CLR Report No 04 - Sampling Strategies for Contaminated Land

Department of the Environment, Transport and the Regions

Crown copyright 1994

Prepared by The Centre for Research into the Built Environment The Nottingham Trent University 1994

Applications for reproduction should be made to HMSO

Printed and published by the Department of the Environment 1994

DEPARTMENT OF THE ENVIRONMENT

CONTAMINATED LAND: Identification, assessment and control

SAMPLING STRATEGIES FOR CONTAMINATED LAND

THE CENTRE FOR RESEARCH INTO THE BUILT ENVIRONMENT, THE NOTTINGHAM TRENT UNIVERSITY

This report is one of a series of reports financed under the contaminated land research programme of the Department of the Environment. The current series deals with: information needed to assess risks; procedures for categorising and assessing risks; and evaluation and selection of remedial methods.

The purpose of the reports is to provide regulators, developers and other interested parties with authoritative and researched advice on how best to identify and assess the problems contamination can pose and what can be done to tackle them. They cannot, however, address the specific circumstances of each site. Every site is unique. Anyone using the information in a report must, therefore, make appropriate and specific assessments of any particular site or group of sites. Neither the Department nor the authors can accept liability for the use or interpretation of the contents of any report.

General guidance on assessing contaminated land and developing remedial solutions which is complementary to the series is provided by the Construction Industry Research and Information Association (CIRIA).

The Department acknowledges with thanks the help of Consultants In Environmental Sciences Ltd.

CLR 4 prepared by:

The Centre for Research into the Built Environment

The Nottingham Trent University

Burton Street

Nottingham NG1 4BU

1 Introduction

This report sets out guidance on formulating sampling strategies at sites where contamination is known or suspected to be present. These strategies include the number of sampling points, their location on site, and the depth of samples to be taken.

The guidance is intended to help design efficient and cost-effective strategies to:

* discover whether contaminants are present on a site;
* establish the distribution of contaminants on a site;
* locate concentrations of contaminants which could lead to unacceptable risks to human health or the environment (“hot spots”); and
* establish the size and shape of such concentrations of contaminants

Guidance is included on using information obtained from desk studies and visual site inspections in designing sampling strategies.

The report discusses sampling patterns, uniform and variable density sampling, and multi-stage sampling.

This report is intended to help site investigators with appropriate training to formulate sampling strategies for individual sites. Only those with such training, and with knowledge of the site to be investigated, can judge whether a particular strategy is appropriate to that site.

2 The Role of Sampling in Site Assessment

Regulatory authorities often rely on detailed site assessment reports to make judgements on issues such as health and safety, waste disposal, planning and groundwater protection. Unless sampling is sound a site assessment report will be unreliable. The penalties for poor sampling can be severe, including

unacceptable risks to site workers and subsequent site users, and expensive unplanned remedial work. Almost all contaminated site assessments will involve the phases outlined in the following paragraphs.

Desk Study

A preliminary investigation should always include desk study of site history, and neighbouring land uses (Ref. 1); a review of the geology, hydrogeology and surface hydrology of the site and its surroundings; and a visual inspection of the site (Ref.2).

Visual Inspection

The visual inspection is used to identify and confirm the positions of former structures (buildings, pits, lagoons, underground storage tanks etc.) and to locate site hazards such as power cables, unsafe structures and contaminants visible at the surface. Visible evidence of the condition of soil and vegetation should be carefully recorded at this stage. Some limited soil sampling at or near the surface (e.g. with a hand auger) can be undertaken.

Initial Working Hypothesis

Information gathered during the desk study and visual inspection will be used to develop an initial working hypothesis covering likely contaminants, their sources, their present spatial distribution, and their potential for migration. At this stage, and in the light of the proposed assessment or remedial work, it is necessary to consider the consequences of failing to detect a contaminant hot-spot if it exists. A decision needs to be made on the largest hot-spot that could be accepted or dealt with economically if it were missed in sampling. This *critical hot-spot size* is an important design parameter.

Starting Sampling Strategy

The investigator is now in a position to formulate a starting sampling strategy, including the number of sampling points, their location on the site, and the number of samples and depth profile at each sampling point.

Implementation of Sampling Programme and Phasing

Each trial pit or bore hole provides new information on the soil profile. Flexibility is important. The initial working hypothesis may change as new information is gained, and the sampling programme may need to change in response. Preliminary decisions on the chemical analyses required may also need to be modified.

If a multi-stage programme has not been designed from the outset, data interpretation may suggest the need for a second stage of sampling and analysis.

Reporting

The content of a report will depend on the brief but should contain at least the following information in relation to sampling:

* The overall objectives of the assessment;
* The investigation strategy and the reasons for choosing it;
* The methods, techniques and standards used in carrying out the site investigation;

3 Choosing a Sampling Pattern

Three main sampling patterns are discussed in the contaminated land literature. These are the square grid, stratified random and simple random patterns, as shown in Figure 1. The square grid is by far the most popular and has obvious practical advantages.

The mathematical theories on optimisation of spatial sampling are discussed in Annex 1

. Analysis of these theories suggests that efficient sampling design should satisfy four conditions:

(i) it should be stratified (that is, the area to be sampled should be partitioned into regular sub-areas);
(ii) each stratum (sub-area) should carry only one sampling point;
(iii)it should be systematic;
(iv)sampling points should not be aligned.

The simple random pattern satisfies only condition (iv). The square grid satisfies all but condition (iv) and stratified random patterns satisfy all but condition (iii). The disadvantage of the square grid is its much reduced ability to detect elongate hot-spots whose long axis is parallel to the grid axis. The weakness of the stratified random sampling pattern is its tendency for uneven sampling.

The herringbone sampling pattern (Ref. 3) overcomes these disadvantages and yet retains the practical advantage of being easy to set out on site (Figure 2). This pattern satisfies all four conditions in paragraph 3.2.

Setting out the Sampling Pattern

The repeat unit of the herringbone pattern can be generated from a unit square grid by offsetting points one-quarter of a unit as shown in Figure 3. A family of herringbone designs can be generated by offsetting points by other fractions. However, computer experiments show that the one-quarter unit offset provides the most efficient search pattern. This type of herringbone pattern can be generated by four square grids each of the same orientation and spacing. Thus it is almost as easy to set out on site as a single square grid.

Performance of the Sampling Pattern

Computer experiments with a randomly located circular hot-spot (Figure 4) show that the square grid and herringbone patterns perform well, the stratified random pattern is relatively inefficient and the simple random pattern is very inefficient. Square grid performance with randomly located elongate hot-spots falls off very badly when the hot-spot is aligned, or hearty so, to the grid directions (Figure 5). The herringbone pattern performs best overall and is only slightly affected by hot-spot orientation.

Figure 6 confirms the superiority of the herringbone pattern for detecting elongate hot-spots in unfavourable orientations.

Simple geometry suggests that an equilateral triangular grid should be more efficient than a square grid. This is because, for sampling neighbourhoods of unit area, no point is more than 0.62 units from a sampling point in a triangular array, compared with 0.71 units for a square array. It is possible to generate an offset triangular grid analogous to the herringbone transformation from a square grid. However, for most purposes, the efficiency gains from triangular systems probably do not compensate for their inconvenience in generation or setting out.

4 Designing a Sampling Plan

How Much is Known?

Desk study and visual inspection may reveal no evidence for likely hot-spot locations, or evidence of such an inconclusive nature that the investigator deems it prudent to treat all parts of a site a shaving an equal chance of containing a hot-spot. In these circumstances, uniform density sampling is appropriate and is discussed below in paragraphs 4.5-4.8.

At the opposite extreme, preliminary investigation will sometimes reveal such strong evidence for contamination that it largely dictates the sampling plan, at least for part of a site. Sampling will be highly focused to reflect that evidence and the investigator's experience and is not discussed further here.

In most cases matters are less clear cut. The investigator may suspect that some parts of a site are more likely to contain hot-spots than others but the weight of evidence or degree of suspicion that this is the case falls between the two extremes described above. In circumstances where it can be assumed that all parts of a site do not have an equal chance of containing a hot-spot, efficiency of sampling designs can

be improved by use of variable density sampling based on weight-of-evidence scores. This is discussed in paragraphs 4.9-4.14.

