DEPARTMENT OF ENGINEERING MANAGEMENT

Staggered-level designs for response surface modeling

Heidi Arnouts & Peter Goos

UNIVERSITY OF ANTWERP
Faculty of Applied Economics
City Campus
Prinsstraat 13, B.226
B-2000 Antwerp
Tel. +32 (0)3 265 40 32
Fax +32 (0)3 265 47 99
www.uantwerpen.be

FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF ENGINEERING MANAGEMENT

Staggered-level designs for response surface modeling

Heidi Arnouts & Peter Goos

RESEARCH PAPER 2013-027
NOVEMBER 2013

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp, Belgium
Research Administration – room B.226
phone: (32) 3 265 40 32
fax: (32) 3 265 47 99
e-mail: joeri.nys@uantwerpen.be

D/2013/1169/027

Staggered-level designs for response surface modeling

Heidi Arnouts[1] and Peter Goos[1,2]

[1]University of Antwerp, Belgium

[2]Erasmus University Rotterdam, the Netherlands

Abstract

In industrial experiments, there are often restrictions in randomization caused by equipment and resource constraints, as well as budget and time restrictions. Next to the split-plot and the split-split-plot design, the staggered-level design is an interesting design option for experiments involving two hard-to-change factors. The staggered-level design allows both hard-to-change factors to be reset at different points in time, resulting in a typical staggering pattern of factor level resettings. It has been shown that, for two-level designs, this staggering pattern leads to statistical benefits in comparison to the split-plot and the split-split-plot design. In this paper, we investigate whether the benefits of the staggered-level design carry over to situations where the objective is to optimize a response, and where a second-order response surface model is in place. To this end, we study several examples of D- and I-optimal staggered-level response surface designs.

Keywords: D- and I-optimality criterion, cost, response surface model, split-plot design, split-split-plot design, staggered-level design.

1 Introduction

Many industrial experiments involve two categories of factors: easy-to-change factors, whose levels are reset independently for each run, and hard-to-change factors , whose levels are not reset independently for each run. This can be due to equipment and resource constraints, but also to budget and time restrictions. For situations in which there is only one hard-to-change factor or in which the levels of all hard-to-change factors need to be reset at the same time, the split-plot design is the appropriate design option.

The design and analysis of split-plot industrial experiments has received considerable attention in the literature in recent years. Huang, Chen and Voelkel (1998), Bingham and Sitter (1999), and Bingham, Schoen and Sitter (2004) focused on the construction of two-level regular fractional factorial split-plot designs using the minimum aberration criterion. Kulahci and Bisgaard (2005) use Plackett and Burman designs to construct non-regular two-level split-plot designs. Vining, Kowalski and Montgomery (2005), and Parker, Kowalski and Vining (2006, 2007a, 2007b) discussed equivalent-estimation split-plot response surface designs for which ordinary least squares estimation leads to the same estimates as generalized least squares estimation. Goos and Vandebroek (2001, 2003, 2004) and Jones and Goos (2007) constructed D-optimal split-plot designs using point-exchange and coordinate-exchange algorithms, respectively. Macharia and Goos (2010) and Jones and Goos (2012a) showed a range of D-optimal and D-efficient designs that also result in equivalent OLS and GLS estimators. Finally, Jones and Goos (2012b) compared the performance I-optimal and D-optimal split-plot designs for response surface models, while Mylona, Goos and Jones (2013) present a new optimality criterion for selecting split-plot designs.

For situations in which there are at least two hard-to-change factors or two classes of hard-to-change factors, the split-split-plot design is the most common and known design configuration. The literature on industrial split-split-plot designs is, however, rather limited. Trinca and Gilmour (2001) discussed the design and analysis of multi-stratum experiments, special cases of which are split-plot and split-split-plot designs. Schoen (1999) constructed an orthogonal split-split-plot design in a combinatorial way by joining fractional factorial designs in order to create the desired nesting structure. Jones and Goos (2009) discussed a coordinate-exchange algorithm to compute D-optimal split-split-plot designs.

Arnouts and Goos (2012) left the well-trodden path of split-plot and split-split-plot designs and presented a new type of two-level design for situations in which there are two classes of hard-to-change factors, the number of settings of the hard-to-change factors is not dictated by the physicalities of the experiment (such as oven or batch sizes), and the runs are conducted under homogeneous circumstances. Contrary to the split-split-plot design, this new design option allows the two classes of hard-to-change factors to be reset independently at different points in time. Since the structure of the new design option requires the hard-to-change factors to be reset alternatingly, the new design option is named a staggered-level design. Arnouts and Goos (2012) compared the performance of D-optimal staggered-level designs to the performances of D-optimal split-plot and split-split-plot designs in case of a main-effects-plus-two-factor-interaction-effects model. It turned out that the D-optimal staggered-level design was not only the most cost-efficient design of the three options but also statistically the most efficient option.

In this article, we introduce the staggered-level design as a cost-efficient and statistically efficient alternative for split-plot and split-split-plot response surface designs. We do not only focus on D-optimal designs, which maximize the determinant of the information matrix, but we also generate I-optimal designs, which minimize the average prediction variance. The consideration and use of I-optimal designs is a logical choice since the goal in response surface experimentation is usually to make predictions. Moreover, Hardin and Sloane (1993) showed that D-optimal completely randomized response surface designs perform poorly in terms of the I-optimality criterion, while I-optimal designs perform reasonable well with respect to the D-optimality criterion, when the design region is cuboidal. When the experimental region is spherical, the differences between D- and I-optimal designs are less pronounced, but generally still in favor of I-optimal designs. Jones and Goos (2012b) observe a similar pattern for split-plot response surface designs.

In this article, we first review the characteristics of the staggered-level design option, using a two-level design. Next, we describe the model used for data from staggered-level designs and define the D-optimality criterion as well as the I-optimality criterion. Finally, we demonstrate the benefits of the staggered-level design option for response surface models using three examples. Additionally, we compare D- and I-optimal staggered-level designs.

2 The staggered-level design option

To explain the characteristics of the staggered-level design, a 16-run two-level staggered-level design is shown in Table 1. This design was constructed for an experiment with two hard-to-change factors, w and s, and two easy-to-change factors, t_1 and t_2, where the goal was to estimate a main-effects-plus-two-factor-interaction-effects model. In Arnouts and Goos (2012), the first hard-to-change factor, w, was referred to as the class-1 hard-to-change factor, and its levels were assumed to be most difficult to reset. Similarly, the second hard-to-change factor, s, was named the class-2 hard-to-change factor, and its levels were assumed to be less difficult to reset. In this article, we adopt the same notation and nomenclature. Unlike Arnouts and Goos (2012), we do not assume that class-1 hard-to-change factors are harder to change than class-2 hard-to-change factors. This is because the staggered-level design is a cost-efficient and statistical efficient option when the class-1 and class-2 hard-to-change factors are equally hard to change as well.

The key feature of the staggered-level design in Table 1 is the fact that the level of the class-2 hard-to-change factor, s, is reset at different points in time than the level of the class-1 hard-to-change factor, w. In the 16-run staggered-level design in Table 1, the runs are divided in four subsets of size four by the settings of the class-1 hard-to-change factor, w. The runs are also divided in subsets through the settings of the class-2 hard-to-change factor s. The latter division begins and ends with a subset of runs half as large as the subsets defined by the settings of w, i.e. with subsets of size two. This results in an alternating pattern for the settings of the hard-to-change factors.

In Arnouts and Goos (2012), this new design option was compared to the split-plot and split-split-plot design, the two alternative design options for an experimental scenario involving two hard-to-change factors described in the literature. First of all, it turned out that, of the three possibilities, the staggered-level design was the most cost-effcient design due to the smaller number of settings of the hard-to-change factors. For the experimen-

Table 1: 16-run staggered-level design for a model including main effects and two-factor interactions of two hard-to-change factors, w and s, and two easy-to-change factors t_1 and t_2. Horizontal lines indicate time points at which a hard-to-change factor's level is reset.

Run	w	s	t_1	t_2
1	-1	1	1	1
2	-1	1	-1	-1
3	-1	-1	1	-1
4	-1	-1	-1	1
5	1	-1	1	1
6	1	-1	-1	-1
7	1	1	1	-1
8	1	1	1	-1
9	-1	1	-1	1
10	-1	1	1	-1
11	-1	-1	1	1
12	-1	-1	-1	-1
13	1	-1	1	-1
14	1	-1	-1	1
15	1	1	-1	-1
16	1	1	1	1

tal situation given in Table 1, the split-plot design would require eight settings of both hard-to-change factors, since both hard-to-change factors are reset at the same time in a split-plot design. In a split-split-plot design configuration, the class-2 hard-to-change factor is reset whenever the level of the class-1 hard-to-change factor changes, as well as at several other time points. For the 16-run experiment, a split-split-plot configuration would lead to four settings of the class-1 hard-to-change factor, w, and eight settings of the class-2 hard-to-change factor, s.

The comparison of the three alternative designs also revealed that the staggered-level design was statistically the most efficient design of the three in terms of the D-optimality criterion. In general, in comparison to the split-plot design, the main effect of the class-2 hard-to-change factor, s, as well as the interaction effect of both hard-to-change factors, is estimated more precisely from the staggered-level design, even though this design option involves a smaller number of settings of both hard-to-change factors. Compared to the split-split-plot design, the staggered-level design generally leads to a more precise estimation of the main effect of the class-1 hard-to-change factor, w, and the interaction effect of the class-1 and class-2 hard-to-change factor, while having fewer settings of the class-2 hard-to-change factor.

