Teaching Response Surface Methodology by Simulation

R Hugh Morton - Palmerston North, New Zealand

1. Introduction

Any tertiary level statistics course in the design and analysis of industrial experiments will include aspects of experimental design, hypothesis testing, and regression analysis. Nowhere is the interplay between these three factors more evident than in response surface methodology (RSM). In the technological context this methodology is applied to study the measured yield or output of a system as it varies in response to the changing levels of one or more physical input variables. The experimental design aspect deals with the choice of suitable variables and their various levels, and regression analysis enables a mathematical form to be fitted to the observations to "model" the varying yield. Hypothesis testing is relevant to both these processes.

2. The course

The general purpose of RSM is usually one or both of the following: to gain some insight into the operation of a system and/or to seek some optimum settings for the variables involved. Implicit therefore are the requirements that the mathematical form of the model be reasonably simple and permit optimisation, and that the regression experimental design be both large enough to provide estimates of all relevant model parameters, and small enough to be efficient in terms of cost and resource utilisation. Such criteria normally dictate the use of first and second order polynomial regression models, based on factorial (or fractional factorial) designs of the 2^n (or 2^{n-k}) series, and on central composite designs.

A module to teach RSM forms a substantive part of a half course on the design and analysis of industrial experiments for final year undergraduate BTech students at Massey University. Minimal prerequisites include a first level introductory statistics course, a second level course in statistical inference, and a third level course on

regression and experimental statistics. The first part of the course covers the 2^n and 3^m factorial experiment series and their fractions, together with screening designs and the foldover theorem. RSM follows, and the course concludes with several case studies from industry.

Since RSM itself is well covered in the literature (Hill and Hunter, 1966; Mead and Pike, 1975; Morton, 1983 and Myers et al., 1989), this paper focuses on the use of computer simulation as a teaching aid within the course. The computer program simulates a simple stochastic thermodynamic reaction, fits a response surface, correlates a path of steepest ascent, and plots contour diagrams; all consequent upon a suitably chosen design of experimental observation points.

The program author is Dr Ian Boag of the Department of Production Technology, Massey University.

3. The simulation and analysis program

The advantages of using a computerised simulation procedure for the teaching of (sometimes large) numbers of students are self-evident. It saves time, is simpler, utilises less resources in manpower, machinery and money, and should provide a better understanding of what is being taught, than a real experiment. Lack of reality is of course its main disadvantage. The students appreciate its ease of use and understanding, the logical progression through the methodology, and its graphical output.

The simulated thermodynamic reaction involves three input variables; X_1, the temperature of the reaction in ^{o}C, usually $30 \leq X_1 \leq 70$; X_2, the duration of the reaction in hours, usually $1 \leq X_2 \leq 2$; and X_3, the concentration of reagent A, usually $0.5 < X_3 < 1$; all three of which are under control of the student experimenter. Two response yields of interest are obtained as output; Y_1, the % conversion of reagent A; and Y_2, the % yield of the desirable product B. (A third response, Y_3, the production rate, is simply the product $X_3Y_2/100$.) The objective is to maximise Y_2, with Y_1 as close to 100% as is possible.

Students are required, as an assignment, to:

(i) design a suitable RSM experiment;
(ii) simulate the experiment and record the results;
(iii) analyse these results by fitting response surfaces;
(iv) plot these surfaces, singly and in overlay;
(v) comment, given that optimisation is the objective;
(vi) write up and submit their work for assessment.

A suitable starting point is provided by the "plant operators". Students are aware that iterations through the first three or four stages are highly likely, but also that efficiency penalties will be levied if it takes too many experimental runs to locate an optimum.

4. Program subroutines

The first subroutine, "streac", opens a spreadsheet into which the experimental

design points in X_1, X_2 and X_3 can be inserted and into which the simulated reaction will insert outputs Y_1, Y_2 and Y_3. The particular design may be for example, a 2^2F in X_1 and X_2 with three added centre points, X_3 being fixed at a concentration of 0.8.

A second subroutine, "ana22", opens another spreadsheet to accept the simulated output from the 2^2F and fit a first order polynomial surface in two variables. This surface can be examined, and the residuals at the centre points used to gauge the adequacy of the first order fit.

