Sequential Sampling to Contour an Uncertain Function

Daniele Veneziano and Peter K. Kitanidis

Sampling exploration of uncertain functions to locate critical contour levels is most effective if sampling decisions are made sequentially. A simple sequential exploration strategy, based on pseudo-Bayesian second-moment analysis, is proposed and compared with non-sequential systematic sampling. Repeated application to functions simulated pseudorandomly from stationary random processes on the line and on the plane indicates uniform superiority of the sequential strategy. The method is particularly advantageous when the function of interest, f(X), has an uncertain trend, and in general when the random process that quantifies prior uncertainty on f(X) is highly correlated.

KEY WORDS: Geostatistics, resource estimation, exploration, sampling.

INTRODUCTION

A frequent objective of mineral exploration is to locate those portions of an ore body which have a grade higher than that which supports profitable mining (grade cutoff or pay limit). In the case of coal mining, the objective is more often to identify areas where the thickness of the coal layer exceeds a certain minimum value which makes excavation possible. These objectives of exploration are especially significant when the grade (thickness) of large portions of the ore body is lower than the cutoff value. In this then economically advantageous to practice some form of selective mining instead of mining all the mineralized regions comprising the ore body.

If selective mining is practical, then one can use sampling exploration to locate the boundaries and to estimate the average grade and the tonnage of each block that can be profitably mined. The problem of estimating the average grade and the tonnage of each block given point samples is widely treated in the ge-
statistics literature (e.g., Krige, 1978; Journel and Huijbregts, 1978), whereas
the problem of devising efficient sampling strategies to contour uncertain func-
tions (grade, thickness) has received little attention and is the subject of this work.

Assume that the objective of sampling exploration is to contour an uncertain
function-state of nature \( h(x) \) at a given level \( \ell(x) \). \( x \) is the location vector
of a point in \( n \)-dimensional space \( \mathbb{R}^n \) \( (n \leq 3) \) and contouring is inside a finite
region \( D \) of \( \mathbb{R}^n \). With reference to the previous examples, \( h(x) \) could be the
grade of an ore of a mineral product, or the thickness of a mineralized layer at
\( x \) (\( x \) is a point of the soil in the first case, a point of the geological plane
in the second case). Then, \( \ell(x) \) is the minimum value of \( h(x) \) that makes excavation
advantageous, that is, the grade or the thickness cutoff. Sampling to contour an
uncertain function at a given level is referred to here as "contouring exploration."

Contouring exploration is of interest in areas other than mining: dams, slopes,
estuaries, and other earth structures are safe if shear resistance of the soil,
\( h(x) \), is not smaller than a critical value \( \ell(x) \) over extended regions of space.
Weak regions may cause failure unless the soil is upgraded \( h(x) \) increased
or the design is modified \( \ell(x) \) reduced. Other applications are to pollution
control of surface water bodies, aquifers, or the atmosphere. In this last case, if
the degree of contamination is time dependent, then \( x \) should include time as
one of its components.

A fundamental aspect of contouring exploration is that sampling decisions
are best made sequentially. Therefore, instead of looking for an optimal-sampling
network, one should identify an optimal-sampling strategy which generates different
networks of observation points, depending on the function \( h(x) \).

In the method that is proposed herein, information on \( h(x) \) is given ini-
tially through a prior random field. After each sample, the random field is modi-
fied through Bayes' theorem and the posterior field is processed by a decision
rule which either locates an additional sample point in \( D \) or else discontinues
exploration. The procedure is not optimal in a theoretical decision sense (any
simple procedure would seemingly be very complicated) but is simple, versatile,
and robust. More sophisticated strategies are possible; some were actually tested
by the writers, but the improvement in performance did not seem to justify the
increase in computation.

Contouring at level \( \ell(x) \) produces a bicolored map: one color where
\( h(x) > \ell(x) \) and a different color where \( h(x) < \ell(x) \). The problem can readily
be generalized to maps with \( K \) colors \( (K > 2) \) if \( K - 1 \) levels are given,
\( r_1(x) < r_2(x) < \cdots < r_{K-1}(x) \). Throughout exploration and at the end of it, the
map is random in the sense that the explorer is uncertain about the color at each
point of \( D \). The purpose of sampling is to reduce this uncertainty until the
expected benefit of continuing exploration is exceeded by its cost. Mapping errors
and measures of contouring uncertainty are defined in the next section. There-

