[A version of this manuscript has been published *Mathematical Geology* 22:4 (1990) 417-455]

VARIANCE OF GEOSTATISTICIANS

Evan J. Englund

Different individuals will take different approaches to the analysis and interpretation of data. This study attempted to quantify the effect of such individual differences on the quality of geostatistical spatial estimates. Identical spatial data sets were sent to twelve investigators, who independently analyzed the data and produced spatial interpolations. The results varied considerably. Differences in the interpolations could be attributed to differences in choice of methodology, differences in data interpretation, and in a few cases, errors in procedure. The potential differences in economic and societal costs between decisions based on "good" vs. "bad" interpolations warrant a systematic approach to the identification and testing of interpolation methods.

Keywords: geostatistics, kriging, interpolation.

INTRODUCTION

Many types of environmental problems involve the collection and interpretation of spatial data. These may range from local site assessments to regional, or even global investigations. A common factor is spatial interpolation; that is, measurements are made at a number of sample locations and used to estimate values at nearby unsampled points.

While some spatial interpolations are used only for identifying trends or patterns in the data, in many cases interpolated estimates are used directly for making decisions. For example, a contour line drawn at a lead concentration of 1000 ppm may define the portion of a site to be remediated. In such cases, the quality of decisions based on spatial interpolations or estimates is directly related to the quality of the estimates. The quality of the estimates, in turn, depends on the data and the interpolation process.

A variety of spatial interpolation methods are available, ranging from completely subjective manual contouring of data, to completely automated "black box" interpolation by computer. Some of the more commonly used interpolation methods include:

- Nearest neighbor (Polygon method)
- Inverse distance weighted averaging
- Splines
- Polynomial trend surfaces
- Kriging

Most of these are actually classes of methods, with a number of variations available to the investigator.

The spatial interpolation process introduces sources of variability which involve subjective judgement on the part of the investigator. These include:

- Choice of interpolation method
- Deletion of outliers
- Data transforms
- Interpretation of spatial correlation structure

For many situations the U. S. Environmental Protection Agency (EPA) provides extensive guidance aimed at assuring data quality, particularly in the areas of equipment and procedures for sampling, sample preparation, and chemical analysis. At present, , however, there is little guidance relating to the selection and effective use of interpolation procedures. The development of performance-based guidelines would require a costly and time

consuming effort with carefully designed experiments. Before embarking on such an effort, it makes sense to ask whether it is necessary; that is the basic objective of this study. Will estimation procedures performed by different individuals exhibit significant differences in results? Will the economic consequences of differences in decision quality warrant the detailed testing and evaluation necessary to prepare appropriate guidance?

Previous Studies of Spatial Interpolation Methods

Most studies of different interpolation methods have been done in the mining industry. Before kriging became widely known in the United States, Hewlett (1964) compared two interpolation methods and a statistical approach to computing ore reserves for the Silver Bell mine in Arizona. Knudsen, et. al. (1978), and Raymond (1979,198?) compared kriging with the polygonal and inverse distance methods. More recently (Verly and Sullivan, 1984) , investigators have begun comparing the various types of kriging.

Mining production records and detailed production sampling do not provide truly exhaustive information for comparison studies. This difficulty has lead investigators such as Brooker (1978) to compute large simulated deposits for use in comparison studies.

Dahlberg (1972) explored variability in manual contouring by having a twelve-point data set contoured by thirty geologists. Dahlberg compared the manual results with a computer-drawn contour map, and suggested that computer contouring provided an unbiased view of the data which could measure the degree of subjective bias in manual contours.

EXPERIMENTAL PROCEDURES

Approach

A sample data set was drawn from each of two large, "exhaustive" spatial data sets. The same sets of data were sent to twelve Ш. independent investigators, who were asked to provide "geostatistical" estimates for specified sets of cells, or blocks, covering the sampled areas. Objectives of the study were:

- To obtain qualitative information on the range of methodology employed by the investigators.
- To quantify the variability of the estimates.
- To quantify the variability of certain measures of the quality of selection decisions made from the estimates.