The efficiency of sampling designs can be improved (i.e reduction in the number of samples without a significant reduction in the probability of locating a hot-spot) by use of multi-stage sampling techniques. These are discussed in paragraphs 4.15-4.22.

Uniform Density Sampling

As stated in Section 3, some sampling patterns are much less efficient than others. The number of sampling points needed to achieve a given probability of success in locating a hot-spot depends primarily on the hot-spot size but also on the shape of the hot-spot.

Calculation of the number of sampling points needed to ensure a given probability of hitting hot spots is shown in Box 1. In the examples shown the probability of hitting a hot-spot is set at 0.95 so that, if the hot-spot were to exist as specified, the chosen sampling strategy would succeed in hitting it 19 times out of 20.

Box 1 Calculation of the Number of Sampling Points Required

| The number of sampling points needed to ensure 0.95 probability of hitting hot-spots of various sizes with a herringbone sampling pattern where all parts of the site have an equal probability of concealing a hot-spot, is shown in Figure 7. The curves in Figure 7 can all be fitted to the equation

$$N = kA/a$$

where N is the number of sampling points, A is the total site area and a is the hot-spot area.* The constant k can be termed the shape constant and takes the following values for various shapes of hot-spot:

Circular ...	k = 1.08
Plume shaped	k = 1.25
Elliptical (aspect ratio 4: 1, oriented at 0° or 90° to the grid directions)	k = 1.80

Assessors should choose a k value to reflect their belief in the likely hot-spot shape, or the level of protection required against the possibility of elongate hot-spots. A conservative choice, say $k = 1.5$, would be prudent when no reasonable assumptions can be made about hot-spot shape.

For example, if a circular hot-spot is assumed occupying 5% of the total site area, then $k = 1.08$, $A/a = 20$ and $N = 22$ for a herringbone pattern.

Choice of N sampling points determines a nominal square-grid spacing of $d = \sqrt{(A/N)}$.

Site assessors will use their own judgement to adjust d or N if necessary to suit particular site shapes. |
|---|

Number of sampling points can be very large on moderately large sites. For example, if it is decided that the appropriate hot-spot size on a 2 ha site (equivalent to just over two football pitches) is one percent of the site area (i.e. about the same size as a singles tennis court) then, using k = 1.5 as shape factor, 150 sampling points are called for. If a depth profile is established by taking, say, 4 samples at every third sampling location with 1 sample only at the others, the total number of samples rises to 300. In practice, sampling numbers can be reduced by choosing a larger nominal hot-spot in first stage sampling, with subsequent stages of sampling building on the information derived from earlier stages.

Calculation of sample numbers becomes more complicated when it is assumed that more than one hot-spot exists on a site. For five hot-spots each occupying 1% of the total site area, the number of samples needed to detect one hot-spot with 95% probability increases to 45, compared to 22 required to detect one hot-spot occupying 5% of the total site area, while 137 sampling points would be required to give the same probability of hitting all five hot-spots.

Calculation of the appropriate number of samples is more complicated for multiple hot-spots and is discussed in Ref. 3. In practice, it is rarely possible to specify the number of hot-spots expected on a site. However, the problem can usually be overcome by multi-stage sampling.

Variable Density Sampling

Experienced investigators will often suspect that some parts of a site are more likely to contain a hot-spot than others. They may then wish to design a sampling strategy that reflects their degree of suspicion about the hot-spot's likely location. One approach to this type of sampling design is to partition the site into subareas and then to score each subarea (say, on a 0 to 10 scale) to reflect degree of suspicion or strength of belief as to where the hot-spot is likely to be.

Figure 8 shows an example of this type of scoring, and the corresponding sampling plan. The degree of suspicion is expressed relative to that in the most favoured subarea which is taken to be 1. Sampling density for the most favoured subarea is calculated such that the probability of hitting a hot-spot if it exists in that subarea is 0.95. Other subareas carry lower sampling point densities to reflect the investigator's lower expectation (i.e. relative degree of suspicion less than 1), that the hot-spot is located in these areas.

The relevant calculations of sample size and hit probabilities for variable density sampling are shown in Box 2. As the calculations show, if the investigator's judgement is correct the 0.95 probability of success in detecting a hot-spot is achieved with fewer sampling points and lower cost. But if the investigator is wrong the penalty is a reduced probability of success in locating the hot-spot (see probability values adjacent to Figure 8B).

Probability that no Hot-Spot Exists

So far in this report, the measure of a successful sampling scheme has been the probability of hitting a specified hot-spot *if it exists.* But a sampling scheme may fail to locate a hot-spot because the hot-spot does not exist. After a failure to find a hot-spot, it is sensible to establish the probability that this was because a hot-spot of at least the specified size does not exist. For example, a sampling criterion based on, say, an 0.95 after-sampling probability that a specified hot-spot does not exist is in some ways more appropriate than one based on an 0.95 probability of hitting a hot-spot if it exists. The relationship between these two probabilities, and the investigator's prior degree of suspicion that a hot-spot exists, is shown in Figure 9. This allows the investigator to adjust sampling density according to his or her before-sampling suspicion about whether or not a hot-spot of the type specified exists. Note that for a before-sampling probability of 0.5 the sample size that would give 0.95 probability of hitting the hot-spot *if it exists* would also give an after-sampling probability of 0.95 that the hot-spot does not exist if no hot-spot had been hit.

---

*The formula used to calculate *N* assumes that there is only one hot-spot within the area covered by the sampling grid. For more than one hot-spot see paragraph 4.8.

An example is given in Box 2 and Figure 10 showing the relationship between before-sampling probability (PrB), hit probability *if the hot-spot exists* (PrH) and number of sample locations (N) in a herringbone search pattern. In this example, the hot-spot is an ellipse with 4:1 aspect ratio orientated parallel to the grid directions.

Box 2 Calculation of Sample Size and Hit Probabilities for Variable Density Sampling

The Sample sizes and hit probabilities discussed in paragraphs 4.9-4.11 and shown in Figure 8 are calculated as follows:

(1) Compute N as in Box 1.

(2) If S_i is the degree of suspicion score given to subarea i, and S_{max} is the maximum score given to any subarea, the number of sample locations n_i to be assigned to subarea i is given by

$$n_i = S_i \; A_i \; N \;/\; S_{max} \; A_t$$

where A_i is the area of subarea i and A_t is the total site area.

(3) The probability PrH of hitting the hot-spot (if it exists) now depends on the subarea as follows:

$S_i = S_{max}$: PrH = 0.95

$S_i < S_{max}$: PrH = $c_1 k_i + c_2 k_i^2 + c_3 k_i^3$

in which $k_i = a n_i / A_i$ where a is the hot-spot area and the coefficients depend on the hot-spot type as follows:

	C_1	C_2	C_3
Circle	1.052	0.130	- 0.262
4:1 ellipse	1.362	- 0.749	0.159

The sample sizes and hit probabilities discussed in paragraphs 4.12-14 and shown in Figure 10 are calculated as follows:

(1) For each subarea i the investigator assigns a before sampling probability $(PrB)_i$ that a hot-spot exists in that subarea, and also sets the required after sampling probability, PrA (say, PrA = 0.05 so that, if a hot-spot is not located, it could be said with 95% confidence that such a hot-spot does not exist).

(2) If the hot-spot exists in subarea i the probability of hitting it is given by

$$(PrH)i = 1 - \frac{PrA\,[1-(PrB)i]}{(PrB)i\,[1-PrA]}$$

(3) The sample size necessary to achieve this $(PrH)_i$ can now be read off graphs (Figures 20 and 21) for the appropriate relative hot-spot size.

The before-sampling specification can be used to vary the sampling density from one part of a site to another.

Figure 10A shows a site that has been partitioned into three subareas each carrying a different before-sampling probability. Each number reflects the investigator's degree of suspicion about the likelihood of a hot-spot (of at least the size specified) existing in that subarea.

Figure 10B shows the corresponding sampling plan with sample point densities calculated to reflect the different before-sampling probabilities, but giving the same after-sampling probability of PrA = 0.05 for all subareas; ie if no hot-spot is detected one could state with probability 1-PrA = 0.95 that a hot-spot of (at least) the specified size does not exist. However, if the hot-spot does exist the probability of hitting it depends on subarea as shown adjacent to Figure 10B

.

