On the Analysis of Balanced Two-Level Factorial Whole-Plot Saturated Split-Plot Designs

Marcus B. Perry\textsuperscript{1}, Gary R. Mercado\textsuperscript{2}
Department of Information Systems, Statistics & Management Science
The University of Alabama
300 Alston Hall
Box 870226
Tuscaloosa, AL 35487
USA

James R. Simpson\textsuperscript{3}
Group OA, 53d Test Management Group
United States Air Force
Eglin AFB FL, 32542
USA

Accepted Version

\textsuperscript{1}Corresponding author. Ph: (205) 348-9864, Email: mperry@cba.ua.edu
\textsuperscript{2}Ph: (205) 348-9864, Email: grmercadovelasco@crimson.ua.edu
\textsuperscript{3}Ph: (850) 882-0619, Email: James.Simpson@eglin.af.mil
Abstract

This paper considers an experimentation strategy when resource constraints permit only a single design replicate per time interval, and one or more design variables are hard-to-change. The experimental designs considered are two-level full or fractional factorial designs run as balanced split-plots. These designs are common in practice and appropriate for fitting a main effects plus interactions model, while minimizing the number of times the whole-plot treatment combination is changed. Depending on the postulated model, single replicates of these designs can result in the inability to estimate error at the whole-plot level, suggesting formal statistical hypothesis testing on the whole-plot effects is not possible. We refer to these designs as balanced two-level whole-plot saturated split-plot designs. In this paper we show that, for these designs, it is appropriate to use ordinary least squares (OLS) to analyze the subplot factor effects at the “intermittent” stage of the experiments (i.e., after a single design replicate is run); however, formal inference on the whole-plot effects may or may not be possible at this point. We exploit the sensitivity of OLS in detecting whole-plot effects in a split-plot design and propose a data-based strategy for determining whether to run an additional replicate following the intermittent analysis, or whether to simply reduce the model at the whole-plot level to facilitate testing. The performance of the proposed strategy is assessed using Monte Carlo simulation. The method is then illustrated using wind tunnel test data obtained from a NASCAR Winston Cup Chevrolet Monte Carlo stock car.

Keywords: Design of experiments, limited resources, randomization restrictions, split-plot design.
1 Introduction

Restrictions in randomization are commonplace in experimentation practices and often result from having experimental variables that are hard or costly to change during the course of the experiment. In most cases, the randomization restriction will lead to a split-plot treatment structure, where experimental factors are classified as either whole-plot or subplot factors. The whole-plot factors correspond to those factors that are difficult or costly to change, whereas the subplot factors correspond to those experimental factors that are relatively easy to change. As a consequence, the subplot factor effects are estimated with greater precision than the whole-plot factor effects, suggesting the power to detect factor effects at the subplot level is greater than the power to detect factor effects at the whole-plot level.

Due to the costs associated with changing the settings of the hard-to-change factors, two-level full factorial split-plot designs are a popular choice by practitioners. These designs will minimize the number of times the whole-plot factors are changed throughout the course of the experiment. If overall experimental cost is proportional to the total number of experimental runs\(^4\), then fractions of the two-level factorial split-plot designs are often considered (c.f. Bisgaard\(^1\), Bingham and Sitter\(^2\), and Loeppky and Sitter\(^3\)). To complicate matters, practitioners are often only afforded a single replicate of the design, at least initially. As a consequence, for a main effect and two-factor interaction model, if the number of whole-plot factors is less than 3 (which is often the case in practice), then a single replicate of a full factorial split-plot design will result in the inability to estimate the experimental error at the whole-plot level, thus, prohibiting formal statistical testing on the whole-plot factor effects. Further, a single replicate of a

\(^4\)In addition to the costs associated with changing levels of the hard-to-change factors, an economical run size is also desired.
two-level fractional factorial split-plot design will often result in a fully saturated design, prohibiting formal statistical testing on all factor effects.

To illustrate the case for the two-level full factorial design, consider the balanced $2^4$ split-plot design shown in Table 1, where $x_1$ and $x_2$ are hard-to-change factors and $x_3$ and $x_4$ are easy-to-change factors. A balanced split-plot design is one having the same number of subplots per whole-plot, and thus, for this design, there are 4 total degrees of freedom available at the whole-plot level and 12 total degrees of freedom available at the subplot level. Suppose a main effect and two-factor interaction model is postulated, then there are 4 whole-plot model degrees of freedom (including the intercept) and 7 subplot model degrees of freedom. Although there are $12 - 7 = 5$ degrees of freedom available for estimating error at the subplot level, there are $4 - 4 = 0$ degrees of freedom available for estimating error at the whole-plot level. Thus, unless an external estimate of the variability between whole-plots is available, formal statistical testing on the whole-plot effects is not possible. To avoid saturation, one can consider reducing the number of whole-plot model terms, or one can run additional whole-plots so that whole-plot error degrees of freedom are made available. The former will often require proper justification, while the latter is least desirable, or “a last resort” since running additional whole-plots is assumed to add significant cost to the overall experiment.

Consider now the two-level fractional factorial case. Table 2 shows a $1/2$ fraction of the $2^4$ split-plot design shown in Table 1, with $I = x_1x_2x_3x_4$ and 8 estimable columns (including the intercept). Note that if three-factor interactions and higher are negligible, then all main effects are free from aliasing, but two-factor interactions are aliased with each other. Thus, the design has resolution $IV$. For this design, $x_1$, $x_2$ and $x_1x_2 = x_3x_4$ are whole-plot factors and $x_3$, $x_4$, $x_1x_3 = x_2x_4$, $x_2x_3 = x_1x_4$ are subplot factors. Also, note that fractionating the design does not reduce the number of whole-plots, only the
total number of runs. Thus, since the design is balanced, there are 4 total whole-plot degrees of freedom and $8 - 4 = 4$ total subplot degrees of freedom. Further, there are 4 estimable columns at the whole-plot level (including the intercept) and 4 estimable columns at the subplot level. As a consequence, the design is saturated at both the whole-plot and subplot levels. That is, there are no degrees of freedom available to estimate whole-plot or subplot error, prohibiting formal statistical testing on all the factor effects. Several approaches to the analysis of these types of designs is discussed extensively in Loeppky and Sitter$^{(3)}$.

In this paper we consider balanced two-level factorial (full or fractional) split-plot designs that are saturated at the whole-plot level; however, degrees of freedom are available to estimate error at the subplot level. The design given in Table 1 falls into this category, as does the design in Table 3, which is the same two-level fractional factorial

Table 1: A single replicate of the $2^4$ factorial split-plot design with 2 HTC and 2 ETC factors.

<table>
<thead>
<tr>
<th>wp</th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>x1x2</th>
<th>x1x3</th>
<th>x1x4</th>
<th>x2x3</th>
<th>x2x4</th>
<th>x3x4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2: A single replicate of the $2^{4-1}\text{IV}$ fractional factorial split-plot design with 2 HTC and 2 ETC factors. The design generator is $I = x_1x_2x_3x_4$.

<table>
<thead>
<tr>
<th>wp</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_1x_2$</th>
<th>$x_1x_3$</th>
<th>$x_2x_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

design given in Table 2, except it has been augmented with subplot center runs to permit the estimation of subplot error.

