Response Surface Methodology in Agriculture

Linda M. Haines

School of Mathematics, Statistics and Information Systems, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa

Response surface methodology (RSM) is concerned with the modelling of one or more responses to the settings of several explanatory variables. The nature of the function relating the responses to the variables is assumed to be unknown and the function or surface is modelled empirically using a first- or a second-order polynomial model. The broad aims of RSM are to investigate the nature of the response surface over a region of interest and to identify operating conditions associated with maximum or minimum responses. RSM is generally conducted in three phases, as emphasized in Myers and Montgomery (2002). Phase 0 involves the screening of explanatory variables to identify those which have a significant effect on the responses, phase 1 is concerned with the location of optimum operating conditions by conducting a sequence of suitable experiments and phase 2 involves the fitting of an appropriate empirical model, usually a second-order polynomial model, in order to examine the nature of the response surface in the vicinity of the optimum. The fundamentals of RSM are set out in the seminal papers of Box and Wilson (1951) and Box and Draper (1959) and in the books by Box and Draper (1987), Khuri and Cornell (1987) and Myers and Montgomery (2002). Further developments are drawn together in three key review articles, namely those of Hill and Hunter (1966), Myers, Khuri and Carter (1989) and Myers, Montgomery, Vining, Borror and Kowalski (2004). It is clear from these articles that research into RSM within academia continues to flourish and that the associated techniques are used extensively in industry.

It has long been perceived that the RSM approach, having as it does an intrinsically sequential nature, is not particularly appropriate in the agricultural setting. Specifically, experiments in agriculture are invariably constrained to be long-term and have as their primary focus the identification of key factors of interest. Mead and Pike (1975) review the role of RSM in agriculture but, in so doing, emphasize the use of nonlinear models to accommodate biological data rather than of the empirical models traditionally used in RSM. Edmondson (1991) provides an interesting application of RSM to greenhouse experiments and, in addition, presents some valuable insights into the use of RSM within an agricultural as opposed to an industrial setting. In addition Khuri and Cornell (1987, Section 5.6) analyze an experiment on snap bean yield conducted using a central composite design. However such papers are rare and there is surprisingly little interest in RSM in agricultural applications within the mainstream statistical literature. On balance it is clear that while certain approaches within RSM are not appropriate for an agricultural setting, there is nevertheless a wealth of knowledge embedded within the broad field of RSM which can be drawn upon with advantage by agriculturalists.

In the first part of this talk the main features of response surface modelling in a phase 2 experiment are introduced by means of an agricultural example. Specifically the dependence of the yield of sugar cane on varying amounts of the nutrients, nitrogen, phosphorus and potassium, can be modelled empirically using a second-order polynomial model (Mapham, 1975) and this scenario is considered here. A design appropriate to the model setting was formulated and developed in a series of conference presentations and notes by Rojas (1962, 1971, 1973) and is termed the San Cristobal design after the sugar milling town of that name in Mexico. The design, a variant of the central composite design widely used in RSM, comprises a 2^k factorial design together with centre and axial points and is structured to accommodate a control with the nutrient amounts set to zero. Some results for comparing the performance of the San Cristobal design with that of other designs such as the 3^k factorials are presented in the thesis of Mapham (1975) and are summarized in the paper by Dicks (1999). In the present study the San Cristobal design is evaluated more broadly in terms of three key areas of interest. First the interpretation and the variability of the parameter estimates are discussed and the related concept of D-optimality is introduced. Second criteria for evaluating the San Cristobal design based on the standardized prediction variance are appraised. Some attention is given to the "single number" criteria of G- and IV-optimality and in addition more flexible tools, such as variance dispersion graphs, are introduced. Third the robustness of the design towards model misspecification, the occurrence of outliers and error control is discussed.

The second part of this talk is devoted to an appraisal of the role of RSM in agricultural experimentation. The example presented above demonstrates that designs taken from the RSM paradigm can be used to good effect in a traditional agricultural setting and this point is further underscored by the work of Khuri and Cornell (1987) and of Edmondson (1991). In the present paper the advantages to be gained by using designs based on the central composite design, such as the San Cristobal design, are highlighted and attention is drawn to the fact that these designs require few points and that information on quantitative variables is used effectively. Furthermore many agricultural experiments involve responses to the explanatory variables which are binary or count in nature and which can thus be modelled within the generalized linear model framework. For example, a researcher may be interested in the potency of the combination of two insecticides and, specifically, in their synergistic or antagonistic action. Response surface techniques are eminently suitable to such situations but this area of application has only recently attracted attention in the RSM literature (Myers, 1999). In the present paper the role of RSM in the context of generalized linear models is reviewed and scope for further development discussed.

In summary, the field of RSM is well-researched and established within the industrial context and researchers in agriculture and related disciplines could well draw with advantage on the broad framework provided by this methodology in order to design and analyze their experiments.

References

Box G.E.P., Draper N.R. (1959), "A basis for the selection of a response surface design." *Journal of the American Statistical Association*, **54**, 622–654.

Box G.E.P., Draper N.R. (1987), *Empirical Model-Building and Response Surfaces.* Wiley, New York.

Box G.E.P., Wilson K.B. (1951), "On the experimental attainment of optimum conditions ." *Journal of the Royal Statistical Society, Series B*, **13**, 1–45.

Dicks H.M. (1999), "The use of the San Cristobal design in soil calibration studies." *Journal of Tropical Forest Resources*, **15**, 43–52.

Edmondson R.N. (1991), "Agricultural response surface experiments based on fourlevel factorial designs." *Biometrics*, **47**, 1435–1448.

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

Khuri A.I., Cornell J.A. (1987), *Response Surfaces.* Marcel Dekker, New York.

Mapham W.R. (1975), *Some Biometrical Aspects of Soil Calibration.* M.Sc. thesis, University of Natal, South Africa.

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

Myers R.H. (1999), "Response surface methodology - current status and future directions." (with discussion) *Journal of Quality Technology*, **31**, 30–44.

Myers R.H., Khuri A.I., Carter W.H. (1989), "Response surface methodology : 1966-1988." *Technometrics*, **15**, 301–317.

Myers R.H., Montgomery D.C. (2002), *Response Surface Methodology.* Wiley, New York.

Myers R.H., Montgomery D.C., Vining G.G., Borror C.M., Kowalski S.M. (2004), "Response surface methodology : a retrospective and literature survey." *Journal of Quality Technology*, **36**, 53–77.

Rojas B.A. (1962), "The San Cristobal design for fertilizer experiments." *Proceedings of the 11th Congress of the International Society for Sugar Cane Technology*, 197–203.

Rojas B.A. (1971), "The orthogonalized San Cristobal design." *Proceedings of the 14th Congress of the International Society for Sugar Cane Technology*, 1085–1093.

Rojas B.A. (1973), "The San Cristobal design - annotation." Unpublished note.