The study was exploratory in scope and essentially uncontrolled. Investigators were not given information about the nature of the areas being estimated, other than sample locations and values. They were requested to provide "geostatistical" estimates, and were required to have prior geostatistical experience. However, kriging was not specifically required, leaving room for investigators to use alternative methods. Within the general category of geostatistical estimation, the investigators were free to apply any combination of techniques for eliminating outliers, transforming data, stratifying data, evaluating spatial structure, and computing the interpolated estimates.

positives at various action levels, $\mathsf{etc.}\textsf{;}$ and that results would The investigators were informed that they were participating in a comparison study based on an exhaustively known data set; that a variety of quality measures would be used to evaluate the results, including "mean and variance of the error distribution, false

be published, identified only by code number with participants acknowledged alphabetically. Investigators were unaware of the true values for the blocks.

The Walker Lake Data Set

The data sets used in this study are based on the "Walker Lake" data set, described in detail by Srivastava (1988). The Walker Lake set contains 78,000 values in a 260 x 300 array (Figure 1). The data are highly positively skewed, discontinuous, and exhibit a spatial correlation structure.

Area A

Two rectangular subsets of the Walker Lake data set were prepared as site surrogates. Area A, in the northeast portion of Walker Lake, includes 110 rows and 180 columns, for a total of 19,800 values (Figure 2) . Scaling, multiplication and inversion served to disguise the data set; an investigator with prior knowledge of the Walker Lake set would not easily identify Area A from a relatively small set of samples. The histogram in Figure 3, of 3000 values drawn at random from the Area A data set, illustrates the highly skewed nature of the data distribution.

Four directional variogram and the average omnidirectional variogram in Figure 4 illustrate the spatial correlation structure of Area A. The variogram are computed from all the possible pairs of values from the Area A data set. They show a small but distinct anisotropy consistent with the observed pattern in Figure 2.

Area A has been subdivided into 198 square blocks, each containing 100 data values. The true average block values were computed from the 100 contained data values, and will be compared with corresponding block estimates.

Area B

The Area B data set (Figure 5) in the southeast portion of the Walker Lake area, contains 26,600 values in 190 rows and 140 columns. The values were derived by taking the square root of the Walker Lake values. The histogram (Figure 6) illustrates the resulting less-skewed distribution.

Exhaustive variogram for Area B are shown in Figure 7. Like the variogram for Area A, they show a distinct spatial structure. Area B data have been composite into 266 square blocks.

Representativeness of the Surrogate Site Data Sets

If results from surrogate data sets are to be extrapolated to real-world situations, the surrogate data sets must exhibit realistic characteristics. Of greatest importance for spatial interpolation are skewness and spatial correlation structure. Computed skewness values for several actual environmental data sets are listed below:

- 3.8 (Lead Texas)
- 5.2 (Lead Texas)
- 2.7 (Lead Texas)
- 2.5 (Cadmium Pennsylvania)
- 2.7 (Dioxin New Jersey)
- 1.0 (Cs137 Nevada)
- 3.6 (Ra227 New Mexico)

The skewnesses of 3.3 and 1.2 for Area A and Area B indicate that these data sets are representative of the frequency distributions found in environmental sampling.

Deciding whether spatial correlation structure is representative is more difficult. Variogram of lead concentration in three Dallas Texas sites (Figure 8; after Brown, et. al., 1985) have shapes similar to those from Areas A and B.

The Sample Data Sets

The sample data set from Area A contains 126 samples. Thirty are at random locations in the southwest portion of the area, while the other 96 are on a regular rectangular grid (Figure 9). Each sample location falls within one of the 19,800 cells in the Area A data file. The sample values are the exact values of their corresponding cells; no simulated sampling or analytical errors were added. Figure 10 shows the sample histogram.

The sample set for Area B contains 190 samples on a stratified random grid (Figure 11) . Sample values were assigned in the same manner as for Area A. Figure 12 shows the sample histogram.

When taking relatively small sample sets from areas of high variability, there is always a chance of obtaining a very unrepresentative sample. Because the current study is intended to examine the variability of the investigators' responses to "ordinary" data, an attempt was made to insure a reasonably representative sample. Four different sample data sets fitting the descriptions above were drawn for each of the two areas. The sets were ranked in order of their proximity to the true mean, median, and standard deviation, and the sample set with the best

combined ranking was selected.

