PB83-206979

Preparation of Soil Sampling Protocol: Techniques and Strategies

Environmental Research Center
Las Vegas, NV

May 83

U.S. DEPARTMENT OF COMMERCE
National Technical Information Service
NTIS

EPA-600/4-83-020
May 1983

PB83-200979

PREPARATION OF SOIL SAMPLING PROTOCOL: TECHNIQUES AND STRATEGIES

by

Benjamin J. Mason
ETHURA
McLean, Virginia 22101

under subcontract
to

Environmental Research Center
University of Nevada, Las Vegas
Las Vegas, Nevada 89154

Cooperative Agreement Number: CR808529-01-2

Project Officer

Robert D. Schonbrod
Exposure Assessment Research Division
Environmental Monitoring Systems Laboratory
Las Vegas, Nevada 89114

Environmental Monitoring Systems Laboratory
Office of Research and Development
U. S. Environmental Protection Agency
Las Vegas, Nevada 89114

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/4-83-020

2.

3. RECIPIENT'S ACCESSION NO. [illegible]

4. TITLE AND SUBTITLE
PREPARATION OF SOIL SAMPLING PROTOCOL: TECHNIQUES AND STRATEGIES

5. REPORT DATE May 1983

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S) Dr. Benjamin J. Mason
Ethura
McLean, VA 22101

8. PERFORMING ORGANIZATION REPORT NO

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Environmental Research Center
University of Nevada, Las Vegas
4505 S. Maryland Parkway
Las Vegas, NV 89154

10. PROGRAM ELEMENT NO. CCRL1A

11. CONTRACT/GRANT NO. CR808529-01-2

12. SPONSORING AGENCY NAME AND ADDRESS
Environmental Monitoring Systems Laboratory
U.S. Environmental Protection Agency
P.O. Box 15027
Las Vegas, NV 89114

13. TYPE OF REPORT AND PERIOD COVERED Final

14. SPONSORING AGENCY CODE EPA/600/07

15. SUPPLEMENTARY NOTES
Prepared as a product of the Exposure Assessment Research Project

16. ABSTRACT

This report sets out a system for developing soil sampling protocols that can be used to meet the needs of the environmental scientist working under a number of situations. The body of the report discusses the factors that influence the selection of a particular sampling design and the use of a particular sampling method. Statistical designs are discussed along with the appropriate analysis of the data. Three appendices are included. One is a discussion of the steps that must be taken to arrive at the desired protocol. The remaining appendices present two examples of protocols; one for a shallow spill situation and the second for a deep contamination plume. A technique called kriging is presented as an approach for handling the analysis of data collected during a soil sampling program. This technique allows the researcher to develop maps of the pollution levels and to assign a statistical precision to the data at each point on the map. The data presented on these two maps can be evaluated to identify areas where additional samples are needed to reach a desired level of precision for the area covered by the maps.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Sampling, Investigations, Soil Surveys, Pollutants, Design Criteria, Statistical Analysis.	monitoring network design, geostatistical methods	1406 1407

18. DISTRIBUTION STATEMENT
Release to Public

19. SECURITY CLASS *(This Report)* N/A

20. SECURITY CLASS *(This page)* N/A

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

NOTICE

The information in this document has been funded wholly or in part by the United States Environmental Protection Agency under assistance agreement number CR 808529-01-2 to the Environmental Research Center of the University of Nevada, Las Vegas, it has been subject to the Agency's peer and administrative review, and it has been approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ABSTRACT

This report sets out a system for developing soil sampling protocols that can be used to meet the needs of the environmental scientist working under a number of situations. The body of the report discusses the factors that influence the selection of a particular sampling design and the use of a particular sampling method. Statistical designs are discussed along with the appropriate analysis of the data. Three appendices are included. One is a discussion of the steps that must be taken to arrive at the desired protocol. The remaining appendices present two examples of protocols; one for a shallow spill situation and the second for a deep contamination plume. A technique called kriging is presented as an approach for handling the analysis of data collected during a soil sampling program. This technique allows the researcher to develop maps of the pollution levels and to assign a statistical precision to the data at each point on the map. The data presented on these two maps can be evaluated to identify areas where additional samples are needed to reach a desired level of precision for the area covered by the maps.

This report was submitted in partial fulfillment of the requirements under Cooperative Agreement No. CR808529-01-2 by ETHURA under subcontract to the Environmental Research Center of the University of Nevada, Las Vegas and under the sponsorship of the U. S. Environmental Protection Agency.

CONTENTS

Appendices

FIGURES

TABLES

SECTION 1

INTRODUCTION

Environmental assessments are designed to evaluate the impacts of chemical pollutants upon man and his environment. Considerable effort has been expended in developing protocols for use in monitoring for air and water borne pollutants. The complexities encountered in sampling the soil system have been a major handicap in the development of field procedures. The Office of Pesticide Programs has promulgated a document outlining procedures for sampling for the National Soil Pesticide Survey Program (Office of Pesticide Programs, 1976). This document has proven to be useful in the sampling of soils on a nationwide basis but cannot be used effectively in many situations where hazardous wastes are encountered.

The system presented below is designed to provide the environmental scientist with a means for developing a protocol that will satisfy the soil sampling needs of agencies such as the U. S. Environmental Protection Agency (EPA). The methods presented will provide the scientist with adequate tools for producing reliable estimates of the spatial distribution of soil borne pollutants. The techniques described in this report have been borrowed from a number of areas of soil science. Each section provides information that will enable a person that is knowledgeable in the behavior of pollutants with the necessary tools for acquiring soils data that will meet most of the needs encountered in environmental monitoring work.

In order to use this document, the environmental scientist should have a familiarity with the general properties of soils and have some idea of the behavior of pollutants in the environment. A chemist and a statistician must be available for consultation from the outset of any soil sampling study in order to provide the technical input into the selection processes that are needed in developing the protocol. The field scientist must be given latitude in modifying the protocol to meet unusual conditions not covered during planning; but, these judgement calls must be made in concert with the chemist and the statistician.

Scientists working for the EPA will need to follow procedures in the areas of chain-of-custody and quality assurance that are more rigid than those likely to be required by a non-regulatory agency. Sections dealing with these areas are included in this report; but they have been made brief because documents are available from the EPA National Enforcement Investigations Center (NEIC) that cover these areas in considerable detail. Chemical analytical methods are not discussed in this report even though these methods are factors that must be considered in selecting the appropriate sample collection techniques.

The basic format of this report is designed to allow the environmental scientist to arrive at the selection of a sampling protocol by answering a series of questions found in Appendix A. The answers to the questions will then lead to one of several approaches that can be used in a particular study setting. The sections of Appendix A are arranged in a chronological order leading from initiation of the study through to the statistical analysis of the data and point back to one of the sections in the main body of the report. Sections 2, 3, and 4 outline some of the basic concepts needed to plan and initiate a soil sampling study. Section 5 sets out the types of background data that are needed in order to properly evaluate the situation in a particular soil sampling study area. Section 6 presents statistical designs, Section 7 the methods for collecting samples and Section 8 data analyses.

The sections on sample design and data analyses present an approach that allows the scientist to handle and present the data that has been acquired through a sampling program designed to estimate the levels and distribution of pollutants found in the soil environment. This approach, called kriging, has been used in evaluating the distribution of radioactivity at a number of sites both in the United States and on the Enewetak Atoll. The method was developed in France for use in ore evaluation and has recently been used in the United States by a number of scientific disciplines. The use of the technique is still in its infancy in the environmental fields but appears to hold considerable promise as a tool for evaluating pollution patterns. The major advantage of kriging is in the ability to develop estimates of concentrations over a geographic area and also provide a measure of the confidence limits to be placed on the data at any point. The statistical errors at each point can be plotted on a map which shows the isopleths for the error terms. The resulting map allows one to identify where additional sampling sites would be most beneficial in increasing the reliability of the data set.

SECTION 2

THE SOIL SYSTEM

Soil sampling as presented in this report encompasses the entire mass of unconsolidated mantle of weathered rock and loose earth material lying above solid rock. This definition is that used by the engineer rather than the agriculturalist (Soil Science Society of America, 1965).

The physical and chemical characteristics of the soil system influence the transformation, retention and movement of pollutants through the soil. Clay content, organic matter content, texture, permeability, pH, and cation exchange capacity will influence the rate and route of migration. These factors must be considered in the process of designing a sampling plan. The agricultural worker considers these factors but does not focus the sampling design on them because a farmer is interested in how much lime or fertilizer to apply to a field and not in the avenues of movement of that fertilizer through the soil system. Little consideration is given to the spatial variability of a field. Occasionally a farmer may fertilize two different soil types at different rates if the yield gains and fertilizer cost savings can justify the time and effort required. In such a case the soil scientist sampling the farmer's fields may take separate samples of each soil type. The environmental scientist on the other hand is interested in a number of possible types of pollution and routes of migration. This cannot be addressed with a single sample or a single composite sample. Therefore, some form of statistical sampling design must be used to evaluate the pollution in a soil system. This report will present several options for sampling soils that are available for evaluating pollution migration.

Environmental sampling must take into consideration one of the key characteristics of the soil system -- extreme variability. Cline (1944) noted that even though it was common knowledge that analytical errors are much less than sampling errors, little attention has been given to studies that would provide data for developing a sound sampling procedure. Cline further noted that, "the limit of accuracy is determined by the sample not by the analysis". The limited knowledge about soil variability changed very little until about two decades ago.

Nielsen and his associates (Nielsen et al., 1973; Warrick et al., 1977; and Vieira et al., 1981) have begun to study the spatial variability of the soil system in an attempt to develop reliable predictions of water movement through the soil. Campbell (1979) makes note of the fact that, in 1915, J. A. Harris discussed the effects of soil variability on the results of experiments. Campbell has used this knowledge in evaluating approaches to delineating soil mapping units (Campbell 1977, 1978, 1979, 1981). Rao et al. (1979) reviews other work where spatial variability was considered.

Petersen and Calvin (1965) also note: "Soil properties vary not only from one location to another but also among the horizons of a given profile. The horizon boundaries may be more distinct than are the surface boundaries of a soil classification unit. Here, also, however, zones of transition are found between adjacent horizons". The magnitude of sampling errors between layers of soil tends to be less than the magnitude of sampling errors in a horizontal direction. Disturbed or plowed soils are reported to be more variable than virgin soils in most cases (Chapman and Pratt, 1961).

One measure of variation is the coefficient of variation (CV)*. Coefficients of variation for soil parameters have been reported ranging from as low as 1 to 2% to as high as 850%. White and Hakonson (1979), for example, noted that the CV for plutonium in the soils of a number of test sites ranged from 62% to 840%. Mathur and Sanderson (1978) reported coefficients for natural soil constituents (i.e., part of the soil itself) varying from 5.6% to 75.2%. Harrison (1979) evaluated four phosphorus properties of soil and reported CV values ranging from 11% to 144% with the highest values being for available P. Hindin et al. (1966) reported a CV of 156% for insecticide residue concentrations in a square block of soil that was 30 inches on a side.

Mausbach et al. (1980) reported on a study conducted by the Soil Conservation Service (SCS) laboratory in Lincoln, Nebraska. Matched pairs of samples were collected from areas within a soil series. The samples were stratified by a number of factors in order to reduce the variability. The samples were selected from the modal phase of the series and were collected at distances that ranged from 2 to 32 km from the other member of the pair. The authors note that the literature indicates that up to half of the variability between similar soils may occur within a distance

---

* $CV = \pm s/\bar{y} \times 100$

where CV = coefficient of variation in %

s = standard deviation of sample

$\bar{y}$ = mean of sample

of one meter. (Studies are underway at Lincoln to determine variability within this one meter distance.) Mausbach et al. (1980) reported that in their study of the variability within a soil type the CV's for physical properties ranged from 9 to 40% for loess, 23 to 35% for glacial drift, 33 to 47% for alluvium and residuum, 18 to 32% for the A and B horizons, and 33 to 51% for the C horizons. The CV's for the chemical properties tended to be higher ranging from 12 to 50% for Alfisols, 4 to 71% for Aridisols, 6 to 61% for Entisols, 10 to 63% for Inceptisols, 9 to 46% for Mollisols, 16 to 132% for Spodosols, 10 to 100% for Ultisols, and 8 to 46% for Vertisols.

The variation that seems to be inherent in the data collected from any soil sampling study must be taken into consideration during the design of a sampling plan for whatever the purpose of the study. Technologies designed to take the variation into account must be employed in anv soil sampling plan. This includes the sampling design, the collection procedures, the analytical procedures and the data analyses.

SECTION 3

INITIATING THE SOIL SAMPLING STUDY

The identification of the components needed in a specific soil sampling protocol begins with a clear statement of the objectives of the study and follows through a series of steps that are required to select the sampling design, the tools to use, the size of crew, etc. Each step is characterized by a number of decisions that must be made before the protocol can be finalized. A number of the questions that must be answered in making the decisions are given in Appendix A in order to assist the reader in making the appropriate selections. The key components of any soils study are discussed in the sections that follow this brief introduction to the protocol development process.

A further assumption has been made that the scientist, his supervisors and the administrators calling for the study are committed to producing a quality study. This requires that a clear statement of the objectives be made and that adequate time is given for planning and reviewing the study. This report can help shorten the planning time by focusing the reader into the areas where decisions must be made; but, the commitment to succeed must come from those responsible for the study.

3.1 The Objective

The objective statement sets forth the specific goals that are to be met by the sampling program. This statment should be a clear, concise definition of why the study is needed and what questions the study is to answer. It could be as brief as a question such as "How much of the surface area around the accident at the XYZ railroad siding has been contaminated with trichloroethylene (TCE)?"; or, it could be the request for a detailed study in support of an enforcement action that must determine the likely avenues of contamination leading from a hazardous waste site into a community. The important point that is being made is that the goals of the study must be spelled out and must be agreed upon by all parties involved with the study.

As part of the objective statement, the environmental scientist responsible for the study should attempt to obtain a statement of the data reliability desired and the resources committed. If this cannot be clearly given at the outset, some initial goal should be indicated with the final goal selected after a number of iterations during the planning process.

3.2 Data Reliability

The scientist needs to know two things before selecting the components of the study plan -- the confidence level desired and the allowable margin of error to be met by the results.

Too often, soil sampling is done without a clear knowledge of the level of precision that can be met by the study. Laboratory systems have been developed to the point where considerable confidence can be placed in the results produced by a quality laboratory. Quality assurance is maintained throughout the studies. The sources of uncontrolled variation are too great for a field study to meet the precision found in a laboratory. The only alternative is to select a level of confidence that is acceptable and attainable within the limits of the resources available for the study.

The reliability expressed by the confidence level states the level of precision of the results generated by the study. Three confidence levels are normally used by the scientific community. These are usually expressed as $\pm$ 1 standard deviation, $\pm$ 1.96 standard deviations and $\pm$ 2.58 standard deviations which covers 68%, 95% and 99% of the total population, respectively. Another way to state this is to say the probability is 0.32 (or 1 in 3) that the value is outside of one standard deviation on either side of the mean; 0.05(or 1 in 20) that the value is outside of 1.96 standard deviations; or 0.01 (or 1 in 100) that the value is outside of 2.58 standard deviations. Where results must be absolute, a 99% confidence level should be used. Where resources are limiting or reliability is not of paramount importance, the 68% confidence level may be acceptable. Environmental sampling often attempts to attain a level of 95% confidence. The actual level is not as important as the fact that the level is known and agreed upon before the study is started.

Environmental studies can often be conducted in phases with an increase in reliability attained as each phase is completed. The design of the study should allow for this by designating the confidence level for each phase and planning the study so that each phase produces an answer. For example, the first phase of many soil sampling programs is a pilot or exploratory study. If properly designed, this can produce results with a known precision and thus avoid the problems of the "quick and dirty" look at a soil study area. The confidence level for this phase might be the 68% level. The second phase might be more definitive whose study design makes use of the data generated by the pilot study. The results are expected to reflect this increased knowledge; therefore, the study should be designed to acquire the data to meet this precision. The confidence level might be increased to 95% for this phase. There may be particular situations where it is necessary to reach a 99% confidence level. An example where this might be required would be a litigation case where there was a possibility that a home had been contaminated by chemicals from an abandoned hazardous waste landfill. The emotions and the liabiity force the scientist to attempt to meet a higher level of reliability.

The second item listed, the margin of error, is needed in determining the number of samples required to meet the precision specified above. This is often expressed as a percentage error that the scientist is willing to accept or it may be the difference that he hopes to detect from the study. The value designated should reflect the importance and use to be made of the numbers. Clean-up of a chemical spill might require a less precise estimate of pollutant levels than would a study attempting to identify subtle variations in background levels. In the spill case, the scientist wants to know if a soil sample is contaminated or not. This situation could change, however, if the spill was of a known carcinogen located next to a city's water treatment plant. The margin of error chosen is combined with the confidence level to derive an estimate of the number of samples required. The smaller the margin of error; the larger the number of samples required.

3.3 Resources

The data reliability is often modified by the time and money available for the study. An attempt should be made to estimate as precisely as possible the amount of resources that can be committed to a particular study. This commitment should be spelled out and monitored on a regular basis. The resources needed include not only money but personnel, laboratory capacity, equipment and time. Where time and staff are both in short supply the reliability of the study will have to be lower in order to meet the schedule. Planning for the study must also include resources for the handling of paper work and chain-of-custody along with shipping costs of samples where these are not handled directly by the sampling crew.

SECTION 4

TYPES OF SOIL POLLUTION SITUATIONS

There are essentially four major types of sampling situations that the environmental scientist is likely to encounter. These are:

- o Large area studies where pollution is in the surface layers, e.g., in support of an ambient monitoring effort.
- o Large area studies where pollution has moved down into the soil profile, e.g., assessing the impact from a major industrial complex.
- o Localized area studies where the pollution is in the surface layers, e.g., sampling around a recent hazardous chemical spill.
- o Localized plume studies where the major source of contamination is below the surface at some depth, e.g., sampling near a leaking hazardous waste disposal site.

The environmental scientist should attempt to determine which of these categories exist in the area to be sampled. Factors such as length of time that the site has been contaminated, the type of pollutant, the type of soil and the past use of the area must all be considered in determining which study category to use. The sampling methods and the statistical designs used in each case are slightly different. These are briefly covered in the paragraphs presented below.

4.1 Large Areas Characterized by Shallow Pollutant Deposition

The basic characteristic for this category of study is that the pollutant covers a wide area and is expected to be primarily on the surface. Studies in this class are usually applicable to surveys of ambient monitoring or monitoring downwind of a major pollutant source that has been present for a short time span. The pollutants may have migrated into the soil up to one to two feet but they tend to be located only in the surface layers of the soil. Where organic matter is present in these layers, penetration may only be a few inches. Hand tools are usually used to collect samples. The designs are usually either the stratified random or the systematic grid designs. In cases where there has been penetration and where a large number of samples are to be collected, the scientist may desire to use a power driven sampler. In this latter case, the samples may be sectioned into several layers located at depth below the surface. This will enable the scientist to evaluate the movement of the pollutant through the soil mass if this is desirable.

The fact that the pollution covers a large area, encompassing a number of soil types and topographic features, makes the use of stratified sampling desirable. The strata can be identified by soil type, by position on a slope, by parent material and in some cases by aspect and vegetation type. Compositing is often used to further reduce the variability and to reduce the cost. The large area covered by the study creates a large number of samples thus cost reduction techniques are important. The National Soils Pesticide Monitoring Program of the Office of Pesticide Programs in EPA is an example of a study conducted at this scale (OPP, 1976).

4.2 Large Areas Characterized by Deep Pollutant Deposition

These study situations are similar in character to those presented in section 4.1 with the exception of the depth of pollutant migration. The sources are such that one would expect the chemical pollutants to move down into the soil profile to a considerable depth. Sources might be major industrial complexes that existed prior to implementation of the Clean Air Act, major agricultural areas where pesticide useage has been in effect for a number of years, world wide nuclear fallout, lead contamination from automobile exhaust or benzo-a-pyrene in an urban area. The pollutants would be those that are not considered to be reactive with the surface soil constituents and would likely be soluble in water to some extent. Metals that are known to chelate with soil organic matter may fall into this class in some cases.

