Application of geostatistics to expedited site characterization

Greg Alan Stenback
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Civil Engineering Commons, Environmental Sciences Commons, and the Statistics and Probability Commons

Recommended Citation
Stenback, Greg Alan, "Application of geostatistics to expedited site characterization" (1996). Retrospective Theses and Dissertations. 11419. https://lib.dr.iastate.edu/rtd/11419

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Application of geostatistics to expedited site characterization

by

Greg Alan Stenback

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Geotechnical Engineering)

Major Professors: Bruce H. Kjartanson and John M. Pitt

Iowa State University

Ames, Iowa

1996
This is to certify that the Doctoral dissertation of

Greg Alan Stenback

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College
## TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** vii

**ABSTRACT** viii

1. INTRODUCTION 1
   1.1 Background 1
   1.2 Assess and Develop Application of Geostatistics to ESC 1
   1.3 Overview and Critical Review of ESC Process 2
   1.4 Organization 2

2. LITERATURE REVIEW 3
   2.1 Environmental Site Characterization: Objectives, Approaches and Methodologies 3
      2.1.1 USDOE Expedited Site Characterization (ESC) 3
      2.1.2 Data Quality Objectives Process (DQO) 8
      2.1.3 Observational Method 14
      2.1.4 Streamlined Approach For Environmental Restoration (SAFER) 15
      2.1.5 Field Assessment Screening Team (FAST) 19
      2.1.6 M^ Expedited Site Characterization 20
      2.1.7 Superfund Accelerated Cleanup Model (SACM) 21
   2.2 Similarities and Differences: ESC, DQO, SAFER, FAST, M^3, SACM 22
   2.3 Ames Laboratory ESC: Current State of Practice 25
      2.3.1 Ames Laboratory ESC Overview 25
      2.3.2 Ames Laboratory ESC Phase 1 31
      2.3.3 Phase 1 Subsurface Characterization Technologies 32
      2.3.4 Ames Laboratory ESC Phase 2 33
      2.3.5 Phase 2 Subsurface Characterization Technologies 34
      2.3.6 Ames Laboratory ESC Concluding Remarks 36
   2.4 Statistical Techniques for Site Characterization 37
      2.4.1 Classical Sequential Analysis 38
      2.4.2 Kriging and Classical Geostatistics 39
      2.4.3 Bayesian Methods and Decision Analysis 42
      2.4.4 Adaptive Site Selection and Stopping Rules 46
      2.4.5 Evaluating Contaminant Concentrations in the Presence of Non-Detect Data 50
      2.4.6 Concluding Remarks on Statistical Methods for Site Characterization 52

3. ASSESSMENT AND DEVELOPMENT OF GEOSTATISTICAL TOOLS FOR ESC; FORMER MANUFACTURED GAS PLANT SITE 56
   3.1 The Decision Making Process 56
   3.2 Detailed Case Analysis 57
      3.2.1 FMGP Site in Marshalltown, IA, Phase 1 57
         3.2.1.1 Geology, Phase 1 58
         3.2.1.2 Analytical Screening Data, Phase 1 60
      3.2.2 FMGP Site in Marshalltown, IA, Phase 2 72
         3.2.2.1 Laser Induced Fluorescence and Geoprobe Electrical Conductivity, Phase 2 74
         3.2.2.2 Geology, Phase 2 78
6.3.1 ESC Phase 1 256
   6.3.1.1 Geophysics, Geology, and Hydrogeology 256
   6.3.1.2 Contamination 256
6.3.2 ESC Phase 2 258
   6.3.2.1 Direct Push Technologies 258
   6.3.2.2 Contamination 259
6.3.3 Post Phase 2 Recommendations 261
6.4 Recommendations for Further Work 262
6.5 Concluding Remarks 262

REFERENCES 265
I would like to thank Dr. Bruce Kjartanson for his invaluable help in providing focus and suggestions to improve this thesis, and for his patient, but persistent, push to bring this thesis to its completion. I would like to thank Dr. John Pitt for the inspiration he provided in getting me involved in this work in the first place. I would like to thank Dr. Al Bevolo for providing the opportunity for me to work with the Ames Laboratory Expedited Site Characterization team on several projects, and for the guidance and background he brought to this work. I would like to thank Dr. Noel Cressie for his support and the ample time he has given me to enhance my understanding of the statistical tools utilized in this work. I would like to thank each of my committee members for the suggestions they made, which without question, have greatly improved this thesis. I would like to thank my family for the sacrifices they made to allow this work to be completed. Finally, any shortcomings or errors in this thesis are solely my own.
ABSTRACT

This thesis develops geostatistical methods for use in the USDOE Expedited Site Characterization (ESC) process with focus on application to sites with contaminated soil and groundwater. Statistical methods for on-site sample location selection for geological and environmental sampling, characterizing uncertainty in the geologic and contaminant models, modeling the spatial distribution of contamination in the presence of non-detect data, determining when sufficient data have been collected, and post investigation analyses of geological and environmental data are given. These statistical methods are applied to data from an ESC demonstration at a former manufactured gas plant where soils were contaminated with polynuclear aromatic hydrocarbons and an ESC project at the Savannah River Site in South Carolina to characterize metal, pesticide, and volatile organic contamination in groundwater. A recommended approach to the use of geophysical methods and direct push technologies, in conjunction with geostatistical methods, for the characterization of the geologic environment and contaminant spatial distribution is developed. A summary of the USEPA Data Quality Objectives process, the USDOE Streamlined Approach for Environmental Restoration, and the USEPA Site Accelerated Cleanup Model is given, with focus on how ESC fits into the environmental restoration process.
1. INTRODUCTION

1.1 Background

The primary intent of this thesis is to assess and develop geostatistical tools that may be incorporated into the U.S. Department of Energy's (USDOE) Ames Laboratory Expedited Site Characterization (ESC) process. A secondary purpose is to provide a comprehensive review and summary of the current ESC process and to summarize and compare several similar hazardous waste site characterization technologies.

1.2 Assess and Develop Application of Geostatistics to ESC

The major thrust of the work presented here is an assessment and development of statistical methods for their potential incorporation into the ESC process. In particular, the use of statistical tools to aid the on-site decision-making process for placement of subsequent samples on the basis of currently available information, when to stop sampling in light of the expected benefit and reduced uncertainty or increased information of additional sampling, and estimation of which portions of a site may need to be remediated are addressed. Incorporated into this is a quantitative assessment of uncertainties associated with both the spatial distribution and magnitude of contaminant concentrations and geologic uncertainties associated with contacts between geologic strata, continuity of geologic media, and properties of those media which will impact remedial design and decision making.

Due to the non-random manner in which samples are selected in the USDOE ESC process, I have avoided the hypothesis testing framework discussed in the U.S. Environmental Protection Agency (EPA) Data Quality Objectives (DQO) process. However, consistent with the DQO process, I have incorporated some probabilistic methods to address relevant uncertainties. This has been accomplished largely through the use of statistical spatial analysis.
1.3 Overview and Critical Review of ESC Process

I provide a comprehensive review and summary of the USDOE Ames Laboratory Expedited Site Characterization (ESC) process. This is supplemented with a review of, and comparison to, two similar site characterization methods: M³ ESC and Field Assessment Screening Team (FAST) methodology. The USDOE Streamlined Approach for Environmental Restoration (SAFER) process, U.S. Environmental Protection Agency (EPA) Data Quality Objectives (DQO), and EPA Superfund Accelerated Cleanup Model (SACM) processes are also reviewed as they all impact the USDOE ESC process. A separate review of the observational method developed by Karl Terzaghi and "formalized" by Ralph Peck (1969) for geotechnical applications is given as it is an inherent part of the SAFER process.

1.4 Organization

This work is organized into six chapters. The first chapter is this introduction.

Chapter 2 is a literature review of ESC, DQO, the observational method, SAFER, FAST, M³, and SACM, with a comparison of their similarities and differences. This is followed by a detailed review of the current Ames Laboratory ESC process. Lastly, Chapter 2 contains a brief review of statistical methods for site characterization as they pertain to the major thrust of this document.

Chapters 3 and 4 provide an assessment and development of statistical tools for the ESC process. This includes application of these tools to data from the Ames Laboratory ESC demonstration at Marshalltown, Iowa in 1994 and the Ames Laboratory ESC project at the D-area oil seepage basin at the USDOE Savannah River site in South Carolina during 1995.

Chapter 5 gives an assessment of site characterization methodologies, including geophysical and direct push technologies, used in the ESC process.

Chapter 6 provides recommendations for the use of geostatistical, geophysical, and direct push technology methods in the ESC process.
2. LITERATURE REVIEW

2.1 Environmental Site Characterization: Objectives, Approaches and Methodologies

A traditional environmental site characterization typically involves use of drilling equipment for the installation of many randomly located monitoring wells and collection of soil and ground-water samples. This is followed by an off-site contract laboratory program (CLP) chemical analysis of soil and water samples with subsequent data analysis attempting to define the type, spatial, and temporal extent of contamination. After a field mobilization, data collection, and data analysis, a typical outcome may be that sufficient uncertainty exists to warrant establishing additional sample points for collection and analysis of more environmental samples. This process may be repeated several times with successive field mobilizations until sufficient data are obtained to perform a risk assessment and feasibility study. Such investigations may degenerate into "plume chases" wherein the goal shifts away from the determination of exposure pathways to locating the leading edge of the plume (Sara, 1994, pages 10-18 to 10-20). The traditional methods may also include the use of some geophysical techniques such as ground penetrating radar, electromagnetic conductivity, magnetometric surveys, gravity, and radiological surveys. Such traditional site characterization approaches may be costly and slow due to the multiple field mobilizations and interim data analysis periods, a large amount of non-detect data from poorly placed monitoring wells, and shipping and analysis costs associated with CLP laboratory sample analysis. In an effort to reduce the site characterization time and costs, alternative site characterization approaches, as discussed in the following sections, have been developed.

2.1.1 USDOE Expedited Site Characterization (ESC)

The Expedited Site Characterization (ESC) process was developed and tested by J. Burton at the U.S. Department of Energy's (USDOE) Argonne National Laboratory and is now being applied by several USDOE organizations including Argonne National Laboratory, Ames Laboratory, and Los
Alamos National Laboratory (Purdy et al., 1995). Site characterizations are conducted where hazardous wastes have been, may have been, or might be, released and where there is potential for contamination to be released into the environment impacting the health of humans or ecosystems. The product of a site characterization is a conceptual model describing the geology, hydrogeology, contaminant source and spatial distribution, contaminated media, transport pathways and mechanisms, exposure routes, and potential or actual receptors with sufficient accuracy and precision to be able to perform a reliable risk assessment. The objective of the site characterization is to reduce the uncertainty of the conceptual model sufficiently so that the decision maker(s) can make a correct decision regarding appropriate future land use and the need for remedial action.

The ESC field investigation is typically composed of two phases, called Phase 1 and Phase 2. Information regarding the land use history from sources such as public records and personal interviews with current and former site workers, local residents, EPA, state, county and city officials, geologic and hydrologic reports, local well data, prior site investigations, soil surveys, and aerial photographs are all used to develop a preliminary site conceptual model (Purdy et al., 1995).

The ESC process is carried out by a team of professionals with expertise in the areas of civil engineering, geology, hydrogeology, geochemistry, geophysics, computer science, statistics, regulations, health and safety, quality assurance and quality control, and analytical chemistry. The team composition is specific to the site and may include disciplines other than those mentioned above (Purdy et al., 1995). The team members must work together in data analysis, interpretation, and integration, and to insure that ambiguities arising from conflicting data are resolved. The Ames Laboratory team has found it advantageous to incorporate appropriate regulators into the process from its early stages to insure that applicable regulatory requirements are met without unnecessary rework and so that the regulators concerns can be addressed during the site characterization process.