This statistical advantage of the staggered-level design is a consequence of the specific ordering of the subsets of runs determined by the level settings of the hard-to-change factors. That ordering ensures that the two levels of the class-2 hard-to-change factor can be compared with each other within each of the four subsets created by the settings of the class-1 hard-to-change factor. Similarly, the two levels of the class-1 hard-to-change factor can be compared with each other in the three large subsets formed by the settings of the class-2 hard-to-change factor. Neither the split-plot nor the split-split-plot design possess this characteristic.

3 Statistical model and optimality criteria

3.1 Model

In the experimental designs considered in this article, there are three types of factors. There is one class-1 hard-to-change factor, w, one class-2 hard-to-change factor, s, and there are $v_1 = v - 2$ easy-to-change factors, $t_1, \ldots, t_{v_1}$. The easy-to-change factors are reset independently for each run, even when the factors' levels are the same in consecutive runs. As mentioned in the previous section, the experimental runs are partitioned in two ways, one for each class of hard-to-change factor. In the model, we include random effects δ_i, $i = 1, \ldots, r$, for each of the r independent settings of the class-1 hard-to-change factor to capture the dependence between runs for which this hard-to-change factor is not independently reset. To capture the dependence between runs for which the class-2 hard-to-change factor is not independently reset, we also include random effects γ_j, $j = 1, \ldots, g$, in the model for each of the g independent settings of the class-2 hard-to-change factors.

For a response surface model, the kth response obtained at the ith setting of w and the jth setting of s can then be written as

$$\begin{aligned} Y_{ijk} =& \beta_0 + \beta_w w_i + \beta_s s_j + \sum_{l=1}^{v_1} \beta_{t_l} t_{lk} + \beta_{ws} w_i s_j \\ &+ \sum_{l=1}^{v_1} \beta_{wt_l} w_i t_{lk} + \sum_{l=1}^{v_1} \beta_{st_l} s_j t_{lk} + \sum_{l=1}^{v_1-1} \sum_{m=l+1}^{v_1} \beta_{t_l t_m} t_{lk} t_{mk} \\ &+ \beta_{ww} w_i^2 + \beta_{ss} s_j^2 + \sum_{l=1}^{v_1} \beta_{t_l t_l} + \delta_i + \gamma_j + \varepsilon_k. \end{aligned} \tag{1}$$

This model is utilized in Webb et al. (2004) for data from experiments involving several hard-to-change factors.

By using the model in Equation (1), we assume that the runs of the staggered-level design mainly suffer from treatment error which is *"error due to our inability to replicate a treatment from one application to the next"*. This is the first of five possible sources of random error that can occur when running an experiment listed by Hinkelmann and Kempthorne (2008). In Equation (1), the inability to replicate a treatment can be due to the class-1 hard-to-change factor(s), the class-2 hard-to-change factor(s), and the easy-to-change factor(s). The errors introduced by resetting the levels of these types of factors are δ_i, γ_j and ε_k. The assumption that the treatment error is the dominating source of errors means that, for example, we assume that the staggered-level design is not spread over different days. This would cause a change in the physical state of the experimental units, introduce a second type of error and necessitate the inclusion of block effects in Equation (1). Hinkelmann and Kempthorne (2008) refer to this second type of error as state error.

In matrix notation, Equation (1) is written as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_\delta \boldsymbol{\delta} + \mathbf{Z}_\gamma \boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{2}$$

where $\mathbf{Y}$ is the $n \times 1$ vector containing the n responses of the experiment, $\boldsymbol{\beta}$ is a $p \times 1$ vector that contains the $p = v(v+1)/2$ model parameters, $\mathbf{X}$ is the $n \times p$ model matrix, $\mathbf{Z}_\delta$ is the $n \times r$ matrix with (i, j)th entry equal to 1 if the ith run is conducted at the jth setting of the class-1 hard-to-change factor and equal to 0 otherwise, $\mathbf{Z}_\gamma$ is the $n \times g$ matrix with (i, j)th entry equal to 1 if the ith run is conducted at the jth setting of the class-2 hard-to-change factor and equal to 0 otherwise, $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are the $r \times 1$ and $g \times 1$ vectors containing the random effects associated with the independent settings of the class-1 and the class-2 hard-to-change factors, respectively, and $\boldsymbol{\varepsilon}$ is the $n \times 1$ vector of random errors.

It is assumed that

$$\mathrm{E}(\boldsymbol{\delta}) = \mathbf{0}_r, \mathrm{cov}(\boldsymbol{\delta}) = \sigma_\delta^2 \mathbf{I}_r,$$

$$\mathrm{E}(\boldsymbol{\gamma}) = \mathbf{0}_g, \mathrm{cov}(\boldsymbol{\gamma}) = \sigma_\gamma^2 \mathbf{I}_g,$$

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n, \mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2 \mathbf{I}_n,$$

and that

$$\mathrm{cov}(\boldsymbol{\delta}, \boldsymbol{\gamma}) = \mathbf{0}_{r \times g}, \mathrm{cov}(\boldsymbol{\delta}, \boldsymbol{\varepsilon}) = \mathbf{0}_{r \times n} \text{ and } \mathrm{cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{g \times n},$$

where $\mathbf{0}_c$ and $\mathbf{I}_c$ represent a c-dimensional zero vector and identity matrix, respectively, and $\mathbf{0}_{c \times d}$ is a zero matrix of dimension $c \times d$.

Under these assumptions, the variance-covariance matrix of the responses, $\mathbf{Y}$, is

$$\begin{aligned} \mathbf{V} &= \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_\delta^2 \mathbf{Z}_\delta \mathbf{Z}_\delta' + \sigma_\gamma^2 \mathbf{Z}_\gamma \mathbf{Z}_\gamma' \\ &= \sigma_\varepsilon^2 (\mathbf{I}_n + \eta_\delta \mathbf{Z}_\delta \mathbf{Z}_\delta' + \eta_\gamma \mathbf{Z}_\gamma \mathbf{Z}_\gamma'), \end{aligned} \tag{3}$$

where η_δ and η_γ are the variance ratios $\sigma_\delta^2/\sigma_\varepsilon^2$ and $\sigma_\gamma^2/\sigma_\varepsilon^2$ for the class-1 and class-2 hard-to-change factors, respectively. The larger these ratios, the stronger the runs conducted at the same setting of the class-1 and/or class-2 hard-to-change factors are correlated.

The statistical model in Equation (1) generalizes the split-plot and the split-split-plot model. For the model to reduce to the split-plot model, it is necessary that $\mathbf{Z}_\delta = \mathbf{Z}_\gamma$. This means that the class-1 and class-2 hard-to-change factor are reset at the same points in time. In that case, the variance components σ_δ^2 and σ_γ^2 cannot be estimated separately. Only their sum is estimable. This does not occur in split-split-plot designs, where $\mathbf{Z}_\gamma = \mathbf{I}_r \otimes \mathbf{1}_{c_1}$ and $\mathbf{Z}_\delta = \mathbf{I}_g \otimes \mathbf{1}_{c_2}$, with $\otimes$ the Kronecker product, $\mathbf{1}_{c_i}$ a c_i-dimensional vector of ones, c_1 and c_2 the number of runs in a whole plot and a subplot, respectively, and $n = rc_1 = gc_2$.

Under the assumption of normality, the maximum likelihood estimator of the unknown model parameter vector $\boldsymbol{\beta}$ is the generalized least squares estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}, \tag{4}$$

with variance-covariance matrix

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{5}$$

The information matrix on the unknown parameter vector $\boldsymbol{\beta}$ is given by

$$\mathbf{M} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}. \tag{6}$$

3.2 Optimality criteria

D-optimality criterion

The D-optimality criterion is probably the most commonly used criterion when constructing optimal designs. Designs that are D-optimal are usually said to also perform well in terms of other design criteria. In addition, the criterion is invariant to linear transformations of the design matrix. Consequently, the criterion is invariant to the scale or the coding of the variables. The D-optimality criterion seeks designs that maximize the determinant of the information matrix in Equation (6).

To compare two designs with information matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ in terms of the D-optimality criterion, we use the relative local D-efficiency $(|\mathbf{M}_1| \, / \, |\mathbf{M}_2|)^{1/p}$. A relative D-efficiency larger than one means that Design 1 is better than Design 2 in terms of the D-optimality criterion.

In general, the D-optimal designs as well as the relative performance of two designs depend on the relative magnitude of the variance components through the variance-covariance matrix $\mathbf{V}$. This is why we use the adjective "local" in the term relative local D-efficiency. Arnouts and Goos (2012) observed that the performance of the staggered-level designs is not sensitive to the variance ratios η_δ and η_γ. This is why we do not use a Bayesian approach, which explicitly accounts for uncertainty concerning the values of η_δ and η_γ during the design construction, in this paper. Instead, we just perform a sensitivity study for each experimental situation we discuss in the following sections.

I-optimality criterion

The I-optimality criterion seeks designs that minimize the average prediction variance

$$\frac{\int_\chi \boldsymbol{f}'(\boldsymbol{x})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\boldsymbol{f}(\boldsymbol{x})d\boldsymbol{x}}{\int_\chi d\boldsymbol{x}} \tag{7}$$

over the experimental region χ. Since the goal in response surface experimentation usually is to make predictions, it makes sense to also consider this optimality criterion when constructing staggered-level response surface designs. Just like the D-optimality criterion, the I-optimality criterion is also invariant to linear transformations of the design matrix.

