NISS

Spatial Sampling and the Environment

Dennis D. Cox, Lawrence H. Cox, and
Katherine B. Ensor

Technical Report Number 38
October, 1995

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

Spatial Sampling and the Environment

Dennis D. Cox

Department of Statistics

Rice University

Email: dcox@stat.rice.edu

Lawrence H. Cox

National Exposure Research Laboratory

U.S. Environmental Protection Agency

Email: COX.LARRY@epamail.epa.gov

Katherine B. Ensor

Department of Statistics

Rice University

Email: kathy@stat.rice.edu

Abstract

In this paper we summarize research issues for spatial environmental sampling stemming from a NISS/USEPA workshop held September 21–22, 1994 at Chapel Hill, NC.

Key Words and Phrases : experimental design, spatial statistics, kriging, environmental monitoring, multi-phase sampling.

1 Introduction.

The problem of choosing a spatial sampling design arises in many contexts in environmental statistics, e.g., determining where to take measurements at a hazardous waste site for the purpose of surface soil characterization and remediation, for planning large scale studies such as the National Resources Inventory (NRI) or the Environmental Monitoring and Assessment Program (EMAP), or for selecting locations for air quality monitoring stations. To address some of these issues, the National Institute of Statistical Sciences (NISS), the Center for Statistical Ecology and Environmental Statistics of the Pennsylvania State University, and the U.S. Environmental Protection Agency (US EPA) jointly organized a Workshop on Spatial Sampling for the Environment that was held September 21–22, 1994 in Chapel Hill, NC. The program covered the following topics:

(1) The theory of spatial statistics and spatial sampling designs with consideration for issues such as design criteria, robustness, and computation;

(2) Tradeoffs between design and analysis;

(3) Adaptive and multi–phase sampling;

(4) Applications to brush fire forecasting, site remediation, atmospheric monitoring, and natural/environmental resources measurement and assessment.

In this paper we report on issues stemming from the workshop and examine topics and questions which merit further investigation. Before launching into any theory, it is instructive to look at some environmental applications which were discussed at the workshop.

1.1 Surface Soil Remediation.

In surface soil remediation it is usually necessary to characterize the pre-remediation and/or post-remediation distribution of contaminant concentrations. It may also be necessary to determine which portions of the site require remediation. An exhaustive sampling procedure may not be feasible, often because of the high cost of obtaining the samples (e.g., when it is necessary to use drilling equipment to obtain cores). Sampling design in this setting entails balancing the costs of acquiring the information with the costs of making mistakes as a result of insufficient information. Typical issues to be considered include the spatial distribution of contamination, area and depth heterogeneity, physical extraction of the sample, and performing the chemical analysis.

An instructive example presented at the workshop by Max Morris is based on work by the late Toby Mitchell (West, Siegrist, Mitchell, Pickering, Muhr, Greene, and Jenkins 1993). A sampling plan was undertaken to select of sites at which to drill soil sample cores for the purpose of estimating the distribution of Volatile Organic Compounds (VOC's) in a hazardous waste site. It was projected that the entire site was to be remediated, but the remediation method was untested. The purpose of the study was to characterize the preremediation distribution of contaminants and to select regions of the site for demonstration of the proposed remediation technology. This design is discussed in Section 3.6.

A similar example that was presented at the workshop by E. J. Englund (Englund and Heravi (1993)) involved the use of conditional simulation (i.e., simulation from a spatial model fit to the observations) to compute sampling locations for the purpose of estimating where the contaminant concentration exceeds a given level. The sampling procedure was iterative: an initial spatial model was fit to the data; conditional simulations were conducted using the estimated model; the expected loss

was computed; the sampling units with the highest expected loss were selected for measurement; and, the process was repeated.

1.2 National Resources Inventory and Related Studies.

The National Resources Inventory (NRI) is a survey conducted periodically by the Soil Conservation Service (SCS) of the U.S. Department of Agriculture with the assistance of the Survey Section of the Statistical Laboratory at Iowa State University. The goals of the survey are to assess soil characteristics, land use, erosion, and conservation needs for all nonfederal land in the U.S., Puerto Rico, and the U.S. Virgin Islands. The SCS began conducting surveys of this nature in 1934, but NRI surveys of 1982, 1987, and 1992 were conducted on a much larger scale and with broader assessment goals than any previous efforts. The NRI is based on a probability sample similar to ones used in inference about finite populations.

A related study is the Environmental Monitoring and Assessment Program (EMAP). EMAP is a multi-agency program of the U.S. government organized by the U.S. Environmental Protection Agency. EMAP involves extensive data collection on the status and trends of ecological resources, including wetlands, lakes and streams, forests, and agroecosystems. Some EMAP data collection is based on a multi-stage probability sampling design. The probability sampling strategies for NRI and EMAP are discussed in Section 3.1.

1.3 Air Quality Monitoring.

Regional air quality monitoring is typically done with expensive equipment placed at monitoring stations located at fixed sites in the region of interest. Monitoring networks need to be evaluated for factors including accuracy and cost–effectiveness. As monitoring data are frequently used for other, non–regulatory, purposes, such

as estimating and modeling human exposure to various pollutants, estimation of trends in pollutant concentrations (Gao, Sacks, and Welch 1994), and investigation of source-receptor relationships (Spiegelman and Dettner 1993), the appropriateness of such networks and their limitations for these purposes also needs to evaluated. The design of an air quality monitoring network offers challenges not encountered in other environmental applications. For one, the sites are generally fixed in space (although mobile monitoring stations may also be utilized) and observations are taken for a long period of time. The rapid transport of pollutants and their relationship with other variables such as meteorological variables raise other issues.

2 Spatial Statistics: Theory

There are two disparate approaches to spatial statistical inference, the *model based* and *design based* approaches. The model based approach begins with a stochastic process model for the spatial process and uses this model for prediction after estimating unknown parameters. The design based approach typically proceeds from a simple random sample or other probability based method of generating observation sites and then bases inferences on the randomness of the design. Design based inferences are generally elementary from a statistical point of view and are arguably more robust across variables of interest and across time than model based inferences. However, one expects that a good model based analysis can do no worse than a design based analysis and can perform substantially better when there is spatial correlation.

Some model based methods incorporate design based approaches. For example, Englund and Heravi (1995) use design based methods to select a grid within which an (unobserved) sample location is selected randomly. Qian (1995) selects (unobserved) sample locations by successive random sampling from conditional posterior t-distributions. We give a brief synopsis of the model based approach (see Cressie

1991) in what follows.