Figure 2.1 shows a simple flowchart of the ESC process based on the flowchart presented in the USDOE ESC course (Purdy et al., 1995). The ESC process starts when a client discovers a
problem at a site that is of sufficient magnitude to initiate the ESC process. Prior to the initiation of phase 1 work, the ESC provider must be certain that the dynamic work plan is acceptable to the client and appropriate regulatory agencies. To the extent possible, data collected during both phase 1 and phase 2 is reduced, interpreted, and integrated in the field to refine the conceptual model and plan subsequent field activities and data sampling locations. Both the phase 1 and phase 2 field investigations take about two to four weeks to complete. Ideally, after phase 1 and prior to phase 2, the ESC team further analyzes the phase 1 data if necessary, refines the conceptual model, issues a phase 1 report, and develops a phase 2 work plan. The time between phase 1 and phase 2 is generally about four to six weeks, depending largely on the time requirements for analysis of off-site analytical work and reduction of geophysical data.

As described by Purdy et al. (1995), traditional characterization methodology has focused on finding the contaminants and mapping the extent of the contaminants in the groundwater on the basis of water samples taken from monitoring wells. Monitoring well placement may follow either a gridded or random pattern over the entire site. Once contamination is detected, it is assumed to flow in the down-gradient direction of the aquifer. The characterization tasks are divided up into
independent activities done by separate contractors and are conducted in phases each taking a few months to a year to accomplish. Well location, sampling protocols, screen lengths, and sometimes screen depths are set in a rigid workplan (Purdy et al., 1995). A short term budget is fit to each assessment activity generally resulting in higher long term costs, while breaking the characterization into a series of separate tasks resulting in a lengthy time required to complete the site assessment (USDOE, 1994).

In contrast to the so-called "traditional" site characterization approach, the ESC process seeks to understand the site geology and hydrogeology so that a proper assessment of contaminant fate and migration pathways can be made. Analytical measurements are made in the field with a real understanding of the problem so that the contaminant source area(s) and extent of the contaminant plume(s) can be determined. With this information, ESC can be used to most effectively place and utilize drilled wells by determining the minimum number and optimal placement of monitoring wells and screen depths and lengths (Wyrick from Purdy et al., 1995).

ESC is a scientific methodology (Purdy et al., 1995) based on the assumption that all features of the conceptual model that are important to the decision maker must be tested, and that each critical hypothesis be supported by confirmatory measurements. As new data become available, the conceptual model is either supported or modified to fit the facts. The investigator is forced to honor the data until it is clear that the model agrees with the measured data. The characterization plan is dynamic and must be flexible to change during the investigation process.

Judgment based, or biased, environmental sample locations are selected on the basis of conceptual contaminant pathways. These samples are most likely to provide a direct test of the conceptual model. This strategy should result in a more cost effective, i.e. fewer total samples, site characterization than the "hit and miss" approach of a random or grid sampling plan which are subject to the risk of sample placement in areas that provide no useful information. The samples are analyzed on-site and analytical results are available within sample handling and analysis time (usually several hours or by the next day) and are used to either support or indicate a deviation from
the conceptual model. In any event, the analysis results are used to aid the decision-making process for selection of future sample locations during the investigation process. The natural constraints that geology and hydrology place on the system are used to reduce the search area (Purdy et al., 1995).

The USDOE Argonne National Laboratory's research on environmental site characterization aimed to accomplish four objectives: decrease cost, save time, be less intrusive, and be accepted by the regulatory authorities (Burton et al., 1993). Characterization only needs to continue until the conceptual model is sufficiently certain to satisfy the remedial action decision authority. Once a desired certainty threshold or minimal uncertainty level is achieved, continued characterization adds to the total cost without adding any meaningful information to the conceptual site model. The cost of ESC field sampling has been reported to be from one tenth to one fifth of the cost of traditional site characterization while the time required for this activity is about one thirtieth the time required for traditional methods (Burton et al., 1993).

Figure 2.2 (modified from Purdy et al., 1995) illustrates the total cost impact of site characterization and remediation as a function of uncertainty in the conceptual model. Geologic and hydrogeologic systems are often complex due to geologic heterogeneity and hydrogeologic variations with time giving rise to complicated patterns of contaminant distribution and transport in the subsurface. Failure to adequately characterize the site geologic and hydrogeologic conditions and pinpoint contaminant sources leaves large uncertainties in the conceptual site model which will (almost certainly) lead to remediation failure. Additional characterization will reduce uncertainty, but if adequate characterization is not achieved, remediation overdesign to compensate for the conceptual model uncertainty will add unnecessary costs. Further characterization will reduce uncertainty to the point that an optimal remediation design can be achieved, as shown by the region of low total cost in Figure 2.2. Continued characterization beyond the region where an optimal remedial design can be achieved will further reduce uncertainty without adding information which is useful to the remedial design while continuing to accumulate characterization costs.
2.1.2 Data Quality Objectives Process (DQO)

The Data Quality Objectives (DQO) process was designed by the U.S. Environmental Protection Agency (EPA) to provide a series of planning steps utilizing the scientific method to ensure that the type, quality and quantity of environmental data used in decision making are appropriate for the intended application (EPA, 1993). The DQO process was developed for use at Superfund sites to minimize unnecessary expenditures related to data collection while ensuring that adequate data are collected to make the required site specific remediation decisions. The DQO process is a seven step, possibly iterative, process providing the site manager with a criteria to determine when the data are sufficient for site decisions, thereby providing a stopping rule to indicate when sufficient data have been collected. The process incorporates statistical design and hypothesis testing to manage uncertainties which are inherent in the site characterization process.

The seven steps of the DQO process are (EPA, 1993, page 2, Figure 1):

1. **State the Problem** - summarize the contamination problem that will require new environmental data, and identify the resources available to resolve the problem.
2. **Identify the Decision** - identify the decision that requires new environmental data to address the contamination problem.

3. **Identify Inputs to the Decision** - identify the information needed to support the decision and specify which inputs require new environmental measurements.

4. **Define the Study Boundaries** - specify the spatial and temporal aspects of the environmental media that the data must represent to support the decision.

5. **Develop a Decision Rule** - develop a logical "if ... then ..." statement that defines the conditions that would cause the decision maker to choose among alternative actions.

6. **Specify Limits on Decision Errors** - specify the decision maker's acceptable limits on decision error which are used to establish performance goals for limiting uncertainty in the data.

7. **Optimize the Design for Obtaining Data** - identify the most resource-effective sampling and analysis design for generating data that are expected to satisfy the DQOs.

In addition to defining the spatial and temporal boundaries of the study, step 4: Defining the Study Boundaries also provides guidance on partitioning the site into media of concern and stratifications of the site which may be useful in the design of data collection or decision making strategies. These may include physical constraints, soil or bedrock stratigraphy, time frame to which the study will apply, exposure units such as the size of a potential future home built on the remediated site, or remediation units such as the amount of soil remediated by, say, a vapor extraction design or the amount of soil removed by one pass of a bulldozer.

The expected benefits of use of the DQO process include improved sampling and analysis design, reduced costs, quicker site characterization, and better decision making with some control over managing decision errors. The scientific method is utilized, whereby a hypothesis is formed followed by an experimental or data collection stage where data is collected to either support or refute the hypothesis. The legal defensibility of site decisions is enhanced by providing a complete record of the decision process, criteria, and conclusions (EPA, 1993).
A site decision maker is identified. A scoping team with representative members knowledgeable about project areas including quality assurance, sampling, analytical chemistry, modeling, risk assessment, biology, ecology, remediation, administration and management, and statistics is created. Each member of the scoping team will participate in all steps of the DQO process. The team develops a site conceptual model that describes the known and suspected sources of contamination, potential or known migration and exposure pathways, and potential or known human and environmental receptors. Development of the site conceptual model should include a compilation, organization, and interpretation of historical data and development of a written summary supported by maps and cross-sections describing the site contaminant sources, release mechanisms, migration routes, and potential receptors. This should provide a foundation to identify data gaps and focus on where potentially unacceptable contamination may or may not exist, and whether additional data need to be collected to support a decision regarding remediation or future use of the site. Data gaps are described by the USDOE (1993) as any insufficiency in information to support the baseline risk assessment, identification of applicable or relevant and appropriate requirements (ARARs), and development and evaluation of remedial alternatives. The baseline risk assessment is based on calculated estimates of carcinogenic risk and non carcinogenic hazard potential which are compared to regulatory acceptable risk and hazard potential criteria, and are used to determine if the site poses an unacceptable risk to human health or the environment (USDOE, 1996).

Four basic decisions are identified in the DQO process (EPA, 1993). First, an early assessment decision to determine whether or not the site poses a potential threat to human health or the environment. Second, an advanced assessment decision, Phase I, to determine whether the concentration of contaminants of potential concern exceed the preliminary remediation goals or other applicable or relevant and appropriate requirements (ARARs). Third, an advanced assessment decision, Phase II, to determine the extent of contamination, as the volume and media of material
that exceeds contaminant concentration action levels. Fourth, a cleanup attainment decision to
determine whether the final remediation levels or removal action levels have been achieved.

The site manager and decision makers would like to know the true state of various features
at the site. Because the site will be sampled at only a finite set of locations, however, it is not
possible to know the true state of the various site features of interest. Accordingly, some inference
from the sample to the site must be made. This is complicated by the presence of a variety of
potential errors, including improper sampling procedures, improper sample handling and storage,
analytical error, data entry errors, and so forth, which increase data variability and may introduce
bias into the data. To minimize such errors, the DQO process calls for a pre-specification of quality
assurance and quality control measures to be implemented and documented as the site investigation
proceeds.

Decision errors may be controlled in a statistical hypothesis testing framework. Figure 2.3
illustrates an example decision scenario under the assumption that the site has an average
contaminant concentration less than the action level until it is shown statistically that the average
contaminant concentration exceeds the action level. A decision whether the site has contamination
which exceeds the action level or not must be made. A false negative decision is made if the site
manager concludes the site is "clean" when in fact it is contaminated beyond the action level. A
false positive conclusion is reached when the site manager decides the site is contaminated beyond
the action level when in fact the average contaminant concentration is less than the action level.
Other specified conditions on which decisions may be based include maximum allowable
concentrations or a maximum proportion of the site which exceeds the action level. In general, false
positive error rates and false negative error rates with conditions under which they apply are pre
specified and should be chosen with some consideration of the costs and health or ecological risks
associated with either incorrect decision.

A data collection strategy needs to be developed. This might be a random sampling pattern,
a systematic grid sample, or a stratified sampling plan. For some field investigations, a non-
probabilistic (judgmental) sampling plan is acceptable (EPA, 1993, page 38), however, the EPA (1993, page 1-27) states that if any statistical conclusions are required, then judgmental sampling is not applicable. While this is true for the specific statistical tests described in EPA (1989, 1992a, and 1993), we shall see that it is not true in general. A statistical design relating the measured values to a true value should be developed and a statistical method for analyzing the resulting data should be developed. All statistical assumptions should be documented. If a statistical hypothesis test is planned, an initial assessment of the expected variability might be made so that a sample size which will satisfy the false positive and false negative error rates can be determined. If the required sample size is too large, the team may decide to increase one or both the decision error rates to reduce the number of samples and associated sampling and analytical costs.

Statistical methods for analysis of analytical data based on probabilistic (random) sampling, normal distribution theory, and the assumption that the site is contaminated (the null hypothesis) until
demonstrated to be clean (the alternative hypothesis) are given in EPA (1989). Methods to search for hot spots, generally based on circular or elliptically shaped regions of high contaminant concentrations, along with a brief introduction to kriging are given in EPA (1989). Also discussed in the EPA (1989) document is a sequential data analysis procedure whereby samples are collected from random locations within the site and a statistical comparison is made following each datum collected resulting in a decision to favor the null hypothesis, favor the alternative hypothesis, or conclude that insufficient evidence exists to support either hypothesis and that another sample should be taken. Selecting sample locations randomly is essential to insure that the data are representative of the population of interest and are serially uncorrelated so that the data may be treated as independent and identically distributed. This sample collection process continues until sufficient evidence to support either the null or alternative hypothesis is obtained.

Statistical methods for analysis of analytical data based on probabilistic sampling using nonparametric theory and the assumption that the site is clean until proven contaminated are given in EPA (1992a). Assuming the site is clean until proven contaminated results in a different statistical test than assuming the site is contaminated until proven clean. Accordingly, it is possible for one data set to yield two different conclusions (the site is clean or the site is contaminated) depending on which assumption the statistical test is based upon. All of these statistical tests may be designed to control both false positive and false negative error rates, however, the definition of a false positive and false negative depends on the initial assumption upon which the statistical test is based.

