

# A Modified Path of Steepest Ascent for Split-Plot Experiments

SCOTT M. KOWALSKI

*Minitab Inc., State College, PA 16801*

CONNIE M. BORROR

*University of Illinois, Urbana, IL 61801*

DOUGLAS C. MONTGOMERY

*Arizona State University, Tempe, AZ 85287-5906*

Response surface methodology is applied often in industrial experiments to find the optimal conditions for the design factors. Response surface methodology involves three main steps: screening potential factors, seeking a region around the global optimum using the path of steepest ascent, and estimating the true model for the response. Many of these experiments involve randomization restrictions and can be thought of as split-plot experiments. This paper investigates the path of steepest ascent within a split-plot structure. Three methods are proposed for calculating the coordinates along the path.

KEY WORDS: Lagrange Multiplier; Response Surface Methodology; Split-Plot.

## Introduction

INDUSTRIAL experiments often involve two types of factors: those that are hard to change and those that are easy to change. The experiments are performed by fixing the levels of the hard-to-change factors and then running all or some of the combinations of the easy-to-change factor levels. Then, a new setting in the hard-to-change factors is selected and the process is repeated. With replication, this type of randomization leads to two error terms and is known as a split-plot experiment. The hard-to-change factors are called the whole-plot factors and the easy-to-change factors are called the subplot factors.

A split-plot design is typically more efficient than

a completely randomized design (CRD) because  $\sigma_{\text{sub}}^2 < \sigma_{\text{CRD}}^2 < \sigma_{\text{whole}}^2$ , where  $\sigma_{\text{sub}}^2$  is the subplot error variance and  $\sigma_{\text{whole}}^2$  is the whole-plot error variance (see Box and Jones (1992)). Often times, split-plot designs are also more cost effective than CRDs. By nature, changing the hard-to-change factors represents a larger cost than changing the easy-to-change factors. In a split-plot experiment, these factor levels are changed less often than in a CRD, which results in reduced cost.

The goal of many industrial experiments is to find the levels of the design factors that optimize a response. This area of industrial statistics is known as response surface methodology (RSM), as described by Box and Wilson (1951). Most applications in RSM are sequential in nature. There are generally three steps in RSM, which are

1. A screening experiment—the objective is to reduce the many candidate factors to a relatively few potentially important factors that influence the response. This allows experiments in the second and third steps to be more efficient and use fewer runs.
2. Using information from step 1 to move through

---

Dr. Kowalski is a Technical Trainer. He is a Member of ASQ. His email address is skowalski@minitab.com.

Dr. Borrer is an Assistant Professor in the General Engineering Department. She is a Member of ASQ. Her email address is cborror@express.cites.uiuc.edu.

Dr. Montgomery is a Professor in the Department of Industrial Engineering. He is a Fellow of ASQ. His email address is douglas.montgomery@asu.edu.

the experimental region in an attempt to get closer to the optimum. The most common approach is called steepest ascent.

3. Optimization or process mapping—after a small region around the optimum has been identified using step 2, then a model that approximates the true response function is estimated. This model is typically a second-order model to approximate curvature near the optimum and can be used to determine optimal conditions on the design factors.

The sequential nature of RSM involves using information from previous steps at each subsequent step, allowing the investigator to learn more about the particular process being studied. RSM is frequently applied in an iterative manner. For more details on RSM, see Myers and Montgomery (1995) and Khuri and Cornell (1996).

A split-plot structure can complicate all three steps of a RSM problem. Many researchers have worked on this topic in recent years. Lucas and Ju (1992) used a simulation study to investigate the use of split-plot designs in industrial experiments. Their results confirm that split-plot designs produce increased precision on the subplot factors while sacrificing precision on the whole-plot factors. Box (1996) explains that completely randomized experiments are often impractical in industry and indicates that split-plot experiments are often very efficient and easier to run. To illustrate the difference in the analysis, Lucas and Hazel (1997) run the well-known paper-helicopter experiment as a CRD and as a split-plot.

Letsinger, Myers, and Lentner (1996) introduce birandomization designs (BRD), which refer to designs with two randomizations and thus are similar to split-plot designs. Vining, Kowalski, and Montgomery (2004) propose central composite designs specifically designed for split-plot experiments. They also provide a general proof that shows for certain design conditions the ordinary least squares coefficient estimates are equivalent to the generalized least squares coefficient estimates for second-order models. Kowalski, Vining, Montgomery, and Borrer (2004) modify the proposed central composite designs of Vining, Kowalski, and Montgomery (2004) to model both the process mean and variances within a split-plot structure.

