

3665

Received w/Ltr Dated 2/1/90

CAMPBELL

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-306

TRIPLE

TITLE BOOTSTRAPPED MODELS FOR INTRINSIC RANDOM FUNCTIONS

AUTHOR(S) Katherine Campbell, S-1

PRESENTED TO: The MGUS-87 Conference.
Redwood City, CA, April 13-15, 1987.

By acceptance of this article the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution or to allow others to do so for U.S. Government purposes.

The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO 836 R4
ST NO 2629 5/81

9002070405 900201
PDR WASTE
WM-11 PDC

102.8

ENCLOSURE

Bootstrapped Models for Intrinsic Random Functions

Katherine Campbell

Statistics and Applied Mathematics Group

Los Alamos National Laboratory

Los Alamos, New Mexico 87544

ABSTRACT

The use of intrinsic random function stochastic models as a basis for estimation in geostatistical work requires the identification of the generalized covariance function of the underlying process, and the fact that this function has to be estimated from the data introduces an additional source of error into predictions based on the model. This paper develops the sample re-use procedure called the "bootstrap" in the context of intrinsic random functions to obtain realistic estimates of these errors. Simulation results support the conclusion that bootstrap distributions of functionals of the process, as well as of their "kriging variance", provide a reasonable picture of the variability introduced by imperfect estimation of the generalized covariance function.

KEY WORDS: regionalized variables, kriging, interpolation, sample re-use, nonparametric error estimates, confidence intervals.

INTRODUCTION

One of the most frequently cited problems with the geostatistical estimation technique known as "kriging" is the necessity of using an estimated model for the covariance function (semi-variogram, generalized covariance function). This problem is alluded to, for example, by Philip and Watson (1986), who cite in particular earlier work of Hardy (1977), and by Armstrong (1984a), among others. Brooker (1986) concluded that the kriging variance is "robust to most errors likely to be made in semivariogram model selection" for the class of models included in his parametric study and a particular regular block-sample geometry, with the exception that it could be quite sensitive to the incorrect choice of the nugget value. Diamond and Armstrong (1984) defined a metric on the space of variogram functions and provided estimates of the stability of the kriged estimates and the kriging variance relative to changes in the variogram function as measured by this metric. As their metric is based on the ratio between two variogram functions, less (absolute) variability is allowed within a δ -neighborhood of a given variogram where that variogram is small, and almost always it is smallest for separations near zero; this confirms the sensitivity of kriging results to the choice of nugget observed by Brooker.

Motivated by this earlier work, this paper explores a technique for quantifying the error resulting from imperfect model estimation when the only data at hand are the sample observations. With the increasing availability of high-speed computers, a number of computation-intensive, non-parametric methods for addressing this problem have been developed in the statistical literature in recent years. These sample-reuse methods--the "jackknife", cross-validation and the "bootstrap"--are reviewed and compared by Efron

(1982). Applications in geostatistics include Chung (1984), Dubrule (1983), Campbell (1986) and Solow (1985). This paper expands on the preliminary work of Solow (1985) which applied the bootstrap, introduced by Efron (1979), to estimates of functionals of a stationary random process.

INTRINSIC RANDOM FUNCTIONS AND KRIGING

This brief review will serve to introduce the notation used in this paper. For greater detail the reader is referred to the seminal paper of Matheron (1973) and the more practical exposition by Delfiner (1979).

The data, $\underline{z} = \underline{z}(\underline{x}) = (z(x_1), \dots, z(x_N))$, are samples at known points x_i of one realization $z(x)$ of a stochastic process $Z(x)$. Probabilistic assumptions about $Z(x)$ might include joint normality of the random variables $Z(x)$ indexed by a spatial variable x (in d -dimensional Euclidean space) or some form of generalized stationarity. In this paper it is assumed that Z is an intrinsic random function of order k (a k -IRF), which means, first, that (at least locally) its drift $E(Z(x))$ is polynomial in x of order k , and secondly, that if α is a measure with compact support in d -space with the property that

$$\int p(x) d\alpha(x) = 0$$

for every polynomial $p(x)$ of order less than or equal to k , then αZ , defined by

$$(\alpha Z)(s) = \int Z(t+s) d\alpha(t) \quad (1)$$

is a stationary random process (indexed by s in d -space). Below only measures a with finite support ("k-increments") are considered, and it will be convenient to replace the integral notation of Eq. (1) with inner product notation:

$$\int Z(t) da(t) \rightarrow \langle \underline{a}, \underline{Z} \rangle \stackrel{\text{def}}{=} \sum_{i=1}^T a_i Z(t_i), \quad (2)$$

where $\{t_1, \dots, t_T\}$ contains the support of a . Second moments of aZ are assumed to exist, but no explicit normality assumptions are made. If a and β are any two k-increments, and $\{t_1, \dots, t_T\}$ contains the union of their supports, then the covariance of the random variables $\langle \underline{a}, \underline{Z} \rangle$ and $\langle \underline{\beta}, \underline{Z} \rangle$ is given in terms of the generalized covariance function $K(h)$ of the IRF Z (defined for $h \geq 0$) by

$$\sum_{i,j=1}^T a_i \beta_j K(|t_i - t_j|), \quad (3)$$

where $|\cdot|$ denotes the usual Euclidean norm in d -space.

The kriging problem is to estimate some functional $f(Z)$ conditional on the observations \underline{z} . Typical functionals are simply the identity evaluated at a point x_0 (interpolation) or integrals of Z over a volume centered at x_0 (block estimation). Estimators of $E \{ f(Z) \mid \underline{z}(x) \}$ are restricted to be linear in the data, i.e., to be of the form

$$\hat{f} = \sum_{i=1}^N \lambda_i z(x_i) = \langle \underline{\lambda}, \underline{z} \rangle. \quad (4)$$

In addition, the kriging algorithm provides an estimate of the "kriging variance"

$$\sigma^2 = E \{ (\hat{f} - f(Z))^2 \mid \underline{x} \}. \quad (5)$$

If the generalized covariance function K were known, this problem would have a best (in the sense of minimizing (5)) linear unbiased solution \hat{f}_T , and the usual kriging algorithm would provide $\hat{\lambda}_T$ such that $\hat{f}_T = \langle \hat{\lambda}_T, \underline{z} \rangle$, together with the associated $\hat{\sigma}_T^2$. In practice, of course, K is unknown and must be estimated from the data, and the usual procedure is to use this estimate \hat{K} in place of K in the kriging equations to obtain a suboptimal solution \hat{f}_K (with weights $\hat{\lambda}_K$) and $\hat{\sigma}_K^2$.

Unfortunately, the estimation of the generalized covariance function K is highly problematic. There have been numerous efforts in recent years to improve on the method of Delfiner (1979) which was critically reviewed by Starks and Fang (1982). These include efforts at making the procedure more robust against non-normality of the data (see for example the discussion by Armstrong, 1984b) as well as other types of estimation procedures (e.g., Kitanidis, 1985). Nevertheless, all methods simulated by this author are plagued by problems of bias and large variability, even within an unrealistic simulation framework where the form of the true function is known and correctly modeled. Under these circumstances, additional effort to quantify the error introduced by imperfect modeling seems worthwhile.

A REVIEW OF THE BOOTSTRAP

A readable introduction to the bootstrap is provided by Efron and Tibshirani (1986). The literature on the subject has grown rapidly since its

introduction by Efron (1979), and widespread application has hardly awaited the establishment of more formal underpinnings such as provided by Bickel and Freedman (1981).

