COMMUN. STATIST.-THEORY METH., 20(2), 417-439 (1991)

ON CASCADING LATIN HYPERCUBE DESIGNS AND ADDITIVE MODELS FOR EXPERIMENTS

BY MARK S. HANDCOCK

IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, U.S.A.

Key words and phrases: experimental design; statistical inference; response surface; spatial statistics.

ABSTRACT

This paper is concerned with aspects of the design and analysis of computer experiments. It has been motivated by issues in the experimental design of integrated-circuits. Suppose we wish to model the behavior of a complex process as a function of several factors. An appealing approach is to model the response of the system as a stochastic process (Sacks, Welch, Mitchell and Wynn (1989)). Often much of the variation in the response can be accounted for by an additive function in each of the factors. In this paper we consider a promising additive stochastic model proposed by Stein (1989).

We introduce Cascading Latin hypercube designs as an efficient basis for statistical inference on the stochastic structure when the number of initial inputs is large. Comparisons are made to simple random designs, Faure (1982) designs and ordinary Latin hypercube designs. The results provide insight into the optimal design under such models.

1. Introduction

Many experiments that aim to improve the quality of products can be placed within the framework of understanding the relationship between a response $y(\cdot)$ and a d-vector of input factors x. For example, $y(x)$ might be the propagation delay of an experimental logic circuit dependent upon the physical design parameters. The experimenter can choose the design points so as to optimize, in some sense, over the response surface or better understand the influence of each of the input factors. There is an extensive and rich literature on this subject (Box, Hunter and Hunter (1978), Box and Draper (1987)).

Recent research has considered modeling the response $y(x)$ as a realization from a random field $Y(x)$. This approach integrates ideas from spatial statistics and response surface methodology. In the forefront of this approach is work on the design and analysis of computer experiments (cf., Currin, Mitchell, Morris and Ylvisaker (1988) and Sacks, Schiller and Welch (1989)). The paper of Sacks, Welch, Mitchell and Wynn (1989) and it's discussants, provide a introduction to the ideas.

This paper investigates aspects of the design and analysis of experiments for quality improvement within the framework of Sacks, Welch, Mitchell and Wynn (1989).

2. Modeling approach

Let $y(x)$ be a real-valued response to a d-vector of input factors $x \in R \subset \mathbb{R}^d$. We choose to model $y(x)$ as a realization from a random field $Y(x)$. Although $y(x)$ may be a deterministic function of x, this approach can be rationalized within a Bayesian framework (cf., Currin, Mitchell, Morris and Ylvisaker (1988)). From a pragmatic perspective it is essential to choose a class of random fields that is flexible enough to capture the behavior of $y(x)$. In order to focus on other modeling issues, we will assume $Y(x)$ is a Gaussian random field.

The experiment consists of choosing n sets of values of the design factors $x_1, \ldots, x_n$ according to a design criterion, and then determining the responses

to these set of factors: $\{y(x_1), \ldots, y(x_n)\}' = Y$. We address only single-step designs, perhaps considered as one stage of a sequential design.

A central issue in the random field approach to modeling response surfaces is the choice of the class of covariance structures for $Y(x)$. For a Gaussian random field we need only specify a mean function and covariance structure to specify the random field completely. Suppose $Y(x)$ has mean

$$E\{Y(x)\} = f'(x)\beta,$$

where $f(x) = \{f_1(x), \ldots, f_q(x)\}'$ is a known vector function, β is a vector of unknown regression coefficients. Furthermore, the covariance function is represented by

$$\operatorname{cov}\{Y(x), Y(w)\} = \alpha K_\theta(x, w) \qquad \text{for } x, w \in R$$

where $\alpha > 0$ is a parameter, $\theta \in \Theta$ is a $q \times 1$ vector of structural parameters and Θ is an open set in $\mathbb{R}^p$. The division is purely formal as θ may also determine aspects of scale.

2·1 The Matérn Class of Covariance Functions

In this section we describe a general class of covariance functions that can provide a sound foundation for the parametric modeling of Gaussian random fields. The spectral density on $\mathbb{R}^d$ has the general form

$$\alpha \frac{\Gamma(\theta_2 + \frac{1}{2}d)}{\Gamma(\theta_2)\pi^{\frac{1}{2}d}} \cdot \frac{(4\theta_2/\theta_1^2)^{\theta_2}}{\left[(4\theta_2/\theta_1^2)^{\theta_2} + \lambda^2\right]^{\theta_2 + \frac{1}{2}d}}$$

The corresponding isotropic covariance functions have the form

$$K_\theta(x) = \frac{1}{2^{\theta_2 - 1}\Gamma(\theta_2)} \left(\frac{2x\sqrt{\theta_2}}{\theta_1}\right)^{\theta_2} \mathcal{K}_{\theta_2}\left(\frac{2x\sqrt{\theta_2}}{\theta_1}\right)$$

where $\theta_1 > 0$ is a scale parameter controlling the range of correlation and $\theta_2 > 0$ is the parameter controlling the smoothness of the field. $\mathcal{K}_{\theta_2}$ is the modified Bessel function of order θ_2 discussed in Abramowitz and Stegun (1964), §9.

This class of covariance functions is motivated by smooth spectral densities, a wide range of behaviors characterized, interpretable parameters and a number of interesting subclasses. The Exponential class corresponds to the sub-class with smoothness parameter $\theta_2 = 1/2$, that is

$$K_E(x) = \exp(-x/\theta_1).$$

The sub-class defined by $\theta_2 = 1$ was introduced by Whittle (1954) as a model for two dimensional fields. Note that the covariance functions are always positive, so that the class is inappropriate for fields with negative correlations. A general treatment is given in the seminal work by Matérn (1986).

The calculation of $\mathcal{K}_{\theta_2}$ for non-integral θ_2 is quite difficult. All calculations of $\mathcal{K}_{\theta_2}$ in this work use the Amos (1986) algorithm. While the calculation is expensive relative to the other forms of covariance functions, this cost is negligible compared to the other computing costs involved in the analysis.

