SPATIAL AUTOCORRELATION: IMPLICATIONS FOR SAMPLING AND ESTIMATION

Evan J. Englund
U.S. Environmental Protection Agency, P.O. Box 93478, Las Vegas, NV 89193-3478

ABSTRACT

Spatial autocorrelation is a normal result of the physical and chemical processes which operate in the environment. All measurable environmental parameters will exhibit spatial autocorrelation at some scale. This not only causes technical problems In sampling and estimation, but leads to more fundamental problems In communication. Terms such as “chemical concentration”, "representative sample”, and “frequency distribution” which are commonly used and well understood in laboratory situations, can become essentially meaningless when applied to environmental measurements without an explicit statement of the spatial scale (support) being considered.

A simulated spatially autocorrelated distrlbution is used to illustrate the changes in concentration, frequency distribution, and sample quality associated with changes in support. Variograms computed from the simulated data illustrate the relationship between spatial variability and standard QA/QC. Practical suggestions are made for sampling and estimation of spatially autocorrelated sites.

INTRODUCTION

The purpose of environmental sampling programs, like other samplings, is to use the information obtained from the sample to make inferences about the larger population from which the sample is drawn. For example, an industrial waste lagoon may be sampled to determine whether its mean concentration exceeds an allowable maximum value for a particular chemical. In other cases, sampling may be done to determine whether the level of pollution has increased or decreased from previous levels at a site, or to identify the locations of polluted areas. A typical sequence of events in an environmental investigation might include the design of a sampling plan, collection of samples, laboratory chemical analysis, Interpretation of data, and finally, a decision based on the interpretation.

Decisions made by various federal, state, and local agencies may require remedial, preventative, or punitive measures which may have substantial consequences in terms of human health as well as economics. It is obviously important that environmental investigations be conducted in such a manner as to ensure that decisions are based on the best possible information. It should be emphasized here that the information on which decisions are made is not the sample data itself, but the interpretation of the data; that is, the estimates of the characteristics of the larger population which are made from the data.

Unfortunately, the interpretation of field data is made difficult by the nature of the environment itself, which is not always amenable to the methods of sampling and data analysis designed for use under more controlled conditions such as laboratory experiments. Although the same sampling and data analysis terminology is used in both the laboratory and the field, the precise meanings of many terms are significantly different. The resulting ambiguities can lead to problems in both the design of sampling programs, and the interpretation of the results. This paper will explore the causes of these problems, attempt to clarify the terminology, and offer some practical suggestions for sampling design and data interpretation.

SPATIAL AUTOCORRELATION

Spatial autocorrelation is the basic cause of the problems to be discussed. All measureable environmental parameters exhibit spatial (and temporal) autocorrelation at some scale. This means that over some range of distances, measurements will tend to be more similar to measurements taken nearby than to measurements taken farther away. For practical purposes, stating that a phenomenon is spatially autocorrelated is equivalent to stating that it is not uniformly distributed, but autocorrelation is more easily quantifled In statistical terms. The term spatial correlation is often used interchangeably with spatial autocorrelation, but the former implies that the measured values are correlated with their locations, as temperature, for example, is correlated with distance from the equator.

The physical and chemical processes that control the fate and transport of chemicals In the environment do not operate at random, although most events include what may be considered random processes, In the sense that they are too complex to be predicted in detail. Random and deterministic processes may operate simultaneously at several scales to produce the phenomenon being measured. The measurement of precipitation (with or without acid), provides a good illustration. Regional weather patterns produce the conditions necessary for rainfall over a large area. Within the area, individual clouds or clusters of clouds form, apparently at random locations, due to local fluctuations in temperature and humidity. If one collected pairs of rain-gauge measurements taken at various spacings, one would expect to observe strongly correlated readings at separations much smaller than the size of the average cloud, weaker correlation at regional-scale separations, and essentially zero correlation for separations larger than the size of the region. Although contiguous rain gauges would be expected to show the greatest correlation, it would not be perfect, due to the fact that the two guages may not receive identical amounts of rain (spatial variability), and even if they did, the readings may not be identical (measurement error).

After a major rainfall, for example, we might be asked to estimate the average precipitation over a watershed from a given set of rain gauge readings, in order to estimate the watershed’s contribution to a flood crest. Alternately, we might want to know how many gauges we would need in order to predict flooding with a specified degree of accuracy. The methodology commonly known as geostatistics (Matheron, 1963; Journel and Huijbregts, 1978) was developed in the mining industry to deal with questions of local ore grade estimation, and has since been shown to be

generally applicable to most situations where spatial autocorrelation is present.

A basic assumption in geostatistical analysis is that the spatial autocorrelation exhibited in a set of measurements can be represented by an underlying autocorrelation function which is valid over the region of interest. For many environmental phenomena, this assumption can be intuitively related to the controlling processes.

The variogram is one commonly used method for quantifying spatial autocorrelation (Fig.l). Experimental variograms are computed from sample data by examining all possible data pairs, grouping them by distance classes, and computing a variance for each distance class using the standard paired-sample formula. A theoretical model is then fitted to the experimental data. A variogram contains exactly the same information as a plot of correlation coefficients for the same distance classes (one can be transformed into the other by inverting the plot and resealing the y-axis).

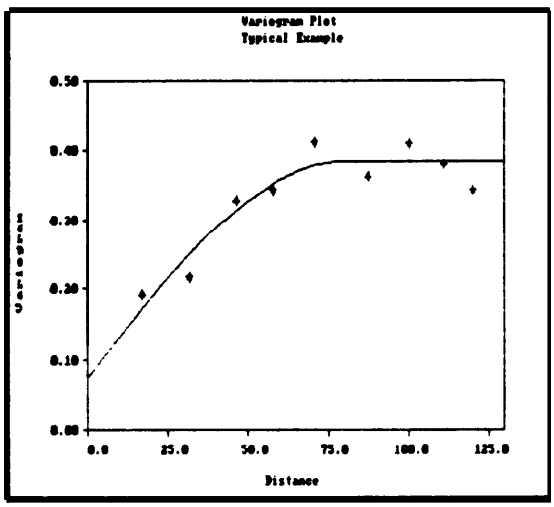


Fig. 1. Typical variogram plot and fitted model. Variances are computed from paired sample differences for pairs in successive distance classes and plotted against distance. The fitted model exhibits commonly observed features: a random component or “nugget” at the y-axis intercept, and an increase in variance with distance up to a maximum “range” of autocorrelation.