The basic idea of the bootstrap is very simple. Let $\hat{\theta}(y)$ be an estimate of a parameter (or parameters) $\mu(F)$. (Here I employ the notation of Efron and Tibshirani, 1986.) This means that there is a well defined algorithm for computing $\hat{\theta}$, given observations y generated by a stochastic mechanism F characterized by unknown parameter(s) μ . The available observations y are used to generate an empirical estimate \hat{F} of F (which can be, for example, the usual non-parametric empirical distribution function, a smoothed version of this, or some parametric form). Simulation from \hat{F} is used to evaluate the properties of the estimator $\hat{\theta}$. Specifically, random samples are drawn from \hat{F} (each is called a bootstrap sample, denoted by y^*) and the algorithm is applied to compute $\hat{\theta}(y^*)$. (Fifty to two hundred bootstrap samples are sometimes recommended, but more are needed for computing bootstrapped confidence intervals.) The sample statistics of the $\hat{\theta}(y^*)$ are then used to estimate the corresponding properties of the original estimator $\hat{\theta}(y)$. For example, the bias of $\hat{\theta}(y)$ may be estimated by

$$E_{\hat{F}} \hat{\theta}(y^*) - \mu(\hat{F}), \quad (6)$$

where $E_{\hat{F}} \hat{\theta}(y^*)$ is approximated by the mean of the bootstrapped values. The standard error is approximated by the square root of the sample variance of the bootstrapped values.

For the case of independently, identically distributed (i.i.d.) observations y from an unknown distribution F the above program is easily carried out. For observations generated by a more complicated model, such as a

regression model, modifications are needed to transform y to a nominally i.i.d. set of random variables so that an empirical \hat{F} may be computed. Similarly, when the observations are generated by an intrinsic random function (including the possibility of non-constant drift) a method for reducing the correlated observations $z(x)$ to an approximately i.i.d. sample must be found. This is of course exactly the same problem faced in the original work of Matheron (1973), and the reason for introducing the notion of k -increments. Details for the present problem will be given in the following section.

What parameter(s) should be bootstrapped? As Diamond and Armstrong (1984) observe, it is not the estimate of the generalized covariance function $K(h)$ which is of interest so much as the effect of misspecification of this function on prediction algorithms such as the kriging algorithm. Recall that the kriged weights $\hat{\lambda}_K$ and the kriging variance $\hat{\sigma}_K^2$ depend on the data z only through \hat{K} and the observation points x (as suggested by the notation of Eq. (5)), whereas \hat{f}_K , computed by inserting $\hat{\lambda}_K$ in Eq. (4) in place of λ , of course depends explicitly on the observed values. Therefore, once the first problem of obtaining an empirical distribution \hat{F} from which we can simulate is solved, \hat{F} will be used to generate bootstrap samples $z^*(x)$ (at the original observation points x) and the prescribed estimation procedure will be used to obtain $\hat{K}(z^*)$. Then the usual kriging algorithm provides λ^* and $\hat{\sigma}^{*2}$ based on $\hat{K}(z^*)$. However, the original $z(x)$ will be used, together with λ^* , to compute \hat{f}^* via Eq. (4).

In part the bootstrap procedure will be assessing a particular algorithm for obtaining \hat{K} . Clearly a requirement for such an algorithm in the present context is that it be completely automatic; methods which rely on graphical interaction with the user are impractical. In this paper Kitanidis' (1985) "minimum-variance" estimator is used, constrained by quadratic programming to

provide positive-definite estimates. (Thus constrained the estimator is no longer unbiased, of course; in fact simulation suggests considerable bias.) This estimator requires a preliminary estimate of K , from which the algorithm can be applied iteratively to move towards the solution \hat{K} , which is a fixed point. Such a preliminary estimate could be provided by Delfiner's method (1979), and simulation experience suggests that convergence is fairly rapid; five iterations generally estimate the fixed point to three or four significant figures. In order to minimize computation in the nested simulation framework described below, however, the algorithm is started with a constant initial estimate, followed by a single step towards the fixed point. The bootstrap technique described here could similarly be used with any automatic algorithm.

BOOTSTRAPPING A k -IRF

Solow (1985) treated the simplest case of an intrinsic random function, namely a stationary process, for which the condition including Eq. (1) holds for all measures α with compact support; this is sometimes called an intrinsic random function of order -1 . The mean is assumed to be zero. (In practice, the sample mean is subtracted from all observations.) The positive definite covariance matrix C of the observations $\underline{Z} = (Z(x_1), \dots, Z(x_N))$ can be decomposed into the product of a matrix Φ and its transpose, $C = \Phi\Phi^T$ (Solow uses a Cholesky decomposition where Φ is lower triangular, but this is not essential), and the N components of the vector $\Phi^{-1}\underline{Z}(x)$ are uncorrelated with mean zero and common variance one. (In general, of course, this does not imply that they are independently identically distributed, without additional assumptions of normality.) To bootstrap estimates of parameters associated with the process, the unknown covariance matrix C is replaced by an estimate \hat{C}

$= \hat{\Phi}\hat{\Phi}^T$, and $\underline{u}(\underline{x}) = \hat{\Phi}^{-1}\underline{z}(\underline{x})$ is used to generate an empirical distribution function \hat{F} . Bootstrapped samples $\underline{z}^*(\underline{x})$ are then obtained from random samples $\underline{u}^*(\underline{x})$ of \hat{F} by the transformation $\underline{z}^*(\underline{x}) = \hat{\Phi}\underline{u}^*(\underline{x})$ and used to generate bootstrapped covariance functions \hat{C}^* and estimates of related functionals of the process.

The situation becomes slightly more complicated for an intrinsic random function of order greater than or equal to zero, because now there is, in addition to correlation among the observations, the possibility of unknown polynomial drift, so the sample covariance matrix of the observations \underline{z} can not be estimated directly. The solution, of course, is to work with k -increments \underline{a} supported on the observation points \underline{x} . If there are N observation points, then there are $M = N - C(k+d, k)$ linearly independent k -increments supported on \underline{x} , where $C(m, m')$ denotes the binomial coefficient, the number of ways of selecting subsets of size m' from a set of size m . Let $\underline{a}_1, \dots, \underline{a}_M$ be M linearly independent k -increments. Let A denote the $M \times N$ matrix whose $(i, j)^{\text{th}}$ element is the j^{th} element (coefficient) of \underline{a}_i . Let \hat{K} be the estimate of the generalized covariance function based on the original data $\underline{z}(\underline{x})$, and estimate the $M \times M$ covariance matrix with elements $C_{ij} = \text{Cov} \{ \langle \underline{a}_i, Z \rangle, \langle \underline{a}_j, Z \rangle \}$ by using \hat{K} in Eq. (3). Now, following Solow, factor $\hat{C} = \hat{\Phi}\hat{\Phi}^T$ and compute $\underline{u} = \hat{\Phi}^{-1}(A\underline{z})$ to obtain M approximately independent and identically distributed random variables \underline{u} . These are used to generate an empirical distribution function \hat{F} .

To complete the bootstrap algorithm, draw a random sample \underline{u}^* of size M from \hat{F} and compute $(A\underline{z})^* = \hat{\Phi}\underline{u}^*$ and finally $\underline{z}^* = A^{-}(A\underline{z})^*$, where A^{-} is any generalized inverse of A . Then, as outlined previously, compute $\hat{K}(\underline{z}^*)$, apply the kriging equations to get bootstrapped kriging weights $\underline{\lambda}^*$, kriging variance

σ^{*2} , and the bootstrapped estimate $f^* = \langle \underline{\lambda}^*, \underline{z} \rangle$. Of interest will be the sample distributions of f^* and σ^{*2} .