Consider the behavior at the origin of these covariance functions. Define

$$c(\theta) = \theta_2^{\theta_2} / \left[\theta_1^{2\theta_2} \Gamma(\theta_2) \Gamma(\theta_2 + 1) \right].$$

If θ_2 is an integer,

$$K_\theta(x) = 1 + 2(-1)^{\theta_2} c(\theta) \cdot x^{2\theta_2} \log x + \{ \sum_{j=1}^{\infty} a_j(\theta) x^{2j} \} + x^{2\theta_2+2} \log x \{ \sum_{j=0}^{\infty} b_j(\theta) x^{2j} \}$$

where $\{a_j(\theta)\}_{j=0}^{\infty}$ and $\{b_j(\theta)\}_{j=0}^{\infty}$ are functions of θ alone. If θ_2 is not an integer then,

$$K_\theta(x) = 1 + \pi c(\theta) \cdot x^{2\theta_2} / \sin(\pi\theta_2) + \sum_{j=1}^{\infty} d_j(\theta) x^{2j} + x^{2\theta_2+2} \{ \sum_{j=0}^{\infty} f_j(\theta) x^{2j} \}$$

where $\{d_j(\theta)\}_{j=0}^{\infty}$ and $\{f_j(\theta)\}_{j=0}^{\infty}$ are functions of θ alone. As θ_2 directly controls the smoothness of $K_\theta(x)$ at the origin, it controls the smoothness of the underlying field in the sense that a field with this covariance function is $\lceil \theta_2 \rceil$ times (mean-square) differentiable (Cramèr and Leadbetter (1967)). As θ_2 increases the field becomes increasingly locally smooth. In addition as $\theta_2 \to \infty$, the field approaches an infinitely differentiable field with correlation

function,

$$K_{\theta_1,\infty}(x) = e^{-x^2/\theta_1^2} \tag{2.1}$$

If $Z(x)$ is a Gaussian random field on $\mathbb{R}^d$ with covariance function $K_\theta(x)$ then it satisfies the stochastic partial differential equation (Whittle (1963)):

$$\Big[\ \theta_1^2 \sum_{j=1}^{d} (\partial^2/\partial^2 x_i) - 4\theta_2\ \Big]^{\theta_2 + d/4} \cdot Z(x) = \sigma dW(x)$$

where $\sigma^2 = 2^{4\theta_2}\Gamma(\theta_2 + d/2)(\sqrt{\pi}\theta_1)^d/\Gamma(\theta_2)$ and $W(\cdot)$ is the d dimensional Wiener random field. If $\theta_2 + d/4$ is an integer then this gives a physical basis for the covariance. If $\theta_2 + d/4$ is not an integer then the interpretation is more problematic. This equation has motivated Jones (1989) to use the member with $\theta_2 = 1$ to model Aquifer Head data and is commonly used in hydrology (Mejía and Rodríguez-Iturbe (1974), Creutin and Obled (1982)).

2·2 Inference for the Structure of $Y(x)$

Traditionally, one estimates α and θ by either likelihood methods or various ad hoc approaches. Mardia and Marshall (1984) is a standard reference in the statistical literature. Usually the predictor and the behavior of the prediction error are themselves estimated by 'plugging-in' these estimates for the underlying values in the defining equations.

The log-likelihood of α, θ and β having observed Y is,

$$L(\alpha, \beta, \theta; Y) = -(1/2)\Big[\ n\ln(2\pi\alpha) + \ln(|K_\theta|) + (1/\alpha)(Y - F\beta)' K_\theta^{-1}(Y - F\beta)\ \Big] \tag{2.2}$$

where $K_\theta = \{K_\theta(x_i, x_j)\}_{n\times n}$, $F = \{f_j(x_i)\}_{n\times q}$ and the dependencies upon n have been suppressed.

$L(\alpha, \beta, \theta; Y)$ involves the inverse and determinant of the $n \times n$ matrix K_θ. As K_θ is a covariance matrix it is positive definite and so, in principle these operations present no difficulties. As the number of observations increases the condition number of K_θ increases, so that so that numerical stability becomes important. In calculating the log-likelihood it is unnecessary to invert K_θ directly. All that is needed is the log determinant of K_θ and a quadratic form. These may be determined from the Cholesky factorization and solving

linear systems in the Cholesky triangle using back-substitution. This is more efficient and numerically stable than calculating the inverse directly. However the Cholesky factorization still requires $n^3/6+O(n^2)$ operations. Unless there is regularity in the sampling design, inducing a patterned structure in K_θ, it is unclear if more efficient algorithms exist.

3. Additive models for $Y(x)$

Often the dimension of x is of the same order as n, and moderately large $(10-100)$. In addition it is common for little to be known about the form of the relationship between the input factors and the response. In such circumstances the detection and measurement of complex relationships between the response and the input factors is difficult. An empirical observation is that, more often than not, much of the variation in $y(x)$ can be explained by the separate influence of each factor ("main effects"). This suggests the model

$$Y(x) = Y_1(x_1) + \ldots + Y_d(x_d) + Z(x)$$

where $\{Y_i(x_i)\}_{i=1}^d$ are real-valued Gaussian random fields over $\{x_i : x \in R\} \subset \mathbb{R}$ and, $Z(x)$ is a Gaussian random field to describe the non-additive components of $Y(x)$. This additive model was introduced by Stein (1989). For definiteness, we consider here the situation where the unidimensional $Y_i(x_i)$ are stationary with Matérn covariance functions $\alpha_i K_{(\theta_{1i},\theta_2)}(x_i)$. Hence $Y(x)$ has covariance function

$$K_\theta(x,w) = \sum_{i=1}^{d} \alpha_i K_{\theta_i}(x_i - w_i) + \mathrm{cov}\{Z(x), Z(w)\} \tag{3.1}$$

The central issue is inference for the parameters $\{\alpha_i, \theta_{1i}, \theta_2\}$. It is important to realize that the information in a single realization for $\{\theta_{1i}\}$ is limited. In fact if we observed the $y(x)$ everywhere we would still not be able to determine the $\{\theta_{1i}\}$ exactly, even though $\{\alpha_i, \theta_2\}$ could be determined with almost surely. The "increasing information" rationale for the optimality of maximum likelihood estimation does not hold in this setting (Handcock (1989)).