Sampling must be done with power equipment but the depths of sampling are usually not as deep as those presented in Section 4.4. Composite samples are not recommended if they can be avoided because of the information that is lost during the mixing. The migration of the chemical through the soil is usually one of the factors of primary concern in these situations.

The statistical designs are somewhat similar to those used in the cases discussed in Section 4.1. Stratification is desirable in most cases. Systematic grid sampling can prove to be effective if local variation is not of interest. Kriging of the results usually does not work too well because the range identified by the variogram is usually around one to two kilometers thus the intensity of sampling required to use this technique requires too many samples to be cost effective. If local variation is of interest, the environmental scientist may desire to identify specific local areas where information is needed and conduct a study on a smaller scale in those areas. This allows the intensity of sampling to be increased and still not run up the costs of the study.

An example of a study that is similar to the last situation would be the soil monitoring done in support of the clean-up of plutonium contaminated soils found on Enewetak Atoll (Barnes 1978).

4.3 Localized Areas of Surface Contamination

This category of sampling is probably one of the most likely to occur in the future. Spills that result from industrial or transportation accidents, fires, or unexpected leaks from storage containers usually pollute the soils in local areas near the source. The primary purpose of the sampling efforts are to map the extent of pollution and to determine the effectiveness of the clean-up operations. In some cases the polluted areas can be identified by color or other observable indicators. If the Emergency Response Team is able to be on-site within a short time after the accident or spill, sampling may be very rapid and can be done by either grid, simple random sample or, in some specialized cases, stratified sampling. Judgement sampling may be permitted in some cases but this method should not be used if follow-up monitoring is to be done. A simple random or stratified sampling plan can be done with a few samples and provide a basis for measuring the degree of clean-up that has been accomplished; thus, nothing can really be gained by using judgement sampling.

The samples are usually collected with a King tube sampler and can be composited in some cases. The fact that the pollutants found at spill sites are limited to a few identified chemicals greatly reduces the costs of analysis. Therefore, individual samples can be collected and analyzed without prohibitive costs. The scientist may desire to collect and composite a number of individual samples at each sampling location, thus reducing the variation in the results of the study. The use of individual samples allows the investigation teams to make more definitive conclusions about the pollution deposition but this is often not justified in these limited study areas.

4.4 Localized Areas Characterized by a Deeply Penetrating Plume

The EPA is currently involved in a number of studies where groundwater plumes have migrated from landfill sites or from leaking storage tanks. These plumes tend to follow the groundwater flow and tend to occur at considerable depth below the surface. Sampling is often limited to the collection of groundwater samples rather than soil samples because of the relative ease in sample collection and analysis compared to the sampling and analysis of soils. The results obtained from water samples tend to be more uniform than data from soil samples.

The location of the pollutant plume will probably require some form of phased approach that will be done in conjunction with the geologists. Sampling requires heavy drilling equipment and frequently is limited in scope due to the costs involved with obtaining the samples. Split spoons are usually used to obtain a core of the soil from depth and provide a means of obtaining a relatively undisturbed sample of the soil for use in both physical and chemical analyses. The cores obtained by this method are easily viewed at the time the sample is removed from the split spoon. Records or logs of the strata observed can be used for selecting samples for analysis and for interpretation of the results obtained.

The statistics for this type of sampling is complicated by the fact that depth plays a prominent role in the analysis of the results of the sampling. Regression may be required along with stratification with depth if any meaningful interpretation is to be made from the results. Frequently in deep sample collection situations a layer with high permeability will be encountered. Some flexibility must be included in the design to allow these samples to be collected if they are encountered. A phased study approach will allow these areas to be identified and incorporated into the final sampling phases.

Mapping of the contaminant plume becomes quite important in many enforcement cases. This fact should encourage the use of some form of systematic sampling array. Kriging can be used effectively in this kind of situation. The mathematics required for three dimensional kriging analysis may preclude the use of this technique for other than selected layers through the plume.

Sample replication, quality assurance, and decontamination all become more difficult in these cases because the study is actually being conducted below ground beyond the sight of the researcher. Cost may force the researcher to compromise on the precision of the numbers obtained. The fact that a number of samples may be collected from each hole drilled also compounds the cost and complexity of the operation. Careful planning, however, can overcome many of these problems.

It should be apparent that these four categories of studies do not cover all possible situations. There are many instances where a hybrid between two or more of them may occur. It is believed that the use of these four categories will allow the environmental scientist to adapt to meet the situations encountered in the field. The field researcher should be aware that modifications may have to be made in the plan to meet the unusual situations that often occur in field sampling. He should, however, not make changes without carefully planning the changes and consulting with the chemist and the statistician before altering the designs and techniques used.

SECTION 5

REVIEW OF BACKGROUND DATA

Any monitoring effort requires a familiarity with the study area. Too little time is usually spent in preliminary data collection, evaluation and planning. It is difficult, if not impossible, to undertake a reliable soils study without review of existing data. The sources presented below should be evaluated and studied prior to developing a plan. The areas below are presented in order to draw the scientist's focus onto those types of data that will reveal the potential location of pollutants and help evaluate their migration through the environment. Combined with site visits and interviews with local citizens a good grasp of the situation can be gained.

Libraries, museums, governmental agencies, public agencies, data bases and researchers are all sources of information that can be accumulated prior to finalizing the study plan. Often the local citizens can provide information that is not available in any of the normal research channels. The environmental scientist working on abandoned hazardous waste sites will find that often the public citizen is one of the most useful sources of unpublished data. They have often lived in the area and are familiar with the operation of the site and may even provide insight into the types of chemicals and the methods of disposal at the site. The scientist working in these cases must become a detective. Any piece of information that will help determine how and where the pollutants may migrate is useful in planning the study. Each piece of information must be sifted and evaluated in an attempt to determine how the soil system responds to such factors as flooding, movement and use.

The following listing of information is only partial and reflects the author's own experience. Each researcher should be able to use this listing as a starting point from which to develop the needed data for other studies.

5.1 Historical Data

The scientist should attempt to collect all available documents dealing with the study area including newspaper accounts, if time permits. The more informed the investigator is, the better his grasp of the situation. The result should be a knowledgeable study that addresses the pollutant problem in the context of the soil system in the study area. Historical data can help answer questions about the sources of pollution, routes of migration, uses of the area, or any data that will aid in designing a study that will acquire the necessary data. The kinds of information will vary with the site; but, in general, they deal with the history of use of the area, historical drainage patterns, groundwater flow and use, and environmental and health problems associated with the study area.

Wildlife biologists and other conservation workers familiar with the natural environment in the study area along with hunters, conservation groups and scout groups can prove to be valuable sources of information about the wildlife and vegetation changes that can reflect the impacts of pollution in the area.

Stream gauging station operators, boating clubs and sportsmen are valuable sources of information about the possible routes of migration for groundwater and pollution. Often they have noted changes in color, sediment loading, algal blooms, etc., that indicate chemicals are entering the streams in the area. This becomes especially important when abandoned hazardous waste sites are the source of pollution.

Local authorities, such as fire, police, health, engineering, highway and maintenance departments, tax departments, forestry and conservation workers can all provide valuable information on prior land use. Where spills have occurred the local fire department are often able to provide information on the movement of the spilled materials. This is especialy important if they have used any particular countermeasures on unusually toxic chemicals.

The U. S. Soil Conservation Service (SCS) along with the Cooperative Extension Service and the Agricultural Stabilization and Conservation Service (ASCS) have frequent contact with the local community and are often in the rural areas. They are interested in the soil system and are usually qualified to assist in obtaining the kinds of information that are needed about not only the history of the area but also the presence and effects of pollution. The staff of all three of these groups can usually identify the local historians in the community. SCS and ASCS both maintain files of aerial photographs of the area. These files often go back for a number of years and can give information on the uses of the area along with changes in soil character with time.

Basically the environmental scientist is attempting to reconstruct the situation with time. An attempt must be made to determine where the pollution came from, how long it has been present, where it has gone in the past, and what effects it had. Any information that will aid in answering these types of questions will assist in developing a meaningful study plan.

5.2 Geological Data

The geological character of the area is important not only for determining the routes of migration of soil pollutants but also as a factor in any attempt to stratify the area into homogenous soil types. Parent materials and bedrock can often play an important part in determining how the pollutants will react in the soil.

The U.S. Geological Survey (USGS), the Corps of Engineers and the Bureau of Reclamation all maintain information on stream conditions and stream flow. These agencies are valuable sources of data about the history of the stream channels, about dredging of channels in the streams and about flooding. These factors may play an important part in determining the rate and route of pollution migration. Groups such as the Tennessee Valley Authority, the Colorado River Commission and the Great Lakes Basin Commission have environmental scientists on their staffs that are often able to provide insight into the environmental setting of the streams and lakes in the area.

The USGS has produced many reports on the geology of parts of the U. S. Their staffs are knowledgeable on rock formations, drainage, groundwater flow and quality, and can provide maps and remote sensing data in many cases. The USGS field geologists often work closely with the various state agencies that cover areas such as mining, groundwater, construction and environmental geology. These scientists are usually familiar with the settings where studies are to be conducted; in fact, they have often been the first persons contacted when a problem with groundwater has occurred.

Any information that will tell the scientist about the nature of the bedrock, the groundwater elevations, the direction of groundwater flow and the sources of recharge to the aquifer should be acquired prior to developing the final study plan.

5.3 Soils Information

As was mentioned above, the SCS, the ASCS and the Cooperative Extension Service (County Agent) are three of the best sources of information on soils in an area and should be the first point of contact before any other soils data searches are undertaken. The state offices and the local offices of the SCS maintain information on the status of the agricultural system in the areas under their responsibility. The SCS soils reports are a good place to develop a familiarity with the soil types in the study area.

Most states maintain an agricultural school that is closely aligned with the U. S. Department of Agriculture's various offices. The Soils Departments of the Land Grant Universities are in close contact with SCS and are often closely involved in agricultural soils analysis work. Their files often contain valuable information on the nature of the soils in an area and they often know of problems that have surfaced in the past. Some of the universities have maintained samples of soils from past studies. These can, on occasion, provide a valuable insight into past pollution levels if the samples have been property maintained.

Any data that will assist in determining the soil properties, chemical composition, amount of organic matter, rates of percolation into the soil, crop history, type and amounts of clay, drainage patterns within the soil and spatial variability in the study area can be a valuable asset when time comes to interpret the results of the study as well as during the planning phases of the study.

5.4 Environmental Studies

Other scientists often are interested in the same areas where the environmental scientist is attempting to determine the levels of soil pollution. These studies often provide valuable insight into the problem, the system and are possible sources of information. Frequently the geologist working on a groundwater problem will have information on pollutant migration and soil properties that can prove to be valuable. The well driller's log books kept when exploratory borings are made for construction of highways can be used to augment the data collected by the soil investigator.

Universities in the area frequently have accumulated data as part of thesis projects and other research studies that can be used to increase the understanding of the soil system.

Where the EPA or one of its state counterparts has been investigating a particular pollution incident, the data accumulated by them along with any analyses should be consulted prior to undertaking the study. This search for data at the state level should include both the environmental agencies and the health agencies. Where the data is archived will depend upon the state involved. Each has a slightly different organizational structure.

Environmental impact statements are a gold mine of information that can save considerable time for the field scientist. Studies where highways and canals, etc., have been the subject of the EIS can greatly increase the information available for planning with little cost involved on the part of the researcher.

The investigator is attempting to find information on the pollutants, routes of migration, and effects of that migration. Therefore, any environmental study that has been undertaken in the past can provide the keys to preparing a viable study plan.

5.5 Legal Cases

Where legal action is pending at a particular location, data often is available through the various enforcement channels. This type of information is sensitive and often difficult to use due to chain-of-custody and confidentiality. Frequently government agencies will share data with each other under normal conditions but when court action is involved or possible, data is difficult to obtain and even more difficult to use in an open forum.

Where a case is closed, considerable data may be available in the various enforcement offices and in the court proceedings. This is available and can usually be obtained if the need exists. The time involved can be extensive but the data may well be worth the effort if the soils study being planned has the potential for creating controversy or of being used in litigation.

5.6 Remote Sensing

Imagery obtained from either aircraft or satellite can prove to be valuable in determining the impacts of pollutants and in identifying routes and effects of migration. Old landfill sites can often be identified from archived aerial photography which are perhaps one of the best historical records available. The

USEPA's Environmental Monitoring Systems Laboratory in Las Vegas, Nevada (EMSL-LV) is the best resource available for pollution-oriented imagery. They are knowledgeable about sources of existing imagery and also can assist in obtaining new imagery. Photographs taken in conjunction with accidents or chemical spills are a valuable resource for determining the areas where samples should be taken.

The following sources can often provide information on available imagery.

- Agricultural Stabilization and Conservation Service
- Bureau of Reclamation
- Colorado River Commission
- EROS Data Center in Sioux Falls, SD
- National Aeronautics and Space Administration
- National Archives and Record Service
- National Oceanic and Atmospheric Administration
- National Park Service
- National Weather Service
- Tennessee Valley Authority
- U.S. Air Force
- U.S. Army Corps of Engineers
- U.S. Army Map Service
- U.S. Coast and Geodetic Survey
- U.S. Commodity Stabilization Service
- U.S. Forest Service
- U.S. Geological Survey
- U.S. Soil Conservation Service

SECTION 6

STATISTICAL DESIGNS

6.1 Background for Statistical Sampling Plans

This section outlines the basic statistical designs that are available for use in soil sampling. The procedure for selecting the appropriate design is covered in Appendix A.

The purpose of any soil sampling program is to obtain information about, or a constituent of the soil. The information obtained from the study should be representative of the soil system in the study area if it is to be useful to the scientific community. Much of the data collected for soil systems in the past have been based upon samples collected according to sampling plans designed for agricultural systems. These were patterned after reports such as Cline (1944), the Soil Conservation Service's Soil Survey Manual (1951) or monographs similar to the USDA's Handbook #60 - Diagnosis and Improvement of Saline and Alkali Soils (Richards, 1969). These resources provide guidance on soil sampling but the approaches often provided are not adequate for studies dealing with soil pollution.

Soil pollution studies require that sampling results provide input into decisions that often have profound health and economic consequences. The environmental scientist desires to determine an average concentration, a measure of data reliability, the direction of movement and the location of any "hot spots" that are likely to create an undue hazard to the public or the environment. The sampling designs used must provide this type of information with maximum reliability and minimum cost. Limited laboratory capacity for conducting sophisticated analyses such as those for pesticides, priority pollutants and TCDD, the potential for litigation, and public awareness further force the environmental scientist into a detailed planning mode. When the time and expense invested in analysis, data handling, and reporting are considered, it makes little sense to invest resources in a study that lacks the planning needed to produce reliable results. Statistical designs must be incorporated into the plan at the outset; thus, the statistician and the environmental scientist must work together throughout the study. This facilitates the design of data forms, analysis of the data and interpretation of results when the study is completed.

The need for a valid statistical plan cannot be over emphasized. It is essential to know the expected variability and confidence limits of both the analytical methods used and the sampling designs employed. The sampling designs must take the natural variability of the soil system into consideration. Too often there has been a tendency to do a "quick and dirty" study with no design. A few grab samples taken at some point of interest may provide some information, but, more often than not, the results of these studies eventually come home to haunt the scientist.

The following section outlines the types of designs that can be used in soils work. Four basic sampling approaches are presented below -- simple random, stratified random, systematic and judgement sampling. Judgement sampling is included in this discussion in order to complete the list of options; however, it is not recommended as a viable approach in most pollution control work.

In this report the term sample is used to describe the individual sample of soil collected at a specific sampling site or location. The sampling site is that location within the study area that is chosen by some random procedure to be the location from which to collect a particular sample of soil. One can look upon the soils in the study area as an assemblage of all possible samples that might be collected from the area. Sampling theory is based upon the selection of some subset of the total number (N) of samples by a random selection process. The object of the sampling effort is to collect a prescribed number (n) of the individual samples at randomly selected locations. The number of samples needed to estimate the pollution level with a prescribed precision will depend upon the magnitude of the variation within the soil system. In a relatively homogeneous soil, a small number of samples may be adequate to satisfy the information needs of the scientist. A greater variation will require more samples to reach the same level of precision.

One technique that can help reduce the effects of the variation upon the statistical analysis of the data is to divide the sampling area into smaller, more homogeneous sub-areas called strata. These strata are defined by some identifiable boundry that is based upon topography, soil chemical or physical properties, or some stratigraphic feature. The identification of the strata in an area required a pilot study or prior knowledge of the area if it is to be effective. The soil population to be sampled should be subdivided into sampling units that are as homogeneous as possible. The different sources of variation within the population should be sampled if valid inferences are to be made about pollution patterns found in the study area (Petersen and Calvin, 1965). This division into relatively homogeneous sub-areas allows the statistician to remove a portion of the variation and thus reduce the statistical error term in the statistical tests. Stratification often allows a study to be conducted with fewer samples or allows the study to reach conclusions with a higher statistical precision. Generally speaking, the more stratification of the area, the greater the increase in precision. Petersen and Calvin (1965) noted, however, that "The precision increases at a decreasing rate as the strata are divided more and more until a point is reached where no further gain in precision is obtained."

Environmental pollution behavior often is difficult to understand without some means of graphic display of the spatial relationship of the data. Maps have provided a useful means of viewing and grasping the data collected in soil sampling studies. Most mapping techniques use some form of data grid to plot the analytical results. The use of the grid, which is a systematic sampling design, has evolved as a result of a desire to provide sampling coverage for portions of the entire study area rather than at only certain randomly selected points. Stratification by soil types offers some improvement over a simple random sample (see discussions of these designs presented below) but the grid provides the most uniform coverage of the study area. (Actually the grid pattern is nothing more than a systematic or uniform stratification of the area into blocks or sub-areas.) The variance obtained by systematic sampling is often less than that derived from simple random sampling. Where plumes of pollution are expected, this approach appears to be the only reliable method for locating the plume and measuring concentrations in the plume. Another advantage of the systematic sample grid is that the data can be easily mapped by most computer plotting routines. The technique called kriging is most effective when used in conjunction with a systematic sampling plan (see below for discussion on kriging).

The stratified random sample plan and the systematic sample plan can be considered to be refinements in statistical designs whose purpose is to make the survey more efficient. This efficiency may result from either obtaining a smaller sampling error with the same number of samples or from reducing the number of sample units required to produce a specified sampling error. Where the scientist knows little about the area to be sampled, a preliminary study may be required. This preliminary study should be either a simple random design or a systematic design with a coarse grid. If kriging is to be used, a transect of sufficient length is necessary in order to conduct the calculations.

Compositing of a number of subsamples is another technique that is often used to reduce the effects of variation and thus increase the precision of the numbers obtained during the sampling study. One of the major advantages of compositing is the gain in precision obtained at no increase in analytical cost. Frequently soil scientists will collect a large number of samples from a farmer's field. These are then mixed in the laboratory and an analysis performed on the composited sample. If a number of subsamples are analyzed from the composite sample, the range of the value obtained decreases in proportion to the square root of the number of sampling units contributing to the composite sample (Cline, 1944).

The problem with compositing of samples is the fact that only an unbiased estimate of the mean is obtained. Additional data on individual samples must be collected to augment the composite sample data. If any statistic other than the mean is required, a single composite sample is completely inadequate (Cline, 1944).

Compositing to reduce costs assumes that the soil is homogeneous and therefore the number of analysis required can be reduced. Compositing done to reduce the variability in the data acknowledges that the variability is present but chooses to overcome this by smoothing the effects of the variation. Pollution studies often use composite samples in order to reduce costs but in the process the very data desired is lost. The environmental scientist is looking for the presence of chemicals. If a small area of contaminated soil is composited with a large volume of uncontaminated soil the resulting analysis often is below the minimum detection limit of the analytical methods; thus, valuable data is lost and an erroneous conclusion may be reached.

With this basic background on the statistical concepts used in soil sampling in mind, each of the statistical designs are discussed in the following subsections. The first two subsections are based primarily on Petersen and Calvin (1965) and Cochran (1965) while the remaining sections are a compendium of information from a number of the sources listed in the Reference Section.