SOME SIMULATION RESULTS

The remarks in this section are based on a simulation study consisting of one hundred simulations of a first-order IRF with generalized covariance function $K(h) = -10|h|$. Thirty observations were generated for each simulation, at the points x_1, \dots, x_{30} shown in Figure 1 with solid symbols. The form of the Kitanidis algorithm previously described was used to estimate the three parameters of the general model

$$K(h) = c_0 \delta(h) - b_1|h| + b_3|h|^3, \quad (7)$$

(where δ is the delta function), starting the iteration with $c_0 = 0$, $b_1 = 1$ and $b_3 = 0$. (As Kitanidis observes in his 1985 paper, what counts in the selection of the initial estimate is the ratio of the coefficients, so in fact the correct ratio is offered as a starting point. This does not keep the algorithm from converging rapidly to its biased fixed point, however.) For each simulation, the covariance matrix of twenty-seven k -increments $A_{\underline{z}}$, estimated using \hat{K} , was factored and twenty-seven approximately independent and identically distributed numbers \underline{u} were computed as described above. The usual empirical distribution function, with mass $1/27$ at each of these points, was used for \hat{F} , and one hundred bootstrapped samples \underline{z}^* were generated from \hat{F} for each simulation. The same Kitanidis algorithm applied to each \underline{z}^* yielded a bootstrapped estimate \hat{K}^* .

The functionals of \hat{K} that were considered were simply $f(Z) = Z(x_p)$ at the five points $P = A, B, C, D$ and E shown in Figure 1 with open symbols.

Point A is very close to one of the data points, B and C are surrounded by observations, D is only partly surrounded, and E is well outside the observed area. For each point the following quantities were computed:

$\hat{f}_T, \hat{\sigma}_T^2$, using the true generalized covariance function $K(h) = -10h$;

$\hat{f}_K, \hat{\sigma}_K^2$, using the estimated generalized covariance function \hat{K} ;

and one hundred values of

$\hat{f}^*, \hat{\sigma}^{2*}$, using the bootstrapped generalized covariance functions \hat{K}^* .

(All thirty observations were used in the kriging equations.)

As observed previously, the constrained Kitanidis algorithm (all of the coefficients in Eq. (7) must be non-negative) is not unbiased. Figure 2 is a histogram of the one hundred simulated values of the nugget effect C_0 ; they range up to 13.3 and only about half are at or very close to zero. The linear coefficients b_1 are shown in Figure 3; they have an average value of about 8. and include some zeros. This bias actually has very little effect on \hat{f}_K , which in most cases is very close to \hat{f}_T , as shown in Figure 4 for point C. The effect on $\hat{\sigma}_K^2$ is more substantial (Figure 5); the kriged values deviate from the optimal values by up to a factor of five or six, and the algorithm has a downward bias.

As an estimate of bias, however, the natural bootstrapped estimate

$$\hat{B}_f = \overline{\hat{f}^*} - \hat{f}_K \quad \text{for} \quad B_f = \hat{f}_K - \hat{f}_T, \quad (8)$$

where the bar denotes the average over all bootstrap samples, is not very satisfactory. When B_f is practically zero, \hat{B}_f is often relatively large; when B_f is significantly different from zero, \hat{B}_f is as likely as not to have the wrong sign. Typical is Figure 6, a plot of \hat{B}_f vs. B_f at point C. It might be thought that this apparent arbitrariness of sign is due to the influence of a few extreme values in the distribution of bootstrapped values; some bootstrap distribution, such as the one in Figure 7, have long tails or other non-normal behavior. However, if the median of the distribution is used in place of the mean in Eq. (8), the result is generally an estimate of bias that is of the wrong sign (Figure 8), so a more robust estimate of the center of the bootstrap distribution is of little help. In fact, the center of the bootstrap distribution is usually quite close to the optimal value, regardless of the bias of the kriged value.

On the other hand, \hat{B}_σ , defined by

$$\hat{B}_\sigma = \log \overline{\hat{\sigma}^{2*}} - \log \hat{\sigma}_K^2, \text{ estimating } B_\sigma = \log \hat{\sigma}_K^2 - \log \hat{\sigma}_T^2,$$

is negative almost every instance, regardless of the sign of B_σ ; see Figure 9. This is a reflection of the bias of the algorithm remarked in Figure 5; just as the estimated kriging variance underestimates the true value on the average across simulations, so the bootstrap simulation underestimates the kriged value.

Thus it appears that at least for this problem bias estimation is poor, and bias "correction" (i.e., subtracting the bias estimate from the kriged estimate) decidedly inadvisable. This agrees with a remark by Efron and Tibshirani (1986; p. 61).

More interesting is the question of the increase in variability in the kriged estimate as a result of using an estimated generalized covariance function. Confidence intervals for \hat{f}_T and $\hat{\sigma}_T^2$ can be constructed for each simulation, and their predicted coverage compared with the observed coverage in the one hundred trials. One hundred bootstrap samples are considered insufficient to construct good confidence limits, according to Efron and Tibshirani. They recommend a minimum of 250 samples for the percentile method (described below), and 1000 samples for the more complicated bias-corrected percentile methods. These sample sizes are impractical in a simulation of this size. Nevertheless, some interesting observations emerge even with the smaller bootstrap samples used here.

The first method for constructing confidence intervals is called the percentile method (Efron and Tibshirani, 1986; Efron, 1982). Here a symmetric $100(1-2\gamma)\%$ confidence interval is constructed simply by taking as endpoints the γ^{th} and $(1-\gamma)^{\text{th}}$ quantiles of the empirical distribution of the bootstrapped values of the statistic in question. Table I shows the actual coverage of the one hundred 90% confidence intervals constructed using the percentile method for each of the estimated quantities, together with the number of simulations for which the optimal value (i.e., \hat{f}_T or $\hat{\sigma}_T^2$) fell below or above the computed confidence interval. The confidence intervals for \hat{f}_T are extremely conservative, attaining at least 97% coverage at all five points. Confidence intervals for $\hat{\sigma}_T^2$ are somewhat liberal and markedly nonsymmetric.

One problem may be that the quantity bootstrapped was not approximately pivotal (see Hinkley, 1986), although the use of the logarithm of σ^2 was an attempt to move closer to a pivotal quantity for this scale parameter. The second method for computing confidence intervals is a fairly automatic technique for correcting this problem (Efron, 1982, p. 83). This "bias-corrected"

percentile method involves making an adjustment for the fact that \hat{f}_K or $\hat{\sigma}_K^2$ is not the median of its bootstrapped distribution. Table 2 shows the coverage of bias-corrected 90%-confidence intervals. The bias-corrected confidence intervals are now liberal for the point estimates \hat{f}_T , but fairly symmetric with respect to failures on the low and high sides. Failures, unsurprisingly, are associated with poor bias estimates. Bias-corrected confidence intervals for the kriging variance $\hat{\sigma}_T^2$ have about the same coverage as percentile-method intervals and failures are slightly more symmetric, although the computed intervals still tend to be low. The bias-corrected intervals are also five to ten percent narrower than intervals computed by the percentile method.

A third method for constructing confidence intervals mentioned by Efron and Tibshirani is a modification of the bias-corrected percentile method which uses an estimate of skewness based on the bootstrap distribution. As bootstrap distributions are frequently highly skewed, this further adjustment might correct the observed asymmetry of failures for the intervals for $\hat{\sigma}_T^2$.