For inference about $Y(x)$ and prediction we will need to explicitly take into account the uncertainty in the model. The Bayesian paradigm is a natural

one for expressing this uncertainty. Using Bayesian methods it is straightforward to derive the (posterior) predictive distributions for linear functionals of $Y(x)$, as in Handcock and Stein (1989). We will use such predictive distributions in §5.

4. Designing experiments under an additive model for $Y(x)$

There is a substantial literature on the design of computer experiments based on given and natural design criteria. For an overview see Sacks and Ylvisaker (1984, 1985). We are interested in designs that exploit the additive structure (3.1). Our criterion is two fold: first, the efficient estimation of the covariance function of $Y(x)$, and second, the optimal prediction of linear functionals of $Y(x)$. Latin hypercube designs are a natural candidate under these conditions. They have been shown by Stein (1987) to be appropriate for additive functions of inputs in the multidimensional integration setting. This randomized procedure is "space filling": spreading out the design points in the design space.

Consider the parameters in the Matérn class. The response for design points close together will provide information useful for the estimation of the scale (α) and smoothness parameters (θ_2). The response for design points far apart will provide useful information for estimating the regression parameters and the range (θ_{1i}). Hence we expect Latin hypercube designs to provide information about the mean and range, while the scale and smoothness may be poorly estimated. Other "low discrepancy" design such as Faure (1982) and Halton (1960) sequences tend to have similar properties.

4·1 Cascading Latin hypercube Designs

In this section we introduce a generalization of Latin hypercube designs to enhance estimation of the scale and smoothness of the additive random field. An example will motivate the design. Consider the two dimensional 27 point Latin hypercube design in Figure 1. The design space $[0,1] \times [0,1]$ is divided into a 27×27 grid of cells. The design points are chosen so that exactly one point is in each row or column of the grid, and placed randomly within that cell. Here, as elsewhere, an argument can be made for placing the design point at the center of the cell. Such designs will be called "centered".

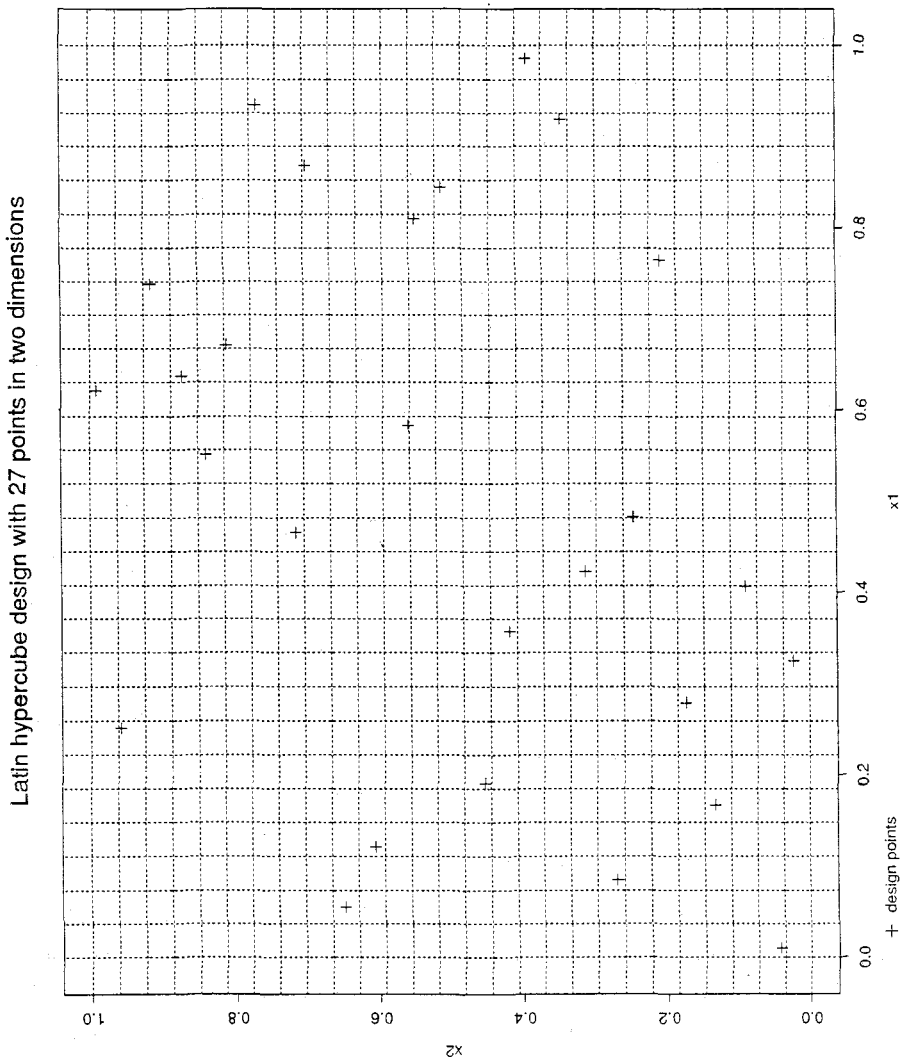

Latin hypercube design with 27 points in two dimensions
x2
0.0
0.2
0.4
0.6
0.8
1.0
x1
+ design points

FIG. 1

Consider the 27 point design in Figure 2. The 9 circles are a centered Latin hypercube on a 9×9 grid in $[0,1] \times [0,1]$. Around each of these circles is a 3 point Latin hypercube design, based on the localized 3×3 square of cells. We note that this is still a Latin hypercube, as there is exactly one design point per row and column on the 27×27 grid of cells. However the design has the additional advantage of ensuring that there are design points close together.