For regular experimental regions, the calculation of the expression in Equation (7) is not computationally involved. If there are v quantitative experimental variables and the experimental region is $[-1, +1]^v$, then the denominator, representing the volume of the experimental region, equals 2^v. The numerator involves the variance of prediction, $\boldsymbol{f}'(\boldsymbol{w})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\boldsymbol{f}(\boldsymbol{x})$, which is a scalar. Consequently,

$$\boldsymbol{f}'(\boldsymbol{x})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\boldsymbol{f}(\boldsymbol{x}) = \text{tr}[\boldsymbol{f}'(\boldsymbol{x})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\boldsymbol{f}(\boldsymbol{x})].$$

Using this result, Jones and Goos (2012b) showed that the formula for the average prediction variance can be rewritten as

$$2^{-v}\text{tr}\left[(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\int_{\chi}\boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}'(\boldsymbol{x})d\boldsymbol{x}\right]. \tag{8}$$

The integral in Equation (8) is applied to a matrix of monomials, i.e. one-term polynomials. Therefore, it can interpreted as the matrix of integrals of these monomials. Let

$$\mathbf{B} = \int_{\chi}\boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}'(\boldsymbol{x})d\boldsymbol{x},$$

then the average prediction variance can be written as

$$2^{-v}\text{tr}[(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}]. \tag{9}$$

If the experimental design region is $\chi = [-1, +1]^v$, then the integrals in $\mathbf{B}$ are quite simple. Hardin and Sloane (1993) pointed out that the matrix $\mathbf{B}$, called the moments matrix, has a very specific structure for a response surface model and a cuboidal design region:

$$\mathbf{B} = 2^v\begin{bmatrix} 1 & \mathbf{0}'_v & \mathbf{0}'_{v^*} & \frac{1}{3}\mathbf{1}'_v \\ \mathbf{0}_v & \frac{1}{3}\mathbf{I}_v & \mathbf{0}_{v\times v^*} & \mathbf{0}_{v\times v} \\ \mathbf{0}_{v*} & \mathbf{0}_{v^*\times v} & \frac{1}{9}\mathbf{I}_{v^*} & \mathbf{0}_{v^*\times v} \\ \frac{1}{3}\mathbf{1}_v & \mathbf{0}_{v\times v} & \mathbf{0}_{v\times v^*} & \frac{1}{45}(4\mathbf{I}_v + 5\mathbf{J}_v) \end{bmatrix}, \tag{10}$$

where v is the number of factors, $v^* = v(v-1)/2$ is the number of two-factor interaction effects and $\mathbf{J}_v$ is a $v \times v$ matrix of ones.

To compare two designs with an average prediction variance of P_1 and P_2, we can calculate the relative local I-efficiency I-efficiency $= P_2/P_1$. A relative local I-efficiency larger than one indicates that Design 1 is better than Design 2 in terms of the average variance of prediction. Since the I-optimality criterion depends on $\mathbf{V}$ through the variance-covariance matrix of the parameter estimates, $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$, the I-optimal design and the I-efficiency of one design relative to another depend on the variance ratios η_δ and η_γ. We therefore also investigate the sensitivity of the I-optimal staggered-level designs constructed in this article to the values of η_δ and η_γ.

4 28-run design

We start the comparison of the three design types, i.e. the staggered-level, the split-plot and the split-split-plot design, in the context of a 28-run experiment to estimate a response surface model in two hard-to-change factors, w and s, and two easy-to-change factors, t_1 and t_2. To generate optimal designs, we implemented a coordinate-exchange algorithm that seeks the D- and I-optimal factor levels for any given staggering structure. When using the algorithm, input on the magnitude of the variance ratios η_δ and η_γ is required. For now, we assume that both variance ratios equal one. Later in this section, we study the dependence of the constructed staggered-level designs on these variance components.

4.1 Design options

To have enough degrees of freedom to estimate all fixed parameters and all variance components, the split-plot designs, which we use as benchmarks, involve seven level settings of both hard-to-change factors. The D- and I-optimal split-plot designs for this situation are shown in Table 2. The staggered-level design with seven level settings of the class-1 hard-to-change factor w and eight level settings of the class-2 hard-to-change factor s is another cost-efficient configuration for this experimental situation. The D- and I-optimal staggered-level designs are shown in Table 3. The least cost-efficient option of the three in this experimental situation is the split-split-plot design, the D- and I-optimal versions of which are shown in Table 4, with seven level settings of w and 14 level settings of the class-2 hard-to-change factor s. The statistical comparison between the three design options is summarized in Table 5, assuming all variance components are equal to one.

4.2 Comparison in terms of the D-optimality criterion

In terms of cost efficiency, the split-plot designs are slightly better than the staggered-level designs with one fewer level setting of s. The statistical results, however, are strongly in favor of the staggered-level design. The D-optimal split-plot design has a D-efficiency of only 77.3% relative to the D-optimal staggered-level design. This is mainly due to an imprecise estimation of the main effects and quadratic effects of both hard-to-change factors, as well as their interaction effect, when using the split-plot design. For the split-plot design, the main effect of w, the main effect of s and their two-factor-interaction effect are confounded with the random effects contained within both $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$. Arnouts and Goos (2012) explained that, because of its specific ordering of the hard-to-change factor levels, this is not true for the staggered-level designs. Instead, the main effect of w is only confounded with $\boldsymbol{\delta}$ and not with $\boldsymbol{\gamma}$. The main effect of s is confounded with $\boldsymbol{\gamma}$ and only to a small extent with $\boldsymbol{\delta}$. This results in more precise estimates of these effects.

As mentioned before, the 28-run split-split-plot design is the least cost-efficient option of the three, but this does not result in a larger D-efficiency. This is due to the fact that the main effect and quadratic effect of the class-1 hard-to-change factor w, as well as the interaction effects between the hard-to-change facors w and s and between the easy-to-change factors t_1 and t_2 are estimated more precisely from the D-optimal staggered-level design. This results in the D-optimal split-split-plot design being 8% less D-efficient than the D-optimal staggered-level design.

4.3 Comparsion in terms of the I-optimality criterion

In terms of the I-optimality criterion, the split-plot design performs poorly in comparison to the staggered-level design. The I-optimal split-plot design has an I-efficiency of only 52% relative to the I-optimal staggered-level design. This means that the split-plot design

Table 2: D- and I-optimal 28-run split-plot designs for estimating a response surface model in two hard-to-change factors w and s, and two easy-to-change factors t_1 and t_2, assuming η_δ and η_γ equal to one. The column labelled WP indicates the seven whole plots.

Run	WP	D-optimal				I-optimal			
		w	s	t_1	t_2	w	s	t_1	t_2
1	1	-1	1	1	1	0	-1	0	1
2	1	-1	1	-1	1	0	-1	1	0
3	1	-1	1	0	-1	0	-1	-1	-1
4	1	-1	1	1	-1	0	-1	0	0
5	2	-1	0	1	-1	-1	1	0	-1
6	2	-1	0	0	1	-1	1	0	0
7	2	-1	0	-1	-1	-1	1	-1	1
8	2	-1	0	-1	0	-1	1	1	1
9	3	1	-1	-1	1	0	0	1	-1
10	3	1	-1	1	-1	0	0	0	0
11	3	1	-1	1	1	0	0	0	0
12	3	1	-1	-1	-1	0	0	-1	0
13	4	-1	-1	1	-1	1	-1	0	-1
14	4	-1	-1	-1	-1	1	-1	0	0
15	4	-1	-1	-1	1	1	-1	1	1
16	4	-1	-1	1	1	1	-1	-1	1
17	5	1	1	-1	0	-1	-1	-1	1
18	5	1	1	0	1	-1	-1	1	-1
19	5	1	1	1	-1	-1	-1	-1	-1
20	5	1	1	-1	-1	-1	-1	1	1
21	6	1	0	0	-1	0	0	-1	0
22	6	1	0	1	1	0	0	0	0
23	6	1	0	1	0	0	0	0	1
24	6	1	0	-1	1	0	0	0	-1
25	7	0	1	1	1	1	1	1	0
26	7	0	1	-1	1	1	1	-1	-1
27	7	0	1	0	0	1	1	0	1
28	7	0	1	-1	-1	1	1	1	-1

Table 3: D- and I-optimal 28-run staggered-level designs for estimating a response surface model in two hard-to-change factors w and s, and two easy-to-change factors t_1 and t_2, assuming η_δ and η_γ equal to one.