Bisgaard and Steinberg (1997) look at the design and analysis of prototype experiments. They present

examples that use split-plot designs and show clearly how to carry out a two-stage analysis of the data. Bisgaard (2000) uses two-level fractional factorials to construct split-plot designs and gives general expressions for deriving alias structures based on the group structure of the arrays.

Huang, Chen, and Voelkel (1998) and Bingham and Sitter (1999a) discuss minimum-aberration (MA) designs for split-plot experiments, where both the whole-plot and subplot factors have two levels. They provide methods for determining the MA designs and then give tables for various combinations of whole-plot and subplot factors. Some design issues with two-level fractional-factorial split-plot experiments, including where to split and where to fractionate, are presented in Bingham and Sitter (2001). Theoretical justification of these types of split-plot designs is given in Bingham and Sitter (1999b).

Schoen (1999) proposes a method for manipulating the division of contrasts over the different error strata of two-level experiments with nested errors. Goos and Vandebroek (2001) propose an exchange algorithm for obtaining  $D$ -optimal split-plot designs. They show that the design matrices for the  $D$ -optimal split-plot designs and  $D$ -optimal CRDs are typically different and that split-plot experiments are often more efficient than CRDs.

Kowalski (2002) considers split-plot experiments in robust parameter design. He constructs 24-run designs in two ways: using the properties of a balanced incomplete block design and by semifolding a 16-run design. One section in Myers, Montgomery, Vining, Borrer, and Kowalski (2004) is devoted to a review of the current literature on industrial split-plot experiments.

The previous papers present methods for screening experiments and for optimization. As far as the authors know, no one has looked at step 2 of a RSM study that is carried out as a split-plot experiment. The goal of this paper is to address the method of steepest ascent when the experiment is conducted as a split plot.

## Method of Steepest Ascent

The method of steepest ascent involves moving through the experimental region along a path that yields increases in the response. After a first-order model has been fit, the regression coefficients from this model are used to determine the coordinates along the path. The movement in  $x_j$  along the path

of steepest ascent is proportional to the magnitude of the regression coefficient  $\hat{\beta}_j$ , with the direction based on the sign of the coefficient. Consider the fitted first-order model

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k.$$

The path of steepest ascent moves a distance  $r$  away from the design center in the direction of the maximum response with a spherical constraint  $\sum_{j=1}^k x_j^2 = r^2$ .

The maximization procedure of the response function uses Lagrange multipliers. Let

$$Q(x_1, x_2, \dots, x_k) = \hat{\beta}_0 + \sum_{j=1}^k \hat{\beta}_j x_j - \lambda \left( \sum_{j=1}^k x_j^2 - r^2 \right),$$

where  $\lambda$  is the Lagrange multiplier. The maximization requires taking the partial derivatives of  $Q$  with respect to each  $x_j$  as well as with respect to  $\lambda$ . Setting these partial derivatives equal to 0 and solving yields  $x_j^* = \hat{\beta}_j / 2\lambda$  (for  $j = 1, 2, \dots, k$ ), which gives the coordinate of  $x_j^*$  on the path of steepest ascent.

To calculate the value of  $x_j^*$ , a value for  $\lambda$  must be determined, which is usually done by choosing a particular variable and an amount of change in the variable,  $\Delta$ . Using  $\Delta$ , the value for  $\lambda$  can be found, and then the new settings in the remaining  $k - 1$  variables can be found. In general, the first step along the path is  $m\Delta$ , where  $m \geq 1$ . Most times,  $m = 1$ ; however, if you believe that you are far removed from the optimum, a larger value for  $m$  might be used. Assuming  $m = 1$ , additional experiments are performed along the path at points corresponding to increments of distances  $2\Delta$ ,  $3\Delta$ , and so on. See Myers and Montgomery (1995, pp. 184–186) and Khuri and Cornell (1996, pp. 158–160) for more details. Sztendur and Diamond (2002) extend the calculations for confidence regions for the path of steepest ascent to include heterogenous errors, nonorthogonal designs, and generalized models.

### Calculating the Steepest Ascent Path in a Split-Plot Experiment

Suppose the experimental situation is that of a split-plot. Also, assume that a screening experiment has been performed in step 1, that only a few active factors have been identified, and that some of these factors are hard to change and some are easy to change. Consider the hard-to-change factors as the whole-plot variables, denoted by  $z_i$ , and the easy to change factors as the subplot variables, denoted by

$x_j$ . Then the fitted first-order model will be of the form

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^a \hat{\delta}_i z_i + \sum_{j=1}^b \hat{\gamma}_j x_j.$$

This model is then used to calculate the path of steepest ascent.