In general a **Cascading Latin hypercube design** of $n \equiv \prod_{k=1}^{L} n_k$ points in d-dimensions with levels $(n_1, \ldots, n_L)$ is a n_L Latin hypercube design about each point in the $(n_1, \ldots, n_{L-1})$ centered Cascading Latin hypercube design. The usual Latin hypercube design is the special case with a single level. Experience indicates that two or three levels is enough and that the lower levels should consist of only two or three points. Note that a d-dimensional Cascading Latin hypercube is always a b-dimensional Cascading Latin hypercube in a projected cube in b of the original factors. The value of the Cascading Latin hypercube will tend to increase as the dimension d increases relative to n.

5. AN EXAMPLE: PREDICTING THE OVERALL MEAN

In this section we compare the performance of Cascading Latin hypercube designs with other commonly used designs. As an objective we will predict the overall mean of the response:

$$\mu_Y \equiv \frac{1}{|R|} \int_R Y(x) dx$$

where $|R|$ is the volume of R. As μ_Y is a linear functional of $Y(x)$ we can use the mode of the predictive distribution $\mathrm{pr}(\mu_Y | Y)$ as a point predictor, and the predictive standard deviation as a measure of uncertainty in the prediction. We will not describe this procedure here, but refer to Handcock and Stein (1989).

As a concrete example consider an additive Matérn random field, $Y(x)$, over $R = [0,1]^d$ with $\alpha_i = 1$, and $\theta_{1i} = \theta_1$, $i = 1, \ldots d$. We can generate a realization from $Y(x)$ on a very fine grid, so that we essentially observe

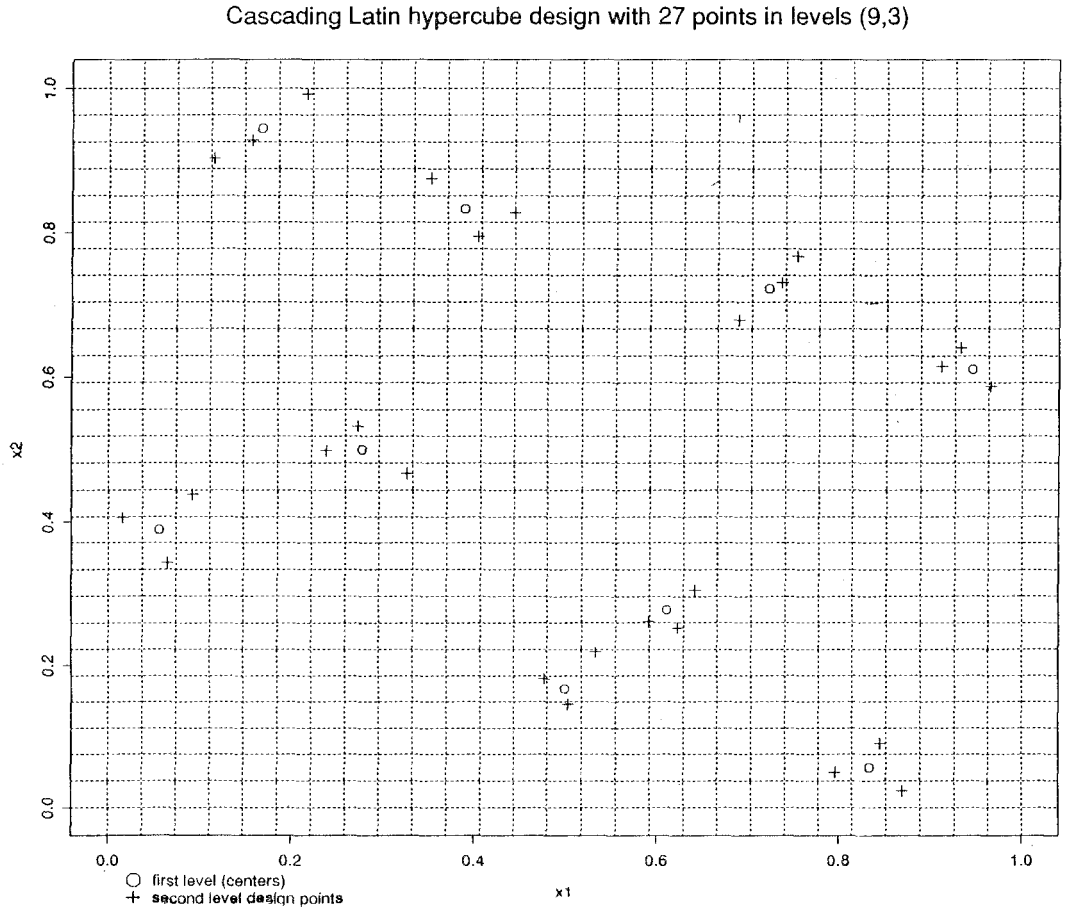

Cascading Latin hypercube design with 27 points in levels (9,3)
x2
x1
0.0
0.2
0.4
0.6
0.8
1.0
first level (centers)
second level design points

FIG. 2

$y(x)$ continuously on R. Figure 3 represents the three two-dimensional projections of a realization with $\theta = (\frac{1}{2}, \frac{1}{2})$. This realization is continuous, but not differentiable. Under the usual standards for response surfaces, this realization is quite "rough". Figure 4 represents a realization with $\theta = (\frac{1}{2}, 2\frac{1}{2})$. This realization is twice differentiable, and is "smooth" relative to the previous example.

By increasing θ_1, and hence the level of correlation, more realistic fields can be produced. Figure 5 represents a two level Cascading Latin hypercube sample for a realization based on $\theta = (1, \frac{1}{2})$.

