analysis* **33**: 25–43.

Trinca, L. A. and Gilmour, S. G. (2001). Multi-stratum response surface designs, *Technometrics* **43**: 25–33.

Vining, G. G., Kowalski, S. M. and Montgomery, D. C. (2005). Response surface designs within a split-plot structure, *Journal of Quality Technology* **37**: 115–129.