Optimize Your Process-Optimization Efforts

Highly efficient, statistically based methods can identify the vital few factors that affect process efficiency and product quality.

Mark J. Anderson and Patrick J. Whitcomb
Stat-Ease, Inc.
2021 East Hennepin Ave., Suite 480, Minneapolis, MN 55413
Telephone: 612.378.9449, Fax: 612.378.2152

(Writer's note: this is the first of a two-part series on design of experiments. Part two provides details on a relatively new technique called mixture design -- a powerful technique for improving formulations.)

What would you do if confronted with "opportunities" like these?

1. A chemist synthesizes a new wonder compound in the laboratory. The marketing people anticipate big profits but only if the product can be made economically at high yields. The chemist provides a very basic recipe sheet with suggested conditions. Your job is to scale up the process as quickly as possible.

2. A competitor makes a minor, but noticeable, improvement to their product while simultaneously reducing price. The sales group fears that your product may get knocked out of the market. Your job is to fine tune the existing plant process and get more yield with better product quality.

In either case, you probably would first try to gather fellow experts and identify all possible variables that may affect yields and product quality. An exhaustive list might include dozens of potential factors - many more than you could possibly investigate. Some of these variables can't be controlled: Be sure to record their values. Other variables won't be given much priority: Hold these at fixed levels. Still, you're likely to be left with five or more possible control factors. Now what do you do?

The traditional approach to experimentation requires you to change only one factor at a time (OFAT). However, the OFAT approach doesn't provide data on interactions of factors, a likely occurrence with chemical processes. An alternative approach called "two-level factorial design" can uncover critical interactions. This statistically based method involves simultaneous adjustment of experimental factors at only two levels: high and low. The two-level design offers a parallel testing scheme that's much more efficient than the serial approach of OFAT.

By restricting the tests to only two levels, you minimize the number of experiments. The contrast between levels gives you the necessary driving force for process improvement. You don't need to run three levels of every factor until you get close to the optimum. At the beginning of your investigation many factors must be considered. Doing all combinations of every factor at three levels would produce a prohibitively large number of runs.

Strategy of experimentation

The basics of two-level factorial design are well-documented.[1,2] To illustrate their application to chemical engineering, we will investigate a process optimization of waterborne polyurethane.[3] The chemical engineers who performed the experiments wanted small particles to ensure a stable aqueous dispersion. This case study illustrates a very basic two-phase strategy for experimenters:

> Phase 1. Use two-level factorial designs as screening tools to separate the vital few factors (including interactions) from the trivial many that have no significant impact.
>
> Phase 2. Follow up by doing an in-depth investigation of the surviving factors. Generate a "response surface" map[4] and move the process to the optimum location.

Part two of this series of articles on design of experiments will cover the optimization via response surface methods (RSM). This article focuses on the screening phase. It presents the vital tools of statistical design of experiments (DOE) via two-level factorial design. We will do this from an engineering perspective, with an emphasis on the practical aspects.

Fractional factorials for maximum efficiency

When screening factors, you need not run the full combination of high and low levels: often a fraction will do. The polyurethane experimenters chose to run a half fraction. They studied five factors (Table 1) in sixteen experiments which can be symbolized mathematically as 2^{5-1}. It provides sound estimates of main effects and two-factor interactions if you can assume that three-factor or higher interactions will not occur. Generally this is a safe assumption, but you should always make confirmation runs to verify experimental findings. Table 2 shows high-resolution design options for five or more factors. (Run full factorials for four or fewer factors if you want high resolution of effects). The designs can be constructed with the aid of a textbook, or better yet, with a statistical software package, most of which offer design of experiments capabilities[5,6].

You will find designs available with as little as k+1 runs, where k equals the number of factors you want to test. For example you could test 7 factors in 8 runs, or 15 factors in 16 runs. However, these "saturated" designs provide very poor resolution: main effects will be confused with two-factor interactions. We advise that you avoid running such low resolution designs.