Because split-plot experiments contain two-types of factors, we propose the use of  $\Delta_i$  to represent the increment of movement in the whole-plot factors and  $\Delta_j$  to represent the increment of movement in the subplot factors. There are several reasons for the choice of two distinct  $\Delta$ 's:

- It is consistent with the way the experiment was carried out, i.e., in a split plot, which separates the two types of factors.
- The hard-to-change factors have a different standard error than the easy-to-change factors; therefore, the largest regression coefficient across all factors could be misleading.
- By their very nature, some factors are hard to change, and so choosing the step size for the path based on the regression coefficient of an easy-to-change factor may lead to a level that is not possible for one of the hard-to-change factors. For example, it may be easier to change temperature in steps of, say  $10^\circ\text{F}$  than to have an easy-to-change factor dictate that temperature be changed in steps of, say,  $6.42^\circ\text{F}$ . On the other hand, having a hard-to-change factor dictate the step size in the easy-to-change factors may require the steps to be larger than desired and may force the path to the boundary of the region.

Although the experimenter can choose to have  $\Delta_i = \Delta_j = \Delta$ , in this paper, we will use two separate  $\Delta$ 's for the reasons mentioned above.

Let  $r_1$  represent the radius for the whole-plot factors,  $\sum_{i=1}^a z_i^2 = r_1^2$ , and let  $r_2$  represent the radius for the subplot factors,  $\sum_{j=1}^b x_j^2 = r_2^2$ . Then there are two equations involving Lagrange multipliers,

$$Q(z_1, z_2, \dots, z_a) = \hat{\beta}_0 + \sum_{i=1}^a \hat{\delta}_i z_i - \lambda \left( \sum_{i=1}^a z_i^2 - r_1^2 \right)$$

and

$$L(x_1, x_2, \dots, x_b) = \sum_{j=1}^b \hat{\gamma}_j x_j - \theta \left( \sum_{j=1}^b x_j^2 - r_2^2 \right),$$

where  $\lambda$  and  $\theta$  are the Lagrange multipliers. Because we are assuming a first-order model with no inter-

actions between the whole-plot and subplot factors, these two equations can be solved separately. Therefore, taking partial derivatives of  $Q$  with respect to each of the  $z_i$  and  $\lambda$  will result in the equations

$$\frac{\partial Q}{\partial x_i} = \hat{\delta}_i - 2\lambda z_i = 0 \quad i = 1, 2, \dots, a$$

and

$$\frac{\partial Q}{\partial \lambda} = - \sum_{i=1}^a z_i^2 + r_1^2 = 0.$$

Likewise, similar equations will result from taking partial derivatives of  $L$  with respect to each  $x_j$  and  $\theta$ . Hence, the solutions will be  $z_i^* = \hat{\delta}_i/2\lambda$  and  $x_j^* = \hat{\gamma}_j/2\theta$ . Therefore, two variables,  $z_i^*$  and  $x_j^*$ , and their amount of change must be chosen to determine the path. There will be two centers or bases; one for the whole-plot factors (Base1) and one for the subplot factors (Base2). Each base will be  $(0, 0, \dots, 0)$  in the coded variables.

### Example

Consider an experiment in which interest lies in examining the image quality of a printing process. A screening experiment has identified five important factors: (A) *blanket type* (in terms of thickness), (B) *paper type* (in terms of thickness), (C) *cylinder gap*, (D) *ink flow*, and (E) *press speed*. Figure 1 is a simple diagram of this part of a printing press.

Changing the cylinder gap, ink flow, and press speed is a very simple procedure and merely consists of making an adjustment on a control panel while the printing press is still running. Therefore, these are the easy-to-change factors in the experiment. Changing the blanket type and paper type, on the other hand, requires the press to be stopped and a manual replacement of the blanket and/or paper type. Thus, these two factors are hard-to-change factors. A completely randomized design would run the 32 treatment combinations in a random order, requiring frequent stopping of the press so that the blanket type and/or paper type could be changed. Instead, a split-plot design is used with two whole-plot variables,  $z_1$  and  $z_2$ , and three subplot variables,  $x_1$ ,  $x_2$ , and  $x_3$ . The levels in the natural variables are provided in Table 1. Let

$$\hat{y} = 60 + 4.2z_1 + 6.8z_2 + 1.4x_1 - 3.6x_2 + 2.2x_3 \quad (1)$$

be the fitted first-order model. Suppose a change of 2 natural units in  $z_2$  (the whole-plot variable with the largest effect), equivalent to 1 coded unit, is chosen for determining the path in the whole-plot variables. Then  $\lambda = 6.8/2(1) = 3.4$ , and the correspond-

