Classes of Second-Order Split-Plot Designs

DATAWorks 2018
Springfield, VA

Luis A. Cortés, Ph.D.
The MITRE Corporation

James R. Simpson, Ph.D.
JK Analytics, Inc.

Peter Parker, Ph.D.
NASA

22 March 2018
Outline

➢ Objectives
➢ The Split-Plot Design
➢ Second-Order Split-Plot Designs
➢ Cartesian Product Split-Plot Design
➢ Design Evaluation Criteria
➢ Cartesian Product Design Performance
➢ Summary and Conclusions
Objectives

➢ Explore the state-of-the-art published second-order split-plot designs
  ✓ OLS-GLS Equivalent Estimation
  ✓ Minimum Whole Plot
  ✓ Optimal Designs

➢ Explore a systematic approach for constructing new second-order split-plot designs, with and without blocking

➢ Introduce the Second-Order Cartesian Product Split-Plot Design

➢ Compare the performance of standard second-order split-plot designs and the new Second-Order Cartesian Product Split-Plot Design

We illustrate practical strategies to help practitioners overcome some of the challenges presented by second-order split-plot design
The Split-Plot Design
Background

➢ An **experiment** is a test in which systematic changes are made to the input variables of a process or system to measure the corresponding changes in the output variables.

➢ **Experiment design** is the systematic integration of scientific strategies for gathering empirical knowledge about a process or system using a statistical approach to problem solving for planning, designing, executing, and analyzing an experiment.

➢ Fisher laid down the fundamental principles of experiment design: factorization, replication, **randomization**, and local control of error.

➢ Box and Wilson catalyzed the application of experiment design to industrial applications.

➢ **Response surface methodology** is a sequential strategy of experimentation that incorporates statistical methods to fit a low-order Taylor series approximation to the true underlying mechanism.

➢ Unfortunately, it is not always possible to carry out an experiment while adhering simultaneously to all of the principles of experiment design.

The regression model (second-order example)

\[ y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \beta_{jj} x_j^2 + \sum_{i<j}^{k} \beta_{ij} x_i x_j + \varepsilon \]
The Split-Plot Design
Split-Plot Structure

<table>
<thead>
<tr>
<th>Block Randomization</th>
<th>Whole-plot Randomization</th>
<th>Sub-plot Randomization</th>
<th>Observational Unit Randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>( z_1 )</td>
<td>( x_2 )</td>
<td>( z_2 )</td>
</tr>
<tr>
<td>( x_1 )</td>
<td>( z_1 )</td>
<td>( x_2 )</td>
<td>( z_2 )</td>
</tr>
<tr>
<td>( \gamma_{1111} )</td>
<td>( \gamma_{1111} )</td>
<td>( \gamma_{1121} )</td>
<td>( \gamma_{1121} )</td>
</tr>
<tr>
<td>( \gamma_{1121} )</td>
<td>( \gamma_{1121} )</td>
<td>( \gamma_{1121} )</td>
<td>( \gamma_{1121} )</td>
</tr>
<tr>
<td>( \gamma_{1211} )</td>
<td>( \gamma_{1211} )</td>
<td>( \gamma_{1211} )</td>
<td>( \gamma_{1211} )</td>
</tr>
<tr>
<td>( \gamma_{1221} )</td>
<td>( \gamma_{1221} )</td>
<td>( \gamma_{1221} )</td>
<td>( \gamma_{1221} )</td>
</tr>
<tr>
<td>( \gamma_{2111} )</td>
<td>( \gamma_{2111} )</td>
<td>( \gamma_{2121} )</td>
<td>( \gamma_{2121} )</td>
</tr>
<tr>
<td>( \gamma_{2121} )</td>
<td>( \gamma_{2121} )</td>
<td>( \gamma_{2121} )</td>
<td>( \gamma_{2121} )</td>
</tr>
<tr>
<td>( \gamma_{2211} )</td>
<td>( \gamma_{2211} )</td>
<td>( \gamma_{2211} )</td>
<td>( \gamma_{2211} )</td>
</tr>
<tr>
<td>( \gamma_{2221} )</td>
<td>( \gamma_{2221} )</td>
<td>( \gamma_{2221} )</td>
<td>( \gamma_{2221} )</td>
</tr>
</tbody>
</table>