5·1 Evaluating Cascading Latin hypercube Designs

To explore the value of Cascading Latin hypercube designs, we designed a simulation experiment. We consider only 27 point designs in $d = 10$ dimensions generated by

a) simple random designs

b) Faure sequences

c) Latin hypercube designs

d) two level cascading Latin hypercube designs

e) three level cascading Latin hypercube designs

Each of the procedures was used to generate design points and the response measures for a given realization of the above additive model. Then θ is estimated using maximum likelihood. The results from 10 repetitions of this process on a single realization from the "rough" process ($\theta = (\frac{1}{2}, \frac{1}{2})$) are presented graphically in Figure 6. The cross-lines represent the smoothness and range for the underlying random field. It is important to note that the single realization may not reflect exactly these values. In particular, note the strong bias in the estimates of the range, regardless of the design procedure used. The values with smoothness of 10 represent particular designs where the maximum likelihood estimator is the limiting infinitely differentiable process of (2.1).

We see that 27 points in 10 dimensions from a non-differentiable process is insufficient to rule out much smoother random fields. Table I provides a numerical summary of the results.

two dimensional projections of a 10 dimensional Additive Matern random field
scale = 1, range = 1/2, smoothness = 1/2

two dimensional projections of a 10 dimensional Additive Matern random field
scale = 1, range = 1/2, smoothness = 2 1/2
response
dimension 2
dimension 1
response
dimension 3
dimension 1
response
dimension 3
dimension 2

FIG. 4

two dimensional projections of a 10 dimensional Additive Matern random field
scale = 1, range = 1, smoothness = 1/2

FIG. 5

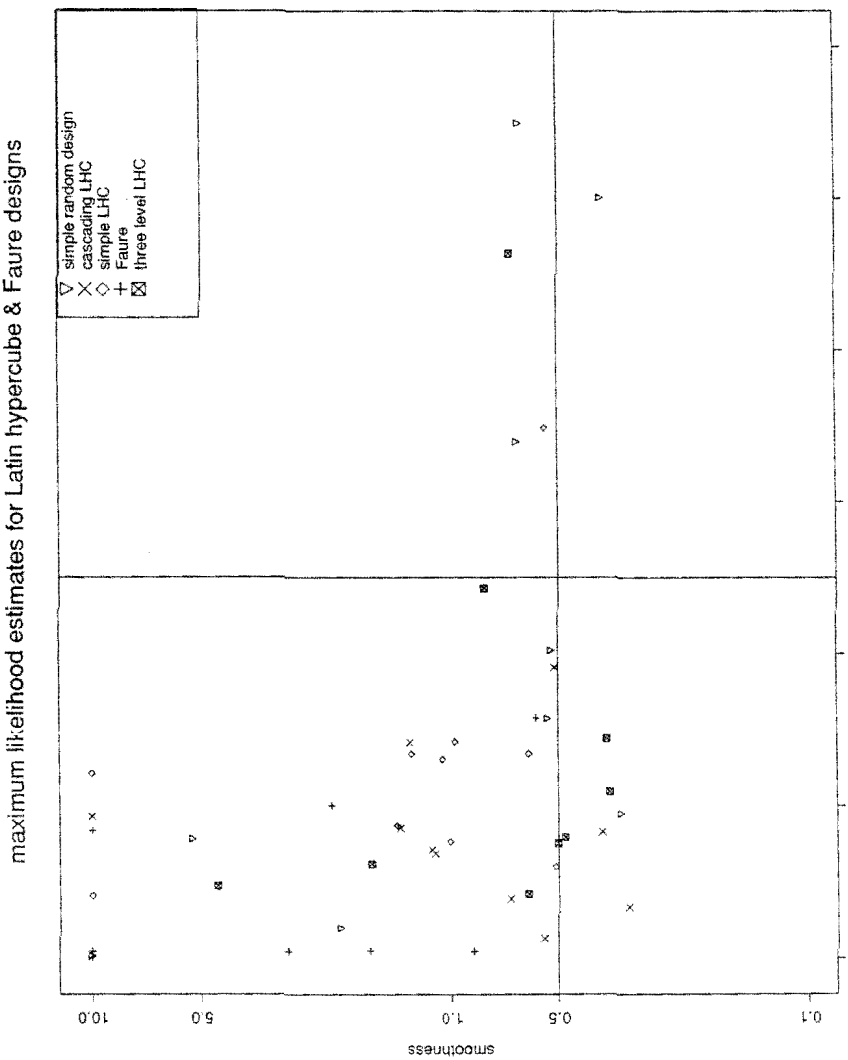

maximum likelihood estimates for Latin hypercube & Faure designs
simple random design
cascading LHC
simple LHC
Faure
three level LHC
smoothness
range
10.0
5.0
1.0
0.5
0.1
0.0
0.2
0.4
0.6
0.8
1.0
1.2

FIG. 6

TABLE I

Summary of simulation experiment for a "rough" realization

	Design Criterion					
Factor	SRS	Faure	Simple LHC	two level LHC	three level LHC	Truth
number finite	8	4	9	10	9	10
median range	0.25	0.01	0.25	0.15	0.15	0.5
median smoothness	0.64	6.42	1.03	0.90	0.55	0.5
var for range	4.60	5.90	5.06	5.47	5.04	NA
var for smoothness	14.43	28.85	8.30	13.31	3.57	NA

Note: The table summaries, for each design procedure, characteristics of 10 repetitions of the design on a single realization from a random field with $\theta = (1, 0.5, 0.5)$.

FIG. 6

All design procedures tend to overestimate the degree of smoothness. They all do about as well in estimating the range, while the three level Cascading Latin hypercube does substantially better in estimating the smoothness. For a given realization the maximum likelihood estimates exhibit similar biases, although across realizations the range is both over and under estimated.