There are three main contributions of this work. First, we provide a general overview of the analysis of split-plot designs using the generalized least squares (GLS) estimates of the model coefficients with restricted maximum likelihood (REML) estimates of the variance components, to include the addition of constraints on the estimated variance components if desired. Second, we show that intermittent analysis of the subplot effects using ordinary least squares (OLS) is appropriate for the designs considered herein, where the term "intermittent" implies after the first design replicate is run. If needed, and when time and/or resources are made available in the future, an additional replicate can be run in a separate block and an appropriate analysis subsequently performed on all effects (i.e., subplot and whole-plot). Lastly, we provide a data-based strategy that exploits the sensitivity of the OLS analysis in detecting whole-plot effects, and recommend this strategy for determining whether to run an additional replicate following the intermittent analysis, or whether the model can simply be reduced at the whole-plot level, either of which would facilitate testing of the whole-plot effects. Obviously, the former is most
desirable since it permits a pure-error estimate of the whole-plot error variance; however, it is also the most costly and will need to be justified. The proposed strategy ultimately seeks to provide this justification.

Table 3: A single replicate of the $2^4 - 1$ fractional factorial split-plot design with 2 HTC and 2 ETC factors (augmented with subplot center runs). The design generator for the factorial points is $I = x_1x_2x_3x_4$.

<table>
<thead>
<tr>
<th>wp</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_1x_2$</th>
<th>$x_1x_3$</th>
<th>$x_2x_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The remainder of this manuscript is organized as follows. In the next section we review the linear mixed-effects model used in modeling data obtained from split-plot experiments, including model specification, estimation, and inference. We discuss a maximum likelihood (ML) approach to the analysis of a split-plot design given a priori knowledge is available on the ratio of the variance components in the linear mixed model. We then review the REML approach to analysis when this ratio is unknown. Subsequently we discuss our data-based strategy for determining whether to run an additional replicate of the design, with the aim of facilitating formal statistical testing on the subplot and whole-plot effects. We discuss the performance of this approach, assessed via Monte Carlo simulation, to include trade-offs for the reduction in expected cost.
illustrate the application of our strategy to automobile wind tunnel experiments, with
the objective of characterizing the aerodynamic performance of a NASCAR Winston
Cup Chevrolet Monte Carlo stock car. Finally, we close with a summary and discussion.

2 Linear Model for Split-Plot Design

In this section we discuss the linear mixed-effect model used in analyzing data obtained
from a split-plot design, including model specification, estimation, and inference. It
should be noted that the estimation methods discussed in this section are not limited to
balanced two-level WPS split-plot designs; but more generally, they are applicable when
the design matrix has full rank.

2.1 Model Specification

The linear model for the split-plot design is given by

\[ \mathbf{y} = \mathbf{X}\beta + \mathbf{M}\delta + \epsilon \]  

(1)

where

\( \mathbf{y} \): \( N \times 1 \) vector of responses,

\( \mathbf{X} \): \( N \times p \) known design matrix of full rank,

\( \beta \): \( p \times 1 \) unknown vector of fixed effects,

\( \mathbf{M} \): \( N \times w \) known indicator matrix,

\( \delta \sim MVN(\mathbf{0}_{w \times 1}, \sigma^2_\delta \mathbf{I}_{w \times w}) \): \( w \times 1 \) random vector of errors,

\( \epsilon \sim MVN(\mathbf{0}_{N \times 1}, \sigma^2_\epsilon \mathbf{I}_{N \times N}) \): \( N \times 1 \) random error vector
and $w$ denotes the number of whole-plots, $p$ denotes the number of fixed-effect terms, and $N = \sum_{i=1}^{w} m_i$ denotes the total number of subplots, where $m_i$ is the number of subplots in whole-plot $i$. The matrix $M$ takes the form

$$M = \text{blkdiag}(1_{m_1}, 1_{m_2}, ..., 1_{m_w}),$$

where $\text{blkdiag}$ implies a block diagonal matrix. It is assumed that $\text{Cov}(\delta, \epsilon) = 0_{w \times N}$, and thus, it is easily shown that $E(y) = X\beta$ and

$$\text{Var}(y) = \sigma^2 \left\{ I_{N \times N} + \eta MM' \right\}$$

where $\eta = \sigma^2 / \sigma^2_\epsilon$ denotes the variance ratio. For more details on the linear mixed effects model discussed here, the reader is referred to Hocking\[^4\].

### 2.2 Model Estimation and Inference

In this subsection we discuss estimation and inference on the the fixed-effect component of the general model in eq.(1). We first consider the case where, although the variance components are unknown, the variance ratio is assumed to be known. Subsequently, we discuss the more likely case where the variance ratio is unknown and must be estimated from the experimental data.

#### 2.2.1 Case I: Known Variance Ratio

Of interest is the estimation and statistical testing of the fixed effect $\beta$. To accomplish this, we shall define the following matrices

$$D = I_{N \times N} + \eta MM'$$

$$F = \eta MM' D^{-1}$$

$$H = X(X'(I_{N \times N} - F)X)^{-1}X'$$
and note that the log-likelihood function of the unknown parameters can be written as

$$\ell(\beta, \eta, \sigma^2_\epsilon | y) = -\frac{1}{\sigma^2_\epsilon} (y - X\beta)'(I_{N \times N} - F)(y - X\beta) - N \log_e(\sigma^2_\epsilon) - \log_e(|D|).$$

Suppose that $\eta$ is known, but the variance components $\sigma^2_\delta$ and $\sigma^2_\epsilon$ are unknown. Then it is easy to show that the estimators for $\beta$ and $\sigma^2_\epsilon$ that maximize the log-likelihood function are

$$\hat{\beta}(\eta) = (X'(I_{N \times N} - F)X)^{-1}X'(I_{N \times N} - F)y \quad (4)$$

and

$$\hat{\sigma}^2_\epsilon(\eta) = \frac{y'Qy}{N} \quad (5)$$

where

$$Q = [(I_{N \times N} - H(I_{N \times N} - F))'(I_{N \times N} - F)(I_{N \times N} - H(I_{N \times N} - F))]. \quad (6)$$

However, the estimator for $\sigma^2_\epsilon$ given above generally produces biased estimates. Note that, if $\eta$ is known, then

$$E(y'Qy) = \sigma^2_\epsilon tr(Q) + \sigma^2_\delta tr(QMM')$$

$$= \sigma^2_\epsilon \{ tr(Q) + \eta tr(QMM') \}$$

$$= \sigma^2_\epsilon \text{rank}(Q).$$

Thus, to correct for this bias we can divide by $\text{rank}(Q) = N - p$ instead of $N$, producing the estimator

$$\hat{\sigma}^2_\epsilon(\eta) = \frac{y'Qy}{N - p}. \quad (7)$$

It is easily shown that

$$E[\hat{\beta}(\eta)] = \beta \quad (8)$$
and

$$Var[\hat{\beta}(\eta)] = \sigma^2 C$$  \hspace{1cm} (9)

where the \( p \times p \) matrix \( C \) is defined by

$$C = [X'(I_{N \times N} - F)X]^{-1}. \hspace{1cm} (10)$$

One can test the hypothesis \( H_0: \beta_h = 0 \) versus \( H_1: \beta_h \neq 0 \) by computing

$$t_0 = \frac{\hat{\beta}_h}{\sqrt{\hat{\sigma}^2(\eta)C_{hh}}} \hspace{1cm} (11)$$

where \( C_{hh} \) denotes the \( h^{th} \) diagonal entry of \( C \). Then under the null hypothesis \( H_0 \), the test statistic in equation (11) follows the \( t \) distribution with \( \text{rank}(Q) = N - p \) degrees of freedom.