SAMPLE SUPPORT

When spatial autocorrelation is present, the physical dimensions of the sample become important considerations. The term 'support', defined as the “size, shape, and orientation of the physical sample taken at a sample point” (Starks ,1986), is used to avoid confusion with the statistical size (number of observations) of a sample. The support of a soil core, for example, would be its diameter and length. In the case of a rain gauge measurement, the support would include the diameter of the orifice, and the time of accumulation.

When the sample support changes, the statistical properties of the sample set such as the variance of samples and the sampling variance also change. These changes make sample support a critical element In the design of a sampling program. The concept of support applies equally to the stage of data interpretation. The support on which decisions will be made is rarely the same as that of the samples. The choice of decision support can significantly affect the outcome of an analysis, and should be considered before a sampling program is undertaken. The idea of support and its ramifications will be developed further in the following sections.

CHEMICAL CONCENTRATION

The term ‘chemical concentration’ is meaningless in the absence of a specified support. Atoms and molecules represent the smallest scale at which elements and compounds can be said to exist. If samples are taken at this scale, the true concentration of any substance within the sample will be a discrete binary phenomenon - the concentration will be either 100% or 0%. Any larger sample is made up of a mixture of discrete components, and its true concentration will be the sum of the weight or volume of the analyte divided by the total weight or volume of the sample; i.e., the average over the sample support. Usually the entire volume of a sample is not measured directly, but the measured concentration of a subsample is used to estimate the the mean concentration of the original sample. Great care is taken during the preparation of a sample to ensure uniform mixing, and if necessary, to reduce the particle size of the material so that any subsample used for analysis will have a true mean concentration very close to the true mean of the entire sample. If the mixing is effective, neither the size of a subsample nor its location should have a significant effect on the outcome of the analysis...for practical purposes, one subsample is as good as any other.

In the field, where the area being investigated cannot be uniformly mixed, the situation is quite different. Layers, crystals, clumps, or other high concentrations of a substance of ten occur such that if a given sample had been taken at a slightly different location (sometimes only a fraction of the sample size away), a significantly different true sample concentration would be found. The classic example of this was observed in placer gold deposits, where the presence or absence of a single, miniscule gold nugget in a sample would make the difference between an assay indicating high-grade ore, and one indicating waste. This led to the interesting term ‘nugget effect’ often being applied to the y-axis discontinuity in variogram models.

Like changes in location, changes in support also result in changes In concentration. The true value of a point sample is 0 or 100%; the true value of any larger volume centered on the point is the mean concentration over the volume. If the dimensions of a sample are increased or decreased; If the shape of the sample is changed, say from a sphere to a cube; or if the orientation of a non-spherical shape is changed: the sample will contain a different set of molecules, and probably a different true concentration. For any point in space which represents a potential sample 'location', an infinite number of possible sample supports exist centered on the point, and an infinite number of possible true concentrations which can be said to be the concentration at that location. Obviously, we must conclude that any reported measurement of chemical concentration in the environment is essentially meaning-

less unless the support is also reported. Likewise, a statement such as "remedial action will be taken if the concentration of cadmium in soil exceeds 500 ppm" is also meaningless unless the support is specified.

A SIMULATED EXAMPLE

An example based on a simple computer simulation serves to illustrate the support problem. A blank computer screen is 'polluted' with a 'realistic-looking' pattern of pixels (Fig. 2). The algorithm which was used to generate the pattern first selected 25 points at random in the central part of the screen. Each of these points was used as the center of a cluster of up to 2000 points scattered around the center at random angles and approximately normally distributed distances (sum of three uniformly distributed random values). Blank pixels were initialized with a value of zero, and incremented by one each time they were hit by the point generator. A color terminal can be used for a more effective display than Fig. 2, because pixel values greater than one can be represented by various color codes. The details of the algorithm are not crucial. Any algorithm which conditions the outcome of a random process on a prior outcome of the same or another random process, can be used to generate spatially autocorrelated patterns. Variations on the drunkard's walk (e.g., two steps forward, three steps back) are effective.

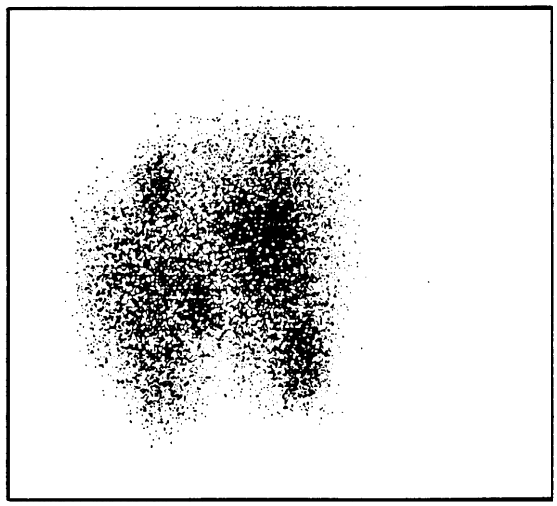
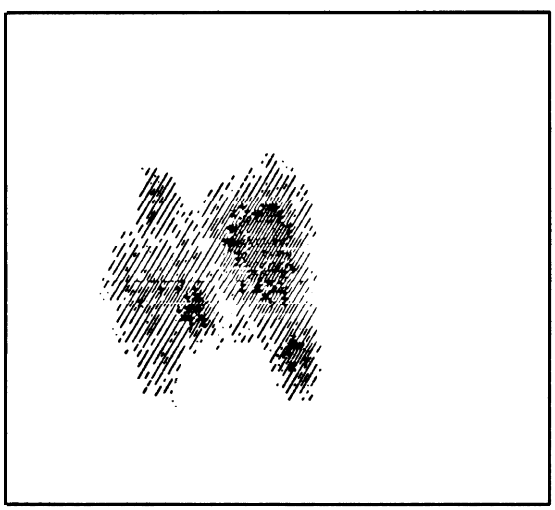


Fig. 2. A simulated, spatially autocorrelated distribution of "pollutant".