2.1 Stochastic Process Model.

A *spatial process* is a real (or vector valued) process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{D}\}$ where $\mathcal{D} \subset \mathbb{R}^k$ for some k, typically $k = 2$ or $k = 3$. There are various models that are widely used for Z, but the following will suffice for our discussion. Assume the mean function or *trend* is of the form

$$E[Z(\boldsymbol{s})] \;=\; \mu(\boldsymbol{s}) \;=\; \sum_{j=1}^{p} \beta_j B_j(\boldsymbol{s}), \tag{1}$$

where the B_j are given functions. For instance, a constant unknown mean is obtained by taking $p = 1$ and $B_1 \equiv 1$, and a mean function which is linear is specified by $p = k + 1$, $B_1 \equiv 1$, and $B_{j+1}(\boldsymbol{s}) = s_j$ for $1 \leq j \leq k$. Write the covariance function as

$$\mathrm{Cov}[Z(\boldsymbol{s}), Z(\boldsymbol{t})] \;=\; C(\boldsymbol{s}, \boldsymbol{t}). \tag{2}$$

Virtually all models make use of a stationary covariance function, i.e., one of the form

$$C(\boldsymbol{s}, \boldsymbol{t}) \;=\; c(\boldsymbol{s} - \boldsymbol{t}),$$

where $c(\cdot)$ is known as the *covariogram*, which is related to the popular *variogram* 2γ by $2\gamma(\mathbf{h}) = 2[c(\mathbf{0}) - c(\mathbf{h})]$. Possible covariance functions include

$$C_1(\boldsymbol{s}, \boldsymbol{t}) \;=\; \sigma^2 \exp\left[-\sum_{i=1}^{k} \theta_i (s_i - t_i)^2\right], \tag{3}$$

$$C_2(\boldsymbol{s}, \boldsymbol{t}) \;=\; \omega^2 \exp\left[-\psi \|\boldsymbol{s} - \boldsymbol{t}\|^p\right], \tag{4}$$

$$C_3(\boldsymbol{s}, \boldsymbol{t}) \;=\; \begin{cases} c_s[1 - (3/2)\|\boldsymbol{s} - \boldsymbol{t}\|/a + (1/2)(\|\boldsymbol{s} - \boldsymbol{t}\|/a)^3] & \text{if } \|\boldsymbol{s} - \boldsymbol{t}\| \leq a, \\ 0 & \text{if } \|\boldsymbol{s} - \boldsymbol{t}\| \geq a, \end{cases} \tag{5}$$

where in (5) it is assumed that the dimension k of the variables $\boldsymbol{s}$, $\boldsymbol{t}$ is no more than 3. In the above, σ^2, θ_i, ω^2, ψ, c_s, and a are positive and $0 < p \leq 2$. Note that all of these are stationary covariances, and C_2 and C_3 are also isotropic, i.e. depend only

on $\|\boldsymbol{s} - \boldsymbol{t}\|$. Anisotropic covariances are useful when variables (components of $\boldsymbol{s}$) are not equally important (Sacks, Welch, Mitchell, and Wynn 1989). Other candidate covariance functions may be found in Cressie (1991).

Covariances such as those in (3) and (4) were used by several workshop presenters including Ylvisaker and Morris. See Sacks et al. (1989) for other applications of these particular covariance functions. The covariance in (5) is known as the "spherical model" in geostatistics (equation (2.3.8), p. 61 of Cressie 1991).

Construction of (optimal) linear estimators and their conditional variances do not require specific distributional assumptions, whereas nonlinear estimation and prediction regions do. Almost always, the Gaussian distribution is used.

Is it reasonable to model the chemical concentration at a hazardous waste site as a random process? Such an approach is basically Bayesian – one can often select the mean and covariance functions to reflect properties one believes the process satisfies. For instance, sample paths of a process with covariance as in (3) are infinitely differentiable and have a different variation in different directions (depending on the magnitudes of the θ_i's). The sample paths of a process with covariance as in (4) have common variability in all directions and variable roughness controlled by the parameter p. For instance, a Gaussian process with $p = 1$ in $k = 1$ dimension is an Ornstein–Uhlenbeck process which is nowhere differentiable and approaches a white noise process as $\psi \to \infty$. From a pragmatic perspective, such spatial stochastic process models have been successfully used in numerous applications (see, e.g., Cressie 1991; Currin, Mitchell, Morris, and Ylvisaker 1991; or Gao et al. 1994).

2.2 Prediction.

We will suppose that the data consist of observations $Z(\boldsymbol{s}_1)$, $Z(\boldsymbol{s}_2)$, ..., $Z(\boldsymbol{s}_n)$ of the process at sites $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n\}$ in a region $\mathcal{D}$. A list of possible objectives is given

next. Below, $\mathcal{D}_0$ is a given subregion of $\mathcal{D}$.

(O–1) Given $\boldsymbol{t} \in \mathcal{D}$, predict (or estimate) the value of the process $Z(\boldsymbol{t})$ at a point $\boldsymbol{t}$.

(O–2) Predict the the integral

$$I = \int_{\mathcal{D}_0} Z(\boldsymbol{t})\, d\boldsymbol{t}$$

or the average, $I/|\mathcal{D}_0|$.

(O–3) Predict the integral of a nonlinear function of the process, i.e.

$$J = \int_{\mathcal{D}_0} g(Z(\boldsymbol{t}))\, d\boldsymbol{t}.$$

A typical example would be $g(z) = \exp[z]$.

(O–4) Predict the maximum of the process over $\mathcal{D}_0$, i.e.

$$\max_{\boldsymbol{t} \in \mathcal{D}_0} Z(\boldsymbol{t}).$$

(O–5) Estimate the unknown subset of $\mathcal{D}$ where the process exceeds some given threshold, i.e.

$$\mathcal{D}(\lambda) = \{\boldsymbol{t} : Z(\boldsymbol{t}) \geq \lambda\}.$$

In each of the above estimation problems, one would of course desire prediction intervals as well as point predictions.

Objective (O–1) is fundamental. Contour plots based on such point predictions are popular with practitioners (see, e.g., Flatman 1984). Objective (O–2) arises, e.g., in soil abatement/remediation if $Z(\boldsymbol{s})$ is the contaminant concentration at $\boldsymbol{s}$ and one wishes to estimate the total amount of contaminant in a certain area. However, it may be more reasonable to take Z as the logarithm of the contaminant concentration, in which case (O–3) becomes relevant. In air quality monitoring, regulatory requirements are based on the maximum ambient ozone levels observed at the monitoring

stations, but the true maximum concentration (i.e., (O–4)) is also needed in order to model human or ecological exposure in the region. It is also important to estimate where the concentration exceeds some threshold, thus giving rise to (O–5). Objectives (O–4) and (O–5) are related to location of "hot spots" in soil abatement applications (Chapter 10 of Gilbert 1987).