Choice of the type of data, screening or definitive, should not be made until the DQO process is complete (EPA, 1993, page 42). The distinction between the definitive and screening data types is based on level of quality assurance (QA) and quality control (QC) with a greater level of QA/QC associated with definitive data. Levels of QA/QC for each data type are listed in EPA (1993, pages 42 - 44).
2.1.3 Observational Method

The observational method was formalized for geotechnical engineering applications by Peck (1969). Peck believed that Karl Terzaghi formulated the method in his own approach to problems in applied soil mechanics. The observational method is discussed here because it is an inherent part of the SAFER process which is described in the following section.

The engineer utilizing the observational method gathers sufficient information to establish the general nature of the soil deposits at the project site. Then the most probable conditions and the most undesirable realistic deviations from these conditions are assessed. A working hypothesis based on the most probable conditions is developed. A course of action for all foreseeable significant departures from the most probable conditions is developed. Critical quantities to be observed during the project implementation are selected, and, as the project proceeds, these observations are compared with the working hypothesis and the most unfavorable conditions. Finally, modifications to the design to suit the observed conditions are implemented as the field work progresses.

While the observational method was developed in a geotechnical framework, the methodology has clear application in environmental site investigations in terms of characterizing the site geology and hydrogeology, defining the contaminant types, sources, extent, and transport behavior and migration routes, and in developing and implementing remediation technologies. The full value of the observational method is not realized unless the site manager is thoroughly conversant with the problem, can alter the design or investigation as new information is obtained, and has the authority to act quickly (Peck, 1969).

Morgenstern (1994) discusses use of the observational method in geotechnical environmental applications where uncertainty is potentially large and robustness in engineering design is desirable. A design is considered robust if it is adaptable to a wide range of potential conditions at little cost. Morgenstern notes that in this terminology, the observational method appears to emphasize robustness. D'Appolonia (1990) discusses a "monitored decision" process
whereby the observational method is extended to provide continuity between the investigation, design, project implementation, and operational life of a facility. D'Appolonia notes that a need for design changes often results from uncertain soil conditions which become known over time, or from regulatory changes or societal interventions that pose new facility requirements.

Holm (1993) provides an excellent overview of the remedial investigation/feasibility study (RI/FS) process and use of the observational method throughout the RI/FS and remedial action process, with focus on the observational method during the construction and operation phase of the remedial action. The approach Holm describes is very much in accord with the SAFER method described in the following section. Holm stresses that the remedial investigation activities should overlap with feasibility study activities because it focuses the investigation on those activities that are necessary and sufficient for remedial planning, and avoids collecting data that may be interesting in defining site conditions but are of little or no relevance to potential action. The interaction between the RI and FS activities suggested by Holm is also apparent in the EPA (1988) Guidance for Conducting Remedial Investigations and Feasibility Studies Under CERCLA. Additionally, the feedback loop(s) described by Holm relating field observations from the construction, operation, and monitoring of the remedial design to both the conceptual model upon which the engineering design is based and potential unfavorable conditions are essentially those illustrated in the SAFER process "Implementation" step shown in Figure 2.4. The observational approach is being utilized only if the implementation can be modified based on the results of such monitoring. Holm concludes that application of the observational method requires that the remedial process be viewed as a continuum whereby action is planned and implemented with consideration for the uncertainties involved.

2.1.4 Streamlined Approach For Environmental Restoration (SAFER)

The Streamlined Approach For Environmental Restoration (SAFER) methodology was developed by the U.S. Department of Energy (USDOE) to aid environmental restoration efforts under conditions of significant uncertainty (USDOE, 1993). SAFER was developed primarily by
integrating the EPA DQO process with the observational method. SAFER was developed for the USDOE's environmental restoration effort in a manner which is compatible and compliant with environmental regulations including the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) and the Resource Conservation and Recovery Act (RCRA) (USDOE, 1993).

The SAFER methodology is used to streamline the RI/FS process and to manage changes in the selected remedy (USDOE, 1993). The SAFER process recognizes that inherent uncertainty in the site conditions, remedial technology performance, and regulatory requirements are all factors in the environmental restoration process. As such, SAFER is applicable for identifying and managing uncertainty throughout the Remedial Design/Remedial Action (RD/RA) process. Decisions depend on what is measured, how it is measured, and the identification of sources, and possible quantification, of uncertainty in data analysis. Decision rules are generally "If . . ., then . . ." statements that establish decisions or actions to be taken on the basis of data evaluations. Such rules may involve statistical hypothesis tests, such as those described in Section 2.1.2 above. Development of such decision rules forces a focus on the need for particular types of data and tends to reduce data collection to an essential minimum (USDOE, 1993).

The SAFER process has four essential elements (USDOE, 1993). First, there is a conceptual model which facilitates focus on the problem, potential solutions, and highlights uncertainty. The conceptual model is a dynamic picture of current site understanding, should help make data gaps and uncertainties apparent, and requires periodic refinement as new information is obtained. A data gap is said to be any insufficiency in information to support the baseline risk assessment, identification of ARARs, and development and evaluation of remedial alternatives (USDOE, 1993). The conceptual site model provides a written and pictorial representation of key site conditions including likely contaminant sources and their physical and chemical characteristics, contaminated media, transport pathways, human and environmental and receptors.
The second SAFER element includes planning and designing in a dynamic, possibly iterative manner, based on the output of earlier steps. The third element is to recognize uncertainty and develop techniques to manage it. Uncertainty exists as incomplete knowledge of site conditions, inability to predict remedial technology performance, and changing or unclear regulatory requirements. The fourth element involves the role of stakeholders, including the project team and other persons or groups interested in or affected by a RI/FS conducted at a DOE facility. Stakeholders play instrumental roles in identifying applicable regulatory frameworks, establishing budgets and schedules, determining information requirements, and contributing to the effective management of uncertainty.

Figure 2.4 shows a flowchart of the SAFER process with three stages: planning, assessment and selection, and implementation. This figure is adapted from USDOE (1993, Submodule 7.1, Figure 1), illustrating the SAFER Framework. As shown in Figure 2.4, the SAFER process considers the entire environmental restoration process from discovery of a contamination problem and development of a conceptual model to site characterization and remediation.

Development of contingency plans for reasonable deviations from the most probable conditions is a primary means by which uncertainty is addressed under SAFER. The SAFER process recognizes that there may arise unreasonable deviations which were judged to be unlikely by the stakeholders, and for which a contingency plan was not identified. A monitoring plan is developed to detect deviations between the site conceptual model and reality. Detection of unreasonable deviations may identify data gaps which need to be filled by further data collection.

The remediation decision is formally established in the Record of Decision (ROD) and cannot be modified without altering a legally binding document. If deviations are detected during the Remedial Action stage such that the stakeholders determine that the remediation decision rule requires revision, a modification to the ROD may be required (USDOE, 1993).
Planning
- Start
- Develop/Refine the Conceptual Model
  - Focus on the Problem and Reasonable Deviations
  - Develop/Refine Remedial Objectives for the Problem
  - Develop/Refine Initial Decision Rules

Assessment and Selection
- Collect and Assess Data
  - Site Adequately Characterized?
    - Yes
      - Develop and Evaluate Alternatives
      - Refine Decision Rules (if needed)/Collect Additional Data (if needed)
      - Select Remedy and Contingency Plans (Document Remediation Decision Rule)
    - No
      - Deviation Detected?
        - Yes
          - Refine Remediation and Monitoring Decision Rules
        - No
          - Remediation Decision Rule Satisfied?
            - Yes
              - Remediation Decision Rule Satisfied?
                - Yes
                  - Finished
                - No
            - No

Implementation
- Design Remedy, Contingency Plans and Monitoring Plan
  - Construct and Operate Remedy
    - Deviation Detected?
      - Yes
        - Implement Contingency Plan and Monitor Conditions
      - No
        - Remediation Decision Rule Satisfied?
          - Yes
            - Remediation Decision Rule Satisfied?
              - Yes
                - Finished
              - No
          - No

Figure 2.4. SAFER process flowchart, adapted from USDOE (1993).
2.1.5 Field Assessment Screening Team (FAST)

The Field Assessment Screening Team (FAST) technology is managed by Martin Marietta Energy Systems, Inc., for the Department of Energy (Nickelson, 1995). The FAST concept integrates innovative site characterization technologies to expedite the characterization process for hazardous and/or radioactive waste sites giving optimized characterization quality with reduced field time and lower costs than a traditional site characterization. The system components include use of intrusive sampling based on (but not limited to) direct push technology, use of a field mobile laboratory equipped to analyze for expected site contaminants, computer-assisted design/geographic information system data management with interactive three-dimensional graphics presentation, a global positioning system to determine sample coordinates, and a telecommunication linkup for data exchange with remote sites. The FAST technology is designed to determine the spatial extent of soil and/or groundwater contamination with one field mobilization of the field team.

The direct push technology unit of choice is a Geoprobe with sampling capability. The field laboratory equipment requirements are determined by the analytical needs at the site. Typically, a gas chromatograph perhaps coupled with a mass spectrometer for organic contaminant analysis, X-ray fluorescence or inductively coupled plasma for metal contaminant analysis, and other instruments for pesticide or inorganic contaminant analysis, or water or soil characteristics such as pH, Eh, TOC, or CEC, as needed are utilized. Each analytical system downloads to an on-board computer that records, stores, and transmits the analytical data.

The FAST system is applicable to the DQO process (Nickelson, 1995). With the availability of near real-time data (within several hours), the site manager maintains current data describing the site conditions for optimum on-site decision making (use of today's data to guide tomorrow's work) and may invoke a stopping rule when appropriate. The focus of the FAST approach is principally on the contaminants — to determine what and where the problem is — with soil or rock core samples taken to define the site geology as needed (Nickelson, 1996).
2.1.6 M³ Expedited Site Characterization

Tindall (1995, page 773) describes ESC using the M³ approach as consisting of a three step approach. First, a "massive" sampling effort is carried out whereby samples are collected using a grid over an area of concern and the samples are analyzed on a daily basis using on-site methods resulting in field screening type data. The second step uses a "moderate" sampling effort whereby split samples from about 10 percent of the step 1 samples from suspected "clean" locations, with possibly fewer samples from contaminated locations, together with some boundary samples (which delineate clean from contaminated areas) are analyzed on-site using contract laboratory program (CLP) type methods to develop field quantitation data with an agreed upon level of QC; these are used to provide verification of the step 1 data. The third step is to select a "minimum" number of sample splits, about 20 percent of the step 2 samples from each of the clean, contaminated, and boundary location samples, to be sent to an off-site laboratory whereby CLP-type data are generated and used to provide verification of the step 2 data. No discussion is given in the event that the verification data from steps 2 or 3 fail to confirm the quality of the step 1 or step 2 data, respectively.

Expedited site characterization using the M³ approach seeks to quickly and cost effectively identify and classify potential areas of concern as either "clean" or contaminated. This approach intends to allow potentially responsible parties to save resources by eliminating further action of those areas found to be clean. The large number of non-detect samples often submitted for CLP analyses is reduced while the high degree of uncertainty associated with a lack of comprehensive scoping in traditional site investigations is supposed to be overcome by the massive step 1 sampling effort used in the M³ approach (Tindall, 1995).

All data is loaded into an on-site data management system with three-dimensional visualization capabilities. These data may be viewed on-site or transmitted to remote locations via modem as desired. Data gaps may be seen on a daily basis and accounted for on succeeding days during the site investigation. In this way it is possible that only one field investigation effort may be
needed with remobilization costs either reduced or eliminated resulting in an overall acceleration of the investigation process.

Tindall (1996) describes the M³ approach as a part of the restoration process whereby the DQO process is used to define which investigation activities are carried out. The focus of M³ is on the characterization of the site contaminants. While characterization of the site geology may be identified as necessary in the DQO process, the characterizations of the site geology and contaminant distribution are separate, perhaps concurrent, activities. While discussion of M³ may begin during step three (identify inputs to the decision) of the DQO process, it will generally be developed fully during step seven (optimize the design for obtaining data) of the DQO process.