after, rules of Bayesian updating that are relevant to the problem are given. Using
these definitions and rules, a strategy of sequential sampling is proposed.
Statistical evaluation of the strategy is limited to bicolored maps \( (K = 2) \) on the
real line \((n = 1)\) but extension to multicolored maps \( (K > 2) \) poses no special
problem. An example of contouring exploration in the plane \((n = 2)\) is also pre-
sented. Evaluation is through numerical simulation; that is, a repeated
application of the procedure to functions \( h(x) \) generated as independent realizations
of the prior random process. Exploration statistics of the sequential rule
can be shown to be very favorably similar to statistics when the realization functions
are sampled at regular intervals.

ACCURACY OF CONTOUR MAPS

Given a function state of nature, \( h(x) \), a threshold function \( \ell(x) \), and a
region of contouring exploration \( D \), let the "excursion set \( D^* \), the "contouring set
\( D^c \), and the crossing set \( D^c \), be defined as

\[
D^* = \{ x : h(x) > \ell(x), x \in D \} \\
D^c = \{ x : h(x) < \ell(x), x \in D \} \\
D^c = \{ x : h(x) = \ell(x), x \in D \} = D - D^c - D^c
\]

Typically, \( D^c \) is a network of \((n - 1)\)-dimensional surfaces in \( \mathbb{R}^n \),
e.g., a network of lines in the plane, with zero Euclidean measure in \( \mathbb{R}^n \). If this is the case then \( D^c \) contains no external point of \( h(x) \), then \( D^c \) is also the boundary set of
\( D^c, D^c = \partial D^c \), and of \( D^c, D^c = \partial D^c \). It follows that the contouring problem can be
stated equivalently as a "excursion problem" (find \( D^* \)), or as a "excursion problem" (find \( D^c \)), or as a "crossing problem" (find \( D^c \)). This obvious equivalence
does not imply absence of a preferred formulation. In fact, the objective of
contouring exploration is typically to discover relatively rare "excursions or
"excursions, where by \( \ell(x) \) "excursion we mean a connected and isolated subset
of \( D^c \), see Fig. 1. In the mathematical formulation one may, without
losing generality, restrict attention to relatively rare "excursions; hence to the
estimation of their number, \( N^* \), and of their boundaries, \( \partial D^c \) \((n = 1, \ldots, N)\).

![Fig. 1. Excursions and "excursions in
the plane.](image-url)
Upon termination of sampling, the actual map \( M = (D^*, \hat{D}^*) \) is estimated as \( \hat{M} = (\hat{D}^*, \tilde{D}^*) \), where \( \tilde{D}^* \) is the union of \( \hat{N}^* \) estimated "excursions, \( \hat{D}_i \) \( (i = 1, \ldots, \hat{N}^*) \). The terminology that follows will be useful in quantifying the degree of similarity between \( \hat{M} \) and \( M \).

1. A "excursion \( D_i^* \) is said to be detected by \( \hat{M} \) if \( D_i^* \cap \hat{D}_i \neq \emptyset \) (Fig. 2a);
2. \( D_i^* \) is said to be identified by \( \hat{M} \) if \( D_i^* \cap \hat{D}_j \neq \emptyset \) for some \( j \) and \( D_i^* \cap \hat{D}_k = \emptyset \) \( \forall k \neq j \) (Fig. 2b);
3. \( D_i^* \) is said to be isolated by \( \hat{M} \) if \( D_i^* \cap \hat{D}_j = \emptyset \) for some \( j \) (Fig. 2c);
4. \( D_i^* \) is said to be matched by \( \hat{M} \) if \( D_i^* \cap \hat{D}_j \neq \emptyset \) for some \( j \) (Fig. 2d).

If an excursion is matched, then it is also isolated, if it is isolated then it is also identified, and if it is identified then it is also detected. Hence, if \( N^*_1, N^*_2, N^*_3, \) and \( N^*_4 \) are, respectively, the number of detected, identified, isolated, and matched "excursions, then the ratios

\[
\gamma_i = \frac{N_i^*}{\hat{N}^*}, \quad i = 1, 2, 3, 4
\]  

(2)

\( (\gamma_i = 1 \text{ if } \hat{N}^* = 0) \) are in nonincreasing order. Interchange between actual and estimated maps leads to the definition of four other similarity ratios, all analogous to the \( \gamma_i \). They are

\[
\hat{\gamma}_i = \frac{\hat{N}_i^*}{\hat{N}^*}, \quad i = 1, 2, 3, 4
\]  

(3)

with \( \gamma_i = 1 \) if \( \hat{N}^* = 0 \). The maps \( M \) and \( \hat{M} \) coincide if and only if

(i) \( \hat{N}^* = N^* \) (all "excursions in \( D^* \) are discovered), and

(ii) \( \gamma_4 = 1 \) (their locations, \( \partial \hat{D}^* \), are inferred exactly).