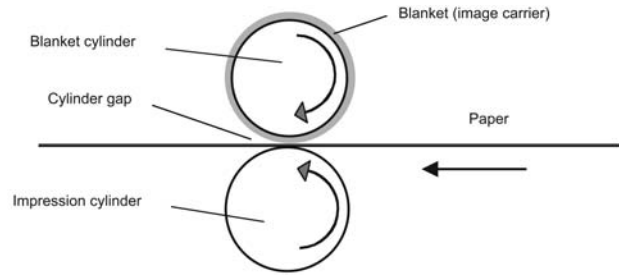


FIGURE 1. The Printing Press.

ing change in  $z_1$  is  $4.2/2(3.4) = 0.618$ . Now suppose that a change of 5 natural units in  $x_2$  (the subplot variable with the greatest effect), equivalent to 1 coded unit, is chosen for determining the path in the subplot variables. Then  $\theta = -3.6/2(1) = -1.8$  and the change in  $x_j^*$  is  $\hat{\beta}_j/2(|\theta|)$ . So the change in  $x_3$  is  $2.2/2(1.8) = 0.611$  and the change in  $x_1$  is  $1.4/2(1.8) = 0.389$ . Table 2 shows the steps for the path of steepest ascent in the coded variables.

### Running the Path of Steepest Ascent in a Split-Plot Experiment

When the experiment is completely randomized, runs are typically carried out one at a time on the path of steepest ascent, which would not be realistic in the split-plot setting. When setting up the order of the split-plot runs, careful consideration has to be given to the restricted randomization. After a setting in the whole-plot factors is determined, several subplots should be run. One must decide how to run the various whole-plot settings as well as which subplot settings will go in each whole-plot. We propose three methods for carrying out the method of steepest ascent in a split-plot experiment. We will use the model in Equation (1) to illustrate the methods. Also, throughout the paper, we will use  $m = 1$  for the first step along the whole plot and subplot paths. Typically, there are only 4–6 steps along the path before the response starts to drop off.

TABLE 1. Levels of the Factors in Natural Variables

	$z_1$	$z_2$	$x_1$	$x_2$	$x_3$
Low	10	4	20	5	2
High	20	8	40	15	6

TABLE 2. Steps in Coded Variables Along the Path of Steepest Ascent for a Split-Plot

	$z_1$	$z_2$		$x_1$	$x_2$	$x_3$
Base1	0	0	Base2	0	0	0
$\Delta_i$	0.618	1.0	$\Delta_j$	0.389	-1.0	0.611
Base1+ $\Delta_i$	0.618	1.0	Base2+ $\Delta_j$	0.389	-1.0	0.611
Base1+2 $\Delta_i$	1.236	2.0	Base2+2 $\Delta_j$	0.778	-2.0	1.222
Base1+3 $\Delta_i$	1.854	3.0	Base2+3 $\Delta_j$	1.167	-3.0	1.833

**Method 1**

This method fixes a whole plot setting and carries out the subplot path. Then, using the result from the subplot path it carries out the whole plot path with fixed subplot settings. Using  $m = 1$ , the first whole-plot setting is Base1 +  $\Delta_i$ . Within this setting of the whole-plot factors, many subplots can be used for the subplot path. For example, say 6 subplots: Base2 +  $\Delta_j$ , Base2 + 2 $\Delta_j$ , ..., Base2 + 6 $\Delta_j$  run in random order. This will determine the subplot setting with the highest response. Now steps must be made along the path in the whole-plot factors. The whole-plot path starts with Base1 +  $\Delta_i$  and Base1 + 2 $\Delta_i$ . In each of these whole-plot settings, some replicates, say 4, of the highest response setting in the subplot factors from the previous subplot path are run. The average of these four subplot runs will serve as the response value for the whole-plot path. Continue running new whole-plot settings along the path until the average response values drop off. This factor setting along the whole-plot path and the previously determined subplot settings from the subplot path could be used as the center for any follow-up experimentation.