It is possible that one might know the variance ratio, particularly if historical data from the same system or apparatus was available. However, when exact knowledge is lacking, the experimenter might produce a rough guess, or theorize about the magnitude of the variance ratio\(^5\).

### 2.2.2 Case II: Unknown Variance Ratio

If the variance ratio is unknown, then there must be sufficient whole-plot and subplot error degrees of freedom available to permit estimation. Assuming this is the case, typical \textit{model dependent} approaches to variance component estimation are maximum likelihood (ML) and restricted maximum likelihood (REML), where the latter method is often preferred as it produces unbiased estimates of the variance components (assuming the model is not underspecified). To estimate the variance ratio using the experimental data,\(^5\)

\(^5\)The assumption here is that it is much easier to guess at the variance ratio instead of guessing at the individual variance components.
we recommend evaluating\(^6\)

\[
\ell(\eta|\mathbf{y}) = -(N - p) \log_e(\mathbf{y}'Q\mathbf{y}) - \log_e(|\mathbf{D}|) - \log(|\mathbf{X}'(\mathbf{I} - \mathbf{F})\mathbf{X}|) 
\]

(12)

across the range of possible values of \(\eta\), and retain the value that achieves the maximum, or

\[
\hat{\eta} = \arg \max_{\eta} \{\ell(\eta|\mathbf{y})\} 
\]

(13)

where \(\hat{\eta}\) denotes the maximum residual likelihood estimate of the variance ratio (c.f., Corbeil and Searl\(^5\)). To find \(\hat{\eta}\), one can search the parameter space using any one of several readily available constrained nonlinear optimization algorithms\(^7\) (or simply evaluate eq.(12) at discrete points along a line), with the constraint that the matrix \(\mathbf{D}\) be positive definite, or equivalently,

\[
\eta > -\frac{1}{\lambda^*}
\]

where \(\lambda^*\) denotes the maximum eigenvalue of \(\mathbf{MM}'\). The analysis is then carried out as described in the previous subsection by replacing \(\eta\) with its estimate \(\hat{\eta}\). Notice that the proposed parameterization reduces the search space to a single dimension, relative to the (two-dimension) iterative approach to variance component estimation discussed in most textbooks on linear mixed-effects models (c.f. Hocking\(^4\)). Additionally, the proposed parameterization easily permits the addition of constraints on the range of the variance ratio, to include, e.g., non-negativity constraints.

It is important to note that, since the variance ratio is estimated in this case, the reference distribution for testing factor effects is no longer \(t_{N-p}\). For balanced two-level

\(^6\)If one prefers to maximize the full log-likelihood (as opposed to the residual log-likelihood), one can evaluate \(\ell(\eta|\mathbf{y}) = -N \log_e(\mathbf{y}'Q\mathbf{y}) - \log_e(|\mathbf{D}|)\).

\(^7\)Since all that is required to evaluate eq.(12) is the capability to perform basic matrix operations (i.e., matrix addition, multiplication, inversion, and determinant), one can analyze such a design using Microsoft Excel, where the maximization of eq.(12) can be accomplished using the free Solver add-in.
factorial split-plot designs, although the distribution of the test statistic under the null hypothesis is still $t$, the degrees of freedom will depend on whether the factor effect being tested is a subplot or whole-plot effect. For unbalanced two-level factorial split-plot designs, the distribution of the test statistic is unknown, and thus, approximations can be employed using, e.g., Satterthwaite\cite{6} and Kenward and Roger\cite{7}.

For a balanced $2^k$ (or $2^k-p$) split-plot design, it is straightforward to determine the degrees of freedom available at the whole-plot and subplot levels. Let $w$ denote the number of whole-plots, $N$ denote the total number of subplots, $p_1$ denote the number of whole-plot terms (including the intercept), and $p_2$ denote the number of subplot terms, then $df_{sp} = N - w$ and $df_{wp} = w$, where $df_{sp}$ and $df_{wp}$ denote the degrees of freedom available at the subplot and whole-plot levels, respectively. The error degrees of freedom at the subplot and whole-plot levels are given by $df_{sp, error} = N - w - p_2$ and $df_{wp, error} = w - p_1$, respectively. Thus, if $\hat{\beta}_h$ is an estimated subplot effect, then under the null hypothesis $H_0$: $\beta_h = 0$, the test statistic

$$t_0 = \frac{\hat{\beta}_h}{\sqrt{\hat{\sigma}_c^2(\hat{\eta})\hat{C}_{hh}}}$$

follows the $t_{df_{sp, error}}$ distribution, and if $\hat{\beta}_h$ is an estimated whole-plot effect, then $t_0$ follows the $t_{df_{wp, error}}$ distribution. The reader is referred to Vining and Kowalski\cite{8} for more details regarding exact inference for these types of designs, as well as $2^{nd}$-order designs.

In the next section we consider the important class of two-level factorial WPS split-plot designs and discuss our strategy for justifying the cost associated with running an additional design replicate in efforts to facilitate formal testing on the whole-plot effects. Subsequently, results of a Monte Carlo simulation study used to assess the reduction in expected cost and power of the proposed strategy are discussed. We then apply
our strategy to automobile wind tunnel tests in efforts to characterize the aerodynamic performance of a NASCAR Winston Cup Chevrolet Monte Carlo stock car.

3 Proposed Analysis Strategy

In this section we discuss our proposed strategy for analyzing balanced two-level factorial WPS split-plot designs, with the aim of facilitating formal statistical testing on the subplot and whole-plot effects, while reducing the expected cost of experimentation. For a balanced two-level factorial WPS split-plot design, it is shown in the Appendix that $H$ and $F$ defined in Section 2 can be written as:

$$H = X(X'X)^{-1}X' + \eta MM'$$

and

$$F = a(\eta)MM'$$

where

$$a(\eta) = \frac{\eta}{1 + m\eta}$$

and $m$ denotes the number of subplots per whole-plot. Consequently, it is further shown in the Appendix that $Q$ reduces to

$$Q = I_{N \times N} - X(X'X)^{-1}X', $$

which is not a function of the variance ratio. Further, since the model is not saturated at the subplot level,

$$\hat{\sigma}^2_{\beta_{sp}} = \frac{y'Qy}{N(N-p)}$$

and

$$\hat{\sigma}^2_{\beta_{wp}} = \hat{\sigma}^2_{\beta_{sp}} (1 + m\eta)$$

14
where $\hat{\sigma}^2_{\beta_{sp}}$ denotes the estimated variance of the subplot effect estimates and $\hat{\sigma}^2_{\beta_{wp}}$ denotes the estimated variance of the whole-plot effect estimates.