Run	D-optimal				I-optimal			
	w	s	t_1	t_2	w	s	t_1	t_2
1	-1	0	-1	1	0	0	0	0
2	-1	0	0	-1	0	0	1	-1
3	-1	-1	-1	0	0	1	0	-1
4	-1	-1	1	1	0	1	1	1
5	1	-1	1	-1	1	1	1	-1
6	1	-1	-1	1	1	1	-1	0
7	1	1	1	0	1	0	0	-1
8	1	1	-1	-1	1	0	0	1
9	-1	1	1	-1	-1	0	-1	1
10	-1	1	-1	1	-1	0	1	0
11	-1	-1	-1	-1	-1	-1	1	1
12	-1	-1	1	0	-1	-1	0	-1
13	0	-1	0	-1	1	-1	-1	1
14	0	-1	-1	1	1	-1	1	0
15	0	1	1	1	1	0	1	1
16	0	1	-1	0	1	0	-1	-1
17	1	1	0	-1	0	0	-1	0
18	1	1	-1	1	0	0	0	0
19	1	-1	1	1	0	1	-1	1
20	1	-1	-1	-1	0	1	1	0
21	-1	-1	0	1	-1	1	-1	-1
22	-1	-1	1	-1	-1	1	0	1
23	-1	1	1	1	-1	-1	-1	0
24	-1	1	-1	-1	-1	-1	1	-1
25	1	1	1	-1	0	-1	-1	-1
26	1	1	0	1	0	-1	0	1
27	1	0	-1	0	0	0	0	0
28	1	0	1	1	0	0	1	1

Table 4: D- and I-optimal 28-run split-split-plot designs for estimating a response surface model in two hard-to-change factors w and s, and two easy-to-change factors t_1 and t_2, assuming η_δ and η_γ equal to one. The column labelled WP indicates the seven whole plots, while the column labelled SP indicates the 14 subplots.

Run	WP	SP	D-optimal				I-optimal			
			w	s	t_1	t_2	w	s	t_1	t_2
1	1	1	1	-1	-1	-1	0	1	0	0
2	1	1	1	-1	1	1	0	1	-1	-1
3	1	2	1	1	-1	1	0	0	1	0
4	1	2	1	1	1	-1	0	0	0	-1
5	2	3	-1	1	1	-1	-1	-1	1	-1
6	2	3	-1	1	-1	1	-1	-1	-1	0
7	2	4	-1	-1	-1	-1	-1	1	0	1
8	2	4	-1	-1	1	1	-1	1	1	-1
9	3	5	-1	-1	1	-1	0	0	0	1
10	3	5	-1	-1	-1	1	0	0	0	0
11	3	6	-1	1	0	1	0	1	-1	0
12	3	6	-1	1	-1	-1	0	1	1	1
13	4	7	-1	1	1	1	0	-1	0	0
14	4	7	-1	1	-1	0	0	-1	-1	1
15	4	8	-1	0	1	0	0	0	0	-1
16	4	8	-1	0	0	-1	0	0	1	0
17	5	9	1	1	1	0	1	-1	-1	0
18	5	9	1	1	-1	-1	1	-1	1	1
19	5	10	1	-1	-1	1	1	1	1	-1
20	5	10	1	-1	1	-1	1	1	-1	1
21	6	11	1	0	0	1	1	-1	1	-1
22	6	11	1	0	-1	0	1	-1	0	1
23	6	12	1	1	0	-1	1	0	-1	-1
24	6	12	1	1	1	1	1	0	0	0
25	7	13	0	-1	1	-1	-1	-1	1	1
26	7	13	0	-1	0	0	-1	-1	-1	-1
27	7	14	0	0	-1	1	-1	0	-1	1
28	7	14	0	0	0	0	-1	0	0	0

Table 5: Variances of estimates of fixed model parameters along with the D- and I-efficiencies for the 28-run split-plot designs in Table 2, the split-split-plot designs in Table 4 and the staggered-level designs in Table 3 when $\eta_\delta = 1$, $\eta_\gamma = 1$ and $\sigma_\varepsilon^2 = 1$.

Effect	Split-Plot (Table 2)		Split-Split-Plot (Table 4)		Staggered-Level (Table 3)	
	D-optimal	I-optimal	D-optimal	I-optimal	D-optimal	I-optimal
Intercept	4.838	1.175	2.130	0.781	3.225	0.824
w	0.376	0.569	0.297	0.459	0.222	0.348
s	0.570	0.569	0.157	0.217	0.215	0.372
t_1	0.046	0.069	0.051	0.070	0.048	0.063
t_2	0.044	0.062	0.051	0.077	0.049	0.063
ws	0.566	0.573	0.164	0.303	0.099	0.266
wt_1	0.054	0.103	0.056	0.106	0.054	0.095
wt_2	0.050	0.085	0.056	0.105	0.054	0.098
st_1	0.062	0.088	0.057	0.092	0.055	0.102
st_2	0.058	0.075	0.057	0.097	0.054	0.097
t_1t_2	0.052	0.091	0.137	0.154	0.065	0.108
w^2	3.412	3.420	2.296	1.117	1.848	0.889
s^2	1.717	3.960	0.936	0.599	1.346	0.703
t_1^2	0.267	0.190	0.321	0.260	0.331	0.214
t_2^2	0.340	0.208	0.320	0.203	0.328	0.207
D-efficiency	0.773	0.657	0.920	0.788	1.000	0.809
I-efficiency	0.327	0.523	0.619	1.025	0.491	1.000

should be run twice to produce comparable prediction variances to the staggered-level design. This result is again to a large extent due to an imprecise estimation of the main effect and the quadratic effect of both hard-to-change factors w and s, as well as their interaction effect. However, the less precise estimation of the intercept also contributes to the poorer I-efficiency for the split-plot design. As a matter of fact, when multiplying the variance-covariance matrix of the parameter estimates with the matrix $\mathbf{B}$ in Equation (10) to calculate the average prediction variance, the variance of the estimate of the intercept is multiplied with a factor of 2^4 and therefore has a major impact on the I-efficiency of a design.

The least cost-efficient split-split-plot design turns out to best design option of the three in terms of the I-optimality criterion. In comparison to the I-optimal staggered-level design, the intercept as well as the main effect and the quadratic effect of s are estimated more precisely from the split-split-plot design. The I-optimal staggered-level design, however, produces more precise estimates of the main effect and the quadratic effect of the other hard-to-change factor, w. Ultimately, the I-optimal split-split-plot design is 2.5% more I-efficient than the I-optimal staggered-level design.

4.4 Fraction of Design Space plot

In Figure 1, we use a Fraction of Design Space (FDS) plot to compare the predictive performance of the three I-optimal designs assuming the variance ratios η_δ and η_γ are equal to one. In the plot, the dotted line shows the predictive performance of the I-optimal staggered-level design. The solid line corresponds to the I-optimal split-plot design and the dashed line shows the predictive performance of the I-optimal split-split-plot design. Each point on the horizontal axis of the FDS plot corresponds to a fraction of the design space. The vertical axis covers the range from the minimum prediction variance to the maximum prediction variance, relative to σ_ε^2. Suppose, for example, that the point (0.75, 1.1) is on the FDS curve. Then, the variance of prediction, relative to σ_ε^2, is less than or equal to 1.1 over 75% of the design region. For a design to be good in terms of predictive performance, its FDS curve should be as low as possible. This actually means that the design results in small prediction variances in large fractions of the experimental region.

Figure 1 confirms that the split-plot design is indeed the worst option of the three in terms of predictive performance and that there is a negligible difference between the cost-efficient staggered-level design and the more expensive split-split-plot design.

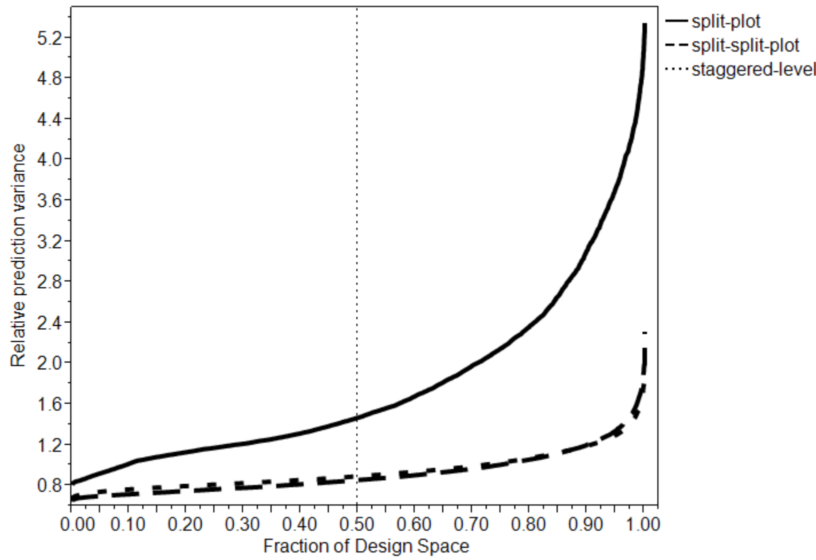

Figure 1: Fraction of Design Space (FDS) plot for the 28-run I-optimal split-plot, split-split-plot and staggered-level designs, constructed assuming $\eta_\delta = 1$ and $\eta_\gamma = 1$.

4.5 Comparison between D-optimal and I-optimal staggered-level design

A comparison of the D- and I-optimal staggered-level designs in Table 3, in terms of their factor level settings, reveals a greater emphasis on the center of the experimental region for the I-optimal design. In the D-optimal design, the class-1 hard-to-change factor w is set at its middle level only once, the class-2 hard-to-change factor s twice, and there are no center runs. On the other hand, in the I-optimal design, w has three settings at its middle level, s has four of them and there are three center runs. This greater emphasis on the center of the experimental region is typical of I-optimal designs, and is also discussed by Hardin and Sloane (1993) and by Jones and Goos (2012b) for completely randomized desings and split-plot designs .