More bootstrap samples smooth out the bootstrap distributions considerably and improve the estimation of tail percentiles, as well as the estimation of the percentile of the kriged value, which plays a central role in the bias-correction method. Rarely, a significant shift in the computed confidence interval can result, but more typically there is little change, as seen in a comparison of Figures 10 and 11.

RANDOMIZATION

There are a number of points in the procedure described above where arbitrary choices are made: in the selection of A (i.e., of M linearly independent k -increments), in the choice of a generalized inverse A^- , and finally in the factorization of \hat{C} : if U is any unitary matrix, $UU^T = I$, then

$$\hat{C} = \hat{\Phi} U U^T \hat{\Phi}^T = \hat{\Phi}' \hat{\Phi}'^T. \quad (9)$$

The choice of a generalized inverse is immaterial because if A^\dagger is another generalized inverse of A , then $A^\dagger(Az)^*$ differs from $A^-(Az)^*$ only by a polynomial of order k , and the algorithm for estimating $K(h)$ should produce the same result using either z^* . If B is another $M \times N$ matrix of linearly independent k -increments, then $B = QA$ for some non-singular $M \times M$ matrix Q , and if $\hat{C}_A = \hat{\Phi} \hat{\Phi}^T$, where \hat{C}_A is the estimated covariance matrix of Az , then $\hat{C}_B = Q \hat{\Phi} \hat{\Phi}^T Q^T$. If this factorization of \hat{C}_B is used, then $\underline{u} = \hat{\Phi}^{-1} Q^{-1} (Bz) = \hat{\Phi}^{-1} (Az)$ as before. However, an alternative choice of the factorization will result in a different vector \underline{u} and a different \hat{F} .

The magnitude of the effect of the choice of factorization was investigated using a small simulation study and found to be comparable to the Monte Carlo sampling effect (i.e., the use of only a finite number of bootstrap samples from \hat{F}). Therefore it might be neglected. However, it might also prove useful. At the expense of a slight increase in computation, randomization of factorization can be achieved by generating random unitary matrices U for Eq. (9) and redefining $\underline{u} = (\hat{\Phi} U)^{-1} (Az)$ for each bootstrap sample, or, more reasonably, for each set of perhaps ten samples. The principal effect of this modification appears to be a widening of the bootstrap distribution and thus of confidence intervals; compare Figures 11 and 12. In this instance the upper end of the 90% confidence interval was shifted just enough to cover the "true" value $\hat{\sigma}_T^2$.

SUMMARY

The Matheronian theory of k -increments as a tool permitting the estimation of the generalized covariance function of an intrinsic random function

from a single realization has been applied here to reduce the observations from a single realization to an approximately i.i.d. collection of linear combinations of the observations. These can be resampled to generate new sets of "bootstrapped" observations, with which the estimation procedure can be repeated. According to bootstrap theory, the distribution of the resulting bootstrapped estimates (be they the parameters of the estimated generalized covariance function itself or functionals of the estimated process such as kriged interpolations, block estimates or kriging variances) should approximate the corresponding true distribution, and thus provide a non-parametric way to estimate such measures of error of the original estimator as bias or standard error, as well as generating confidence intervals for them.

Qualitatively, the simulation results reported here support this theory, at least insofar as suggesting that the spread of the bootstrap distributions does indeed reflect the variability introduced by imperfect estimation of the generalized covariance function required for kriging. Bias estimation is unreliable, but bias-corrected confidence intervals are only slightly too narrow (considering that the simulations used only one hundred bootstrap samples per simulation, compared to the minimum of one thousand recommended for this particular type of computation). The arbitrariness of one step of the procedure, namely the choice of factorization for the covariance matrix of the k -increments used, might profitably be exploited to widen the computed confidence intervals slightly.

REFERENCES

- Armstrong, M., 1984a, Problems with Universal Kriging: *Math. Geol.*, v. 16, p. 101-108.
- Armstrong, M., 1984b, Improving the Estimation and Modelling of the Variogram: in *Geostatistics for Natural Resources Characterization* (G. Verly et al., eds.), D. Reidel Publishing Company, Dordrecht, p. 1-19.
- Bickel, P. J. and Freedman, D. A., 1981, Some Asymptotic Theory for the Bootstrap: *Ann. Stat.*, v. 9, p. 1196-1217.
- Brooker, P. I., 1986, A Parametric Study of Robustness of Kriging Variance as a Function of Range and Relative Nugget Effect for a Spherical Semivariogram: *Math. Geol.*, v. 18, p. 477-488.
- Campbell, K., 1986, Kriging for Interpolation of Sparse and Irregularly Distributed Geological Data: Los Alamos National Laboratory Technical Report LA-UR 86-1894.
- Chung, C. F., 1984, Use of the Jackknife Method to Estimate Autocorrelation Functions (or Variograms): in *Geostatistics for Natural Resources Characterization* (G. Verly et al., eds.), D. Reidel Publishing Company, Dordrecht, p. 55-69.
- Delfiner, P., 1979, The Intrinsic Model of Order k : Notes for a short course at Battelle Seattle Research Center, Seattle, October, 1979.
- Diamond, P. and Armstrong, M., 1984, Robustness of Variograms and Conditioning of Kriging Matrices: *Math. Geol.*, v. 16, p. 809-822.
- Dubrulle, O., 1983, Cross Validation of Kriging in a Unique Neighborhood: *Math. Geol.*, v. 15, p. 687-699.
- Efron, B., 1979, Bootstrap Methods: Another Look at the Jackknife: *Ann. Stat.*, v. 7, p. 1-26.

- Efron, B., 1982, The Jackknife, the Bootstrap and Other Resampling Plans: Society for Industrial and Applied Mathematics, Philadelphia, 92 p.
- Efron, B. and Tibshirani, R., 1986, Bootstrap Methods for Standard Errors, Confidence Intervals and Other Measures of Statistical Accuracy: Statistical Science, v. 1, p. 54-77.
- Hardy, R. L., 1977, Least Squares Prediction: Photogram. Eng. Remote Sensing, v. 18, p. 445-448.
- Hinkley, D., 1986, Comment on Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis by C. F. J. Wu: Ann. Stat., v. 14, p. 1312-1316.
- Kitanidis, P. K., 1985, Minimum-Variance Unbiased Quadratic Estimation of Covariances of Regionalized Variables: Math. Geol., v. 17, p. 195-208.
- Matheron, G., 1973, The Intrinsic Random Functions and Their Applications: Adv. Appl. Prob., v. 5, p. 439-468.
- Philip, G. M. and Watson, D. F., 1986, Matheronian Geostatistics--Quo Vadis?: Math. Geol., v. 18, p. 93-117.
- Solow, A. R., 1985, Bootstrapping Correlated Data: Math. Geol., v. 17, p. 769-775.
- Starks, T. H. and Fang, J. H., 1982, On the Estimation of the Generalized Covariance Function: Math. Geol., v. 14, p. 57-64.