A third subroutine, "dosteep", opens another spreadsheet to accept the coefficients of the fitted first order model and calculates its steepest ascent path. A series of points along the path will be presented, together with their associated simulated output yields. A new region of experimentation can then be determined in which it may be likely that, due to curvature of the surface, a second order central composite design may be necessary to achieve an acceptable second order polynomial fit.

This central composite design is simulated in two orthogonal blocks using the "streac" subroutine twice more, the first block of which would be a 2^2F such as above. The second block comprises the star points, and additional centre points, of the design.

A fourth subroutine, "anacc", opens a further spreadsheet to accept both blocks of simulated output and calculate a fitted second order polynomial.

Finally, a fifth subroutine, "contua", opens another spreadsheet to accept the coefficients of the fitted (first or) second order polynomial and prepares and plots its contour diagram. These plots, such as the second order surface in Figure 1, can be examined for shape, location of optimum, etc., and can be overlaid.

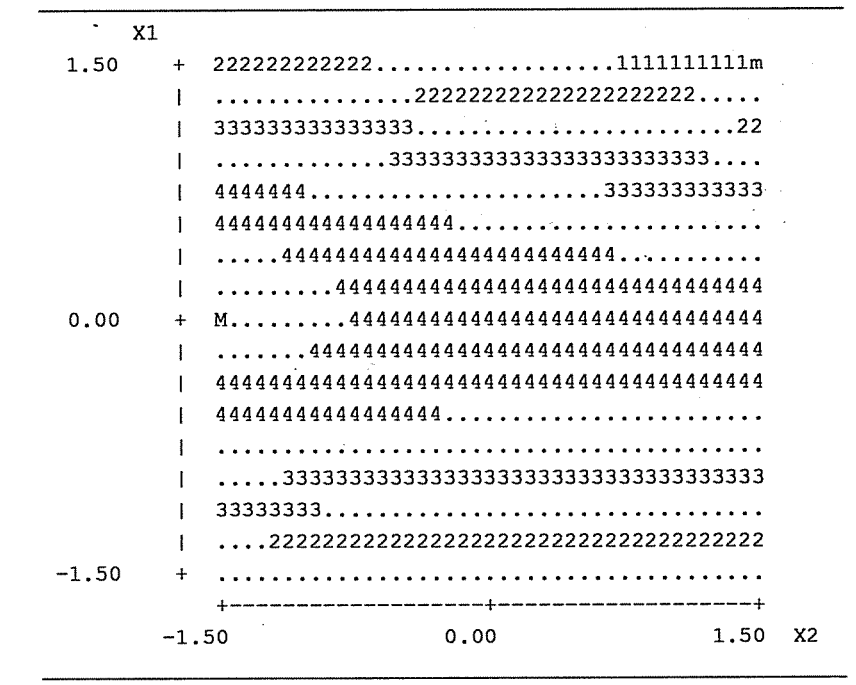

FIGURE 1

Figure 1 reveals the characteristics of the simulated system in respect of output response Y_2, the yield of product B for fixed concentrations of reagent A. It is a gently sloping ridge, aligned somewhat negative diagonally in the temperature/time plane. Optimal product yields in the range 90-94% can be obtained with low reagent concentration (0.5), temperature in the range 40-45°C, and duration times of 1.1 to 1.3 hours. The % conversion of reagent A (Y_1) is obtainable at 100% in all such cases. Higher temperature/shorter time (and vice versa) combinations are also feasible. Increasing the reagent concentration tends to lower the yields and shifts the ridge towards generally higher temperature/time combinations, a demonstrably less efficient area of operation. The stochastic nature of the system is such that the slope direction of the crest of the ridge and/or its diagonal alignment in the temperature/time plane will almost certainly vary from student to student.

5. Conclusion

This simulation program has been used successfully for several years to teach RSM as part of a course on the design and analysis of industrial experiments for final-year undergraduates. It has been well received by students and has also been utilised in off-campus workshops in industry.

References

Hill, W J and Hunter, W G (1966) A review of response surface methodology : a literature survey. *Technometrics* **8**, 571-590.

Mead, R and Pike, D J (1975) A review of response surface methodology from a biometric viewpoint. *Biometrics* **31**, 803-851.

Morton, R H (1983) Response surface methodology. *The Mathematical Scientist* **8**, 31-52.

Myers, R H, Khuri, A I and Carter, W H (1989) Response surface methodology : 1966-1988. *Technometrics* **31**, 137-157.