2.1.7 Superfund Accelerated Cleanup Model (SACM)

The Superfund program operates within a complex set of perceived needs, changing laws, and public concerns and expectations. These factors have created a new focus to improve the Superfund program. The new Superfund paradigm, called the Superfund Accelerated Cleanup Model (SACM), was developed by the EPA to streamline the Superfund process in an effort to speed cleanup of hazardous waste sites in a more cost effective manner (EPA, 1992b). The SACM is designed to be simple and flexible to allow fast risk reduction, realistically achievable cleanup goals in reasonable time frames, and focus on protection of people and the environment. During the last ten to fifteen years large amounts of money have been spent executing separate contracts, designing sampling programs, mobilizing sampling teams, and modifying health and safety plans (EPA, 1992b). The SACM combines whole steps in this redundant process and seeks to eliminate redundancy in data collection programs by taking into consideration data and information generated from previous studies, minimize the number of mobilizations for field work, and eliminate unessential down time between steps in the characterization and remedial process (EPA, 1993).

In addition to integration of site assessment steps to save time and resources, the SACM seeks to develop and expand the use of new cleanup technologies that can quickly reduce
contamination more efficiently with fewer resources (EPA, 1994). The SACM priority is to correct the worst sites first to maximize protection of human health and the environment. The SACM stresses early action to reduce immediate risk, long-term cleanup commitment to environmental restoration, and development of community relations and public involvement throughout the process. From the "Planning" stage of SAFER (see Figure 2.4) it is clear that SAFER incorporates the early action aspect of SACM. A team of Regional EPA staff determines the focus of assessment activities to meet the site data needs. According to EPA (1993), data collection proceeds until there is enough information for the remedial decision team to make a response decision, however, it is not clear what criteria is used to determine when sufficient data have been collected.

2.2 Similarities and Differences: ESC, DQO, SAFER, FAST, M³, SACM

Table 2.1 shows a comparison of ESC, FAST, M³, SAFER, and SACM. Because the DQO process was designed to insure that the correct type, quality, and quantity of data are collected for the intended decision making process and does not endorse a specific site characterization method, it was not included in the comparison shown in Table 2.1. Because the SAFER and SACM methods have a focus which encompasses the entire environmental restoration process and neither emphasize a particular site characterization methodology, several rows under the SAFER and SACM column headings in Table 2.1 are not applicable.

ESC, DQO, SAFER, FAST, M³, and SACM all seek to speed the environmental restoration process and reduce costs. Both SAFER and SACM strive to integrate the characterization and restoration steps to save time and resources. The ESC, FAST, and M³ methodologies seek to carry out the intensive site investigation in one field mobilization thereby eliminating additional mobilization costs often associated with traditional site characterization.

The ESC, DQO, SAFER, FAST, and M³ processes all seek to manage uncertainty. The DQO and SAFER processes deal directly with managing uncertainty as a primary driving force in their methodology. The Ames Laboratory ESC process manages uncertainty directly by
Table 2.1. Comparison of ESC, FAST, M^3, SAFER, and SACM.

<table>
<thead>
<tr>
<th>Aspect of the Process</th>
<th>ESC</th>
<th>FAST</th>
<th>M^3</th>
<th>SAFER</th>
<th>SACM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus of the Process</td>
<td>RI, geology, hydrogeology, &amp; contaminant characterization</td>
<td>RI, contaminant characterization, geology &amp; hydrogeology are secondary</td>
<td>RI, contamination characterization</td>
<td>PA/SI, RI/FS, and RD/RA</td>
<td>PA/SI, RI/FS, and RD/RA</td>
</tr>
<tr>
<td>Manages Uncertainty</td>
<td>By use of DQO</td>
<td>By use of DQO</td>
<td>By massive sampling effort and use of DQO</td>
<td>By use of DQO</td>
<td>Not specifically mentioned</td>
</tr>
<tr>
<td>Applicable to DQO</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Uses dynamic site conceptual model</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Not specifically mentioned</td>
</tr>
<tr>
<td>Number of field mobilizations</td>
<td>Usually two</td>
<td>One or two</td>
<td>One or two</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Uses on-site decision making</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Uses on-site chemical analyses</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Focus on geology &amp; hydrogeology first</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Senior multidisciplinary team in the field</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA - not applicable, while these aspects of the process may be compatible with SAFER and SACM, SAFER and SACM are not specific site characterization methods do not directly address this level of detail.
incorporation of both DQO’s and SAFER. FAST is also applicable to the DQO process while the M\textsuperscript{3} process uses the DQO process in conjunction with a large number of screening samples to overcome the uncertainty associated with lack of scoping in a traditional sampling program.

The ESC, DQO, SAFER, FAST and M\textsuperscript{3} processes stress the development of a dynamic site conceptual model which is updated periodically as new data become available. The ESC, FAST, and M\textsuperscript{3} characterization methods utilize direct push technologies, on-site chemical analyses, and three-dimensional graphics for data visualization. Both ESC and FAST use data results with rapid turnaround time to guide the on-site selection of future sample locations, and M\textsuperscript{3} looks for data gaps to be filled in on succeeding days during the on-site investigation.

The DQO process calls for characterization of the spatial distribution of the site contaminants. Both ESC and FAST seek to determine the spatial distribution of contaminants by utilizing current data to guide future sample locations. ESC stresses development of the geologic and hydrogeologic models in the initial stages of the investigation to enhance the understanding of the contaminant spatial distribution. The M\textsuperscript{3} process seeks to quickly determine which portions of a site are clean and which are contaminated by utilizing a massive sampling effort using on-site screening analytical methods followed by selected verification samples and sampling to fill data gaps. The USDOE ESC process appears to be the most comprehensive site investigation in its effort to develop a site geology and hydrogeology model so that a proper assessment of contaminant transport and fate can be made in conjunction with a characterization of the site contaminant source and spatial distribution and proper data collection for remedial action decisions.

These approaches can generally be viewed as an evolution of the hazardous waste site characterization process with the intent to develop better, faster, safer, and cheaper approaches. The site characterization approach described by EPA (1988) includes a review of existing data, identification of ARARs and DQOs, discovery and quantification of hazardous substances and waste sources, geophysical surveys and geological/hydrological investigations, installation of observation wells or monitoring stations, sampling over a wide area to quantify contaminant distributions and
migration pathways, and integration of all data into an assessment of site characteristics and effects on the environment and human health. Perhaps the greatest recent changes to this basic characterization methodology are the advances in direct push technologies (DPTs), chemical analytical field methods, and data management techniques which have made possible synergistic integration of the characterization phases, analysis of data on-site as it is collected so that on-site decisions can be made to guide subsequent sampling and reduce the total amount of sampling by better placement of samples.

2.3 Ames Laboratory ESC: Current State of Practice

This section provides an overview of the Ames Laboratory ESC process and discusses specific phase 1 and phase 2 activities. A summary is given with discussion of ESC in view of the possible temporal changes in site conditions and the time required to complete documentation of the field work and results so that remedial action decisions can be made in a manner which expedites the entire environmental restoration process.

2.3.1 Ames Laboratory ESC Overview

The goal of the Ames Laboratory ESC is twofold: to demonstrate and field the ESC methodology first developed at Argonne National Laboratory (Burton, 1993) as described in section 2.1.1 above, and to evaluate and promote both innovative technologies and state-of-the-practice technologies for site characterization and monitoring (Bevolo, Kjartanson, and Wonder, 1996). Accordingly, existing tools and methods (i.e. analytical, geophysical, sampling strategy, etc.) used in the ESC process are not set in stone, and as new tools are developed in the future, they may be utilized and compared with existing tools, validated, and eventually incorporated into the ESC process. In this regard, the ESC process itself is evolving in an effort to find the most efficient and cost effective methods to perform adequate site characterization, and seeks to achieve regulatory acceptance of tools which can be shown to provide reliable data. Emphasis is given to the use of a
dynamic work plan, coordination of nonintrusive geophysical and minimally intrusive geologic and
hydrogeologic sampling, on-site collection and analysis of screening and definitive chemical
contaminant data, and on-site data management, integration, and interpretation to aid further
development of the conceptual site model. The Ames Laboratory ESC has been fielded to
characterize subsurface polycyclic aromatic hydrocarbon (PAH) contamination at a Former
Manufactured Gas Plant (FMGP) site in Marshalltown, Iowa, radioactive contamination at the St.
Louis Airport, St. Louis, Missouri, chlorinated organic solvents, pesticides and metals at the D-area
oil seepage basin (DOSB) at the USDOE Savannah River Site in South Carolina, and petroleum
contamination associated with a refinery in Czechowice, Poland.

The Marshalltown, IA, FMGP ESC demonstration was carried out to demonstrate, evaluate
and compare several emerging technologies with state-of-the-practice technologies and was not
intended to fully characterize the site. In fact, this FMGP site had already been characterized and a
Remedial Investigation report written (BVWST, 1992). My role in the FMGP ESC demonstration
consisted primarily of working with the Kansas City District Army Corps of Engineers Site
Characterization and Analysis Penetrometer System (SCAPS) personnel to gain an understanding of
the SCAPS operation and output, and to coordinate the SCAPS activities with respect to the needs
of the ESC demonstration program. During the investigation, I was able to provide some input
regarding subsequent CPT push locations. My post phase 2 field work consisted of work toward
preparation of three reports for Ames Laboratory. The first of these reports describes the SCAPS
operation and data provided within the FMGP ESC project framework (Stenback et al., 1994). The
second report provided a summary of analytical data analysis to compare five different soil PAH
extraction methods which were used on site, a comparison of the use, output, and interpretation of
the SCAPS, Geoprobe soil electrical conductivity probe, and Geoprobe soil sampling results, and
some lessons learned during the phase 2 FMGP demonstration (Stenback and Kjartanson, 1994).
The third report provided a closer examination of the analytical soil PAH extraction method data in
light of some additional information provided by the McLaren/Hart Environmental Engineering
Corporation of Warren, NJ, who were responsible for collecting this data (Stenback and Kjartanson, 1995).

The Savannah River Site (SRS) ESC demonstration was carried out as a true ESC investigation program with the intent to characterize the DOSB site (WSRC, 1995), i.e. this was a unit assessment RI activity. The DQO and SAFER processes were applied to the ESC investigation at the Savannah River site. My role in the SRS ESC project consisted of participation in the development of the site conceptual model prior to the phase 2 work, on-site technical support to assist with interpretation of the CPT and geoprobe conductivity results aimed at defining the site stratigraphy and contaminant distribution, and assist with selection of subsequent CPT, geoprobe, and environmental sample locations.

While one of the goals of the ESC methodology is to reduce site characterization costs, the up-front initial costs may be high due to the intensity of the activities and the variety of expertise and geophysical and analytical tools used. As such, the ESC methodology is most applicable to large or heavily contaminated sites, such as Superfund sites with possible major contamination and potentially costly remediation, and where the landowners and/or other potentially responsible parties can absorb the relatively high cost. Because of the potential high cost, a small property owner or business, such as a local dry cleaner or gasoline station, may not be able to implement a full ESC program. Nevertheless, the ESC methodology and characterization tools are applicable even to such small sites. An accelerated site characterization guide (ASTM, PS 3-95, 1995) has been developed for petroleum releases.

The Ames Laboratory ESC process is typically carried out in two sequential phases: Phase 1 and Phase 2. The objective of phase 1 is to pull together all relevant site specific information, develop and refine the site conceptual model (SCM) with non intrusive and minimally intrusive investigation techniques, develop a site specific list of contaminants of potential concern (COPCs) if not done already, and create the phase 2 work plan. The primary focus of phase 1 is an assessment of the site geology and hydrology aspects of the SCM. There is often a one or two month break
between phase 1 and phase 2 to allow time for the analysis and interpretation of the geophysical and analytical data, identification of the COPCs for the phase 2 work, and refinement of the SCM that will guide the phase 2 efforts. Phase 2 includes an intensive, on-site, characterization to further advance and verify the site geology and hydrogeology models and to define the source and spatial distribution of the COPCs. The outcome of phase 2 is an updated site conceptual model describing the site geology, hydrogeology, contaminant source(s), contaminant spatial distribution and transport pathways, and other information as negotiated during phase 1 that might be useful to the remedial action decision authority. Geology and hydrogeology are emphasized as they play a major role in the fate and transport of contaminants — both factors which play key roles in placement of monitoring wells and potential remediation strategies. Figure 2.5 illustrates the ESC process.