Most commonly used predictors of $Z(\boldsymbol{t})$ for (O–1) are linear unbiased predictors. Letting $\hat{Z}(\boldsymbol{t})$ denote a predictor of $Z(\boldsymbol{t})$, we say $\hat{Z}(\boldsymbol{t})$ is linear if it is of the form

$$\hat{Z}(\boldsymbol{t}) = \sum_{i=1}^{n} a_i Z(\boldsymbol{s}_i)$$

for some coefficients a_i which depend on $\boldsymbol{t}$. Unbiasedness is equivalent to the requirement that $\sum_{i=1}^{n} a_i B_j(\boldsymbol{s}_i) = 0$, $1 \leq j \leq p$, where the B_j are given in (1). The Best Linear Unbiased Predictor (BLUP) of $Z(\boldsymbol{t})$ (also known as the kriging predictor) is the one which minimizes the variance of the prediction error. It is given by

$$\begin{aligned} \hat{Z}(\boldsymbol{t}) &= \begin{bmatrix} \boldsymbol{v}(\boldsymbol{t})^t & \boldsymbol{B}(\boldsymbol{t})^t \end{bmatrix} \begin{bmatrix} V & F \\ F^t & 0 \end{bmatrix}^{-1} \boldsymbol{z} \\ &= \boldsymbol{B}(\boldsymbol{t})^t \hat{\beta} + \boldsymbol{v}(\boldsymbol{t})^t V^{-1} (\boldsymbol{z} - F\hat{\beta}) \end{aligned}$$

where the superscript t denotes matrix transpose, and

$$\begin{aligned} \boldsymbol{v}^t(\boldsymbol{t}) &= \begin{bmatrix} C(\boldsymbol{t}, \boldsymbol{s}_1) & C(\boldsymbol{t}, \boldsymbol{s}_2) & \cdots & C(\boldsymbol{t}, \boldsymbol{s}_n) \end{bmatrix} \\ \boldsymbol{B}^t(\boldsymbol{t}) &= \begin{bmatrix} B_1(\boldsymbol{t}) & B_2(\boldsymbol{t}) & \cdots & B_p(\boldsymbol{t}) \end{bmatrix} \\ V_{ij} &= C(\boldsymbol{s}_i, \boldsymbol{s}_j), \quad 1 \leq i, j \leq n \\ F_{ij} &= B_j(\boldsymbol{s}_i), \quad 1 \leq i \leq n, \quad 1 \leq j \leq p \\ \hat{\beta} &= (F^t V^{-1} F)^{-1} F^t V^{-1} \boldsymbol{z} \\ \boldsymbol{z} &= \begin{bmatrix} Z(\boldsymbol{s}_1) & Z(\boldsymbol{s}_2) & \cdots & Z(\boldsymbol{s}_n) \end{bmatrix}. \end{aligned}$$

Note that $\hat{\beta}$ is the generalized least squares estimator of β. Discussion of covariance estimation is deferred to Section 2.3.

The mean squared error of prediction is

$$\begin{aligned} MSE(t) &= MSE(t \mid s_1, s_2, \ldots, s_n) && (6) \\ &= E[(Z(t) - \hat{Z}(t))^2 \mid z] && (7) \\ &= C(t,t) - \begin{bmatrix} v(t)^t & B(t)^t \end{bmatrix} \begin{bmatrix} V & F \\ F^t & 0 \end{bmatrix}^{-1} \begin{bmatrix} v(t) \\ B(t) \end{bmatrix}. && (8) \end{aligned}$$

Note that $MSE(t)$ does not depend on the actual observed values z, but only on the observation sites (or design) s_1, s_2, ..., s_n and the covariance. Also, we do not need to know the vector of regression coefficients β, and can compute $MSE(t)$ up to a scale if we specify the covariance up to a scale. If one is willing to assume a Gaussian process for Z, then using the formulae above we can obtain prediction intervals for expressing the accuracy of $\hat{Z}(t)$. For instance, a 95% prediction interval would be $\hat{Z}(t) \pm 1.96\sqrt{MSE(t)}$. These factors make $MSE(t)$ a handy tool for spatial design.

We will of course typically be interested in predicting the process at more than one point. It is useful to know that the conditional covariance of the prediction error at two sites is given by

$$\begin{aligned} & Cov[Z(t_1) - \hat{Z}(t_1), Z(t_2) - \hat{Z}(t_2) \mid z] \\ &= \begin{bmatrix} C(t_1,t_1) & C(t_1,t_2) \\ C(t_2,t_1) & C(t_2,t_2) \end{bmatrix} \\ &\quad - \begin{bmatrix} v(t_1) & v(t_2) \\ B(t_1) & B(t_2) \end{bmatrix}^t \begin{bmatrix} V & F \\ F^t & 0 \end{bmatrix}^{-1} \begin{bmatrix} v(t_1) & v(t_2) \\ B(t_1) & B(t_2) \end{bmatrix}. \quad (9) \end{aligned}$$

For (O–2), it is not hard to show that the BLUP of an integral is

$$\hat{I}(\mathcal{D}_0) = \int_{\mathcal{D}_0} \hat{Z}(t)\, dt, \quad (10)$$

which has mean squared prediction error

$$E\left[\left(\int_{\mathcal{D}_0}\{\hat{Z}(t)-Z(t)\}\,dt\right)^2\,\middle|\,z\right] \tag{11}$$

$$= \int_{\mathcal{D}_0\times\mathcal{D}_0} E[\{\hat{Z}(t)-Z(t)\}\{\hat{Z}(s)-Z(s)\}\,|\,z\,]\,dtds. \tag{12}$$

The situation for (O–3), (O–4), and (O–5) is somewhat more complicated as they are not linear functionals of the sample path. For instance, consider (O–4). A naive estimator of $\sup_{t\in\mathcal{D}_0} Z(t)$ is the maximum of the BLUP, i.e. $\sup_{t\in\mathcal{D}_0}\hat{Z}(t)$, but this estimator will typically underestimate the true maximum with high probability (e.g,, predicting the future maximum of a discrete time Gaussian random walk with no drift).

A general approach for estimation of nonlinear functionals is conditional simulation (Englund and Heravi 1993). We describe a conditional simulation approach for (O-4) assuming a Gaussian process. As pointed out by Handcock and Stein (1993), assuming an improper prior on the coefficient vector β in the mean function, the BLUP gives the conditional mean and (9) gives the conditional covariance. One may then simulate the process conditional on the observations on a very fine grid and obtain from each simulation a maximum on the grid. The empirical distribution of the simulated maxima can be used to approximate the posterior distribution of the true maximum on the grid. The mean or median of the simulated maxima may be used as a point estimate of the true maximum. The 95th percentile of the simulated maxima provides an upper 95% prediction bound on the true maximum on the grid.

