Comment

A. Owen, J. Koehler and S. Sharifzadeh

We have been running computer experiments related to semiconductor process design and recently switched over to the paradigm described by the authors. We have found it to be more flexible than response surface methodology in handling deterministic responses.

The Bayesian approach suggests how to interpolate, extrapolate, assess uncertainty and construct designs. To what extent do the Bayesian answers make sense, if one does not hold the prior belief? The authors cite several works in which connections are drawn between well accepted interpolation methods and various priors and give an example in which the uncertainty assessment is accurate. It would be very interesting if the uncertainty assessments were reasonbly accurate for a large class of underlying functions. Have the authors investigated this point? We doubt that the Bayesian method will help in extrapolation (which we suspect should be avoided) and thus are worried that the optimal designs sometimes concentrate near the center of the design space.

Our main comments are directed at the design problem and at estimation of the parameters of the covariance model. Our applications have 5 to 10 input variables and a like number of outputs. The programs we use are fast enough to make it feasible to consider 50 or more runs.

Before addressing the design and estimation issues, we wish to point out that ideas from exploratory data analysis have a role to play in computer experimentation. The authors (with their coworkers) have plotted contours, trajectories and the additive main effects (mentioned in Section 6) of the response functions. We think their contributions are noteworthy and look forward to further developments. When there are many response variables, care should be taken in optimizing a functional of the responses without first considering the tradeoffs among competing goals. The approach taken in Sharifzadeh, Koehler, Owen and Shott (1989) is to evaluate the model functions at thousands of input points and to explore the resulting data set with interactive graphics, in this case S (Becker, Chambers and Wilks, 1988).

A. Owen is Assistant Professor of Statistics, J. Koehler is a Ph.D. candidate in Statistics, and S. Sharifzadeh is a Ph.D. candidate in Electrical Engineering, all at Stanford University. All three are affiliated with the Center for Integrated Systems. Their mailing address is: Department of Statistics, Stanford University, Stanford, California 94305.

DESIGN ISSUES

In the authors' Figure 1, the design points are all quite close to the center. We share the misgivings of the authors, suspecting that this leads to a robustness problem. Extrapolation by conditional expectation depends to a far greater degree on the covariance function used than does interpolation. Thus outside of the convex hull of the data, the predicted values will depend strongly on hard-to-verify properties of the model.

We have been using low discrepancy sequences, mentioned in Section 7.5, as designs. These designs are constructed so that the empirical measure of the design points is close in a Kolmogorov-Smirnov sense to the uniform measure on the cube. These should be good designs in the case of large θ, when estimation is difficult. Johnson, Moore and Ylvisaker (1988) characterize the optimal designs in the large θ limit. Minimizing the maximum distance from a point in the cube to a design point leads to their version of G optimality and maximizing the minimum distance between two sample points leads to their version of D optimality. Low discrepancy sequences (such as Halton-Hammersley sequences) tend to have small, but not minimal, maximum distances from points in the cube.

We have found that sometimes some of the θ_j appear quite small while others are large. That is a response variable is heavily dependent on a few of the d inputs and not very sensitive at all to the others. We may not know in advance which input variables are the important ones or, more commonly, each output variable may depend most strongly on a different small set of inputs. This opens up the possibility of reducing the dimension of the problem by considering the response as a function of the most important inputs, possibly with some noise due to the other inputs. For instance, in our first experiment the thickness of a layer of SiO_2 only depended on the oxidation temperature. Unfortunately, our design (an all-bias design) only used three distinct values of the temperature in 43 runs. If our experiment had had 43 nearly equispaced temperature values, the results would have been more informative.

Low discrepancy designs have the added benefit that when projected onto a cube defined by a subset of the original variables they are still nearly uniform.