TABLE II

Summary of prediction based on a "rough" realization

	Design Criterion					
Factor	SRS	Faure	Simple LHC	two level LHC	three level LHC	Truth
mean prediction	0.45	0.40	0.50	0.47	0.43	0.45
RMS prediction error	0.22	0.18	0.11	0.12	0.06	NA

Note: The table summaries, for each design procedure, prediction characteristics of 10 repetitions of the design on a single realization from a random field with $\theta = (1, 0.5, 0.5)$.

How well do these designs do when we wish to predict the overall mean? Figure 7 represents the predictive mode and standard deviation for each of the designs in Figure 6. The "truth" can be directly determined from the generated realization, $\mu_Y = 0.447$. A numerical summary is given in Table II.

The mean is the mean of the predictive mode over the 10 repetitions. For this realization, the three level Cascading Latin hypercube does better. Typically, the two and three level Cascading Latin hypercube designs are similar, while the Faure and simple random design are extremely poor.

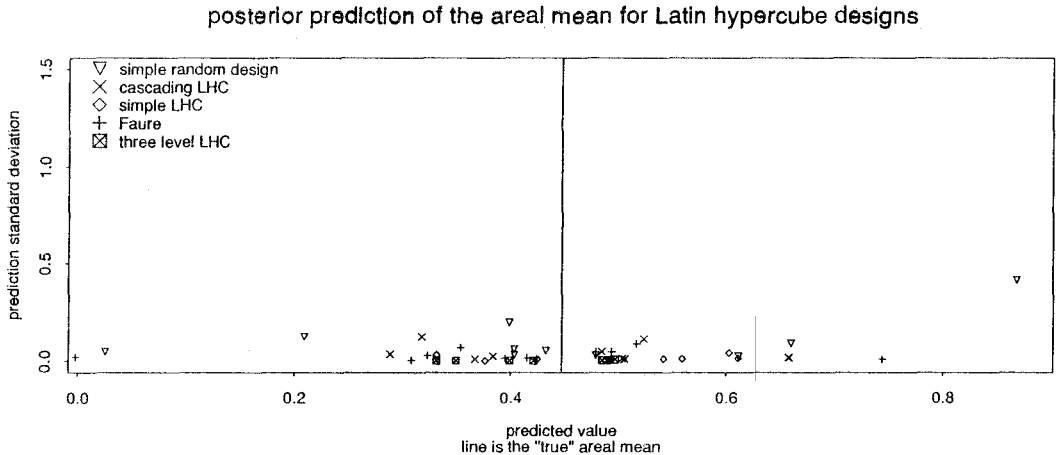
Figure 8 graphically represents an analogous experiment on a "smooth" realization with $(\theta = (\frac{1}{2}, 2\frac{1}{2}))$. Again the smoothness and range are overestimated. The graphical summary of the prediction is given in Figure 9, and numerical summary in Tables II and III. Again the two and three level Cascading Latin hypercube designs are similar, while the Faure and simple random design are substantially weaker. Note that the prediction error are substantially smaller, reflecting the relative smoothness of the realization.

TABLE III

Summary of simulation experiment for a "smooth" realization

	Design Criterion					
Factor	SRS	Faure	Simple LHC	two level LHC	three level LHC	Truth
number finite	9	4	10	9	9	10
median range	1.66	0.01	1.66	2.01	1.79	0.5
median smoothness	3.00	10.00	2.47	2.52	2.66	2.5
sd for range	0.98	0.11	0.68	0.92	0.43	NA
sd for smoothness	2.62	4.26	0.88	2.59	2.38	NA

Note: The table summaries, for each design procedure, characteristics of 10 repetitions of the design on a single realization from a random field with $\theta = (1, 0.5, 2.5)$.


posterior prediction of the areal mean for Latin hypercube designs
simple random design
cascading LHC
simple LHC
Faure
three level LHC
prediction standard deviation
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
predicted value
line is the "true" areal mean

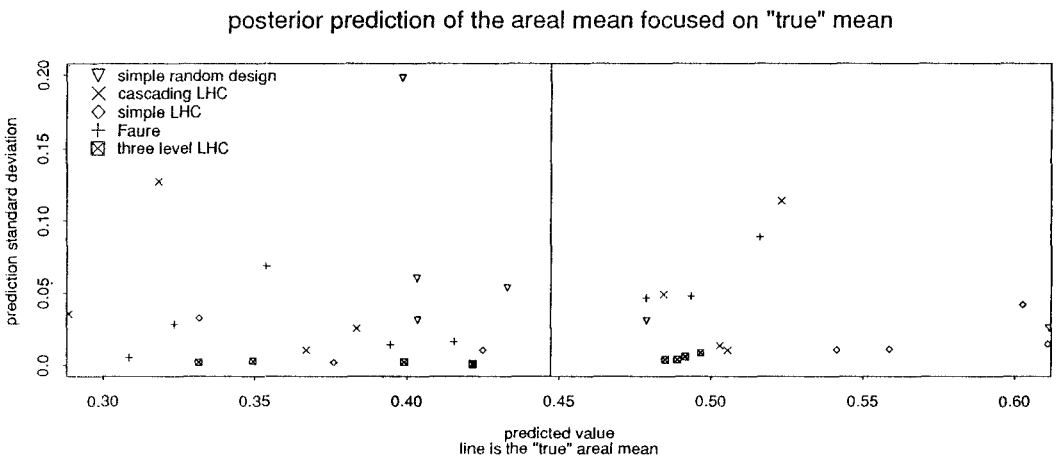

posterior prediction of the areal mean focused on "true" mean
simple random design
cascading LHC
simple LHC
Faure
three level LHC
prediction standard deviation
0.0
0.05
0.10
0.15
0.20
0.30
0.35
0.40
0.45
0.50
0.55
0.60
predicted value
line is the "true" areal mean