Selection of Investigators

Investigators were selected through bids. Sixteen qualifying low bids were accepted; four subsequently withdrew. The minimum requirements for an investigator to participate in the study were a B.S. degree in a scientific or technical field, and some prior training or experience in geostatistics. The 12 investigators are listed below in alphabetical order:

Randal Barnes, Univ. of Minnesota, Ph.D. - Mining Engineering

- Istvan Bogardi, Univ. of Nebraska, Ph.D. Civil Engineering (with Andras Bardossy, Ph.D. - Mathematics)
- David Bowles, Utah State Univ., Ph.D. Water Resources & Hydrology
- James Carr, Univ. of Nevada-Reno, Ph.D. Geological Engineering
- Robert Enwall, Lockheed Engineering and Sciences Co., M.S. Geology
- Marshall Hardy, Applied Research Associates, M.S. , Probability & Statistics
- William Harper, Resource International, Ph.D. Industrial & Systems Engineering

Jonathon Istok, Oregon State Univ., Ph.D. - Civil Engineering (with George Weaver, M.S. - Statistics)

Gerald Jalkanen, Univ. of Arizona, M.S. - Mining Engineering

Y. C. Kim, Univ. of Arizona, Ph.D. - Operations Research & Statistics

Stan Miller, Univ. of Idaho, Ph.D. - Geology

A. W. Warrick, University of Arizona, Ph.D. - Soil Physics

DATA ANALYSIS PROCEDURES

Each investigator provided a set of 198 local block estimates for Area A and 266 for Area B. Each block estimate was compared to the true mean value of the block, and various measures of estimation quality were computed:

Population Measures of Estimation Quality

The population measures are the classical statistical measures of the quality of a set of estimated values. Mean error measures bias of the set of estimates; error variance measures lack of precision of the set of estimates; and "Pearson's r" measures correlation between the true and estimated values.

Conditional Measures of Estimation Quality

Conditional measures evaluate the quality of the set of decisions made when blocks are selected for remediation if their estimated values exceed a specified concentration threshold or action level. The measures will be computed at several different action levels.

Number of False Positives (or False Negatives) - A count of the blocks which are estimated to be above (below) the action level, but which are actually below (above).

False Positive (or False Negative) Deviations - A "false positive (negative) deviation" is the difference between the true value and the action level for a misclassified block. This reflects an assumption that estimation error - the difference between the estimated and true values - is irrelevant once a block has been misclassified. Deviations are summed over all false positive (negative) blocks.

Selection Efficiency - This measure is the ratio (in percent) of the total contaminant in the set of n blocks selected for remediation with the total contaminant in the n highest blocks . It essentially compares what was actually cleaned up with the most that could have been cleaned for the same cost.

Estimation Efficiency - This measure is the ratio (in percent) of the total quantity of contaminant in the n selected blocks to the total quantity estimated to be in those n blocks.

Cost (or loss) functions assign a net economic cost to the set of block remediation decisions. The basic assumption is that society pays a constant unit remediation cost for each block cleaned, and also pays a less easily defined cost (health effects, ecological damage, etc.) for each contaminated block left uncleaned. The total cost of a set of remediation decisions is thus block remediation costs plus the cost to society from unremediated blocks . Three societal cost functions are used, which define cost as proportional to concentration, to concentration squared, or to the log of concentration.

These cost functions assume that the specified action level is society's best estimate of the breakeven point, where the cost of cleaning a block is exactly equal to the cost of not cleaning it. From this viewpoint, a "safety margin" built into an action level allows for the estimated cost of psychological effects on the populace. Ш.

The total cost (L) is expressed in units of constant "block

remediation cost". The concentrations (C) are expressed in units of "action level" by dividing the concentration by the action level. Thus, all cost function curves must pass through the point (1,1) as shown in Figure 13. The equations for the functions are:

The cost for a set of 198 or 266 block decisions is computed by applying the constant remediation cost (1.0) to all blocks estimated to be above the action level, and one of the three 'societal' cost functions to all blocks estimated to be below the action level. To convert the cost to dollars, multiply by the block remediation cost in dollars. Note that this does not include the cost of sampling programs. The goal of any sampling and remediation program would obviously be to minimize the total cost , including the sampling cost. Thus the improvement in decision quality obtained from more samples may or may not be warranted, depending on the sampling cost.

RESULTS AND DISCUSSION

Investigators have been assigned numbers from 1 to 12 in random order, and results from investigator number 1 will be referred to as Al and B1 for areas A and B, respectively.