6.2 Simple Random Sampling

A random sample is any sample in which the probabilities of selection are known. Random samples are selected by some method that uses chance as the determining factor for selection. The chance mechanism used may range from a simple "toss of the coin" to the use of a random number table. The choice can be one of convenience as long as chance is the means used to make the final sample selection. The chance of selection of any individual in the population can be calculated using the laws of probability. The random sample by definition is free of selection bias.

Simple random sampling is a limiting case of random sampling. In simple random sampling of soils, the chances of selection of any particular segment of the soil system must be the same, in other words each member of the soil population must have an equal probability for selection. If a two inch core sampler is used to sample the soil, the total number of possible samples is determined by dividing the total area of the study boundaries by the cross sectional area of the soil core. For example a one square mile area would contain approximately 1,278,000,000 individual soil samples (640 X 43,560 X 144)/ (1 X 1 X π).

Simple random sampling is the basis for all probability sampling techniques used in soil sampling and serves as a reference point from which modifications to increase the efficiency of sampling are evaluated. Simple random sampling in itself may not give the desired precision because of the large statistical variations encountered in soil sampling; therefore, one of the other designs may be more useful. Where there is a lack of information about the area to be studied or about the pollution distribution, the simple random sampling design is the only design other than the systematic grid that can be used effectively.

In soil sampling, the unit of soil taken from the area is usually a volume of soil, i.e., a core, a cube of soil or a shovel full of soil. Occasionally there is a need to determine the deposition of a particular pollutant on a per unit area basis. In this case a known area of soil is collected for the sample. This has been done with radioactive fallout in the past. Results of soil sampling programs are usually expressed in either a per unit volume, per unit area, or per unit weight. The bulk density of the soil is the common denominator for all three of these units. If conversions between these units are planned, several measurements of bulk density should be made.

6.2.1 Determination of the Number of Samples Required

The procedures used in this section for determining the number of samples required to meet a predetermined precision is the basis for the allocation of samples to a strata in the stratified random design and can be used to determine the number of sample points required in the systematic sample design; therefore, the information in this section will only be presented once in this report.

The number of samples required to obtain a given precision with a specific confidence level can be obtained from equation 6.2.1 if some measure of the variance can be obtained from either a preliminary experiment, a pilot study, or from the literature. References such as Beckett and Webster (1971) and Mausbach et al. (1980) can be used as a first approximation of the variance of soil samples. This can be used to develop the initial estimate of the number of samples needed. A preliminary study will then further refine this number once an estimate for the variance of the soils in the specific area is known.

$$n = t_\alpha^2 s^2 / D^2 \qquad (6.2.1)$$

Where D is the precision given in the specifications of the study; s^2 is the sample variance and t is the two-tailed t-value obtained from the standard statistical tables at the α level of significance and (n-1) degrees of freedom. D is usually expressed as $\pm$ a specified number of concentration units (i.e. $\pm$ 5.00 ppm). Equation 6.2.1 can also be written in terms of the coefficient of variation (CV). This conversion yields Equation 6.2.2.

$$n = (CV)^2 t_\alpha^2 \ / \ p^2 \qquad (6.2.2)$$

where
- n = number of samples
- CV = coefficient of variation
- $\bar{y}$ = mean of the samples
- p = allowable margin of error expressed as a percent ($D/\bar{y}$).
- t = the two-tailed t value obtained from standard statistical tables at the α level of significance and at (n-1) degrees of freedom.

Since the t-value is dependent upon the number of degrees of freedom, it is necessary to use an iterative approach to arrive at the number of samples to use. Curves can be prepared that plot the number of samples against the coefficient of variation and thus avoid the bother of the iterations.

Use of this equation assumes that the population is normally distributed and that less than 10% of all possible samples in the study area are being collected. The latter criteria is seldom exceeded in soils sampling. (In those very limited situations where this may be the case, the finite population correction must be applied. This correction, (N-n)/N, is multiplied by the variance obtained from the sampling experiment.)

The environmental scientist can gain more information on the sampling error if more than one sample is taken at each location. These replicates are used to provide a measure of the sample to sample variation.

6.2.1.1 Cost of Collection

Determination of the number of samples in the above section is based upon the coefficient of variation of the sample population. There are many cases where the number of samples required by this method is not acceptable because of the cost of sample collection, the cost of analysis or limitations imposed by the lack of available laboratory capacity to handle the analyses. The following paragraph outlines a means for integrating the costs with the precision of the estimates obtained by the sampling program.

The total cost of soil studies often follows a linear form of equation similar to equation 6.2.3. (After Petersen and Calvin, 1965)

$$C = C_o + nC_s + nC_a \qquad (6.2.3)$$

where
- n = number of samples
- C = total costs
- C_o = overhead or fixed costs
- C_s = cost of sampling
- C_a = cost of analysis

The equation is used with equation 6.2.2 to arrive at the number of samples that will satisfy the budget and still have an identified precision.

EXAMPLE: Samples costs (Cs) are \$1800 for collection, preparation and shipping. Fixed costs (Co) are \$15,000. Analytical costs (Ca) are \$2,000. The budget for the study is \$75,000. The estimated coefficient of variation obtained from another similar study conducted nearby was 25%. The precision desired on the results is $\pm$10% at a 95% confidence level.

Equation 6.2.2 indicates that approximately 26 samples should be taken. Equation 6.2.3 however indicates the following:

$$C = C_o + nC_s + nC_a$$
$$C = C_o + n(C_s + C_a)$$
$$n = (C - C_o)/(C_s + C_a)$$
$$n = (75{,}000 - 15{,}000)/(1{,}800 + 2{,}000)$$
$$n = 15.8 \quad \text{which is rounded down to 15}$$

The effects of the budgetary constraints must be resolved either by reducing the precision or else by increasing the budget. Assuming that the budget cannot be changed, equation 6.2.2 will again be used to arrive at the t value that would result from the use of the smaller number of samples.

$$n = (CV)^2 t^2/p^2$$
$$15 = (25)^2 t^2/10^2$$
$$t^2 = 2.4000$$
$$t = t_\alpha^2 = 1.5491$$

This value for t is obtained from the statistical t-tables for a two-tailed t-test for 14 degrees of freedom (n-1). This value indicates that the significance level for the test would have to drop to 85% with the smaller number of samples or the allowable margin of error (p) would have to be increased.

6.2.2 Location of Sampling Points

Once the number of samples is determined their location can be planned. A map of the study area is overlain with a grid of an appropriate scale. The starting point of the grid should be randomly selected rather than located for convienence. This can be accomplished by selecting four random numbers from a random number table. The first two numbers locate a specific grid square on the overlay. The second two identify a point within that grid square. This point is fixed on the map and the entire grid shifted so that the lower right corner of the original grid square lies on the point chosen. This procedure is simple and fast. Using this technique avoids the questions that are often raised about bias. A second alternative is to select two map coordinates at random. This becomes the starting point for the grid used in sampling. All lines are then laid out on the map overlay starting at that point.

The grid prepared in this fashion becomes the basis for the selection of the sample locations. Using the number of samples

(n) determined in Section 6.2.1, n pairs of random numbers are selected from a random number table. These pairs of numbers become the X and the Y coordinates of the sample location. This procedure is the basis for locating sampling points in all of the methods where random samples are collected. In situations similar to Love Canal, a house lot is the area to be represented by each sample. The grid intersections can be used to locate the houses to sample or a listing of the houses can be prepared and the individual samples identified by a random number procedure.

6.3 Stratified Random Sample

Prior knowledge of the sampling area and information obtained from the background data can be combined with information on pollutant behavior to reduce the number of samples necessary to attain a specified precision. The statistical technique used to produce this savings is called stratification. Basically it operates on the fact that environmental factors play a major role in leaching and concentrating pollutants in certain locations. For example, a pesticide that is attached to clay particles may accumulate in stream valleys because of soil erosion from surrounding agricultural lands. The agricultural land may have lost most of the pesticide because of the same erosion. Stratification in this case might be along soil types or along elevational changes. Soil types are frequently used as a means of stratification, especially if they are quite different in physical and chemical properties.

Examples of factors used for stratification are:

- Soil type - Comus silt loam and Baile silt loam (The Comus contains mica that is known to bind a number of pollutants.)
- Texture - Sandy loam and clay loam
- Drainage - Stream bottom, valley slope and ridgetop
- Uses - Cropland and fence rows
- Practices - No till cropland and plowed land
- Horizons - A horizon and C horizon (Surface (A) usually has more organic matter)

The whole purpose of stratification is to increase the precision of the estimates made by sampling. The stratified random sampling plan should lead to this increased precision if the strata are selected in such a manner that the units within each stratum are more homogenous than the total population. Stratification must remove some of the variation from the sampling error or else there is no benefit from the effort.

In general, the more stratification, the greater the increase in precision. As was mentioned earlier, Petersen and Calvin (1965) have pointed out that the benefit of stratification has a limit where the law of diminishing returns takes over and no further gain in precision is encountered.

At least two samples must be taken from each stratum in order to be able to obtain an estimate of the sampling error. The number of sampling units is usually allocated according to a proportion based on the land area covered by each stratum. (i.e. if the area of soil in one stratum is 25% of the total study area, then 25% of the samples would be taken from that stratum.) Proportional allocation is used in soil sampling work primarily because the variance within a general area tends to be constant over a number of soil types. A pilot study would allow the scientist to determine if this is in fact the case. If the variances are materially different, the allocation must be on the basis of optimum allocation.

The procedure used once the number of samples is determined is the same as that outlined in Section 6.2 for the simple random design. Each stratum is handled as a separate simple random sampling effort.

6.4 Systematic Sampling

The systematic sampling plan is an attempt to provide better coverage of the soil study area than could be provided with the simple random sample. The method is easy to use therefore it has been popular in many cases. Samples are collected in a regular pattern (usually a grid or a line transect) over the areas under investigation. The starting point is located by some random process; then all other samples are collected at regular intervals in one or more directions. The orientation of the grid lines should also be randomly selected. This creates problems however when a pollution plume is the subject of the investigation. The orientation of the grid should be such that the lines in one direction are parallel to the general trace of the plume. This is especially important if kriging is going to be used.

The spacing on the grid also becomes important if regionalized variable theory (this is the basis of kriging) is used to design the study. The theory is based upon the spacing of data points along the grid lines. The samples must be close enough to provide a measure of the continuity of the location to location variation within a soil sampling unit. If on the other hand a measure of the mean and variance of the population is the focus of the grid sampling array, the samples must be placed outside of the "range" of the variance of each point. This allows the environmental scientist to collect samples that are not influenced by the regionalized variables. Beckett and Webster (1971) indicate that about 50% of the reported variation occurs within the first few meters of a point. This would indicate that the range beyond which kriging is not effective probably lies at a distance of approximately ten meters or greater. Beyond this distance the mean and variance of the population are the only parameters that can be determined.

A number of studies are reviewed by Petersen and Calvin (1965) that have compared systematic sampling with a simple random or a stratified random sampling plan. The results favored the systematic sampling in nearly all cases. The optimum sampling is obtained with a triangular grid design located over the area, but the square grid is almost as efficient. The fact that the square grid is probably easier to set out in the field would suggest that a square or rectangular grid should be used unless there is some reason for desiring to optimize the placement of sampling points

The systematic sampling plan is ideal when a map is the final product. This provides a uniform coverage of the area and also allows the scientist to have points to use in developing the map. (Most mapping techniques use a grid to generate the points for plotting isopleths of concentration, etc.)

The location of the grid on the area would be according to the procedures outlined in Section 6.2.2. At each grid intersection samples would be collected according to one of the methods outlined in Section 7. It is desirable to collect duplicate samples at some of the locations in order to provide a measure of the sampling error. This will increase the precision of the estimate of concentration and also allow the researcher to check the reliability of the sampling at the same time.

There are two problems that may limit the use of this design. First, the estimation of the sampling error is difficult to obtain from the sample itself unless double sampling is used at a number of sites. The variance cannot be calculated unless some method of mean successive difference test is used to evaluate the data. The second problem area concerns the presence of trends and periodicity in the data. Both of these create problems when the direction of the grid aligns with the pattern in the data. Soil sampling seldom encounters the cyclic pattern to a degree that a problem is created. Trends however are common in soil pollution work. That is the whole purpose for the sampling in many cases. There are a whole array of methods available for handling the analysis of data from sequential sampling. An excellent reference for soil scientists working in this area is a book by John C. Davis (1973) entitled Statistics and Data Analysis in Geology. Davis spends considerable time discussing the analysis of sequences of data. Techniques such as least squares analysis, regression, filtering or time-trend analysis, autocorrelation, cross correlation, Fourier transformations, map analysis , nearest neighbor analysis, cluster analysis, contouring, trend surface analysis, double Fourier series and moving averages are presented. Kriging and multivariate analysis are also discussed. A valuable addition to this text is a series of Fortran computer subroutines for conducting most of these analyses.

Yates (1948) and Quenouille (1949) present excellent reviews of the use of systematic sampling from a statistical point of view.

6.5 Judgemental Sampling

This technique is often used with one of the other methods in order to cover areas of unusual pollution levels or where effects have been seen in the past. The problem with the approach is that it tends to lead to sloppy science and to wrong conclusions. The scientist's own bias is built into the sampling effort and the data therefore often suspect. Where the data has a potential for litigation or where there is a likelihood of emotional reactions to the results, this system should be absolutely avoided. A simple random design with a known precision can be developed that will allow the scientist to determine the presence of pollutants without the risk of creating problems that cannot be handled. If it is essential that judgemental sampling be used, duplicate or triplicate samples should be taken in order to have some measure of precision.

6.6 Phasing the Study

Often data is not available for use in planning a study in a particular area. This type of situation leads to a phased approach. The first phase of the study might be a simple random study design with a 68% confidence level. The results of this study would then be used to design a more definitive study with a 95% confidence level. This latter study could use a stratified random design or a systematic sampling grid. The grid design would allow the researcher to analyze the data using kriging and thus find where additional samples are needed to further refine the sampling design so that the entire area is covered at the 95% confidence level.

Careful planning can provide data at each phase of the study that can stand the scrutiny of the scientific and legal communities and at the same time not place all of the resources into a one shot study that does not meet the situation at a particular site. Planning takes time, but it will pay off in the long run by providing the data needed at a precision that is acceptable to most scientists. The use of phases can greatly help in this process by allowing the data to grow as the awareness of the study situation evolves.

6.7 Control Areas

Control sites are used quite often in major soils studies especially if the study is attempting to determine the extent and presence of local pollution. Sites for controls must be as representative as possible of the study area. A careful survey of the area should be made prior to the final selection. In most cases it is desirable to spend as much time searching out data on the control as on the study area. The purpose of the control area is to serve as a base line against which the results of the soil sampling study can be compared.

Soil type should be the main factor chosen in selecting the control but factors such as depth to groundwater, location in relation to pollution sources and vegetation type all should be taken into consideration in making the selection. Where pollution sources are being studied the ideal selection would be a control site that only differs from the study area by the lack of the pollution source under investigation. This is seldom possible but every attempt should be made to reduce the factors that are different.

SECTION 7

SAMPLE COLLECTION

There are two portions of the soil that are important to the environmental scientist. The surface layer (0-15 cm) reflects the deposition of airborne pollutants; especially those recently deposited pollutants. Pollutants that have been deposited by liquid spills, or by long term deposition of water soluble materials may be found at depths ranging up to several meters. Plumes emanating from hazardous waste dumps or leaking storage tanks may be found at considerable depths. The methods of sampling each of these are slightly different; but, all make use of one of two basic techniques. Samples can either be collected with some form of core sampling or auger device; or, they may be collected by use of excavations or trenches. In the latter case the samples are cut from the soil mass with spades or short punches. The American Society for Testing and Materials (ASTM) has developed a number of methods that have direct application to soil sampling. These often need to be modified slightly to meet the needs of the environmental scientist that requires samples for chemical analyses since the ASTM methods are designed primarily for engineering tests. The techniques that are utilized should be closely coordinated with the analytical laboratory in order to meet the specific requirements of the analytical methods used.

The methods outlined below are for the collection of soil samples alone. At times it is desirable to collect samples of soil water. In these cases use can be made of some form of suction collector. The statistical designs would be the same no matter which of the soil water collectors was used. In those cases where suction devices are used, the sampling media is water and not soil even though the samples are a good reflection of soluble chemicals that may be moving through the soil matrix. These methods are not discussed in this report.

7.1 Surface Sampling

Surface soil sampling can be divided into two categories -- the upper 15 cm and the upper meter. The very shallow pollution such as that found downwind from a new source or at sites of recent spills of relatively insoluble chemicals can be sampled by use of one of the methods listed in Section 7.1.1. The deeper pollutants found in the top meter are the more soluble, recent pollutants or those that were deposited on the surface a number of years ago. These have begun to move downward into the deeper soil layers. One of the methods in 7.2 should be used in those cases.

7.1.1 Sampling with a Soil Punch

A number of studies of surface soils have made use of a punch or thin walled steel tube that is 15 to 20 cm long to extract short cores from the soil. The tube is driven into the soil with a wooden mallet; the core and the tube are extracted; the soil is pushed out of the tube into a stainless steel mixing bowl and composited with other cores. Two alternates are the short King-tube samplers or the tube type density samplers used by the Corps of Engineers. (These sampling devices can be supplied by any field equipment company or by agricultural equipment companies.) The latter sampler is machined to a predetermined volume and is designed to be handled and shipped as a soil-tube unit. A number of similar devices are available for collecting short cores from surface soils.

The soil punch is fast and can be adapted to a number of analytical schemes provided precautions are taken to avoid contamination during shipping and in the laboratory. An example of how this method can be adapted would be to use the system to collect samples for volatile organic chemical analysis. The tubes could be sealed with a Teflon plug and coated with a vapor sealant such as paraffin or, better yet, some non-reactive sealant. These tubes could then be decontaminated on the outside and shipped to the laboratory for analyses.

7.1.2 Ring Sampler

Soil engineers have a tool that can be purchased from any engineering equipment supply house that can be used to collect larger surface samples. A seamless steel ring, approximately 15 to 30 cm in diameter, is driven into the soil to a depth of 15 to 20 cm. The ring is extracted as a soil-ring unit and the soil removed for analysis. These large cores should be used where the results are going to be expressed on a per unit area basis. This allows a constant area of soil to be collected each time. Removal of these cores is often difficult in very loose sandy soil and in very tight clayey soils. The loose soil will not stay in the ring. The clayey soil is often difficult to break loose from the underlying soil layers thus the ring must be removed with a shovel.

This device has not been used extensively for collecting samples for chemical analysis but the technique should offer a useful method for collecting samples either for area contamination measurements or for taking large volume samples.

7.1.3 Scoop or Shovel Sampling

Perhaps the most undesirable sample collection device is the shovel or scoop. This technique is often used in agriculture but where samples are being taken for chemical pollutants, the inconsistencies are too great. Samples can be collected using a shovel or trowel if area and/or volume are not critical. Usually the shovel is used to mark out a boundary of soil to be sampled. The soil scientist attempts to take a constant depth of soil but the reproducibility of sample sizes is poor; thus the variation is often considerably greater than with one of the methods listed above.

7.2 Shallow Subsurface Sampling

Precipitation may move surface pollutants into the lower soil horizons or move them away from the point of deposition by surface runoff. Sampling pollutants that have moved into the lower soil horizons requires the use of a device that will extract a longer core than can be obtained with the short probes or punches. Three basic methods are used for sampling these deeper soils:

- Soil probes or soil augers
- Power driven corers
- Trenching

The soil probe collects 30 or 45 cm of soil in intact, relatively undisturbed soil cores whereas the auger collects a "disturbed" sample in approximately the same increments as the probe. Power augers can use split spoon samplers to extract cores up to 60 cm long. With special attachments longer cores can be obtained with the power auger if this is necessary.

The requirement for detail often desired in research studies or in cases where the movement of the pollutants is suspected to be through very narrow layers cannot be met effectively with the augers. In these cases some form of core sampling or trenching should be used.

