Application of Statistics for process/product Development

Joshy, C.G.
Scientist, Fish Processing Division
ICAR- Central Institute of Fisheries technology, Cochin
cgjoshy@gmail.com

Every product/Process is designed with an intent purpose to solve/modify an existing problem. In most cases, any product/process development involves tests or experiments, since the product/Process is not well understood, and the desired response can't be guaranteed. Experimental design and analysis has been used to improve the performance of product/Process given the inherent noise in the various responses of interest. In development of new product/Process, research & development groups conduct experiments, develop models and finally optimize the responses related to the performance of the new product/process being developed.

Response surface methodology (RSM) is such a collection of statistical and mathematical techniques useful for developing, improving and optimizing processes. It also has important applications in the design, development and formulation of new products, as well as in the improvement of existing product designs. The most extensive applications of RSM are in the industrial world, particularly in situations where several input variables potentially influence some performance measure or quality characteristic of the product or process.

In general, suppose that the scientist or engineer is concerned with a product, process or system involving a response Y that depends on controllable input factors x_1, x_2,------x_p .The relationship between Y and the x's is defined as

$$Y = f(x_1, x_2, \cdots\cdots\cdots, x_p) + \varepsilon$$

Where the form of the true response function *f* is unknown and perhaps very complicated, and ε is a term that represents other sources of variability not explained or accounted by *f*. Thus ε includes effects such as measurement error on the response, other sources of variation that are inherent in the process or system, the effect of other variables, and so on. We treat ε as a statistical error term with mean zero and constant variance i.e. $\varepsilon \sim N(0, \sigma^2)$, then

$$E(Y) = E[f(x_1, x_2, \cdots\cdots\cdots, x_p)] + E(\varepsilon)$$

$$= f(x_1, x_2, \cdots\cdots\cdots, x_p)$$

Because the form of the true response function f is unknown, we must approximate it. In fact, successful use of RSM is critically dependent upon the experimenter's ability to develop a suitable approximation for *f*. Usually, a low order polynomial in some relatively small region of the independent variable space is appropriate. In many cases a first order or a second order model is used. The major objectives and applications of RSM are

1. To determine and quantify the relationship between response variables and settings of a group of experimental factors (independent variables) i.e. Mapping a response surface over a particular region of interest
2. To find the settings of experimental factors that produces the best value or the best set of values of the response variables i.e. Optimization of the responses

The major steps involved in RSM to improve an existing process/product or formulation of new product are

1. Formulation of experimental design in terms of independent variables
2. Formulation of hypothesis
3. Execution of experiments and generation of experimental data
4. Development of empirical model to predict the response variables in terms of independent variables
5. Model adequacy checking and testing of hypothesis
6. Optimization of response variables in terms of independent variables

Formulation of Experimental Design

Factorial designs are widely used in experiments involving several factors (independent variables) to investigate the main and interaction effects of the factors on response variables. The factorial designs can be classified into two groups viz: symmetrical and asymmetrical factorial experiments. A good response surface design should possess the properties *viz.*, detectability of lack of fit, the ability to sequentially build up designs of increasing order and the use of a relatively modest, if not minimum, number of design points. Examples on some experimental situations, where response surface methodology can be usefully employed are

Example 1: To Optimize the high pressure process parameters viz: pressure, ramp rate and holding time and to see its effect on high pressure treated Indian white prawn. The levels of various factors are

	Factors	**Levels**
1.	Pressure (MPa)	150, 250, 350
2.	Ramp Rate	300, 400, 500
3.	Holding Time(Min)	5, 10, 15

Example 2: For value addition to the agriculture produce, food-processing experiments are being conducted. In these experiments, the major objective of the experimenter is to obtain the optimum combination of levels of several factors that are required for the product. To be specific, suppose that an experiment related to osmotic dehydration of the banana slices is to be conducted to obtain the optimum combination of levels of concentration of sugar solution, solution to sample ratio and temperature of osmosis. The levels of the various factors are the following

	Factors	Levels
1.	Concentration of sugar solution	40%, 50%, 60%, 70% and 80%
2.	Solution to sample ratio	1:1, 3:1, 5:1, 7:1and 9:1
3.	Temperature of osmosis	25^0C, 35^0C, 45^0C, 55^0C and 65^0C

In this situation, response surface designs for 3 factors each at five equispaced levels can be used.