Variability in Methodology

Because the data sets used by each investigator were identical, variability in the estimated values is due to variability in methodology. Potential sources of variability include the computer hardware (10 of 12 investigators used IBM-compatible Pc's; 2 used VAX's) and software (two investigators used pre release copies of Geo-EAS (Englund and Sparks, 1988) ; the others all used unique software systems) , plus the various choices and interpretations discussed below.

Interpolation Methods and Data Transforms

Of the twelve investigators who submitted interpolated results, eleven used some combination of variogram analysis (or equivalent) and kriging. Several varieties of kriging (Journel and Huijbregts, 1978) were used, including ordinary kriging, log kriging, disjunctive kriging, and indicator kriging. One investigator used a trend surface method, fitting a fifth order polynomial surface to the sample values.

One of the "ordinary kriging" results for Area A is also a "classical statistical" estimate. Investigator number 4 interpreted the spatial structure of the data as random noise, i.e., as a "nugget only" model. With such a model, the kriged estimate for any point is the mean of the samples used; the investigator simply assigned sample mean to each block.

In the data tables, results have been grouped by estimation method. The sample mean estimator is listed as a separate method.

Spatial Structure Analysis

#

Table 1 lists the variogram models used by the investigators. Because the variogram model controls the computation of kriging weights, differences in interpretation of the variogram can lead to major differences in results. The pure "nugget" model (A4) is an extreme example. All of the other investigators found clear spatial structures in Area A; most used the "spherical model" function with a random or nugget component of 10 - 20% of the total sill, and a range on the order of 250 - 500 units. Two investigators used models with no nugget component, representing the other extreme of the variogram model spectrum.

The magnitude of the nugget component relative to the maximum variogram model value is a major factor affecting the amount of smoothing which occurs during kriging. In general, the lower the

nugget component, the less smoothing, and the more the resulting block estimates will honor the nearest sample values. Two investigators used model functions other than the spherical, and three investigators interpreted the data as being anisotropic, that is, showing greater correlation in one direction than another.

The Area B variogram models show less variability than Area, probably because the data are less skewed and there are more samples.

Examples of omnidirectional variogram of the untransformed sample data from Areas A and B (Figures 14 and 15) illustrate the interpretation problem. The variogram computed from sample sets are a rather poor reflection of the exhaustive variogram shown in Figures 4 and 7.

Investigator 4 had to disregard the first (lowest) value in the Area A variogram in order to conclude that no structure existed. This is not as unreasonable as it may appear, because the points on a variogram are not of equal weight. Closely spaced samples are rare. Only a relatively small number of sample pairs are represented in the first point, and those only from the area where the random samples were taken. This is a common problem in variogram modeling - the most crucial portion of the variogram is defined by the least data.

Search Parameters

Kriging and other interpolation methods often use some form of moving window procedure to select a subset of samples in the immediate neighborhood of the estimated point. This is done for practical considerations of computer time and precision rather than on theoretical grounds. The choice of which samples to include in making an estimate can have a major impact on the resulting value. Most search algorithms for selecting the sample subset include some form of circular or elliptical neighborhood, plus a minimum and maximum number of samples to use within that neighborhood.

Table 2 lists the search parameters used by the investigators, again showing considerable variability. Investigator number 7, for example, always uses the closest 15 samples regardless of distance for Area A; investigator number 12 only uses samples within a radius of 225 from the block center, and of those, only the closest 10; and investigator number 3 uses up to 24 samples within a 300 x 600 ellipse. Comparing the parameters for areas A and B indicates that the investigators tend to maintain similar Ш. search strategies from area to area. This may be due in part to the particular algorithms employed in the various kriging programs.

Variability of Block Estimates

Variability in Spatial Patterns

Shaded maps (Figures 16,17,18,and 19) provide a means of comparing the spatial patterns of true and estimated block values. The four shades correspond to the four quartiles of the true block values.

Area A maps show nine estimates with similar patterns, and three which stand out as distinctly different. The most obvious "outlier" pattern is the A4 sample mean estimate; followed by A5, the trend surface estimate. The third "outlier" is A3, one of the log krigings. It has a pattern similar to the other krigings, but biased toward high values. Bias is sometimes introduced in the process of back-transforming estimates after log kriging; this may have happened in this case.