2\(^2\) x 2\(^2\) Randomized Complete Block Split-Plot Design

- Encountered in situations when the levels of one or more experimental factors are impractical or inconvenient to change
- Leads to restriction in randomization of the experimental run order
- Replicates, or blocks, are divided into parts called whole-plots, which are further divided into parts called sub-plots
- Defining feature - two sources of error
  - \( \sigma^2 \) - whole plot variance
  - \( \sigma^2 \) - sub-plot variance
- Typically \( \sigma^2 \) / \( \sigma^2 \) \( \geq \) 1.0
- Effects of factors that have more experimental units (more information) can be estimated with better precision
- Design, execution, and analysis strategies need to be modified
- Faster but more costly than a completely randomized design – typically the whole-plot factors are costly-to-change
- Limitations in the resources available to complete the experiment in heterogeneous settings leads to a blocked split-plot structure or a split-split-plot design

Split-plot structures are commonplace in industrial experiments
Second-Order Split-Plot Designs
State-of-the-Art Designs (without Blocking)

  - Provided OLS-GLS equivalent estimation central composite and Box-Behnken designs
  - Recommended options for obtaining balanced split-plot designs
  - Derived the necessary conditions to obtain OLS-GLS equivalent estimation designs

  - Provided a method for redistributing the overall center point runs in Vining, Kowalski, and Montgomery (2005)
  - Proposed methods for constructing balanced OLS-GLS equivalent estimation central composite and Box-Behnken minimum whole-plot designs

  - Provided balanced and unbalanced minimum whole-plot designs from Box and Draper, Notz, and Hoke hybrid minimal point designs

  - Provided a D-optimal design algorithm to redistribute the overall center point in Vining, Kowalski, and Montgomery (2005)
  - Provide an algorithm for estimating the pairwise correlation coefficients between second-order split-plot model terms
Second-Order Split-Plot Designs
State-of-the-Art Designs (with Blocking)

  ✓ Provided second-order orthogonally blocked OLS-GLS equivalent estimation split-plot designs from Vining, Kowalski, and Montgomery (2005) and second-order orthogonally blocked minimum whole-plot split-plot designs from Parker, Kowalski, and Vining (2007b)
  ✓ Expanded Box and Hunter (1959) second-order orthogonal blocking general conditions to split-plot designs

  ✓ Provided candidate split-plot designs and constructed balanced second-order blocked split-plot designs using designs provide by Dey (2009)
  ✓ Provided candidate split-plot designs and constructed unbalanced second-order blocked split-plot designs using designs provide by Zhang et. al. (2011)
**Cartesian Product Split-Plot Design**

**Design Construction Example ($\Gamma = \Theta = 1$)**

Whole-plot array ($p = 2$) in a central composite design arrangement

<table>
<thead>
<tr>
<th>$W_{1,2}$</th>
<th>$S_{1,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_1$</td>
<td>$x_1$</td>
</tr>
<tr>
<td>$z_2$</td>
<td>$x_2$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$-\beta$</td>
<td>$-\alpha$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>$0$</td>
<td>$-\beta$</td>
</tr>
<tr>
<td>$0$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>$0$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

Sub-plot array ($q = 2$) in a central composite design arrangement

<table>
<thead>
<tr>
<th>$W_{1,2}$</th>
<th>$S_{1,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$0$</td>
</tr>
<tr>
<td>$1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$1$</td>
<td>$0$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$-\alpha$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>$0$</td>
<td>$-\beta$</td>
</tr>
<tr>
<td>$0$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>$0$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

Cartesian product:

$W_{\Gamma,p} \times S_{\Theta,q} = \{(z_1, z_2, \ldots, z_p), (x_1, x_2, \ldots, x_q)\} \ (z_1, z_2, \ldots, z_p) \in W_{\Gamma,p} \ (x_1, x_2, \ldots, x_q) \in S_{sa,q}$

Replicate one whole-plot for whole-plot variance estimation
Cartesian Product Split-Plot Design
Design Construction Example (Γ = Θ = 1)
Cartesian Product Split-Plot Design
Design Construction Example ($\Gamma = \Theta = 2$)

<table>
<thead>
<tr>
<th>$W_1$</th>
<th>$W_2$</th>
<th>$S_1$</th>
<th>$S_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_1$</td>
<td>$z_2$</td>
<td>$x_1$</td>
<td>$x_2$</td>
</tr>
<tr>
<td>$z_1$</td>
<td>$z_2$</td>
<td>$x_1$</td>
<td>$x_2$</td>
</tr>
<tr>
<td>$z_1$</td>
<td>$z_2$</td>
<td>$x_1$</td>
<td>$x_2$</td>
</tr>
<tr>
<td>$z_1$</td>
<td>$z_2$</td>
<td>$x_1$</td>
<td>$x_2$</td>
</tr>
</tbody>
</table>

$$
W_1 \begin{array}{cc}
-1 & -1 \\
1 & -1 \\
-1 & 1 \\
1 & 1 \\
0 & 0 \\
\end{array}
W_2 \begin{array}{cc}
-\beta & 0 \\
\beta & 0 \\
0 & -\beta \\
0 & \beta \\
0 & 0 \\
\end{array}
S_1 \begin{array}{cc}
-1 & -1 \\
1 & -1 \\
-1 & 1 \\
1 & 1 \\
0 & 0 \\
\end{array}
S_2 \begin{array}{cc}
-\alpha & 0 \\
\alpha & 0 \\
0 & -\alpha \\
0 & \alpha \\
0 & 0 \\
\end{array}
$$

$$
\begin{array}{cc}
\text{W1 } x S_1 \\
\text{W2 } x S_2 \\
\text{D(2C2, 2C2)-1} \\
\text{W1 } x S_2 \\
\text{W2 } x S_1 \\
\text{D(2C2, 2C2)-2} \\
\end{array}
$$
## Design Evaluation Criteria
### Split-Plot Design Evaluation Criteria

**Myers, Montgomery, and Anderson-Cook (2009)**

<table>
<thead>
<tr>
<th></th>
<th>Split-Plot Design Evaluation Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Provide a good fit of the model to the data</td>
</tr>
<tr>
<td>2.</td>
<td>Allow a precise estimation of the model coefficients</td>
</tr>
<tr>
<td>3.</td>
<td>Provide a good prediction over the experimental region</td>
</tr>
<tr>
<td>4.</td>
<td>Provide an estimation of both whole-plot variance and sub-plot variance</td>
</tr>
<tr>
<td>5.</td>
<td>Detect lack-of-fit</td>
</tr>
<tr>
<td>6.</td>
<td>Check the homogeneous variance assumption at the whole-plot and sub-plot levels</td>
</tr>
<tr>
<td>7.</td>
<td>Consider the cost in setting the whole-plot and sub-plot factors</td>
</tr>
<tr>
<td>8.</td>
<td>Ensure the simplicity of the design</td>
</tr>
<tr>
<td>9.</td>
<td>Ensure simple calculations</td>
</tr>
<tr>
<td>10.</td>
<td>Be robust to errors in control of factor levels</td>
</tr>
<tr>
<td>11.</td>
<td>Be robust to outliers</td>
</tr>
</tbody>
</table>

1. **Provide a good fit of the model to the data**
   - A small number of runs
   - Sequential assembly to higher order
   - Blocking

2. **Allow a precise estimation of the model coefficients**
   - Pairwise correlation coefficients between model terms