As a result of this greater emphasis on the center of the experimental region, the intercept and the quadratic effects are estimated more precisely from the I-optimal staggered-level design, while the D-optimal staggered-level design produces more precise estimates of the main effects and two-factor interaction effects.

Evaluating the D-optimal staggered-level design in terms of the I-optimality criterion and the I-optimal staggered-level design in terms of the D-optimality criterion confirms that

Table 6: D- and I-efficiencies resulting from misspecification of the η_δ and η_γ assuming these variance ratios and σ_ε^2 are equal to one when constructing a 28-run staggered-level design.

	D-efficiency			I-efficiency		
	η_δ			η_γ		
η_γ	0.1	1	10	0.1	1	10
0.1	0.979	0.981	0.981	0.951	0.976	0.987
1	0.983	1.000	0.995	0.996	1.000	0.995
10	0.981	0.986	0.988	0.994	1.000	0.991

the D-optimal response surface design performs poorly in terms of the I-optimality criterion, while the I-optimal design performs reasonable well with respect to the D-optimality criterion in case of a cuboidal experimental region. The D-optimal staggered-level design has an I-efficiency of only 49.1% relative to the I-optimal staggered-level design, while the I-optimal staggered-level design has a D-efficiency of 80.9% relative to the D-optimal staggered-level design. The same kind of pattern in the D- and I-efficiencies can be observed for the split-plot and split-split-plot designs. This can be seen in Table 5.

4.6 Sensitivity study

In Section 3, we already mentioned that the D-optimal as well as the I-optimal staggered-level design depends on the variance ratios η_δ and η_γ. All optimal designs in Section 4 so far have been constructed assuming both variance ratios are equal to one. It is interesting to investigate how sensitive the D- and I-optimal staggered-level designs are to these assumptions. Therefore, we computed D- and I-optimal 28-run staggered-level designs for a 3 by 3 grid of η_δ and η_γ values from 0.1 to 1 to 10. The results of this sensitivity study are reported in Table 6. When η_δ and η_γ are one, the D-efficiencies of the designs found, relative to the optimal design, range from 97.9% to 99.5%, while the I-efficiencies range from 95.1% to 100%. So, a misspecification of the variance ratios η_δ and η_γ has no major impact on the efficiency of the staggered-level response surface design obtained.

5 20-run design

In the previous section, it turned out that the split-plot design, even though it had only one fewer level setting of the class-2 hard-to-change factor than the staggered-level design and was the most cost-efficient design option, perfomed poorly in terms of statistical efficiency. In this section, we therefore consider an experimental situation in which the split-plot design is no longer the most cost-efficient option. Suppose there is a budget to perform 20 runs to estimate a response surface model in two hard-to-change factors, w

Table 7: D- and I-optimal 20-run split-plot designs for estimating a response surface model in two hard-to-change factors w and s, and two easy-to-change factors t_1 and t_2, assuming η_δ and η_γ are equal to one.

Run	WP	D-optimal				I-optimal			
		w	s	t_1	t_2	w	s	t_1	t_2
1	1	1	-1	-1	-1	0	0	-1	1
2	1	1	-1	1	0	0	0	0	0
3	2	-1	-1	-1	1	-1	1	1	1
4	2	-1	-1	1	-1	-1	1	-1	0
5	3	-1	-1	1	1	1	-1	0	1
6	3	-1	-1	-1	-1	1	-1	1	-1
7	4	-1	1	-1	1	1	1	-1	1
8	4	-1	1	0	-1	1	1	1	-1
9	5	1	1	1	-1	-1	0	0	1
10	5	1	1	0	1	-1	0	1	-1
11	6	-1	1	-1	-1	0	0	1	0
12	6	-1	1	1	1	0	0	0	-1
13	7	1	0	1	1	1	0	-1	0
14	7	1	0	-1	0	1	0	1	1
15	8	1	-1	-1	1	-1	-1	1	1
16	8	1	-1	0	-1	-1	-1	-1	-1
17	9	0	0	1	-1	0	-1	-1	1
18	9	0	0	0	0	0	-1	0	0
19	10	0	1	1	0	0	1	-1	-1
20	10	0	1	-1	-1	0	1	0	0

and s, and two easy-to-change factors, t_1 and t_2.

5.1 Design options

To have enough degrees of freedom to estimate all fixed parameters and all variance components, the benchmark split-plot designs involve ten settings of both hard-to-change factors. The D- and I-optimal split-plot designs for this situation are shown in Table 7. The split-split-plot and the staggered-level designs are more cost-efficient design options, since they require fewer level settings of the hard-to-change factors. The D- and I-optimal split-split-plot designs, consisting of only five whole plots and ten subplots, are shown in Table 8. The staggered-level designs in Table 9 are, in this scenario, the most cost-efficient options, with five settings of the class-1 hard-to-change factor w and only six settings of the class-2 hard-to-change factor s. Table 10 provides an overview of the results obtained for the variance of the parameter estimates and the relative efficiencies assuming all variance

Table 8: D- and I-optimal 20-run split-split-plot designs for estimating a response surface model in two hard-to-change factors w and s, and two easy-to-change factors t_1 and t_2, assuming η_δ and η_γ are equal to one.

Run	WP	SP	D-optimal				I-optimal			
			w	s	t_1	t_2	w	s	t_1	t_2
1	1	1	1	-1	0	-1	0	-1	0	1
2	1	1	1	-1	1	1	0	-1	1	-1
3	1	2	1	1	-1	1	0	0	0	0
4	1	2	1	1	-1	0	0	0	-1	0
5	2	3	-1	1	-1	1	-1	-1	-1	-1
6	2	3	-1	1	1	-1	-1	-1	1	1
7	2	4	-1	-1	-1	-1	-1	1	1	-1
8	2	4	-1	-1	0	1	-1	1	-1	1
9	3	5	1	1	0	1	1	1	-1	1
10	3	5	1	1	1	-1	-1	1	1	-1
11	3	6	1	0	-1	0	1	0	0	-1
12	3	6	1	0	1	-1	1	0	1	0
13	4	7	0	0	1	1	1	-1	-1	0
14	4	7	0	0	0	0	1	-1	1	1
15	4	8	0	-1	1	0	1	0	0	0
16	4	8	0	-1	-1	1	1	0	-1	-1
17	5	9	-1	1	-1	-1	0	1	1	1
18	5	9	-1	1	1	1	0	1	-1	-1
19	5	10	-1	-1	1	-1	0	0	-1	1
20	5	10	-1	-1	-1	1	0	0	0	0

components equal to one.

5.2 Comparison in terms of the D-optimality criterion

The staggered-level design is not only the most cost-efficient design, but also statistically the most efficient one of the three D-optimal design options. The D-optimal staggered-level design is 12.6% and 6.4% more D-efficient than the D-optimal split-plot design and the D-optimal split-split-plot design, respectively. In comparison to the split-plot design, the good result for the staggered-level design is due to a more precise estimation of the two-factor interaction effect between the hard-to-change factors and a more precise estimation of the quadratic effect of the class-2 hard-to-change factor s, despite the fact that the staggered-level design has a smaller number of level settings for s than the split-plot design. Note that the staggered-level design also yields more precise estimates of the interaction effect between t_1 and t_2 and their quadratic effects.

Table 9: D- and I-optimal 20-run staggered-level designs for estimating a response surface model in two hard-to-change factors w and s, and two easy-to-change factors t_1 and t_2, assuming η_δ and η_γ are equal to one.

Run	D-Optimal				I-Optimal			
	w	s	t_1	t_2	w	s	t_1	t_2
1	0	0	0	0	0	1	0	1
2	0	0	-1	-1	0	1	1	0
3	0	-1	1	0	0	0	1	-1
4	0	-1	-1	1	0	0	0	0
5	-1	-1	1	1	1	0	0	0
6	-1	-1	0	-1	1	0	1	1
7	-1	1	1	1	1	-1	-1	1
8	-1	1	-1	1	1	-1	1	-1
9	1	1	1	1	-1	-1	-1	1
10	1	1	0	-1	-1	-1	1	-1
11	1	-1	-1	-1	-1	1	-1	-1
12	1	-1	0	1	-1	1	1	1
13	-1	-1	-1	0	1	1	-1	0
14	-1	-1	1	-1	1	1	0	-1
15	-1	1	1	1	1	0	0	0
16	-1	1	1	-1	1	0	-1	-1
17	1	1	1	-1	0	0	0	0
18	1	1	-1	1	0	0	-1	1
19	1	0	1	-1	0	-1	1	1
20	1	0	-1	0	0	-1	-1	-1

Table 10: Variances of estimates of fixed model parameters along with the D- and I-efficiencies for the 20-run split-plot designs in Table 7, the split-split-plot designs in Table 8 and the staggered-level designs in Table 9 when $\eta_\delta = 1$, $\eta_\gamma = 1$ and $\sigma_\varepsilon^2 = 1$.