FIG. 7

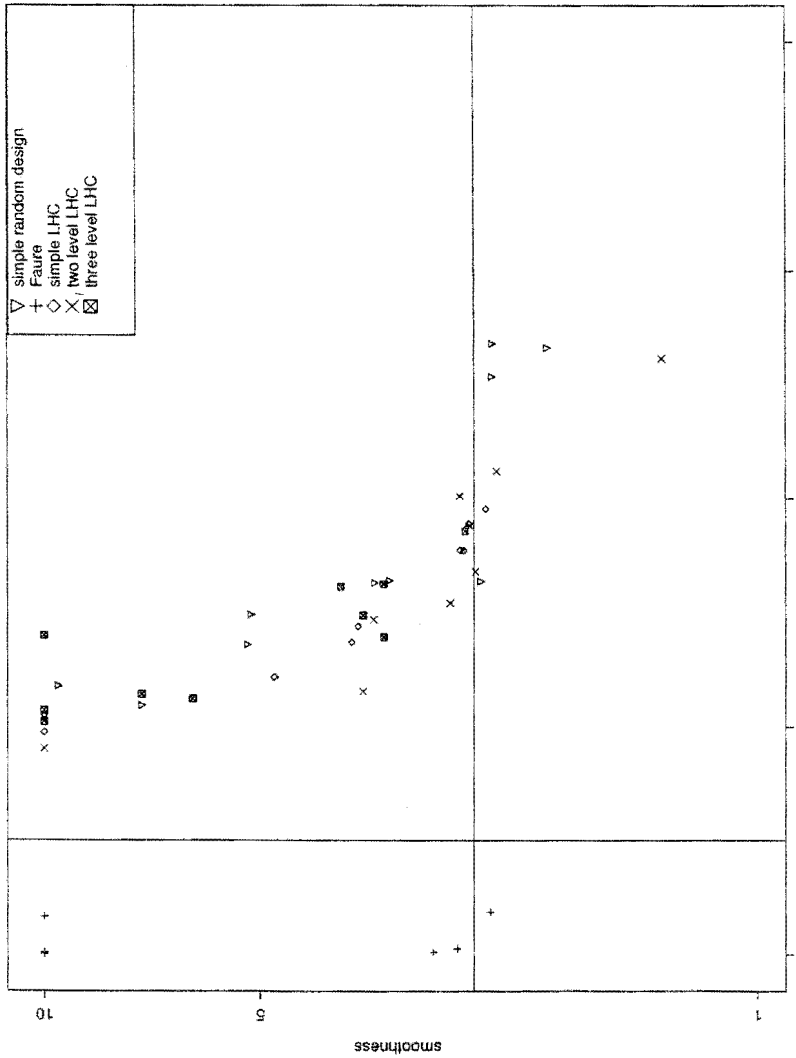
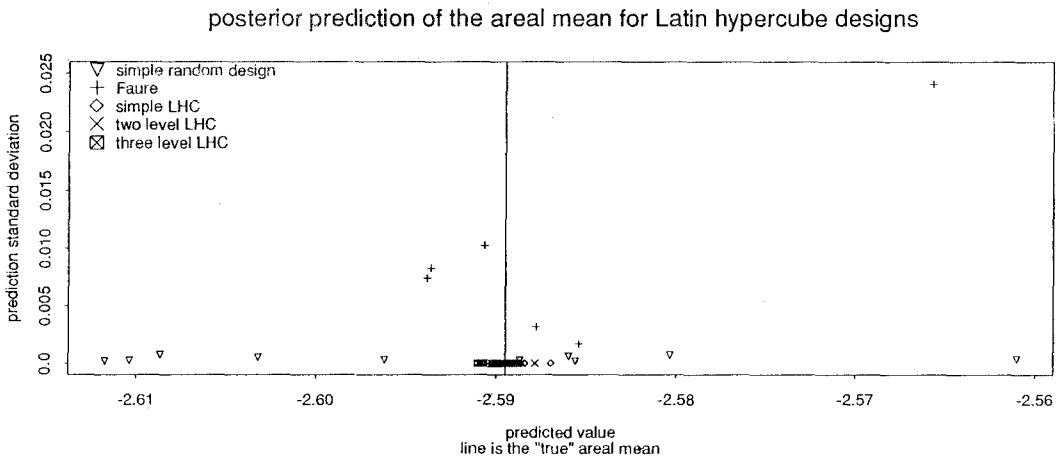

maximum likelihood estimates for Latin hypercube & Faure designs
simple random design
Faure
simple LHC
two level LHC
three level LHC
smoothness
range
10
5
1
0
1
2
3
4

FIG. 8


posterior prediction of the areal mean for Latin hypercube designs
simple random design
Faure
simple LHC
two level LHC
three level LHC
prediction standard deviation
0.0
0.005
0.010
0.015
0.020
0.025
-2.61
-2.60
-2.59
-2.58
-2.57
-2.56
predicted value
line is the "true" areal mean

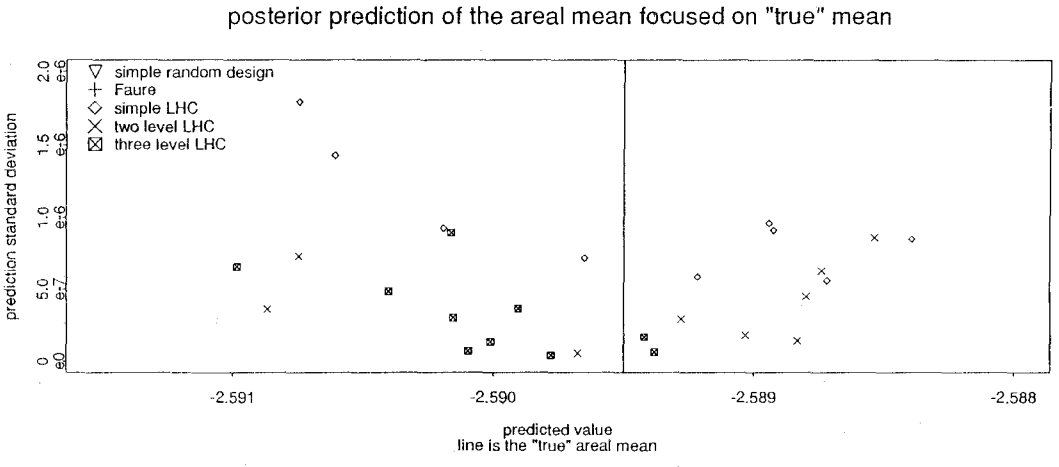

posterior prediction of the areal mean focused on "true" mean
simple random design
Faure
simple LHC
two level LHC
three level LHC
prediction standard deviation
0
5.0 e-7
1.0 e-6
1.5 e-6
2.0 e-6
-2.591
-2.590
-2.589
-2.588
predicted value
line is the "true" areal mean