TABLE 1. COVERAGE OF NOMINAL 90% CONFIDENCE INTERVALS FOR \hat{f}_T AND $\hat{\sigma}_T^2$ CONSTRUCTED BY THE PERCENTILE METHOD (100 SIMULATIONS)

FUNCTIONAL	OPTIMAL VALUE			MEDIAN WIDTH OF INTERVAL
	BELOW INTERVAL	WITHIN INTERVAL	ABOVE INTERVAL	
\hat{f}_T at:				
A	0	99	1	.765
B	1	99	0	.431
C	0	100	0	.417
D	1	99	0	.407
E	0	97	3	1.455
$\hat{\sigma}_T^2$ at:				
A	5	95	0	1.396
B	1	86	13	.414
C	1	84	15	.413
D	0	84	16	.416
E	0	92	8	.489

TABLE 2. COVERAGE OF NOMINAL 90% CONFIDENCE INTERVALS FOR \hat{f}_T AND $\hat{\sigma}_T^2$ CONSTRUCTED BY THE BIAS CORRECTED PERCENTILE METHOD

FUNCTIONAL	OPTIMAL VALUE			MEDIAN WIDTH OF INTERVAL
	BELOW INTERVAL	WITHIN INTERVAL	ABOVE INTERVAL	
\hat{f}_T at:				
A	9	79	12	.440
B	8	86	6	.327
C	7	91	2	.335
D	7	86	7	.410
E	7	82	11	1.380
$\hat{\sigma}_T^2$ at:				
A	17	81	2	1.276
B	2	88	10	.389
C	3	86	11	.378
D	3	84	13	.378
E	2	90	8	.455