Equation (18) suggests that for balanced two-level factorial WPS split-plot designs, the variance of the subplot effect estimates is not dependent upon the variance ratio. Consequently, tests can be performed on the subplot factors by setting $\eta = 0$ and carrying out the analysis as described in Section 2.2.1. This is equivalent to analyzing the experimental data using standard OLS. The $p$-values corresponding to the tests on the subplot factors can be interpreted in the traditional sense (i.e., as strength of evidence against the null hypothesis $H_0 : \beta_{sp} = 0$). However, the $p$-values corresponding to the whole-plot factors should not be interpreted this way. Although they lack standard interpretation, the $p$-values at the whole-plot level can be useful in determining whether or not the model can be reduced at the whole-pot level, or the design should be augmented with additional runs, either of which would facilitate formal statistical testing on the whole-plot factor effects.

From equation (19) it is clear that, for a given response vector $y$, as $\eta$ get larger, then so does the estimated variance of the $\hat{\beta}$’s at the whole-plot level. This suggests the corresponding $p$-values for the $t$-tests on the whole-plot factors will also increase with $\eta$. As a result, if an effect at the whole-plot level is not significant at $\eta = 0$, then in lack of additional response data, it will most certainly not be significant for any $\eta > 0$. Consequently, any whole-plot terms having $p$-values greater than the $\alpha$-level of significance can be viewed as candidates for removal. If there are whole-plot terms with corresponding $p$-values less than the $\alpha$-level of significance, this suggests that these terms might very well be significant; however, knowledge of $\eta$ is required to be certain of their significance at a given confidence level. To facilitate the estimation of $\eta$, the model matrix can be modified by deleting columns of $X$ that correspond to those whole-plot
model terms that are strong candidates for removal. Or, if the OLS analysis suggests that no whole-plot terms are strong candidates for removal, then the design should be augmented with an additional replicate so that the whole-plot error can be estimated. Figure 1 shows a graphical representation of the proposed strategy.

Figure 1: Proposed strategy for analyzing balanced two-level WPS split-plot designs.

Note that the strategy in Figure 1 suggests initially expending minimal resources by running only a single design replicate and performing an “intermittent” analysis using
OLS. Then, results of this analysis are used to determine whether or not additional resources should be expended on another design replicate to facilitate formal statistical testing on the whole-plot effects. In essence, the strategy suggests that, for a given whole-plot term, if the more sensitive OLS analysis fails to detect the presence of an effect with a single design replicate, then the cost of running the additional replicate needed to detect the effect (which, in fact, may not even exist) is not justified.

It is expected that the proposed strategy will result in a decrease in the expected cost of the experiments; however, this reduction in cost will depend on the variance ratio $\eta$ and the magnitude of the whole-plot effect (say, $\gamma$). To shed some light, note that when $\eta$ is large (and for any magnitude of $\gamma$), the estimate for the variance of $\hat{\gamma}$ will be significantly underestimated using OLS, resulting in a false inflation of the test statistic $t_0 = \hat{\gamma}/se(\hat{\gamma})$. This implies that when $\eta > 0$ and OLS is used to assess the significance of a whole-plot effect, the probability of detecting an effect will significantly increase, even if the effect does not exist. As a result, in the proposed strategy, the probability of augmenting the design with an additional replicate will increase with $\eta$, and thus, so will the expected cost. Also, since the test statistic $t_0$ is, on average, more inflated when $|\gamma| > 0$ than when $\gamma = 0$, the probability of detecting such an effect (using either OLS or GLS) will increase with $|\gamma|$, and thus, in the proposed strategy, the probability of augmenting the design with an additional replicate also increases with $|\gamma|$.

It is important to note that the additional replicate should be run in a separate block so as to isolate any effects that may exist between blocks. If additional whole-plots are required, then one can model the block effect as a random effect and adopt the model:

$$ y = X\beta + M_1\delta + M_2\nu + \epsilon $$  \hspace{1cm} (20)

where,
\( M_1: N \times b \) known indicator matrix for blocks,

\( M_2: N \times w \) known indicator matrix for whole-plots,

\( \delta \sim MVN(0_{b \times 1}, \sigma^2_{\delta} I_{b \times b})\): \( b \times 1 \) random vector of errors at block level,

\( \nu \sim MVN(0_{w \times 1}, \sigma^2_\nu I_{w \times w})\): \( w \times 1 \) random vector of errors at whole-plot level,

\( \epsilon \sim MVN(0_{N \times 1}, \sigma^2_\epsilon I_{N \times N})\): \( N \times 1 \) random vector of errors at subplot level

and

\( b = \) total number of blocks,

\( m_i = \) number of whole-plots in block \( i \),

\( w = \sum_{i=1}^{b} m_i \) total number of whole-plots,

\( n_{ij} = \) number of subplots in \( j^{th} \) whole-plot of block \( i \),

\( n_i = \sum_{j=1}^{m_i} n_{ij} = \) number of subplots in block \( i \),

\( N = \sum_{i=1}^{b} \sum_{j=1}^{m_i} n_{ij} = \) total number of subplots,

where \( M_1 \) and \( M_2 \) take the forms:

\[
M_1 = blkdiag(1_{n_1}, 1_{n_2}, \ldots, 1_{n_b})
\]

\[
M_2 = blkdiag(1_{n_{11}}, \ldots, 1_{n_{1m_1}}, 1_{n_{21}}, \ldots, 1_{n_{2m_2}}, \ldots, 1_{n_{b1}}, \ldots, 1_{n_{bm_b}}).
\]

For this model it is assumed \( Cov(\delta, \nu) = 0_{b \times w}, Cov(\delta, \epsilon) = 0_{b \times N}, \) and \( Cov(\nu, \epsilon) = 0_{w \times N}, \) and thus \( E(y) = X\beta \) and

\[
Var(y) = \sigma^2_\epsilon \{ I_{N \times N} + \eta_1 M_1 M_1' + \eta_2 M_2 M_2' \}\]
where \( \eta_1 = \sigma_y^2 / \sigma_\epsilon^2 \) and \( \eta_2 = \sigma_y^2 / \sigma_\epsilon^2 \) denote the variance ratios.