Effect	Split-Plot (Table 7)		Split-Split-Plot (Table 8)		Staggered-Level (Table 10)	
	D-optimal	I-optimal	D-optimal	I-optimal	D-optimal	I-optimal
Intercept	2.209	1.007	2.502	1.192	2.432	1.365
w	0.371	0.463	0.482	0.707	0.354	0.517
s	0.370	0.429	0.233	0.346	0.353	0.423
t_1	0.075	0.115	0.075	0.093	0.074	0.110
t_2	0.079	0.117	0.079	0.097	0.079	0.102
ws	0.401	0.648	0.257	0.478	0.112	0.216
wt_1	0.091	0.136	0.091	0.114	0.091	0.168
wt_2	0.085	0.137	0.085	0.138	0.083	0.171
st_1	0.092	0.148	0.091	0.118	0.091	0.158
st_2	0.083	0.141	0.083	0.127	0.085	0.152
t_1t_2	0.280	0.273	0.186	0.253	0.156	0.209
w^2	2.168	1.185	2.607	1.598	2.355	1.393
s^2	2.080	1.110	1.443	0.967	1.551	1.038
t_1^2	0.687	0.444	0.537	0.430	0.515	0.390
t_2^2	0.607	0.413	0.539	0.528	0.517	0.390
D-efficiency	0.888	0.802	0.940	0.816	1.000	0.837
I-efficiency	0.784	1.159	0.746	1.005	0.767	1.000

Compared to the split-split-plot design, the statistical benefits of the staggered-level design are mainly the more precise estimation of the main effect and quadratic effect of the class-1 hard-to-change factor w as well as the two-factor interaction effect between both hard-to-change factors. The split-split-plot design allows a more precise estimation of the main effect of the class-2 hard-to-change factor s, due to the larger number of level settings of this factor.

5.3 Comparison in terms of the I-optimality criterion

The large number of level settings of the hard-to-change factors in the split-plot design option has a positive influence on the I-efficiency of the design. More particularly, it turns out that the I-optimal split-plot design, the design option with the largest number of level settings of the hard-to-change factors, is in fact the most I-efficient design. The more cost-efficient I-optimal staggered-level and I-optimal split-split-plot design are about 16% less I-efficient than the I-optimal split-plot design. Especially for the staggered-level design, this result is quite good, since a decrease of 50% and 40% in the number of level settings of w and s, respectively, leads to a decrease in I-efficiency of only 16%. So, the staggered-level design produces reasonably precise predictions despite the fact it is much more cost efficient.

A thorough comparison of the I-optimal staggered-level design and the I-optimal split-plot design reveals that the better result of the split-plot design in terms of the I-optimality criterion is mostly due to the more precise estimation of the intercept, as shown in Table 10. As mentioned before, the variance of the estimate of the intercept is important when calculating the I-optimality criterion value. The I-optimal split-plot design also scores better in terms of the estimation precision of the main effect and the quadratic effect of w.

In spite of the smaller number of level settings of the class-2 hard-to-change factor s, the I-optimal staggered-level design is equally good as the I-optimal split-split-plot design in terms of the I-optimality criterion. The staggered-level design allows a more precise estimation of the main effect and quadratic effect of the class-1 hard-to-change factor w and the two-factor interaction effect between both hard-to-change factors, whereas the split-split-plot design allows a more precise estimation of the intercept and the main effect and quadratic effect of the class-2 hard-to-change factor s.

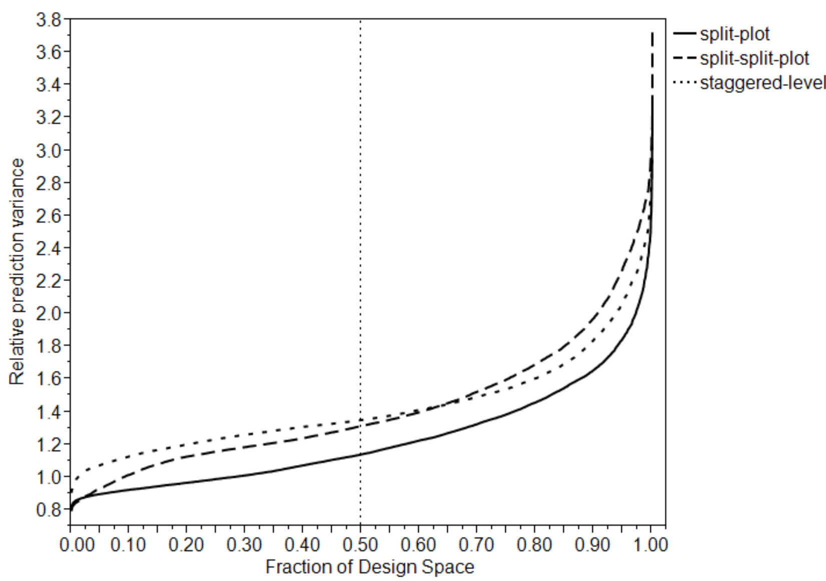

Figure 2: Fraction of Design Space (FDS) plot for the 20-run I-optimal split-plot, split-split-plot and staggered-level designs, constructed assuming $\eta_\delta = 1$ and $\eta_\gamma = 1$.

5.4 Fraction of Design Space plot

The FDS plot in Figure 2, constructed assuming that η_δ and η_γ equal one, confirms that the expensive I-optimal split-plot design performs better than the cost-efficient I-optimal split-split-plot and the I-optimal staggered-level design. Perhaps the most striking observation is that, in about 62.5% of the experimental region, the I-optimal split-split-plot design results in more precise predictions than the I-optimal staggered-level design. For the remaining fraction of the design space, the I-optimal staggered-level design has a lower prediction variance. The difference in precision of prediction between the two designs is, however, limited throughout the entire design region.

5.5 Comparison between D-optimal and I-optimal staggered-level design

Again, the I-optimal staggered-level design puts a greater emphasis on the center of the experimental region. This results in a more precise estimation of the intercept and all quadratic effects compared to the D-optimal staggered-level design. The latter, on the other hand, allows a more precise estimation of the main effects and two-factor interaction effects. In the I-optimal staggered-level design in Table 9, two runs (the center run and an axial run) are replicated, whereas the D-optimal staggered-level design has only one replicate (see rows 7 and 15).

Finally, the D-optimal staggered-level design performs reasonable well in terms of the I-optimality criterion. It has an I-efficiency of 76.7% relative to the I-optimal staggered-level design, while the I-optimal design has a D-efficiency of 83.7% relative to the D-optimal staggered-level design.

5.6 Sensitivity study

To conclude the discussion on the 20-run designs, we investigated the sensitivity of the D- and I-optimal 20-run staggered-level design to the assumption that both variance ratios η_δ and η_γ are equal to one. The results of the sensitivity study are reported in Table 11. When η_δ and η_γ are one, the I-efficiencies of the designs found, relative to the optimal design, range from 97.2% to 100%. For the D-efficiencies, the differences are a bit larger sinces these efficiencies range from 94.7% to 99.7%. Thus, the efficiencies of the staggered-level response surface design are not substantially influenced by a misspecification of the variance ratios η_δ and η_γ

Table 11: D- and I-efficiencies resulting from misspecification of the η_δ and η_γ assuming these variance ratios and σ^2_ε are equal to one when constructing a 20-run staggered-level response surface design.

	D-efficiency			I-efficiency		
	η_δ			η_γ		
η_γ	0.1	1	10	0.1	1	10
0.1	0.955	0.989	0.995	0.972	0.990	0.980
1	0.997	1.000	0.996	0.991	1.000	0.997
10	0.947	0.947	0.989	0.991	0.978	1.000

6 36-run design

6.1 Design options

In the last example, we consider an experiment with one class-1 hard-to-change factor w, one class-2 hard-to-change factor s, and three easy-to-change factors t_1, t_2 and t_3. Suppose there is a budget to perform 36 runs. A cost-efficient staggered-level design with a sufficient number of degrees of freedom to estimate all fixed effects and variance components is one with six settings of w and seven settings of s. The designs in Table 12 are the D- and I-optimal staggered-level designs for estimating a response surface model assuming both variance ratios η_δ and η_γ equal to one.

The split-plot alternatives we consider are the 36-run split-plot designs with nine whole plots shown in Table 13. These design options have enough degrees of freedom to estimate all fixed parameters and variance components and offer reasonable benchmarks for the staggered-level designs in terms of experimental cost. Another option would be a split-plot design with twelve whole plots, but this option is too expensive in comparison to the staggered-level designs. The split-split-plot alternatives we consider are the 36-run split-split-plot designs in Table 14 with six whole plots and twelve subplots.

In Table 15, the variances of the parameter estimates of all the presented design options are compared assuming that $\eta_\delta = 1$, $\eta_\gamma = 1$ and $\sigma^2_\varepsilon = 1$, as well as the values of the D- and I-optimality criteria. For both optimality criteria, the results are in favor of the staggered-level designs, despite the fact that the design involves the smallest number of level settings of the hard-to-change factors.

6.2 Comparison in terms of the D-optimality criterion

In comparison to the D-optimal staggered-level design, the main effects and quadratic effects of both hard-to-change factors as well as their interaction effect are estimated less

Table 12: D- and I-optimal 36-run staggered-level designs for estimating a response surface model in two hard-to-change factors w and s, and three easy-to-change factors t_1, t_2 and t_3, assuming η_δ and η_γ are equal to one.