Note that Figure 2.5 shows the Phase 1 and Phase 2 field activities continuing until the associated field objectives have been met. In practice, there are contractual agreements with technology suppliers and budget considerations which come into play. Additionally, determining when sufficient data have been collected to adequately characterize a site is a standing problem and is not always clear while in the field. Our experience has been that data may be generated faster than it can be properly interpreted in the field.

The makeup of the ESC team is variable, depending on the specific needs of the project (Purdy et al., 1995). The team will include a project leader, and may include members with expertise in geology, geohydrology, geochemistry, geophysics, civil engineering, analytical chemistry, computer science, statistics, health and safety, regulations, and quality assurance and quality control. Other personnel, particularly stakeholders such as the landowner, other potentially responsible parties, and regulatory agencies, may be included as needed at a specific site. During phase 1 the team will assess any prior information available and determine a course of action to follow. They will decide which, if any, geophysical, analytical, conventional drilling, or other tools should be fielded to enhance or verify the existing data. All work is carried out by experts in the appropriate field. The team that plans the site work also manages the site work.
ESC Client procures ESC Provider

ESC Technical Manager selects Core and Auxiliary ESC Team Members

ESC Core Team collects, archives, evaluates, analyzes, integrates and interprets existing data – ESC Core Team visits the site

ESC Core Team develops initial site model and identifies questions that must be answered and hypotheses that must be tested

ESC Core Team selects characterization methods for potential site characterization use

ESC Core Team develops Phase 1 work plan considering:
- Regulatory framework
- Site description and history
- Analysis of existing data
- Data Quality Objectives of the technical program
- Schedule
- QA/QC plan
- Health and safety
- Data management plan
- Community relations

Is Phase 1 Work Plan acceptable to client and regulatory authority?

Yes

Continued on next page

No

Revise Phase 1 Work Plan

Figure 2.5. Generalized flowchart of the ESC process.
ESC team collects field data to enhance site geology and hydrogeology conceptual models and to develop a list of contaminants of concern

Daily data reduction, integration, visualization, and interpretation to refine the site conceptual model while in the field

Have Phase 1 field objectives been met?  

Phase 1

Have Phase 1 field objectives been met?  

ESC Core Team writes Phase 1 report and develops the dynamic technical approach and DQO's for Phase 2

Is Phase 1 report and Phase 2 Work Plan acceptable to client, regulatory authorities, and other stakeholders?

Yes

ESC Team collects field data focusing on distribution of subsurface contaminants and geology

Daily data reduction, integration, visualization, and interpretation to refine the site conceptual model in the field

Phase 2

Have Phase 2 field objectives been met?

No

Plan next days measurements

Revise Phase 1 report or Phase 2 Work Plan as necessary

Yes

Write Phase 2 report describing final site model

Is Phase 2 report acceptable to client, regulatory authority, and other stakeholders?

Yes

Are data sufficient for a risk assessment?

No

Collect needed data

Post Phase 2 Work

No

Client uses modeling and risk assessment to choose course of action

Figure 2.5. (continued).
2.3.2 Ames Laboratory ESC Phase 1

Phase 1 begins with some identified reason for a potential threat to human health and the environment due to environmental contaminants. The landowner may enter into a contractual agreement with the ESC technology provider to characterize a site. At this stage the ESC team enters phase 1 and begins to collect and review any relevant site information, including historic land use, regional and local geological and geohydrological reports, soil surveys, aerial photographs of the site, local municipal well data, site soil boring logs, and site soil and/or water sample analyses. Interviews with current or former site workers, property owners, or nearby residents may be performed. The boundaries of the site to be investigated are clearly defined, often according to property lines, but also through negotiation with the site owners, federal, state, or local regulatory agencies, and possibly adjacent property owners. Ames Laboratory has sought to involve appropriate regulatory agencies as soon as possible to insure that their concerns are understood, and incorporated into the work plan right from the start. One or more visits to the site are made to observe the site surface conditions including terrain, vegetation, signs of vegetative stress, areas of discolored soil, surface soil or fill materials, etc. and work obstructions such as buildings, fences, material storage areas, power lines, trees, wetlands, etc.

The outcome of the phase 1 activities includes a site specific list of COPCs, a site geologic and contaminant conceptual model, and a work plan for the phase 2 activities. The site specific list of COPCs is developed on the basis of a pre-established decision rule based on ARARs or risk based concentrations. The concentration levels and spatial distribution of the contaminants of concern will be investigated during phase 2. Any uncertainties in the geologic environment needed to clarify the conceptual model should be identified during phase 1 so that the phase 2 activities can address these issues. The nature and uncertainties of the problem are brought into clear focus so that only data that are needed by the remedial action decision makers are collected during the phase 2 activities.
2.3.3 Phase 1 Subsurface Characterization Technologies

**Geology** A suite of geophysical tools, such as ground penetrating radar, seismic reflection and refraction, borehole geophysics, and electromagnetic sensors, may be considered for the potential use at the site, according to need and applicability to existing site conditions. Geophysical tools have been used to locate buried utilities, buried metal or plastic objects, to develop soil stratigraphy, depth to bedrock, or search for subsurface regions with elevated organic or inorganic concentrations (see for example Bevolo, Kjartanson, and Wonder, 1996). If there is not already one or more continuously logged boreholes drilled through the soil units of interest at the site, then several "calibrating" boreholes will be drilled and logged to develop a general site stratigraphic column. This will be used to correlate and calibrate the subsequent phase 2 cone penetrometer and Geoprobe sensor output and to aid the interpretation of the geophysical data collected during phase 1. While the electromagnetic and borehole geophysical data may be interpreted in a preliminary sense on-site, many of the geophysical tools require about one to four weeks of off-site processing for analysis and interpretation.

**Hydrogeology** If groundwater information is lacking, the team may decide to install piezometers during phase 1 to better understand the groundwater system, including water table depth and fluctuations and groundwater flow direction (possibly including characterization of vertical gradients). Slug and/or bail tests may be performed in piezometers and/or monitoring wells to obtain data to estimate the hydraulic conductivity of the screened soil units.

**Contamination** If volatile or semivolatile contaminants are suspected, a grid of passive and/or active soil gas stations may be installed to estimate the location and extent of contamination. Active soil gas samples may be analyzed on-site, however, passive soil gas analyses are performed off-site thus requiring a waiting period (generally two to four weeks) for the chemical concentration data. Soil and/or water (groundwater and/or surface water) samples from both background (locations on-site or nearby which are not affected by site contaminants) and from potentially contaminated areas of the site may be collected and sent to an off-site CLP laboratory for analysis of a full suite of
potential contaminants (e.g. Target Analyte List, TAL, and Target Compound List, TCL), including metals, total cyanide, pesticides, inorganic compounds, organic compounds, and radioactive wastes, as appropriate. Analytical screening methods, such as chemiluminescence, active and passive soil gas, and immunoassay, have been used to locate plumes of volatile or semivolatile soil vapors. The immunoassay and active soil gas results may be obtained in the field. Soil samples, water samples, and passive soil gas collectors are generally sent to an off-site laboratory for analysis requiring about two to four weeks unless special analytical services yielding a faster turn-around time are requested.

**Data Analysis and Integration**  The principle data analysis and integration software has been EarthVision® from Dynamic Graphics, Inc. The EarthVision software provides the ability to view either the site geologic or contaminant model in a three-dimensional view, with the ability to rotate the image on-screen. The software allows a site plan to be overlaid onto any plot, can produce surface or contour plots, and has fence (cross-section) diagram capabilities. Other software, such as Microsoft Excel®, has been used for data handling as needed. Adequate geostatistical software and tools have historically been weak or lacking altogether. EarthVision now has a kriging module with several univariate descriptive statistical measures, spatial measures including h-scatterplots and variograms, and 2-D and 3-D ordinary kriging capability.

**2.3.4 Ames Laboratory ESC Phase 2**

The phase 2 activities comprise an intensive field investigation according to the work plan developed during phase 1. The field work is managed by the ESC team that performed the phase 1 work and developed the phase 2 work plan. The intended output of the phase 2 activities include an updated site geologic and contaminant model with sufficient information to allow remedial decisions to be made with relatively high confidence. The ESC team and a variety of sampling and subsurface exploration tools are brought to the site. A mobile analytical laboratory equipped to test for each contaminant of concern in each media of concern yielding screening or definitive type data, as needed, is brought to the site. Data are available within hours, or perhaps a day at most, from the
analytical and geologic assessment tools. As new data becomes available, it is incorporated into the site conceptual model and is used to aid the selection of subsequent sample locations.

At the Marshalltown FMGP site ESC demonstration, the initial focus was to begin soil sampling for chemical analyses from outside the PAH plume and work in toward the plume center. This approach resulted in a greater density of non detect samples than were necessary to define the eastern edge of the plume, while the western edge of the plume was not well defined due to the presence of buildings on the site. At the Savannah River Site, groundwater contamination was very widespread and the alluvial soils overlying a marine clayey soil at a variable depth of about 45 (±10) feet were heterogeneous resulting in an apparent erratic contaminant distribution. Initial sample placement was near the expected source of contamination. Subsequent groundwater sample locations for chemical analyses were based on an attempt to understand contaminant migration pathways based on available geologic information and prior contaminant concentration measurements. Sample location decisions were based on input from the ESC team members, including the site hydrogeologist and the analytical team manager, with primary authority and ultimate decisions made by the site project manager, the Ames Laboratory ESC principal investigator. The sample location selection process was made difficult by the heterogeneous nature of the soil deposits and by the occasional unexpected high or low measured analyte concentrations. This on-site sample location selection process is one of the ways that the observational method is utilized in the ESC process.

2.3.5 Phase 2 Subsurface Characterization Technologies

Geology The primary tools for phase 2 characterization of the site geology are the minimally invasive geoprobe and cone penetrometer, sometimes called "direct push technologies". Both tools were used at the Marshalltown FMGP site and at the Savannah River DOSB site. The geoprobe was used to push a soil electrical conductivity probe into subsurface soils to develop a depth versus soil electrical conductivity profile for stratigraphic logging using technology developed
by Geoprobe Systems (Christy, Christy, and Wittig. 1994). The cone penetrometer was used to develop a soil behavior classification profile versus depth using the cone tip stress and sleeve friction according to well developed empirical soil classification schemes by Robertson (1989) and Olsen (1988). At the Savannah River site there were two cone penetration trucks, both equipped with a piezocone which provided soil pore fluid pressure data in addition to the tip and sleeve friction stress data; the piezocone pore pressure data can be used to provide additional soil behavior information for the soil classification developed by Robertson. Having two cone penetrometers at the Savannah River site provided a unique opportunity to compare these techniques. Both geoprobe and cone penetration test (CPT) pushes are performed adjacent to several (preferably two or more) continuously logged boreholes to correlate the direct push output with the actual soil type. Both the direct push technologies have the ability to take soil samples allowing a field classification of the soil, however, samples are generally taken for contaminant analysis and not solely for a soil classification. Nevertheless, this sampling capability may provide additional information regarding soil types and soil mass characteristics, such as fractures in clays, root holes, etc., that may not be apparent from the probe sensor output.