The material discussed in Section 2 on estimation and inference can be easily extended to account for the additional random effect due to blocks. In particular, the variance ratios are estimated from

\[
\hat{\eta} = \arg \max_{\eta_1, \eta_2} \{ \ell(\eta_1, \eta_2 | y) \}
\]

subject to the constraints

\[
\eta_1 \lambda_{\ell_1} + \eta_2 \lambda_{\ell_2} > -1
\]

for all \( \ell = 1, \ldots, N \), where \( \lambda_{\ell_1} \) is the \( \ell^{th} \) eigenvalue of \( M_1 M'_1 \), \( \lambda_{\ell_2} \) is the \( \ell^{th} \) eigenvalue of \( M_2 M'_2 \), and

\[
\ell(\eta_1, \eta_2 | y) = -(N - p) \log_e(y'Qy) - \log_e(|D|) - \log(|X'(I - F), X|)
\]

where \( \hat{\eta} = [\hat{\eta}_1, \hat{\eta}_2] \) denotes the restricted maximum likelihood estimates of \( \eta_1 \) and \( \eta_2 \). We note that only \( D \) and \( F \) are modified from their original definitions given in Section 2 and are given by

\[
D = I_{N \times N} + \eta_1 M_1 M'_1 + \eta_2 M_2 M'_2 \]

\[
F = (\eta_1 M_1 M'_1 + \eta_2 M_2 M'_2)D^{-1}
\]

The analysis is conducted in the same way as discussed in Section 2 by replacing the unknown variance ratios with their REML estimates and computing the test statistics as defined in eq.(11). If the design is fully balanced (i.e., same number of whole-plots per block and same number of subplots per whole-plot), then the distribution of the test statistic under the null hypothesis is known to follow the \( t \) distribution. However, as before, the degrees of freedom will depend on whether the factor effect being tested is a whole-plot or subplot factor.
3.1 Performance of Proposed Strategy

In this section we report results of a Monte Carlo simulation study designed to assess the expected performance of the proposed strategy, relative to the benchmark case where two replicates are initially run in a single block. As a test case, we consider the full $2^4$ design with two hard-to-change and two easy-to-change factors. We postulate the full main effects and two-factor interaction model

$$y_{ij} = \gamma_0 + \gamma_1 z_1 + \gamma_2 z_2 + \gamma_{12} z_1 z_2 + \delta_i + \beta_3 x_3 + \beta_4 x_4 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4$$

$$+ \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \epsilon_{ij}$$ (24)

where the $\gamma$'s correspond to the whole-plot effects and the $\beta$'s correspond to the subplot effects. Similarly, the $z$'s represent whole-plot terms, and the $x$'s subplot terms. For the model in equation (24), a single replication of the the $2^4$ split-plot design is then saturated at the whole-plot level since there are no degrees of freedom available to estimate the whole-plot error variance.

To assess performance of the proposed method, we arbitrarily choose to study $\gamma_{12}$ in our simulation model. Since the designs we consider are orthogonal, and our postulated model is a main effects plus interactions model, it makes no difference which whole-plot effect we choose as the basis for our study since each will result in the same performance.

For each simulation run, we randomly generated a response vector $y_1$ from a multivariate normal distribution with prescribed mean function $E(y_1) = Z\gamma + X\beta$, where the matrix $Z$ contains the columns corresponding to the whole-plot factor effects and $X$ contains the columns corresponding to the subplot factor effects. Also, $\gamma$ and $\beta$ are parameter vectors representing the whole-plot and subplot effects, respectively. The variance-covariance matrix of $y_1$ takes the form of that in equation (3). Without loss of generality, we set $\sigma^2_e = 1$ for each combination of $\eta$ and $\gamma_{12}$ considered. The proposed strategy was
then applied to the simulated $y_1$, and the simulation model recorded whether or not a signal was produced, suggesting the effect $\gamma_{12}$ was deemed significant at the 5% level. In addition, the simulation model recorded whether or not the design was augmented with an additional replicate. In these cases, a second response vector, $y_2$, was randomly generated with the same mean and variance function and $y = [y'_1, y'_2]'$ was analyzed in two blocks of four whole plots each using GLS with REML estimates of the variance components. This process was repeated 100,000 times and estimates for the probabilities of detecting and augmenting the design were computed over the 100,000 independent simulation runs. We considered values of $\gamma_{12} = 0.5, 1.0, 1.5, 2.0, 3.0$, and $4.0$ (expressed in standard deviation units of the response), as well as values of $\eta = 0.5, 1.0, 2.0, 5.0$, and $10.0$. Figure 2 shows results of the simulation study where the estimated probability of augmenting the design is plotted versus the magnitude of $\gamma_{12}$ for the various values of $\eta$.

It is clear from examining Figure 2 that for any value of the variance ratio $\eta$, and for smaller magnitudes of $\gamma_{12}$ (say, $\gamma_{12} < 1$), the proposed method can offer some significant cost savings, especially when the magnitude of $\gamma_{12}$ is null. Unfortunately, however, this reduction in cost does not come without a tradeoff. In particular, when $|\gamma_{12}| > 0$, the proposed strategy loses some power to detect small magnitude effects as evidenced by the estimated power curves shown in Figure 3. This is an expected result, however, since when using the proposed strategy, the expected number of design replicates will be less than two (which corresponds to the benchmark alternative), unless the magnitude of $\gamma_{12}$ is very large. Hence, one would expect to see a reduction in power.

Overall, results of the simulation study suggest that the proposed strategy is a viable alternative to the benchmark case of initially committing resources for two design replicates, especially when resources are limited. The proposed strategy offers the greatest
Figure 2: Estimated probabilities of augmenting the design with an additional replicate (and in a separate block) versus the magnitude of the interaction effect $\gamma_{12}$ for various values of the variance ratio $\eta$. The parameter $\gamma_{12}$ is expressed in standard deviation units of $y$. The benchmark performance corresponds to the case where two design replicates are initially run, so that the probability of augmenting the design is 1.
Figure 3: Estimated power (y-axis) versus the magnitude of the whole-plot interaction effect $\gamma_{12}$ (x-axis) for various values of the variance ratio $\eta$. In each plot, $\gamma_{12}$ is expressed in standard deviation units of $y$. The solid line represents the benchmark case where two replicates are run for each of the 100,000 independent simulation runs performed, while the dashed line represents the estimated power corresponding to the proposed method.
benefit to the experimenter when the true magnitude of the whole-plot effect is null, at which point a significant reduction in the expected number of replicates is accompanied by a small reduction in the false alarm rate.

4 Automobile Wind Tunnel Testing

The process of automobile wind tunnel testing has the primary objective of characterizing aerodynamic performance in terms of important factors. In this section we analyze the experiments discussed in Simpson et al.\cite{9}, where wind tunnel tests were conducted on a NASCAR Winston Cup Chevrolet Monte Carlo stock car to characterize changes in the coefficient of drag ($y$) due to the factors: front ride height ($x_1$), rear ride height ($x_2$), yaw angle ($x_3$), and grill tape ($x_4$). Figure 4 shows a picture of the actual car used in the experiments, where the factors front ride height, rear ride height and yaw angle are illustrated. Figure 5 illustrates the factor grill tape, where the high and low levels of the factor correspond to having, or not having, tape placed over the grill, respectively. For these experiments, changing the ride heights is costly and time consuming, and thus, $x_1$ and $x_2$ are denoted the hard-to-change (or whole-plot factors), while $x_3$ and $x_4$ are the subplot factors. Table 4 shows the factors studied in this experiment, along with their type and levels. For extensive detail about these experiments, the reader is referred to Simpson et al.\cite{9}.

Table 4: Design factors studied in this experiment along with their type and levels.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Type</th>
<th>Low level</th>
<th>High level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front Ride Height ($x_1$)</td>
<td>HTC</td>
<td>-0.5in</td>
<td>+0.5in</td>
</tr>
<tr>
<td>Rear Ride Height ($x_2$)</td>
<td>HTC</td>
<td>-1.0in</td>
<td>+1.0in</td>
</tr>
<tr>
<td>Yaw Angle ($x_3$)</td>
<td>ETC</td>
<td>-3°</td>
<td>+1°</td>
</tr>
<tr>
<td>Grill Tape ($x_4$)</td>
<td>ETC</td>
<td>0%</td>
<td>100%</td>
</tr>
</tbody>
</table>
In the following subsection, we demonstrate our analysis approach using a single replicate of the $2^4$ split-plot design. We emphasize the need to run additional whole-plots only as a last resort alternative since the cost associated with doing so is assumed to be substantial. Subsequently, we consider the need for a reduced run size and apply our approach using data collected from a $2^4_{-1}$ split-plot design augmented with subplot center runs.