(4)

The ratios in Eqs. 2 and 3 quantify the degree of agreement between "excursions in \( \hat{M} \) and \( M \). However, they give little indication about the fraction of \( D \) that is misclassified (unless both conditions of Eq. 4 are satisfied). For this last purpose one should use different indicators. With \( \mu(B) = \int_{\mathbb{R}^n} \mu \) the Euclidean measure of a set \( B \) in \( \mathbb{R}^n \), the following ratios can be taken as measures of mismatch between \( \hat{M} \) and \( M \):

\[
\mu_1 = \mu(D^* \cap \hat{D}^*)/\mu(D^*) = \text{misclassified fraction of } D^*
\]

\[
\mu_2 = \mu(D^* \cap \hat{D}^*)/\mu(D^*) = \text{misclassified fraction of } D^*
\]

\[
\mu_3 = \mu(\hat{D}^* \cap D^*)/\mu(\hat{D}^*) = \text{misclassified fraction of } \hat{D}^*
\]

\[
\mu_4 = \mu(\hat{D}^* \cap D^*)/\mu(\hat{D}^*) = \text{misclassified fraction of } \hat{D}^*
\]

(5)

If \( \hat{M} = M \), then all the \( \mu_i \) are zero. The ratios of Eqs. 2 and 3 and those of Eq. 5 are complementary indicators of similarity between \( M \) and \( \hat{M} \). For example, \( \gamma_i \) and \( \hat{\gamma}_i \) are the same for the pairs of maps in Figs. 3a, b, whereas the \( \mu_i \) are different in the two cases. All these ratios have values between 0 and 1. The higher \( \gamma_i \) and \( \hat{\gamma}_i \) and the lower the \( \mu_i \), the better the agreement between \( M \) and \( \hat{M} \).

**BAYESIAN UPDATING OF UNCERTAINTY**

Denote by \( H(M) \) the random field that quantifies uncertainty on \( h(X) \) at the beginning of exploration, and by \( H_i(X) \) the revised random field after sampling at \( X_1, \ldots, X_i \). Given \( r \) samples, a sequential contouring strategy is a rule that operates on \( H_i(X), \mu(X), \) and \( D \) to produce one of two actions: (1) terminate exploration and estimate \( M \), or (2) continue exploration by sampling at \( X_{i+1} \).

The sampling rule of the next section is very simple in that it operates on \( H_i(X) \) and \( \mu(X) \) only through the dimensionless function

\[
\beta_i(X) = \frac{\mu_i(X) - \mu(X)}{\mu_i(X)}, \quad X \in D
\]

(6)
in which \( m_r(X) = E[H_r(X)] \) and \( \sigma_r^2(X) = \text{var}[H_r(X)] \). Hence, \( \beta_r(X) \) is the algebraic distance of the mean of \( H_r(X) \) from the contouring level \( r(X) \), in units of standard deviations \( \sigma_r(X) \). If exploration terminates after \( r \) samples, then one can use this distance function to estimate \( D_r^* \); for example

\[
D_r^* = \{ X: \beta_r(X) > 0, \quad X \in D \}
\]

(7)

The probability \( P_m(X) \) that such an estimator misclassifies \( h \) at \( X \) depends on the probability distribution of \( H_r(X) \). For a Gaussian posterior field, \( P_m(X) = \Phi[\beta_r(X)] \) with \( \Phi \) cumulative standard normal distribution function. For the same field, \( P_m(X) \) is maximum where \( \beta_r(X) \) is minimum and has value 0.5 on the boundary of \( D_r^* \). The Gaussian misclassification probability is at two different stages of the function in Fig. 4a are shown in Figs. 4c and 4e. Figure 4c is for \( r = 3 \) and corresponds to the distance function \( \beta_r(X) \) in Fig. 4b; Fig. 4e is for \( r = 20 \) and corresponds to \( \beta_r(X) \) in Fig. 4d. These cases were generated while using the sequential sampling rule to be described in the next section.