Consider the model fit in Equation (1). Then, as an example of Method 1, first the runs in Table 3 are carried out in random order. The subplot path

TABLE 3. First Whole-Plot Run With Setting Base1 +  $\Delta_i$  Using Method 1

Subplot runs	$x_1$	$x_2$	$x_3$	Response
Base2 + $\Delta_j$	0.389	-1	0.611	69
Base2 + 2 $\Delta_j$	0.778	-2	1.222	73
Base2 + 3 $\Delta_j$	1.167	-3	1.833	75
Base2 + 4 $\Delta_j$	1.556	-4	2.444	73
Base2 + 5 $\Delta_j$	1.945	-5	3.055	70
Base2 + 6 $\Delta_j$	2.334	-6	3.666	66

gives the setting Base2 + 3 $\Delta_j$  because it has the highest response (75). Next, more whole-plot settings are used, all with the same subplot setting Base2 + 3 $\Delta_j$ . The average response of the subplots will be the whole-plot response. From Table 4, the whole-plot response drops off after Base1 + 2 $\Delta_i$ . Thus, this example produces the setting of Base1+2 $\Delta_i$  in the whole-plot variables and the setting of Base2 + 3 $\Delta_j$  in the subplot variables as the center of the follow-up experiment. The design point in coded units for this setting is  $z_1 = 1.236$ ,  $z_2 = 2.0$ ,  $x_1 = 1.167$ ,  $x_2 = -3$ , and  $x_3 = 1.833$ .

The number of steps required in this particular example, using the split-plot method 1 approach, was five. It is important to keep in mind the split-plot nature in determining the steps to termination of the path. In this example, even though 18 runs were carried out, the path terminates after only 2 steps in the whole-plot variables and 3 steps in the subplot variables.

**Method 2**

Method 2 is essentially the reverse of method 1. Instead of determining the ideal subplot setting first

TABLE 4. New Whole-Plot Runs With Subplot Setting Base2 + 3 $\Delta_j$  Using Method 1

Subplot setting	Whole-Plot Settings		
	Base1+ $\Delta_i$	Base1+2 $\Delta_i$	Base1+3 $\Delta_i$
Base2 + 3 $\Delta_j$	76	79	74
Base2 + 3 $\Delta_j$	75	77	75
Base2 + 3 $\Delta_j$	76	77	74
Base2 + 3 $\Delta_j$	75	76	73
Average	75.5	77.25	74.0

TABLE 5. Whole-Plot Runs With Subplot Setting Base2 + Δ<sub>j</sub> Using Method 2

Subplot setting	Whole-Plot Settings		
	Base1+Δ <sub>i</sub>	Base1+2Δ <sub>i</sub>	Base1+3Δ <sub>i</sub>
Base2+Δ <sub>j</sub>	70	73	71
Base2+Δ <sub>j</sub>	72	75	73
Base2+Δ <sub>j</sub>	69	72	70
Base2+Δ <sub>j</sub>	69	72	70
Average	70.0	73.0	71.0

and then finding the whole-plot setting, method two finds the whole-plot setting first and then determines the subplot setting. Therefore, one would start by running two whole-plot settings, Base1 + Δ<sub>i</sub> and Base1 + 2Δ<sub>i</sub>, with each containing, say, 4 replicates of Base2 + Δ<sub>j</sub>. Using the average of the subplot runs as the response, continue running whole-plot settings along the whole-plot path until the average response drops off. This will determine the whole-plot setting. Then, in this whole-plot setting, say, 6 subplots can be run: Base2 + Δ<sub>j</sub>, Base2 + 2Δ<sub>j</sub>, . . . , Base2 + 6Δ<sub>j</sub>. These can be used to determine where the subplot responses drop off. The setting from this subplot path and the previously determined whole-plot setting from the whole-plot path could serve as the center for any follow-up experimentation.

Using Equation (1) for an example of method 2, the runs in Table 5 are carried out, resulting in Base1 + 2Δ<sub>i</sub> for the whole-plot variables. Next, subplot settings are run in random order, all with the same whole-plot setting Base1 + 2Δ<sub>i</sub> from the whole plot path. From Table 6, the subplot response drops off after Base2 + 3Δ<sub>j</sub>, so Base1+2Δ<sub>i</sub> in the whole-plot variables and Base2 +3Δ<sub>j</sub> in the subplot variables

TABLE 6. Subplot Settings at the Best Whole-Plot Setting Base1 +2Δ<sub>i</sub> Using Method 2