**Hydrogeology** Little effort was made to further understand the hydrogeology at the Marshalltown FMGP site while only a moderate amount of additional hydrologic data was collected at the Savannah River Site during phase 2. At the Marshalltown FMGP site, the hydrogeology had been characterized during the BVWST Remedial Investigation, and further characterization of the hydrogeology was not part of the ESC phase 2 work there. At the Savannah River site, six piezometers were installed surrounding the DOSB prior to ESC phase 1, and were monitored for groundwater depth on approximately a monthly basis for a time period beginning about one year prior to the ESC phase 2 field work. Together with historic data from four prior piezometers at the site, this data provided the depth to groundwater, seasonal and longer term natural variations in water table depth, and the prevailing direction of surface groundwater flow.
The cone penetrometer piezocone data provided a rough estimate of depth to groundwater. In addition, a series of pore pressure dissipation tests were conducted to enable an estimation of the hydraulic conductivity of the fine grained soils. In a few cases, the pore pressure dissipation tests were carried out long enough to allow the soil pore pressure to stabilize so that a static pore pressure versus depth assessment could be made. This could potentially be used to characterize vertical hydraulic head gradients in the soil, however, comparison of data from the two piezocones on-site indicated a discrepancy between the two instruments which was not noticed until after the phase 2 investigation. This discrepancy introduced some analysis complications for this data; more will be said about this in Chapter 3.

Contamination Contamination in soil and/or groundwater is assessed on-site in mobile analytical laboratories using EPA SW-846, or other methods approved by the regulatory authorities. Soil and groundwater sampling was performed at both the Marshalltown FMGP and Savannah River DOSB Sites by using the geoprobe as the method of choice. The geoprobe is mounted on the back of a pick-up truck, and is relatively fast, produces little contaminated waste material to contain, and is relatively maneuverable over uneven terrain and in tight spaces.

Data Analysis The principal data analysis and integration software has been EarthVision® and is used as described above. Other software, such as Microsoft Excel®, has been used for data handling as needed. Adequate geostatistical software has historically been lacking. Approaches and methods for on-site data analysis include fence diagrams and 2-D and 3-D iso-contour plots of geologic and groundwater data and iso-concentration plots of chemical data.

2.3.6 Ames Laboratory ESC Concluding Remarks

The ESC process strives to achieve a rapid site characterization by condensing the process into two shortened investigation phases with a goal of reduced site characterization cost. The result of such a short term intensive focus on a site provides a "snapshot" of the site during the time period in which the investigation takes place. However, physical processes in the environment are
continually changing, e.g. surface groundwater flow may change directions from time to time depending on prevailing hydrogeological conditions, dissolved contaminants move with the groundwater so that plumes may grow or change shape over time, an influx of rainwater may dilute contaminants, volatile compounds in the soil may enter the aqueous phase or may escape to the atmosphere, non-aqueous phase liquids move in a complex relationship with the soil and groundwater, and so forth. Accordingly, we must bear in mind that changes with time may not be well characterized with the ESC method, or any other rapid site characterization technology, and that significant changes in the actual site conditions may occur after the site is "characterized."

One outcome of the ESC process is a recommendation for the placement of several monitoring wells so that expanding plumes, or changing contaminant concentrations can be assessed over time. Early actions taken to remediate a site, as under the SACM, may prove to be the most beneficial in view of the fact that site conditions may change substantially with time after the site investigation is performed. One might conclude that the site is not fully characterized until a remedial investigation report is written, or at least until the data generated during the site investigation is converted into useful information which can be used to make remedial decisions. Accordingly, a proper evaluation of the ESC methodology must consider the time required to interpret the data and disseminate the information to the remedial decision authority so that some action can be taken.

2.4 Statistical Techniques for Site Characterization

This section describes several statistical techniques that are applicable to on-site decision making and the characterization of spatial phenomena. Statistical analysis of spatial data is not new, has borrowed a good deal from mining applications, and has been applied to a variety of soil and groundwater contamination problems. The following sections briefly describe sequential data analysis, kriging, some Bayesian approaches, adaptive sample location selection, stopping rules, and treatment of non-detect analytical data. These techniques will be useful in the characterization
of the geology, hydrogeology, and contaminant distribution at a site, as well as in the development of the on-site decision making process and stopping rules.

2.4.1 Classical Sequential Analysis

Sequential data analysis provides a statistical approach to analyzing data as it is collected for hypothesis testing or parameter estimation about the population from which the data is drawn. It may be assumed that the data are independent and identically distributed from one collection time to the next, however, procedures for data which are not identically distributed, and possibly not independent exist (for example, see Ghosh and Sen, 1991). Sequential procedures for data which are not independent, not identically distributed, or both, require that the joint density function for the sample be known, or at least estimated, so that a proper likelihood ratio can be constructed (Ghosh and Sen, 1991, page 48). Sequential procedures are generally designed to control the rate of falsely concluding the null hypothesis is correct when in fact it is not, and the rate of falsely concluding the alternative hypothesis is correct when in fact it is not. At the outset of a sequential analysis, the number of observations that will be made to reach a decision is random and depends on the accumulating data. This differs from, perhaps more common, fixed sample size statistical procedures whereby a fixed number of observations are collected, followed by an hypothesis test or parameter estimation procedure used to make an inference about the population from which the samples where drawn. A major advantage of the sequential procedure is that when both sequential and fixed sample size procedures are applicable to a given problem, the sequential procedure reduces the expected number of observations, and therefore the expected cost, to reach a decision for given decision error rates (Ghosh and Sen, 1991).

In a sequential hypothesis test, the data may be collected and analyzed one at a time, or in groups, such as five points at a time. At each data collection period, a statistic is calculated and is used to arrive at one of three conclusions: 1) accept a null hypothesis, such as the site contaminant mean concentration exceeds the cleanup standard, 2) accept an alternate hypothesis, such as the
site contaminant mean concentration is below the cleanup standard, or 3) decide there is insufficient
data to support or refute the null hypothesis and that more data needs to be collected. It is possible
that the decision to collect more data is reached again and again, so that the data collection process
could continue indefinitely. To avoid this situation, a practical upper limit on the sample size, or
number of observations to be made, is decided upon before the data collection begins. If the upper
sample size is reached before a decision in the sequential analysis is made, then an alternate
decision rule is applied so that a conclusion to favor either the null or the alternate hypothesis, can
be made, and the data collection is terminated. A host of sequential analysis methods for a variety
of hypotheses and assumptions is given by McWilliams (1989) and Ghosh and Sen (1991).

2.4.2 Kriging and Classical Geostatistics

Kriging is a data modeling approach to analyzing spatial data which utilizes spatial
correlations or dependencies observed in the data. The product of a kriging analysis is a statistical
predictive model for a spatial process. The model includes a prediction variance that may be used
to characterize the uncertainty associated with the prediction. Kriging is sometimes called a classical
geostatistical procedure.

Kriging uses either a spatial covariance model or a semivariogram model to characterize the
spatial dependence; ideally this model is based on data but one could make an educated guess to
define the correlation structure, as may be done with Bayesian updating (discussed below), and
perform spatial prediction calculations using the kriging equations. Kriging is based on some
stationarity assumptions (see Cressie, 1991, for example). A constant mean over the region of
interest is assumed (this is first order stationarity); however, if a trend is thought to exist, it can be
removed prior to the analysis or accommodated by using universal kriging equations. The modeler
may assume second-order stationarity whereby the covariance between two spatial locations is a
function of the distance (vector quantity) between the points but does not depend on the location of
those points. A less stringent assumption is made by the intrinsic hypothesis whereby the variance
of the difference between spatially separated values, \( \text{Var}(Z(s_1) - Z(s_2)) = 2\gamma(s_1 - s_2) \), where \( Z(s_j) \) is the observed value at location \( s_i = (x_i, y_i) \), for \( i = 1 \) or \( 2 \), is a function only of the separation distance (a vector quantity, \( s_1 - s_2 \)) but is not dependent on the location of those points. The intrinsic hypothesis is less stringent than the second-order stationarity hypothesis because the semivariogram, \( \gamma(\cdot) \), may exist when the covariance does not (see Cressie, 1991, for example).

If the covariance exists, then the kriging equations can be constructed in terms of either the covariance function or the semivariogram (see Cressie, 1991, for example). If \( C(0) = \text{Var}(Z()) \) is the variance of the process, and \( C(h) \) is the covariance between values separated by distance \( h \), then for a second-order stationary process, \( \gamma(h) = C(0) - C(h) \). In terms of the semivariogram \( \gamma(h) \), for data \( Z(s_j), i = 1,\ldots,n \), the ordinary kriging predictor is (Cressie, 1991, pages 121-122)

\[
\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i). \tag{2.1}
\]

The weights \( \lambda_1,\ldots,\lambda_n \) can be obtained from

\[
\lambda_0 = \Gamma_0^{-1} \gamma_0,
\]

where

\[
\lambda_0 = (\lambda_1,\ldots,\lambda_n,m)^T,
\]

\[
\gamma_0 = (\gamma(s_0 - s_1),\ldots,\gamma(s_0 - s_n),1)^T, \text{ and}
\]

\[
\Gamma_0 = \begin{cases} 
\gamma(s_i - s_j), & i = 1,\ldots,n, j = 1,\ldots,n \\
1, & i = n+1, j = 1,\ldots,n \text{ and } j = n+1, i = 1,\ldots,n \\
0, & i = n+1, j = n+1.
\end{cases}
\]

The kriging prediction variance is given by

\[
\sigma_k^2(s_0) = \lambda_0^T \gamma_0 \tag{2.2}
\]

where \( ^T \) is the matrix transpose operator, and \( m \) is a Lagrange multiplier that ensures \( \sum_{i=1}^{n} \lambda_i = 1 \) so that \( 2.1 \) is an unbiased predictor. The semivariogram is calculated by
\[ \hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{i=1}^{N(h)} (Z(s_i) - Z(s_i + h))^2 \]  \hspace{1cm} [2.3]

where \(|N(h)|\) is the number of data pairs separated by distance \(h\).

Kriging (and Bayesian updating) honors the data in that the equations will reproduce the data values when predicting to a location which has been sampled. If there is no correlation structure in the \(Z\) process, then the kriging equations reduce to a prediction equal to the average of all the data with a kriging variance equal to the variance of all the data at each non-sampled location. Cressie (1991, page 128) shows how measurement error, common in analytical data, can be incorporated into the kriging equations.

Universal kriging attempts to model a nonstationary mean structure by modeling the mean trend as an unknown linear combination of known functions. In other words, the mean is modeled as a linear combination of known functions \( \{f_0(s), f_1(s), \ldots, f_p(s)\} \) defined over the spatial coordinates, \(s\), of the region of interest, however, the weighting parameters defining the linear combination are unknown. These functions are incorporated into the universal kriging equations to give a predictor with minimized mean-square prediction error (see Cressie, 1991, for details). Cressie (1986) developed a median polish kriging algorithm whereby a median polish algorithm is used to approximate the nonstationary mean and a kriging model is developed from the residuals defined by the data minus the median polish surface. The kriged residual surface is added back to the median polish surface to define the median polish kriged surface. Cressie (1986) shows that median polish kriging provides a relatively outlier resistant, nearly bias-free way of kriging in the presence of drift, and provides results as good as the mathematically optimal, but operationally difficult, universal kriging. Details of the median polish kriging procedure are described by Cressie (1986 and 1991).

Philip and Kitanidis (1989) model some Wolfcamp aquifer hydraulic head data where there appears to be some trend in the hydraulic head. They assume a locally stationary mean and use only data nearby to each kriged grid point in the kriging interpolation process using a variogram estimated from data spaced approximately perpendicular to the direction of the apparent trend.
Isaaks and Srivastava (1989, page 343) suggest that restricting the search neighborhood to points nearby the kriged point brings the actual statistical properties closer to the model so that the stationary random function conceptualization becomes more plausible.

Much of the kriging analysis and conclusions are dependent upon the assumptions and modeling approach used. This is shown clearly by Englund (1990) whereby markedly different results were obtained by 12 different investigators who took different approaches to analyzing identical data sets. Englund concludes that failure to use appropriate interpolation techniques will result in significantly increased remediation costs from the probable thousands of contaminated sites across the U.S. for which spatial interpolation from sample data will be required. Englund finds no spatial interpolation methodology to be best, but concludes that deciding which measure of estimation quality is most relevant to the particular circumstances of a site investigation is crucial to selecting the "best" interpolation method.