Multi-Stage Sampling

Sampling in two or more stages almost always results in much better spatial definition of contaminants for a given total number of samples. The basic idea of multistage sampling can be illustrated by reference to

Figure 6

, which shows that about 30 sampling points are required to give 0.95 probability of hitting the elliptical hot-spot with a herringbone search. But a first-stage sampling of only 15 points would still give 0.65 probability of hitting the hot-spot; a second-stage sampling plan could then fully exploit the additional information (which might include a hot-spot hit).

Careful analysis of first-stage results is particularly important when multiple small hot-spots are anticipated (see paragraph 4.8).

Two-stage sampling allows use of a formal geostatistical procedure for demonstrating how the sampling density and configuration affect the precision with which contaminant concentrations at various points can be estimated. It is first necessary to estimate the degree of spatial dependence between measured concentrations in different samples which is done using a *semivariogram* (see Box 3).

Box 3 Construction of a Semivariogram

(1) The semivariance is equal to the average of the squared differences between sample values spaced a distance *h* apart. When *h* = 0 the value at each point is being compared with itself and hence the semivariance is zero. As *h* increases, the values being compared tend to be less and less closely related; eventually they become uncorrelated and their squared differences become equal to the variance around the average concentration of contaminants for the whole site (i.e. the conventional variance). Hence the ideal semivariogram looks like

Figure 11

. The distance at which the semivariance approaches the conventional variance is referred to as the *range* of the variable; it defines the neighbourhood within which all values of the variable are related to one another.

(2) Some semivariograms for clay content in soil are shown in

Figure 12

. Note that, although the semivariance at *h* = 0 is by definition zero, measured semivariances often tend towards some positive value as *h* approaches zero. This value is known as the *nugget variance,* and the phenomenon (which is very common) is called the *nugget effect.* (The terms derive from gold mining, gold nuggets being both rare and much smaller than the spacing between sampling cores.) The nugget effect therefore refers to spatially dependent variation that occurs over distances much smaller than the shortest sampling interval. With wide sampling intervals all the variance might appear as nugget variance (e.g.

Figure 13

). However, a much smaller sampling interval usually resolves the true shape of the semivariogram (

Figure 14

), which in this example has a range of 40 - 50m.

Having constructed a semivariogram, the form of a regionalised variable is estimated by the method known as *kriging.* At its simplest, kriging involves no more than estimating the weighted average of the contaminant concentration within an area or volume, the weighting being chosen to ensure that the estimates are unbiased and have minimum variance.*

For the present purpose the relevant point is that the weighting depends only on the configuration of the sampling points and on the semivariogram, not on the observed values. So if the semivariogram is known from first-stage sampling, it is possible to estimate how the precision in a second stage of sampling will vary with sampling density. First-stage information therefore allows the design of a second-stage sampling scheme to satisfy a given data quality objective.

The procedure can be illustrated by the following example. The simplest configuration is that of a square area being sampled by square grids representing increasing numbers of sampling points of 4, 9, 16, 25 etc. An ideal linear semivariogram is assumed and then, for each grid, the kriging variance is calculated for different grid spacing, and the results are plotted as in Figure 15 (note that 'standard error' is simply the square root of the kriging variance). Plotting the minimum standard error against number of samples then gives a graph (Figure 16) which enables specification of the number of samples needed to achieve a given precision. The computation is more difficult for real semivariograms but the principle is unchanged.

To show how the method can be applied to a contaminated site, some volatile solids data from a small part of a landfill site (Ref. 4) are used in a worked example (see Box 4). The sampling configuration is shown in Figure 17.

Box 4 Example of a Semivariogram

```
+-------------------------------------------------------------------------------------------------------------
-----------------------------+
 | Note:  Semivariograms are not statistically robust with respect to outlying values, which should
therefore be                      |
 |        excluded (from the semivariogram calculation only). For this purpose values that are greater
than 2.5                        |
 |     standard errors from the sample mean are excluded.
|
 | (1)    First, semivariograms are constructed in four orientations (0°, 45°, 90°, 135°) to test for
directional                       |
 |    differences in spatial autocorrelation. As no such effects are apparent, an omnidirectional
|
 |    semivariogram is constructed which thus includes the maximum possible number of points at each
|
 |        separation distance. The result (

Figure 18

) is typical of small data set semivariograms, in that it shows                    |
 |        a progressive rise in semivariance up to a certain separation distance; thereafter the plot
becomes                           |
 |        erratic as the number of samples for each distance becomes too small for a stable variance
calculation.                       |
 | (2)    The experimental semivariogram is used to guide the choice of a semivariogram model for the
site. The                         |
 |        model is arrived at by trial and error using a back estimation technique (i.e. for each trial
model, data                       |
 |        points are removed one at a time and the model used to estimate these “unknown” values; the
model                             |
 |       giving the lowest back-estimation error is chosen as that best representing the spatial structure
of the                        |
 |        data). The model chosen in this way is shown as a smooth curve on

Figure 18

. It indicates that                                  |
 |       percent volatile solids values are spatially autocorrelated over a relatively short range (about
15 m), and                    |
 |        that half the total variance is due to very short range variability that cannot be resolved
even with a 4 m                   |
 |      sampling grid.
|
 |(3)  The semivariogram model and data point configuration can now be used to show how the estimation
|
 |       variance (or its square root, the standard error) varies as a function of the number of samples
(Figure                           |
 |        19). In this example, about 12 samples are sufficient to estimate percent volatile solids at
unknown                           |
 |       locations with a standard error of 8.5. But, thereafter, there is little to be gained by taking
additional                    |
 |     samples; even 60 samples only reduces the standard error to 7.5.
```

|
+----------
----------+

Site investigators should be aware that use of semivariograms and kriging can lead to substantial savings in analytical costs, especially on large sites. This is particularly true when there are several key contaminants that are themselves correlated.

For example, on many gas-works sites the key contaminants for remedial design are phenols and polyaromatic hydrocarbons (PAHs), the soil concentrations of which are often strongly correlated. However, phenols are relatively inexpensive to analyse for, whereas PAHs are much more expensive. A two-stage sampling plan can be devised in which phenols are analysed for at all first-stage locations while PAHs are only analysed at a subset of those locations. The phenols semivariogram can then be used in conjunction with the phenols - PAHs correlation to derive a PAHs semivariogram which can be used in turn to guide second-stage sampling density. This technique is known as co-kriging (Ref. 5).

5 The Depth of Sampling

There are no hard and fast general rules governing the final depth to which trial holes should be sunk, or the depth intervals at which samples should be taken. The nature of the site, its use and environmental setting will all influence the depth sampling plan.

In selecting depths for sampling, investigators should pay particular attention to made-ground, highly coloured material or ground which appears poorly consolidated. Some contaminated strata are very thin. For example, spent oxide from town gas manufacture was often spread in layers as thin as 5cm on the ground to reoxidise. Such material is often easy to identify by its characteristic blue colour.

Investigators should consider the significance of a potentially undetected layer of contaminated material in the ground and the chances of obtaining a sample of it in selecting the number of samples to be taken at different depths.

For many uses, site users would be most at risk from ingestion of inhalation of surface and near-surface soil.For this reason sampling strategies are often designed to provide most information on near-surface contaminant levels, especially when:

* groundwater is not at risk because the water table is relatively deep and/or the intervening geology is relatively impermeable;
* the use does not involve excavations that might place site workers at risk;
* the likely contaminants would not damage buried building materials or services.

In other circumstances, careful depth-profiling is needed, particularly on sites where contaminants may have been introduced at depth rather than percolating down from the surface, e.g in old fill or from underground tanks or pipes. Such heavy contamination at depth may give no indication of its presence at or near surface level.

If soil remediation proves to be necessary on any part of a site, a second stage of focused depth sampling will almost certainly be needed to provide adequate information for remedial design.

6 Conclusions

Spatial sampling theory shows that an efficient sampling design should be systematic, stratified and unaligned.The herringbone sampling pattern satisfies these conditions and is easy to set out on site.

Computer experiments with randomly located hot-spots (small areas of a site containing high concentrations of contaminants) confirm that the herringbone sampling pattern has a higher probability of locating hot-spots than other patterns using the same number of sampling points, or can achieve the

---

* The theory is mathematically complex and is set out in Ref. 5. Definitions of kriging and co-kriging are given in Annex 2

.

same probability of hitting hot-spots with a smaller number of sampling points. This is especially true for elongate (i.e. plume-shaped or elliptical) hot-spots.