When the pattern has been generated, we have an area subdivided into pixel-sized units, for each of which we know the exact pollution concentration. The pixel scale of support is considered to be smaller than any support site we would be interested in, but since we have exact knowledge at that scale, we can now combine pixels into any larger support areas we choose, and compute the exact average pollution concentrations over these areas.

Figs. 3-7 illustrate the results of this kind of change In support. In Fig. 3, the screen area was divided into support blocks of 5x5 pixels, and the mean pixel value was computed for each block. The block means were grouped into class intervals and represented as shaded patterns on the map. The histogram scaled by area, and univariate statistics, are also shown for the set of non-zero value blocks. Figs. 4-7 repeat this process at supports of 10x10, 20x20, 40x40, and 80x80 pixels, respectively.



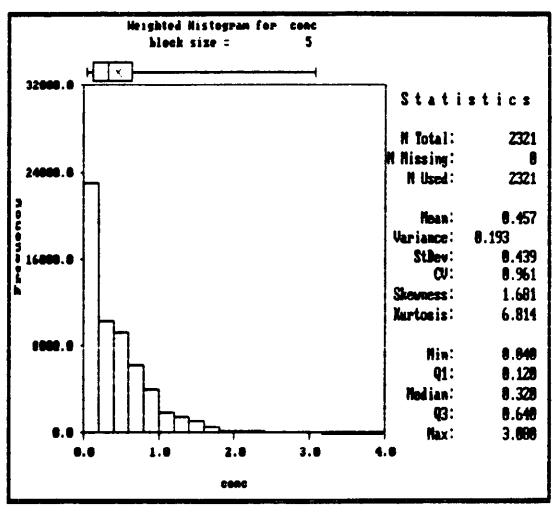
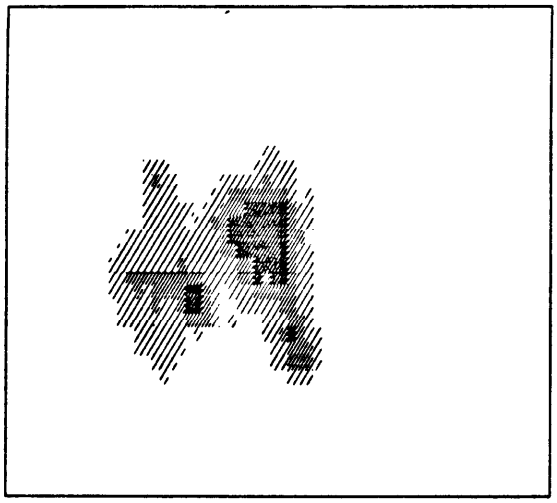


Fig. 3. Map and histogram of true "concentrations" of simulated pollutant averaged over 5x5 pixel blocks. Darker shades on map represent higher values: Blank (<0.4 units per pixel); Light shade (0.4 to 0.8); Medium shade (0.8 to 1.2); Dark shade 01.2).



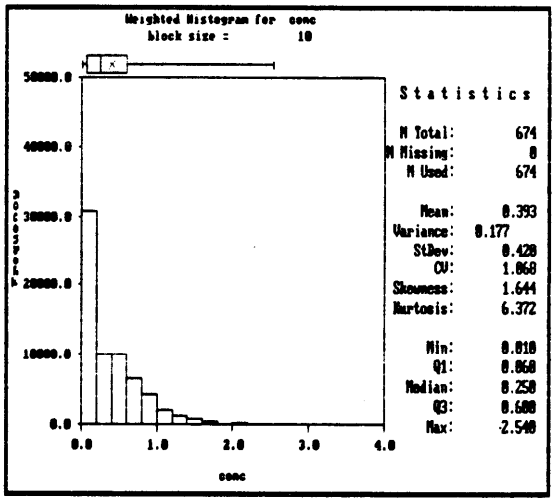
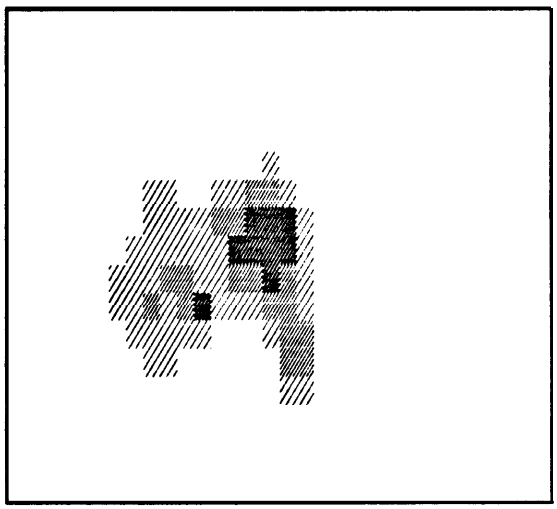


Fig. 4. Map and histogram of true "concentrations" of simulated pollutant averaged over 10x10 pixel blocks. Shades represent the same values as in Fig. 3.