For (O–3), the idea is basically the same – compute J for each simulated sample path. One approach for (O–5) is to compute

$$p(t) = P\left[Z(t)>\lambda|\,Z(s_1),Z(s_2),\ldots,Z(s_n)\right]$$

for each $t\in\mathcal{D}$ from the conditional normal distribution of $Z(t)$ given the data (no simulation required). Then use $\{t : p(t) > 1/2\}$ as an estimate of the set $\mathcal{D}(\lambda)$.

2.3 Estimation of the Covariance Function.

It is only necessary to know the covariance up to scale factor in order to compute the BLUP. For instance, if one uses the covariance function C_1 in (3), then it is necessary to know the θ_i's but not σ^2 to compute $\hat{Z}(t)$. One will typically not be willing to specify values a priori for the covariance parameters. The most commonly used methodology for getting around this difficulty is to estimate the unknown covariance parameters and then plug these estimated values into the BLUP computation. Classically, the covariance (or equivalently the variogram) has been estimated by method of moments, but various difficulties arise (pp. 69–70 of Cressie 1991). With modern computing, it is easy to fit a parametric model such as those in (3) through (5) by maximum likelihood (assuming a Gaussian model) (Sacks et al. 1989). A nonparametric approach to covariance estimation, presented by Keh-Shin Lii, avoids prespecification of the covariance model by estimating the spectral density (Chow and Lii 1992). Chow, Lii, and Fujioka (1992) illustrate the method using fire weather data.

As pointed out by Le and Zidek (1992), "classical kriging methodology fails to incorporate uncertainty about the covariance in the prediction error variance. This deficiency leads to ... seemingly valid decisions or regulatory actions which are ... unjustified." These authors present a conjugate prior Bayesian approach which does include uncertainty about the covariance. However, their analysis is restricted to a preselected finite grid of sites and does not extend to a spatial continuum. Handcock and Stein (1993) position priors on the covariance in a manner more consistent with spatial statistics and give further results on the coverage probability of prediction intervals. Qian (1995) applies this technique to an environmental water quality problem.

3 Sampling Design.

Classically, spatial sampling designs have either been derived from probability sampling or from "geometric" designs. Optimal design in the spatial setting has only recently received much attention. Other sampling design strategies are in use as well.

3.1 Probability Sampling.

In the probability sampling approach the domain of interest is taken as the "population," and it may be further split into "strata," etc. See Provost (1984) or Chapters 4 through 9 of Gilbert (1987). Of course, the inference from data collected by a probability sampling design can be design based, but can also be model based. For instance, to estimate mean or total contamination as in (O–2) from a "simple random sample" (i.e., sites selected independently and uniformly in $\mathcal{D}_0$), the mean concentration of the sample would be an unbiased estimator (w.r.t. the design) of the mean concentration of the region. A model based estimator would involve kriging the data (which utilizes the locations of the sites as well as the observed concentrations at the sites). If one uses a nonrandom design, then only model based inferences are valid.

As mentioned in the introduction, the National Resources Inventory survey is based on a probability sample. The 1992 NRI used a stratified 2 stage area sample. Strata are essentially 1/3 of a township – a 2×6 mile region. Strata were subdivided into 0.5×0.5 mile Primary Sampling Units (PSU's). In all, approximately 300,000 PSU's were selected for the 1992 NRI. Sampling rates vary somewhat and are determined by landscape/use characteristics, with special consideration for regions that don't conform to the standard Public Land Survey, irrigated regions in the West, and Alaska. The nominal rate involves drawing 2 PSUs at random per stratum and then 3 points selected randomly within each selected PSU. The 3 points are not selected independently, however: the first point is randomly selected within a PSU;

then the other two points are selected in a manner which separates them from the first point and from each other. The data collected include some PSU-wide data and detailed data collected at the 3 points. Both the design and the selected PSUs have remained relatively constant since the 1982 survey. Special modules are appended to the survey in particular areas such as those where land use is known to have changed significantly.

Some samples for the U.S. Environmental Monitoring and Assessment Program will be collected at long term research sites selected purposively based on their ecological importance rather than randomly. The probability sample based portion of EMAP is designed as follows. The EMAP sampling strategy is multistage and involves aspects of stratified and systematic sampling. The first stage of the design is a triangular grid covering the conterminous United States. The grid is randomly situated over the U.S. land mass; the inter-point distance along the grid is approximately 27 km and the ratio of area to number of grid points is approximately 635 km^2 per grid point. The grid design is appropriate to the task of measuring ecological resources whose position does not change over the time of the survey and to the need for repeated sampling and reporting over time. The multistage design permits tailoring a design to the resource(s) of interests and purposes of the reporting. Data may be collected at a random sample of grid points during the first stage, on the basis of which informed choices for the definition, stratification, etc. of second and lower stage units can be made. In preparation for the second stage, a randomly placed hexagonal template is constructed over the region. The typical size of the template is sixteen hexagons per grid point.

EMAP distinguishes between two types of ecological resources: extensive and discrete. Extensive resources include large rivers and streams, and forests and Great Lakes, represented by lengths and areas, respectively. Discrete resources include small lakes, represented as points. A first stage sample of grid points identifies a

corresponding set of resources containing or coincident with the grid points. At the second stage, the centers of randomly selected hexagons are used to identify a random sample of discrete resources or a random set of locations for collecting data on extensive resources. As reporting is typically done on a regional basis, it is undesirable to oversample spatially contiguous areas. This is partially controlled by organizing first stage units within groups of contiguous hexagonal templates, randomizing the order of groups, templates, and first stage grid points separately, and applying random-start systematic sampling to the resulting linear list of first stage units. See Stevens (1994) for a complete description of the EMAP probability based design.

3.2 Geometrical Designs.

Probability sampling designs offer the advantages of simplicity and robustness (from the lack of dependence on a model), but can suffer from a lack of efficiency. According to Cressie (p. 319, Cressie 1991), "Regular (random and nonrandom) sampling plans are usually more efficient than simple and stratified random sampling ... where efficiency here is measured in terms of the average and maximum kriging variances."

Many practitioners of geostatistical environmetrics have used regular or geometric designs such as a triangular, square, or hexagonal grid. See for instance the "bull's eyes" in Figure 5 of Flatman (1984), which center on the observation sites, as noted in Figure 1 of Journel (1984), indicating the use of a square sampling grid. In a study by Olea (1984), two design criterion functions (IMSE and MMSE, introduced below) were compared for various geometric designs and some probability sampling designs. His results suggest that the hexagonal tiling design is universally best (among those he considers) for both criteria. Yfantis, Flatman, and Baher (1987) found that equilateral triangles provide the most efficient design for kriging among regular triangles, rectangles and hexagons. An historical example of an optimal geometric

design (in a very special setting) may be found in Dalenius, Hajek, and Zubrzycki (1960).