Area A maps also reveal an interesting difference between the ordinary kriging patterns (A2, A8, and A9) and the unbiased log kriging patterns (Al, A6, M, and A12) . High estimates from log kriging seem to form stronger NE-SW trends, while ordinary krigings present a more "patchy" appearance. Because Area A data is nearly log-normal, one might expect the log krigings to be

clearly superior to ordinary kriging, but this is not obvious from the map patterns. Interestingly, both the disjunctive (A1O) and indicator (All) krigings, which use highly transformed data, have patterns which resemble ordinary kriging.

Area B patterns also show three "outliers." The trend surface estimate B5 is again clearly different, and the log kriging B3 is once again biased. One of the ordinary krigings, B6, is more smoothed than the other krigings, suggestive of a relatively high nugget term in the variogram model. However, the variogram models and search parameters provided by the investigator do not confirm this. Perhaps an error in the input parameters for the kriging program is responsible.

The nine other krigings all show quite similar patterns, and unlike Area A, there is no obvious difference between the log (B7 and B12), ordinary (Bl, B2, B4, B8, B9, Bll), and disjunctive (B1O) krigings.

Individual Block Estimates

Table 3 lists values for 15 blocks selected at random from each area to illustrate the variability of individual block estimates. Means and standard deviations are listed, both before and after excluding obvious outlier values. Even after excluding outliers,

standard deviations are relatively high, particularly in Area A where standard deviations for many blocks exceed 50% of the mean. Variability in Area B is lower, with standard deviations ranging from about 5 - 20% of the mean after excluding outliers.

Population Statistics

Table 4 contains univariate statistics for the sample data sets, and true block and estimated block values. Differences between the sample and true block means reflect the bias of the particular sample set drawn. Differences in standard deviation, however, are primarily due to the difference in support (physical size) between the samples and blocks. A block value is the mean of 100 samplesize points. The distribution of block means should have a lower standard deviation, and be less skewed than the original data. Because kriging is a regression technique, the standard deviations of the kriged block estimates are lower than those of the samples.

Estimation Quality: Population Measures

Measures of the overall quality of the populations of block estimates are provided in Table 5. Means and standard deviations for estimation errors are listed, with rankings based on proximity to the ideal target values (mean and standard

deviation equal to zero). Observed values and their corresponding rankings are separated by colons (:) .

Normalized error means and standard deviations are also shown in Table 5. Normalized errors are computed by dividing the observed estimation error by the kriging standard deviation. If all of the underlying assumptions are satisfied and the spatial structure is perfectly known, normalized kriging errors should exhibit a perfect normal distribution. Many investigators did not provide standard deviations, especially in Area A, arguing that the highly skewed distributions made them unreliable. When provided, however, they were reasonable approximations over the entire set of kriged blocks. Because of the missing data, normalized errors were not ranked.

The correlation coefficient between true and estimated values is another measure of the quality of a population of estimates. Correlations and their rankings are included in Table 5.

Estimation Quality: Conditional Measures

Conditional measures of estimation quality evaluate the ability of the estimates to distinguish between blocks with true values above or below some specified cutoff or action level. A good estimator should be consistently good at all possible action levels:

Conditional measures in this study have been evaluated at four different cutoffs for each area, representing approximately the 25th, 50th, 75th, and 90th percentiles.

False Positives and Negatives

Tables 6 - 9 present false positive and negative measures for the four cutoffs. The tables list the numbers of positives and negatives, the numbers of false positives and negatives, the percentages of positives and negatives that are false, and the sums of the false positive and false negative deviations. Rankings of the estimates are also provided.

If only false positives or false negatives are considered, the three measures (number, percent, and sum of deviations) appear to be redundant because their rankings are very similar.

A problem with false positive and false negative measures is that they tend to vary inversely. A high bias, for example, will result in few false negatives, but many false positives. Adding the positive and negative measures is a possible solution, but false positives and false negatives may not be equally bad.

Cost Functions and Other Measures

Cost functions provide a combined measure of the cost to society from remediating selected blocks plus the cost of failing to remediate unselected blocks. They describe the variability in total economic impact among the various spatial estimates. For a perfect set of block estimates, the cost function scores represent the minimum cost outcome for society. For a particular sampling and specified action level, the best interpolation is the one closest to this minimum.

Tables 10 - 13 present the cost function results and rankings along with the selection and estimation efficiency measures described earlier. Also shown are the scores which would be obtained if selection were made on the true block values (perfection); if all blocks were selected regardless of estimated value (clean-everything) ; and if none of the blocks were selected regardless of estimated value (do-nothing). A successful spatial estimate should have a cost function score lower than either of the latter two. Note that for very high and very low action levels, many estimates fail this test.