In general response surface methodology is useful for all the factorial experiments in agricultural experimental programme that are under taken so as to determine the level at which each of these factors must be set in order to optimize the response in some sense and factors are quantitative in nature.

Examples of experimental design setup for RSM

1. All the factorial experiments where the factors are quantitatively measured
2. Central Composite Design
3. Box-Behnken Design
4. Simplex lattice mixture design
5. Simplex centroid mixture design
6. D-optimal design

Development of Empirical Models

In practice the mathematical form of '*f*' discussed in the introduction is not known; we, therefore, often approximate it, within the experimental region, by a polynomial of suitable degree in variables x_{iu} (independent variables). The adequacy of the fitted polynomial is tested through the usual analysis of variance. Polynomials which adequately represent the true input-response relationship are called **Response Surfaces** and the designs that allow the fitting of response surfaces and provide a measure for testing their adequacy are called **response surface designs**. If the function '*f*' is of degree one in x_{iu}'s *i.e.* the response can be represented as

$$y_u = \beta_0 + \beta_1 x_{1u} + \beta_2 x_{2u} + \ldots + \beta_v x_{vu} + e_u$$

And we call it a first-order response surface in $x_1, x_2, \ldots, x_v$.

The second-order (quadratic) response surface can be represented as

$$y_u = \beta_0 + \sum_{i=1}^{v} \beta_i x_{iu} + \sum_{i=1}^{v} \beta_{ii} x_{iu}^2 + \sum_{i=1}^{v-1} \sum_{i'=i+1}^{v} \beta_{ii'} x_{iu} x_{i'u} + e_u$$

This functional form has many applications in most of the agricultural experiments

The analysis of variance table for a second order response surface design is given below.

Analysis of variance for second order response surface

Source	d.f.	S.S.
Due to regression coefficients	$2v+\binom{v}{2}$	$\hat{b}_0\sum_{u=1}^{N} y_u+\sum_i \hat{b}_i\left(\sum_{u=1}^{N} x_{iu}y_u\right)+\sum_i \hat{b}_{ii}\left(\sum_{u=1}^{N} x_{iu}^2 y_u\right)$ $+\sum\sum_{i\neq i'} \hat{b}_{ii'}\left(\sum_{u=1}^{N} x_{iu}x_{i'u}y_u\right)-CF$
Error	$N-2v-\binom{v}{2}-1$	$By\ subtraction = SSE$
Total	$N-1$	$\sum_{u=1}^{N} y_u^2 - CF$

In the above table CF = correction factor = $\frac{(Grand\ Total)^2}{N}$. For testing the lack of fit the sum of squares is obtained using (2.16) and then sum of squares is obtained by subtracting the sum of squares due to pure error from sum of squares due to error. The sum of squares due to lack of fit and sum of squares due to pure error are based on $N'-2v-\binom{v}{2}-1$ and $N-N'$ degrees of freedom respectively.

The lack of fit is tested using the statistic $F=\frac{SS_{LOF}/(N'-p)}{SS_{PE}/(N-N')}$

where N is the total number of observations, N' is the number of distinct treatments and p is the number of terms included in the model. SS_{PE} (sum of squares due to pure error) has been calculated in the following manner: denote the l^{th} observation at the u^{th} design point by y_{lu}, where $l=1,\ldots,r_u(\geq 1),\ u=1,\ldots,N'$. Define $\bar{y}_u$ to be average of r_u observations at the u^{th} design point. Then, the sum of squares for pure error is

$$SS_{PE}=\sum_{u=1}^{N'}\sum_{l=1}^{r_u}(y_{lu}-\bar{y}_u)^2 \qquad (2.16)$$

Then sum of squares due to lack of fit (SS_{LOF}) = sum of squares due to error - SS_{PE}

It is suggested that in the experiments conducted to find an optimum combination of levels of several quantitative input factors, at least one level of each of the factors should be higher than the expected optimum. It is also suggested that the optimum combination should be determined from response surface fitting rather than response curve fitting, if the experiment involves two or more than two factors.