3.3 Optimal Designs.

Given a spatial model as in the previous section, and given a (linear) method of prediction, one can formulate various optimal design criteria. We state here only design criteria that admit a reasonable mathematical formulation. Assume that $\hat{Z}(\boldsymbol{t})$ is a linear unbiased predictor of $Z(\boldsymbol{t})$, which may or may not be the BLUP. Then the mean squared error of the prediction depends only on the covariance of the process. We assume that the covariance has been specified, at least up to a scale factor. We will now enumerate several design criteria.

(D–1) The Integrated Mean Squared Error (IMSE) criterion is to minimize

$$IMSE(\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n) \;=\; \int_{\mathcal{D}} E[\{\hat{Z}(\boldsymbol{t}) - Z(\boldsymbol{t})\}^2]\, d\boldsymbol{t}.$$

This may be modified by including a weight function in the integral or integrating over another domain.

(D–2) The Maximum MSE (MMSE) criterion is to minimize

$$MMSE(\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n) \;=\; \sup_{\boldsymbol{t} \in \mathcal{D}} E[\{\hat{Z}(\boldsymbol{t}) - Z(\boldsymbol{t})\}^2].$$

(D–3) The entropy criterion is to maximize

$$H\,(\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n) \;=\; E\,[-\log\{f(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2), \ldots, Z(\boldsymbol{s}_n))\}]\,,$$

where f denotes the probability density function of $(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2), \ldots, Z(\boldsymbol{s}_n))$.

We will speak of an IMSE optimal design to be one which minimizes the criterion in (D–1), and similarly for the other criteria. In most cases, the computation of

an actual design will require numerical optimization, and evaluation of any of the criterion functions for a specific design will typically involve numerical integration or optimization, so obtaining optimal designs is by no means simple.

The IMSE design criterion (D–1), which minimizes average kriging variance, has been widely used (see e.g., Sacks and Ylvisaker 1985; Sacks, Schiller, and Welch 1988; and Sacks et al. 1989), perhaps because it is easy to formulate and intuitively appealing. Basically, one seeks the design that does "best" on the average, where "best" is measured by smallness of the mean squared prediction error. Similarly, the MMSE design criterion (D–2) seeks to minimize the worst case. See Johnson (1996) for discussion of this criterion and its similarities with the entropy criterion.

The entropy criterion is usually considered in the setting where there are only finitely many sites available. Assuming $Z(t)$ is a Gaussian process, maximization of H is equivalent to maximization of the determinant of the covariance of $\boldsymbol{Z}(\boldsymbol{S})$ (Shewry and Wynn 1987). For this reason, the entropy optimal designs are also known as D–optimal designs in analogy with linear models. In a Bayesian framework, one can include uncertainty about unknown parameters in the entropy criterion (Le and Zidek 1994). The entropy criterion has been used by Shewry and Wynn (1987) and Currin et al. (1991). The proponents of the entropy criterion claim that it is appropriate for situations where there is not a single objective (e.g., Le and Zidek 1994), as is the case with most environmental data.

As emphasized at the workshop by Ylvisaker, one should not only optimize the design for prediction error within the model but should also address the issue of model validation. As discussed previously, classical kriging ignores the error inherent in estimating the variogram and the propagation of this error through the prediction.

3.4 Multi-Phase Sampling.

Englund and Heravi (1995) investigated design optimization for soil remediation, using a linear cost function comprising: 1) per remediation unit remediation cost (clean-up is done, or not, within each of several predetermined remediation units within the site); 2) per sampling unit sampling cost; and, 3) a per remediation unit environmental cost per unit of concentration for unremediated false negatives. Specifically, the region is divided into remediation units of $100\,m^2$. Using a spatial model of contaminant concentration, the total concentration in each remediation unit can be estimated by kriging. If the estimated contaminant concentration exceeds an action level, then the unit is remediated at a remediation cost of \$10,000. Otherwise, there is a cost of \$10,000 per unit of concentration per remediation unit cost for in remediation units which are not remediated. Finally, the cost per sample is \$500. Costs are designed so that false positives and false negatives incur higher cost than correct decisions, and that costs for false postives and false negatives are equal (to \$10,000 per unit) at the action level. One noisy surrogate site model (similar to that encountered with natural trace elements in soil) and one smooth site model (similar to that representing geological structural surfaces) of 19,800 points each were simulated for the analysis. Three alternative sampling approaches, each involving a total of N sampling units, were investigated: one-phase sampling, two-phase sampling ($N_1 < N$ units in the first phase), and N-phase (sequential) sampling.

The one-phase sampling algorithm simply partitions the site into N regions of equal area and samples one point randomly per region. After selecting N_1, the two-phase algorithm first repeats the one-phase algorithm, but only for the first phase sample of N_1 units. The first phase sample values are kriged to estimate soil contamination concentration Z_i and associated standard error s_i in each remediation unit. Contamination concentration in remediation unit i is assumed to follow a triangular

distribution with mean Z_i and half-width $3s_i$, from which the expected loss L_i due to a false negative decision for remediation unit i can be calculated relative to a predetermined action level. The $(N_1 + 1)$st sample unit is selected randomly within the remediation unit with highest expected loss. The estimated kriged standard deviations are updated, and the procedure is repeated to select the next sample unit, etc., until a total of N sample units have been assigned. Final kriged estimates per remediation unit and estimated total remediation cost are then computed. Beginning with default estimates of Z_i and s_i equal to the action level and the population standard deviation, respectively, the N-phase algorithm constructs the first, second, etc., samples in the same manner as the two-phase algorithm, except that the concentration estimates are updated along with the standard deviations as each new sample unit is selected.

Englund and Heravi (1995) report results based on intensive simulation of the three sampling algorithms over the two simulated surrogate site models. They concluded that, for fixed N, better decisions result from more sampling phases, but only marginally so. This, combined with limitations in current technology for real–time analysis of samples and related quality assurance problems, favors few-phase sampling. For one-phase designs, they determined that equal probability sampling is marginally inferior to designs based on iterative minimization of the maximum kriging variance. Major conclusions were that the optimum number of samples, N, is independent of the number of phases, and that best results are obtained when 70-80% of the samples are allocated in the first phase of a two-phase design (indeed, designs based on only 10-20% of the samples were deemed counter-productive). This enables a simple rule of thumb approach to cost optimization: first, select N based on cost optimization for a one-phase design; second, allocate 70-80% of the N units to the first phase. Englund and Heravi (1993) report an earlier approach to these problems involving site-specific models based on conditional simulation.