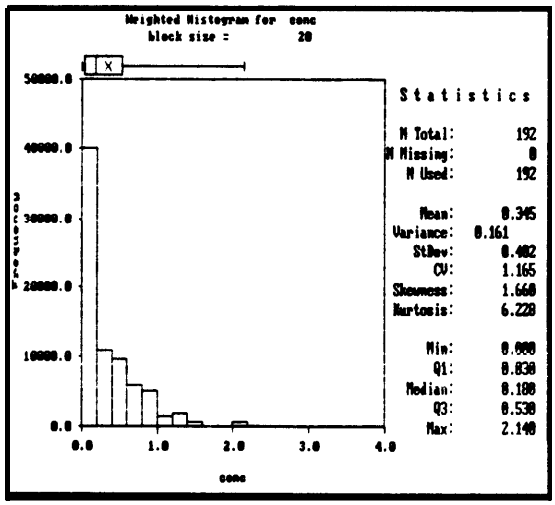
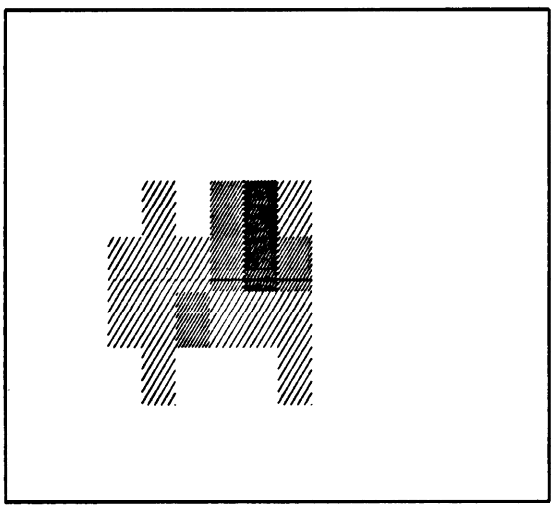


Fig. 5. Map and histogram of true "concentrations" of simulated pollutant averaged over 20x20 pixel blocks. Shades represent the same values as in Fig. 3.



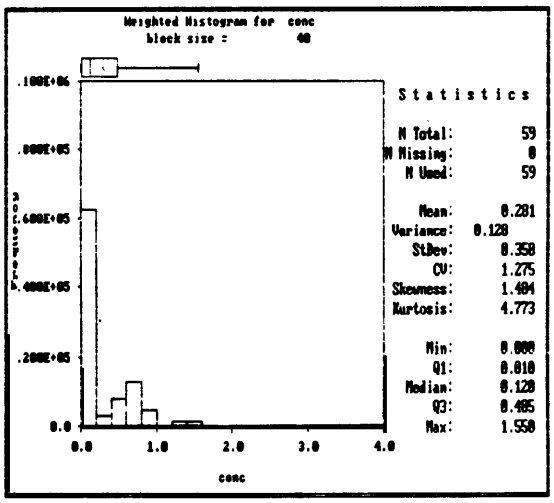


Fig. 6. Map and histogram of true "concentrations" of simulated pollutant averaged over 40x40 pixel blocks. Shades represent the same values as in Fig. 3.

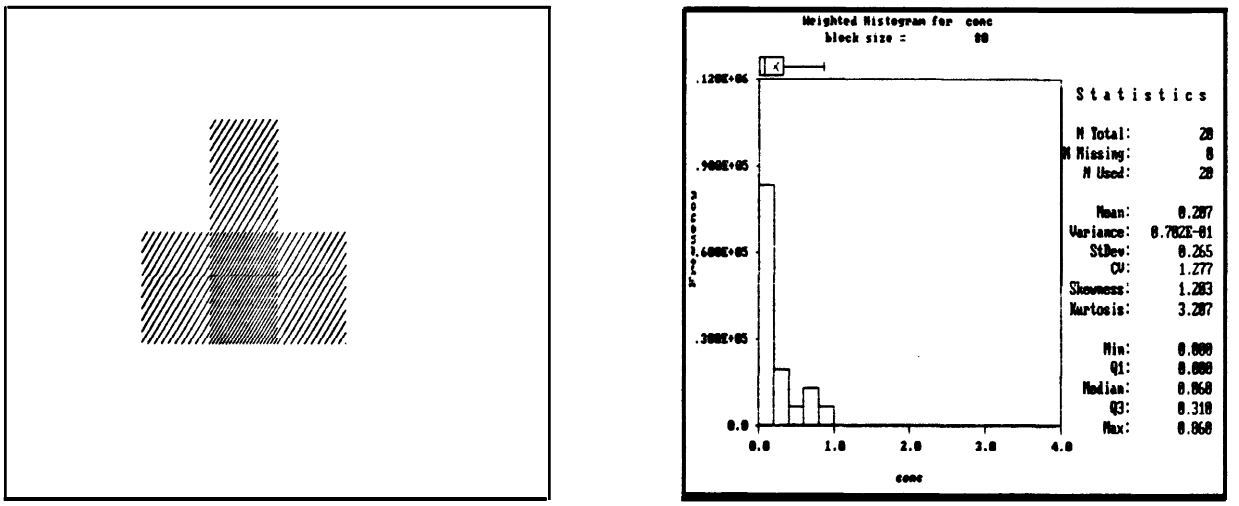


Fig. 7. Map and histogram of true "concentrations" of simulated pollutant averaged over 80x8 pixel blocks. Shades represent the same values as in Fig. 3.

The most significant thing to remember when examining these results is that these are not estimates of the underlying distribution. Each of these is the true distribution for the specified support and location. Additional, equally true distributions can be generated by shifting the grid origin, by filling in the missing block sizes (6x6, 7x7,...), by changing the shape (4x6, 8x12,...), etc. Continuing our analogy between the simulation and a polluted site, we must ask ourselves, as site investigators, which truth do we want to know? Or perhaps more realistically, which truth do we need to know, and what do we have to do to know it with sufficient accuracy?

If we want to take some remedial action, say, by cleaning up those areas in excess of an action level of 1.2, we must first specify the remediation support. If we specify 5x5 pixel blocks as in Fig. 3, our goal would be to identify as polluted, and clean up, all of the darkest blocks. If, on the other hand, we specify the 80x80 pixel blocks, we would find that there is nothing to clean up at all. If a support is not specified along with an action level in the applicable regulations, the site investigator is faced with the problem of establishing one, either explicitly or implicitly. Along with the flexibility comes the cost of ambiguity: long meetings spent arguing over semantics; or perhaps extensive litigation.