The selection efficiency measure does not require any economic model. It could also be called a recovery factor. Once the decision, has been made to remediate n blocks, then the best possible outcome is that the blocks selected are the n most contaminated blocks. Selection efficiency is the percentage of the contamination content of the n worst blocks actually contained

in the n selected blocks. A strongly biased estimate with a high correlation coefficient could still rank high in selection efficiency.

Estimation efficiency defines how close the estimated mean of the selected blocks is to the true mean of those blocks. Although desirable, this is not important for most environmental decisions.

Discussion

For all measures of estimation quality, scores obtained by the various estimation procedures vary considerably. Site assessors clearly need to be concerned about the methods used to evaluate data, in addition to methods for collecting and chemically analyzing samples.

Variability of the estimates in this study is due in part to insufficient sampling. If enough more samples were taken, the estimates would converge. However, sampling costs and time constraints are limiting factors in real-world sampling, and when a decision must be made based on the currently available data, it makes sense to use data analysis methods which make the best possible use of the data. Surrogate contaminant data sets, and measures of estimation and decision quality provide a useful tool for studying the effectiveness of interpolation methods, as well

as other significant factors such as sampling designs and data quality.

Although the nature of this study makes it impossible to draw definitive conclusions about the various methods, as each is a unique combination of options, some tentative judgments about overall performance may be suggested.

Trend surface is a poor spatial estimator compared to kriging, although it is possible that a higher-order surface might have produced better results.

The sample mean is a poor spatial estimator compared to kriging.

Ordinary kriging is a relatively consistent, good estimator, even for highly skewed distributions.

Log kriging can be a good estimator for highly skewed distributions, perhaps better than ordinary kriging, but the quality seems to be more variable - the back transform details may be critical.

Disjunctive kriging, based on only two examples, appears to be a good estimator. The results did not differ significantly from ordinary kriging.

Indicator kriging, though used only once, produced very good results.

CONCLUSIONS

Variability in spatial estimation methodology has a significant effect on the quality of the estimates, and on the quality of decisions based on the estimates.

When estimated values are compared to true values, different estimation methods produce markedly different results. In Area A, for example, correlation coefficients between true and estimated values ranged from .00 to .78, while in Area B they ranged from .36 to .75.

When cost functions quantify the combined economic cost of decisions to remediate and not remediate, the relative differences between the highest and lowest costs, measured at several different action levels, ranged from 4 - 75 % (with one extreme of 980%) in Area A, and from 4 - 29% in Area B. There are probably thousands, of contaminated sites in the United States alone for which spatial interpolation from sample data will be required. Total remediation costs for such sites could easily reach billions of dollars. Failure to use appropriate interpolation techniques

will result in significantly increased costs.

No single spatial estimate was consistently best or worst for all quality measures. The rank order of the spatial estimates changes significantly when different measures are used. Deciding which measure of estimation quality is most relevant to the particular circumstances of a site investigation is crucial to selecting the "best" interpolation method.

NOTICE

Although the research described in this article has been supported by the United States Environmental Protection Agency, it has not been subjected to Agency review and no official endorsement should be inferred.