The number of sampling locations depends on specifying a critical hot-spot size and shape, and a desired probability of locating the hot-spot. With herringbone sampling the number of locations needed to ensure a 95% hit probability for a randomly located hot-spot is given by $N = k\,A/a$ where k is the hot-spot shape factor(which typically varies between 1 and 2), a is the hot-spot area and A is the total site area.

Where it is not assumed that hot-spots are randomly located, subjective judgements about the likely location of a hot-spot can be used to design sampling plans which require fewer samples to achieve the same probability of locating a hot-spot. For some sites this approach could lead to appreciable cost savings. It should, however, be used with caution because, if the subjective judgement is wrong, the probability of locating hot-spots will be substantially reduced.

Sampling in two or more stages almost always results in much better estimates of pollutant concentrations at various points for a given total number of samples. It also enables construction of semivariograms, which model the degree of spatial dependence between samples, from which contaminant concentrations can be estimated at any point by a technique known as kriging.

When two or more substances are analysed on the same site, and their values are strongly correlated (which, for example, is often the case for phenols and PAHs on former gas-works sites), adequate spatial characterisation can be achieved at substantially lower cost by using the technique known as co-kriging.

It is recognised that some of the techniques described in this report will be difficult for many site investigators to implement. Current research on sampling strategies, funded by the Department, is focused on developing a computer based decision support system. This will help users to design spatial sampling patterns that reflect fine-scale variations in prior information without losing the statistical confidence properties of more conventional sampling patterns.

7 References

Annex 1. Optimisation of Spatial Sampling

Spatial sampling theory

There is substantial literature on spatial sampling, usually based on a complex mathematical presentation of theory. Most of the theoretical research has been directed at the problem of finding optimal sampling patterns for estimating the mean value of a spatial variable. Nevertheless, this forms a useful guide for the related but distinct problem of efficient search strategies for hot-spots. Key results are summarised below.

---

1 RPS Consultants *Documentary Research on Industrial Sites* (Department of the Environment 1994)

2 Applied Environmental Research Centre Ltd *Guidance on Preliminary Site Inspection of Contaminated Land* (Department of the Environment 1994)

3 Ferguson, C.C. 'The statistical basis for spatial sampling of contaminated land'. *Ground Engineering* 25/5 (June 1992) 34 - 38.

4 Fawcett, T., Brewer, A., Cowan, L. and Westlake, K. *Assessment of Landfill Sample Variability.* Draft report to the Department of the Environment. (Harwell: AEA Environment and Energy, 1993).

5 Journel, A.G. and Huijbregts, C.J. *Mining Geostatistics* (London: Academic Press, 1978).

6 Cochran, W.G. Relative accuracy of systematic and random samples for a certain class of populations. *Ann. Math, Statist.* 17, 164-177 (1946).

7 Quenouille, M.H. Problems in plane sampling. *Ann. Math. Statist.* 20, 355-375 (1949).

8 Bellhouse, D.R. Some optimal designs for sampling in two dimensions. *Biometrika* 64, 605-611 (1977).

9 Schoenfelder, C. and Cambaris, S. Random designs for estimating integrals of stochastic processes. *Ann. Statist.* 10, 526-583 (1982).

10 Barrett, J.P. and Philbrook, J.S. Dot grid area estimates: precision by repeated trials. Journal of Forestry 68, 149-151 (1970).

11 Bellhouse, D.R. Area estimation by point-counting techniques. *Biometrics* 37, 303-312 (1981).

12 Martin, R.J. A subclass of lattice processes applied to a problem in planar sampling. *Biometrika* 66, 209-217(1979).

13 Olea, R.A. (ed.) *Geostatistical Glossary and Multilingual Dictionary* (New York: Oxford University Press, 1991).

Progress in sampling theory is all but impossible without making some assumptions about the structure of spatial correlation in the underlying population. Cochran (Ref. 6) introduced the concept of a super population (in one dimension) with a convex autocorrelation function, which requires that nearby samples are in general closer in value than more widely separated samples.

Quenouille (Ref. 7) extended this to two dimensions, and introduced three general classes of sampling scheme.

Class 1 Sampling units are aligned in both directions

Class 2 Sampling units are aligned in the row direction and unaligned in the column direction

Class 3 Sampling units are unaligned in both directions

Bellhouse (Ref. 8) showed that (a) there can be no optimal fixed sample size design even when the correlation function is convex; (b) any design in Class 1 is less efficient than its associated design in Class 2; (c) in each of Quenouille's three classes, the optimal design for the class is systematic.

Quenouille (Ref. 7) showed that the optimal design in Class 3 is more efficient than the optimal design in Class 1. Schoenfelder and Cambaris (Ref. 9) showed that for any simple random design there is always a better stratified design.

Barrett and Philbrook (Ref. 10) advocated a multiple random start systematic sampling design; but Bellhouse (Ref. 11) showed that a single random-start design is always more efficient than multiple random start designs.

Together these results point to the four conditions for an efficient sampling design itemised in paragraph 3.2.

Rectangular sampling grids

Martin (Ref. 12) studied the problem of planar sampling from a population characterised by the general correlation structure

$$x_{ij} = \lambda x_{i-1,j} + \nu x_{i,j-1} - \lambda\nu x_{i-1,j-1} + \epsilon_{ij}$$

which was the structure used by Quenouille (Ref. 7) and referred to by him as a degenerate Markov process of the first order. In this equation ν is the correlation function in the row direction i, and λ the correlation function in the (orthogonal) Column direction, j. If h and k are, respectively, the spacings between sampling points in the i and j directions, Martin shows that estimation is most efficient (gives minimum mean square error) when $\lambda^h = \nu^k$. Grid spacings should therefore be chosen such that

$$\frac{h}{k} = \frac{\log\nu}{\log\lambda}$$

provided that the ration $\log\nu/\log\lambda$ can be estimated approximately.

Annex 2. Definitions

The following definitions are taken, with minor abridgements, from reference 13.

Kriging:

A collection of generalized linear regression techniques for minimizing an estimation variance defined from a prior model for a covariance or a generalised covariance of order ν that may or may not have constraints to eliminate bias. In contrast to classical linear regression, kriging takes into account stochastic dependence among the data. Nonlinear kriging is a linear regression on some appropriate nonlinear transformation of the original data and includes lognormal kriging, multigaussian kriging, indicator kriging, probability kriging, and disjunctive kriging.

All kriging algorithms result in systems of linear equations involving structural function models and sometimes additional linear constraints. These systems of equations correspond to the normal equations of classical regression.

The estimation technique is named after D. G. Krige, in honour of his empirical studies of South African gold mines that contributed to the formulation of the technique.

Co-kriging:

Estimation of a regionalised variable using not only observations on the variable, but also data on one or more additional related regionalised variables defined over the same field. Co-kriging can be used to reduce the estimation variance when the variable of interest has been undersampled, or for the simultaneous estimation of several regionalised variables sampled over the same field.

Figure 1. Three sampling designs. A, regular square grid; B, stratified random; C, simple random.

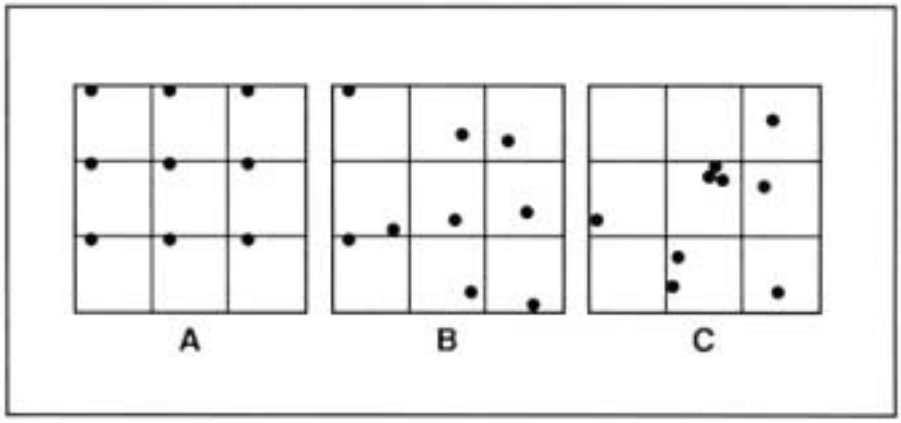


Figure 2. Herringbone sampling design.