Optimization of Response

The result of model-building procedure is an equation. Once the model is developed, the next stage is to optimize the process. Different type of optimization methods are

1. Method of steepest ascent/descent
2. Method of graphical evaluation of response surface plot
3. Method of desirability function analysis
4. Method of genetic algorithm

Method of steepest ascent/descent

The experimental design, model building procedure and sequential experimentation that are used in searching for a region of improved response constitute the method of steepest ascent. The method of steepest ascent contains the following steps to optimize the system in the initial region of $x_1,x_2,\text{-------},x_p$

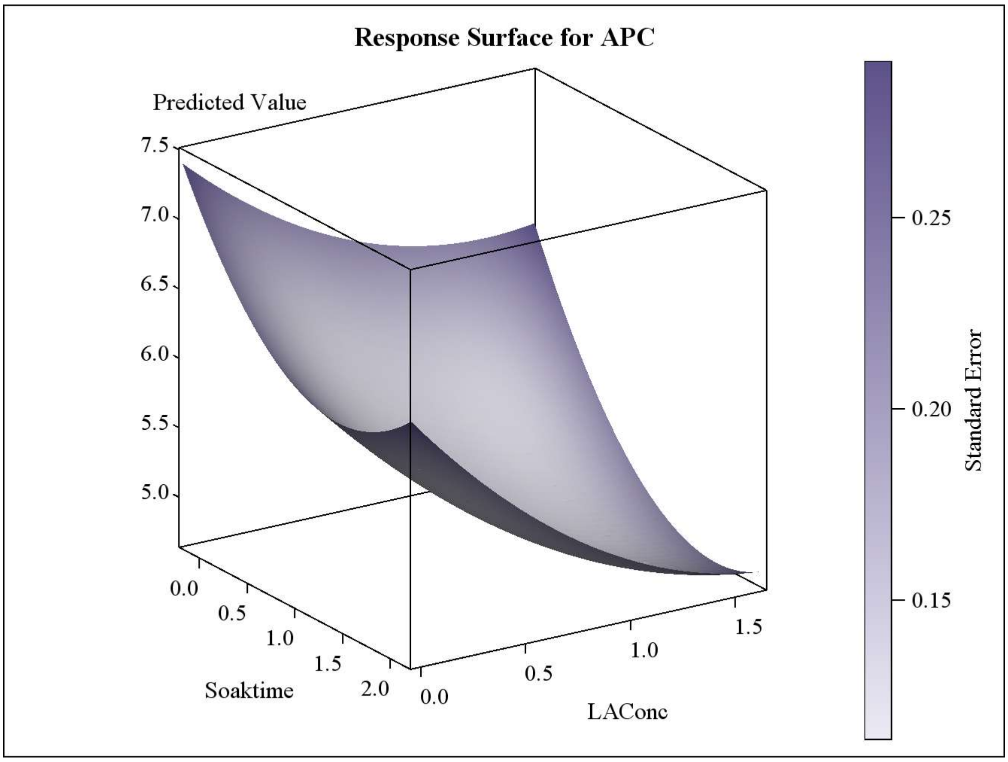
1. Fit a first order model (a plane or hyperplane) using either factorial or response surface design
2. Compute a path of steepest ascent if maximizing the response or steepest descent if minimizing the response. The path of steepest ascent is computed to get maximum increase in the response and path of steepest descent is computed to get maximum decrease in the response.
3. Conduct the experimental runs along the path. That is do either single response or replicated runs, and observe the response value and the result will normally show improving values of the response. At some region along the path the improvement will decline and eventually disappear.
4. At point where an approximation of the maximum (or minimum) response is located on the path, a base for second experiment is chosen.
5. A second experiment is conducted, and another first order model is fitted to the data. A test of lack of fit is made. If the lack of fit is not significant, a second path based on the new model is computed. This is often called a mid-course correction. It is quite likely that the improvement will not be as strong as that enjoyed in the first path. After improvement is diminished, one typically has a base for conducting a more elaborate experiment and a more sophisticated process optimization.