3.5 Adaptive Sampling

Adaptive sampling methods (Thompson 1992) are useful for characterizing spatial phenomena that are clustered, or rare, or both. A typical example is to estimate population size within a region of animal species or ecological resource that may be rare but tends to cluster. Adaptive sampling begins here with a first phase sample, such as a random sample (or complete enumeration) within each unit of a simple random sample of equal area subregions. If the phenomenon is observed in a sample region, neighboring regions are added to the sample and examined. Based on theory developed by Thompson (1992), revised probabilities of selection for all included units can be computed, leading to consistent estimators of population size. In many cases, these estimators exhibit efficiencies many times greater than that of simple random sampling (with the equivalent total number of units).

In addition to ecological resource estimation, adaptive sampling applies naturally to estimation of communicable disease. Here, relationships may be spatial not only in the traditional (geographic sense) but also in terms of familial or occupational relationships. It appears to us that adaptive sampling could be used in conjunction with trajectory-based optimization for hot-spot identification and estimation in environmental applications. However, the only application of this sort of which we are aware is that of Johnson (1996), discussed below (and the details of the adaptive design reported therein are sketchy). It may be that current theory needs to be extended (e.g., adaptive sampling combined with line-transect sampling) to further use of adaptive sampling in the spatial context.

3.6 Other Design Strategies.

Still other categories of designs are "judgemental" and "hybrid" designs. In pp. 101-102 of Barth and Mason (1984), there is a good discussion of using random,

judgmental, and systematic design methodologies in a study. The example presented there concerns leaching of a contaminant from a known point source into ground water used for irrigation of crop land. The authors recommend establishing a radial grid system centered at the source, selection of rays emanating from the source at regular angular intervals (as in a systematic design) with extra such rays aligned with the hydraulic gradient (incorporating one's judgment about the direction the contaminant will most likely move), and random selection of sites along the selected rays.

Bayesian methods may also be employed. Qian (1995) applies Bayesian kriging to water quality assessment. Johnson (1996) combines Bayesian and geostatistical methods with adaptive sampling in search of low cost designs for waste site characterization.

In practice, it is sensible to employ a mixture of design strategies as a hedge against misspecification of the spatial model and to provide necessary data to estimate and investigate the covariance on a variety of scales. The design constructed by Mitchell (West et al. 1993) for estimating the pre-remediation distribution of VOC's at a hazardous waste site may be called a hybrid design. It includes a large scale geometric design of more or less regularly spaced bore holes combined with a "local" design of more closely spaced bore holes. This sampling plan facilitated modeling of local, midrange and longrange variability.

3.7 Extracting Representative Soil Samples.

In population sampling, "representative sample" is synonomous with "probability sample." In environmental and other spatial applications, the representativeness of probability designs can be overshadowed by lack of information on correlation structure. In materials sampling such as soil sampling for waste site characterization, rep-

resentativeness is often more in the eye of the expert (e.g., the risk assessor) than in a statistical formula, and judgmental sampling may be favored (although this situation is improving, see, e.g. US EPA (1994)). The main obstacles to obtaining representative samples in soil sampling are: area and depth heterogeneity, physical extraction of the sample (e.g., core diameter), mixing of extracted core (e.g., repeated quartering), quantity sent to laboratory and quantity analyzed, analytical method used (e.g., dissolved, filtered), and lack of understanding of the effects of spatial variability on the part of risk assessors. The three principal methods for extracting representative soil samples are to sample "hot spots" only, Gy sampling, and composite sampling.

Hot spots are often selected using expert or prior knowledge, such as knowldege of sources of the contamination or topography or by visual inspection. This may be augmented by random grid sampling. A pattern for the contamination (e.g., elliptical) may be assumed, as well as the relative size of hot spots to grids. To the extent that true hot spots are located, hot spot sampling addresses the problem of area heterogeneity. Hot spot sampling is simple to execute, often yields a large number of samples, and is supported by well-documented procedures (Gilbert 1987). However, it can be costly, and, for Superfund applications, results must be reinterpreted in terms of average contamination.

Gy sampling focuses on the physical aspects of extracting soil samples, including particle size (Gy 1982; Pitard 1989). Gy's theory is based on a hierarchy of sampling errors, e.g., delimitation error, extraction error and preparation error. Delimitation error arises from physical collection and particle size determination; extraction error includes factors such as sampling from moving media; and, preparation error deals with errors introduced in processes such as crushing, drying and screening physical samples. Gy sampling addresses the problems of mixing extracted core, quantity sent to laboratory, and amount extracted for analysis but may suffer from empirical biases and assumptions about particles.

Composite sampling involves the mixing of several physical samples prior to extraction of the final physical sample for laboratory analysis. This sampling strategy addresses area and depth heterogeneity and problems of mixing the extracted core. Boomer, Erickson, Swanson, Kelso, Cox, and Schultz (1985) and Boswell and Patil (1990) concluded that approximately square groups should be used to create composites. The interplay between the physical and spatial aspects of soil sampling was demonstrated in a simulation study of Radium-226 contamination by Williams, Leggett, Espegren, and Little (1989): consistent with theory, composite sampling demonstrated improved performance over random sampling, but less than theory would predict.

4 Research Issues.

The preceding discussion raises several issues and questions for research in spatial design and analysis.

4.1 Design.

Geometric and other regularly spaced spatial designs have been studied for several decades (see, e.g., Dalenius et al. 1960). These designs, including space-filling designs (Johnson, Moore, and Ylvisaker 1990) are optimal for some spatial problems. It is conjectured that such designs may be "good enough" for many problems, particularly when combined with trenchant methods for spatial analysis. If this is true, then regular designs would be robust over a range of applications and evaluation criteria. These notions need to be quantified and evaluated. Specific research questions include evaluating the robustness of regular designs across different criteria (e.g., maximum or average concentration, entropy) and in specific environmental settings such as

hazardous waste site characterization and air quality monitoring. Also important is the ability of regular designs to detect change, i.e., trends.

Asymptotic designs are based on paradigms for the asymptotic behavior of certain criteria, e.g., the expected number of samples per unit area remains constant, or goes to zero, or goes to infinity; the ratio of signal variance to error variance remains constant, or goes to zero, or goes to infinity. It would be useful to identify environmental problems that fit such paradigms (e.g., signal variance to error variance goes to zero as size of region increases), and to determine if there are asymptotic designs that are optimal or "good enough" for environmental applications.