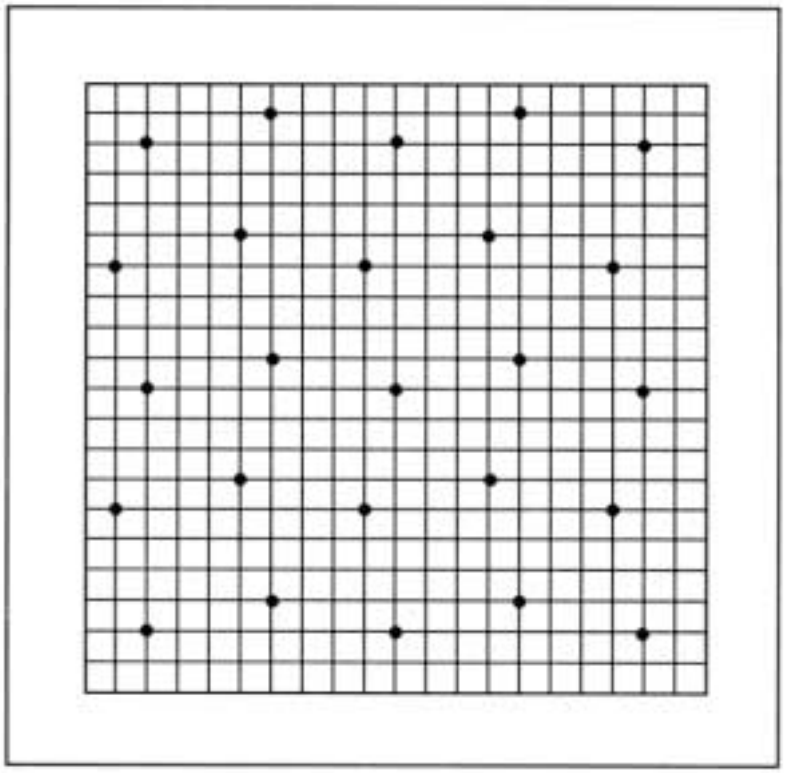


Figure 3. Derivation of herringbone pattern (open circles) by offsetting unit square grid points (dots) by one-quarter of a unit.

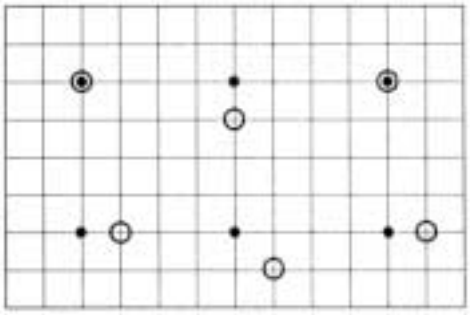
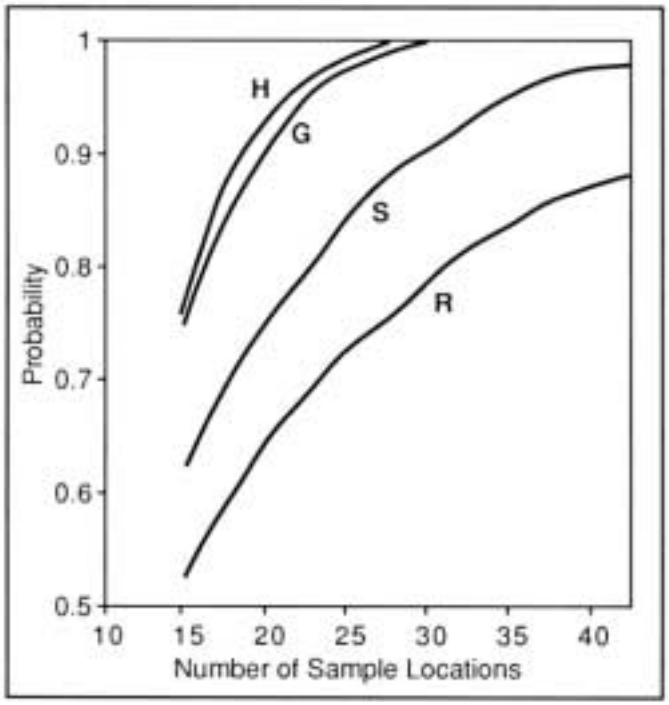
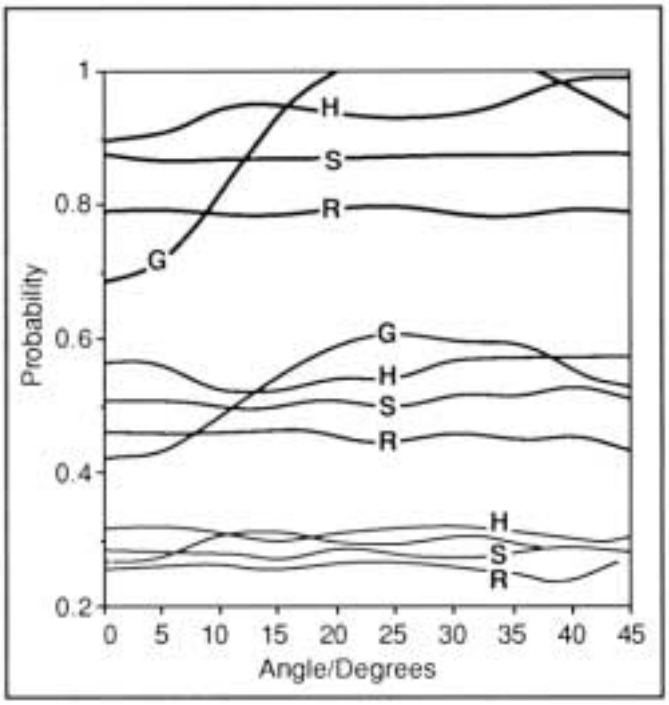


Figure 4. Performance of five sampling designs for detecting circular target occupying 5% of total site area.



Key: R = simple random; S = stratified random; U = stratified systematic unaligned; G = regular (square) grid; H = herringbone.

Figure 5. Performance of five sampling designs for detecting an ellipse (aspect ratio 4:1) as a function of orientation.



Relative size of target is 5% (top), 2% (middle), and 1% (bottom) of total site area. Number of sampling locations is 30. Key as Figure 4.

Figure 6. Performance of square grid and herringbone sampling designs for detecting targets of various shapes. Relative size of each target is 5% of total site area.

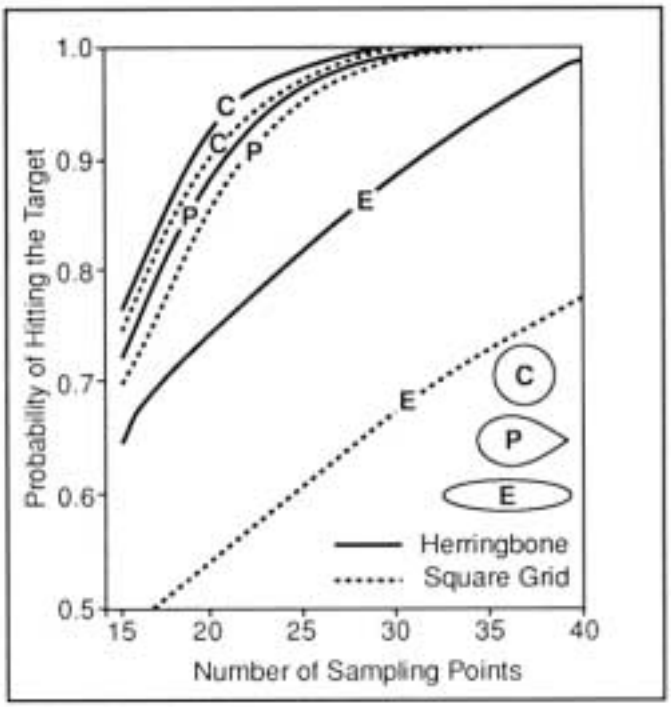


Figure 7. Number of sampling locations needed to ensure 0.95 probability of success in hitting targets of different relative sizes.