The specific design layout for the polyurethane (PU) case is shown in Table 3. Columns A through D are laid out according to a standard order that can be obtained from any textbook on design of experiments. This portion of the matrix represents a full two-level design for four factors in sixteen runs. In order to get the additional factor E into the design, the four-factor interaction ABCD is used. Multiply the A, B, C and D columns and you find the product in column E:

$$E = ABCD \tag{1}$$

Statisticians call this relation between E and ABCD an "alias". You cannot differentiate between the two: an observed difference in response due to a change in E could really be caused by ABCD, or even some combination of E and ABCD. But as

discussed earlier, you can ignore three factor or higher interactions, so don't worry. It turns out that with the half fraction, <u>every</u> effect will have another effect aliased with it. These are shown under the column labeled "Alias" in Table 4. All of these aliased effects are third order or higher so they can be ignored. If you run a 1/4th fraction, each effect will have three aliases, effects in an 1/8th fraction each have seven aliases, and so on.

Note the balanced array of plus (high) and minus (low) levels in the test matrix (Table 3). Each column contains eight pluses and eight minuses. The matrix offers a very important statistical property called "orthogonality" which means that factors are not correlated. If you just collected happenstance data from a plant, it is highly unlikely you would get an array of factors like this. You would probably find that factors such as temperature and pressure go up and down together. As factors become more and more correlated, the error in estimation of their effects becomes larger and larger. That's not good.

Orthogonal test matrices make effect estimation neat and easy. For example, the effect of factor D is calculated by simply averaging the responses at the plus level and subtracting the average at the minus levels.

$$Effect = Mean\ Y_+ - Mean\ Y_- \tag{2}$$

Applying a transformation to satisfy statistical assumptions

Notice in Table 3 that the response varies by nearly an order of magnitude, from 40 to 389 nanometers. In situations like this, statisticians routinely perform a *transformation* of the response, most commonly with a logarithm. Chemical engineering students do the same thing when they use special graph paper, such as log scale, to get the data to come out in a straight line. Contractive transformations, such as the log or square root, counteract a very common relationship: the true standard deviation (σ) increases as the true mean (μ) goes up. Statisticians express this as a power law[1]:

$$\sigma = fn(\mu^{\alpha}) \tag{3}$$

In the ideal case, there will be no relationship between standard deviation and the mean, so the power (α) will be zero. This satisfies an important statistical assumption - constant variance. If you cannot satisfy this assumption, consider a transformation. Ideally, your engineering or chemistry knowledge will guide you in the selection of an appropriate transformation. However, if you can't predict what the relationship should be, just try the log or square root or some other transformation. We got a better statistical fit from the square root transformation. The resulting responses can be seen in the last column of Table 4. Later in this article we will discuss how to validate your model, whether transformed or not, via analysis of the residuals.

Using statistical principles to pick significant factors

You now know how to calculate effects. It seems obvious that you should be pick the largest ones and run with those - right? Wrong! How do you know where to make the cut off? What if none of the effects are real, and you've just measured results due to random error? Somehow the vital few significant factors must be screened out of the trivial many that occur due to chance. You can do this easily with a graph called a "half-normal plot". Simply rank the absolute value of the effects from low to high. Then assign cumulative probability values according to the following formula.

$$P_i = 100(\frac{i - 0.5}{m}) \hspace{4cm} (4)$$

where i = rank

m = number of effects

Now plot the effects versus their assigned probability on half-normal graph paper. Then find the group of near-zero effects and draw a line through them, but don't include the points after the bend in the 'dogleg'. Anything significant will fall off to the right of the line. Figure 1 shows the half-normal plot of effects for the polyurethane case. Significant effects are labeled. The near-zero effects fall on a straight line - exhibiting normal scatter. These insignificant effects can be used to estimate experimental error.

If you want to be conservative, consider replicating the design to get estimates of "pure" error. Be sure that you go back through all of the steps, e.g., charge the reactor, bring it up to temperature, line it out, take samples and do the analysis. Don't just re-analyze or re-sample and put these in as replicated because you won't get a fair estimate of the pure error. Also, be sure to randomize the run order of your entire design, including replicates. Otherwise you leave yourself open to "lurking factors", such as ambient temperature or catalyst degradation, that could confound your factor estimates.

Another way to add replication is to put a "centerpoint" in your design. This is a set of conditions at the midpoint of every factor level. For example the centerpoint for the polyurethane experiment would be at 75 ppm catalyst, 3.65 acetone/PU, 40 degrees Centigrade, 925 rpm and 3 milliliters per minute addition rate. Most experimenters repeat the centerpoint several times mixed in randomly with the remaining design points. In addition to an estimate of pure error from the replication, you can then estimate "curvature" in your system. Significant curvature indicates that your response behaves in a non-linear fashion. Then you will need to run additional factor levels and employ response surface methods. Issues related to curvature will be discussed in more detail in part two of this series on DOE. In most cases you will find that curvature is not significant, which means that you can rely on the two-level design.