Subplot runs	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	Response
Base2 + Δ <sub>j</sub>	0.389	-1	0.611	73
Base2 + 2Δ <sub>j</sub>	0.778	-2	1.222	76
Base2 + 3Δ <sub>j</sub>	1.167	-3	1.833	78
Base2 + 4Δ <sub>j</sub>	1.556	-4	2.444	75
Base2 + 5Δ <sub>j</sub>	1.945	-5	3.055	73
Base2 + 6Δ <sub>j</sub>	2.334	-6	3.666	69

can be used as the center for any follow-up experimentation. The design point for this setting would be z<sub>1</sub> = 1.236, x<sub>2</sub> = 2.0, x<sub>1</sub> = 1.167, = -3, and x<sub>3</sub> = 1.833. In this example, the same termination point is achieved from methods 1 and 2, although this will not always be the case.

It should be noted that both methods 1 and 2 rely heavily on the assumption of no interaction between whole-plot and subplot factors. However, when the method of steepest ascent is applied in the usual completely randomized RSM experiments, a strict first-order model is assumed to be adequate. Therefore, the assumption of no interaction between whole-plot and subplot factors is not an additional assumption, just consistent with typical RSM situations.

### Choosing Between Methods 1 and 2

There is no simple way to choose between methods 1 and 2 in an optimum fashion, where, by optimum, we mean choosing the method that would result in the most rapid movement to the region of the maximum. The optimum choice depends on the starting point in the true response surface and the shape of this surface. These are, of course, unknown to the experimenter. We have drawn many types of surfaces and used different starting points. Applying methods 1 and 2 to these surfaces resulted in very little difference in terms of total number of runs needed. Figures 2 and 3 illustrate the two methods on one such surface for one starting point (the bold symbols indicate steps along the whole plot path and the + symbol indicates steps along the subplot path). Figure 2 uses method 1, while Figure 3 illustrates method 2. Both

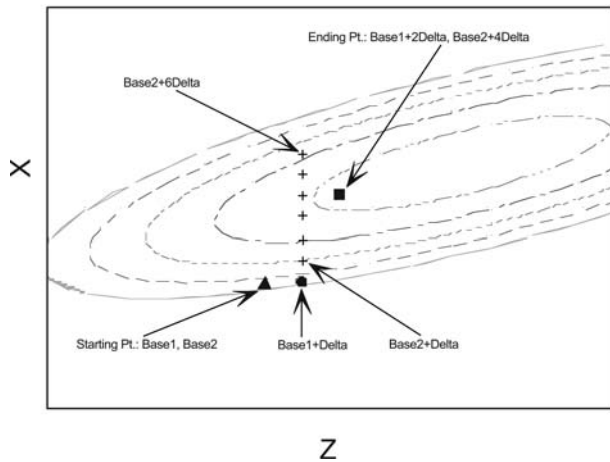


FIGURE 2. Trajectory Using Method 1.

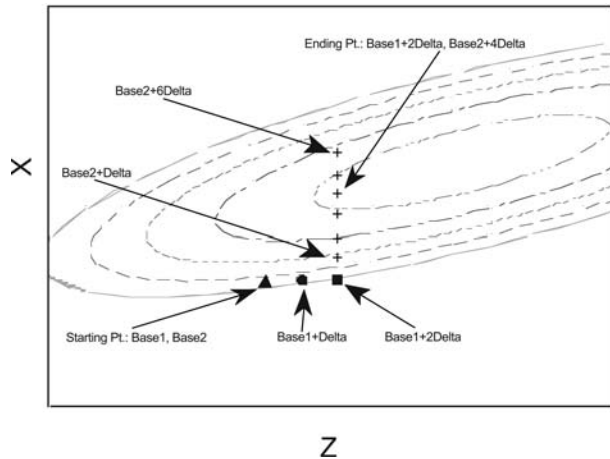


FIGURE 3. Trajectory Using Method 2.

methods require 2 steps along the whole plot path and 4–5 steps along the subplot path.

From a practical point of view, if the largest standardized regression coefficient (coefficient/standard error of the coefficient) belongs to a hard-to-change factor, method 2 seems like an obvious choice. This would allow the experimenter to make great leaps in the hard-to-change factors toward the optimum and then minor refinements in the easy-to-change factors. On the other hand, if the largest standardized regression coefficient belongs to an easy-to-change factor, method 1 seems like a better choice because movement in the easy-to-change factors should provide the quicker ascent.