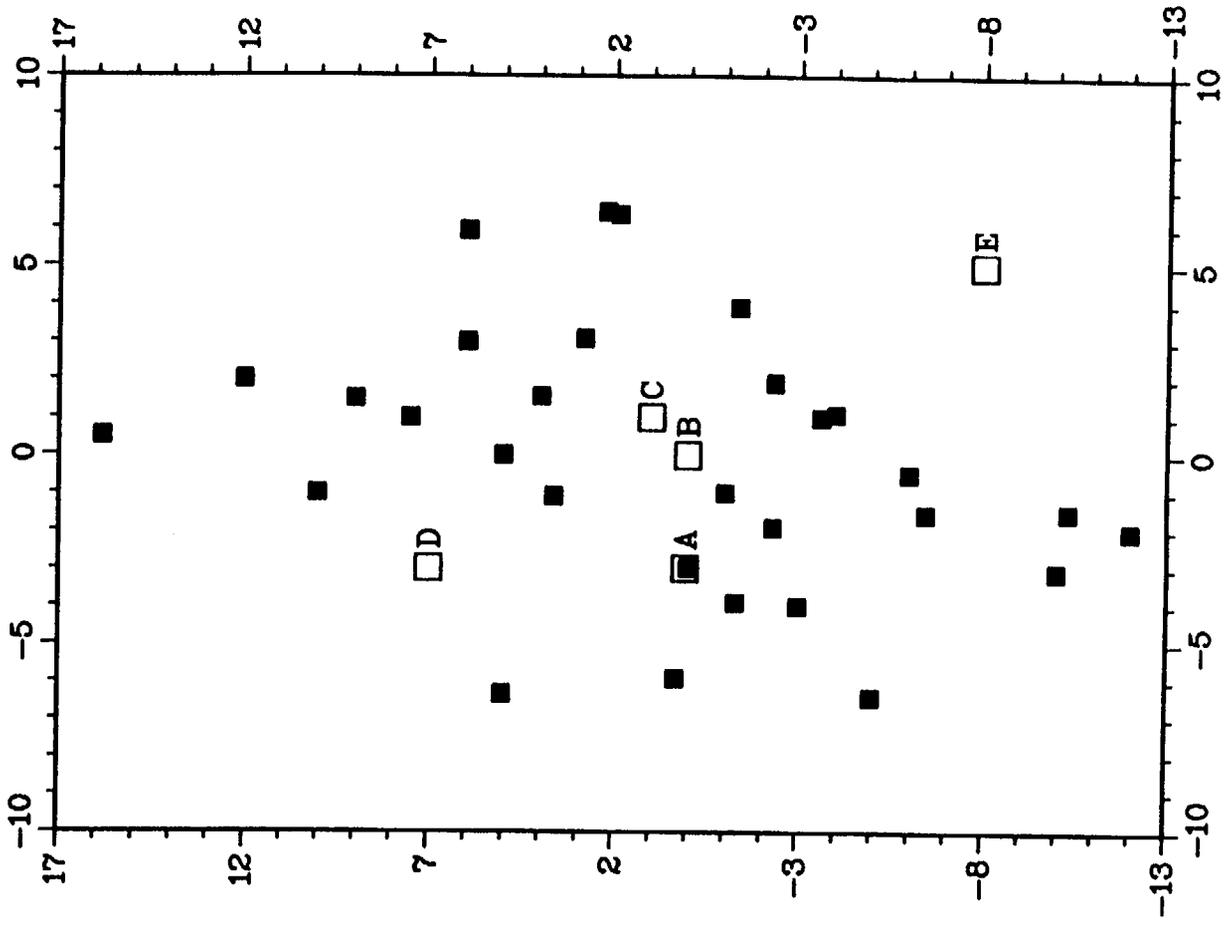
FIGURE CAPTIONS

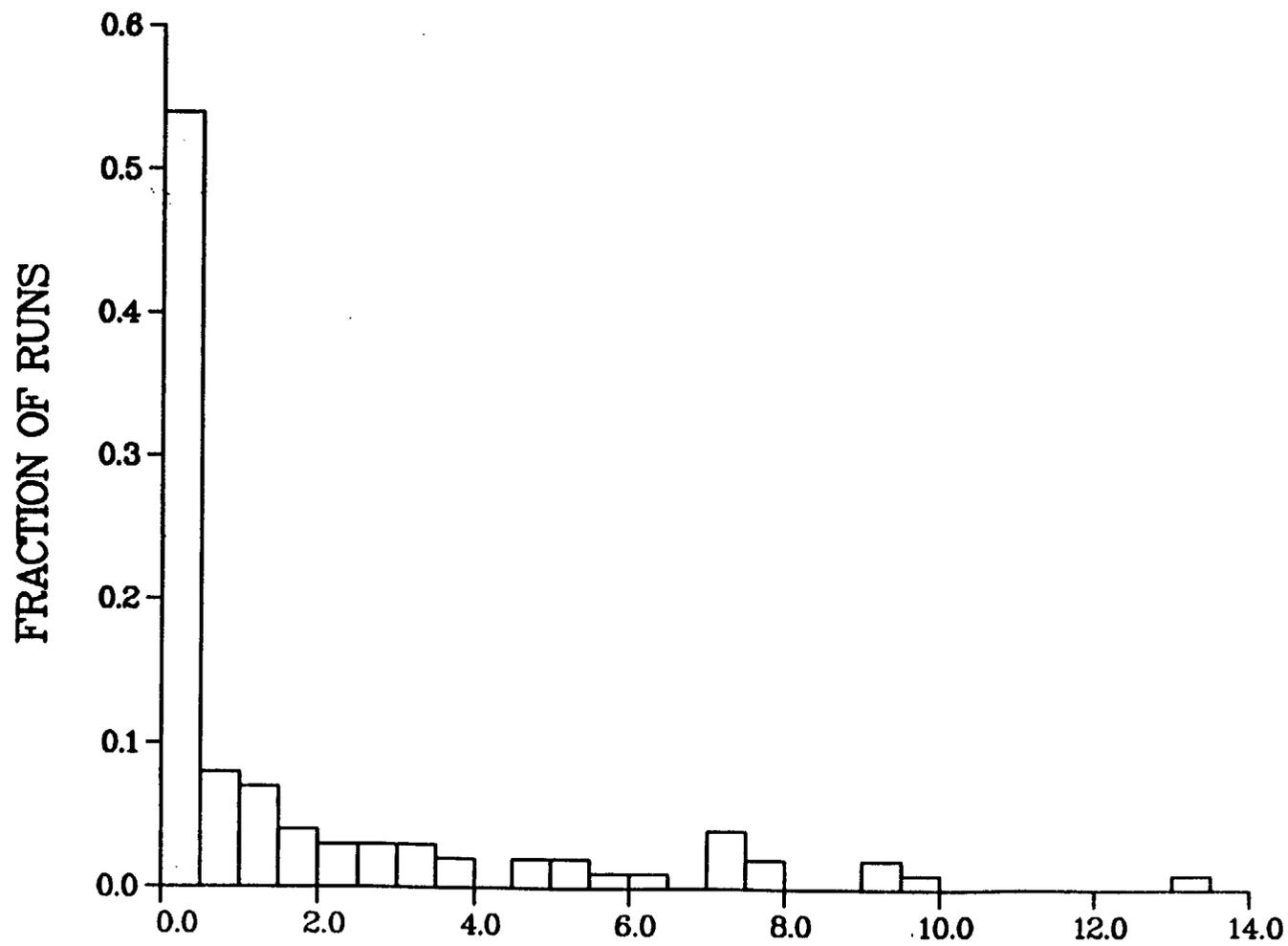
- Fig. 1. Thirty "observation" points used in the simulations are shown using solid symbols. Kriged estimates are computed at the five labeled points shown by open symbols.
- Fig. 2. Histogram of one hundred estimates based on simulated data of the nugget effect, C_0 . The true value, used to generate the simulations, was zero.
- Fig. 3. Histogram of one hundred estimates of the coefficient of the linear term, b_1 . The true value was 10.
- Fig. 4. Histogram of one hundred values of the bias $\hat{f}_K - \hat{f}_T$ of the estimate at point C.
- Fig. 5. Histogram of one hundred values of the bias $\log \hat{\sigma}_K^2 - \log \hat{\sigma}_T^2$ of the kriging variance at point C.
- Fig. 6. The bootstrapped bias of the point estimate, computed as the average of $\hat{f}^* - \hat{f}_K$ over one hundred bootstrap samples, vs. the actual bias $\hat{f}_K - \hat{f}_T$ at point C.
- Fig. 7. Histogram of the bootstrap values \hat{f}^* at point B for one simulation. The optimal value \hat{f}_T and the original kriged value \hat{f}_K are shown as heavy lines.
- Fig. 8. The bootstrapped estimate of the bias of the point estimate, computed as the median of $\hat{f}^* - \hat{f}_K$ over one hundred bootstrap samples, vs. the actual bias $\hat{f}_K - \hat{f}_T$ at point C.
- Fig. 9. The bootstrapped estimate of the bias of the kriging variance, computed as the average of $\log \hat{\sigma}^{2*} - \log \hat{\sigma}_K^2$ over one hundred bootstrap samples, vs. the actual bias $\log \hat{\sigma}_K^2 - \log \hat{\sigma}_T^2$ at point C.
- Fig. 10. Histogram of one hundred bootstrapped values of the kriging variance at point C for one simulation. $\hat{\sigma}_T^2$ and $\hat{\sigma}_K^2$ are shown as solid lines, and

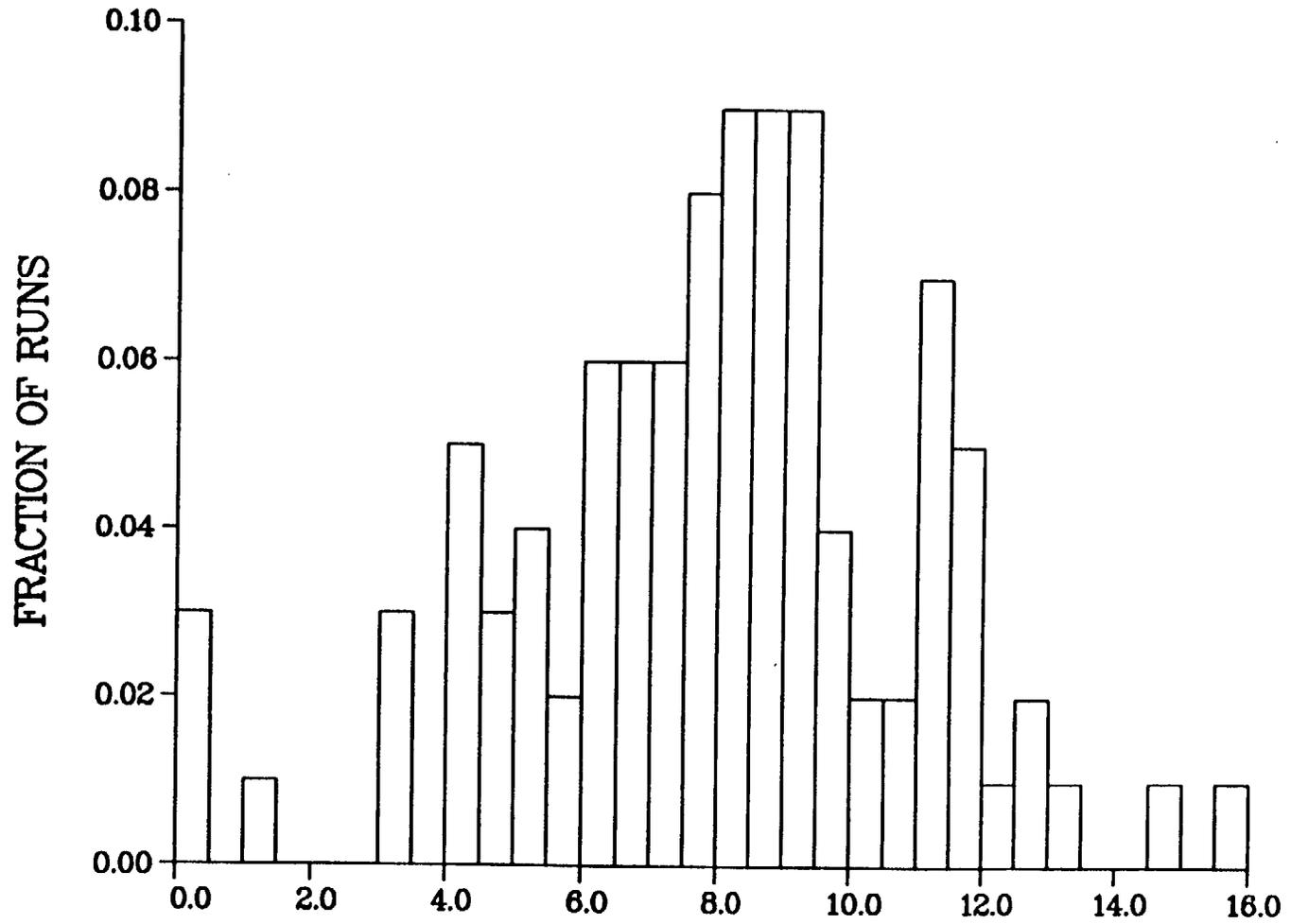
the boundaries of the bias-corrected 90% confidence interval are shown as dashed lines. The same factorization of the covariance matrix was used throughout.

Fig. 11. Histogram of one thousand bootstrapped values of the kriging variance at point C for the simulation illustrated in Figure 10. The same factorization of the covariance matrix was used throughout.