FREQUENCY DISTRIBUTIONS

As is evident from the maps and histograms in Figs. 3-7 when the support changes, the frequency distribution also changes. Within the domain defined by the screen boundaries, as the support dimension increases, the variance of the distribution decreases. This is particularly important in the context of regulation and remediation, because, as was suggested above, increasing the support dimension can often reduce (or eliminate) the amount of material above an action level. It is interesting to note that support is frequently specified in the time dimension for air quality monitoring and regulation; it is common to refer to 1-hour averages, 24-hour averages, etc., recognizing that these would give quite different results if compared against a single concentration limit.

From a practical standpoint, as mentioned above, specifying the support associated with an action level will reduce ambiguity in site assessment and enforcement. At a more fundamental level, support is a significant factor in the establishment of action levels. Environmental actions are ultimately decisions about risk abatement, and the risks associated with chemical substances are in turn functions of toxicity and exposure, both of which involve questions of support. If two one-hectare lots each have mean concentrations of 1000 ppm lead in the upper ten centimeters of soil, are the risks associated with them significantly different if the distribution of one-square-meter supports for one lot is normal with a standard deviation of 10, while the other is log-normal with a standard deviation of 800? If remedial action were practical only at the hectare support, should the action decision be based on the mean concentration over the hectare, on the frequency distribution of some smaller support within the hectare, or on some combination?

An observation worth noting about the frequency distributions In Figs. 3-7 is that although the variance of the distributions decreases as the size of support increases, the shape of the histogram seems to remain relatively constant. If the spatial distribution were random rather than autocorrelated, the variance would be expected to decrease faster as the support increased, and the shape of the histogram would be expected to converge more rapidly toward the normal. The persistence of histogram shape in spatially autocorrelated fields suggests that affine correction of variance (Journel and Huijbregts, 1978, p. 471) may be a useful tool for estimating the frequency distribution of one support, given the frequency distribution of another.

REPRESENTATIVE SAMPLES

'Representative' is one of the most misused terms in environmental sampling (Splitstone, 1987). In this case, the effects of spatial autocorrelation in the field combine with inconsistency In the use of the term 'sample' itself, resulting in a great deal of confusion. A statistical sample of n observations is sometimes said to be representative if the mean, and or other parameters of its frequency distribution are 'similar' to the population from which it was drawn. In a somewhat more abstract sense, the distribution might be called representative if it was properly drawn (that is, if each member of the population had an equal chance of being included in the sample), even though it is not similar to the population.

An environmental sample such as a soil core, collected In the field and sent to a lab for analysis, is not a 'sample' in the commonly used statistical sense. It might be more appropriately called a member of a population which has been included in a sample of that population. A set of n such environmental samples, or observations, make up a statistical sample which can be representative of the population of possible observations of the same support within the sampled domain. It can not, however, be directly representative of any other support population, except for the population mean, which is constant within the domain regardless of support. In the context of ordinary random statistics, it is therefore difficult to ascribe any representativeness to an individual environmental sample other than being a member of a representative set at that support.

In the case of environmental field investigations, however, the whole point of a sampling plan is often to obtain samples (observations) that are representative In a different sense, namely, representative of their local areas. This spatial kind of representativeness is also hard to define. Intuitively, what we are really after is a measurement at a location that is close to the local mean concentration, but as we have seen, the local mean is a function of the support. Thus, until we define the support which we want our observation to represent, we can't say anything about how good the observation is.

Given a set of n observations within a domain, a geometric support neighborhood can be defined around each observation as the set of all points closer to it than to any other observation (i.e., a Voroni polygon). Voroni polygons are an approximation of the neighborhoods with which the observations are most correlated, and give an idea of the spatial resolution of the set of observations.

APPLICATIONS

In most of the above discussion, we have dealt with the uncertainties and ambiguities involved in taking and measuring samples of the physical environment, and the inherent limitations of relating measured values back to some 'real' characteristic of the environment. In spite of these problems, most of which have not been adequately dealt with theoretically, the world marches on. Site Investigators are still faced with the necessity of collecting samples in the field, and using them to make interpretations and decisions. The remainder of this paper will focus on practical approaches to sampling and data analysis which will help reduce ambiguity, and improve data quality and usefulness. We will continue to use the simulation as an illustrative example.

SAMPLING DESIGNS

One of the common problems facing a site investigator is the layout of a sampling pattern. Assume for the moment that you have a predefined domain that must be sampled and that the number of samples you can take has been fixed by budget constraints. Where should you take the samples? Most available guidance documents such as SW-846 recommend random sampling as the general solution. However, it has been shown (Olea, 1984, Yfantis, et al., 1987) that In the presence of spatial autocorrelation, sampling on a systematic grid will produce a more efficient sampling. If spatial autocorrelation is not present, the regular grid will be no better or worse than a random sample. Because the regular grid is a periodic sampling in space, it is obviously contraindicated when the presence of a spatial periodicity in the phenomenon is suspected at a scale near that of the proposed grid. Fortunately, this situation is not common, and the regular grid, because of its simplicity and effectiveness, can be used in most spatial sampling programs.

The efficiency of the regular grid is a result of the minimization of spatially clustered data which are duplicating each other's information. As a practical matter, small departures from regularity do not have major effects, so that field crews can make offsets from the grid to avoid obstacles without affecting the results of a study. This is particularly true when the autocorrelation function exhibits a large random component (nugget effect) relative to the autocorrelated component. The larger the relative nugget effect, the greater the tolerance area around the regular grid nodes.

SAMPLE SUPPORT AND QA/QC

To illustrate the potential impact of sample support on the quality of an investigation, we will sample our simulation at two different supports, and use the results for interpolation of concentrations. The sample locations are on the regular grid shown in Fig. 8. The first sampling is done at the single pixel support. Each sample is assigned the true value of the pixel at that location; we are assuming no sampling or analytical error. The second sampling is done on the same grid at a 2x2 pixel support. In this case the true mean value of the 4 pixels in the sample is used.