It is possible to combine or switch between the two methods. Suppose while using method 1 there is very little change in the response as the steps along the path in the easy-to-change factors are carried out. This may lead the experimenter to switch to method

2 and implement steps along the path in the hard-to-change factors to move quicker to an optimum point.

### Method 3

In this method, we propose running a certain number of whole plots, each with the same number of subplots. This can be thought of as one experiment that carries out both paths at the same time. It is probably practical to run, say 4 whole plots, each with 6 subplots. The whole-plot settings of  $Base1 + \Delta_i$ ,  $Base1 + 2\Delta_i$ ,  $Base1 + 3\Delta_i$ , and  $Base1 + 4\Delta_i$  should be randomized. In each of these settings, the subplot settings of  $Base2 + \Delta_j$ ,  $Base2 + 2\Delta_j, \dots, Base2 + 6\Delta_j$  should be separately randomized for each whole-plot setting. We assume there is no whole-plot by subplot interaction, which is consistent with RSM applications and the model in Equation (1). There are 2 ways to determine the response setting that will serve as the center of a follow-up experiment. One approach is to first find the average of the subplot settings for each whole-plot setting. These can be compared to determine the maximum. Next, the average of the same subplot settings across the whole plots can be calculated to determine the maximum in the subplot settings. Combining these two gives the appropriate point. Another approach is to simply take the highest response from the 24 runs and make this the new center for a follow-up experiment.

Using Equation (1) as an example, suppose an experiment with 4 whole plots and 6 subplots is chosen, with the resulting runs displayed in Table 7. In this example, the highest response occurs at  $(Base1+2\Delta_i, Base2+3\Delta_j)$ . Also, the highest average subplot setting is  $Base2+3\Delta_j$ , and the highest average whole-plot setting is  $Base1+2\Delta_i$ . The design point at this setting would be  $z_1 = 1.236, z_2 = 2.0, x_1 = 1.167, x_2 = -3, \text{ and } x_3 = 1.833$ . This agreement between

TABLE 7. Runs of the Experiment for Method 3 (bold numbers indicate highest yields)

Subplots	$Base1+\Delta_i$	$Base1+2\Delta_i$	$Base1+3\Delta_i$	$Base1+4\Delta_i$	Average
$Base2+\Delta_j$	70	72	72	69	70.25
$Base2+2\Delta_j$	74	75	73	71	73.25
$Base2+3\Delta_j$	74	<b>79</b>	76	75	<b>76.00</b>
$Base2+4\Delta_j$	73	75	75	71	73.50
$Base2+5\Delta_j$	70	74	73	70	71.75
$Base2+6\Delta_j$	67	70	70	67	68.50
Average	71.33	<b>74.16</b>	73.16	70.50	

the highest overall setting and the highest average settings may not occur in all situations.

The advantage of method 3 compared with methods 1 and 2 is that all the data for the experiment is collected together. This can be very important in some industrial settings. For example, in semiconductor manufacturing, multiple runs may be performed on a single-wafer tool during one time period. These runs involve changing both hard-to-change and easy-to-change factors on the tool. The wafers are then sent to a lab for processing, and it may take up to a week to get the response values back. Therefore, it is important to collect all of the data at one time, and it would be impractical to attempt to collect the measurements sequentially, as would be required with methods 1 and 2.

One possible disadvantage of method 3 is that the response may not drop off after, say, 4 steps in the hard-to-change factors and/or 6 steps in the easy-to-change factors. When using any of these methods, it is useful to have a rule for deciding when to stop conducting experiments along the path. A good rule of thumb is to continue experimenting along the path of steepest ascent until two consecutive runs have resulted in a decrease in the response.

### Summary

Often times, due to cost or time constraints, it is not feasible to completely randomize an experiment. Typically, some factors are hard to change and others are easy to change, which results in a split plot. Many experiments in industry are RSM studies in which the goal is to find the optimal conditions for the design factors. One important step in a RSM study is moving through the region using the path of steepest ascent toward a region where the optimum lies. It would be virtually impossible to run one run at a time in a split-plot setting because of the randomization restrictions. We have introduced two equations involving Lagrange multipliers to handle the different types of factors, whole plot and subplot. After defining how to calculate the two paths, we propose three methods for carrying out the paths of steepest ascent within a split-plot framework. Two of the methods involve sequential experimentation by either fixing a whole-plot setting first, varying the subplot settings, and then varying the whole-plot settings or varying the whole-plot settings with fixed subplot settings and then varying the subplot settings in a fixed whole-plot setting. The third method just runs one experiment by varying both the whole-

plot settings and subplot settings together, which covers cases where all of the data must be collected at one time. Each method has advantages and disadvantages. The best method will depend on the particular experimental situation, the starting point used, the shape of the true response surface, experience, and results from previous experiments.