Run	D-optimal					I-optimal				
	w	s	t_1	t_2	t_3	w	s	t_1	t_2	t_3
1	-1	0	1	1	-1	0	0	0	0	-1
2	-1	0	-1	0	1	0	0	0	-1	0
3	-1	0	0	-1	0	0	0	-1	0	0
4	-1	1	1	-1	-1	0	-1	1	0	1
5	-1	1	-1	1	0	0	-1	0	0	-1
6	-1	1	0	0	1	0	-1	-1	1	0
7	1	1	-1	-1	-1	-1	-1	-1	-1	-1
8	1	1	1	0	1	-1	-1	1	1	0
9	1	1	-1	1	1	-1	-1	0	1	1
10	1	-1	1	-1	-1	-1	1	-1	0	0
11	1	-1	-1	1	-1	-1	1	1	-1	-1
12	1	-1	-1	-1	1	-1	1	1	1	1
13	-1	-1	1	0	0	1	1	0	-1	0
14	-1	-1	0	1	1	1	1	1	1	-1
15	-1	-1	-1	-1	-1	1	1	-1	1	1
16	-1	1	-1	-1	1	1	0	1	1	1
17	-1	1	-1	1	-1	1	0	-1	-1	1
18	-1	1	1	0	0	1	0	0	0	-1
19	1	1	-1	1	1	-1	0	1	-1	1
20	1	1	1	1	-1	-1	0	-1	1	1
21	1	1	1	-1	1	-1	0	0	0	-1
22	1	-1	1	1	1	-1	1	-1	-1	1
23	1	-1	-1	1	-1	-1	1	-1	1	-1
24	1	-1	0	-1	-1	-1	1	1	0	0
25	-1	-1	1	1	-1	1	1	1	-1	1
26	-1	-1	1	-1	1	1	1	0	1	0
27	-1	-1	-1	1	1	1	1	-1	-1	-1
28	-1	1	1	1	1	1	-1	-1	1	-1
29	-1	1	-1	-1	0	1	-1	1	-1	-1
30	-1	1	0	0	-1	1	-1	0	0	1
31	0	1	1	1	0	0	-1	0	-1	0
32	0	1	0	-1	1	0	-1	1	1	-1
33	0	1	-1	0	-1	0	-1	-1	0	1
34	0	0	1	-1	-1	0	0	0	0	0
35	0	0	-1	-1	1	0	0	0	1	0
36	0	0	0	1	0	0	0	1	0	0

Table 13: D- and I-optimal 36-run split-plot designs for estimating a response surface model in two hard-to-change factors w and s, and three easy-to-change factors t_1, t_2 and t_3, assuming η_δ and η_γ are equal to one.

Run	WP	D-optimal					I-optimal				
		w	s	t_1	t_2	t_3	w	s	t_1	t_2	t_3
1	1	1	-1	0	1	1	0	0	0	0	1
2	1	1	-1	1	1	-1	0	0	-1	1	0
3	1	1	-1	1	-1	1	0	0	0	1	0
4	1	1	-1	-1	-1	0	0	0	-1	0	-1
5	2	0	-1	0	-1	-1	0	1	0	0	0
6	2	0	-1	1	1	0	0	1	1	1	0
7	2	0	-1	-1	0	-1	0	1	0	-1	1
8	2	0	-1	-1	1	1	0	1	1	0	-1
9	3	1	1	-1	1	-1	1	0	0	0	0
10	3	1	1	1	1	1	1	0	1	1	-1
11	3	1	1	-1	-1	1	1	0	0	-1	-1
12	3	1	1	1	-1	-1	1	0	-1	0	0
13	4	-1	1	1	-1	0	-1	-1	0	0	0
14	4	-1	1	1	1	1	-1	-1	-1	-1	-1
15	4	-1	1	0	1	-1	-1	-1	1	-1	1
16	4	-1	1	-1	0	1	-1	-1	1	1	-1
17	5	1	-1	-1	-1	1	-1	1	1	1	1
18	5	1	-1	1	-1	-1	-1	1	1	-1	-1
19	5	1	-1	-1	1	-1	-1	1	-1	1	-1
20	5	1	-1	1	0	1	-1	1	-1	0	1
21	6	-1	-1	-1	-1	1	1	1	1	-1	1
22	6	-1	-1	1	-1	-1	1	1	-1	-1	-1
23	6	-1	-1	-1	1	-1	1	1	-1	1	1
24	6	-1	-1	1	1	1	1	1	0	1	-1
25	7	-1	0	-1	-1	-1	1	-1	1	1	1
26	7	-1	0	1	1	-1	1	-1	1	-1	-1
27	7	-1	0	1	-1	1	1	-1	-1	-1	1
28	7	-1	0	0	0	0	1	-1	-1	1	-1
29	8	-1	1	1	0	-1	-1	0	-1	1	1
30	8	-1	1	-1	-1	-1	-1	0	-1	-1	0
31	8	-1	1	-1	1	0	-1	0	0	0	-1
32	8	-1	1	0	-1	1	-1	0	1	0	0
33	9	1	1	1	-1	1	0	0	1	0	1
34	9	1	1	1	1	-1	0	0	0	0	0
35	9	1	1	-1	-1	-1	0	0	0	-1	0
36	9	1	1	-1	1	1	0	0	0	0	0

Table 14: D- and I-optimal 36-run split-split-plot designs for estimating a response surface model in two hard-to-change factors w and s, and three easy-to-change factors t_1, t_2 and t_3, assuming η_δ and η_γ are equal to one.

Run	WP	SP	D-optimal					I-optimal				
			w	s	t_1	t_2	t_3	w	s	t_1	t_2	t_3
1	1	1	-1	1	-1	-1	-1	0	0	0	0	0
2	1	1	-1	1	1	-1	1	0	0	0	1	-1
3	1	1	-1	1	-1	1	1	0	0	1	0	0
4	1	2	-1	-1	-1	-1	-1	0	-1	0	0	-1
5	1	2	-1	-1	1	1	-1	0	-1	-1	1	1
6	1	2	-1	-1	0	-1	1	0	-1	1	-1	1
7	2	3	1	-1	1	1	0	0	0	-1	0	0
8	2	3	1	-1	-1	1	-1	0	0	0	0	-1
9	2	3	1	-1	-1	-1	1	0	0	0	-1	0
10	2	4	1	1	1	-1	1	0	1	0	0	-1
11	2	4	1	1	-1	1	1	0	1	1	1	1
12	2	4	1	1	1	1	-1	0	1	-1	-1	1
13	3	5	-1	1	1	1	0	-1	1	-1	1	-1
14	3	5	-1	1	-1	-1	1	-1	1	0	0	1
15	3	5	-1	1	0	0	-1	-1	1	1	-1	-1
16	3	6	-1	-1	1	-1	1	-1	-1	-1	-1	-1
17	3	6	-1	-1	-1	1	1	-1	-1	1	1	-1
18	3	6	-1	-1	0	1	-1	-1	-1	0	0	1
19	4	7	1	-1	-1	1	1	1	1	1	1	-1
20	4	7	1	-1	1	-1	1	1	1	-1	1	1
21	4	7	1	-1	-1	-1	-1	1	1	0	-1	0
22	4	8	1	1	-1	0	0	1	0	1	-1	-1
23	4	8	1	1	0	-1	-1	1	0	-1	0	0
24	4	8	1	1	1	1	1	1	0	0	0	1
25	5	9	0	0	-1	-1	1	1	-1	1	1	0
26	5	9	0	0	0	1	0	1	-1	-1	-1	1
27	5	9	0	0	1	0	-1	1	-1	-1	1	-1
28	5	10	0	1	0	0	1	1	1	1	-1	1
29	5	10	0	1	-1	1	-1	1	1	-1	-1	-1
30	5	10	0	1	1	-1	-1	1	1	0	1	0
31	6	11	-1	0	-1	1	-1	0	-1	1	1	1
32	6	11	-1	0	0	-1	0	0	-1	1	-1	-1
33	6	11	-1	0	1	0	1	0	-1	-1	0	0
34	6	12	-1	-1	-1	0	0	0	0	1	0	0
35	6	12	-1	-1	1	-1	-1	0	0	0	1	0
36	6	12	-1	-1	1	1	1	0	0	-1	-1	1

precisely from the D-optimal split-plot design. This result is striking since the split-plot design involves more level settings of both hard-to-change factors. So, it seems that, as long as the number of level settings of w in the split-plot design is not twice the number of level settings in the staggered-level design, the split-plot design performs substantially weaker when it comes to estimating the hard-to-change factors' effects. As a result, in this example, the split-plot design has a D-efficiency of 91.5% relative to the D-optimal staggered-level design.

The D-optimal split-split-plot design performs slightly better with a D-efficiency of 95.5%. The biggest differences with the D-optimal staggered-level design are a less precise estimation of the main effect and quadratic effect of w and the interaction effect between both hard-to-change factors from the split-split-plot design and a less precise estimation of the main effect of s from the staggered-level design.

6.3 Comparison in terms of the I-optimality criterion

As in Section 5, the I-optimal split-split-plot design and the I-optimal staggered-level design are almost equally efficient, despite the fact that the I-optimal staggered-level design is more cost efficient than the I-optimal split-split-plot design. This time, the I-efficiency of the split-split-plot design relative to the staggered-level design equals 98.8%. The I-optimal staggered-level design offers a more precise estimation of the main effect and the quadratic effect of w, and the two-factor-interaction effect between w and s, while the I-optimal split-split-plot design offers the advantage of a more precise estimation of the intercept, the main effect and the quadratic effect of s.

The I-optimal split-plot design has an I-efficiency of 89.6% relative to the I-optimal staggered-level design. This is mainly due to an imprecise etimation of the main effect of s, the two-factor interaction effect between both hard-to-change factors as well as their quadratic effects.