7.2.1 Soil Probes and Hand Augers

Two standard tools used in soil sampling are the soil probe (often called a King-tube) and the soil auger. These tools are designed to acquire samples from the upper two meters of the soil profile. The soil probe is nothing more than a stainless steel or brass tube that is sharpened on one end and fitted with a long, T shaped handle. These tubes are usually approximately 2.5 cm inside diameter although larger tubes can be obtained. The cores collected by the tube sampler or soil probe are considered to be "undisturbed" samples although in reality this is probably not the case. the tube is pushed into the soil in approximately 20 to 30 cm increments. The soil core is then removed from the probe and placed in either the sample container or in a mixing bowl for compositing.

The auger is approximately 3 cm in diameter and is used to take samples when the soil probe will not work. The samples are "disturbed"; therefore, this method should not be used when it is necessary to have a core to examine or when very fine detail is of interest to the scientist. The auger is twisted or screwed into the soil then extracted. Because of the length of the auger and the force required to pull the soil free, only about 20 to 30 cm maximum length can be extracted at one time. In very tight clays it may be necessary to limit the length of each pull to about 10 cm. Consecutive samples are taken from the same hole thus cross contamination is a real possibility. The soil is compacted into the threads of the auger and must be extracted with a stainless steel spatula.

Larger diameter augers such as the bucket auger, the Fenn auger and the blade augers can also be used if larger samples are needed. These range in size from 8 to 20 cm in diameter.

If distribution of pollutant with depth is of interest, the augers and the probes are not recommended because they tend to contaminate the lower samples with material from the surface. The probe is difficult to decontaminate without long bore brushes and some kind of washing facility. One alternative is to take several waste cores at each site prior to collecting the actual samples. This allows the probe tube to be cleaned by the scouring action of soil at similar concentrations to those found in the sample taken. This should remove any contamination leftover from previous locations. Where there is a potential for litigation, decontamination is essential to avoid any question about cross contamination. The augers have some of the same decontamination problems but the open thread surfaces allow easier access to the collection surfaces; therefore, they are easier to clean. See Section 7.8 for more detail on decontamination procedures.

One final warning about the use of the hand augers and soil probes. There are many soil scientists with back problems that have resulted from trying to extract a tool that has been inserted too far into the soil. A foot jack is a necessary accessory if these tools are to be used. The foot jack allows the tube to be removed from the soil without use of the back muscles.

7.2.2 Power Augers and Core Samplers

These truck or tripod mounted tools are used for collecting samples to depths greater than approximately 30 cm. Standard ASTM methods for use of these tools are available from the American Society for Testing and Materials or can be found at any college or university library. The methods outlined in Section 7.3 are applicable in this case and will not be discussed further.

7.2.3 Trenching

This method of soil sampling is used to carefully remove sections of soil during studies where a detailed examination of pollutant migration patterns and detailed soil structure are required. It is perhaps the least cost effective sampling method because of the relatively high cost of excavating the trench from which the samples are collected. It should therefore be used only in those cases where detailed information is desired.

A trench approximately 1 meter wide is dug to a depth approximately one foot below the desired sampling depth. The maximum effective depth for this method is about 2 meters unless done in some stepwise fashion. Where a number of trenches are to be dug, a backhoe can greatly facilitate sampling. The samples are taken from the sides of the pit using the soil punch or a trowel.

The sampler takes the surface 15 cm sample using the soil punch or by carefully excavating a 10 cm slice of soil that is 10 cm square on the surface. The soil can be treated as an individual sample or composited with other samples collected from each face of the pit. After this initial sample is taken the first layer is completely cut back exposing clean soil at the top of the second layer to be sampled. Care must be exercised to insure that the sampling area is clear of all material from the layers above. The punch or trowel is then used to take samples from the shelf created by the excavation from the side of the trench. This process is repeated until all samples are taken. The resulting hole appears as a set of steps cut into the side of the trench as is shown in Figure 7.1.

An alternate procedure that is also effective results from using the punch to remove soil cores from the side of the trench at each depth to be sampled (Figure 7.1). Care must be taken to guard against soil sloughing down the side of the hole. A shovel should be used to carefully clean the soil sampling area prior to driving the punch into the trench side.

7.3 Sampling for Underground Plumes

This type of sampling is perhaps the most difficult of all of the soil sampling methods. Often it is conducted along with groundwater and hydrological sampling. The equipment required usually consists of large, vehicle mounted augers and coring devices although there are some small tripod mounted coring units available that can be carried by several men using backpacks.

7.3.1 Usual Procedure for Underground Plume Sampling

The procedure listed here closely follows ASTM method D1586-67 in many respects. The object of the sampling is to take a series of 45.7 cm (18 in) or 61 cm (24 in) undisturbed cores with a split spoon sampler. (Longer cores can be obtained by combining several of the shorter tubes into one long split spoon.) A 15.2 cm (6 in) auger is used to drill down to the desired depth for sampling. The split spoon is then driven to its sampling depth through the bottom of the augered hole and the core extracted.

The ASTM manual calls for the use of a 63.5 kg (140 lb) hammer to drive the split spoon. The hammer is allowed to free fall 76 cm (30 in) for each blow to the spoon. The number of blows required to drive the spoon 15.2 cm (6 in) is counted and recorded. The blow counts are a direct reflection of the density

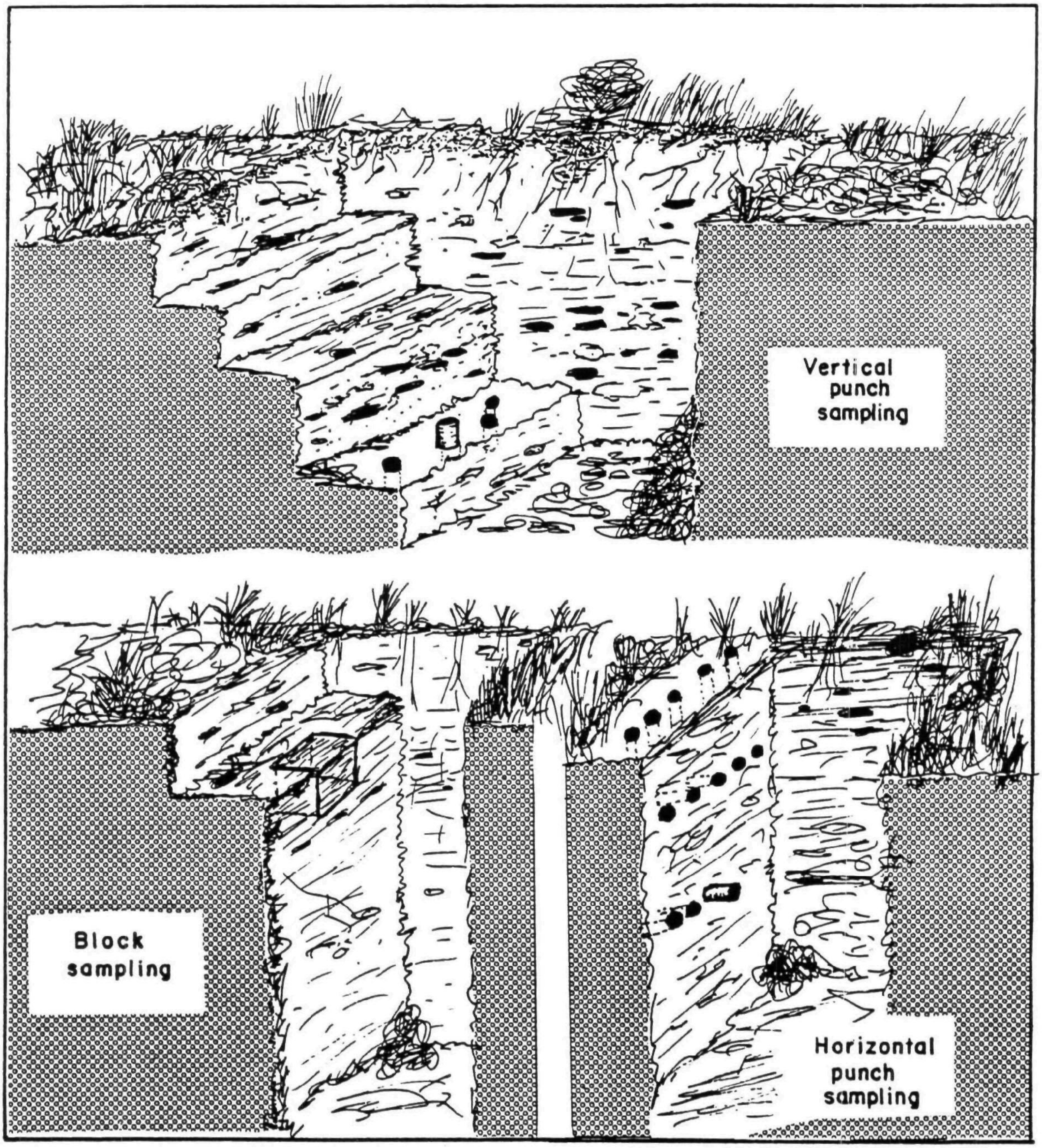

Figure 7.1 Trench sampling techniques

of the soil and can be used to obtain some information on the soil structure below surface. Unless this density information is needed for interpretive purposes, it may not be necessary to record the blow counts. In soft soils the split spoon can often be forced into the ground by the hydraulic drawdown on the drill rig. This is faster than the hammer method and does not require the record keeping necessary to record the blowcounts. Most commercial drilling companies have the equipment and the experience required to conduct this type of sampling with some supervision from the field scientist.

Samples should be collected at least every 1.5 meters (5 ft) or in each distinct stratum. Additional samples should be collected where sand lenses or thin silt and sand layers appear in the profile. This sampling is particularly important when information on pollution migration is critical. Soluble chemicals are likely to move through permeable layers such as sand lenses. This appears to be especially important in tight clay layers where the main avenue of water movement is through the porous sandy layers.

Detailed core logs should be prepared by the technical staff present at the site during the sampling operation. These logs should note the depth of sample, the length of the core and the depth of any features of the soil such as changes in physical properties, color changes, the presence of roots, rodent channels, etc. If chemical odors are noted or unusual color patterns are detected, these should be noted also. Blow counts from the hammer should be recorded on the log along with the data mentioned above.

The procedure using samples collected every 1.5 meters (5 ft) is most effective in relatively homogeneous soils. A variation in the method that is preferred by soil scientists is to collect samples of every distinct layer in the soil profile. Large layers may be sampled at several points if they are unusually thick. A disadvantage of this approach is the cost for the analyses of the additional samples acquired at a more frequent interval. The soil horizons or strata are the avenues through which chemical pollutants are likely to migrate. Some are more permeable than others and are thus more likely to contain traces of the chemicals if they are moving through the soil. Generally speaking the sands and gravels are more prone to contamination than are the clays because of increased permeability. This is especially true out on the leading edges of the plume and shortly after a pollutant begins to move. Low levels found in these layers can often serve as a warning of a potential problem at a later date.

Decontamination of the large equipment required for plume sampling is difficult but necessary if the study is to be useable. The staged sampling using the auger then the split spoon helps reduce the chances of serious cross contamination. The auger carries considerable soil in the threads of the bit. This can only be removed with high pressure hoses.

A disadvantage of this type of sampling is the impact of the vehicle on yards and croplands. Special care must be taken to protect yards, shrubs, fences and crops. The yards must be repaired, all holes backfilled and all waste removed. Plastic sheeting should be used under all soil handling operations such as subsampling, compositing and mixing.

7.3.2 Variations in the Procedures

There are several variations for split spoon sampling. Samples collected from soils below the water table or in very soft soils may require the use of split spoons equipped with retainers in the end of the spoon. The retainer is made with flexible fingers that close over the end of the tube as the spoon is retracted from the soil.

Samples collected for the analysis of volatile organic chemicals pose a problem to the environmental scientist. The volatile chemicals can be lost during transport and handling. One option that may offer a solution to this problem is the use of brass, stainless steel or Teflon liners in the split spoon. Brass liners are available from most engineering and agricultural supply houses. The liners are easily removed when the split spoon is opened. The liner tube can be sealed with Teflon plugs and some form of sealant applied over the plug. The method is currently used for moisture determinations in agricultural and research situations. This system avoids the problems of the loss of chemicals that volatilize into the headspace of the sample jars. The liners can be discarded after analysis if necessary thus reducing the labor costs required to clean the tubes.

The main disadvantages of using this modified system is that no core log can be prepared of the sample. The author was informed that some laboratories are reluctant to develop methods that can make use of samples acquired in this fashion.

7.4 Compositing

Many sample plans call for compositing of the soils collected at a sampling location. This creates a problem from the point of view of the soil scientist. The key to any statistical sampling plan is the use of the variation within the sample set to test hypotheses about the population and to determine the precision or reliability of the data set. As was mentioned earlier, the composite sample provides an excellent estimate of the mean but does not give any information about the variation within the sampling area. Section 7.4.1 discusses one alternative that is a combination of the compositing methods and random sampling with duplication. Three methods that have been used to composite soil are presented below.

7.4.1 Estimating Sample Variance.

The problem with the statistical analysis is found in the lack of duplication within the sampling location. Each subsample is combined into the composite therefore the data that is contained in the subsample is averaged with all other subsamples. The lack of a measure of the sampling error is the cause of the problem confronting the statistician. Multiple samples taken at each location would avoid this problem but costs usually preclude this. A compromise is possible by only analyzing duplicates or triplicates at a percentage of the locations. The exact location is chosen by use of a random number table and should be identified before the study begins. The duplicates should not be two subsamples taken from the same composite sample but should be made up of a second set of subsamples.

Large cores such as those collected by split spoon can be split lengthwise in half. Each half is thus used as part of two separate composite samples. This avoids the time required to take the second set of cores but provides the duplication necessary for calculating the sampling error.

7.4.2 Compositing with a Mixing Cloth

Soil scientists often use a large plastic or canvas sheet for compositing samples in the field. This method works reasonably well for dry soils but has the potential for cross contamination problems. Organic chemicals can create further problems by reacting with the plastic sheet. Plastic sheeting, however, is inexpensive and can therefore be discarded after each sampling site.

This method is difficult to describe. It can be visualized if the reader will think of this page as a plastic sheet. Powder placed in the center of the sheet can be made to roll over on itself if one corner is carfully pulled up and toward the diagonally opposite corner. This process is done from each corner. The plastic sheet acts the same way on the soil as the paper would on the powder. The soil can be mixed quite well if it is loose. The method does not work on wet or heavy plastic soils. Clods must be broken up before attempting to mix the soil.

After the soil is mixed, it is again spread out on the cloth into a relatively flat pile. The pile is quartered. A small scoop, spoon or spatula is used to collect small samples from each quarter until the desired amount of soil is acquired (this is usually about 250 to 500 grams of soil but can be less if the laboratory desires a smaller sample). This is mixed and placed in the sample container for shipment to the laboratory. The waste material not used in the sample should be disposed of in a safe manner. This is especially important where the presence of highly toxic chemicals is suspected.

7.4.3 Compositing with a Mixing Bowl

An effective field compositing method has been to use large stainless steel mixing bowls. These can be obtained from scientific, restaurant, or hotel supply houses. They can be decontaminated and are able to stand rough handling in the field. Subsamples are placed in the bowls, broken up, then mixed using a large stainless steel scoop. The rounded bottom of the mixing bowl was designed to create a mixing action when the material in it is turned with the scoop. Careful observance of the soil will indicate the completeness of the mixing.

The soil is spread evenly in the bottom of the bowl after the mixing is complete. The soil is quartered and a small sample taken from each quarter. The subsamples are mixed together to become the sample sent to the laboratory. The excess soil is disposed of as waste.

7.4.4 Laboratory Compositing

Small sets of samples can often be composited better in the laboratory than in the field. A number of the small surface cores discussed in Section 7.1.1 can be placed in the sample bottle for shipment to the laboratory. These can then be placed in a stainless steel laboratory mixer and mixed to the degree needed by the analytical methods. This technique is the only method that may be useful for obtaining composite samples for some types of soils and may be the best method to use if compositing of samples is necessary in a particular situation.

7.5 Replicate Samples

The quality control program will require duplicate or triplicate samples from a percentage of the sampling sites. These may be collected from the composite or they may be comprised of duplicate sets of samples. The latter is the preferred method.

A question often arises about how to handle the analytical data for these multiple results. All analytical results for the field replicates should be reported. Proper statistical designs can use this data to increase the precision of the estimates made. There is a tendency on the part of many scientists to discard unusual results (outliers) and to average the remainder of the samples. The discussion on soil variability given earlier should point up the problem with this approach. The outliers are probably part of the normal, wide variation seen in soils data. Averaging the numbers in effect throws away data on the sampling error that is needed to determine the reliability of the data collected.

7.6 Miscellaneous Tools

Hand tools such as shovels, trowels, spatulas, scoops and pry bars are helpful for handling a number of the sampling situations. Many of these can be obtained in stainless steel for use in sampling hazardous pollutants. A set of tools should be available for each sampling site where cross contamination is a potential problem. These tool sets can be decontaminated on some type of schedule in order to avoid having to purchase an excessive number of these items.

A hammer , screwdriver and wire brushes are helpful when working with the split spoon samplers. The threads on the connectors often get jammed because of soil in them. This soil can be removed with the wire brush. Pipe wrenches are also a necessity as is a pipe vise or a plumbers vise.

7.7 Record Keeping

One of the vital components of the protocol is to adequately define the records required during the study. Good records become extra important if litigation results from the data collected. Every sample will be questioned in an attempt to either discredit or verify the data depending upon the side of the issue the attorney represents. Some of the records are discussed below.

7.7.1 Log Books

The sample teams should maintain an official log book of the investigation. Observations of the field conditions, equipment used, procedures followed and crew members involved are recorded for each day's sampling. These log books should be bound and all data must be recorded in ink (preferably black ink). Each log book should be maintained by the crew leader and signed by him. No erasures are allowed. When mistakes are made the data is lined out with one line only and the corrected data entered above the incorrect entry or on the next line of the log.

7.7.2 Site Description Forms

These serialized forms record the conditions at each site at the time the samples are collected. A sketch map and photographs of the site should be a part of the description. A Polaroid-type camera should be used so that the pictures of the sites can be checked before leaving the area of the sample collection. These forms and the back of the photographs should be signed and dated by the crew leader responsible for taking the samples. The NEIC site description form should be used in most cases where the USEPA is involved.

7.7.3 Sample Tags

Tags made up according to the specifications provided by NEIC should be printed for use in the soils study. A tag must be prepared for each sample. All data must be included on the tag at the time the sample is collected. Wet samples should be double bagged with the tag in the outer bag. The person collecting the sample should sign the tag.

7.7.4 Chain-of-Custody Forms

This form is perhaps one of the most important as far as the legality of the samples is concerned. Chain-of-custody traces the possession of the sample from its origin through to data analysis. Most field researchers are not accustomed to observing the care needed to insure the safe custody of their samples. The samples must be in the the physical custody of the scientist collecting the sample or else be secured in a facility with controlled, limited access until the samples are signed for and transfered to another responsible party. Samples must not be left unattended in an unlocked vehicle for any reason. There is nothing more disconcerting to technical representatives of the regulatory agencies than to spend hours working with data collected by field teams and then find the data is open to question because the chain-of-custody had been violated. Samples are a valuable resource and should be treated accordingly.

7.8 Decontamination

One of the major difficulties with soil sampling arises in the area of cross contamination of samples. The most reliable methods are those that completely isolate one sample from the next. Freshly cleaned or disposable sampling tools, mixing bowls, sample containers etc. are the only way to insure the integrity of the data.

Field decontamination is quite difficult to carry out, but it can be done. Hazardous chemical sampling adds another layer of aggravation to the decontamination procedures. The washing solutions must be collected for disposal at a waste disposal site. The technique outlined below has been used under field conditions.

7.8.1 Laboratory Cleanup of Sample Containers

One of the best containers for soil is the glass canning jar fitted with Teflon or aluminum foil liners placed between the lid and the top of the jar. These items are cleaned in the laboratory prior to taking them into the field. All containers, liners and small tools should be washed with an appropriate laboratory detergent, rinsed in tap water, rinsed in distilled water and dried in an oven. They are then rinsed in spectrographic grade solvents if the containers are to be used for organic chemical analysis. Those containers used for volatile organics analysis must be baked in a convection oven at 105^{o} C in order to drive off the rinse solvents.

The Teflon or aluminum foil used for the lid liners is treated in the same fashion as the jars. These liners must not be backed with paper or adhesive.