3. **Provide a good prediction over the experimental region**
   - Stability of the prediction variance over the entire design space
   - Prediction variance at the center of the design space
   - Maximum prediction variance
   - Average prediction variance
   - Prediction variance at the 50th percentile of the design space
   - Prediction variance at the 90th percentile of the design space

4. **Provide an estimation of both whole-plot variance and sub-plot variance**
   - Whole-plot replication for estimating the whole-plot variance component
   - Pooling of the sub-plot center runs to obtain an estimation of the sub-plot variance
Cartesian Product Split-Plot Design Performance

\[(\Gamma = \Theta = 1; p = q = 2; \alpha = \beta = 1.414)\]

Pairwise correlation between model terms

Prediction variance profile

Prediction variance vs. Fraction of Design Space (FDS)
- Prediction variance @ 50% FDS - 0.74
- Prediction variance @ 90% FDS - 1.15
- Maximum prediction variance - 3.52
- Average prediction variance - 0.86
- Prediction variance @ overall center - 1.21

Color map on correlations
- No. cells with correlation coefficient \(|r| = 0.0 \to 89/91\)
- No. cells with correlation coefficient \(0.0 < |r| < 0.5 \to 0/91\)
- No. cells with correlation coefficient \(0.5 < |r| < 1.0 \to 2/91\)
Cartesian Product Split-Plot Design Performance
Central Composite and D-Optimal Designs

Vining, Kowalski, and Montgomery (2005) OLS-GLS equivalent

Jones and Nachtsheim (2009) D-optimal design

Second-order sub-array Cartesian product D(2C2, 2C2)-2

$w_j = 12$
$N = 48$

$w_j = 12$
$N = 48$

$w_j = 12$
$N = 60$
Cartesian Product Split-Plot Design Performance Designs for Spherical Regions

- Vining, Kowalski, and Montgomery (2005) OLS-GLS equivalent
- Second-order sub-array Cartesian product D(2C2, 2C2)-1
- Parker, Kowalski, and Vining (2007a) minimum whole plot
- Second-order sub-array Cartesian product D(3C2, 3C2)-5

\[ w_i = 10 \]
\[ N = 40 \]

\[ w_i = 10 \]
\[ N = 50 \]

\[ w_i = 9 \]
\[ N = 45 \]

\[ w_i = 9 \]
\[ N = 36 \]
Cartesian Product Split-Plot Design Performance
Designs for Cuboidal Regions

Parker, Kowalski, and Vining (2007a) minimum whole-plot

Verma et. al. (2012) unbalanced

Second-order sub-array Cartesian product D(2B2, 2B2)-2 balanced
Cartesian Product Split-Plot Design Performance
Blocked Designs


Second-order sub-array Cartesian product D(2C2, 2C2)-1

Wang, Kowalski, and Vining (2009) - Parker, Kowalski, and Vining (2007b) MWP

Verma et al. (2012) balanced

Avg. PV = 0.52

Avg. PV = 0.65

Avg. PV = 0.71

 Avg. PV = 0.68
Summary and Conclusions

Second-Order Cartesian Product Split-Plot Designs

➢ Have desirable features
  – Independent whole-plot and sub-plot sub-arrays form intuitive building blocks
  – Flexible design building structure to accommodate your needs
  – Multiple types of designs can be built and easily compared with combinations of sub-arrays
  – The sub-array structure creates a natural blocking structure

➢ Are an efficient and effective alternative to split-plot design methods such as the Taguchi’s product array design methods
  – Efficient—require about one-half of the number of runs required by Taguchi product arrays
  – Effective—permit modeling second-order effects whereas product arrays model only first-order effects

➢ Are high information-quality
  – The variance of the regression coefficients is low
  – The prediction variance of the regression coefficients is low and stable
  – The aliasing between the terms in the model is low

➢ Perform as well as state-of-the-art second-order split-plot designs
  – Comparable prediction variance
  – Comparable aliasing structure
  – Comparable blocking performance
Bibliography