2.4.3 Bayesian Methods and Decision Analysis

Bayesian updating, as applied below, is based on the assumption that the process under study is at least second order stationary. This means that the process, \( Z \), has a constant mean and variance, with a spatial dependence structure which is independent of the spatial location, \( s = (x,y) \) or \( s = (x,y,z) \). In practice we may model the appropriately transformed residuals of some trend surface function of the data as the stationary process, and then back-transform the model to the data scale. \( Z = Z(s) \) could be a contaminant concentration, stratigraphic surface elevation, stratigraphic layer thickness, hydraulic conductivity, etc. at location \( s \). Initial "prior" statistical distributions for reasonable values of the mean, variance, and spatial correlation structure are made on the basis of any hard or soft data available. This may include experience with similar sites (soft data), measured values (hard data), geophysical data, or any of the usual sources of information gathered to form the phase 1 conceptual model, such as the expected size of the contaminant plume. Then, as field data is collected, the "prior" parameter distributions are updated to "posterior" distributions based on
equations developed from application of Bayes theorem (see Benjamin and Cornell, 1970, for a discussion of Bayes theorem).

Massmann and Freeze (1989, see also Freeze et al., 1990) provide Bayesian updating equations to predict values of the Z process at unsampled locations conditioned on measured data at known sample locations and a prior distribution which may, or may not, be based on prior data. Their equations are based on the assumption that Z(·) is normally distributed with known covariance matrix. For spatial locations s_i, i = 1,..., p, of interest, let the prior vector of means be \{\mu_j\} with autocovariance matrix [\tau_{ij}]. The autocovariance matrix [\tau_{ij}] is the same as the covariance matrix, C(·), discussed in the preceding kriging section 2.4.2. The conditionally updated terms (conditioned on the n observed data) are given by

\begin{align*}
\{\mu_i\}_c &= \{\mu_i\} + \{\tau_{ij}\}B^T(B\{\tau_{ij}\}B^T + E) (Z - B\{\mu_i\}) \\
[\tau_{ij}]_c &= [\tau_{ij}] - [\tau_{ij}]B^T(B[\tau_{ij}]B^T + E)^{-1}B[\tau_{ij}]
\end{align*}

[2.4]

where \(T\) is the matrix transpose operator, Z is an nx1 vector of observed data, B is an nxp (n<p) matrix of zeros and ones satisfying \(Z = B(Z_i) + e\), \(Z_i\) is a px1 vector of unmeasured and measured values at locations s_i, and e is an nx1 vector of unobservable zero-mean measurement errors with covariance matrix E. Note that measurement error variance can be incorporated into equations [2.4] via the measurement error matrix E. As more data is collected, the posterior parameter distributions may be used as new prior distributions, and updated posterior distributions based on [2.4] are developed.

Freeze et al. (1990, 1991, and 1992) produced a four-part series that describes the application of decision analysis to engineering design for projects in which the hydrogeological environment plays an important role. They espouse a more integrated process whereby the investigation, design, and implementation phases of the project are considered throughout the investigation. They couple a decision model based on a risk-cost-benefit objective function, a simulation model for groundwater flow and transport, and an uncertainty model that encompasses both geologic uncertainty and parameter uncertainty. The authors state that such a decision analysis
could lead to cost savings in projects such as design of dewatering systems, waste containment facilities, and contaminated groundwater remediation, and that it is well suited to the design of site investigation programs and monitoring well networks and assessment of the potential worth of additional data from either source. The methodology provides a risk-based approach whereby the risk of failure to meet the design objectives reflects the uncertainty in the technical analysis.

Freeze et al. (1990) carry out the decision analysis in a Bayesian framework in which prior uncertainties are reduced to lesser posterior uncertainties as additional data are collected. Distinction is made between geological uncertainty and parameter uncertainty. Parameter uncertainty refers to the spatial distribution of parameters, such as hydraulic conductivity or porosity, whereas geologic uncertainty addresses uncertainties in the location of boundaries between geologic units and the continuity of those units. The parameter uncertainty is modeled as a stationary stochastic autocorrelated normally distributed process. They note that parameter values vary due to heterogeneity in the geology, whereas uncertainty is in the mind of the analyst. Indicator kriging and search theory are used to define and reduce uncertainty in the geologic model on the basis of drilling and logging of boreholes. The search theory discussed in the work of Freeze et al. (1990) is the same theory discussed in EPA (1989) to detect hot spots, and is suggested in the work of Freeze et al. as a method to detect the presence of discontinuities in aquitards.

As outlined by Freeze et al. (1990), a given project may have several alternative designs, perhaps on the basis of the local geology, and for each design there are possibly several ways in which the project could fail by not satisfying its intended purpose. The decision model is based on an objective function, \( \Phi_j \), defined for each of the \( j = 1, ..., N \) alternatives,

\[
\Phi_j = \sum_{t=0}^{T} \frac{1}{(1+i)^t} \left[ B_j(t) - C_j(t) - R_j(t) \right] 
\]

where \( B_j(t) \) is the benefit of alternative \( j \) in year \( t \) [$], \( C_j(t) \) is the cost of alternative \( j \) in year \( t \) [$], \( R_j(t) \) is the risk of alternative \( j \) in year \( t \) [$], \( T \) is the time horizon [years], and \( i \) is a decimal fraction discount rate. There may not be any benefit, \( B_j(t) \) (income) for a remedial design alternative, but
there are construction and operation costs, $C_j(t)$. The risk, $R_j(t)$, includes the cost of a failure of the design times the probability of failure times a utility function. Conversely, one could write the objective function as the expected costs minus expected benefit. In terms of costs, we wish to minimize the objective function; in terms of benefits, we wish to maximize the objective function.

The design alternatives depend on the geological and geotechnical parameters which are both sources of uncertainty. The probability of failure depends on the design and the geology. One can estimate the required probabilities and evaluate the objective function to determine the expected cost for each possible alternative and outcome. The appropriate action would favor the most cost effective outcome. One alternative that may be considered in the analysis is to collect additional data to reduce geological/geotechnical uncertainties. If the best alternative (most cost effective) turns out to be 'collect more data,' then more data is collected, the model is updated using the new and existing data, new probabilities are estimated, new alternatives are developed, and the objective function is again evaluated. Several iterations may be required before a design decision is made.

The decision analysis process may depend on some assumptions, even educated guesses, regarding the probabilities required, spatial covariances, and statistical distributions. Nevertheless, Freeze et al. (1990) feel that existing information and good engineering judgment can often be used to make reasonable choices, and that such information is incorporated in all analyses, but that the Bayesian analysis incorporates this information in an open and objective way. Lastly, these authors note that the Bayesian updating model could be replaced by a parameter uncertainty model based on kriging, but feel that the Bayesian approach is better suited to the style and needs of engineering design.

James and Gorelick (1994) apply the decision analysis described in the preceding paragraphs to a hypothetical groundwater contamination remediation problem. Their goal is to find the optimum number and best locations for a sequence of observation wells that minimize the expected cost of remediation plus sampling. They conclude that points of greatest uncertainty in plume presence (i.e. where the probability of contamination exceeding a concentration threshold is
near 0.5 based on indicator kriging) were generally poor candidates for sample locations because sampling near the center of the plume provides little information about the plume edges, and randomly located samples were not cost effective. The procedure involves computationally intensive Monte Carlo simulations for the generation of many hypothetical contaminant plumes which are assumed to equally likely represent the real unknown plume; the authors conclude that this computation step may provide some practical limitations with the method. Inputs include remedial design and implementation costs, failure modes and associated costs, and probabilities of failure of the remedial design(s). The method focuses the sampling effort to provide data that is expected to reduce the remediation costs.

2.4.4 Adaptive Site Selection and Stopping Rules

Thompson (1992) provides a review of adaptive and conventional sampling designs, whereby "adaptive sampling" refers to sampling designs in which the procedure for selection of sample units may depend on the locations and values of the variable of interest already observed in the course of the study. The primary purpose of adaptive sampling designs is to use the observed sample characteristics to obtain more precise estimates of the population characteristics for a given sample size or cost than might be obtained with conventional (random, systematic, or stratified random) sampling designs. The basic idea is to take advantage of aggregation tendencies, or spatial correlations, in the sampling process. Accordingly, adaptive sampling designs are necessarily sequential, but go beyond the usual considerations considered in sequential statistics in that the sample unit labels (or spatial locations) make it possible to choose which units to sample next (Thompson, 1992). Adaptive selection may introduce biases into conventional estimators, giving rise to a need for the development of unbiased estimators for adaptive sampling strategies (Thompson 1992). Isaaks and Srivastava (1989) and Deutsch and Journel (1992) address bias in assessing statistical distributions by the use of declustering techniques.
David and Yoo (1993) describe several tools for spatial interpolation, adaptive measurement site selection, and a stop-decision rule for site characterization. For spatial interpolation they use Hardy's Multiquadratic Biharmonic (MQ-B) method which is simply kriging with the distance between sample points used in place of the variogram. However, quantification of the prediction variance requires a proper characterization of the variogram, or covariance, of the process. Their sequential adaptive measurement site selection requires only the locations and measurements for the sites selected "so far," and is based on three ideas: attaining nearly uniformly dispersed measurement sites; selecting the next measurement site which, together with some smaller, earlier, set of measurement sites, would have yielded a map near the current map; and avoiding selection of sites on the boundary of the area of interest.

David and Yoo (1993) select a "next" sample location, \( s_{n+1} \), from the area \( A \) by defining

\[
\mu = \max_{s \in A} \min_{i=1,...,n} \| s_i, s \|, \text{ and with } p \text{ less than, but near, } 1, \text{ let } T \text{ be the set of sites whose distance to the nearest sampled site is between } p \mu \text{ and } \mu:\]

\[
T = \{ s : p \mu \leq \min_{i=1,...,n} \| s_i, s \| \leq \mu \}. \tag{2.6}
\]

They then choose a site from the set \( T \) which, together with some "earlier" smaller set of measurement sites, would have yielded a map close to the current map. This is accomplished as follows: let \( \sigma_j, j=1,...,J, \) be sites in the set \( T \) and let \( \delta_{\tilde{j}} = \int_A \| Z(\cdot) - Z_{\tilde{j}}(\cdot) \| ds \) be the discrepancy between the Hardy map \( Z(\cdot) \) constructed on the basis of \((s_i, Z(s_i)),..., (s_n, Z(s_n))\) and the Hardy map \( Z_{\tilde{j}}(\cdot) \) constructed by replacing \((s_j, Z(s_j))\) by \((\sigma_j, \tilde{Z}(\sigma_j))\), for \( i=1,...,n \), and \( j=1,...,J \), and where \( \tilde{Z}(\sigma_j) \) is the Hardy estimate of \( Z(\sigma_j) \). The next sample location \( s_{n+1} \) is the site \( \sigma_j \) minimizing \( \max_{i=1,...,n} [\delta_{\tilde{i}}] \).

David and Yoo (1993) choose a stopping rule based on a discrepancy function

\[
\delta(n) = \delta(Z^{n/2}(\cdot), Z^n(\cdot)) = \int_A \| Z^{n/2}(s) - Z^n(s) \| ds, \tag{2.7}
\]

David and Yoo (1993) select a "next" sample location, \( s_{n+1} \), from the area \( A \) by defining
\[
\mu = \max_{s \in A} \min_{i=1,...,n} \| s_i, s \|, \text{ and with } p \text{ less than, but near, } 1, \text{ let } T \text{ be the set of sites whose distance to the nearest sampled site is between } p \mu \text{ and } \mu:\]

\[
T = \{ s : p \mu \leq \min_{i=1,...,n} \| s_i, s \| \leq \mu \}. \tag{2.6}
\]

They then choose a site from the set \( T \) which, together with some "earlier" smaller set of measurement sites, would have yielded a map close to the current map. This is accomplished as follows: let \( \sigma_j, j=1,...,J, \) be sites in the set \( T \) and let \( \delta_{\tilde{j}} = \int_A \| Z(\cdot) - Z_{\tilde{j}}(\cdot) \| ds \) be the discrepancy between the Hardy map \( Z(\cdot) \) constructed on the basis of \((s_i, Z(s_i)),..., (s_n, Z(s_n))\) and the Hardy map \( Z_{\tilde{j}}(\cdot) \) constructed by replacing \((s_j, Z(s_j))\) by \((\sigma_j, \tilde{Z}(\sigma_j))\), for \( i=1,...,n \), and \( j=1,...,J \), and where \( \tilde{Z}(\sigma_j) \) is the Hardy estimate of \( Z(\sigma_j) \). The next sample location \( s_{n+1} \) is the site \( \sigma_j \) minimizing \( \max_{i=1,...,n} [\delta_{\tilde{i}}] \).