In the remainder of this section, Bayes' theorem is used to calculate \( \beta_r(X) \) from the first two moment functions of the prior random field and from observations at \( X_1, \cdots, X_r \). The observation variables, \( Z_1, \cdots, Z_n \), are assumed to

be linear random functions of \( h(X) \) at the sampled locations, of the form

\[
Z_i = Z_{o,i} + \beta_i h(X) + \epsilon_i, \quad i = 1, \cdots, r
\]

(8)

in which \( Z_{o,i} \) and \( \epsilon_i \) are given constants (with typical values, \( Z_{o,i} = 0 \) and \( \epsilon_i = 1 \)) and \( \epsilon = [\epsilon_1, \cdots, \epsilon_r]^T \) is a vector of random measurement errors with zero mean and known joint probability distribution.

With only very few exceptions (notably that of jointly Gaussian \( H_0 \) and \( \epsilon \)) calculation of \( m_r(X) \) and \( \sigma_r^2(X) \) is prohibitively tedious. For this reason, and also because the full-distribution properties of \( H_0 \) and \( \epsilon \) may not be known, one may resort to rules that operate only on the first two moments of \( H_0 \) and \( \epsilon \), that is,

\[
m_r(X) = E[H_0(X)], \quad m_r(X, X) = \text{Cov}[H_0(X), H_0(X)]
\]

\[
\Theta_r = E[\epsilon \epsilon^T], \quad \gamma(X, i) = \text{Cov}[H_0(X), \epsilon_i]
\]

(9)

For example, if one uses linear-minimum-variance estimation, or pseudo-Bayesian estimation, or Bayesian estimation with jointly Gaussian \( H_0 \) and \( \epsilon \), or weighted-leastsquares estimation with appropriate weights, one finds in all cases that (Veneziano, 1979).

\[
m_r(X) = m_r(X) + [b_r(X) + \gamma_r(X)]^T[H \Sigma_r H + H \Gamma_r + \Gamma_r^T H + \Theta_r]^{-1}
\]

\[
(Z_r - Z_{o,r} - H m_r)
\]

(10)

and that

\[
\sigma_r^2(X) = [b_r(X) + \gamma_r(X)]^T[\Sigma_r H + H \Sigma_r H + \Gamma_r^T]^{-1}
\]

(11)

in which

\[
b_r(X) = \begin{bmatrix} b_{0,1}(X, X_1) \\ \vdots \\ b_{0,n}(X, X_r) \end{bmatrix}, \quad \gamma_r(X) = \begin{bmatrix} \gamma(X, 1) \\ \vdots \\ \gamma(X, r) \end{bmatrix},
\]

\[
H = \text{diag}(a_i) ; \quad \Sigma_r = \begin{bmatrix} \beta_{0,1}(X, X_1) \\ \vdots \\ \beta_{0,n}(X, X_r) \end{bmatrix},
\]

\[
\Gamma_r = \begin{bmatrix} \gamma(X, i) \\ \vdots \\ \gamma(X, r) \end{bmatrix}, \quad Z_r = \begin{bmatrix} Z_1 \\ \vdots \\ Z_r \end{bmatrix}, \quad Z_{o,r} = \begin{bmatrix} Z_{o,1} \\ \vdots \\ Z_{o,r} \end{bmatrix}, \quad m_r = \begin{bmatrix} m_{0,1}(X_1) \\ \vdots \\ m_{0,r}(X_r) \end{bmatrix}
\]

The first two moments of the random field \( H_0(X) \) are specified on the basis of available prior information. Samples and regional information are utilized to select a model type (the "structure" of the random field) and to estimate its pa-
rameters. In geostatistics, methods have been developed to address these types of model selection and parameter estimation problems (Journel and Huijbregts, 1978). If, relative to the a priori knowledge, the data collected during contouring exploration is highly informative on the structure of $H$, or on its parameters, then the functions $m_{s}(X)$ and $b_{s}(X_{x}, X_{y})$ in Eq. 9 should be periodically updated using all the information available at the time.