7.8.2 Field Decontamination

Sample collection tools are cleaned according to the following procedure.

- Washed and scrubbed with tap water using a pressure hose or pressurized stainless steel, fruit tree sprayer.
- Check for adhered organics with a clean laboratory tissue.
- If organics are present, rinse with the waste solvents from below. Discard contaminated solvent by pouring into a waste container for later disposal.
- Air dry the equipment.
- Double rinse with deionized, distilled water.
- Where organic pollutants are of concern, rinse with spectrographic grade acetone saving the solvent for use in step 3 above.
- Rinse twice in spectrographic grade methylene chloride or hexane, saving the solvent for use in step 3.
- Air dry the equipment.
- Package in plastic bags and/or pre-cleaned aluminum foil.

The distilled water and solvents are flowed over the surfaces of all the tools, bowls etc. The solvent should be collected in some container for disposal. One technique that has proven to be quite effective is to use a large glass or stainless steel funnel as the collector below the tools during flushing. The waste then flows into liter bottles for later disposal (use the empty solvent bottles for this. A mixing bowl can be used as a collection vessel. It is then the last item cleaned in the sequence of operations.

The solvents used are not readily available. Planning is necessary to insure an adequate supply. The waste rinse solvent can be used to remove organics stuck to the tools. The acetone is used as a drying agent prior to use of the methylene chloride or hexane.

Steam cleaning might prove to be useful in some cases but extreme care must be taken to insure public and worker safety by collecting the wastes. Steam alone will not provide assurance of decontamination. The solvents will still have to be used.

7.9 Quality Assurance

Quality assurance in EPA is usually handled by someone other than the sampling team. The field team leader is responsible for insuring that the quality assurance program is carried out correctly, however. The team will be required to take duplicate samples at prescribed intervals and will be required to submit field blanks of all materials used. It would be desirable to prepare a bulk soil for use as a field blank for the soil samples. This will have to be handled very carefully because of the difficulty in finding "clean" soil for use as the blank. Distilled water can be used in lieu of a soil blank. Additional samples such as equipment swipes, rinse water and solvents should be taken on a regular basis to verify the quality of the data obtained from the samples. Procedures for handling quality assurance have been outlined in an interim guideline prepared by the EPA Office of Monitoring Systems and Quality Assurance of the Office of Research and Development (OMSQA, 1980).

7.10 Safety

Toxic chemicals create a hazard for the soil sampling team. The team often is operating above plumes containing mixtures of highly toxic chemicals. The drillers and excavators are in an especially hazardous position. An industrial safety specialist should be consulted prior to undertaking a study of these highly contaminated areas. Physical examinations should be given to the crew on a regular basis unless the sampling team operates only on rare occasions in which case they should have physicals before and after the sampling effort.

Many of the field team members will not want to follow the procedures outlined by the safety officer. This should not be tolerated. This problem seems to be especially acute with the drilling crews. Every effort should be made to provide the teams with adequate training on the use of all safety equipment and recovery procedures prior to going into the field.

SECTION 8

DATA ANALYSIS

The final step in any sampling study is the analysis and interpretation of the data that has been collected. It is not necessary for the field scientist to conduct the data analysis, but his input is necessary if any interpretation of the data is to be made. Impressions and observations obtained during on site activities are needed to adequately determine the actual behavior of the pollutant.

The person doing the data analysis must keep in mind the purpose for which the samples were collected. These purposes can usually be grouped into one of the following categories.

- Estimate the level and variability of a pollutant in a geographic area.

- Determine if the pollution measured is above some standard or is higher than the ambient levels found in the control area.

- Define the areal extent and depth of the pollution and map the pattern of the distribution.

There are statistical tests available for handling data collected by each method discussed in Section 6. Prior to attempting to use any of the methods, a statistician versed in sampling design should be consulted to assure that the appropriate design is being used. An assumption has been made that this was done at the beginning and is not being done at the end of the study.

8.1 Analysis of Data From a Simple Random Design

The simplest analysis of anv of the designs is that for the simple random analysis. This analysis can be done easily on a desk calculator if the number of points is not too great. The first three parameters that should be calculated are the mean, the variance and the confidence interval. Where the results are to be compared to other areas a number of tests are likely to be used. These are discussed below.

8.1.1 Basic Parameters

The mean (Eq. 8.1.1), variance (Eq. 8.1.2) and the confidence interval (Eq. 8.1.3) are calculated by the following equations.

$$\bar{y} = \sum_{i=1}^{n} y_i/n \qquad (8.1.1)$$

$$V(\bar{y}) = \sum_{i=1}^{n} (y_i - \bar{y})^2/n(n-1) \qquad (8.1.2)$$

where y_i = ith observation
n = number of samples
$\bar{y}$ = mean of the samples
$V(\bar{y})$ = variance of the mean

$$L = \bar{y} \pm t_\alpha \sqrt{V(\bar{y})} \qquad (8.1.3)$$

where L = confidence interval
t_α = t-statistic for a two-tailed t-test at the α significance level for (n-1) degrees of freedom.
α = significance level
$V(\bar{y})$ = variance of the mean

8.1.2 Use of t-Test

The data for a control population and the data for the contaminated areas can be compared statistically using either the analysis of variance test or the Studentized t-test. The control population often has less samples than the polluted area thus some adaptation will have to be made in the data to compensate for the inequality in the number of samples in the two treatments. The t-test uses equation 8.1.4 (Li, 1959) to determine if the two sets of samples are different. This equation assumes homogeneity of variances which is most often the case in soils work.

$$t_c = (\bar{y}_1 - \bar{y}_2) / \sqrt{s^2_p \, (1/n_1 + 1/n_2)} \qquad (8.1.4)$$

where t_c = calculated t-value

$\bar{y}_x$ = mean for sample set x

s^2_p = pooled variance calculated by formula 8.1.5

n_x = number of samples in x

$$s^2_p = \left\{ \sum_{i=1}^{n_1} (y_i - \bar{y}_1)^2 + \sum_{j=1}^{n_2} (y_j - \bar{y}_2)^2 \right\} / (n_1 + n_2 - 2) \qquad (8.1.5)$$

The two data sets are considered to be from the same population (i.e., they are equal) if the calculated t-value (t_c) falls outside of the critical region (i.e., it falls within the range $-t_\alpha < t_c < t_\alpha$ where t is the value taken from the t-tables for (n_1+n_2-2) degrees of freedom and at the α significance level).

8.1.3 Analysis of Variance of Simple Random Design

If the analysis of variance (ANOVA) is used to determine the difference between the two sets of sample data, the following table can be used to do the calculations for the ANOVA. The term treatment is a term used by statisticians when handling different data sets. This should be kept in mind when reading the sections that follow. One treatment is the data from the polluted area and the second is the data from the control area.

Table 8.1.1 Analysis of Variance for Simple Random Design

Source of Variation	Degree of Freedom	Mean Squares*
Treatment	$(k-1)$	$((\Sigma T^2/n)-(G^2/\Sigma n))/(k-1)$
Within treatment	$(\Sigma n-k)$	$(\Sigma y^2-(\Sigma T^2/n))/(\Sigma n-k)$
Total	$(\Sigma n-1)$	$(\Sigma y^2-(G^2/\Sigma n))/(\Sigma n-1)$

* T^2/n = The total of each treatment squared then divided by the number of observations in the treatment.
$G^2/\Sigma n$ = the square of the sum of all observations divided by the total number of observations.
y^2 = The sum of the square of each observation.
k = the number of treatments.

To test the hypothesis that the mean for the controls is equal to the mean for the polluted soils, the within treatment mean square is divided into the treatment mean square to give the $F\alpha$ value at the significance level with $(k-1)$ and $(\Sigma n-k)$ degrees of freedom. When the calculated F value is less than the value found in the statistical tables, the hypothesis is not rejected. A calculated F value greater than $F\alpha$ leads to the conclusion that the hypothesis is false and the two treatments are statistically different. The within treatment mean square is s_p^2, therefore, the standard deviation can be obtained by taking the square root of this value.

Occasionally a study will be conducted where no control exists. This is not a recommended practice, but there are situations where there is no alternative. The analysis in this restricted case is limited simply to the mean and the variance of the data.

8.2 Data Analysis for Stratified Random Design

The origin of the stratified random design is similar to the simple random design with the exception that there is stratification of the study area into subareas. The following sections provide the procedures for the analyses of the data that can be conducted with this design.

8.2.1 Basic Parameters for the Stratified Random Design

The calculation of the mean and the variance for the stratified random design can take two forms. Only one of the forms will be presented here. Petersen and Calvin (1965) note that in cases where the variance is common between two strata and proportional allocation is used to assign samples to each strata the calculations for the mean and the variance are simplified. As was mentioned earlier, proportional allocation is common in soils work because the variance tends to be the same over strata that are in close geographic proximity. Petersen and Calvin (1965) and Cochran (1965) can be consulted for the other approach to calculating these parameters.

The mean is calculated by use of equation 8.2.1 and the variance by use of 8.2.2.

$$\bar{\bar{y}} = \left(\sum_{i=1}^{n} y_i / n \right) \qquad (8.2.1)$$

$$V(\bar{\bar{y}}) = s^2_p / n \qquad (8.2.2)$$

where $\bar{\bar{y}}$ = mean of all strata

$V(\bar{\bar{y}})$ = variance of the mean $\bar{\bar{y}}$

n = total number of observations y_i

s^2_p = pooled variance (see equation 8.1.5)

8.2.2 Analysis of Variance for Stratified Random Design

The analysis of variance for the stratified random design is similar to that for the simple random design. Table 8.2.1 presents the ANOVA table for handling the data from a soil sampling study where this design has been used.

Table 8.2.1 Analysis of Variance for Stratified Random Design

Source of Variation	Degree of Freedom	Mean Squares*
Strata	$(h-1)$	$((\Sigma T^2/n)-(G^2/\Sigma n))/(h-1)$
Within strata	$(\Sigma n-h)$	$(\Sigma y^2-(\ T^2/n))/(\Sigma n-h)$
Total	$(\Sigma n-1)$	$(\Sigma y^2-(G^2/\Sigma n))/(\Sigma n-1)$

* T^2/n = Sum of the total of each stratum squared and divided by the number of samples in the stratum.

$G^2/\Sigma n$ = The square of the sum of all observations divided by the total number of observations.

y^2 = The sum of the square of each observation

h = The number of strata.

The similarity between Table 8.1.1 and 8.2.1 can be explained by the fact that the simple random experiment with a control is nothing more than a stratified random experiment with 2 strata. The only difference is in the identifiers for treatment (k) and strata (h).

Tests and interpretation of the results of the ANOVA would be the same as that presented for the simple random design. If the strata are different, it may not be possible to determine which stratum is causing the ANOVA to show the differererence if more than two strata are used. This can be determined, however, by using tests such as the single degree of freedom, regression analysis or by calculating the least significant difference (LSD). This latter test can be used to determine if there is a difference between the stratum means ($\bar{y}$) or the differences between stratum totals (T_h). Inequality 8.2.3 can be used for testing the differences between two stratum means.

$$\left| \bar{y}_1 - \bar{y}_2 \right| > t_{\alpha} \sqrt{2s^2_p / (\sum n)} \qquad (8.2.3)$$

where $\bar{y}_h$ = mean for the h^{th} stratum

t_{α} = t value from statistical tables at the ($\sum n-h$) degree of freedom and the α significance level

$\sum n$ = total number of observations

s^2_p = Pooled variance which is also the within strata mean square from Table 8.2.1

The absolute, $|\bar{y} - \bar{y}|$, is compared to the term on the right side of the inequality; if larger, the difference is significant. Each pair of strata means are compared in this fashion.

8.3 Data Analysis for Systematic Sampling Designs

This design provides uniformly spaced data for the entire study area. The uniform distribution allows reliable maps to be drawn, trend analysis to be conducted and facilitates kriging calculations. The literature (Petersen and Calvin, 1965; Cochran, 1946; Yates, 1948) indicates that the systematic approaches provide an increased precision over the two random designs under most conditions. As was mentioned earlier, cyclic and periodic phenomena can create problems in data analysis. The systematic design is in reality a stratified design with one observation per stratum. The lack of multiple samples taken from each of these stratum precludes the use of analysis of variance because there are no degrees of freedom for use with the within stratum mean square.

A number of tests have been devised for extracting statistical data from these designs. One method analyzes the differences between observations made at adjacent nodes on the grid. Either 4, 6, or 8 differences are generated for each point on the grid. The number depends upon the design of the grid and the location of the particular pair of points. The set of differences become the source of data for calculating a within stratum mean square. The analysis would be the same as that shown in Table 8.2.1.

Regression techniques can also be used. Regression effectively removes the effects of trends over the area and still provides information about the concentration differences and the variance. These tests do not allow the researcher to identify the spatial relationship of the data points, however.

When systematic designs are used, an attempt should be made to evaluate the effects of cyclic patterns that may be present in the data. Autocorrelation analysis can be used to identify if these patterns are present or not. Frequently examination of a plot of the residuals from a regression analysis will also reveal cyclic patterns in the data. If these patterns exist, some effort should be made to remove their effects from the data before analysis is done.

8.3.1 Kriging Analysis of Systematic Data

The technique called kriging was first developed by D.G. Krige (1966) as a means for estimating gold ore reserves in South Africa. The technique develped by Krige was based upon the use of moving averages in handling systematic data. His techniques were further developed by G. Matheron (1971) at the National School of Mines in Paris and have since been expanded into a whole body of knowledge called geostatistics. Matheron called the method kriging in honor of D. G. Krige. This technique has been used in European and South African mining fields for some time but it has only recently begun to be used in pollution control work.

A number of soil scientists have explored its use and found that in cases where a mapping effort involves significant research, economic or political decisions or any kind of analysis where spatial distribution is an important part of a decision making process, the technique provides what is called the best linear estimation of the distribution of a particular soil or rock component. The estimate is unique in that it provides a minimum estimation variance for the available data set and also allows the researcher not only to develop an isopleth map but an error map as well. The technique has advantages over other methods of analyzing spatial data in that trends normally seen in environmental data (especially pollution data) can be removed from the analysis. (The term trend applies to those cases where there is a change in some property along one or more axes of the study area. Pollution from a leaking landfill would show such a trend. The further away from the landfill the lower would be the concentration thus a trend would exist.)

Nielsen and his associates (Nielsen et al., 1973; Warrick et al., 1977; and Vieira et al., 1981) have used the method to study the spatial variability of the soil system in an attempt to develop reliable predictions of water movement through the soil. Campbell (1978) used the technique in an attempt to delineate the boundaries between soil types. Burgess and Webster (1980 a, b) and Webster and Burgess (1980) present an excellent review of the technique. Olea (1975) has produced a number of monographs for the Kansas Geological Survey that provide detailed information on the technique. Associates of Dr. Olea have produced a computer package that can be used to conduct kriging and plot the resulting maps. This program ,called Surface II, is avialable at a number of computer centers around the U.S.

In those cases where kriging has been used in pollution control work, it has met with considerable favor. Madeline Barnes (1978, 1980) has used this technique in a number of major operations where radioactive materials were being removed from the environment. The most notable example was the work done on the Enewetak Atoll in the Pacific Ocean. Nielsen and his associates (Nielsen , et. al., 1973; Vieira et. al., 1981) have used the techniques for studying the variability of a number of soil water properties.

Kriging has an advantage over other statistical tools in that it provides not only the means of evaluating the spatial variability of the soil property but it also provides an estimate of the variance at each point on the map surface. The main disadvantage of kriging appears to be the complexity of the mathematics involved. Although kriging may be difficult to use under some conditions, it often makes the most use of limited available data and thus provides the best answer for the amount of data available.

The statistical basis for kriging is the theory of regionalized variables. Kriging attempts not only to estimate the values of the regionalized variable (the spatially distributed variable) but also to assess the probable error associated with the estimates. A variable is considered to be regionalized if "it varies from one place to another with apparent continuity, but cannot be represented by an ordinary, workable function" (Davis, 1973). The theory is not a new branch of statistics but in reality is an extension of the part of conventional statistics called time series analysis. It is used primarily because "conventional statistical approaches are inadequate for the description of any variable from a natural phenomenon which has a spatial distribution" (Olea, 1975) -- a common occurrence with soil systems. The intrinsic theory behind kriging is explained fully by Olea (1975) who reviews the current status of the method and by Matheron (1969, 1971) who developed much of the theory.

Figure 8.3.1 presents an overview of the procedure for using kriging. The process of the data flow begins with the selection of sampling sites even though kriging is not a sampling method. The first step in the use of kriging is to calculate the semivariance (1/2 of the variance) by use of the following equation:

$$\gamma_{(h)} = 1/2n \cdot \sum_{i=1}^{n} (y_{(i+j)} - y_{(i)})^2 \qquad (8.3.1)$$

where n = the number of pairs of points on a line

$\gamma_{(h)}$ = the semivariance

y_i = the value at point x_i

$y_{(i+j)}$ = the value at a point j distance from x_i along a line passing through x_i

The $\gamma_{(h)}$ values are plotted against the spacing h along the grid line to give the variogram. There are a number of forms of the variogram. Those shown in Figure 8.3.2 are found in a number of the references listed in this report. A number of the patterns are reported by Barnes (1978,1980). Figure 8.3.2 (a) and (b) are the two classic forms of the variogram. All of the other examples are variants of these two forms. Figure 8.3.2 (a) shows the range of influence which is the distance over which the samples are correlated. When the curve fails to pass through the origin as in (c) or (d), a "nugget" effect is seen. This results from sample grid spacing that is too wide to pick up the detail of the system. (The term originated from the gold fields where discrete nuggets of gold were found within blocks of ore.) The limit of the curve in Figures 8.3.2 (a,c,f) is called the sill and represents the variance of the entire system (i.e. there is no longer any correlation or dependence of one sample upon another). Figure 8.3.2 (e) shows pure nugget effect where all samples are independent of each other. Classical random statistics should be used in such situations. Olea (1975) notes that there is an art to fitting the correct model to the variogram. A number of the models can be tried to find the one that best fits the type of data being evaluated.

A second important property of regionalized variables is that of drift. Drift is the trend of the data over the geometric area of the investigation. If $Z(\vec{x})$ is a regionalized variable and the drift is $m(\vec{x})$ then the residual $y(\vec{x})$ is given by equation 8.3.2 (Olea, 1975).