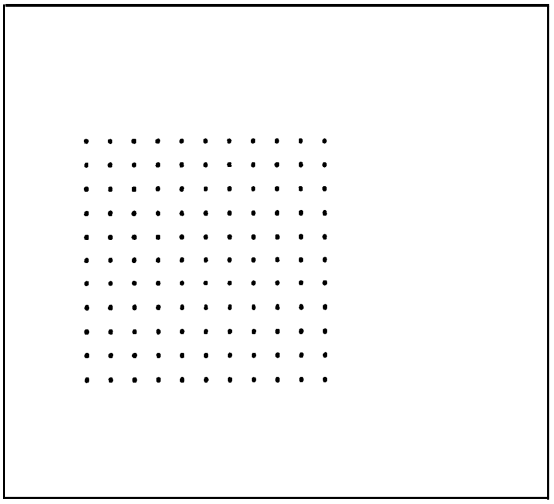


Fig. 8. Location of samples taken from the simulated distribution. True values of single pixels and 2x2 pixel blocks were taken at each point.

Variograms computed from the two sample sets are shown in Fig. 9. Note the dramatic decreases in variance at all distances with the larger support. When considering quality control, we take particular note of the projected y-axis Intercept or nugget, which has dropped from a value of 0.5 to less than 0.1. The nugget value provides an estimate of the total variance associated with taking adjacent, or co-located samples, and includes small-scale spatial variability as well as sampling, preparation, and analytical errors. In the present example, we introduced no measurement errors, so all of the nugget value is the result of spatial variability, and the change in support has a major effect. If, at the other extreme, the nugget were entirely due to sampling and analytical errors, it would not be reduced at all by changing support.

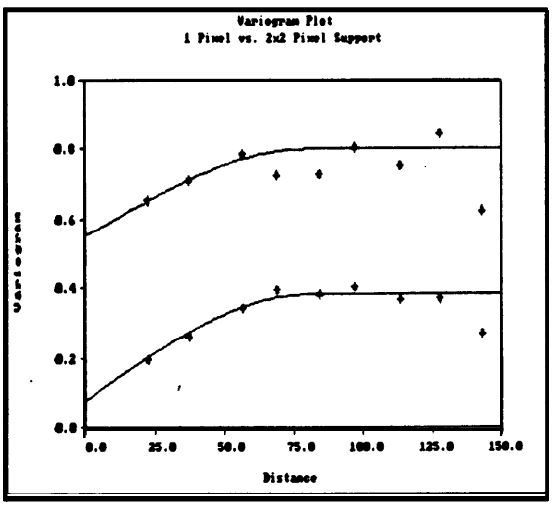


Fig. 9. Variograms computed from simulated samples on single pixel (upper curve) and 2x2 pixel (lower curve) support. Points are experimental values; solid lines are subjectively fitted models.

The approach one takes to improving data quality, therefore, should ideally be dictated by a variance components analysis of the variogram nugget value. The largest variance component should get the most attention. If spatial variance dominates, increase the support by taking larger samples or perhaps by compositing small clusters of samples. If measurement error dominates, look for a more accurate method. When either spatial or measurement variance is very dominant and cannot practically be reduced further, it may be possible to achieve a significant cost saving without seriously affecting overall data quality by going to cheaper methods which moderately increase the lower variance component. For example, if the maximum feasible support still results in a spatial variance component ten times greater than the measurement variance, the measurement variance could be allowed to increase by a factor of five with a resultant increase in the total standard error of only 17%.

The question of quality assurance goes beyond quality control of sample data, to include the adequacy of the program as a whole to achieve the desired objectives. In the case of a site investigation, we are particularly interested in the quality of the spatial estimate made from the sample data, and we must therefore look at sample quantity and location, and the estimation or interpolation procedures, as well as sample data quality. All of these factors can contribute to the quality of the end result, and as always, the most cost effective use of resources is to work at reducing the largest error component.

With our simulated example, we can compare the effects of the sample quality on the overall quality of the data interpretation by interpolating concentrations over the screen area from each of the two sample sets. Using the variogram models from Fig. 9, kriged estimates were computed for 10x10 pixel blocks. The results are shown in Fig. 10, with the blocks shaded according to the same classification used in Fig. 2. Comparison with the true values shown in Fig. 3 illustrates the overall superiority of the 2x2 sample support in defining the general pattern of block concentrations.

Plots of the true vs. estimated concentrations for the two interpolations are shown in Fig. 11, and histograms of the estimation errors in Fig. 12. Note that even though the 2x2 pixel samples provide better estimates, there is still relatively high error which may lead to a large number of false positives and false negatives at most concentration action levels. Further reduction of estimation error would require more samples, larger support, or both.

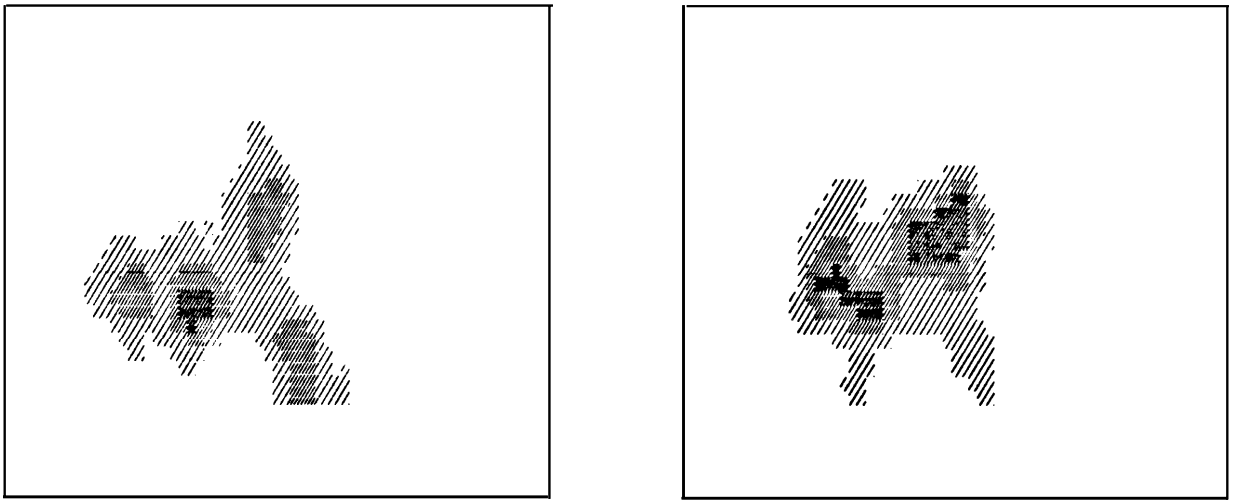


Fig. 10. Maps of kriged estimates for 10x10 pixel blocks using single pixel (left) and 2x2 (right) samples. Compare with map of true block values in Fig. 4.