It follows from Eqs. 6, 10, and 11 that

$$
\beta_{s}(X) = \frac{\beta_{0}(X) + \epsilon_{j}(X) C_{j}^{-1} \beta_{z}}{(1 - \{1/\alpha(x)^{-1}\} \epsilon_{j}(X) C_{j}^{-1} \epsilon_{j}(X))^{1/2}} \tag{12}
$$

in which

$$
e_{j}(X) = b_{j}(X) - \gamma_{j}(X); C_{j} = \Sigma_{j} + \mathbf{H} \mathbf{H}^{T} + \Gamma_{j}^{T} \mathbf{H} + \Theta_{j}^{T}$$

$$
\beta_{z} = \begin{bmatrix}
\beta_{0}(X) \\
\vdots \\
\beta_{z}(X)
\end{bmatrix}
$$

Of course, $m_{s}$ and $\sigma^{2}_{s}$ in Eqs. 10 and 11 are the posterior mean and the posterior variance functions only under particular distribution assumptions on $H$ and on $e$ (e.g., $H$ and $e$ jointly Gaussian). With this limitation understood, simplicity of language suggests that we keep referring to $m_{s}$, $\sigma^{2}_{s}$, and $\beta_{s}$ as the posterior mean, the posterior variance, and the posterior distance function, respectively. An important feature of Eq. 12 is that the matrix under the inversion sign is not a function of $X$, hence, in the calculation of the function $\beta_{s}(X)$, only one matrix of dimension $(r \times r)$ must be inverted. In the inversion one can take advantage of knowing $C_{j}^{-1}$ from the previous step of exploration. Simplifications are possible if $e = 0$ (noiseless sampling); in particular, Eq. 12 reduces in this case to

$$
\beta_{s}(X) = \frac{\beta_{0}(X) + b_{j}^{T}(X) \Sigma_{j}^{-1} \beta_{z}}{(1 - \{1/\alpha(x)^{-1}\} b_{j}^{T}(X) \Sigma_{j}^{-1} b_{j}(X))^{1/2}} \tag{13}
$$

at all unsampled points. At the location of sampling, $\beta_{s}(X_{x}) = +\infty$ if $h(X_{x}) > r(X_{x})$ and $\beta_{s}(X_{x}) = -\infty$ if $h(X_{x}) < r(X_{x})$.

**SEQUENTIAL SAMPLING STRATEGY**

The first objective of Eq. 4 is subordinate to the second objective in importance and in the chronological sense that excursions can be located only after they have been discovered. This dependence suggests a two-phase strategy of exploration, with phase 1 aiming primarily at the first objective and phase 2 aiming primarily at the second objective. Phases alternate in time as $12121 \ldots$ until no sample point is generated by either phase.

**Phase 1**

Suppose that $r$ observations have been made and that the procedure is in phase 1. Since the objective is to discover relatively rare excursions, one might decide to sample next at the location where $P(h(X) > r(X))$ is maximum. However, this policy is satisfactory only if $\beta_{s}$ is nowhere positive, that is, no excursion has been located yet. In the case when $\beta_{s}$ is nowhere negative (a rare event if excursions are themselves rare) the criterion should be reversed and a sample taken at the point where the same probability is minimum. In fact, such a point maximizes the probability of discovering an excursion and also a new excursion, since it may happen that what appeared initially as a single excursion becomes fragmented and joins to one excursion and two excursions. Considerations of this type suggest the rule: *sample at the point of stationary of $\beta_{s}$ where $|\beta_{s}|$ is minimum.*

The foregoing criterion can be used inside $D$. For points on $\partial D$, it was found convenient to assume that $\beta_{s}$ is stationary at those points, in the direction normal to the boundary (in two or three dimensions, stationarity along the boundary must be satisfied before considering a point of $\partial D$ as a candidate for sampling). This provision activates phase 1 in cases when $\beta_{s}$ is almost constant in $D$ with values close to zero and there is no point of stationarity inside $D$. Sampling continues under phase 1 until the value of $|\beta_{s}|$ at the next sample point exceeds a prescribed maximum value, $\beta_{m}$. When this happens, control passes under phase 2.