FIG. 9

TABLE IV

Summary of prediction based on a "smooth" realization

	Design Criterion					
Factor	SRS	Faure	Simple LHC	two level LHC	three level LHC	Truth
mean prediction	-2.59	-2.60	-2.59	-2.59	-2.59	-2.59
RMS prediction error	0.016	0.056	0.001	0.0009	0.0007	NA

Note: The table summaries, for each design procedure, prediction characteristics of 10 repetitions of the design on a single realization from a random field with $\theta = (1, 0.5, 2.5)$.

6. Conclusion

Additive random field models in experimental design are a promising new approach for increasing the efficiency of experiments for product quality improvement. This paper introduces Cascading Latin hypercube designs as a means to exploit the additive nature of the models. Based on a theoretical arguments and a modest simulation experiment, they appear to improve the efficiency of prediction and deserve further study.

Acknowledgment

The author grateful to D. Raghavarao, J. R. M. Hosking and J. E. Kolassa for valuable comments on the manuscript. This work was completed while the author was at the IBM T. J. Watson Research Center, Yorktown Heights, N. Y.

Bibliography

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions*. New York: Dover.

AMOS, D. E. (1986). A Portable Package for Bessel Functions of Complex Argument and Nonnegative Order. *ACM Transactions of Mathematical Software* **12**, 265-273.

BESAG, J. (1975). Statistical Analysis of Non-lattice Data. *Statistician* **24**, 179-195

BOX, G. E. P., HUNTER, W. G., and HUNTER, J. S. (1978). *Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building* New York: Wiley.

BOX, G. E. P. and DRAPER, N. R. (1987). *Empirical Model Building and Response Surfaces* New York: Wiley.

CRAMÉR, H. and LEADBETTER, M. R. (1967). *Stationary and Related Stochastic Processes.* New York: Wiley.

CREUTIN, J. D. and OBLED, C. (1982). Objective Analyses and Mapping Techniques for Rainfall Fields: An Objective Comparison. *Water Resources Research* **18**, 413-431.

CURRIN, C., MITCHELL, T. J., MORRIS, M., and YLVISAKER, D. (1988). A Bayesian approach to the design and analysis of computer experiments. ORNL Technical Report 6498, available from the National Technical Information Service, Springfield, VA, 22161.

FAURE, H. (1982). Discrépance de suites associées à un système de numération (en dimension s). *Acta Arithmetica* **41**, 337-351.

HALTON, J. H. (1960). On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals. *Numer. Math.* **2**, 84-90.

HANDCOCK, M. S. (1989). Inference for Spatial Gaussian Random Fields when the Objective is Prediction. Ph.D. thesis, Department of Statistics, University of Chicago, Chicago, Illinois.

HANDCOCK, M. S. and STEIN, M. L. (1989). A Bayesian analysis of Kriging. submitted to *Technometrics.*

JONES, R. H. (1989). Fitting a Stochastic Partial Differential Equation to Aquifer Head Data. *Stochastic Hydrology and Hydraulics* **3**, 100-105

MARDIA, K. V. and MARSHALL, R. J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* **71**, 135-146.

MARDIA, K. V. and WATKINS, A. J. (1989). On Multimodality of the Likelihood in the Spatial Linear Model. *Biometrika* **76**, 289-95.

MATÉRN, B. (1986). *Spatial Variation* Second Ed, *Lecture Notes in Statistics,* **36.** Berlin: Springer-Verlag.

MEJÍA, J. M. and RODRÍGUEZ-ITURBE, I. (1974). On the Synthesis of Random Field Sampling From the Spectrum: An Application to the Generation of Hydrologic Spatial Processes. *Water Resources Research* **10**, 705-711.

SACKS, J. and SCHILLER, S. (1988). Spatial Designs. Design and Analysis of Computer Experiments In *Statistical Decision Theory and related Topics IV,* Eds. S. S. Gupta A. and J. O. Berger,**2**, pp. 385-389. Springer: New York.

SACKS, J., SCHILLER, S. B., and WELCH, W. J. (1989). Designs for Computer Experiments. *Technometrics* **31**, 41-49.

SACKS, J., WELCH, W. J., MITCHELL, T. J., and WYNN, H. P. (1989). Design and Analysis of Computer Experiments. *Statistical Science* **4**, 409-435.

SACKS, J., and YLIVISAKER, D. (1984). Some model robust designs in regression. *Annals Statistics* **12**, 1324-1348.

SACKS, J., and YLIVISAKER, D. (1985). Model robust designs in regression: Bayes theory In *Proc. of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer,* Eds. L. M. Le Cam and R. A. Olshen,**2**, pp. 667-679. Wadsworth: Monterey.

STEIN, M. L. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics* **2**, 143-151

STEIN, M. L. (1989). Discussion of Design and Analysis of Computer Experiments. by J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn *Statistical Science* **4**, 432-433.

VECCHIA, A. V. (1988). Estimation and Model Identification for Continuous Spatial Processes. *J. Royal Statist. Soc. B, 50,* 297-312.

WHITTLE, P. (1954). On Stationary Processes in the Plane. *Biometrika* **41**, 434-449.

WHITTLE, P. (1963). Stochastic Processes in Several Dimensions. *Bulletin of the International Statistical Institute* **40**, 974-994.