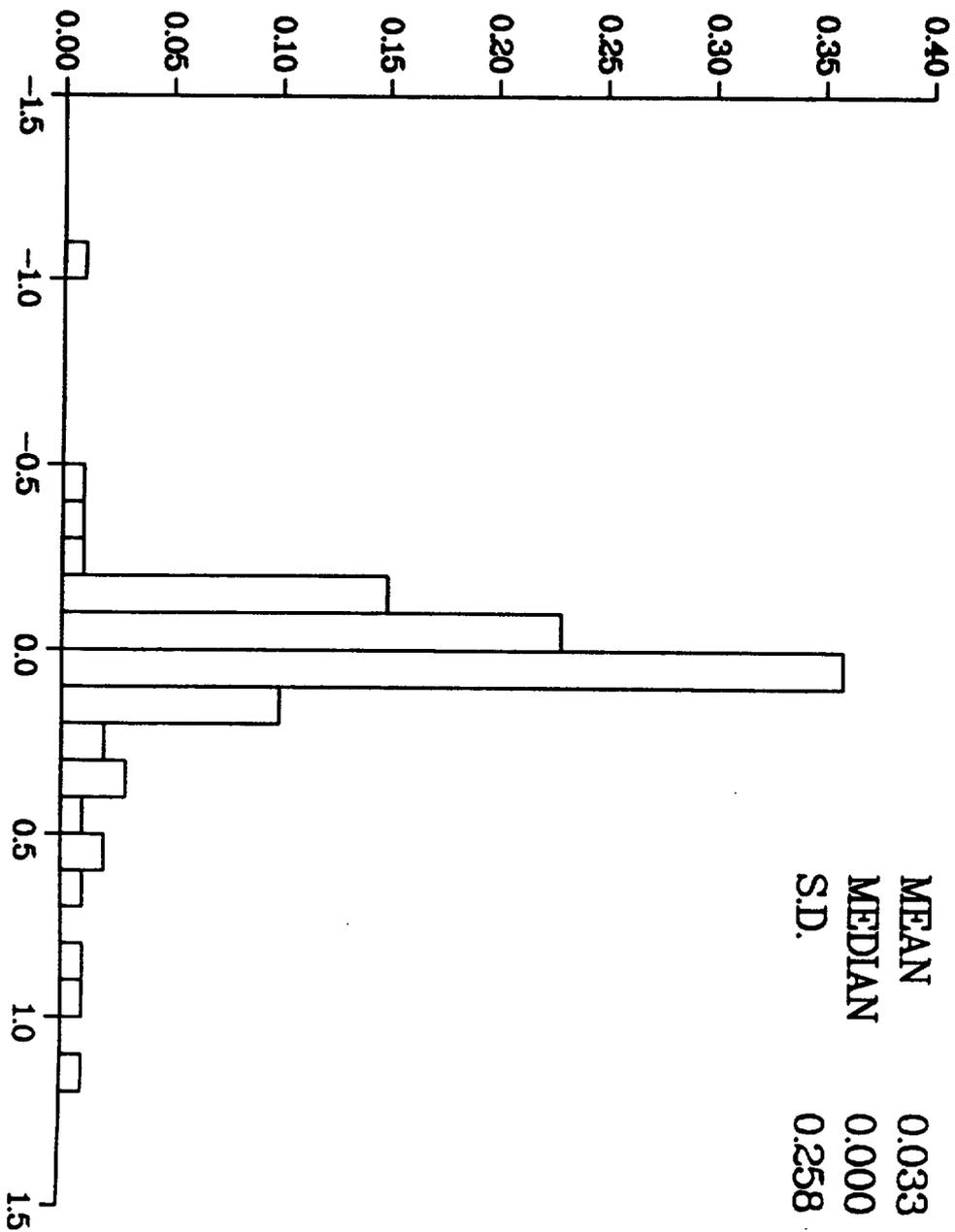
Fig. 12. Histogram of one thousand bootstrapped values of the kriging variance at point C for the simulation illustrated in Figures 10 and 11. A randomly selected factorization of the covariance matrix was used to generate each bootstrap sample.