4.1 Two-Level Full Factorial WPS Split-Plot Design

Figure 6 shows the $2^4$ split-plot design chosen for these experiments, with $x_1$ and $x_2$ (i.e., ride heights) being hard-to-change. The response is coefficient of drag. Since we are interested in main effects and two-factor interactions, then without additional whole-plots, running the design shown in Figure 6 will result in a WPS split-plot design. That is, the model completely consumes the available degrees of freedom at the whole-plot
Figure 5: **Top:** Setting for *grill tape* at low level (0% covered). **Bottom:** Setting for *grill tape* at high level (i.e., 100% covered).
Since the design in Figure 6 has saturated whole-plots, we can draw inference on the
subplot factors by analyzing the design using OLS (i.e., $\eta = 0$). Doing this produced
the results shown in Table 5, suggesting $\beta_3$ and $\beta_4$ (subplot effects) are significant at
the $\alpha = 0.05$ significance level. At the whole-plot level, there is ambiguity with respect
to the whole-plot effects $\beta_1$ and $\beta_2$ (since their $p$-values are less than $\alpha$); however, there
does not appear to be any evidence suggesting $\beta_{12}$ is significant (since its corresponding
$p$-value is greater than $\alpha$). As a result, one can consider removing the $x_1x_2$ term from
the model, and consequently, free up a single degree of freedom at the whole-plot level
which can be used to estimate the whole-plot error. The modified design matrix is shown
in Figure 7, while Table 6 presents the analysis results for this design using the REML
estimate for $\eta$, or $\hat{\eta} = 0.42$. We note that since $\beta_{24}$ was significant at $\alpha = 0.10$, we chose
to retain the $x_2x_4$ column in the modified design matrix; however, at the $\alpha = 0.05$ level

---

**Table 5**

<table>
<thead>
<tr>
<th>HTC</th>
<th>ETC</th>
<th>Fr Ht</th>
<th>R Ht</th>
<th>Yaw</th>
<th>Tape</th>
</tr>
</thead>
<tbody>
<tr>
<td>wp</td>
<td>x1</td>
<td>x2</td>
<td>x1x2</td>
<td>x3</td>
<td>x4</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>

Figure 6: Single replicate of the full $2^4$ split-plot design with the response $y$ being
*coefficient of drag*.
of significance, the results suggest a main effects model is adequate. Notice that only 4 whole-plots were required to complete a formal statistical analysis on all the factor effects, thus, saving valuable resources.

Table 5: Results of analyzing $2^4$ split-plot design as a completely randomized design (i.e., $\eta = 0$). Note that the $p$-values for the whole-plot effects require special interpretation, while the $p$-values for the subplot effects can be interpreted in the traditional way.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\hat{\beta}$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP Effects</td>
<td>$x_1$ 0.0086 0.0000*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_2$ 0.0086 0.0000*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_1x_2$ 0.0007 0.1336*</td>
<td></td>
</tr>
<tr>
<td>SP Effects</td>
<td>$x_3$ -0.0116 0.0000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_4$ -0.0048 0.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_1x_3$ 0.0006 0.2031</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_1x_4$ -0.0004 0.3065</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_2x_3$ -0.0002 0.6462</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_2x_4$ 0.0008 0.0881</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_3x_4$ -0.0006 0.2031</td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Results of analyzing the modified $2^4$ split-plot design using the REML estimate for $\eta$, or $\hat{\eta} = 0.42$.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\hat{\beta}$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP Effects</td>
<td>$x_1$ 0.0086 0.0510</td>
<td></td>
</tr>
<tr>
<td>$df_{uperror} = 1$</td>
<td>$x_2$ 0.0086 0.0510</td>
<td></td>
</tr>
<tr>
<td>SP Effects</td>
<td>$x_3$ -0.0116 0.0000*</td>
<td></td>
</tr>
<tr>
<td>$df_{sperror} = 9$</td>
<td>$x_4$ -0.0048 0.0000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_2x_4$ 0.0008 0.0859</td>
<td></td>
</tr>
</tbody>
</table>

4.2 Two-Level Fractional Factorial WPS Split-Plot Design

Suppose that, in addition to a minimum number of whole-plots, a reduction in the overall run size is desirable. Then one could run a $\frac{2^4}{4}$ fractional factorial split-plot design, such as that shown in Table 2, with defining relation $I = x_1x_2x_3x_4$. As noted
Table 2 shows the design matrix based on initial results given in Table 5.

Fr Ht R Ht Yaw Tape
wp x1 x2 x3 x4 x2x4 Y

1 -1 -1 1 -1 1 0.380
-1 -1 -1 -1 1 0.402
-1 -1 1 1 -1 0.369
-1 -1 -1 1 -1 0.394
-1 1 1 -1 -1 0.393
-1 1 -1 1 1 0.410
-1 1 1 -1 -1 0.402
-1 1 1 -1 1 0.419
-1 1 1 1 1 0.386
1 1 1 1 1 0.414
1 1 1 -1 -1 0.405
1 1 -1 -1 1 0.435
1 1 -1 1 1 0.428
1 -1 1 1 -1 0.383
1 -1 -1 -1 1 0.418
1 -1 -1 1 -1 0.408
1 -1 1 -1 1 0.399

Figure 7: Modified $2^4$ design matrix based on initial results given in Table 5.

earlier, this design has resolution $IV$, suggesting main effects are aliased with three-factor interactions, and two-factor interactions are aliased with other two-factor interactions. It should be clear that, if all estimable columns are of interest, then the design shown in Table 2 is fully saturated, and thus, formal statistical testing is not possible at the whole-plot or subplot levels.

To facilitate formal testing on the subplot factor effects, we ran an additional subplot center within each whole-plot (to maintain balance) as shown in Table 3 (and reproduced in Figure 9). Note that the addition of subplot center runs to this design will not change the property that the estimated variance of the subplot effect estimates is independent of the variance ratio. Additionally, in terms of cost, although adding subplot centers to this design will increase the overall run size to 12 (vs. 8 runs for 1/2 fraction), the number of whole-plots remains unchanged at 4, suggesting a relatively small increase
in cost. The center points for the factors yaw angle and grill tape were \(-1^\circ\) and 50\%, respectively. For the factor grill tape, Figure 8 illustrates how the grill tape was placed over the grill when this variable was set at its center point.

![Figure 8: Picture of grill illustrating the center point setting for grill tape.](image)

<table>
<thead>
<tr>
<th>HTC</th>
<th>ETC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fr Ht</td>
<td>R Ht</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 9: \(2^{4-1}\) fractional factorial split-plot design with \(I = x_1x_2x_3x_4\) and augmented with subplot centers to permit subplot error estimation. The response is coefficient of drag.