Sequential sampling including multi-phase sampling (Englund and Heravi 1995) and adaptive sampling (Thompson 1992) have been applied successfully to environmental problems including waste site characterization and ecological monitoring. Applications of these techniques to air and water quality monitoring should prove fruitful and instructive. Related questions include: What is the role for sequential sampling in the evaluation of regular and other fixed location designs? How can fixed location and sequential designs be combined for environmental applications?

Designs that are optimal with respect to standard criteria (e.g., maximum or average concentration) under standard covariance models (e.g., spherical, Ornstein-Uhlenbeck process) need to be identified and their capabilities and limitations studied. The applicability of space-filling and other process optimization designs to environmental monitoring and assessment problems and the robustness of these designs needs to be investigated. Further, the use of composite sampling, Gy sampling, and other methods for physically constructing samples warrants additional scientific scrutiny.

Furthermore, what role do covariates play in determining optimal designs and effective sample sizes? In general, a suite of contaminants are measured in conjunction with physical site characteristics and meteorological variables, e.g., soil porosity, direction of ground water flow, predominant wind direction, temperature and the location

of the source of the pollutants. An optimal design with respect to one contaminant is not necessarily optimal with respect to all contaminants. How can covariate information assist in determining an optimal and robust design for the primary pollutants?

4.2 Evaluation.

Sources and effects of bias in environmental sampling need to be identified and studied. Techniques for examining correlation structure to determine the effective sample size of a design are needed. The issues surrounding combining design-based (e.g., regular designs) and model-based (e.g., conditional simulation) approaches in spatial design and analysis need to be stated and examined. For example, can soil pollution concentrate data collected on different designs be combined across all or a sample of the hazardous waste sites in a region to provide meaningful regional pollution characterization and remediation cost information?

Most environmental spatial designs will have multiple unforeseen future uses. For example, data from monitoring networks designed for regulatory purposes to detect local maxima are routinely used to model average environmental exposure. Such model biases need to be identified and corrected for, if possible. Robustness of standard spatial designs and analysis procedures across standard evaluation criteria need to be quantified. Methods to evaluate the robustness of designs against realistic changes to covariance structure are needed. These issues are of paramount importance.

Correlation structures are often assumed or approximated in spatial environmental models: estimates of bias or demonstration of model robustness due to such errors are needed. Including uncertainty in the covariance model through Bayesian methods warrants further exploration.

How useful is it and when is it necessary to include replication in environmental

monitoring designs or in designs for evaluating monitoring designs? How useful and when should the precise location of samples be "jiggled" through local randomization in order to capture small scale spatial variation?

Are entropy criteria useful in creating designs that will be robust against other evaluation criteria? It would be beneficial to identify a category of model based designs that performs acceptably across a range of criteria and applications, analogous to the efficiency of equilateral triangular designs among regular designs (Cressie, Gotway, and Grandona 1990; Yfantis et al. (1987)). Can entropy measures be used to determine effective sample size?

4.3 Computation.

Improved computational methods (e.g., for all-subsets kriging) need to be developed to facilitate the evaluation and fine-tuning of spatial designs. Designs that are suboptimal but robust and easy to compute need to be identified.

The computational experience of the first author and others (Sacks et al. (1989)) in searching for optimal IMSE designs supports the results of Olea (1984), namely, that good designs tend to spread points uniformly in the design region. Using standard optimization algorithms with a random initial design, the algorithm usually makes substantial improvement over the initial design with the first few iterations, but quickly reaches a broad valley in the high dimensional search space and takes an extremely long time to converge to the optimal design. Investigation of the progress of the algorithm suggests that the initial gains are made by moving apart points that are too close together, moving some points into regions where the initial random design is sparse, and making some adjustments for the boundaries. This observation suggests that it is relatively easy to obtain a "good" design, but difficult to get the "best" design. The robustness study reported in Sacks et al. (1988) show that some-

times optimal designs for one covariance may work well across a range of covariances. The "good" design may actually prove to be a robust design under variations in the original assumptions.

What is the role of mathematical optimization methods such as dynamic and linear programming in problems such as optimal sequential spatial design and all-subsets kriging? Christakos and Killam (1993), for example, illustrate the use of simulated annealing for spatial sampling design.

What is the role of geographic information systems (GIS) for sample stratification, design, optimization and selection (e.g., in the NRI)? Methods, perhaps involving GIS, for incorporating location and auxiliary and conditional information should be investigated. What is the role of computer graphics and the human/computer interface in spatial sample design for environmental applications? How useful are interactive graphics and GIS as tools for spatial design? One use of these visualization and representation tools would be to model location(s) of pollution source(s) and the probable extent and shape(s) of the plumes. Ensor, Scott, Boeckenhaurer, and Bedient (1995), for example, demonstrate how graphical methods facilitate ground water site characterization and sampling.

Acknowledgments.

Thanks are due to the participants of the 1994 Workshop on Spatial Sampling for the Environment, whose thoughtful comments stimulated many of the issues examined in this article. The work of D. Cox and K. Ensor was supported by the USEPA Cooperative Agreement #CR819638-01-0 with the National Institute of Statistical Sciences. K. Ensor also received funding under USEPA Cooperative Agreement #CR821906. The information in this article has been funded wholly or in part by the United States Environmental Protection Agency. It has been subject to

Agency review and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Barth, D.S., and Mason, B.J. (1984), "Soil Sampling Quality Assurance and the Importance of an Exploratory Study," in *Environmental Sampling for Hazardous Wastes*, Schweitzer, G.E., and Santolucito, J.A., editors, American Chemical Society, Washington, D.C., 97–104.

Boomer, B.A., Erickson, M.D., Swanson, S.E., Kelso, G.L., Cox, D.C., and Schultz, B.D. (1985), "Verification of PCB Spill Cleanup by Sampling and Analysis." EPA Contract 68-02-6938, TC 8501-A37.

Boswell, M.T., and Patil, G.P. (1990), "Composite Sampling Using Spatial Autocorrelation for Palmerton Hazardous Waste Site: A Preliminary Report." *Proceedings of the Workshop on Superfund Hazardous Waste: Statistical Issues in Characterizing a Site: Protocols, Tools, and Research Needs.* Lacayo, H., Nadeau, R.J., Patil, G.P. and Zaragoza, L., editors. pp20-42.

Christakos, G. and Killam, B. (1993), "Sampling Design for Classifying Contaminant Level Using Annealing Search Algorithms," *Water Resources Research*, **29**, 4063-4076.

Chow, I.-S., and Lii, K.-S. (1992), "Spectral Analysis of Random Fields with Random Sampling," *Probabilistic and Stochastic Methods in Analysis with Applications,* Byrnes, J.S., editor, Kluwer, Amsterdam.