Given a valid estimate of experimental error, regardless of the source, standard statistical analyses can then be performed to validate the overall outcome and individual effects. Textbooks provide hand-calculation schemes for doing statistical analysis of two-level factorials, but it's much easier to let a statistical software program to do this work for you.

Be sure you choose a program that provides "residual analysis" capabilities. Residuals are the difference between actual and predicted response. The residuals must be approximately normal. You can check this by plotting residuals on normal or half-normal paper. Figure 2 shows the normal plot of residuals for the polyurethane case. The plot was constructed by ranking the residuals from low to high and assigning probabilities according to Equation 4. A crude but effective way to evaluate this plot is the "pencil test": If you can cover the points with your pencil, then the residuals are well-behaved. In this case, they look fine.

Your statistical software also should offer a plot of residuals versus predicted level such as that shown in Figure 3. This plot exhibits a desirable scatter - no obvious change in variation as the level increases.

If you do see bad patterns on either of these residual plots, such as an "S" shape on the normal plot (Figure 2) or a megaphone shape on the residual versus predicted plot (Figure 3), consider the use of a response transformation. The log transformation often helps. You might also try a square root or one of many other functions. Refer to a DOE textbook for statistical advice on this subject. However, there is no substitute for your process knowledge. This should guide you in selection of a transformation.

Residual analysis also may reveal individual outliers. But be careful, don't delete points unless you can assign a special cause, such as a temporary breakdown in an agitator or the like. Quite often a outlier turns out to be simply an error in data entry. Digits can get transposed very easily.

Interpreting the results

Now you are ready to make your report. Start by making a plot of any significant main effects that are not part of a significant interaction. In the polyurethane case, only factor E stands alone: It does not combine with any other factors in a significant interaction. The effect of factor E can be seen in Figure 4. Remember that each point in this plot represents a contrast between the response averages at low versus high factor levels. In this case there are eight runs at each level, so the results carry a lot of weight. Clearly factor E needs to be set at its low level to get minimal particle size. Next, produce the interaction plots. In this case we found three interactions active: AC, BC and BD, which can be seen in Figures 5,6 and 7, respectively. Notice that the lines on these plots are not parallel. In other words, the effect of one factor depends on the level of the other, so it would be inappropriate to display any of these factors by themselves. For example, on the interaction plot in Figure 7, notice that factor B (acetone/PU ratio) has a much bigger impact when D (agitation) is at the higher rate. Clearly it's best to go with low B and high D. Then from the BC plot (Figure 6) it's clear that factor C should be set at its high level. Finally, from the AC plot (Figure 5), given that C will be set at its plus level, it makes no difference how you set A, just pick a convenient level. Table 5 summarizes the recommended settings for minimizing particle size.

Before you make a final recommendation on the new factor levels, it would be wise to perform confirmation runs. You can predict the outcome with a simple equation that uses the overall average modified up or down depending on the level you set each factor to. Statisticians call this a "coded" equation because you plug in values of plus one for high and minus one for low levels. (A midpoint setting is entered as zero.)

$$Predicted\ Response = Overall\ Average + \sum_{i=1}^{p-1} (\frac{Effect_i}{2})X_i \qquad (5)$$

For the polyurethane case the predictive model is:
$$Y^{1/2} = 11.15 - 0.76A + 2.04B - 1.91C + 0.06D + 1.39E$$
$$+ 0.74AC - 0.95BC + 1.02BD \qquad (6)$$

This model includes factor D to maintain hierarchy of terms: D is one of the "parent" terms for the significant BD interaction. Although factor D may not be important alone, you can see from the BD interaction plot that it does make a difference - depending on the level of B. When B is low, increasing D causes smaller particle size. Conversely, when B is high, increasing D causes larger particle size. These "crossover" type interactions can cause confusion for OFAT experimenters. By using two-level

factorial designs you can screen out these nuggets of information and perhaps accomplish a real breakthrough.

Plugging in the recommended settings in coded form gives a predicted outcome. (Let's choose the high level for A).

$Y^{1/2} = 11.15 - 0.76(+1) + 2.04(-1) - 1.91(+1) + 0.06(+1) + 1.39(-1)$
$+ 0.74(+1)(+1) - 0.95(-1)(+1) + 1.02(-1)(+1)$
$= 5.78$ \hfill (7)

Then to get the response back to the original units of measure, the transformation must be reversed.