REFERENCES

- Brooker, P. I., 1978. Geostatistical Calculations Using True and Estimated Values from a Simulated Deposit. Proc. Australas. Inst. Min. Metall. , No. 268, p. 63-69.
- Browm, K. W., Mullins, J. W., Richitt, E. P. Jr., Flatman, G. T., Black, S. C., and Simon, S. J., 1985. Assessing Soil Lead Contamination in Dallas Texas. Environmental Monitoring and Assessment, V. 5, p. 137-154.
- Dahlberg, E. C., 1972. Aspects of Unbiased and Biased Contouring of Geological Data by Human and Machine Operators. Geol. Sot. Amer. Abstracts, V. 4, p. 482-483.
- Englund, E. J. and Sparks, A. R., 1988. Geo-EAS (Geostatistical Environmental Assessment Software) User's Guide. EPA/600/4-88/033. U.S. EPA Environmental Monitoring Systems Laboratory, Las Vegas. 170 p.
- Hewlett, R. F., 1964. Comparison of the Triangular, Polygonal, and a Statistical Method of Computing Grade and Tonnage of Ore for the Silver Bell Oxide Porphyry Copper Deposit. U.S. Bureau of Mines Report of Investigations 7331.
- Journel, A. G. and Huijgbregts, Ch. J., 1978. Mining Geostatistics. Academic Press, London. 600 p.
- Journel, A. G., 1973. Geostatistics for Conditional Simulation of Orebodies. Economic Geology, V. 69, p 673-687.
- Knudsen, H. P., Kim, Y. C., and Mueller, E., 1978. Comparative Study of the Geostatistical Ore Reserve Method over the Conventional Methods. Mining Engineering, Jan 1978, p. 54-58.
- Matheron, G., 1963. Principles of Geostatistics. Economic Geology, V. 58, p. 1246-1266.
- Raymond, G. F., 1979. Ore Estimation Problems in an Erratically Mineralized Orebody. CIM Bulletin, Jun 1969, p. 90-98.
- Raymond, G. F., 1982. Geostatistical Production Grade Estimation in Mount Isa's Copper Orebodies. Proc. Australas. Inst. Min. Metall., No. 284, P.17-39.
- Srivastava, R. M., 1988. A Non-Ergodic Framework for Variogram and Covariance Functions. SIMS Technical Report No. 114.

------ Number

Figure 1. Shaded map of the Walker Lake data set

Figure 2. Shaded map of the Area A data set.

Figure 3. Histogram of 3000 sample values drawn from the Area A data set.

FIGURE 4

FIGURE 5

FIGURE 7

FIGURE 8

FIGURE 12

FIGURE 13

 $\overline{\mathbf{H}}$

 \mathbf{E}

 $\overline{\mathbf{R}}$

True Value

 \overline{R}

FIGURE 19

TABLE 1b: Variogram Models of Spatial Structure for Area B

| | Search Ellipse # of Samples | | | | | |
|----------------------|-----------------------------|-----|-------------|-----|----|--|
| | | | | | | Major Minor Angle Max Min Data-Transform |
| Ordinary Kriging | | | | | | |
| A2 | 250 | 250 | na | 49 | 17 | none |
| A8 | 350 | 350 | na | all | 17 | none |
| A9 | 400 | 400 | na | 18 | | none |
| Log Kriging | | | | | | |
| A1 | 320 | 320 | na | 24 | | Natural logs |
| A3 | 600 | 300 | N45E | 24 | | Natural logs |
| A6 | 400 | 300 | N45E | 12 | 1? | Natural logs |
| A7 | na | na | na | 15 | 15 | Natural logs |
| A12 | 225 | 225 | na | 10 | 1 | Natural logs |
| Disjunctive Kriging. | | | | | | |
| A10 | nđ | nd | nd | nd | nd | Gaussian |
| Indicator Kriging | | | | | | |
| A11 | 300 | 300 | na | 18 | 2 | Indicator |
| Trend Surface | | | | | | |
| A5 | na | na | na | na | na | Natural log |
| | | | | | | |
| Sample Mean | | | | | | |
| 84 | na | na | na | na | na | none |

TABLE 2a: Interpolation Options for Area A

TABLE 2b: Interpolation Options for Area B

| | Search Ellipse # of Samples | | | | | |
|----------------------|-----------------------------|-----|---------------------------|-----|----|----------------|
| | | | Major Minor Angle Max Min | | | Data Transform |
| Ordinary Kriging | | | | | | |
| 81 | 255 | 255 | na | 24 | | none |
| 82 | 350 | 350 | na | 49 | 1? | none |
| B4 | 300 | 300 | na | 50 | | none |
| 86 | 450 | 321 | N30E | 14 | | none |
| 88 | 225 | 225 | na | all | | none |
| Ŗ9 | 400 | 400 | na | 18 | | none |
| 811 | 350 | 350 | na | 12 | 2 | none |
| Log Kriging | | | | | | |
| B3 | 400 | 300 | N10E | 24 | 1 | Natural logs |
| 87 | na | na | na. | 15 | 15 | Natural logs |
| B12 | 200 | 200 | na | 10 | | Natural logs |
| Disjunctive Kriging. | | | | | | |
| B10 | nd | nd | nd | nd | nd | Gaussian |
| Trend Surface | | | | | | |
| 85 | na | na | na | na | na | Natural logs |
| | | | | | | |
| | | | | | | |