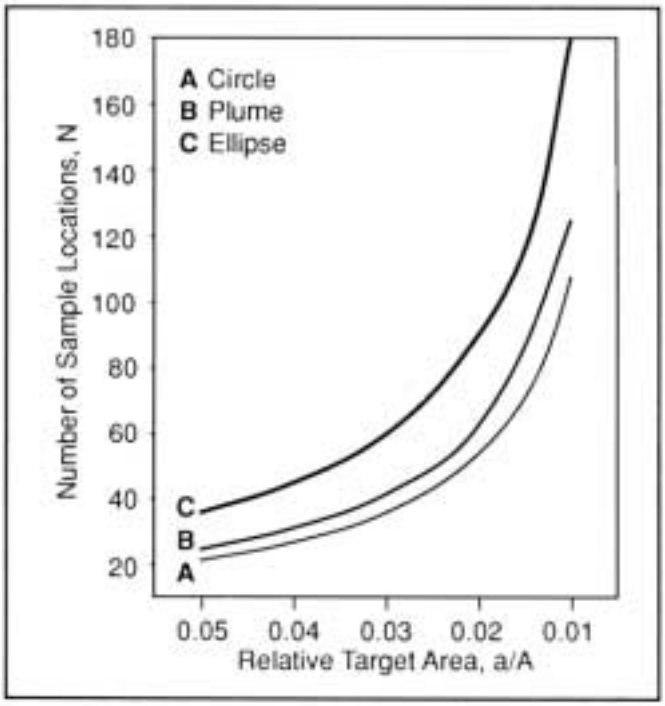


Figure 8. Variable density sampling. Different strength of belief scores (A) imply different sampling densities (B) and variable hit probabilities (see text).

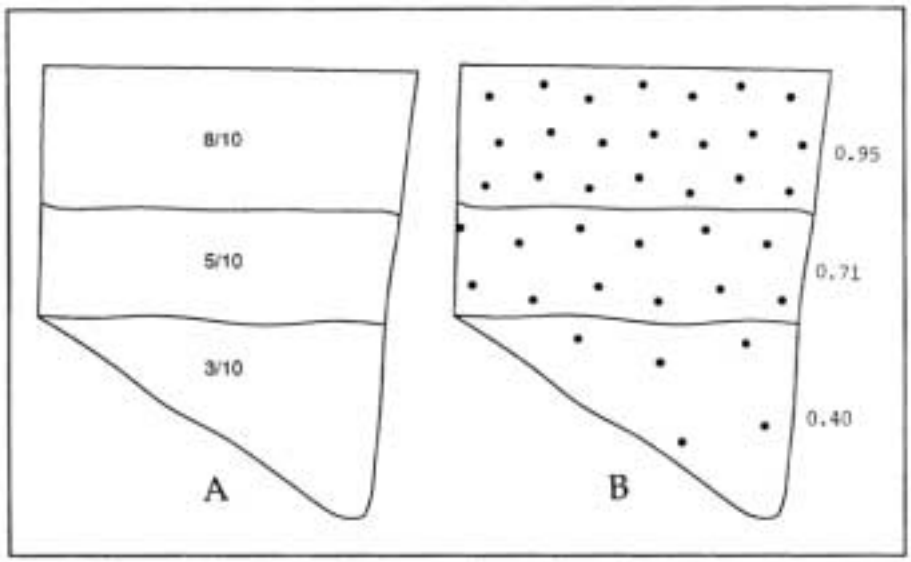


Figure 9. Relationships between a priori (before sampling) probability, PrB, a posteriori (after sampling) probability 1-PrA (labelled curves), and hit probability, PrH.

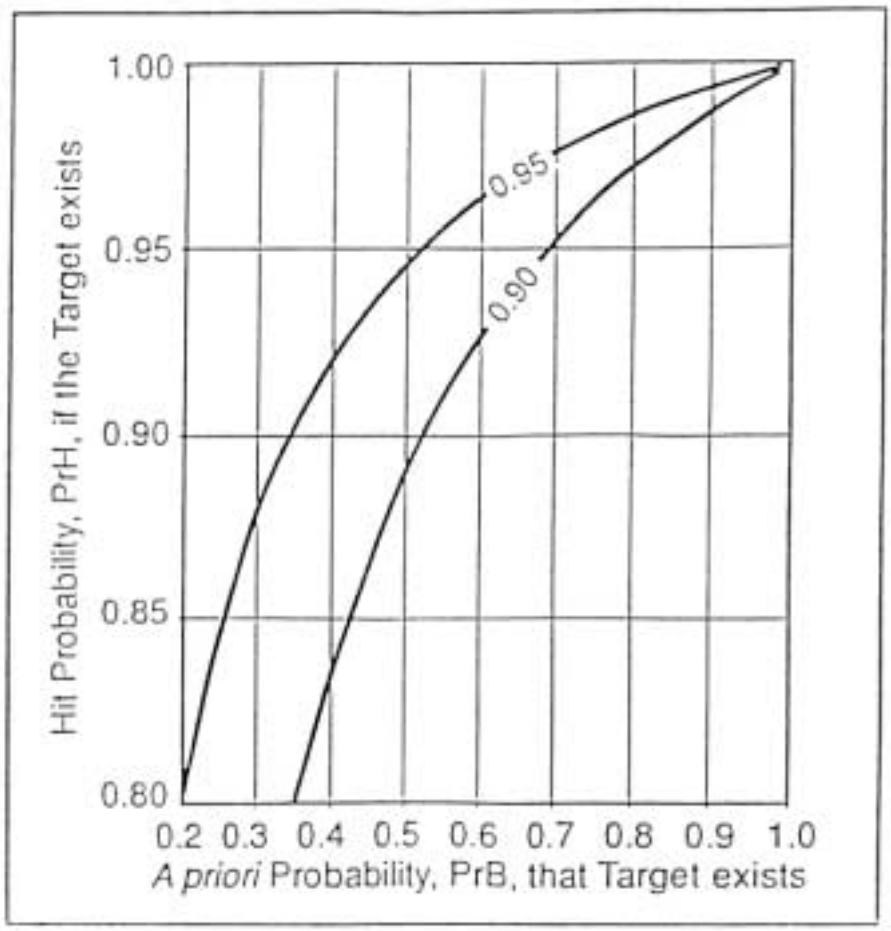


Figure 10. Variable density sampling. Different before sampling probabilities (A) imply different densities (B) and variable hit probabilities (see text).

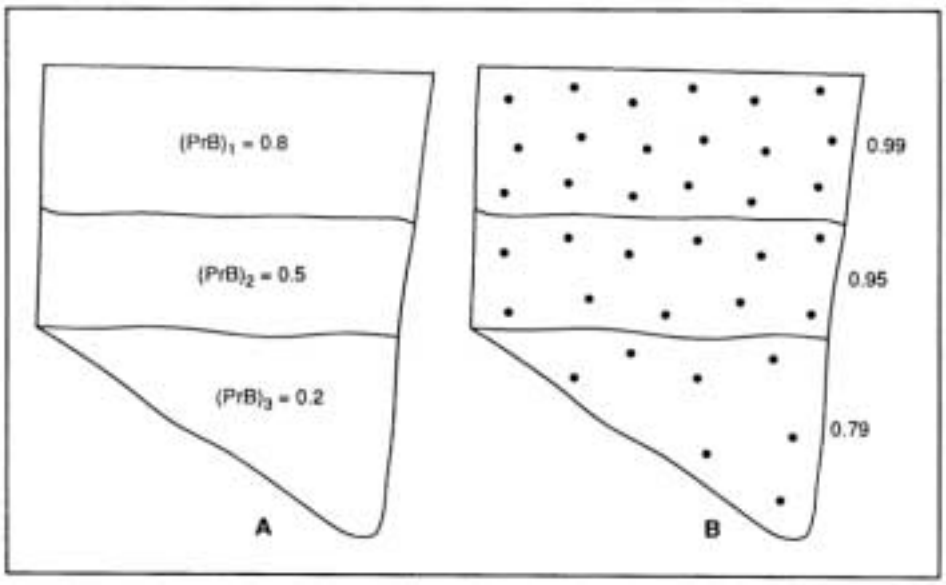


Figure 11. A typical semivariogram.

