Design of experiments strategies

The traditional approach to experimentation requires changing only one factor at a time (OFAT). However, the OFAT approach does not provide data on interactions of factors, a likely occurrence with processes.

Statistically based design of experiments (DOE) provides valid process models. The two-phase strategy of DOE is simple and straightforward:

Phase 1. Use two-level factorial designs as screening tools to separate the vital few factors (including interactions) from those with no significant impact.

Phase 2. Do an in-depth investigation of the surviving factors. Generate a response surface map and move the process to the optimum location.

These sections discuss the two phases of DOE.

■ **Phase 1: Screening with two-level factorials.** Two-level factorial design involves simultaneous adjustment of experimental factors at high and low levels. By restricting the tests to only two levels, the number of experiments is minimized. The contrast between levels provides the driving force for process improvement. This parallel testing scheme is more efficient than OFAT. To obtain high-resolution of effects, run full factorials for four or fewer factors.

Table 1 shows high-resolution fractional design options for five or more factors. Designs can be constructed with a textbook or statistical software.

Designs are available with as little as k+1 runs, where k equals the number of factors to test. For example, 7 factors can be tested in 8 runs, or 15 factors can be tested in 16 runs. However, these "saturated" designs provide very poor resolution because main effects will be confused with two-factor interactions. Running low resolution designs should be avoided.

The run order of the entire design should be randomized. Otherwise, lurking factors, such as ambient temperature or catalyst degradation, could confound the factor estimates. After completing the experiments, standard statistical analyses provide significance tests on the overall outcome and individual effects. Textbooks provide hand-calculation schemes for doing analysis of two-level factorials, but it is much easier to let a statistical software program do this work.

To be conservative, consider putting a

TABLE 1. HIGH-RESOLUTION, TWO-LEVEL FRACTIONAL FACTORIALS.

Factors	No. runs, full two-level design	Fraction applied	No. runs, fraction design
5	32 (2^5)	1/2 (2^{-1})	16 (2^{5-1})
6	64 (2^6)	1/2 (2^{-1})	32 (2^{6-1})
7	128 (2^7)	1/4 (2^{-2})	32 (2^{7-2})
8	256 (2^8)	1/8 (2^{-3})	32 (2^{8-3})
9	512 (2^9)	1/8 (2^{-3})	64 (2^{9-3})
10	1,024 (2^{10})	1/16 (2^{-4})	64 (2^{10-4})
11	2,048 (2^{11})	1/32 (2^{-5})	64 (2^{11-5})

TABLE 2. EXPERIMENTAL DATA FOR EXAMPLE.

Point type	Time (X_1), min	Temperature (X_2),°C	Catalyst (X_3), %	Conversion (Y_1), %	Activity (Y_2)
Factorial	40.00	80.00	2.00	74	53.2
Factorial	50.00	80.00	2.00	51	62.9
Factorial	40.00	90.00	2.00	88	53.4
Factorial	50.00	90.00	2.00	70	62.6
Factorial	40.00	80.00	3.00	71	57.3
Factorial	50.00	80.00	3.00	90	67.9
Factorial	40.00	90.00	3.00	66	59.8
Factorial	50.00	90.00	3.00	97	67.8
Center	45.00	85.00	2.50	81	59.2
Center	45.00	85.00	2.50	75	60.4
Center	45.00	85.00	2.50	76	59.1
Center	45.00	85.00	2.50	83	60.6
Center	45.00	85.00	2.50	80	60.8
Center	45.00	85.00	2.50	91	58.9
Axial	36.59	85.00	2.50	76	53.6
Axial	53.41	85.00	2.50	79	65.9
Axial	45.00	76.59	2.50	85	60.0
Axial	45.00	93.41	2.50	97	60.7
Axial	45.00	85.00	1.66	55	57.4
Axial	45.00	85.00	3.34	81	63.2

Fig. 1. A 2-D contour plot.