### Acknowledgments

The authors would like to thank the Editor and an anonymous reviewer for their insightful comments on an earlier version of this paper.

### References

- BINGHAM, D. and SITTE, R. S. (1999a). "Minimum-Aberation Two-Level Fractional Factorial Split-Plot Designs". *Technometrics* 41, pp. 62–70.
- BINGHAM, D. and SITTE, R. S. (1999b). "Some Theoretical Results for Fractional Factorial Split-Plot Designs". *The Annals of Statistics* 27, pp. 1240–1255.
- BINGHAM, D. and SITTE, R. S. (2001). "Design Issues in Fractional Factorial Split-Plot Experiments". *Journal of Quality Technology* 33, pp.1–15.
- BISGAARD, S. (2000). "The Design and Analysis of Split Plot Experiments". *Journal of Quality Technology* 32, pp. 39–56.
- BISGAARD, S. and STEINBERG, D. M. (1997). "The Design and Analysis of Prototype Experiments". *Technometrics* 39, pp. 52–62.
- BOX, G. E. P. (1996). "Split-Plot Experiments". *Quality Engineering* 8, pp. 515–520.
- BOX, G. E. P. and JONES, S. (1992). "Split-Plot Designs for Robust Product Experimentation". *Journal of Applied Statistics* 19, pp. 3–26.
- BOX, G. E. P. and WILSON, K. B. (1951). "On the Experimental Attainment of Optimum Conditions". *Journal of the Royal Statistical Society, Series B*, 13, pp. 1–45.
- GOOS, P. and VANDEBROEK, M. (2001). "Optimal Split-Plot Designs". *Journal of Quality Technology* 33, pp. 436–450.
- HUANG, P.; CHEN, D.; and VOELKEL, J. O. (1998). "Minimum-Aberation Two-Level Split-Plot Designs". *Technometrics* 40, pp. 314–326.
- KHURI, A. I. and CORNELL, J. A. (1996). *Response Surfaces: Designs and Analyses, 2nd ed.* Dekker, New York, NY.
- KOWALSKI, S. M. (2002). "24-Run Split-Plot Experiments for Robust Parameter Design". *Journal of Quality Technology* 34, pp. 399–410.
- KOWALSKI, S. M.; VINING, G. G.; MONTGOMERY, D. C.; and BORROR, C. M. (2004). "Modifying a Central Composite Design to Model the Process Mean and Variance When There Are Hard-to-Change Factors". Submitted to *Technometrics*.
- LETSINGER, J. D.; MYERS, R. H.; and LENTNER, M. (1996). "Response Surface Methods for Bi-Randomization Structures". *Journal of Quality Technology* 28, pp. 381–397.
- LUCAS, J. M. and HAZEL, M. C. (1997). "Running Experiments With Multiple Error Terms: How an Experiment Is Run Is Important". *ASQC Technical Conference Transac-*

- tions. American Society for Quality Control, Milwaukee, WI. pp. 283–296.
- LUCAS, J. M. and JU, H. L. (1992). “Split Plotting and Randomization in Industrial Experiments”. *ASQC Quality Congress Transactions*. American Society for Quality Control, Nashville, TN. pp. 374–382.
- MYERS, R. H. and MONTGOMERY, D. C. (1995). *Response Surface Methodology, Process and Product Optimization Using Design Experiments*. Wiley, New York, NY.
- MYERS, R. H.; MONTGOMERY, D. C.; VINING, G. G.; BORROR, C. M.; and KOWALSKI, S. M. (2004). “Response Surface Methodology: A Retrospective and Literature Survey”. *Journal of Quality Technology* 36, pp. 53–77.
- SCHOEN, E. D. (1999). “Designing Fractional Two-Level Experiments With Nested Error Structures”. *Journal of Applied Statistics* 26, pp. 495–508.
- SZTENDUR, E. M. and DIAMOND, N. T. (2002). “Extensions to Confidence Region Calculations for the Path of Steepest Ascent”. *Journal of Quality Technology* 34, pp. 289–296.
- VINING, G. G.; KOWALSKI, S. M.; and MONTGOMERY, D. C. (2004). “Response Surface Designs Within a Split-Plot Structure”. *Journal of Quality Technology* (to appear).

