Thus if the dimension can be reduced, the design in the remaining dimensions is still reasonably good. The optimal designs depicted in Johnson, Moore and Ylvisaker (1988) do not tend to project uniformly.

We prefer the sequences of Faure (1982) to the Halton-Hammersley sequences. The Halton-Hammersley sequences are usually based on the first d prime numbers, whereas Faure uses the same prime number (the smallest prime $r \geq d$) on each axis. When $n = r^k$, the Faure sequences exercise each input variable in much the same way Latin hypercube designs do. Moreover for $k \geq 2$ they exercise pairs of input variables in that, for any given pair of inputs, one can partition their domain into r^2 squares and find r^{k-2} points in each square. Similarly there are equidistribution properties for three or more axes. The equidistribution properties of the Halton-Hammersley sequences are different for each marginal subcube, depending on the associated primes. We have found that with $n = r^2$ and $r = 5$ or 7 that the Faure sequences appear to lie on planes in three dimensions. This is alleviated by replacing each digit b in the base r representation of the Faure sequence by $\sigma(b)$ where σ is a permutation of $0, \ldots, r-1$. The permutation does not alter the equidistribution properties. One can inspect three-dimensional scatterplots to make sure that a given permutation is effective.

PARAMETER ESTIMATION

We would like to mention a quick way of estimating $\theta_1 \ldots, \theta_d$ in the covariance given by the authors' equation (9) with $p = 1$. When the function $Y(x)$ is nearly additive, we can estimate the main effects using scatterplot smoothers. This corresponds to the inner loop of the ACE algorithm in Breiman and Freidman (1985). Let g_j denote the estimate of the jth main effect. A very smooth $g_j(\cdot)$ is evidence that θ_j is small and a rough $g_j(\cdot)$ suggests that θ_j is large. The roughness may be assessed by $\mathscr{R}_j = \sum_{i=1}^m (g_j(i/m) - g_j((i-1)/m))^2$ where the domain of g_j has been rescaled to $[0, 1]$. The expected value of $\mathscr{R}_j$ may be expressed in terms of θ_1 through θ_d, for fixed σ. The d equations in d unknowns can be solved iteratively. The likelihood can be used to choose between the answers from several different values of m. This avoids a high dimensional search for $\theta_1, \ldots, \theta_d$. The first time we tried it, we got better parameter values (as measured by likelihood) than we had found by searching. Alternatively it suggests starting values for such a search.

ADDITIONAL REFERENCES

BECKER, R., CHAMBERS, J. and WILKS, A. (1988). *The New S Language.* Wadsworth and Brooks/Cole, Pacific Grove, Calif.

BREIMAN, L. and FRIEDMAN, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *J. Amer. Statist. Assoc.* **80** 580–598.

FAURE, H. (1982). Discrepance de suites associees a un systeme de numeration (en dimension s). *Acta Arith.* **41** 337–351.

SHARIFZADEH, S., KOEHLER, J., OWEN, A. and SHOTT, J. (1989). Using simulators to model transmitted variability in IC manufacturing. *IEEE Trans. Manuf. Sci.* To appear.

Comment

Anthony O'Hagan

The authors are to be congratulated on their lucid and wide-ranging review. Like others before, I have independently rediscovered many of the ideas and results presented here. I therefore sincerely hope that the greater prominence given to those ideas and results by this excellent paper will enable future researchers to start well beyond square one. I first have some comments concerning the derivation of the basic estimator (7), and I will then discuss the model and the practical implementation of the methods from my own experience.

Anthony O'Hagan is Senior Lecturer and Chair, Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom.

The authors mention three derivations of (7). In a classical framework, it is the MLE if the process $Z(\cdot)$ is Gaussian, and relaxing this assumption it is the BLUP, minimizing (2). Thirdly, it is the posterior mean of $Y(x)$ in a Bayesian analysis with a Gaussian $Z(\cdot)$ and a uniform prior on β. It is first worth pointing out that with a proper multivariate normal prior $\beta \sim N(b, B)$ and known σ^2 the posterior mean of $Y(x)$ has the same form as (7), but with $\hat{\beta}$ replaced by the posterior mean of β, i.e.,

$$\tilde{\beta} = (F'R^{-1}F + \sigma^2 B^{-1})(F'R^{-1}F\hat{\beta} + \sigma^2 B^{-1}b).$$

The interpretation of (7), as comprising the fitted regression model plus smoothed residuals, still holds.

We can also dispense with normality in the Bayesian framework, using a similar device to (2). The