Predicted PU particle size = 5.78^3 = 33 nanometers

This compares well with the observed result of 40 for experiment number 14, which happened to be run at the recommended settings.

What's in it for you

The case study on polyurethane illustrates how two-level factorials can be applied to a chemical process with many variables. The design of experiments uncovered several interactions which led to a breakthrough product improvement. The experimenters found that one of the factors, catalyst concentration (A), could be eliminated from further consideration. In part two of this series on DOE, we will take a look at an in-depth optimization on three of the factors critical to the waterborne polyurethane system. The article shows how to set up and analyze a response surface method (RSM) design. Figure 8 shows an example of a three-dimensional response surface graph with contours projected underneath.

If you equip yourself with the basic tools of statistical DOE, you will be in a position to make the most of opportunities such as those presented at the outset of this article. Your reputation will be enhanced and the competitive position of your company advanced.

Nomenclature

k	=	Number of factors
p	=	Number of model parameters, including intercept
PU	=	Polyurethane
X	=	Factor
Y	=	Response
Greek Letters		
α	=	Exponent for power law
μ	=	True mean
σ	=	True standard deviation

Literature Cited

(1)   Box, G.E.P., Hunter, W.G., and Hunter, J.S., *Statistics for Experimenters*, John Wiley & Sons, Inc, New York, 1978.

(2)   Montgomery, D.C., *Design and Analysis of Experiments*, 3rd ed., John Wiley & Sons, Inc, New York, 1991.

(3)   Yang, C.H., Lin, S.M. and Wen, T.C., "Application of Statistical Experimental Strategies to the Process Optimization of Waterborne Urethane," *Polymer Engineering and Science, Vol. 35, No. 8* (April 1995).

(4)   Myers, R.H., Montgomery, D.C., *Response Surface Methodology*, John Wiley & Sons, Inc, New York, 1995.

(5)   "Mathematics, Statistics" section, *1995 CEP Software Directory*, pages 27-30 (supplement to December 1994 issue of *Chemical Engineering Progress*).

(6)   Helseth, T.J., et al, *Design-Ease*, Version 3 for Windows or Macintosh, Stat-Ease, Inc, Minneapolis, 1994 ($395).

Table 1. Factors and Levels for Polyurethane (PU) Experiment

Factor	Description	Low Level	High Level
		–	**+**
A (X_1)	Catalyst (ppm)	0	150
B (X_2)	Acetone/PU (ratio)	2.8	4.5
C (X_3)	Phase-inversion temperature (ºC)	30	50
D (X_4)	Agitation rate (rpm)	350	1500
E (X_5)	Water-addition rate (mL/min)	2.0	4.0

Table 2. High Resolution Two-Level Fractional Factorials*

Factors	# Runs Full 2 Level Design	Fraction Applied	# Runs Fractional 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^{8-3})
10	1024 (2^{10})	1/16 (2^{-4})	64 (2^{10-4})
11	2048 (2^{11})	1/32 (2^{-5})	64 (2^{11-5})

*These designs give clean estimates of all main effects and all, or nearly all, two-factor interactions. For details on construction and properties of these designs see Box, Hunter and Hunter[1] Chapter 12.

Table 3. 2^{5-1} Design Matrix and Data for Polyurethane Case

Standard Order	A	B	C	D	E (=ABCD)	Particle Size (nanometers)	Square root Particle Size
1	−1	−1	−1	−1	+1	196	14.00
2	+1	−1	−1	−1	−1	70	8.37
3	−1	+1	−1	−1	−1	234	15.30
4	+1	+1	−1	−1	+1	257	16.03
5	−1	−1	+1	−1	−1	75	8.66
6	+1	−1	+1	−1	+1	86	9.27
7	−1	+1	+1	−1	+1	101	10.04
8	+1	+1	+1	−1	−1	50	7.07
9	−1	−1	−1	+1	−1	85	9.21
10	+1	−1	−1	+1	+1	76	8.72
11	−1	+1	−1	+1	+1	389	19.72
12	+1	+1	−1	+1	−1	173	13.15
13	−1	−1	+1	+1	+1	70	8.36
14	+1	−1	+1	+1	−1	40	6.32
15	−1	+1	+1	+1	−1	100	10.00
16	+1	+1	+1	+1	+1	202	14.21