Method of graphical evaluation of response surface plot

Evaluation of three dimensional response surface plot will help to identify the optimum values of input factors (say x_1 and x_2) that maximizes (minimizes) the predicted values of the response. Here, x_1 takes on X axis, x_2 takes on Y axis and predicted values of response variables takes on Z axis.

Example: An experiment was conducted to optimize the Lactic acid concentration and soaking time to minimize the APC in Tuna chunk. The second order response surface plot is given below.

From the graph, it can be inferred that the minimum APC was observed at higher levels of Lactic acid concentration and soaking time.



Method of desirability function analysis

The desirability function approach is one of the most widely used methods in industry for the optimization of multiple response processes. It is based on the idea that the "quality" of a product or process that has multiple quality characteristics, with one of them outside of some "desired" limits, is completely unacceptable. The method finds operating conditions x that provide the "most desirable" response values. For each response $\boldsymbol{Y_i(x)}$, a desirability function $\boldsymbol{d_i(Y_i)}$ assigns numbers between 0 and 1 to the possible values of $\boldsymbol{Y_i}$, with $\boldsymbol{d_i(Y_i) = 0}$ representing a completely undesirable value of $\boldsymbol{Y_i}$ and $\boldsymbol{d_i(Y_i) = 1}$ representing a completely desirable or ideal response value.
A desirability function is a useful approach to optimize for simultaneous optimization of multiple responses. The general approach is to first convert each response y_i into an individual desirability function d_i that varies over the range $0 \leq d_i \leq 1$

Where if the response y_i is at its goal or target, then d_i=1, and if the response outside an acceptable region d_i=0. Then the design variables (input variables) are chosen to maximize the overall desirability$D = (d_1 d_2 \cdots\cdots\cdots d_m)^{1/m}$, where there are m responses

If the objective or target T for the response y is a maximum value,

$$d = \begin{cases} 0 & y < L \\ \left(\frac{y-L}{T-L}\right)^r & L \le y \le T \\ 1 & y > T \end{cases}$$

When the weight r=1, the desirability function is linear. Choosing r>1 places more emphasis on being close to the target value, and choosing 0< r <1 makes this less important. If the target for the response is a minimum value,

$$d = \begin{cases} 0 & y < T \\ \left(\frac{U-y}{U-T}\right)^r & T \le y \le U \\ 1 & y > U \end{cases}$$

The two sided desirability function when the target is located between lower (L) and upper (U) limits is defined

$$d = \begin{cases} 0 & y < L \\ \left(\frac{y-L}{T-L}\right)^{r_1} & L \le y \le T \\ \left(\frac{U-y}{U-T}\right)^{r_2} & T \le y \le U \\ 0 & y > U \end{cases}$$

Method of genetic algorithm

GA is an optimization technique based on the principles of genetics and natural selection. Genetic operators, such as selection, crossover, and mutation, are applied to repressor settings while searching for the optimum. In a GA, a search point, a setting in the search space, is coded into a string which is analogous to a chromosome in biological systems. The string/chromosome is composed of characters which are analogous to genes. In a response surface application, the chromosome corresponds to a particular setting of k factors (or regressors), denoted by $x = [x_1, x_2, ..., x_k]'$, in the design space and i^{th} gene in the chromosome corresponds to a x_i, the value of the i^{th} regressor.

Suggested Readings

- Box, G.E.P. and Draper, N.R. (1987). *Empirical model building and response surfaces*. New York, Wiley.
- Myers, R.H. and Montgomery, D.C. (1995). *Response Surfaces Methodology: Process and product optimization using designated experiments.* John Wiley and Sons.