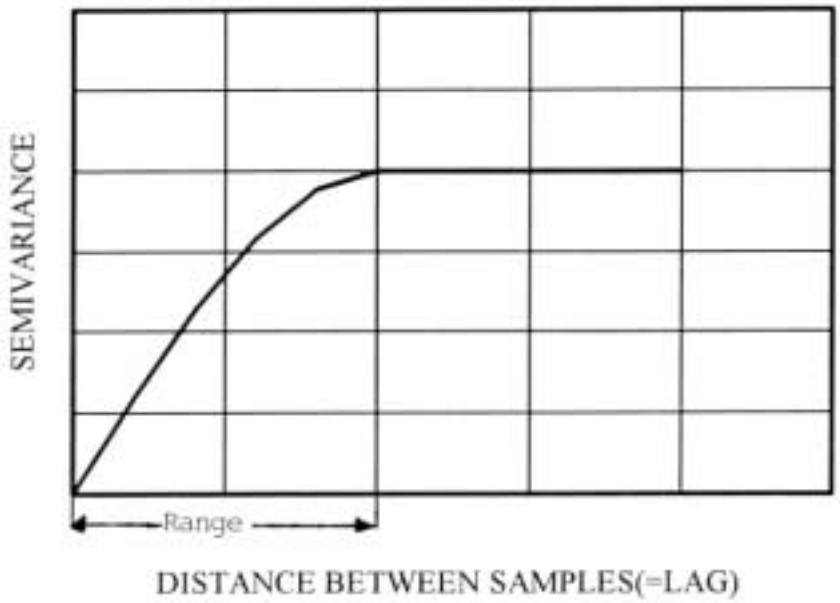


Figure 12. Sample semivariograms of clay content, (a) for top soil and (b) for subsoil. (from Webster and Oliver 1990).

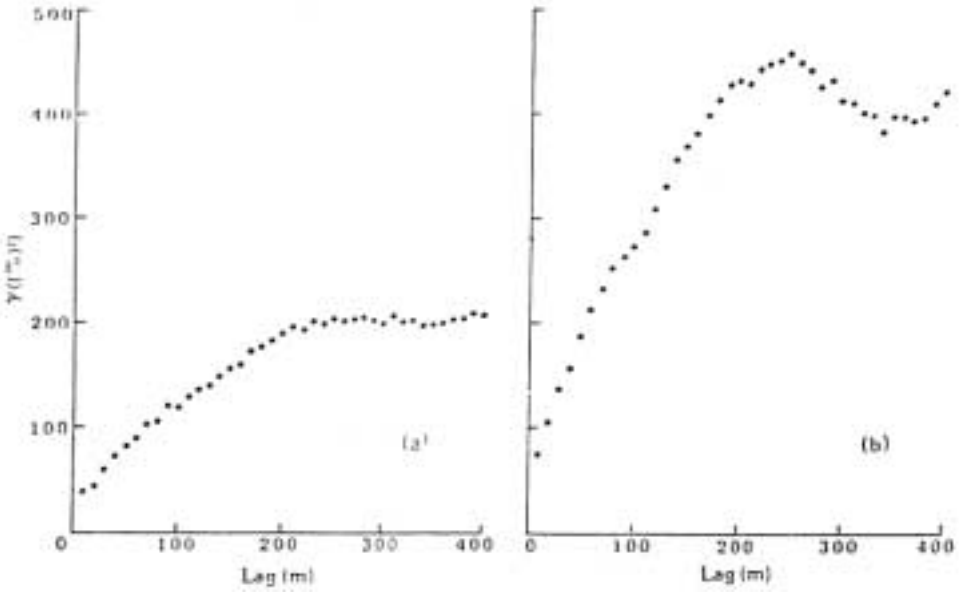


Figure 13. Example of all nugget variance semivariograms (from Webster and Oliver 1990).

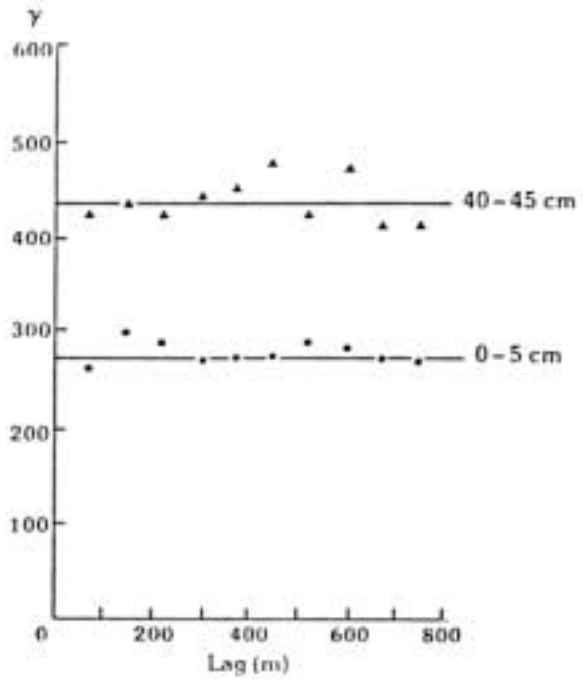


Figure 14. Sample and model semivariograms for sand content from a linear soil survey (from Webster and Oliver 1990).

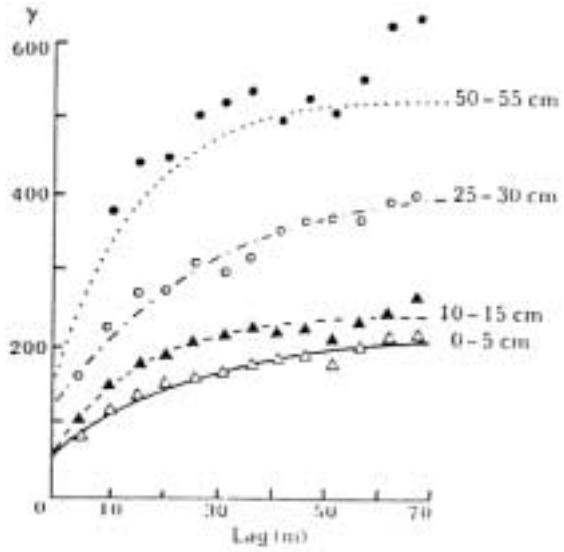


Figure 15. Kriging error for estimating the mean value in a region with samples of different sizes and varying spacing. A linear semivariogram is assumed. (from Webster and Oliver 1990).

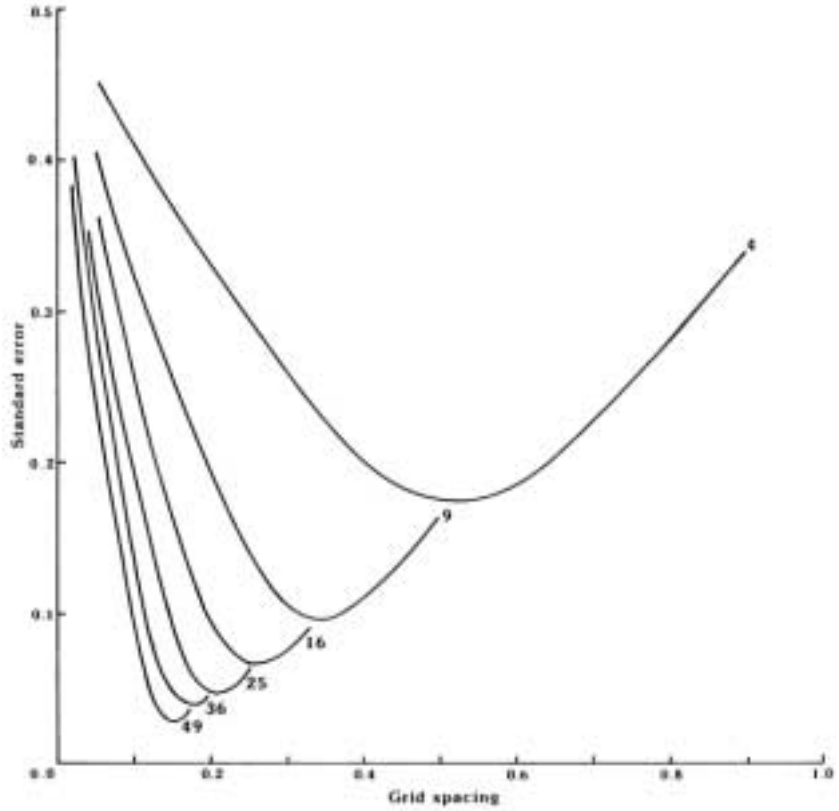


Figure 16. Kriging error against sample size, lower graph. The upper curve is the standard error calculated assuming no spatial dependence (from Webster and Oliver).