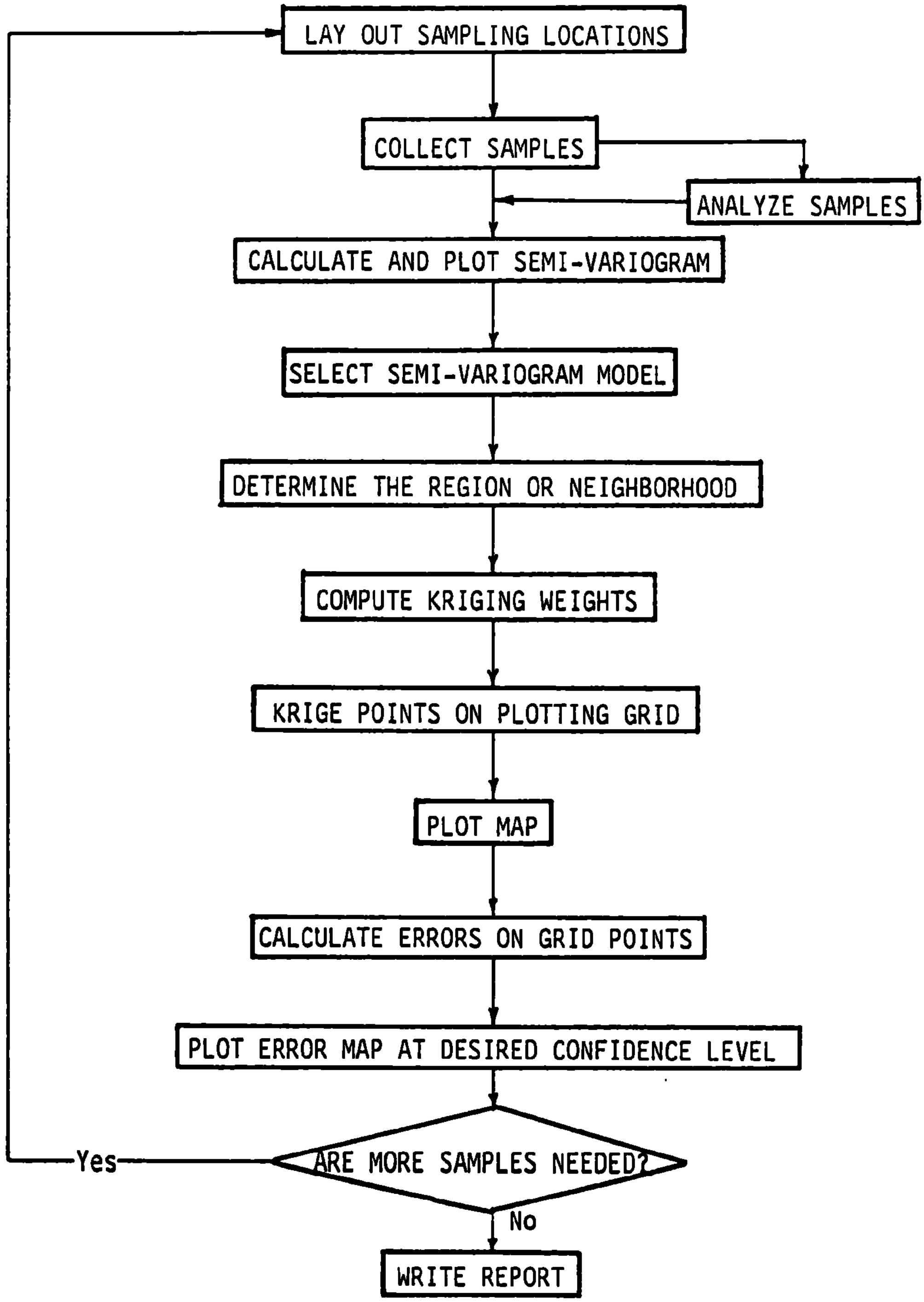

Figure 8.3.1 Flow sheet for kriging analysis

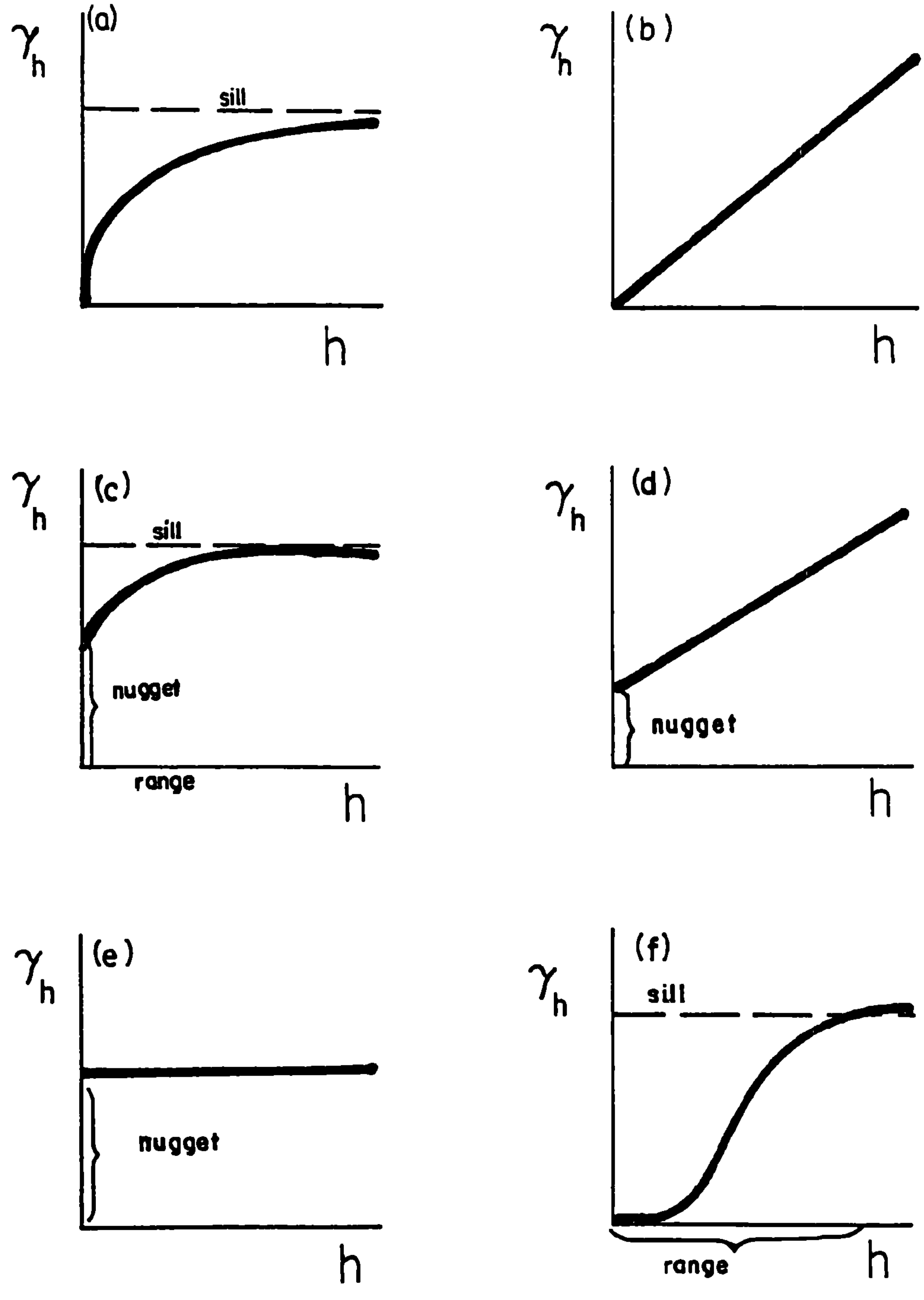

Figure 8.3.2 Examples of semi-variograms developed during kriging.

$$y(\vec{x}) = Z(\vec{x}) - m(\vec{x}) \quad (8.3.2)$$

A variogram of the residuals is plotted in a manner similar to Figure 8.3.2. This represents the variogram with the drift removed from the data. The object of this exercise is to determine the true variance of the data. Where drift is a major factor in the data, it may be necessary to go to some form of multivariate analysis rather than use kriging.

Calculation of the semivariance of the grid data for several directions will allow the researcher to determine if drift is present in the data and also allow other anisotropy to be identified. Anisotropy can often be removed from the data by transposition of the data.

Kriging also can be used with data that was not obtained from a systematic grid pattern but the mathematical manipulations of the data increase considerably. The problem with data not obtained on some linear pattern is that the variogram cannot be calculated. If additional information will allow an estimate to be made of the variogram, it is possible to krige missing points and thus develop a grid for developing isopleths. This type of situation is where the method is really an improvement over other methods of data analysis. The areas not covered by samples are partially covered by samples lying adjacent to the vacant areas. The variance of the points kriged provides an estimate of the precision of any data obtained in this way.

One of the methods that has been used for testing kriging has been to krige a set of points where data is known. Each actual data point is removed, a point is kriged then a comparison made between the actual and the kriged points. Chi-squares can be used to test the two distributions to determine the fit of the data. This comparison is also used to calculate the error term for the kriged values.

This technique is relatively new when compared to classical statistics; therefore, the methods have not been refined to the same degree as those of random sampling theory or analysis of variance. The details of this method are too extensive for inclusion in this paper. The reader is referred to Olea (1975), Barnes (1980), Campbell (1978), Burgess and Webster (1980 a,b), Webster and Burgess (1980), and Rendu (1978) for the details of the method. The Kansas Geological Survey computer programs can be obtained for a fee at a number of computer centers in the U.S. The names of those closest to the researcher can be obtained from the Survey.

REFERENCES

Barnes, M. G. 1978. Statistical Design and Analysis in the Cleanup of Environmental Radionuclide Contamination. Unpublished M. S. Thesis. University of Nevada, Las Vegas, Nevada. 52 pp.

Barnes, M. G. 1980. The Use of Kriging for Estimating the Spatial Distribution of Radionuclides and Other Spatial Phenomena. In Trans-Stat No. 13. Battelle Memorial Institute-Pacific Northwest Laboratory. Richland, Washington. 22 pp.

Beckett, P. H. T., and R. Webster. 1971. Soil Variability: A Review. Soils and Fertilizers. 34:1-15.

Burgess, T. M. and R. Webster. 1980a. Optimal Interpolation and Isarithmic Mapping of Soil Properties. I. The Semi-variogram and Punctual Kriging. The J. of Soil Science. 31(2): 315-332.

Burgess, T. M., and R. Webster. 1980b. Optimal Interpolation and Isarithmic Mapping of Soil Properties. II. Block Kriging. The J. of Soil Science. 31(2):333-342.

Campbell, J. B. 1977. Variation of Selected Properties Across a Soil Boundary. Soil Science Society of America J. 41(3):578-582.

Campbell, J. B. 1978. Spatial Variation of Sand Content and pH Within Single Contiguous Delineations of Two Soil Mapping Units. Soil Science Society of America J. 42(3):460-464.

Campbell, J. B. 1979. Spatial Variability of Soils. Annals of the Association of American Geographers. 69(4):544-556.

Campbell, J. B. 1981. Spatial Correlation Effects Upon Accuracy of Supervised Classification of Land Cover. Photogrammetic Engineering and Remote Sensing. 47(3):355-363.

Chapman, H. D., and P. F. Pratt. 1961. Methods of Analysis of Soils, Plants and Waters. University of California, Riverside, California. 309 pp.

Cline, M. G. 1944. Principals of Soil Sampling. Soil Science. 58:275-288.

Cochran, W. G. 1946. Relative Accuracy of Systematic and Stratified Random Samples for a Certain Class of Population. The Annals of Mathmematical Statistics. 17(2):164-177.

Cochran, W. G. 1965. Sampling Techniques. John Wiley. New York, New York. 413 pp.

Davis, J. C. 1973. Statistics and Data Analysis in Geology. John Wiley. New York, New York. 550 pp.

Harrison, A. F. 1979. Variation of Four Phosphorus Properties in Woodland Soils. Soil Biology and Biochemistry. 11:393-403.

Hindin, E., D. S. May, and G. H. Dunston. 1966. Distribution of Insecticide Sprayed by Airplane on An Irrigated Corn Plot. Chapt. 11 in Organic Pesticides in the Environment. A. A. Rosen and H. F. Kraybill (eds). Advances in Chemistry Series #60. American Chemical Society. Washington, D. C. pp. 132-145.

Krige, D. G. 1966. Two-dimensional Weighted Moving Average Trend Surfaces for Ore Evaluation. J. of South African Institution of Mining and Metallurgy. 66:13-38.

Li, J. C. R. 1959. Introduction to Statistical Inference. Edwards Brothers. Ann Arbor, Michigan. 553 pp.

Matheron, G. 1969. Le Krigeage Universel: Les Cahiers du Centre de Morphologie Mathmatique de Fontainebleau. Fascicule 1. l'Ecole Nationale Superieure des Mines de Paris. Paris. 83 pp.

Matheron, G. 1971. The Theory of Regionalized Variables and its Applications. Les Cahiers du Centre de Morphologie Mathmatique de Fontainbeleau, Fascicule 5. l'Ecole Nationale Superieure des Mines de Paris. Paris. 211 pp.

Mathur, S. P., and R. B. Sanderson. 1978. Relationships Between Copper Contents, Rates of Soil Respiration and Phosphatase Activities of Some Histosols in an Area of Southwestern Quebec in the Summer and the Fall. Canadian J. of Soil Science. 58(5):125-134.

Mausbach, J. J., B. R. Brasher, R. D. Yeck, and W. D. Nettleton. 1980. Variability of Measured Properties in Morphologically Matched Pedons. Soil Science Society of America J. 44(2):358-363.

Nielsen, D. R., J. W. Biggfar and K. T. Erh. 1973. Spatial Variability of Field-measured Soil Water Properties. Hilgardia. 42(7):215-260.

Office of Monitoring Systems and Quality Assurance. 1980. Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans. QAMS-005/80. U. S. Environmental Protection Agency. Washington, D. C.

Office of Occupational Safety and Health. 1979. Safety Manual for Hazardous Waste Site Investigations. U. S. Environmental Protection Agency. Washington, D. C.

Office of Pesticides Programs. 1976. Sample Collection Manual: Guidelines for Collecting Field Samples. U. S. Environmental Protection Agency. Washington, D. C. 39 pp.

Olea, R. A. 1975. Optimum Mapping Technique Using Regionalized Variable Theory. Kansas Geological Survey. Lawrence, Kansas. 137 pp.

Petersen, R. G., and L. D. Calvin. 1965. Sampling. Chapt. 5 in Methods of Soil Analysis. C. A. Black (ed.). American Society of Agronomy. Madison, Wisconsin. pp. 54-72.

Quenouille, M.H. 1949. Problems in Plane Sampling. The Annals of Mathematical Statistics. 20:355-375.

Rao, P. V., P. S. C. Rao, J. M. Davidson, and L. C. Hammond. 1979. Use of Goodness-of-fit Tests for Characterizing the Spatial Variability of Soil Properties. Soil Science Society of America J. 43(2):274-278.

Rendu, J. M. 1978. An Introduction to Geostatistical Methods of Mineral Evaluation. South African Institute of Mining and Metallurgy. Johannesburg, South Africa. 84 pp.

Richards, L. A. (ed.). 1969. Diagnosis and Improvement of Saline and Alkali Soils. U. S. Salinity Laboratory. Riverside, California. Agriculture Handbook No. 60. U. S. Department of Agriculture. Washington, D. C. 160 pp.

Soil Science Society of America. 1965. Glossary of Soil Science Terms. Soil Science Society of America Proceedings. 29(3): 330-351.

Soil Survey Staff. 1951. Soil Survey Manual. Agriculture Handbook No. 18. Soil Conservation Service. U. S. Department of Agri- culture. Washington, D.C. 503 pp.

Vieira, S. R., D. R. Nielsen and J. W. Biggar. 1981. Spatial Variability of Field-measured Infiltration Rates. Soil Science Society of America J. (in-press).

Warrick, A. W., G. J. Mullen, D. R. Nielsen. 1977. Scaling Field-measured Soil Hydraulic Properties Using Similar Media Concept. Water Resources Research. 13(2):355-362.

Webster, R. and T. M. Burgess. 1980. Optimal Interpolation and Isarithmic Mapping of Soil Properties III. Changing Drift and Universal Kriging. The J. of Soil Science. 31(3):505-524.

White, G. C. and T. E. Hakonson. 1979. Statistical considerations and Survey of Plutonium Concentration Variability in Some Terrestrial Ecosystem Components. J. of Environmental Quality. 8(2):176-182.

Yates, F. 1948. Systematic Sampling. Philosophical Transactions of the Royal Society of London. Series A. 241:345-377.

APPENDIX A

SOIL PROTOCOL DEVELOPMENT PROCESS

SECTION A-1

PURPOSE

This appendix is designed to be a guide to assist in the development of a soil sampling protocol. Extensive discussions with soil scientists in the field indicate that it is not possible nor advisable to develop a single protocol that will attempt to meet all situations. The material that follows allows the environmental scientist to progress through a protocol developmental process thus arriving at a protocol that will meet his particular needs. The four major types of study situations addressed in Section 4 are addressed here from the point of view of developing a protocol for use in each type of a setting.

SECTION A-2

PRELIMINARY BACKGROUND INFORMATION

Every field of research has a basic set of knowledge that the worker must be able to handle. The follwing types of reference materials should be read before any attempt is made to sample soils.

Be conversant with terminology in the following areas:

Basic soil science
Sampling theory
Statistical designs
Environmental monitoring
Quality assurance
Basic toxicology
Basic chemistry

Read the following materials:

Federal Register. Vol. 144:69464. December 3,1979.

Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans (OMSQA, 1980).

Safety Manual for Hazardous Waste Site Investigations (OOSH, 1979).

The three articles on kriging written by Burgess and Webster (1980 a, b) and Webster and Burgess (1980).

The article by Beckett and R. Webster (1971).

The article by Mausbach et al. (1980)

The main body of this report.

SECTION A-3

DEVELOPMENT OF THE PROTOCOL

Any protocol that will be developed by the environmental scientist will have the following basic outline:

1. Introduction

2. Objective
   2.1 Goals to be met
   2.2 Reliability that can be placed on the results
   2.3 Precision of the estimates generated
   2.4 Resources to be made available
   2.5 Schedule

3. Definition of the Magnitude of the Problem
   3.1 Known aereal extent of the contamination
   3.2 Known vertical extent of contamination
   3.3 Chemicals that have been identified
   3.4 Known concentrations
   3.5 Toxicity of chemicals
   3.6 Attitude of community toward problem and problem solvers.
   3.7 Identification of type of pattern

4. Selection of Statistical Design
   4.1 Number of samples needed to meet reliability
   4.2 Distribution of sample sites
   4.3 Frequency of sampling
   4.4 Quality assurance

5. Selection of a Sampling Method
   5.1 Type of pollution
   5.2 Depth of contamination
   5.3 Tools
   5.4 Determine how samples to be collected

6. Data Analysis
   6.1 Basic parameters
   6.2 Statistical designs
   6.3 Regression
   6.4 Spatial distribution

A suggested way to use this section is to make up a series of sheets of paper with headings similar to those in the above outline. When the questions are asked and answered make note of the answers on the pieces of paper. When you are through with the exercise you should have a good start on writing your own protocol.

3.1 Background Information.

Read section 5 (pp. 14-19) of the main body of this report; then, assemble all of the following kinds of information that are available.

1. History of the particular study area; its uses and misuses.
2. Any chemicals that have been placed on the site.
3. The source of the pollution that you are attempting to identify.
4. Ownership of the study area both past and present.
5. Any analytical information on samples collected in the area.
6. The results of all investigative reports.
7. All geological and soils information on the study area or the surrounding area.
8. All information on the toxicology of the pollutants studied.
9. The political and community situation.
10. All existing protocols pertinent to the sampling program.
11. Copies of USGS quad maps of the area.
12. Copies of all archived aerial photography of the area.
13. Copies of all current aerial photography.
14. Copies of all documentary photography of the site.
15. Copies of all reports on the site.

In addition to the above information, visit the study area and talk to members of the community that are knowledgeable about the locality and also about the specific area of the study. Attempt to determine likely control areas at this time. Remember you are an investigator at this time and are attempting to learn as much about the study area as you can in order to make your sampling study not only meet the objectives but also to gain as much insight into the pollution situation as is possible.

3.1.1 Questions to Answer.

Attempt to answer the following questions prior to the study.

1. What is the pollution source?
2. What chemicals are in that source?
3. When was it placed there?
4. How was it placed there?
5. Who placed it there?
6. Why was it placed there?
7. What kind of containers was it in?
8. What levels have been found in the environment?
9. How toxic is it and what are the known effects?
10. Has anyone locally been hurt by the pollution? (This is not an epidemiological study but is to let you know what the reactions of the community may be.)
11. How far has the pollution gone?
12. What data is available on its migration?
13. Who determined that migration had occurred?
14. Are the data on concentrations and on migration reliable?
15. What other environmental studies have been conducted in the area?
16. Is the data available for examination?
17. Will the researchers in the area cooperate?
18. Are there any court actions pending?
19. Are there any administrative actions pending?
20. Is there a potential for litigation in the case?
21. What is the public's awareness of the problem?
22. Are they frightened or angry?
23. Who are the spokesmen for the local population?

3.2 Develop an Objective Statement for the Study.

A. Has an objective statement been issued for the study?
YES - Record it and go to 3.3
NO - Review Section 3 of main report and then go to 3.2.1

3.2.1 What is this study attempting to accomplish?

The goals that are needed are explicit, clear statements of what is to be accomplished by the study. It may be nothing more than a question such as "How much of the area around the spill is contaminated?". It may, on the other hand, go to considerable length to determine the levels of contamination at various locations within the study area.

Spell out a definite set of questions that you hope to answer with the study. You should not be hesitant to ask for a more definitive goal if it is not clear what is being asked.. There are often cases where the person ordering the study does not know what is really possible and what is really needed. That is where the scientist must assist in setting the objectives of the study.