Table 4. Estimates of Effects for Polyurethane Case

	Alias	Effect (Square root scale)
A:Catalyst conc*	BCDE	-1.52
B:Acetone/PU*	ACDE	4.08
C:PI Temp*	ABDE	-3.82
D:Agitation	ABCE	0.12
E:Water add*	ABCD	2.79
AB	CDE	0.37
AC*	BDE	1.47
AD	BCE	0.29
AE	BCD	0.54
BC*	ADE	-1.90
BD*	ACE	2.04
BE	ACD	0.84
CD	ABE	0.84
CE	ABD	-0.32
DE	ABC	0.29

*Significant effects based on inspection of normal plot. Statistical analysis revealed that each of these were significant at the 5% level (>95% confidence).

Table 5. Recommended Settings for Minimal Particle Size

Factor	Best Setting	Coded Level
A:Catalyst conc	Pick a level*	−1 to +1
B:Acetone/PU	Low	−1
C:PI Temp	High	+1
D:Agitation	High	+1
E:Water add	Low	−1

*(Factor A not statistically significant factor when factor C set at high level)
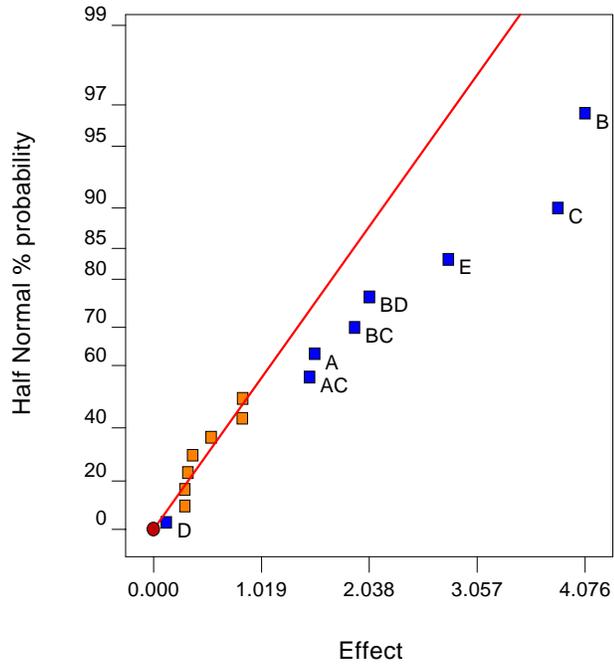
Figure 1. Half-Normal Plot of Effects (after square-root transformation)
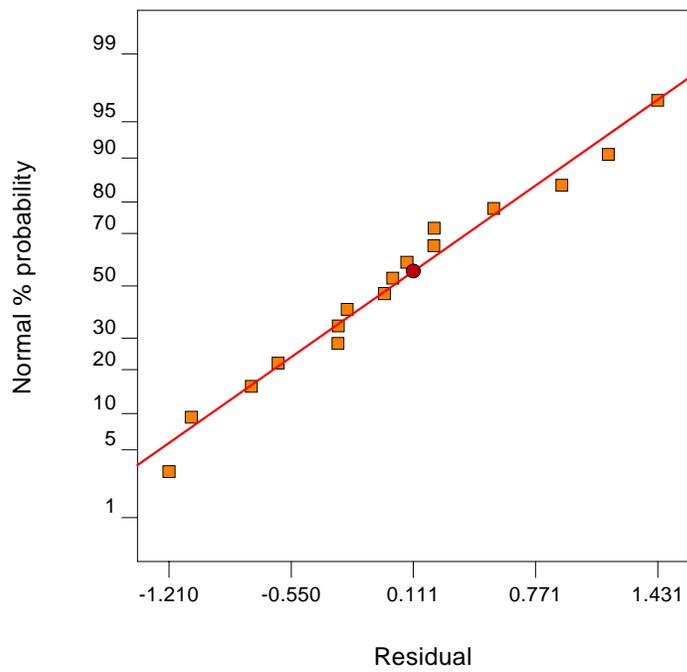
Figure 2. Normal Plot of Residuals
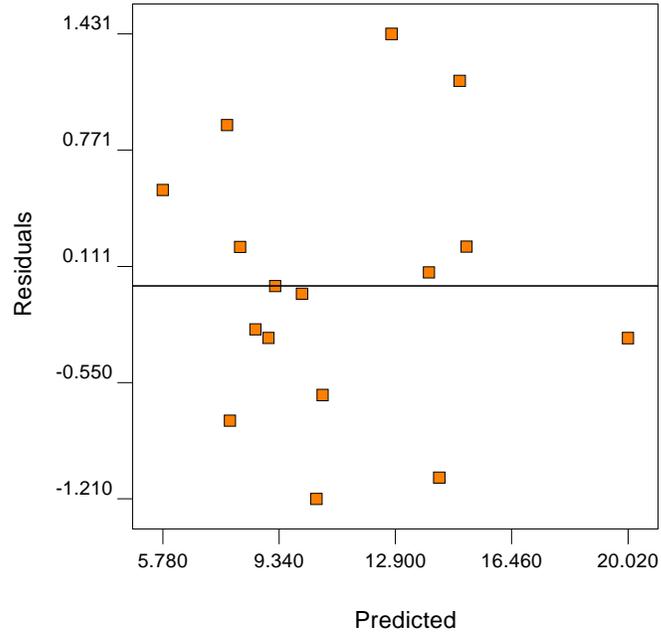
Figure 3. Residuals versus Predicted Level of Response
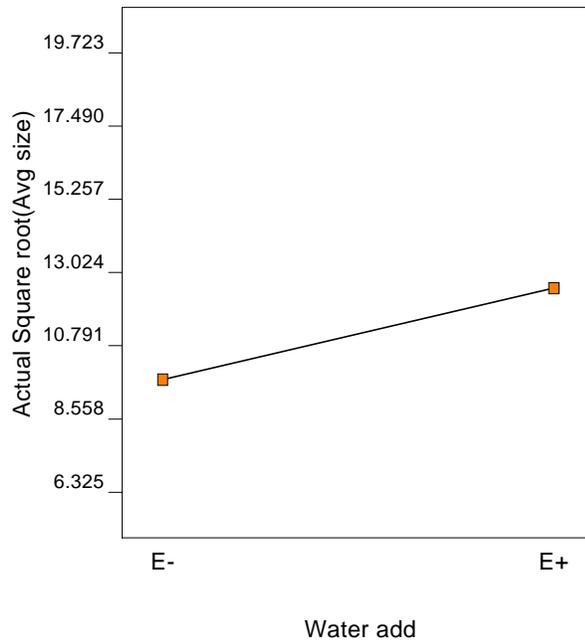
Figure 4. Main Effect Plot for Factor E