Chow, I.-S., Lii, K.-S., and Fujioka, F. (1992), "Multidimensional Spectral Analysis of Monthly Fire Weather Patterns," *5th International Meeting on Statistical Climatology,* Toronto, June 22-26, 1992, 251-254.

Cressie, N. (1991), *Statistics for Spatial Data,* Wiley, New York.

Cressie, N., Gotway, C. A., and Grandona, M. O. (1990), "Spatial Prediction from Networks," *Chemomectrics and Intelligent Laboratory Systems* **7**, 251–271.

Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D. (1991) "Bayesian Prediction of Deterministic Functions, with Applications to the Design and Analysis of Computer Experiments," *J. Amer. Statist. Assoc.*, **86**, 953–963.

Dalenius, T., Hajek, J., and Zubrzycki, S. (1960), "On plane sampling and related geometrical problems," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume I,* University of California Press, Berkeley, California, 125–150.

Englund, E.J., and Heravi, N. (1993), "Conditional Simulation: Practical Aplications for Sampling Design Optimization," in *Geostatistics Troia '92*, Vol. 2, A. Soares, ed., Kluwer, Dordrecht, 613–624.

Englund, E.J., and Heravi, N. (1995), "Multi–phase Sampling for Soil Remediation," to appear, *Environmental and Ecological Statistics 2.*

Ensor, K. B., Scott, D. W., Boeckenhaurer, R., and Bedient, P. B. (1995), "Visualization of Groundwater Contamination Data," *Proceedings of the 27th Symposium on the Interface Between Computing Science and Statistics,* to appear.

Flatman, G.T. (1984), "Using Geostatistics in Assessing Lead Contamination Near Smelters," in *Environmental Sampling for Hazardous Wastes*, Schweitzer, G.E.,

and Santolucito, J.A., editors, American Chemical Society, Washington, D.C., 43–52.

Gao, F., Sacks, J., and Welch, W. (1994), "Predicting the urban ozone levels and trends with semiparametric modeling," Technical Report 14, National Institute of Statistical Sciences, Research Triangle Park, N.C.

Gilbert, R.O. (1987), *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinhold Co., New York.

Gy, P. M. (1982), *Sampling of Particulate Materials*, Elsevier, New York.

Handcock, M.S., and Stein, M.L. (1993), "A Bayesian Analysis of Kriging," *Technometrics*, **35**, pp. 403-410.

Johnson, R. L. (1996), "A Bayesian/Geostatistical Approach to the Design of Adaptive Sampling Programs," *Geostatistics for Environmental and Geotechnical Applications, ASTM STP 1283*, Srivastava, R., Rouhani, S., Cramer, M., and Johnson, V., editors, American Society for Testing and Materials, Philadelphia (in press).

Johnson, M., Moore, L., and Ylvisaker, D. (1990), "Minimax and Maximin Distance Designs," *J. Statist. Planning and Inference*, **26**, 131-148.

Journel, A.G. (1984), "New Ways of Assessing Spatial Distributions of Pollutants," in *Environmental Sampling for Hazardous Wastes*, Schweitzer, G.E., and Santolucito, J.A., editors, American Chemical Society, Washington, D.C., 109–118.

Le, N.D., and Zidek, J.V. (1992), "Interpolation with Uncertain Spatial Covariances: A Bayesian Alternative to Kriging," *J. Multivariate Analysis*, **43**, 351-374.

Le, N.D., and Zidek, J.V. (1994), "Network Designs for Monitoring Multivariate Random Spatial Fields," *Advances in Probability and Statistics*, edited by J. Perez Vilaplana and M.L. Puri, 191-206.

Olea, R.A, (1984), "Sampling Design Optimization for Spatial Functions," *Mathematical Geology*, **16**, 369-392.

Pitard, F.F. (1989), *Pierre Gy's Sampling Theory and Sampling Practice*, Vol. I and II, CRC Press, Boca Raton, FL

Provost, L.P. (1984), "Statistical Methods in Environmental Sampling," in *Environmental Sampling for Hazardous Wastes*, Schweitzer, G.E., and Santolucito, J.A., editors, American Chemical Society, Washington, D.C., 80–96.

Qian, S. S. (1995), "Computing the Area Affected by Phosphorous Runoff in An Everglades Wetland Using Bayesian Kriging," *Proceedings of the 27th Symposium on the Interface Between Computing Science and Statistics,* to appear.

Sacks, J., Schiller, S., and Welch, W. (1988), "Designs for Computer Experiments," *Technometrics*, **31**, 385–399.

Sacks, J., Welch, W., Mitchell, T., and Wynn, H. (1989), "Design and analysis of computer experiments (with discussion)," *Statistical Science* , **4** , pp. 409-435.

Sacks, J., and Ylvisaker, D. (1985), "Model Robust Design in Regression: Bayes Theory," *Proceedings of the Berkeley Converence in Honor of Jerzy Neyman and Jack Kiefer*, eds. L.M. LeCam and R.A. Olshen, Wadsworth, vol. II, pp. 667-679.

Shewry, M., and Wynn, H. (1987), "Maximum Entropy Sampling," *J. Applied Statist.*, **14**, 165-207.

Spiegelman, C.H., and Dettner, S. (1993), "Multivariate Chemometrics: Applying and Developing Receptor Models for the 1990 El Paso Winter PM_{10} Receptor Modeling Scope Study," Chapter 24, in *Multivariate Environmental Statistics*, Patil, G.P. and Rao, C.R., editors, Elsevier Science Publishers BV.

Stevens, D.L. (1994), "Implementation of a National Monitoring Program," *Journal of Environmental Management*, **42**, 1-29.

Thompson, S.K. (1992), *Sampling*, John Wiley and Sons, New York.

US EPA (1994), *Statistical Methods for Evaluating the Attainment of Cleanup Standards*, Vol. 3: Reference–Based Standards for Soils and Salient Media, U.S. Environmental Protection Agency, Washington, DC.

West, O.R., Siegrist, R.L., Mitchell, T., Pickering, D.A., Muhr, C.A., Greene, D.W., and Jenkins, R.A. (1993), "X231-B Technology Demonstration for In Situ Treatment of Contaminated Soil: Contaminant Characterization and Three Dimensional Spatial Modeling," Oak Ridge National Laboratory Technical Report No. ORNL TM–12258.

Williams, L.R., Leggett, R.W., Espegren, M.L., and Little, C.A. (1989), "Optimization of Sampling for the Determination of Mean Radium-226 Concentration in Surface Soil." *Environmental Monitoring and Assessment,* **12**, 83-96.

Yfantis, E.A., Flatman, G.T., and Baher, J.V. (1987), "Efficiency of Kriging Estimation for Square, Triangular and Hexagonal Grids," *Journal of Mathematical Geology* **19**, 183-205.