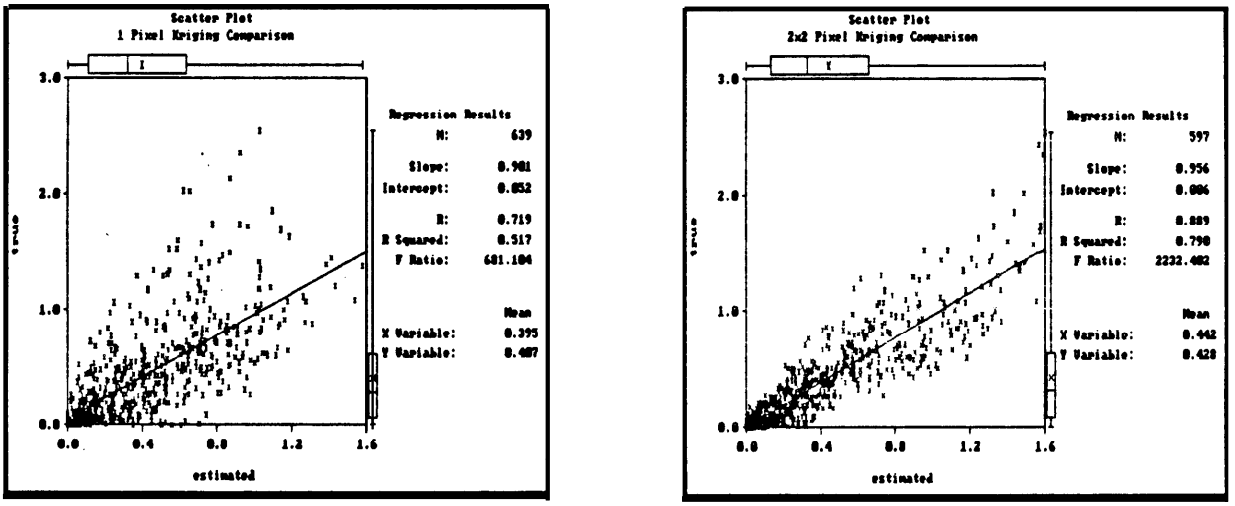


Fig. 11. Plots of estimated vs. true block values and associated regression statistics for the two kriged estimates in Fig. 10. The left plot is from the single pixel case, the right, from the 2x2 pixel case.

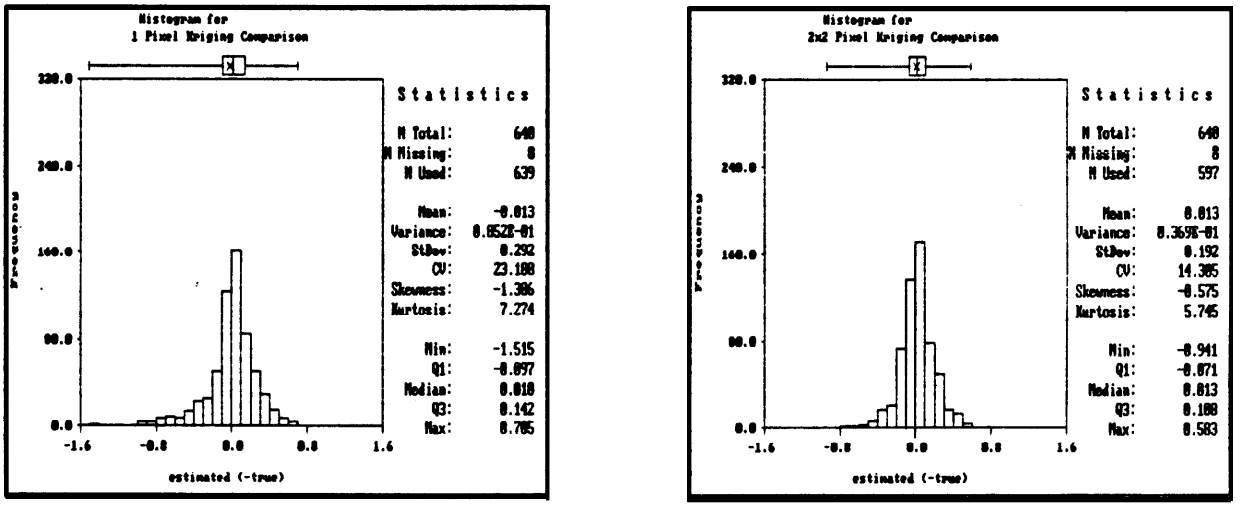


Fig. 12. Histograms of estimation errors (estimated - true) for the single pixel case (left), and the 2x2 pixel case (right).

ISOPLETH MAPS

Isopleth, or contour, maps are a commonly used and effective method for displaying the spatial distribution of chemical concentration. The use of isopleths, however re-introduces some of the ambiguity discussed above regarding support. If we examine the kriged maps in Fig. 10, we can easily visualize the process of generating a contour line by smoothing the stepwise boundary between two shading patterns. This is essentially the same result we would get if we assigned the estimated block average values to the block midpoints and ran the resulting grid through a typical contouring algorithm. It is also comparable to the limit we would approach (very expensively) with kriging as the grid dimension approaches zero. The problem arises because a contour represents the intersections of a plane parallel to the x-y axis and a continuous surface. The existence of one contour line implies the existence of all possible contour lines, and demonstrates that we have in effect estimated individual concentration values at every point. But we showed earlier that point concentrations are binary and discontinuous, and that meaningful statements about concentration require specifying a support. Which support?