Interaction of A:DBTDL conc and C:PI Temp
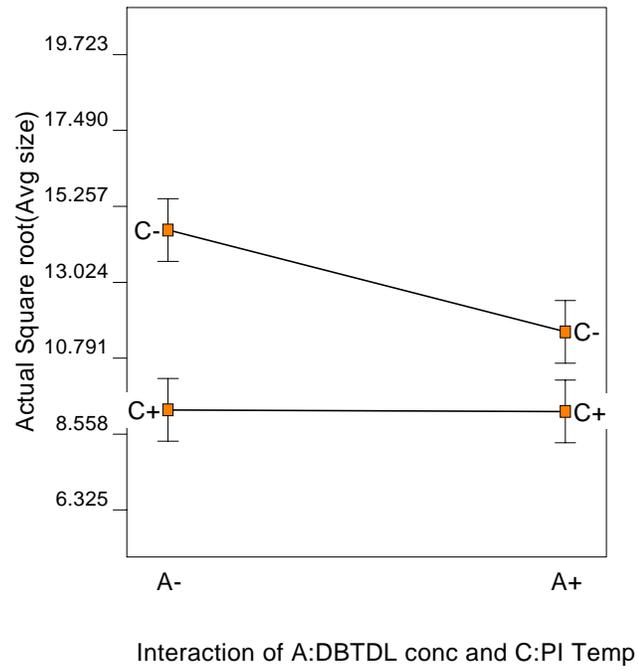
Figure 5. Interpretation Plot for Interaction AC