Decide upon your objectives and write them down. Get an approval on them so that you are sure that you understand what is being asked and to insure that your results will meet the needs envisioned by the administrative order.

Write down your final objective under the appropriate part of the outline then go to 3.2.2.

3.2.2 How close must the estimates be to the real mean value.

Scientists often measure the precision of the methods that they use in a laboratory but they seldom attempt to do so in the field. This is because of the difficulty in defining the variation to the point that they can control it. A level of precision should be delineated for the study. This is usually expressed as a percentage of the mean.

A second way of evaluating the precision of the data is to identify the smallest difference between two samples that you would like to be able to detect. If you are attempting to delineate variations in the pollution in a large area, you would probably want a level that was quite small because of the subtilties of the deposition pattern. On the other hand, in a situation where you are attempting to determine the area contaminated by a major spill, you may not need to be able to detect small differences in order to make the determination that one area is contaminated and another next to it is not.

If a number cannot be decided upon a twenty percent value may not be too far out of line. You are asking that your sampling determine a value that is within twenty percent of the true mean. This would be totally unacceptable for a laboratory method but it reflects the true field situation.

Record your decision and go to 3.2.3.

3.2.3 How reliable must the answers be?

The question of reliability and precision may seem to be the same. What you are attempting to determine is not how close your answers are but what is the probability that the answer you get is the correct one.

Read Section 6.1 of the main report before deciding your answer to this question. You should also read the section on simple random sampling (6.2). Statistical sampling requires that a known confidence level be given. Three are often used by the scientific community. These are the 68%, 95%, and 99% confidence levels or significance levels. Another way to state this is to say that the probability is 0.32 (or about 1 in 3) that the absolute statistical error exceeds one standard deviation around the true mean for the 68% significance level; 0.05 (or 1 in 20) for the 95% confidence level; and .01 or (1 in 100) for the 99% confidence level.

For the results to be absolute, a 99% confidence level should be used. When resources are limited or when reliability is not of paramount importance, the 68% confidence level can often be used. The important point is to know the confidence level before the study is conducted.

When there is no basis for a decision, the 95% confidence level is often used. Record your answer and go to 3.2.4.

3.2.4 What resources are available for the study?

The answer to this question may be developed by an iterative process. The answer can be a very definite limitation if it is known before the planning is done but without it the planning is often unrealistic. You may not be able to answer this question. This depends upon the administrative structure you operate under. Your final determination may not come until the planning is complete and the schedule is known. Resources include not only finances but also personnel, laboratory capacity and equipment.

Record your answer and go to 3.2.5.

3.2.5. When must the study be finished?

The schedule is determined by the resources and vice versa. You will have to set a target date and attempt to meet that date; therefore, be realistic in setting the date.

Record the target date and go to 3.3.

3.3 Determine the Magnitude of the Problem.

You need to know the magnitude of the problem as it is known at this time. The information that you must have at this point covers not only the physical extent of the contamination but also the concentration range, the toxicity of the chemicals involved, and the public relations aspects of the problem. This is not a major data gathering phase of the study but it is a reflection of the needs that may have to be met if information is totally lacking. Read Section 4 of the main report before you answer the following questions.

A. Do you know the study category you will be working in?
YES - Go to Question B.
NO - Go to 3.3.1.

B. Which category will the study area come under?

Large area with pollutant in the surface layers.
Large area with pollutant at deeper layers.
Local area with pollutant in surface layers.
Local area with pollutant in a deep plume.

The category that the study falls into determines the sampling and statistical designs that should be used. After making your decision go to 3.4.

3.3.1 What is the known areal extent of the contamination?

You need to determine the magnitude of the area to be covered. This may be defined by the type of situation you are addressing.

For example if you are looking at worldwide fallout from nuclear testing you are working at one scale; but, if you are working with a small spill site you are at the opposite end of the spectrum. There are two categories that you are attempting to identify:

- o Large areas that range from several square miles to a major portion of the U.S.
- o Local areas, such as the area around an accident site or a landfill site.

Make your determination on the basis of available data and your own knowledge of the problem.

Record your answer and go to 3.3.2.

3.3.2 What is the vertical extent of the contamination?

Information on this part of the study may be more difficult to answer than the question in 3.3.1. Examine drilling logs, investigation reports and information on the behavior of the pollutants in the soil. Those that are soluble are more likely to have moved both horizontally and vertically. Examination of the groundwater flow patterns should tell where the pollution may have gone. Your designs will depend upon this fact.

Is the pollution a soluble, deeply penetrating chemical?
Is the source located deep in the soil profile?
Is the chemical migrating upward to the surface some distance from the suspected source?
Has the chemical been at the site for a considerable period of time?

If you answer yes to any of these questions, then place the chemical in one of the two deeply penetrating categories. Record your answer and go to 3.3.3 through 3.3.7.

3.3.3 What chemicals have been identified in the study area?

3.3.4 What are the concentrations?

3.3.5 Are the chemicals toxic?

3.3.5.1 Are they toxic at the levels seen?
3.3.5.2 Have there been any suspected injuries in the study area?

3.3.6 What is the attitude of the community toward the problem?

3.3.7 Place the study into one of the four classes of studies listed in 3.3.B and go to 3.4.

3.4 Selection of a Statistical Design.

Each category of study design has an associated group of statistical designs that can best be used to address the situation. This section will attempt to direct you to those that have promise of working but the final decision should be done in conjunction with the statistician.

A. Do you know the design that you will use on the study?
YES - Record the information and go to 3.5.
NO - Go to 3.4.1.

3.4.1 How many samples do you need?

See Section 6.2.1 of the main report for answer. After determining the number of samples go to 3.4.2.

3.4.2 What is the distribution of the sample sites?

At this point you may want to delay the final decision on which statistical design you will use until you talk to a statistician. Basically you are being asked to select between the simple random design (Section 6.2 of the main report), the stratified random design (Section 6.3) and the systematic design (Section 6.4). The simple random design locates the points at any location on the study area. The stratified design does the same thing but it places a portion in each stratum. The systematic covers the entire area with a grid or with a series of line transects.

Based upon the objectives of the study and the decisions made in the paragraphs above determine the sample distribution pattern that you will use. If you are more interested in a uniform distribution use a grid design. If you are interested in a totally unbiased sample use some form of the simple random design.

There are a number of variants on each of the designs that can be provided by your statistician. One that may be used is what is called a nested simple random design. This is used when simple random designs are used to locate the sample site but the location of the samples within that site are fixed by geometry or some other limit to randomness. This is the case with soils where the location on the surface is randomly determined but the below surface samples are fixed by the surface location. This is not a problem for the field scientist but it is for the statistician if this fact is not taken into consideration in the analysis.

Record the information on the design chosen and include information on the number of replicates that the laboratory wants along with the number of duplicates your statistician tells you that he needs for the data analysis in 3.6 below. Go to 3.4.3.

3.4.3 What is the frequency of sampling?

The frequency of sampling depends upon the nature of the study. Any study attempting to determine seasonal patterns or patterns in variables that are likely to be associated with the seasons should be designed in such a manner that samples are collected at least once each quarter. Long term trends in ambient soil pollutant levels will only need to be sampled on an annual basis; but, it must be sampled at the same time each year.

The more detailed the information needed the more often the soils should be sampled. Those situations where the pollution is changing rapidly will require more frequent sampling than those situations where the pattern is stable or only slowly changing.

Select the frequency of sampling that meets the circumstances of your study. Record the choice made and go to 3.5.

3.5 Select a Sampling Method.

Read Section 7.0 of the main report. From this select a sampling method that meets your specific need. You must combine the type of pollution with the depth of contamination to arrive at the tools that you will use. Once this is done you select the method by which the samples will be collected. This process combines the tools with the statistical design and the laboratory procedures. If the pollution is deep, use one of the methods in Section 7.3 dealing with underground pollution. If it is shallow then one of the methods in 7.2 or 7.1 will be useful. The decision to use compositing will be determined by the data needs of the statistician in order to answer the questions combined with economics. Use Equations 6.2.1, 6.2.2 and 6.2.3 to assist in making the decisions. Go to 3.6.

3.6 Data Analysis.

The data analysis will be determined by the design that was chosen in 3.5. Section 8 of the main body gives a discussion of the approaches. You need to determine the approach that best fits your needs. Approach the problem from the point of view of the user of the data. The statistician will probably use one of the statistical computer program packages to analyze the data.

The outline that is now filled in should provide you with a fairly detailed set of information that you can incorporate into the final protocol.

APPENDIX B

SAMPLING PROTOCOL FOR SURFACE SOILS

SECTION B-1

OVERVIEW

The purpose of this protocol is to provide an example of a protocol for use in sampling surface soils. The procedures outlined here are not intended to be the final solution to a soil sampling protocol but are only a guide for use by those scientists undertaking to develop a protocol for some particular use. The protocol can be used in cases where the chemicals of interest are expected to be located on the soil surface as would be the case in an accident or chemical spill.

The procedure used in this protocol is simple and fast. Ten 15 cm. long subsamples are to be taken from the surface layer at each sampling station with an Oakfield type tube sampler. The ten subsamples will be composited and sampled for analyses. A triangular grid is used to insure complete coverage of the area.

SECTION B-2

SAMPLING DESIGN

A triangular patterned systematic sampling design will be used to collect the samples. Ten subsamples will be collected at each of the intersections of the grid lines. At ten percent of the stations, three sets of samples will be taken for analysis. The choice of ten percent was purely arbitrary and has no particular significance. The number can be better selected when input from the analytical laboratories is available. The data collected will be analyzed by kriging, if possible. If it is not possible to use kriging, the method of adjacent differences will be used to determine the variation within the grid.

2.1 Minimum number of samples

Calculate the minimum number of samples to collect using equation B-2.1. In order to use the equation you must have the following pieces of information:

- Significance level you desire to use
- Coefficient of variation in %
- Percentage error you will allow
- t-tables

The values are entered into the equation. It is not possible to obtain the t-value without knowing n. The approach used is to assume a value for n then look in the t-table for the t-statistic at (n-1) degrees of freedom. Calculate a new value for n. This value is then used to pick a new t-value. The process is repeated until the starting and ending values for n are the same. Use the next highest integer if fractional values of n result from the calculations.

2.1.1 The Equation

$$n = (CV)^2 t^2 / p^2 \qquad \text{(B-2.1)}$$

Where n = the number of samples; CV = coefficient of variation in %; p = the percentage of error you will allow; and t is the two-tailed t-value obtained from t-tables for the significance level at (n-1) degrees of freedom.

If values are not known for the variables in the equation, a first approximation can be used. Use 95% significance level, 65% coefficient of variation and 20 % for the error you will allow in the numbers. These latter two values may seem high. They are intentionally chosen on the high side to provide a margin for error. The 65% CV is not out of line with data from a number of soils studies. The 20% margin of error is also not unreasonable for soils work. Using these values gives an n of 43. The closest triangular spacing to the n of 43 would be 37 samples. This number is determined from Table B-2.1. This number would reduce the significance level to approximately 93%. This value was determined by inserting the n of 37 into equation B-2.1 and solving for the t value. This calculated t value was then compared with the t-table values in order to determine the level of significance.

If it is essential to have 95% confidence or better then the next highest triangular grid would be chosen. This grid yields 61 sample points. If this number is used, the significance level would be approximately 98%.

2.2 Grid Layout

The triangular grid consists of equilateral triangles laid out to form a hexagon. The design can be elongated if desired but the hexagon is the easiest to determine the number of samples needed to meet a particular area configuration. The hexagon has an equal number of line segments on each side. This fact can be used to assist in determining the number of samples in the array. Figure B-2.1 shows the grid and its coordinates. The number of sample points for each grid size is shown in Table B-2.1.

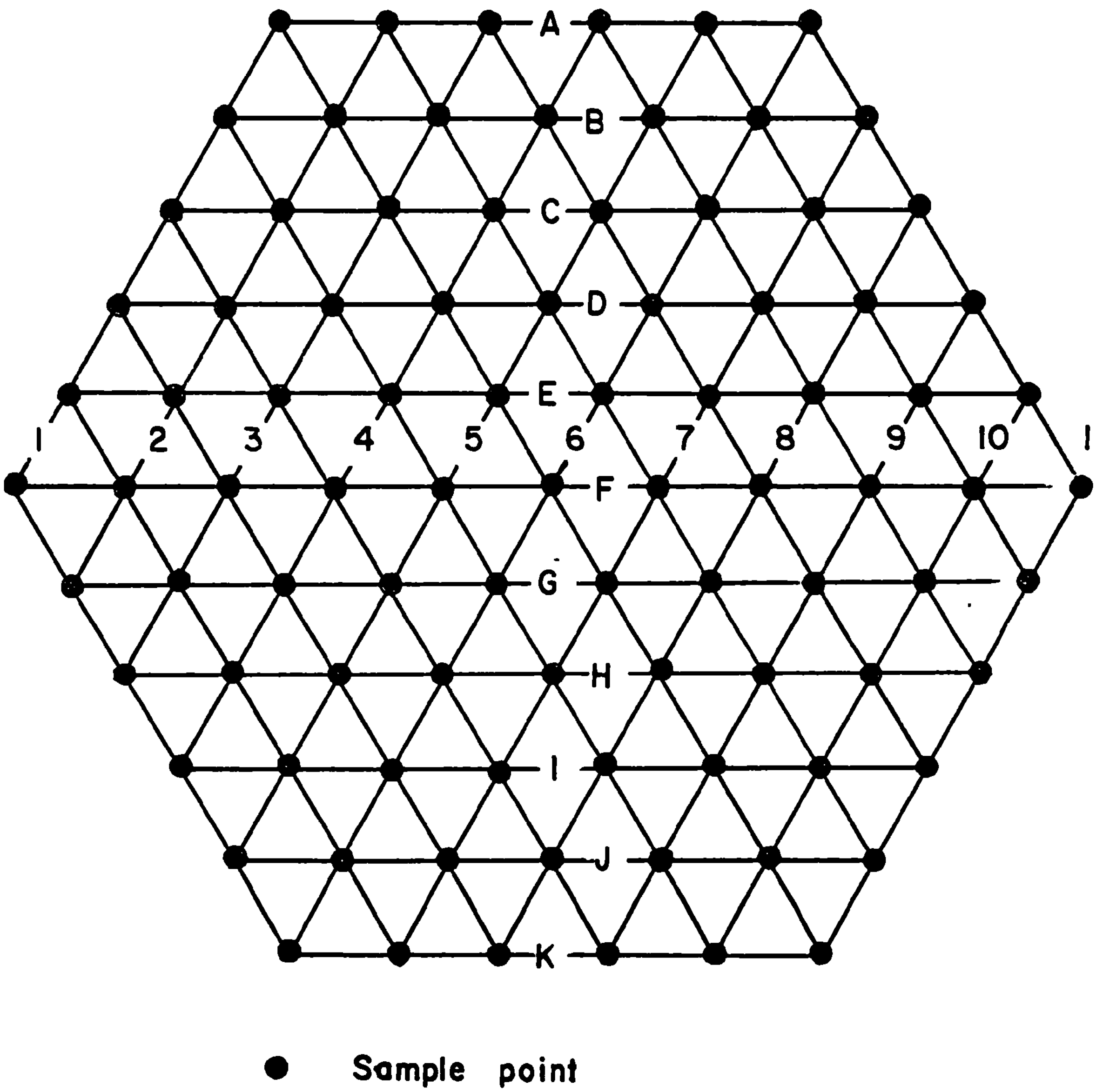

Figure B-2.1. Triangular grid design showing coordinate numbering.

Table B-2.1. Number of Sampling Points in Triangular Grids.

Number of Points on a Side	Number of Samples
2	7
3	19
4	37
5	61
6	91

2.2.1 Selection of a Starting Point

Select a starting point for locating the lower left corner of the hexagon by using either the UTM coordinates on a map or by using a grid overlay. The procedures outlined in Section 6.2.2 of the main body of the report can be used to locate the starting point for the grid. After the starting point is selected, the triangular grid is oriented over the area to be studied in such a manner that one of the nodes is superimposed on the starting point selected above. Samples will be collected at each of the nodes where the lines cross.

2.2.2 Grid Size

The length of the sides of the equilateral triangles is to be determined by examination of the expected pollution area. This may be determined by a pilot study or by examination of aerial photographs and by site visits. The grid should extend beyond the boundaries of the polluted area if it is possible to make a determination of the approximate location of the boundaries before the sampling is done. A set of grids drawn on plastic with different map scales can greatly facilitate the use of this design. The location of the starting point determined in Section 2.2.1 determines where the sampling points will fall but the orientation and the size of the hexagon may have to be determined by trial and error using the plastic overlays. Move the overlay until the desired grid alignment is obtained keeping in mind that one of the nodes must be located on the starting point.

The size of the hexagon can be determined by an alternate procedure. First, estimate the area of the study by placing a square over the known or suspected contamination. This area (A) is entered into equation B-2.2 to determine the length (l) of the sides of the hexagon formed by the triangular grid.

$$A = 2.598\ (l^2) \qquad \text{(B-2.2)}$$

Use of this equation will allow the edges of the resulting hexagon to extend beyond the boundaries of the contamination because the border of the hexagon will fall outside of much of the square used to estimate the area of contamination. Using the size of grid calculated, prepare a plot plan on a map of the area using the starting point as one of the nodes of the grid.

The grid nodes are identified by a number-letter combination as is shown in Figure B-2.1. The east-west rows will be lettered and the diagonal rows will be numbered.

2.2.3 Physical Location of the Grid in the Field

Locate the starting point identified in Section 2.2.1 on the ground using a map or aerial photograph of the site. Locate an east-west line through that point. This line should correspond to one of the lines of the grid that you placed on the map of the site in Sections 2.2.1 and 2.2.2. Measure east along the line the distance (l) calculated with equation B-2.2 or scaled from the map. This is the first sampling location. Continue this process until the appropriate number of points are located along this line. Mark each sample point with a stake. Using a metal tape measure or surveyor's chain, strike an arc from each of the points. Another point is located at the intersection of the arcs. This point is staked. The process is continued until the entire hexagon has been staked out on the ground. The number of sample points located on each side of the hexagon is given in Table B-2.1. Photograph the plot upon completion of staking.

2.3 Compositing of the Samples

At each grid node, take ten individual 15 cm cores with a standard Oakfield tube sampler. These cores are to be placed in a stainless steel pan, then mixed with a stainless steel scoop until the soil appears to be relatively homogeneous. Take small samples from each quarter of the pan. Remix the remaining soil.

Take a second set of subcomposites adding these to the first subcomposite. Continue this process until 250 g of soil is obtained. Dispose of any waste at an appropriate disposal site.

2.4 Replicates

At 10% of the sampling points take triplicate sets of cores (i.e. 3 groups of 10 each); composite these as separate samples -- DO NOT MIX AS ONE COMPOSITE. The same procedure will be used for each composite as was described in Section 2.3 above. These samples will be used to obtain a measure of the variation within each grid. In order to insure uniform coverage with the replicates, the hexagon will be divided into six segments determined by running a line from each corner. This will give six triangles from which to take the triplicate samples. Ten percent of the samples in each small triangle will be randomly selected for replication. Fractional numbers will be rounded to the next higher number.

2.5 Controls and Background Samples

Collect enough samples from an uncontaminated area to equal 20% of the total samples collected in the sampling grid. These samples will be used to establish an environmental baseline and to test the limits of the contamination measured in the grid. (This sets an analytical limit to determine which of the grid samples are contaminated and which are not).

2.5.1 Location of Controls

Locate the control samples upwind from the site in an area of similar soil types and similar land use to the study area. Aerial photographs available in the county SCS office can be used as a starting point for identifying where the best control sites are likely to be found. The candidate areas should be visited and an attempt made to determine the history of the area. You are looking for an area that has as many characteristics in common with the study area as is possible. The ideal control area would be one with no source of pollution but all soil and land use characteristics the same. This is seldom possible so the field team leader must determine the best compromise for the area. Where time permits, a number of pilot study samples could be taken from the candidate control areas.

Care must be exercised when taking these samples to insure that the owners and neighbors understand that the control area is not believed to be contaminated; on the contrary you believe that it is not contaminated. Let them know that you are collecting the samples for a baseline. This precaution avoids a lot of unnecessary anxiety in the community where a spill has occurred. In rural areas, the county extension agent or the soil conservationist can greatly assist in finding cooperative land owners and in reducing the fears of the local community.