In summary, sampling during phase 1 is at points $X_{s+1}$ that satisfy

1. $\beta_{s}$ is stationary at $X_{s+1}$, if $X_{s+1} \in D$, or is stationary along the boundary at $X_{s+1}$, if $X_{s+1} \in \partial D$.
2. $|\beta_{s}(X_{s+1})|$ is the minimum value of $|\beta_{s}|$ over all points that satisfy condition (1).
3. $|\beta_{s}(X_{s+1})| < \beta_{m}$

**Phase 2**

If the conditions of Eq. 14 are not satisfied, (e.g., in the case of Fig. 4b) control is shifted to phase 2, with the objective of refining the boundary of known excursions. If no excursion has been discovered, then exploration terminates. It seems logical, now, to sample at locations where the probability of misclassification is maximum; hence at points of the boundary of $D^*$, since
there \( \beta = 0 \). In order to decide where on \( \hat{D}^* \) to sample, it is observed that if \( \beta \) is a smooth function, then the precision with which the boundary of \( D^* \) is estimated at \( X, X \in \hat{D}^* \) increases with \( |\text{grad } \beta(X)| \). As \( |\text{grad } \beta(X)| \to \infty \), the probability that the actual boundary \( D^* \) crosses any finite neighborhood of \( X \) approaches 1. Uncertainty is largest for \( |\text{grad } \beta(X)| = 0 \). Therefore, a rational sampling rule for phase 2: sample at a point \( X_{*+1} \) that satisfies

\[
\begin{align*}
& (1) \quad \beta(X_{*+1}) = 0 \\
& (2) \quad |\text{grad } \beta(X_{*+1})| = \min_{X; \beta(X) > 0} |\text{grad } \beta(X)|
\end{align*}
\]

(15)

in which \( \beta_m \) is a given constant.

After one sample point has been located by phase 2, control returns to phase 1 for possible discovery of new excursions. The constants \( \beta_m \) and \( \beta_M \) control the accuracy of exploration: they allow one to (indirectly) consider the cost of exploration and to emphasize or deemphasize the importance of discovering new excursions against refining their estimated location.

\section*{EVALUATION}

At the end of exploration, a map estimate \( \hat{M} \) is obtained, for example, from Eq. 7. Because the actual map remains uncertain, one cannot calculate the performance ratios of Eqs. 2, 3, and 5 and the accuracy of \( \hat{M} \) can only be evaluated in probabilistic terms.

As pointed out by Switzer (1969), one should distinguish between the accuracy of a specific estimate \( \hat{M} \) (with \( M \) a random \( \cdots \) map) and the accuracy of a sampling/mapping rule, given only the prior random field \( H_0 \). In the latter case, both \( \hat{M} \) and \( M \) are random maps.

The accuracy of a given estimate after \( r \) measurements can be expressed in terms of the joint probability distribution of the performance ratios that correspond to the posterior random field \( H_0(X) \). Analytical calculation of this distribution is hopeless, although some of its characteristics (e.g., the mean and variance of \( \beta_3, \mu_3 \), and \( \mu_3 \)) are easy to find. For other characteristics one must resort to conditional simulation (to numerical simulation from the posterior random field \( H_0(X) \)), for example, through the procedure of Deffner and Delfhomme (1975).

In evaluating a sampling/mapping rule, the number of sample points \( n \), their locations \( X_1, \ldots, X_n \), and the associated measurements \( Z_1, \ldots, Z_n \) are random variables that ultimately affect the distribution of the performance ratios. Also in this case numerical simulation is a most convenient tool: first one generates functions \( h(X) \) as independent realizations of the prior random field \( H_0(X) \); then one samples each realization and calculates \( \hat{D}^* \) from Eq. 7 and the performance ratios from Eqs. 2, 3, and 5. Finally, distribution characteristics of the performance ratios are estimated as sample statistics. By this method one can also estimate the distribution of the number of sample points \( r \), the distribution of the distance \( d \) traveled during exploration, and the dependence of the performance ratios on \( r \) and \( d \). Finally, one can study the effect of varying the control parameters \( \beta_m \) and \( \beta_M \) or the initial state of uncertainty (random field, \( H_0(X) \)) and compare the performance of different sampling and mapping rules.

The second approach, with both \( H_0 \) and \( \hat{M} \) random maps and only \( H_0(X) \) given, was used to evaluate the strategy of the last section and to compare it with a simpler rule of systematic sampling. In all cases, \( H_0(X) \) was assumed to be stationary and Gaussian in the one-dimensional interval \([0, L]\), with exponential correlation function, \( \rho(x) = \exp \left(-\frac{x}{\delta L}\right) \). Sampling was noiseless \((e = 0)\). The a priori distance function \( \beta_3(x) \) was taken to be constant in \([0, L]\) and these parameters combinations were considered (see Table 1).