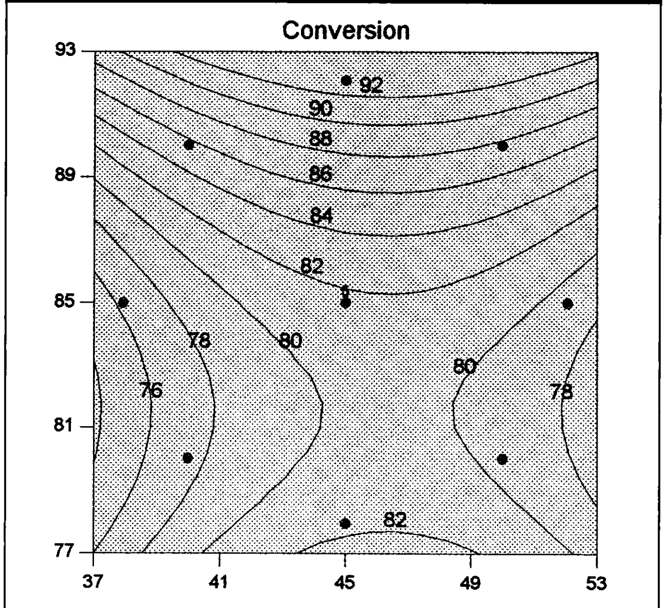
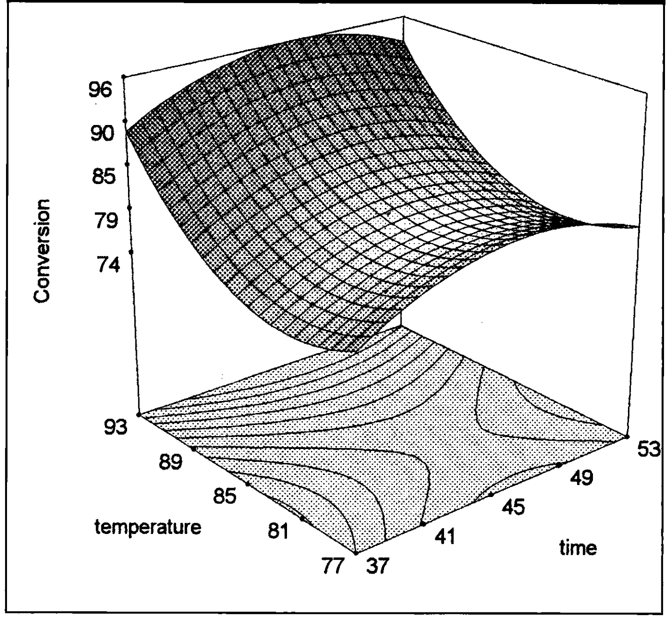


Fig. 2. A 3-D contour plot.



centerpoint in the design. This is a set of conditions at the midpoint of every factor level. For example, if the design requires that time vary from 40 min to 50 min while the temperatures vary from 80°C to 90°C, then the centerpoint setting is 45 min at 85°C. To get an estimate of pure error, the centerpoint is repeated several times, mixed randomly with the remaining design points.

Each time the centerpoint is run, all of the steps are repeated. Including reanalyzed or resampled data in the method as replicates will not produce a fair estimate of the pure error.

The "curvature" in the system can be estimated with centerpoints. In most cases curvature is not significant, which means that the two-level design is reliable. However, if the system is close to peak performance, there will be significant curvature, which indicates that the response behaves in a non-linear fashion. Then, one needs to run additional factor levels and employ response surface methods.

■ **Phase 2. Optimization via response surface methods.** As the optimum is approached, it becomes necessary to do more than two levels of the factors that survive the screening phase. The tool for Phase 2 is response surface methodology (RSM).

Begin by focusing the study on a specific region of interest, identified by prior experimentation as a likely location for the optimum. The narrower the factor ranges are set, the more likely one will be able to approximate the surface with a simple polynomial model. In most cases, a quadratic equation proves to be sufficient. Equation 1 shows a quadratic equation for two factors.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 \qquad (1)$$

where:

Y = Measured response or outputs such as yield, efficiency, conversion (Y_1) and activity (Y_2);

X = Input factors such as time (X_1), temperature (X_2) and catalyst (X_3);

β = Mathematical model coefficients assigned to input factors.