6.4 Fraction of Design Space plot

In the FDS plot in Figure 3, we see that the I-optimal split-split-plot design has a smaller prediction variance than the I-optimal staggered-level design in 60% of the experimental region. This is surprising given that the staggered-level design has the smallest average variance of prediction. It turns out, however, that the split-split-plot design has a substantially larger maximum prediction variance than the staggered-level design. Despite its larger cost, the I-optimal split-plot design generally results in larger prediction variances than the other two designs.

Table 15: Variances of estimates of fixed model parameters along with the D- and I-efficiencies for the 36-run split-plot designs in Table 13, the split-split-plot designs in Table 14 and the staggered-level designs in Table 12 when $\eta_\delta = 1$, $\eta_\gamma = 1$ and $\sigma_\varepsilon^2 = 1$.

Effect	Split-Plot (Table 13)		Split-Split-Plot (Table 14)		Staggered-Level (Table 12)	
	D-optimal	I-optimal	D-optimal	I-optimal	D-optimal	I-optimal
Intercept	6.298	0.944	2.280	0.784	2.279	1.042
w	0.353	0.379	0.363	0.660	0.262	0.324
s	0.353	0.514	0.153	0.207	0.260	0.326
t_1	0.035	0.051	0.040	0.050	0.043	0.045
t_2	0.035	0.054	0.040	0.052	0.040	0.047
t_3	0.034	0.047	0.038	0.046	0.039	0.046
ws	0.359	0.577	0.163	0.312	0.053	0.113
wt_1	0.039	0.061	0.047	0.101	0.046	0.063
wt_2	0.040	0.061	0.044	0.104	0.048	0.062
wt_3	0.039	0.061	0.047	0.088	0.045	0.060
st_1	0.038	0.082	0.045	0.063	0.046	0.064
st_2	0.038	0.072	0.043	0.063	0.047	0.072
st_3	0.038	0.070	0.044	0.056	0.045	0.066
t_1t_2	0.040	0.064	0.056	0.077	0.053	0.079
t_1t_3	0.039	0.061	0.053	0.074	0.058	0.076
t_2t_3	0.042	0.063	0.057	0.072	0.058	0.085
w^2	3.170	1.395	2.312	1.301	1.904	1.101
s^2	3.140	1.235	1.229	0.707	1.158	0.789
t_1^2	0.270	0.182	0.224	0.170	0.223	0.168
t_2^2	0.273	0.170	0.252	0.215	0.221	0.188
t_3^2	0.276	0.201	0.266	0.211	0.230	0.191
D-efficiency	0.915	0.774	0.955	0.789	1.000	0.866
I-efficiency	0.295	0.896	0.636	0.988	0.656	1.000

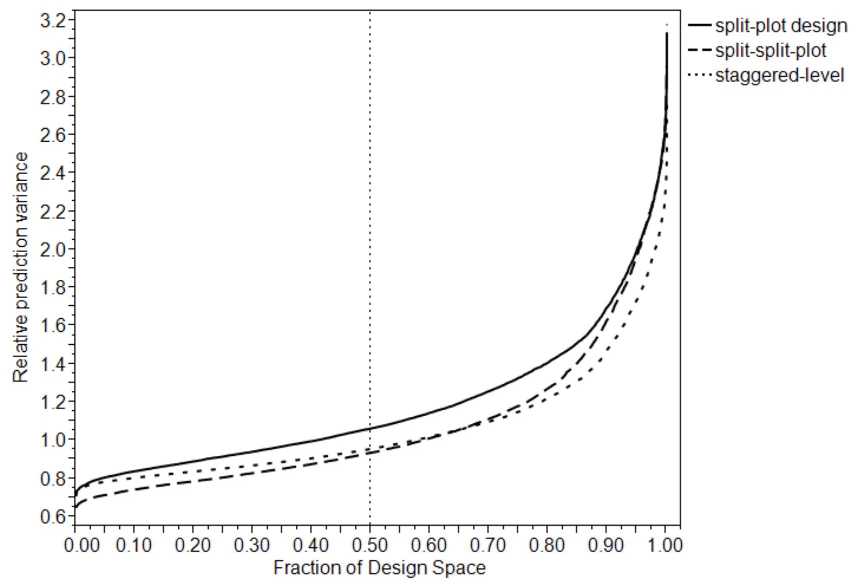

Figure 3: Fraction of Design Space (FDS) plot for the 36-run I-optimal split-plot, split-split-plot and staggered-level designs, constructed assuming $\eta_\delta = 1$ and $\eta_\gamma = 1$.

6.5 Sensitivity study

Finally, a sensitivity study pointed out that a misspecification of the variance ratios η_δ and η_γ has a small impact on the efficiency of the D- and I-optimal 36-run staggered-level designs. The results of the sensitivity study are reported in Table 16. When η_δ and η_γ are one, the D-efficiencies of the designs found, relative to the optimal design, range from 97.6% to 100%. For the I-optimal designs, the I-efficiencies range from 98.1% to 100%.

7 Conclusion

In this paper, we introduced the staggered-level design as an interesting design option for response surface modelling. For the generation of optimal staggered-level designs we focused on the D-optimality criterion as well as the I-optimality criterion. Comparing D- and I-optimal staggered-level designs in different experimental situations showed a greater emphasis on the center of the design region for the I-optimal staggered-level design. This emphasis is typical for I-optimal designs and has been observed for completely randomized designs and split-plot designs too. Consequently, the intercept as well as the quadratic effects are estimated more precisely from I-optimal staggered-level designs. The D-optimal staggered-level designs produce more precise estimates of the main effects and

Table 16: D- and I-efficiencies resulting from misspecification of the η_δ and η_γ assuming these variance ratios and σ_ε^2 are equal to one when constructing a 36-run staggered-level response surface design.

	D-efficiency			I-efficiency		
	η_δ			η_γ		
η_γ	0.1	1	10	0.1	1	10
0.1	0.982	0.985	0.981	0.990	1.000	0.995
1	0.976	1.000	1.000	0.991	1.000	0.998
10	0.976	0.980	0.994	0.981	1.000	0.986

the two-factor interaction effects.

In comparison to the split-plot and the split-split-plot designs, the staggered-level designs turn out to be cost-efficient as well as statistically efficient design options, also in a response surface modeling context. In general, the staggered-level designs turn out to be the most D-efficient design options of the three. For the I-optimality criterion, the statistical differences between the staggered-level design and the split-split-plot design were often negligible. This result is in favor of the staggered-level design, because that type of design is generally the least costly of the two. On the other hand, there were considerable differences in I-efficiency between the staggered-level design and the split-plot design, in favor of the staggered-level design.

References

Arnouts, H. & Goos, P. (2012). Staggered-level designs for experiments with more than one hard-to-change factor, *Technometrics* **54**: 355–366.

Bingham, D. R., Schoen, E. D. & Sitter, R. R. (2004). Designing fractional factorial split-plot experiments with few whole-plot factors, *Journal of the Royal Statistical Society, Series C* **53**: 325–339. Corrigendum, **54**, 955–958.

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

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

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

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

Hardin, R. H. & Sloane, N. J. A. (1993). A new approach to the construction of optimal designs, *Journal of Statistical Planning and Inference* **37**: 339–369.

Hinkelmann, K. & Kempthorne, O. (2008). *Design and Analysis of Experiments: Volume 1, Introduction to Experimental Design*, John Wiley Sons, Inc.

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

Jones, B. & Goos, P. (2007). A candidate-set-free algorithm for generating D-optimal split-plot designs, *Applied Statistics* **56**: 347–364.

Jones, B. & Goos, P. (2009). D-optimal design of split-split-plot experiments, *Biometrika* **96**: 67–82.

Jones, B. & Goos, P. (2012a). An algorithm for finding D-efficient equivalent-estimation second-order split-plot designs, *Journal of Quality Technology* **44**: 363–374.

Jones, B. & Goos, P. (2012b). I-optimal versus D-optimal split-plot resonse surface designs, *Journal of Quality Technology* **44**: 85–101.

Kulahci, M. & Bisgaard, S. (2005). The use of Plackett and Burman designs to construct split-plot designs, *Technometrics* **47**: 495–502.

Macharia, H. & Goos, P. (2010). D-optimal and D-efficient equivalent-estimation second-order split-plot designs, *Journal of Quality Technology* **42**: 358–372.

Mylona, K., Goos, P. & Jones, B. (2013). Optimal design of blocked and split-plot experiments for fixed effects and variance component estimation. To appear in Technometrics.

Parker, P. A., Kowalski, S. M. & Vining, G. G. (2006). Classes of split-plot response surface designs for equivalent estimation, *Quality and Reliability Engineering International* **22**: 291–305.

Parker, P. A., Kowalski, S. M. & Vining, G. G. (2007a). Construction of balanced equivalent estimation second-order split-plot designs, *Technometrics* **49**: 56–65.

Parker, P. A., Kowalski, S. M. & Vining, G. G. (2007b). Unbalanced and minimal point equivalent estimation second-order split-plot designs, *Journal of Quality Technology* **39**: 376–388.

Schoen, E. D. (1999). Designing fractional two-level experiments with nested error structures, *Journal of Applied Statistics* **26**: 495–508.

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

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

Webb, D. F., Lucas, J. M. & Borkowski, J. J. (2004). Factorial experiments when factor levels are not necessarily reset, *Journal of Quality Technology* **36**: 1–11.