For each combination, 50 pseudo-random samples were generated from \( H_0(X) \) using the procedure of Shinozuka and Jan (1972). After sampling and estimation using Eq. 7, the performance vector

\[
\mathbf{Y} = \begin{bmatrix} \gamma \\ \mu \end{bmatrix}
\]

in which

\[
\begin{align*}
\gamma &= \text{vector of } \gamma_i \\
\mu &= \text{vector of } \mu_i \\
\end{align*}
\]

(16)

was calculated for each sample function. With \( y_i \), the value of \( \mathbf{Y} \) for the \( i \)-th function, the mean value and the covariance matrix of \( \mathbf{Y} \) were estimated as

\[
\bar{\mathbf{Y}} = \frac{1}{50} \sum_{i=1}^{50} \mathbf{y}_i
\]

and

\[
\mathbf{S}_\mathbf{Y} = \frac{1}{49} \sum_{i=1}^{50} (\mathbf{y}_i - \bar{\mathbf{Y}})(\mathbf{y}_i - \bar{\mathbf{Y}})^T
\]

(17)

\begin{table}[h]
\centering
\caption{Parameter Combinations for One-Dimensional Contouring Exploration Problems}
\begin{tabular}{|c|c|c|c|}
\hline
Case & \( \beta_0 \) & \( \beta_m \) & \( \beta_M \) \\
\hline
1 & 1.5 & 2.0 & 0.01 \times 111 \\
2 & 1.5 & 2.5 & 0.01 \times 111 \\
3 & 2.0 & 2.5 & 0.01 \times 111 \\
\hline
\end{tabular}
\end{table}
respectively. The mean value vector $\overline{Y}$ includes the expected number of sample points, $\tilde{F}$. For purpose of comparison, all 50 realizations were also sampled systematically (at regular intervals) with a number of observation points equal to the integer closest to $\tilde{F}$. New values of the $y_i$ were then obtained using the map estimate that results from linear interpolation between the observed values of $h$. Finally, new performance characteristics were calculated from Eq. 17.

In all cases, the interval $[0, L]$ was discretized into 100 subintervals and the $i$th subinterval was considered to be entirely in $D^+$ ($D^-$) if its center point was itself in $D^+$ ($D^-$).

**NUMERICAL RESULTS**

Relationships between $\beta_M$ and the expected sample size $\tilde{F}$ are shown in Fig. 5 for the parameter combinations of Table 1. When $\beta_M = 0$, phase 2 is inactive and exploration terminates as phase 1 first terminates. In this case, the accuracy of search and the number of samples depend only on $\beta_m$, and more precisely on $\beta_m - \beta_0$. (This difference is 0.5 in cases 1 and 3 and 1 in case 2.) The increase of $\tilde{F}$ with $\beta_M$ is due to the refinement of the excursion boundaries under phase 2 and, to a smaller extent, to further search for "excursions under new activations of phase 1. Hence the increase of $\tilde{F}$ is larger if the number of identified "excursions is larger (if $\beta_0$ is smaller), as can be seen by comparing cases 1 ($\beta_0 = 1.5$) and 3 ($\beta_0 = 2.0$). The rapid increase of $\tilde{F}$ from $\beta_M = 0.1$ to $\beta_M = 0.3$ indicates that the gradient of $\beta$ at new estimated crossing points is typically between 0.1 and 0.3, irrespective of the contouring level. Although

Fig. 5. Mean sample size $\tilde{F}$ versus $\beta_M$ for the parameter combinations of Table 1.

Fig. 6. Expected performance ratios versus expected sample size for sequential sampling (solid lines) and systematic sampling (dashed lines). Case 1 in Table 1. Numbers on plots are values of $h$. Few additional samples are needed for $\beta_M > 0.3$, these few samples are found to be very effective in improving some of the performance statistics.

Figures 6 through 8 relate average performance ratios to the average sample size $\tilde{F}$ (to the actual sample size $r$ for the case of systematic sampling). They share some common features. One is that the performance statistics of the se-
Fig. 7. Expected performance ratios versus expected sample size for sequential sampling (solid lines) and systematic sampling (dashed lines). Case 2 in Table 1. Numbers on plots are values of $r$.

Sequential strategy are consistently superior to those of systematic sampling, especially when $F$ and $r$ are large. Another interesting feature is that, for given $\beta_0, \beta_m$, and $\beta_d$, the performance ratios for sequential sampling are statistically less variable than in the case of systematic sampling. Increasing $F$ by increasing $\beta_m$ shifts the emphasis of exploration from discovering new excursions to better locating excursions that are already known; hence, the ratios $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$, and $\gamma_6$ which are the most sensitive to accurate estimation of $\beta D^*$, are also those that benefit the most from increasing $F$.