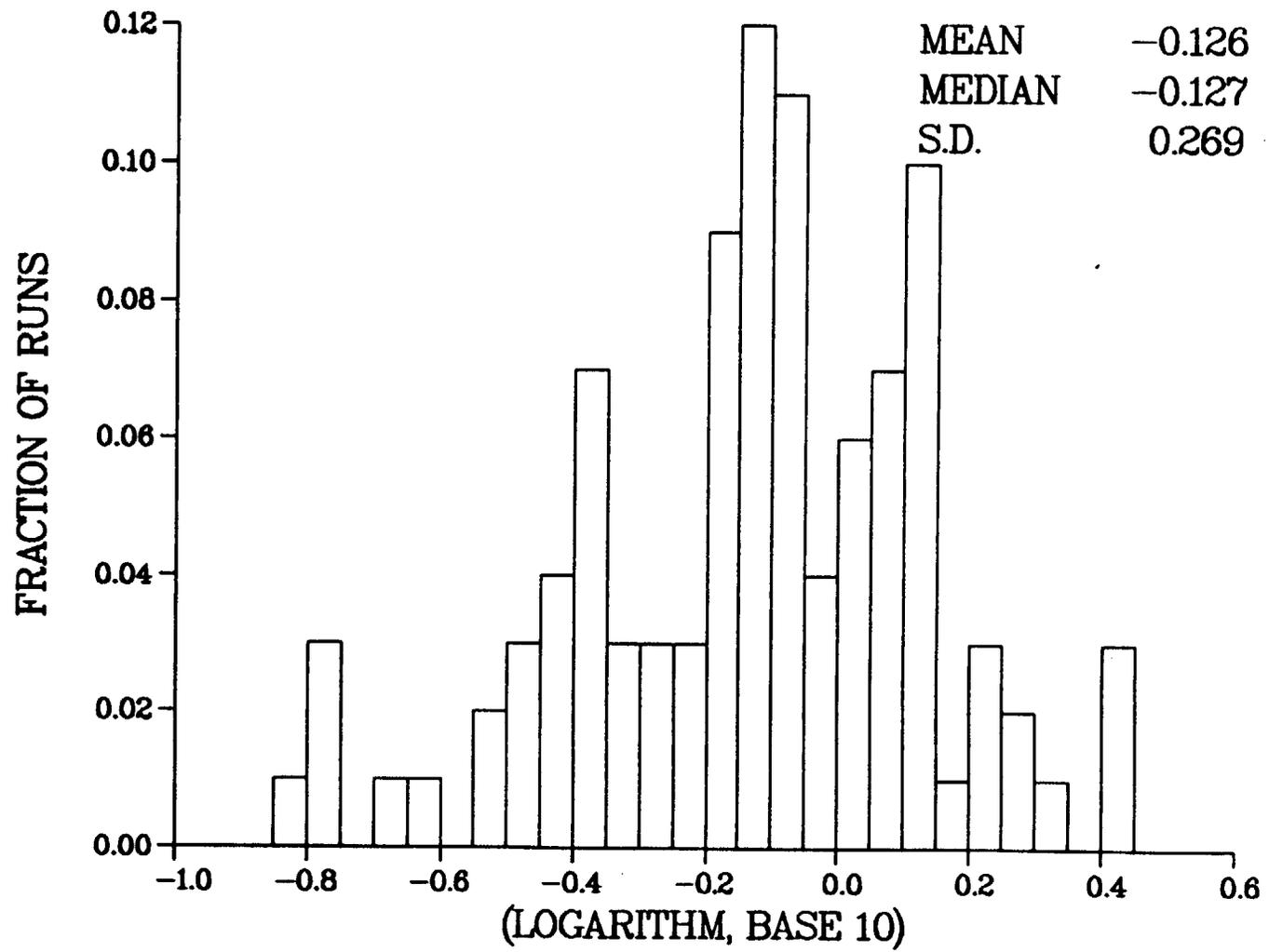


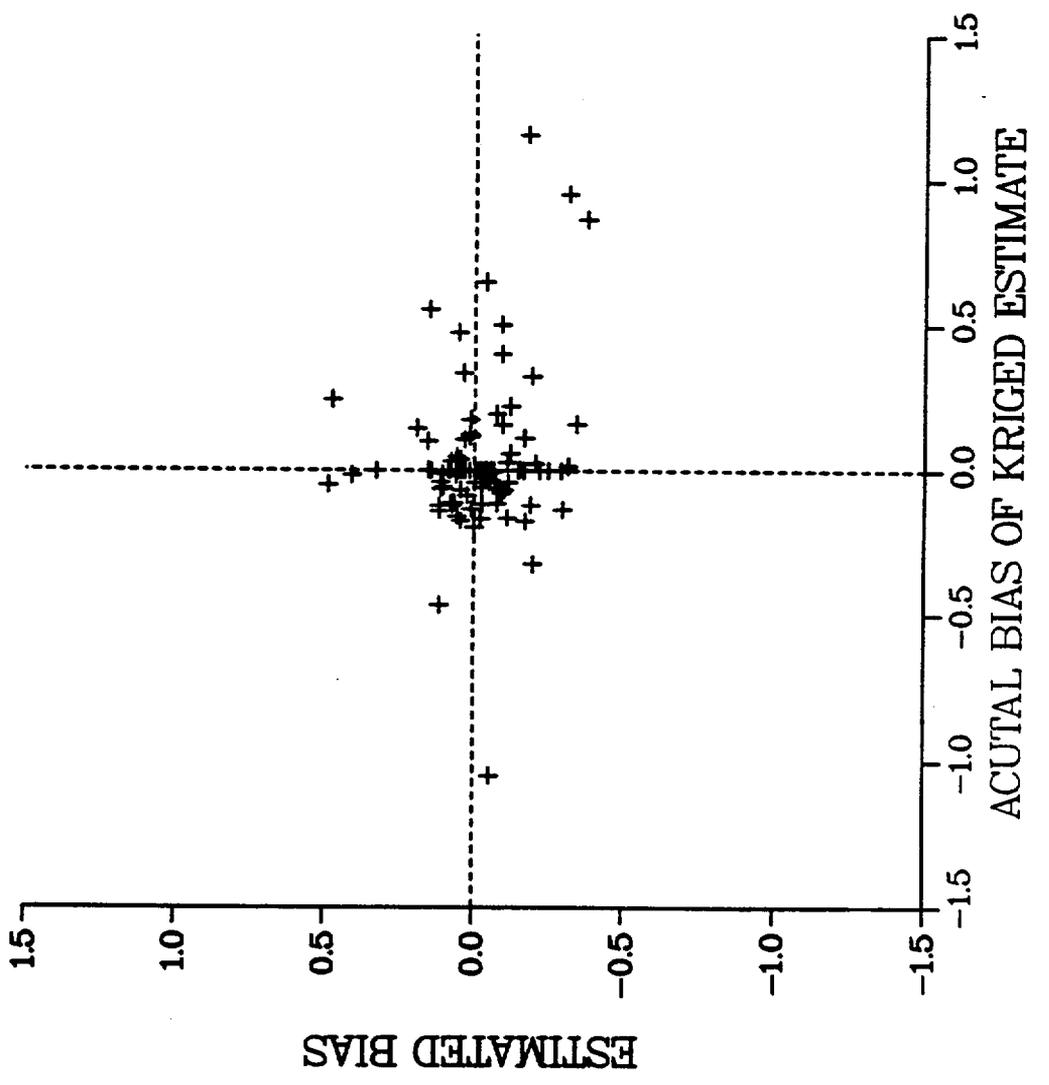


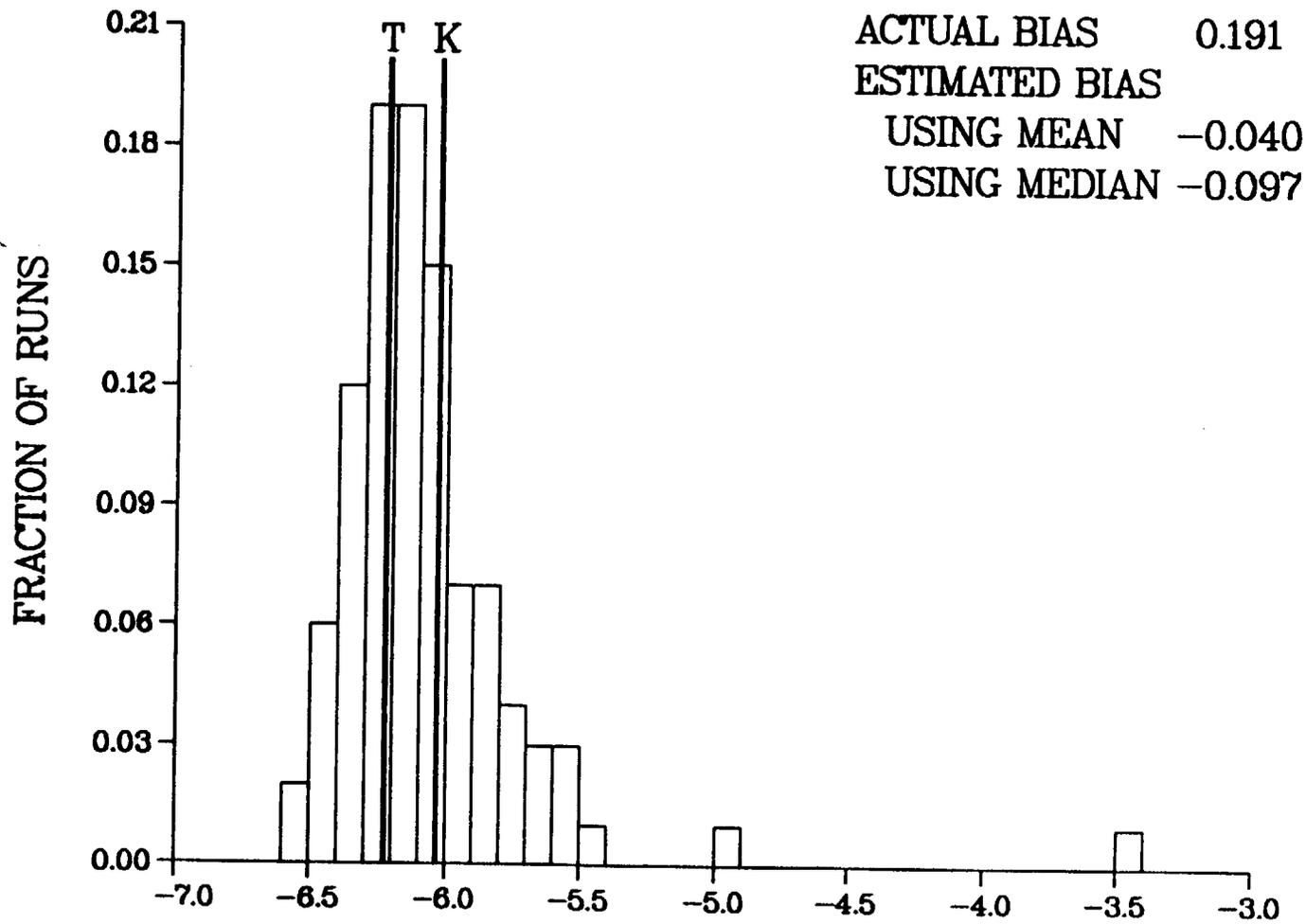


FRACTION OF RUNS









(8)