Assuming we are interested in all estimable columns, then there are zero degrees of freedom available to estimate whole-plot error and 4 degrees of freedom available
to estimate the subplot error. We should note that, since center points were run in the subplot factors only, the expressions for the estimated variances of the subplot and whole-plot effect estimates are slightly different from those given earlier. In particular, let \( N_c \) denote the number of subplot center runs, then

\[
\hat{\sigma}^2_{\hat{\beta}_{sp}} = \frac{y'Qy}{(N - N_c)(N - p)} \tag{25}
\]

and

\[
\hat{\sigma}^2_{\hat{\beta}_{wp}} = \frac{y'Qy}{N(N - p)} (1 + m\eta) \tag{26}
\]

where, as before, the estimated variance of the subplot effect estimates is not dependent upon the variance ratio. As a result, inference can be drawn on the subplot factors by analyzing the design using standard OLS. Further, since, for a given \( y \), the estimated variance of the whole-plot effect estimates increases with \( \eta \), then, as before, the \( p \)-values at the whole-plot level can be useful in determining whether or not the model should be reduced at the whole-plot level, or the design should be augmented with an additional replicate.

Analyzing the design in Figure 9 using OLS produced the results shown in Table 7. Notice that, at the subplot level, the factors *yaw angle* \((x_3)\) and *grill tape* \((x_4)\) are highly significant terms, while the interaction terms appear to be less important. At the whole-plot level, it appears that a good candidate for removal from the model is the interaction term \( x_1x_2 \) since the \( p \)-value computed at \( \eta = 0 \) is much larger than \( \alpha = 0.05 \). Removing this term would make available a single degree of freedom to estimate whole-plot error. Thus, based on results of the intermittent analysis, we retain only the main effects to include in a subsequent analysis, where \( \eta \) is estimated using REML. The results are shown in Table 8.
Table 7: Results of analyzing the $2_4^{k-1}$ design in Figure 9 as a completely randomized design (i.e., $\eta = 0$). The $p$-values for the whole-plot effects require special interpretation, while the $p$-values for the subplot effects can be interpreted in the traditional way.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\hat{\beta}$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP Effects</td>
<td>$x_1$</td>
<td>0.0093</td>
</tr>
<tr>
<td></td>
<td>$x_2$</td>
<td>0.0087</td>
</tr>
<tr>
<td></td>
<td>$x_1x_2$</td>
<td>0.0003</td>
</tr>
<tr>
<td>SP Effects</td>
<td>$x_3$</td>
<td>-0.0111</td>
</tr>
<tr>
<td></td>
<td>$x_4$</td>
<td>-0.0046</td>
</tr>
<tr>
<td></td>
<td>$x_1x_3$</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>$x_2x_3$</td>
<td>-0.0006</td>
</tr>
</tbody>
</table>

Table 8: Results of analyzing the modified $2_4^{k-1}$ design with REML estimate for $\eta$ ($\hat{\eta} = -0.2306$). For this analysis, the $p$-values for all factor effects can be interpreted in the traditional way.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\hat{\beta}$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP Effects</td>
<td>$x_1$</td>
<td>0.0093</td>
</tr>
<tr>
<td></td>
<td>$x_2$</td>
<td>0.0087</td>
</tr>
<tr>
<td>SP Effects</td>
<td>$x_3$</td>
<td>-0.0111</td>
</tr>
<tr>
<td></td>
<td>$x_4$</td>
<td>-0.0046</td>
</tr>
<tr>
<td></td>
<td>$x_1x_3$</td>
<td>0.0014</td>
</tr>
</tbody>
</table>

The results in Table 8 suggest that the main effect terms are by far the most important in characterizing changes in the expected drag coefficient. Suppose that $\alpha = 0.10$, then there would appear to be some marginal significance associated with either the interaction between front ride height ($x_1$) and yaw angle ($x_3$), or the interaction between rear ride height ($x_2$) and grill tape ($x_4$). Without additional experiments, it is not possible to de-alias these effects. One way to break the alias chain is to run the alternate fraction $I = -x_1x_2x_3x_4$ in a separate block. Of course, doing so would require 4 additional whole-plots, doubling the overall experimental cost (assuming balance is maintained with subplot centers). The experiments would then require 8 total whole-
plots and 24 total subplot runs (i.e., 3 subplots per whole-plot). In Table 8, since the $x_1x_3$ term is only marginally significant at $\alpha = 0.10$, this may not be sufficient to justify the added cost associated with running the alternate fraction in this case.

For the sake of illustration, suppose it was deemed important to de-alias the interaction terms $x_1x_3$ and $x_2x_4$, thus, justifying the cost associated with sequentially running the alternate fraction ($I = -x_1x_2x_3x_4$). To account for the random block effect, we assume the linear model specified in eq. (20), where there are two variance ratios, $\eta_1$ and $\eta_2$. The full design is shown in Figure 10 along with the observed responses. Note that for a main effect and two-factor interaction model, there are 3 whole-plot error degrees of freedom and 9 subplot error degrees of freedom, thus, facilitating formal statistical testing on the whole-plot and subplot factor effects. Table 9 shows the results of the analysis, where REML was used to estimate the unknown variance ratios $\eta_1$ and $\eta_2$. By de-aliasing $x_1x_3$ and $x_2x_4$, one can see that the more active of the two interaction effects is that associated with the $x_2x_4$ term. However, the results in Table 9 suggest that, at the $\alpha = 0.05$ significance level, the expected drag coefficient is primarily affected by main effects.

Notice at the $\alpha = 0.05$ significance level, all three cases explored in this section lead to the same main-effects model. More importantly, for both the full and fractional factorial design cases presented, using the proposed analysis approach allowed us to complete a formal statistical analysis on all the factor effects with only 4 whole-plots. This is significant as additional whole-plots are assumed to substantially increase the overall cost of experimentation.
Figure 10: 24 run split-plot design with 8 whole-plots ran in two blocks of 4 each. Response is coefficient of drag.

Table 9: Results of analyzing the design in Figure 10 with REML estimates for $\eta_1$ and $\eta_2$ ($\hat{\eta}_1 = -0.0471$ and $\hat{\eta}_2 = -0.1417$). The $p$-values for all factor effects are interpreted in the traditional way.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\hat{\beta}$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP Effects</td>
<td>$x_1$</td>
<td>0.0090</td>
</tr>
<tr>
<td>$df_{wperror} = 3$</td>
<td>$x_2$</td>
<td>0.0090</td>
</tr>
<tr>
<td></td>
<td>$x_1x_2$</td>
<td>0.0007</td>
</tr>
<tr>
<td>SP Effects</td>
<td>$x_3$</td>
<td>-0.0116</td>
</tr>
<tr>
<td>$df_{sperror} = 9$</td>
<td>$x_4$</td>
<td>-0.0048</td>
</tr>
<tr>
<td></td>
<td>$x_1x_3$</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>$x_1x_4$</td>
<td>-0.0004</td>
</tr>
<tr>
<td></td>
<td>$x_2x_3$</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>$x_2x_4$</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>$x_3x_4$</td>
<td>-0.0006</td>
</tr>
</tbody>
</table>
5 Concluding Remarks