With respect to case 1, case 2 emphasizes the discovery of new excursion sets. This fact causes an increase of $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$, and $\gamma_6$ and a decrease of $\gamma_1, \gamma_2, \gamma_3$, and $\gamma_4$ (more excursions are detected or identified, but fewer of them are isolated or matched; e.g., compare solid-line values in Figs. 6 and 7 for $F = 16$).

The coefficients of variation of $\gamma_i$ vary from about 0.3 for $i = 1, 2, 3$ to about 0.8 for $i = 4$. The coefficients of variation of $\mu_i$ are higher—typically 2 or 3 but sometimes as high as 7—because the expected values $\left[\mu_i\right]$ are close to zero. In general, correlation is positive among the $\gamma_i$, and the $\gamma_i$ is negative among the $\mu_j$ and positive among $\gamma_i$ or $\gamma_j$ and $\mu_j$. 
Prior uncertainty on $h(X)$ often includes a parametric component in the form of a stochastic regression trend. Added to this stochastic trend is a non-parametric random component, for example, in the form of the prior random process used in all previous calculations. Equations 10 through 13 can still be used to simultaneously update both components of uncertainty. The case when the parametric component reduces to a single random variable $h(X)$ was analyzed by numerical simulation and the performance statistics of sequential sampling were compared with those of systematic exploration. It is not surprising that the sequential strategy gave much better results. The reason is that this strategy uses the first few samples to refine the estimate of the spatial mean of $h(X)$; then, if the estimate is high, much additional effort is allocated to discovering likely excursions, whereas if the estimate is low, exploration is soon terminated. Indeed, this would be the strategy of any rational decision maker.

An example of exploration in the plane is shown in Fig. 9. The figure gives contour lines of a realization from a homogeneous Gaussian random field with zero mean, unit variance, and autocorrelation function, $R(e_1, e_2) = \text{corr}[h(X_i, X_j)] = \exp(-|e_1 + e_2|)/2$. The region of exploration $D$ is a square of side length $L$ and the crossing level of interest is $\beta_0 = 1.6$. The sequential strategy, with parameters $\beta_0 = 2.0$ and $\beta_1 = 1.0$, allocated 35 sample points and estimated correctly the darker excursion regions in Fig. 9. The realization happens to be essentially positive in the right half of $D$ and essentially negative in the left half. The procedure "learned" about this apparent trend and, after a few exploratory samples, concentrated the effort in the region with highest expected return. The many samples along the right boundary of $D$ are due to assuming $\beta_1$, stationary in the direction normal to the boundary, at all boundary points. For the numerical calculation, the exploration region was discretized using a square grid of $20 \times 20$ points.

CONCLUSIONS

Contouring exploration of uncertain functions is most effective if decisions are made sequentially. Rigorous optimization of the sampling strategy in a decision-theoretic sense does not appear to be feasible or to lead to practicable solutions; hence the need for simple approximate methods. One such method, based on second-moment analysis, is proposed herein and compared with systematic nonsequential sampling. The main objectives of contouring exploration, (i) discovering new excursions above the level of contouring, and (ii) locating the boundary of known excursion sets, are pursued by the sequential strategy in two alternating phases. Repeated application to functions simulated pseudo-randomly from stationary random processes on the line and on the plane has confirmed the superiority of the sequential strategy with regard to both objectives. Sequential methods are particularly useful when the function of interest has an uncertain trend component and in general when the random process that quantifies prior uncertainty on such a function is highly correlated. All applications in this study are to uncertain functions at space, but the same procedure applies to functions that also vary in time; in particular, to functions the uncertainty of which satisfies a Kalman filter equation in time. This may be the case, for example, in sampling for violation of pollution standards in water bodies or the atmosphere (Kitamis, Veneziano, and Queiroz, 1978).

ACKNOWLEDGMENT

Financial support was provided by the National Science Foundation, under Grant No. ENG-75-19400.
REFERENCES


Veneziano, D., 1975, Statistical estimation and data collection; a review of procedures for civil engineers, in Proceedings, Third international conference on the applications of statistics and probability in soil and structural engineering: Sydney, Australia.