A three-level factorial design provides a sufficient number of experiments to fit a quadratic model. It works well for two factors.

However, for three factors the three-level design requires 27 experiments, an excessive number. The central composite design (CCD) is a better choice for response surface experiments with three or more factors. The CCD is composed of a core two-level factorial surrounded by axial points.

Fig. 1 shows the points' layout for a central composite design on two process factors. Larger CCDs can be divided between two or more blocks if all of the experiments cannot be run together due to lack of time, material or equipment availability.

For example, one could run the factorial portion and then follow-up with the axial points. By running centerpoints with the factorial, one can check for curvature in the response. If there is no significant curvature, the axial points need not be run. More centerpoints should be included with the axial points to tie the blocks together. Then, any constant difference in response due to blocks can be arithmetically removed.

Montgomery[1] and Myers and Montgomery[2] provide details on CCDs and other designs for response surface methods.

Statistical software is handy for setting up RSM designs, and a necessity for analysis. These software packages provide the means to fit the response data to the quadratic polynomial model. A standard analysis of variance (ANOVA) indicates if the fit is statistically significant. The software will provide maps that show the way to optimal process performance. Graphical outputs make it easy to explain how critical factors affect the process. Figs. 1 and 2 show examples of these maps in 2-D and 3-D, respectively.

Example

Table 2 shows factor settings and resulting responses from a CCD on a reactor. Figs. 1 and 2 are derived from this data. The design is comprised of eight factorial rows with all combinations of the 3 factors each at 2 levels. This is a good starting point for screening purposes. The effects can be calculated from the contrast of the average responses at high versus low levels.

The design has six center rows at the midlevel of the factor ranges. The average of the associated responses can be compared to the average of the factorial points for an indication of curvature in the surface. There also are six axial rows from the optimization phase. These are extreme points outside of the original factorial range. They provide information on the shape of the response surface.

To analyze the data, each of the two responses is fitted to a polynomial model via ordinary least-squares regression. The design provides sufficient levels to fit up to a quadratic polynomial such as Equation 1. Each response must be fitted separately. In this case the equations are:

Conversion = 1026.52 - 11.57 * time - 18.70 * temperature + 44.48 * catalyst - 0.076 * time^2 + 0.12 * temperature^2 - 21.01 * catalyst^2 + 0.085 * time * temperature + 4.55 * time * catalyst - 1.55 * temperature * catalyst.

Activity = 6.39 + 0.85 * time + 0.051 * temperature + 4.46 * catalyst.

The second order terms for activity were not statistically significant, so this model is simply linear. These equations then can be used as predictive models to accomplish specific objectives. However, they will be valid only within the factorial range. Anywhere else will be an extrapolation which would need to be verified by additional experimentation.

Based on the fitted model, conversion is maximized when each of the three factors are set at their highest factorial levels (50 min, 90°C and 3% catalyst). Even higher levels of conversion might be found outside of this range via another DOE.

The second response, activity, has a specification of 60 to 66 with a target of 63. When the conditions favoring high conversion (50, 90, 3) are plugged in to the activity model, the predicted response is approximately 67, which falls outside of specification. A trade-off must be made. This can be accomplished via trial and error or with the aid of software. In this case, a desirable tradeoff can be achieved at the following conditions: 47 min, 90°C and 2.68% catalyst that give predicted responses at 91% conversion and 63 for activity.

References

1. Montgomery, D.C., Design and Analysis of Experiments, 3rd ed., John Wiley & Sons, Inc., NY, 1991.

2. Myers, R.H., and Montgomery, D.C., Response Surface Methodology, John Wiley & Sons, Inc., NY, 1995.

3. Kraber, Shari, Whitcomb, Patrick, and Anderson, Mark, "Handbook for Experimenters," Stat-Ease Inc., Minneapolis, MN, 1997.

By Mark J. Anderson, PE, and Patrick Whitcomb, PE, Principals, Stat-Ease Inc., Minneapolis, MN.