Restrictions in randomization are becoming more prevalent in experimental design practice due to having experimental design factors that are difficult or costly to change during the course of the experiment. Many times, the randomization restriction leads to a split-plot execution of the experiments, where the hard-to-change and easy-to-change factors correspond to whole-plot and subplot factors, respectively. In a split-plot design, it is assumed that a large fraction of experimental cost is proportional to the number of whole-plots, and thus, it is desirable to minimize the number of whole-plots, if possible. The two-level (full or fractional) factorial design is a very practical choice since it will minimize the number of times the whole-plot factor level combination is changed. When experimenters are only afforded a single replicate of the design, then, depending on the postulated model, the design may be fully saturated, or saturated only at the whole-plot level. If the design is fully saturated, then formal statistical testing is not possible on any of the factor effects, and thus, the experimenter should consider the approaches to analysis discussed in Loeppky and Sitter [3] if additional experimental runs are prohibited. Often, a fully saturated two-level factorial split-plot design can be converted to whole-plot saturated by simply augmenting the design with subplot centers as was done in the previous section. If the design is saturated only at the whole-plot level, formal statistical testing on the subplot factor effects is possible; however, formal tests are not possible on the whole-plot factor effects since there are no available degrees of freedom to estimate the whole-plot error. In this case, one may use the methods discussed in Loeppky and Sitter [3] to assess the significance of the whole-plot effects, and analyze the subplot effects via standard tests. If a formal statistical test on the whole-plot effects is desired, this will require degrees of freedom for estimating the whole-plot error. Unless
the model can be reduced at the whole-plot level, additional whole-plots will need to be run to facilitate estimation, and thus, testing.

In this paper we considered balanced two-level factorial *whole-plot saturated* split-plot designs, and proposed an analysis strategy that aims to justify the cost for running an additional design replicate in efforts to facilitate formal testing at the whole-plot level. The proposed strategy exploits the sensitivity of the OLS procedure in detecting whole-plot effects by suggesting to run a second design replicate only if an effect is detected in the intermittent analysis. It was shown via Monte Carlo simulation that the proposed strategy yields a reduction in the expected cost of the experiments, especially when the whole-plot effect under scrutiny is truly null. When this effect is not null, then the proposed strategy will lose some power to detect small magnitude effects, relative to the case where two design replicates are initially run. However, this is expected and might be a reasonable tradeoff for the experimenter, particularly during screening experiments where resources are limited and the purpose of running the experiments is merely to characterize response variation.

As a final note, our approach to analysis is in line with the general concept of sequential experimentation discussed by Box\textsuperscript{[10]}, where it was noted that, when possible, it is best to avoid “all encompassing” experiments, which must be planned when the least is known about the system or process under study. Instead, it is best to run smaller sets of experimental runs in sequence. This way, in addition to potential cost savings, one can learn about the process via the current set of experiments, which, hopefully, can aid the experimenter in planning the next set of experiments.
Acknowledgements

The authors acknowledge support from the Summer Excellence in Research Program (SEiRP) at the Culverhouse College of Commerce and Business Administration at The University of Alabama.

References


Appendix

In what follows, it is assumed that the design matrix $X$ consists of a two-level (full or fractional) factorial split-plot design that is saturated at the whole-plot level.

Form of $F$ matrix for two-level factorial WPS split-plot design

Under the assumed form for $X$, the matrix $F$ can be written as

$$F = a(\eta)MM'$$

where

$$a(\eta) = \frac{\eta}{1 + m\eta}$$

and $m$ denotes the number of subplots per whole-plot. To prove this, recall that the $N \times N$ matrix $F$ is defined generally as

$$F = \eta MM'D^{-1},$$

where the $N \times N$ matrix $\eta MM'$ takes the form:

$$\eta MM' = diag(A_1, A_2, ..., A_w)$$
where $A_s$ is a $m \times m$ matrix with the same elements $\eta$, for $s = 1, ..., w$. Then it is easy to verify that the $N \times N$ matrices $D = I_{N \times N} + \eta MM'$ and $D^{-1}$ takes the form:

$$D = \text{diag}(D_1, D_2, ..., D_w)$$

where $D_s$ is a $m \times m$ matrix with elements $(D)_{ii} = 1 + \eta$ and $(D)_{ij} = \eta$, for $s = 1, ..., w$, $i = 1, ..., m$, $j = 1, ..., m$ ($i \neq j$), and thus,

$$D^{-1} = \text{diag}(D_1^*, D_2^*, ..., D_w^*)$$

where $D_s^*$ is a $m \times m$ matrix with elements $(D_s^*)_{ii} = \frac{1+\eta}{1+m\eta}$ and $(D_s^*)_{ij} = -\frac{\eta}{1+m\eta}$, for $s = 1, ..., w$, $i = 1, ..., m$, and $j = 1, ..., m$ ($i \neq j$). It follows that the $N \times N$ matrix $F$ then takes the form

$$F = \frac{\eta}{1+m\eta}MM'$$

**Form of H matrix for two-level factorial WPS split-plot design**

Under the assumed form for $X$, the matrix $H$ can be written as:

$$H = X(X'X)^{-1}X' + \eta MM'.$$

To prove this, recall that the $N \times N$ matrix $H$ is defined generally as

$$H = X(X'(I_{N \times N} - F)X)^{-1}X'.$$

Next, define the $p \times p$ matrix $B = (X'(I_{N \times N} - F)X)^{-1}$, then

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

where

$B_{11}$ is a $w \times w$ diagonal matrix with elements equal to $\frac{1+m\eta}{N}$.
\( \mathbf{B}_{12} \) and \( \mathbf{B}_{21} \) are \((p - w) \times w\) and \(w \times (p - w)\) null matrices, respectively, \( \mathbf{B}_{22} \) is a \((p - w) \times (p - w)\) diagonal matrix with elements equal to \(\frac{1}{N}\).

It then follows that the \(N \times N\) matrix \(\mathbf{H}\) takes the form

\[
\mathbf{H} = \mathbf{X}\mathbf{B}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' + \eta \mathbf{M}\mathbf{M}'.
\]

Form of Q matrix for two-level factorial WPS split-plot design

Under the assumed form for \(\mathbf{X}\), the matrix \(\mathbf{Q}\) can be written as:

\[
\mathbf{Q} = \mathbf{I}_{N \times N} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.
\]

To prove this, recall that the \(N \times N\) matrix \(\mathbf{Q}\) is defined generally as

\[
\mathbf{Q} = [(\mathbf{I}_{N \times N} - \mathbf{H}(\mathbf{I}_{N \times N} - \mathbf{F}))'(\mathbf{I}_{N \times N} - \mathbf{F})(\mathbf{I}_{N \times N} - \mathbf{H}(\mathbf{I}_{N \times N} - \mathbf{F}))].
\]

Then, given the reduced forms of \(\mathbf{F}\) and \(\mathbf{H}\), it is straightforward to show that

\[
\mathbf{H}(\mathbf{I}_{N \times N} - \mathbf{F}) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'
\]

and after simplifying \(\mathbf{Q}\) algebraically we obtain

\[
\mathbf{Q} = \mathbf{I}_{N \times N} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.
\]

which is independent of the variance ratio \(\eta\).