SECTION B-3

SAMPLE COLLECTION ON THE TRIANGULAR GRID

Sampling of shallow soils is simple and straightforward. The approach presented here is designed to provide the researcher with the basic tools for collecting the samples.

3.1 Sampling for Non-volatile Chemicals

At each sampling point lay out a circle with a diameter of either 3 meters or 1/4 of the side of the small triangle, whichever is smaller. Beginning at the northernmost point on this circle, take ten, equally spaced, 15 cm soil cores around the circle using the Oakfield tube sampler Model 22G. Place these cores in the mixing pan and follow the procedure outlined in Section 2.3. If ten cores do not provide an adequate sample, increase the sampling density around the circle. The composited samples are to be placed in precleaned, glass containers fitted with precleaned teflon or aluminum foil lid liners. Use the procedures outlined in Section 7.8 of the main body of this report for cleaning the glassware and the tools used.

When rock or large roots are encountered, move the sampler several inches to one side or the other. Select a direction and follow it during the sampling effort. A new, decontaminated sampling tube is to be used at each sampling location. The mixing bowl, stainless steel scoop, spoon etc., are to be decontaminated between each sampling point. See Section 7.8 of the main report for details on this decontamination procedure.

3.2 Sampling for Volatile Chemicals

Volatile chemicals create a special problem. Samples containing volatiles must be collected in a form that will prevent loss from volatilization. Water samplers use what is called a "headspace jar". These do not work well with soil because it is impossible to exclude all air space with the cores.

Compositing speeds up volatilization therefore it cannot be used as a sampling method with these chemicals. The following procedure is used in lieu of the "headspace jars".

3.2.1 Soil Sampler

One of two types of samplers can be used to obtain cores from the soil. The U.S. Army Corps of Engineers tube density sampler drives a 5.1 cm (2 in) diameter tube into the soil to collect a soil sample. The tubes are machined to a predetermined weight and inside diameter. These tubes can be reused and are small enough for handling during surface sampling efforts.

The second sampler is the short 45.7 cm (10 in) split spoon with ring liners. The spoon with liners can be driven into the soil with a sledge hammer or with a sample jack such as the Soiltest Hydraulic Porta-Sampler . The split spoon retains the soil core inside of a brass liner. A 15 cm long ring liner should be used for the sample. (This same system can be used where a depth profile is desired.) A series of ring liners of desired thickness are included in the split spoon. The soil in each ring is analyzed separately.

3.2.2 Collection of Sample

The tubes, either split spoon or density sampler are forced into the soil to the desired depth. The tubes are extracted and sealed with a precleaned teflon cap, then tightly wrapped with duct tape and coated with a non-contaminating sealing compound. Three samples will be taken at each location.

3.2.3 Transport and Analysis

The sealed soil-tube unit will be shipped to the laboratory in locked ice chests cooled with dry ice. They will be stored in a cooler maintained at 4° C. The tubes will not be opened until time for extraction of the organic chemicals. The tube will be opened and the soil extruded into the extraction vessel or else placed in a tube furnace where the volatile organics can be driven off at the appropriate temperatures. With these relatively short tubes it may be possible to place the soil and the split spoon liner into the tube furnace without having to extrude the soil.

SECTION B-4

RECORDS, SECURITY AND SAFETY

4.1 Records

Maintain a logbook of the operations along with all appropriate site description forms, photographs and maps showing the sampling locations and the collection sites. All samples are to be tagged with NEIC tags and chain-of-custody forms are to be filled out by the sample collector and accompany the samples until the analysis is complete.

4.2 Security

Samples are to be in the physical possession of the sample collector or within his immediate view at all times. If they must be left in a vehicle, the vehicle is to be locked at any time the sample collector is away from the vehicle. The NEIC chain-of-custody procedures are to be rigidly adhered to at all times.

4.3 Safety

The chemicals of concern to the environmental sampler are usually toxic. Some are extremely hazardous. Prior to extensive sampling under this protocol, the field team should determine the nature of the chemicals present and provide adequate safety measures. Gloves should be worn at all times. Make sure that the gloves and safety clothing worn will stand up to the solvents used in decontamination. If volatile chemicals are involved, it may be necessary to wear charcoal filter masks such as the American Optical Organic Vapor Respirator. NEIC and the EPA Safety Officer can advise on the best equipment to use. OSHA and NIOSH can also assist in determining the proper equipment if there are any questions on what should be worn.

4.4 Site Restoration

Experience has shown that often there can be some terrain damage to the study area from vehicle traffic. This damage must be repaired by the sampling crew immediately following the field work. This should be done within a couple of davs of the sample collection in order to avoid problems with the land owners.

SECTION B-5

DATA ANALYSIS

The results of the analysis will be subjected to a series of statistical tests. The following is a suggestion. The final tests should be determined after close consultation with the statistician. Details of the methods are not discussed here because standard reference materials and computer packages such as Stat Pac can be used to conduct the analysis. The steps are listed below.

- Calculate the mean and standard deviation for the control and the polluted area data.
- Use the t-statistic to determine a confidence interval for the background. Use the 95% significance level.
- Determine which samples are above the background level.
- Plot the location and concentrations of the samples on a map.
- Delineate the contaminated area.
- If isopleths are desired, conduct kriging.
- Plot isopleth map.
- Plot error map.
- Calculate the error mean square for differences between adjacent points.

SECTION B-6

STAFFING, EQUIPMENT AND SUPPLIES

The sampling effort outlined in this protocol requires a minimum of two people. A third person can greatly facilitate the work and should be included if possible. Two members lay out the grid and collect the samples. The third member takes the photographs, prepares the site description and map and handles all tagging and record keeping.

6.1 Equipment and Supplies.

- 10 to 12 Oakfield tube samplers, Model 22-g obtained from Soil Test, Inc.
- Borebrush for cleaning.
- 10 to 12 ten-quart stainless steel mixing bowls.
- A U. S. Army Corps of Engineers tube density sampling set with 30 to 40 six-inch sample tubes.
- Safety equipment as specified by safety officer.
- One-quart Mason type canning jars with Teflon liners (order 1.5 times the number of samples. Excess is for breakage and contamination losses.).
- A large supply of heavy-duty plastic trash bags.
- Sample tags.
- Chain-of-custody forms.
- Site description forms.
- Logbook.
- Camera with black-and-white film.
- Stainless steel spatulas.
- Stainless steel scoops.
- Stainless steel tablespoons.
- Caps for density sampling tubes.
- Case of duct tape.
- 100-foot steel tape.
- 2 chain surveyor's tape.
- Tape measure.
- Noncontaminating sealant for volatile sample tubes.
- Supply of survey stakes.
- Compass.
- Maps.

- Plot plan.
- Trowels.
- Shovel.
- Sledge hammer.
- Ice chests with locks.
- Dry ice.
- Communication equipment.
- Large supply of small plastic bags for samples.

APPENDIX C

SAMPLING PROTOCOL FOR CONTAMINANT PLUME

SECTION C-1

OVERVIEW

This protocol can be used for those situations where contaminated groundwater has moved from the point of deposition. The contaminated soils are likely to be located at considerable depth below the surface. The soil scientist is not only interested in the horizontal spatial pattern but also the vertical pattern. Sampling in this situation requires power equipment such as truck mounted augers and coring devices. The cost of sample collection becomes an overriding consideration. Litigation is a definite possibility; therefore, extreme care must be taken to insure the integrity of the samples.

Samples are to be collected with a 60.9 cm (24 in) split spoon sampler operated in conjunction with a 20.3 cm (8 in) auger. This procedure allows discrete samples to be collected from the various strata found below ground level. The nature of the subsoils is such that compositing is not used except in rare cases. The layers of soil provide avenues for migration of the pollutants. These should be sampled individually rather than in a composite. Compositing can be used if the soil is very homogeneous, a rare occasion in most soils. The decision to composite should be made only after preliminary coring has determined that the soils are in fact homogeneous. Sampling in soft soils will require retainers in the end of the split spoon.

A rectangular or square grid is recommended because of the plume. Random samples often fail to reveal the presence of highly contaminated plumes unless a very high sampling intensity is used. The grid overcomes the problems encountered with random samples by covering the entire area of suspected contamination.

SECTION C-2

SAMPLING DESIGN

A rectangular grid pattern is recommended. The long axis of the rectangle should be located along the axis of the plume or suspected plume. Investigation of well logs and consultations with geologists will determine the direction of groundwater flow, thus allowing the plumes general direction to be determined. Samples should also be collected on the upstream side of the source. The author has observed situations when two phases of chemical pollutant would move in opposite directions. Exploratory drilling prior to initiation of the study will provide the necessary information to determine if there is a need for a more complete grid pattern around the total circumference of the source.

A preliminary study is recommended even when there is no reason to suspect that there may be a large area of contamination. This allows the statisticians to develop a complete picture of the data needs before the study is begun. The preliminary study also is a means for the laboratories to prepare to receive the samples and to have a chance to obtain standards and work out potential problems with the analyses before the main load of samples comes into the laboratory.

2.1 Minimum Number of Samples

The large equipment needed for the deep sampling efforts requires considerable cost; therefore, the total number of samples will be controlled by the budget for the study. The resources available for the study therefore must be committed before the study planning is begun in any great detail. The equation presented below (equation 6.2.3 of the main body of the report) can be used to calculate the number of samples.

$$C = C_o + nC_s + nC_a$$

The costs are for the total cost, fixed costs, sampling costs and analytical costs respectively. The equation can be rearranged to yield the number of samples as follows:

$$n = (C - C_o)/(C_s + C_a)$$

Once the number of samples is determined, the calculated n value is entered into equation 6.2.2 which is then solved for the t-value. The rearranged equation appears below.

$$t = \sqrt{(n\ p^2 / CV^2)}$$

The terms are as described in the main body of the report on page 25. The calculated t-value is compared with the t-value in the statistical tables. Interpolate to find the level of significance or probability that the number of samples will provide. The allowable error (or the percentage difference that you desire to detect), (p), can be taken as the confidence interval on a background sample set, or the detection level of the method of analysis, or a known sampling error for the types of situations encountered during plume sampling.

2.2 Grid Layout

The sampling grid should be aligned with the axis of the suspected plume. Where information is not available for determining this, the axis should be elongated with the direction of groundwater flow. The number of grid points should closely fit the number of samples calculated in Section 2.1 above. An attempt should be made to determine the approximate limits of contamination. This can usually be done in conjunction with the hydrogeologists that are usually involved with any study dealing with underground plumes. The rate and extent of migration can be used as a rough estimate of the extent of contamination. The size of the grid cells then will be determined by dividing the total area of suspected contamination by the allowable number of samples.

The coordinates of the grid will be those of any X - Y grid system. An appropriate numbering scheme can be assigned to the grid lines to aid in locating sampling sites and in coding tags and site description forms. The starting point for the grid is located by the same random process that is discussed in Section 6.2.2.

2.3 Control Area

An appropriate control site located in an uncontaminated portion of the area should be obtained for collecting control samples. The soil structure should be as nearly identical to that in the study area as can be found in the immediate vicinity. Care must be taken to insure that there is no communication underground between the control sites and that the control site is not located down wind from any industrial source of the same chemicals. Candidate control areas should be sampled prior to the final selection to insure that there are no unsuspected sources of pollutants likely to cause problems in the analysis and interpretation phases. This preliminary sampling is especially important around major industrial areas where past practices may have contaminated an area.

2.4 Preliminary Study

Many "plume hunt" studies do not have site specific historical data available for use in planning the study. The systematic sampling grid used in this study allows the researcher to work up to the optimum number of samples in stages. This phased approach is recommended if the scheduled completion date will allow the time for these phases to be carried out. Use of equation 6.2.2 can be used to determine the reliability of some percentage of the total number of samples. The particular grid nodes sampled can be randomly selected if there is a desire to sample in a random fashion but a subset of data points can be obtained by using a coarser grid made up of every other grid line or every third grid line.

This use of the grid designed for the main study allows repeated sampling to be done if there is a desire to determine if there are seasonal variations in the data and also allows for kriging analysis to be done on a portion of the samples. Where kriging is done an error map can be generated that will show where additional samples are needed to reach the precision desired.

Where the preliminary study is truly exploratory (i.e., there is no data available other than visual or olfactory evidence of a problem) the use of some form of surrogate analysis is desirable. For example, in some cases total organic halide analysis can be used to locate the plume at a considerably reduced cost when compared to an analysis like the gas chromatographic-mass spectrographic analysis required for the standard Priority Pollutant Analysis series. These surrogate tests can often allow the outer limits of contamination to be found in the field.

SECTION C-3

SAMPLE COLLECTION

Sampling for deep lying plumes of contaminated soil is expensive and at times quite difficult to perform. The approach used here is essentially the ASTM method D1586-67.

3.1 Sampling Equipment

A truck mounted drill rig equipped with A-rod drilling equipment or adaptable to the A-rod connector will be used. Standard 5.1 cm (2 in) split spoon samplers that are 61 cm (24 in) long will be used to collect samples. Samples collected for volatile organic chemical analyses will be contained in brass liners unless stainless steel liners can be obtained. The brass thin walled liners will be used for all samples unless compositing is done for some specific reason. A 20.3 cm (8 in) diameter soil auger will be used to excavate to the sampling depth. All samples will be placed in precleaned one-quart Mason canning jars fitted with a precleaned teflon lid liner. Dry sands, or soft noncohesive soils will require that the split spoon be fitted with a retainer or pocket shoe.

3.2 Non-Volatile Pollutant Sampling

3.2.1 Sample extraction

The split spoon will be attached to the drill rod and forced into the soil to the full depth of the spoon. In difficult soils, a 63.5 kg (140 lb) hammer will be used on the drill rig to drive the spoon into the soil. If the soil is loose, wet or in any way unconsolidated, use a basket retainer in the split spoon.

The spoon is then extracted and turned over to the sampling crew. The drill crew will then attach the auger and drill down to the depth of the first sample's penetration (i.e., to two feet). The soil will be shoveled back from the hole face. The hole will then be cleaned with the auger by increasing the revolutions as the auger is lowered and returned to the surface. A second split spoon will be attached and forced into the bottom of the hole formed by the auger. This sample will be extracted and given to the sampling crew. The process will be continued until the desired depth of sampling is reached. If large rocks are encountered it may be necessary to shift the sample point to a second location. Move the rig one meter north. If this does not work move one meter south of the original hole. The exact location should be noted in the log book and on the core log. The blow counts for the hammer will be recorded on the core log.

3.2.2 Sample Preparation

The split spoons will be opened by the sampling crew. The cores will be carefully split lengthwise with a stainless steel spatula. The color, texture and any unusual features of the core will be noted. Any evidence of chemical contamination will be recorded in the logbook and on the core log sheet. The length of the core should be measured as well as the depth of any textural changes or unusual features present in the core. A standard tape measure will be used for these measurements.

The sample will then be transfered to a precleaned, labeled, glass canning jar fitted with a teflon liner placed next to the sample. An NEIC sample tag will be filled out for each sample. The tag will accompany the sample to the laboratory. The outside of the jar will be cleaned then it will be double bagged with the tag placed in the outer bag.

Any obvious potential routes of migration such as sand lenses, silt layers or old root channels should be sampled separately. These layers are often the first areas to become contaminated and therefore provide an early warning of future problems. Any marked changes in the texture should be sampled as a separate unit. If the number of samples acquired is excessive, every second or third split spoon sample can be taken for analysis.

3.3 Volatile Chemical Sampling

Samples collected for volatile chemicals are often difficult to acquire in a condition suitable for analysis. The less disturbance the greater the chance that the analysis will be meaningful.

Split spoons can be fitted with liners made of brass, stainless steel and in some cases, Teflon. The samples cannot be described on the well log so a considerable amount of interpretive information is lost. The core log will be made from observations of the material removed by the auger. The procedure for taking the sample is the same as that presented above in Section 3.2. The extracted core will be left in the liner and shipped to the laboratory for analysis. The ends of the liner tube will be sealed with a teflon cap, then taped with duct tape and sealed with a non-contaminating sealant. The samples will then be transported to the laboratory in locked ice chests maintained at 4° C with dry ice. The details of sample handling are outlined in the Federal Register (Federal Register. Vol 44:69464. December 3, 1979).

3.4 Security

All samples will be logged, tagged and entered on the standard NEIC chain-of-custody forms. The samples will be maintained either under constant surveillance or locked in a limited access storage area. If it is necessary to place the samples in a vehicle, the vehicle will be kept locked when unattended. The amount of time samples are left unattended in a vehicle should be kept to a minimum. The chain-of-custody form will be signed when samples change hands. The team leader is responsible for the samples and should insure that they are turned over to a responsible party before relinquishing custody of the samples.

3.5 Safety

Underground plume samples often contain highly toxic chemicals. Extreme care should be taken when handling the samples. Follow the EPA safety manual procedures during sampling. It is better to overprotect than to have someone hurt because of a lack of diligence to pursue safety procedures. Protective gloves are a minimum protection for all members of the crew that must handle the soil samples. Crews sampling around sources of volatile organic chemicals should have proper fitting vapor masks available at all times. There are some situations where the safety officer may require the masks to be worn at all times. In areas where the chemical concentrations are known to be low, the masks may be carried and donned only when chemicals are detected by odor or sight. All auger holes are to be refilled with either grout or clean soil.

3.6 Decontamination

Decontamination is a major problem. The augers and split spoons require a pressurized hose and decontamination facility that can handle hazardous wastes. The augers and split spoons are to be pressure washed, scrubbed then rinsed with tap water. If organics are involved, the split spoons are to be rinsed in waste acetone, then methylene chloride. This is followed by a distilled water rinse, a spectrographic grade acetone rinse, then a spectrographic grade methylene chloride rinse. Inorganic chemical sampling can eliminate the solvent rinses. All tools, etc., must be cleaned following the same procedure as the sampling equipment. Save all waste for disposal at a licensed hazardous waste landfill.

SECTION C-4

DATA ANALYSIS

The addition of another variable, depth, further complicates the analysis of the data. Regression techniques often provide the best method for handling this type of sampling data especially in the early stages of an investigation. The preliminary data should be used to determine the optimum number of samples if cost is not a major factor. Kriging can help determine the location of any additional samples that are needed. Careful examination of the preliminary data will often allow the researcher to exclude some samples because of the homogeneity and/or thickness of a particular layer. (This statement is based upon the fact that the more homogeneous the media, the fewer the number of samples that are required to arrive at a conclusion with a predetermined precision.)

The depth variable requires that kriging be conducted in three dimensions. This has not been done although it should be possible. The mathematics would be quite difficult. An alternative would be to krige the data at each layer where samples were collected. (This assumes that sampling was done at a common depth throughout the entire soil mass.) By superimposing one on top of each other the volume of the plume can be observed.

Regression is to be used to determine the variables that are influencing the migration of the pollutants and to make comparisons between layers in the soil and between locations in the study site. The t-test can be used to develop a confidence interval for comparing the contaminated and uncontaminated areas. Use kriging to develop isopleth and error maps for the area.

SECTION C-5
STAFFING, EQUIPMENT AND SUPPLIES

5.1 Staffing

- Drill crew of three men.
- Sampling crew of three men (two sample collectors and one record keeper).
- One runner to acquire supplies and handle decontamination.

5.2 Equipment and Supplies

- Drill rig equipped with A-rod fittings or adaptors and an 8 inch auger.
- 25 two-foot sections of 2 inch split spoon tubing.
- Brass or stainless steel liners.
- Teflon caps for liners.
- Shovels.
- Drums for hauling waste material.
- Grout and clean soil for refilling holes.
- Spatulas.
- One-quart Mason jars (use 1.5 times the number of samples expected.)
- Safety equipment.
- Compass.
- Surveyor's chain.
- Survey stakes.
- Hammer.
- Case of duct tape.
- Non-contaminating sealant.
- All core logs, notebooks, etc.
- Chain-of-custody forms.
- Tags.
- Ample supply of small plastic bags to hold sample bottles.
- Ice chests with pad locks.
- Dry ice.
- Heavy-duty plastic trash bags.
- Communication equipment.