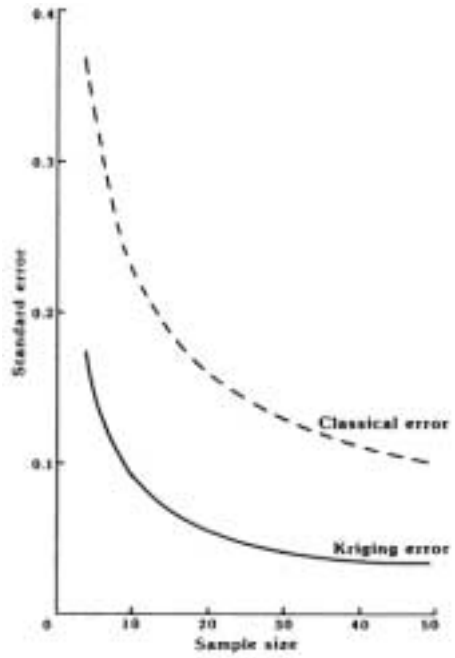


Figure 17. Sample locations at Calvert landfill site (one cell). From Fawcett et al 1993

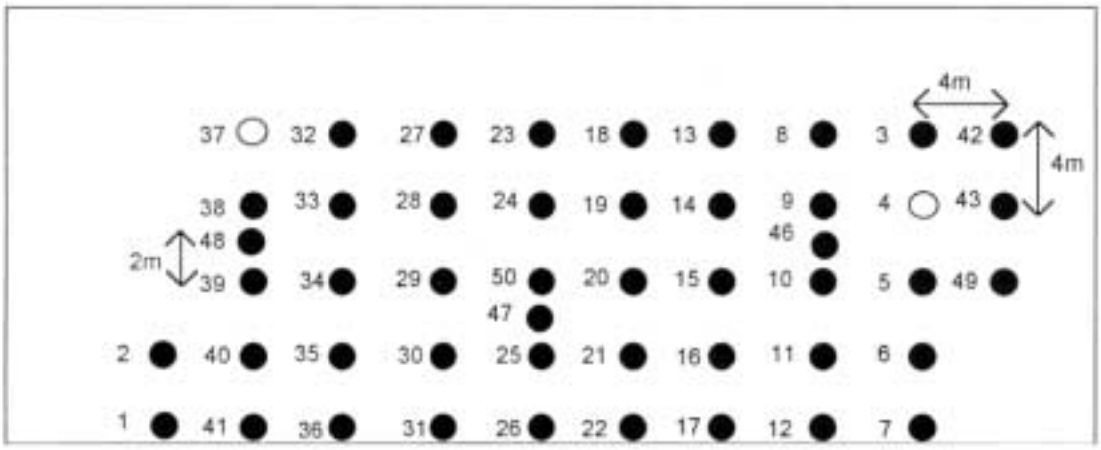


Figure 18. Sample and model semivariograms for percent volatile solids in the Calvert landfill cell.

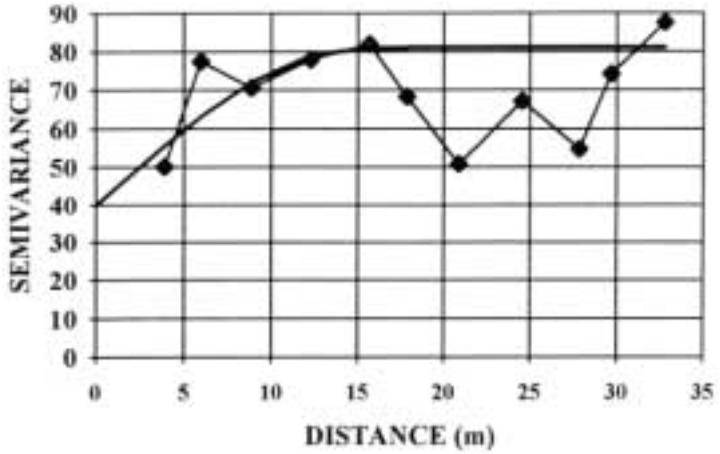


Figure 19. Estimation error of volatile solids in the Calvert landfill cell for different sample numbers.

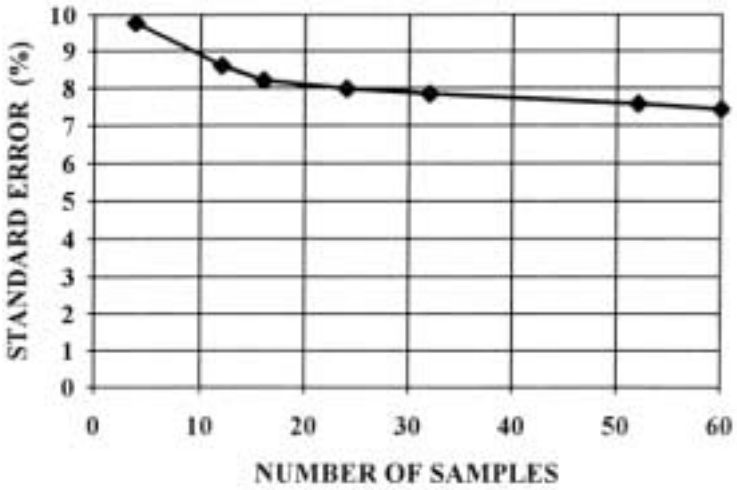


Figure 20. Sample size (labels on curves) necessary to achieve a specified probability of hitting a circular target of prescribed relative area (Target Area/Search Area). See Box 2

.

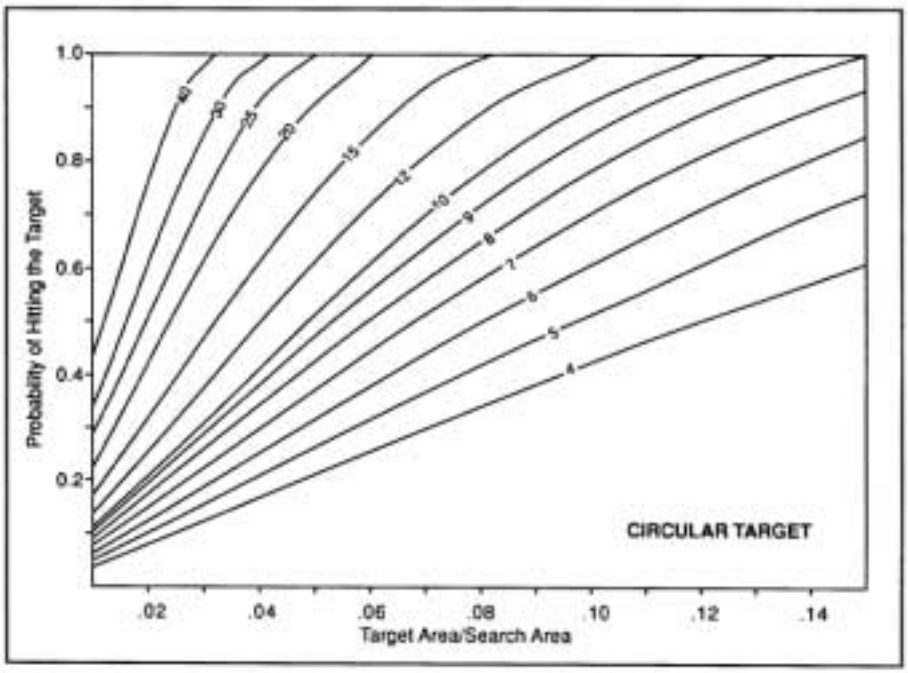


Figure 21. Sample size (labels on curves) necessary to achive a specified probability of hitting an elliptical target (aspect ratio 4:1, oriented parallel to grid direction) of prescribed relative area (Target Area/Search Area). See Box 2

.