Intuitively, we might like to think of the isopleth surface as representing the result of a moving average based on a specified support window. While this rounds good, it doesn't work in practice. When we specified a 10x10 block grid, it was easy to compute the true block means and compare them to the estimated block means. However, when we try to compute a true moving average isopleth based on a 10x10 support window, we find peaks, holes, and irregularities in the isopleth lines occurring at much smaller scales than the 10x10 support. Such a situation is self-contradictory, and it is not obvious how to define a true isopleth, compare it with an estimated one, and determine the goodness of the estimate.

In spite of the difficulty of defining a 'real' isopleth, a case can be made for using isopleths drawn from kriged block estimates as remediation boundaries. Given that kriging is in some sense an optimal estimator, and the goal is to remediate all of the area above an action level, the best approach should be to krige every point in the sampled domain, and remediate all of the points estimated to be above the action level. Contouring kriged block values provide6 a good approximation which is computationally feasable. The point to remember when doing this is that the nice, smooth contour6 may be primarily a regression effect, and that the underlying reality may in fact be very erratic. If we were to clean up an area inside an isopleth boundary, and then take check samples immediately around the boundary , we would expect to see a large number (approximately 30 to 50%, depending on the distribution) of samples exceed the action level, even if the boundary is well estimated. A more practical approach would be to do the check sampling along the boundary before the clean-up, and use the additional data to better define the boundary.

CONCLUSIONS

Chemical compounds in the environment can be spatially autocorrelated at scales ranging from the molecular to global. Autocorrelation is practically significant when it occurs at scales relevant to the problem, that is, between the dimensions of a sample and the dimensions of the domain being investigated. Under such conditions the support, or physical size, shape, and orientation of a sample, becomes a critical factor in the quality of an investigation. The use of QA/QC data for a variance components analysis of the nugget component of a variogram model can be very useful in selecting the most cost-effective approach to additional sampling.

Chemical concentration in the inhomogeneous environment must refer to a specific support to be meaningful. Specifying the support associated with an action level for enforcement or remediation removes a source of ambiguity and potential misunderstanding.

Spatial estimates of concentration such as kriged blocks, and the contour maps derived from them, are smoothed representations of reality, and represent a support more like the sampling grid size than the dimensions of the sample. Site investigators should understand, and be able to explain to the public, why many of the values at the sample support, taken outside an isopleth boundary, are expected to exceed the isopleth value.

NOTICE

The information in this document has been funded wholly by the U.S. Environmental Protection Agency. It has been subjected to Agency review and approved for publication.

REFERENCES

Journel, A. G., and Huijbregts, Ch. J., 1978. Mining Geostatistics. Academic Press, London. 600 pp.

Matheron, G., 1963. Principles of Geostatistics. Economic Geology, Vol. 58, pp 1246-1266.

Olea, R. A., 1984. Systematic Sampling of Spatial Functions. Kansas Geological Survey, Lawrence, Kansas. 57 pp.

Splitstone, D. E., 1987. Sampling Design: Some Very Practical Considerations. (This volume).

Starks, T. H., 1986. Determination of Support in Soil Sampling. Mathematical Geology, Vol. 18, No. 6, pp. 529-537.

Yfantis, E. A., Flatman, G. T., and Behar, J. V., 1987. Efficiency of Kriging Estimation for Square, Triangular, and Hexagonal Grids. Mathematical Geology, Vol. 19, No. 3, pp. 183-205.

Discussion

John Warren
US Environmental Protection Agency

Spatial prediction, as illustrated by these two papers, holds great promise in extracting large amounts of information from situations where extensive data collection is not feasible.

Evan Englund's paper on using kriging in conjunction with the concept of sample support is most interesting and his discussion illustrates clearly the problem of defining "sample" with respect to the physical dimensions of a sample site. The use of "sample support" to avoid confusion with the common statistical understanding of "sample" is to be commended and his definition: "The support on which decisions will be made is rarely the same as that of the samples. The choice of decision support can significantly affect the outcome of an analysis, and should be considered before a sampling program is undertaken." should be included in every Data Quality Objectives (DQO) Program.

The illustration of "sample support" by simulation is particularly useful in demonstrating how the definition of contamination is a function of the particular support chosen. This simulated example should be made part of the Agency's DQO training program as it clearly demonstrates to managers and decision-makers the importance of allocating resources proportional to the ratio between decision-area and sample support *area.*

Evan's discussion on the choice of most efficient or most practicable sampling design for sampling in a spatial domain is very useful for there does not seem to be a general solution applicable to all sites. Each site must have a unique sampling plan and although the Agency's guidance document, SW-846, endorses simple random sampling, the document does recommend a tailored-to-the-site approach. The revised edition Of SW-846 (4th edition with an expected publication date of mid-1990) will provide guidance on most sampling schemes likely to be encountered in spatial Statistics, and also gives guidance on the construction of optimal grid sizes for the effective use of kriging.

Noel Cressie's paper is a nice summarization of spatial statistics and introduces some of the scientific notations essential for discussion of kriging. His comparison of trend surface model with the random field model is very useful and clearly illustrates the differences between the two models.

The strength of the paper lies in the excellent discussion of spatial design of networks, Section 3. The problem of adding a monitoring site to an existing network cannot be solved easily and only by making key assumptions on the intent and use of the monitoring data can any solution be attempted. The concept of "average" prediction error (equation 3.4) has much merit as many of the Agency's monitoring networks are intended to measure mean increases/decreases in pollution levels. It seems clear that this notion of optimal network design can be applied to Agency networks where non-statistical considerations in determining actual sites are relatively unimportant.

There is much that needs to be done in order to effectively bring spatial statistics to the practical level at EPA, and these two papers are an important start. There is definitely an interest in spatial statistics as evidented by the number of people present at the spatial statistics sessions at the annual EPA Conference on Statistics: these two papers should further stimulate this interest.