TABLE 3a: Example Block Estimates from Area A

TABLE 3b: Example Block Estimates from Area B

TABLE 4a: Population Statistics on Block Estimates for Area A

TABLE 4b: Population Statistics on Block Estimates for Area B

| | Percentiles | | | | | | | | |
|----------------------|-------------|---------------------------------------|--------|--|------|------|-----|---|------|
| | | Mean St.D. | Min 25 | | 50 - | 75 — | Max | Skew. Kurt. | |
| Sample Data Set | | | | | | | | 10.58 10.2 0.0 3.3 7.4 16.3 54.8 1.38 5.19 | |
| True Block Values | | | | | | | | 11.03 6.78 0.0 5.8 10.4 14.8 29.2 0.56 2.68 | |
| | | | | | | | | | |
| Ordinary Kriging | | | | | | | | | |
| 81 | | | | | | | | 10.71 6.09 1.4 6.0 9.7 14.2 36.8 1.06 4.23 10.68 5.54 1.9 6.3 9.7 14.0 32.3 1.00 | 3.97 |
| B2 | | | | | | | | | |
| B4 | | | | | | | | 10.68 5.89 1.5 6.1 9.7 13.9 35.6 1.04 | 4.18 |
| 86 | | | | | | | | 10.60 4.51 2.5 7.2 10.0 13.3 26.2 0.81 | 3.59 |
| 88 | | | | | | | | 10.69 5.83 0.8 6.3 9.6 13.9 38.6 1.13 4.87 | |
| 89 | | | | | | | | 10.72 5.55 1.5 6.5 10.0 14.1 33.2 0.96 | 4.02 |
| B11 | | 10.70 5.51 1.8 6.3 9.8 13.9 33.2 0.98 | | | | | | | 3.93 |
| Log Kriging | | | | | | | | | |
| B3 | | | | | | | | 13.78 6.06 4.2 8.9 12.2 18.0 32.3 0.65 2.56 | |
| B7 | | | | | | | | 10.65 6.57 1.0 5.5 9.4 14.0 35.8 1.10 | 4.17 |
| B12 | | 10.85 7.17 1.7 5.2 9.6 14.2 49.2 1.44 | | | | | | | 6.08 |
| Disjunctive Kriging. | | | | | | | | | |
| B10 | | | | | | | | 10.56 4.65 0.7 7.2 10.1 13.1 26.1 0.51 | 3.14 |
| Trend Surface | | | | | | | | | |
| B5 | | | | | | | | 10.81 3.86 0.6 8.4 10.9 13.3 39.9 0.45 5.38 | |
| | | | | | | | | | |

TABLE 5a: Population Quality Measures for Area A Block Estimates

TABLE 5b : Population Quality Measures for Area B Block Estimates

 \sim

TABLE 6a: False Positives and Negatives for Area A
Action Level = 15 (25th Percentile)

TABLE 6b: Faise Positives and Negatives for Area B
Action Level = 6.0 (25th Percentile)

TABLE 7a: False Positives and Negatives for Area A
Action Level = 70 (50th Percentile)

TABLE 7b: False Positives and Negatives for Area B
Action Level = 10.4 (50th Percentile)

TABLE 8a: False Positives and Negatives for Area A
Action Level = 180 (75th Percentile)

TABLE 8b: False Positives and Negatives for Area B
Action Level = 14.8 (75th Percentile)

 $\sim 10^7$

TABLE 9a: False Positives and Negatives for Area A
Action Level = 300 (90th Percentile)

TABLE 9b: False Positives and Negatives for Area B
Action Level = 21.5 (90th Percentile)

TABLE 10a: Loss Functions and Other Conditional Measures of Quality
for Block Estimates from Area A
Action Level = 15 (25th Percentile)

TABLE 11a: Loss Functions and Other Conditional Measures of Quality
for Block Estimates from Area A
Action Level = 70 (50th Percentile)

TABLE 11b: Loss Functions and Other Conditional Measures of Quality
for Block Estimates from Area B
Action Level = 10.4 (50th Percentile)

TABLE 12a: Loss Functions and Other Conditional Measures of Quality
for Block Estimates from Area A
Action Level = 180 (75th Percentile)

* parameter not defined; median rank assigned.

TABLE 13a: Loss Functions and Other Conditional Measures of Quality
for Block Estimates from Area A
Action Level = 300 (90th Percentile)

* parameter not defined; median rank assigned.