Interaction of B:Acetone/PU and C:PI Temp
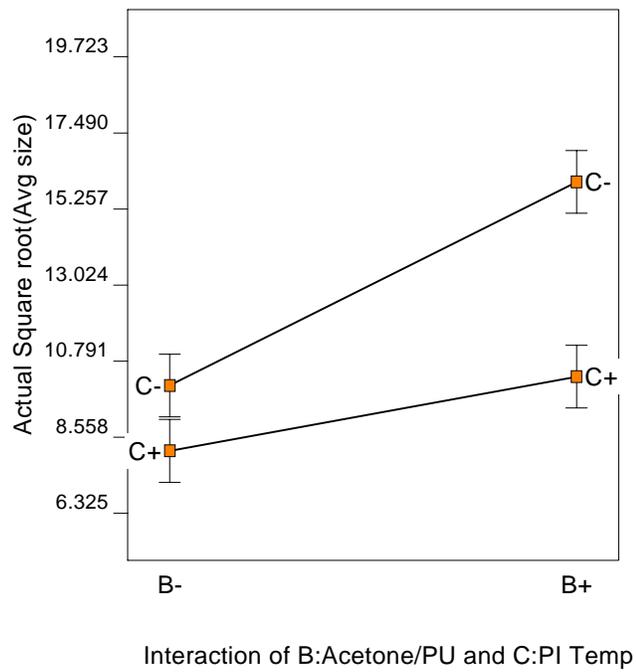
Figure 6. Interpretation Plot for Interaction BC

Interaction of B:Acetone/PU and D:Agitation
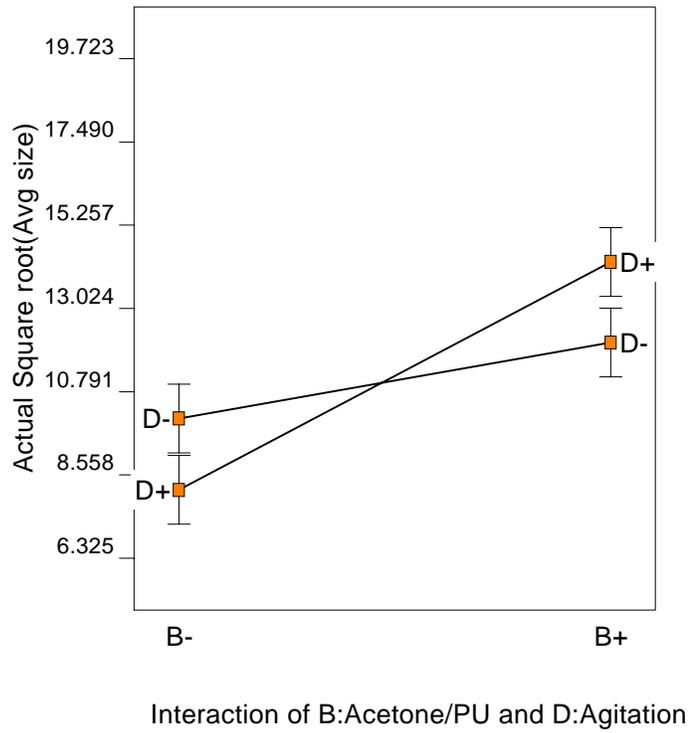
Figure 7. Interpretation Plot for Interaction BD

Water Addition

Acetone/PU
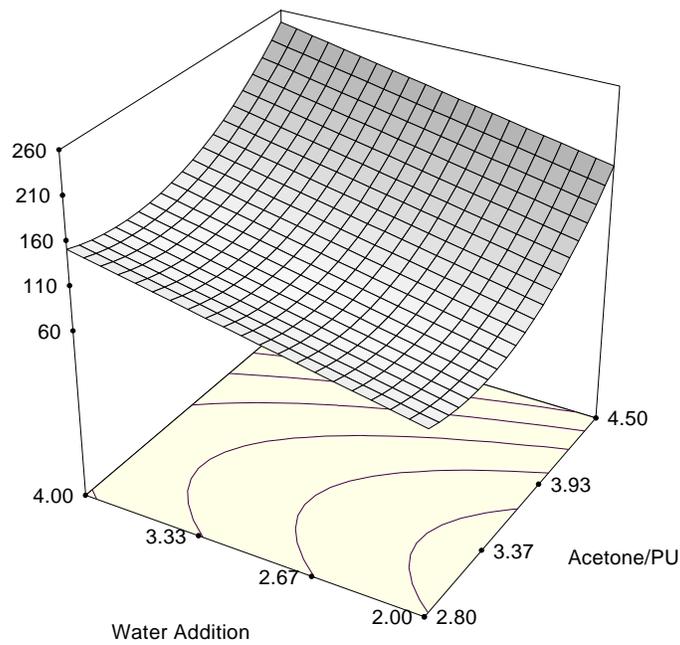
Figure 8. Hypothetical Response Surface Graph of Polyurethane Particle Size