David and Yoo (1993) choose a stopping rule based on a discrepancy function

\[
\delta(n) = \delta(Z^{n/2}(\cdot), Z^n(\cdot)) = \int_A \| Z^{n/2}(s) - Z^n(s) \| ds. \tag{2.7}
\]
where $Z^{(m)}(\cdot)$ represents the surface estimated from the first $m$ measurement sites for $m = n$ or $m = n/2$ (about one half the data), and $A$ represents the area of interest (in the subsequent analyses I have divided the discrepancy function [2.7] by the area, $|A|$, of $A$ so that the discrepancy function has the same units as $Z$). Ideally, we would want to stop sampling when $\delta(n)$ is small. For highly "granular" surfaces for which spatial interpolation is unreliable (perhaps because the spatial continuity occurs at a scale smaller than the data spacing), the discrepancy function will fail to converge, but otherwise, it can be relied on to quantitatively determine whether convergence is taking place. This idea can be used to help determine when enough data of sufficient accuracy have been collected. Note, however, that convergence of successive map surfaces is not equivalent to convergence of predicted and measured values. Successive map surfaces may converge according to a plot of $\delta(n)$ versus number of samples, $n$, with only a small improvement in the ability to predict values at unsampled locations. Uncertainty associated with prediction may be best quantified with a careful analysis that yields a statistical prediction standard deviation and error distribution.

Johnson (1993), has developed an adaptive sampling technique based on sample results generated in the field using a combination of Bayesian and indicator kriging methodologies. The Bayesian analysis allows quantitative integration of soft information with hard (quantitative analytical) data. Soft data can include historical information, geophysical survey data, preliminary modeling results, or personal experience with similar sites. Indicator kriging based on quantitative analytical data is used to update the initial conceptual image. New sampling locations are selected to minimize the uncertainty associated with contaminant extent by selecting sample locations to either maximize the area declared clean, maximize the area declared contaminated, or minimize the area declared as state uncertain. Johnson (1996b) has found that maximizing the area declared clean appears to provide the best sampling strategy to define the extent of contamination. The basic methodology is discussed further in Chapter 3, section 3.2.2.

Englund and Heravi (1994) used a Monte Carlo simulation to compare three phased sampling schemes each using the same total number of observations. Their first scheme selected
all samples at random in one sampling event and then performed an ordinary kriging analysis to estimate the average contaminant concentration for each remediation unit within the site. Remediation units with an average concentration estimated to be above an action level were remediated. The second, two-phase, scheme selected random locations for a fraction of the observations in a first phase and performed an ordinary kriging analysis to predict the average concentration for each remediation unit. A loss function based on decision errors for false positive remediation cost (remediating a "clean" unit) and false negative cost (failing to remediate a contaminated unit) was then used to adaptively select sample locations for a second sampling phase. Second phase samples were placed in units with the highest expected losses. A subsequent ordinary kriging analysis was used to estimate the average contaminant concentration for each remediation unit. The third scheme selected each sample one at a time at random from within the remediation unit with the greatest expected loss. This process continued until the desired number of samples was obtained at which time ordinary kriging was used to estimate the average contaminant concentration for each remediation unit.

The cost structure Englund and Heravi used is illustrated in Figure 2.6. They assumed a $10,000 cost to remediate a remediation unit and a cost that is proportional to the contaminant concentration for unremediated contaminated soils. The cost associated with unremediated soils is difficult, if not impossible, to quantify accurately in practice; this cost includes many intangibles such as non-compliance fines and costs to society which are related to the contamination. The loss function they used is proportional to the distance between the action level and the true concentration (see Figure 2.6); the expected loss was obtained by integrating over the false positive and false negative decision error regions using the ordinary kriging standard deviation and assuming a double triangular kriging error distribution centered at the kriging estimate and varying three standard deviations to each side.

On the basis of total costs, including sampling costs, remediation costs, and costs associated with failing to remediate contaminated units, Monte Carlo simulation lead Englund and Heravi (1994)
to conclude that more sampling phases are better (lower total cost), but selecting sample locations one at a time is only marginally better than the two-phase scheme, and that the optimal total number of samples was found to be independent of the number of phases. Furthermore, collecting about 75 percent of the samples in the first phase of the two-phase scheme is near optimal, while collecting less than 20 percent in the first phase is actually counterproductive. Englund and Heravi note that the true costs associated with not remediating contaminated soils include many intangibles and are difficult to quantify, however, they do not explore the effects of loss functions other than a linear loss using their methodology.

2.4.5 Evaluating Contaminant Concentrations in the Presence of Non-Detect Data

In the characterization of a contaminant plume, sampling along the edges of a contaminant plume where contaminant concentrations are low may yield samples with analytical test results below a method detection limit (MDL) or practical quantification limit (PQL). This practice results in a left
censored probability distribution for the chemical concentration data whereby all values from the true
distribution which are below the MDL are replaced by ND or the detection/quantification limit. Such
censoring of the data adds a complication for the data analyst and additional uncertainty in the
analysis and conclusions. As Cressie (1994, page 140) points out, we should not lose sight of the
larger goals of the study. Exploratory data analysis followed by methods such as statistical spatial
analysis attempt to filter out the measurement error and make inference on true concentrations. This
task is made much more difficult if measurements below the MDL or PQL are replaced by NDs.
Statistical methods to estimate parameters for censored distributions without consideration of spatial
or temporal relationships, such as given by Sara (1994, page 11-50), are available.

Stein (1992) has developed a method to characterize spatial relationships in truncated
spatial data which follow a multivariate normal distribution. As non-detect contaminant concentration
data are common in environmental site characterizations, the method developed by Stein may prove
useful in the analysis of such data. At either an unsampled location or at a location where the
sample result is ND, Stein's method estimates the probability, conditioned on the measured data
(including the ND data), that the contaminant concentration is less than some set of user specified
cutoffs. If the cutoffs are chosen well, the conditional mean and variance (i.e. conditioned on the
data) at unsampled locations can be predicted. The method is based on simulated data values
under the assumption that the process is multivariate normal with known covariance matrix, and that
the data are truncated at zero. Since contaminant data may not be normally distributed, and have a
non-detect truncation value greater than zero, the data may need to be transformed so that they
follow (approximately) a normal distribution, followed by a translation so that the truncation point is
zero. The method has direct application for the estimation of the probability that a contaminant
concentration at some spatial location is less than (or greater than) a cutoff value, such as an action
level or regulatory limit. Details of the algorithm are given by Stein (1992).

A practice which further complicates the non-detect data problem is the occasional dilution of
samples prior to analysis, perhaps to avoid fouling the analytical equipment with samples which
appear heavily contaminated. The effect of this is to raise the analytical detection limit for these samples by an amount equal to the dilution factor, e.g. a ten times dilution will raise a 1 ppm detection limit to 10 ppm. This creates a potential data analysis problem which the method of Stein does not address.

2.4.6 Concluding Remarks on Statistical Methods for Site Characterization

The EPA (1989 and 1992a) hypothesis test and sequential analysis methods are based on measurements at randomly selected sample locations within a particular stratum (not necessarily a geologic strata, but a selected region within the site) so that the data can be treated as statistically independent and identically distributed samples from the population of interest. In contrast, the geostatistical approach may assume the process, e.g. contaminant concentration or geologic strata, has the same mean everywhere, that the variance of deviations from this mean do not depend on the spatial location, and that the data exhibit some spatial correlation which is related to the continuity of the phenomenon of interest. The geostatistical approach does not require statistically independent data but rather seeks to capitalize on the spatial dependence in the data to obtain unbiased prediction at unsampled locations.

Consider a site with one or more sources of contamination. There is likely to be one or more regions of high contaminant concentrations (near the source areas) with contaminant concentrations that decrease away from the source, perhaps with concentrations decreasing less rapidly down-gradient of the source. Due to geologic heterogeneities and distance from source areas, the local variations in concentrations may be greater in some areas than others. Given a sufficient quantity of data, non constant variance over spatial locations can be modeled with the use of a relative variogram (see Isaaks and Srivastava, 1989, and Cressie, 1991). For example, data defining a contaminant plume may contain a substantial proportion of non-detect (ND) values near the plume edges. If sufficient data exists, it might be reasonable to partition the site into two (or more) regions to construct a model of the plume, and a separate model (i.e. no contamination) in those regions
containing only ND data. However, for limited data sets in the absence of external information, the practitioner might simply assume a constant process variance over the entire region of interest.

While ESC (and possibly FAST and M³ as well) is generally incompatible with the statistical hypothesis testing framework described in the DQO process (because sample locations are generally judgmental), the method is compatible with the basic intent of the DQOs. The seven steps (see section 2.1.2) of the DQO process can be applied to an ESC project with perhaps some modification to step 6: limits on decision errors. Probability maps developed from application of geostatistics can serve a similar purpose in understanding decision errors and driving the sampling program to reduce uncertainty. The geostatistical approach more naturally addresses the problem by utilizing spatial correlations to understand the contaminant distribution and factors in the geologic and hydrogeologic environment which control the fate and transport of the contaminants. The DQO type hypothesis testing framework under the assumption of statistically independent and identically distributed data is generally not compatible with judgmental sampling of spatially correlated phenomena.

The geostatistical approach working within the framework of ESC is directly relevant to addressing the four basic decisions identified in the DQO process (see section 2.1.2). The ESC phase 1 encompasses the early assessment decision to determine if the site poses a potential threat to human health or the environment. ESC can directly address the second advanced assessment decision to determine if the concentrations of COPCs exceed the preliminary remediation goals or other ARARs. The third advanced assessment decision, to determine the extent of contamination, as the volume and media of materials that exceed the contaminant action levels is directly assessed by the coupled use of geostatistics within the ESC process. The fourth cleanup attainment decision may be addressed using conventional random sampling and the statistical hypothesis testing framework described by EPA (1989, or 1992a); this is a post-ESC and post-remediation activity.

Journel (1985) notes that a random function \( Z(s) \) defined over a region \( A \) is stationary if its multivariate distribution is invariant by translation within the region \( A \), entailing that the moments (mean, variance, etc.) are also invariant by translation. Stationarity allows statistical inference, but is
a property of the probabilistic model, and not an intrinsic property of the real phenomena under study. The choice of the stationarity of a model may depend on the scale of the observation and the amount of data available. Joumel points out that stationarity, being a choice of the experimenter, can be validated a posteriori by judging whether this choice has been efficient in solving the particular problem at hand.

A summary of the statistical methodologies which I consider further in this thesis is given in Table 2.2. In particular, Table 2.2 focuses on the potential application of these statistical methods within ESC. The sequential statistical methods and hypothesis testing framework discussed in the EPA (1989 and 1992) and DQO documents require some randomization in the sample location selection process and are therefore not compatible with the judgmental sample selection used in the ESC process. Accordingly, these methods are not shown in Table 2.2. While the statistical based method to search for "hot" spots is well developed and appropriate for some environmental site characterizations, the ESC process does not generally collect data according to the sampling schemes resulting from this method. Accordingly, the "hot spot," or search theory method is not included in Table 2.2. The decision analysis developed by Freeze et al. is focused on sampling in conjunction with remediation design and alternatives. Because the current ESC process does not focus on remediation alternatives, I have not looked further at the decision analysis procedure.
Table 2.2. Summary of statistical methods considered for potential use in ESC.

<table>
<thead>
<tr>
<th>Purpose In ESC</th>
<th>Englund and Heravi</th>
<th>David and Yoo</th>
<th>Johnson</th>
<th>Stein</th>
<th>Bayesian Updating</th>
<th>Kriging</th>
<th>Decision Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive sample location selection</td>
<td>Adaptive sample location selection</td>
<td>Adaptive sample location selection</td>
<td>Prediction with ND data</td>
<td>Prediction</td>
<td>Prediction</td>
<td>Selection from among alternate actions</td>
<td></td>
</tr>
</tbody>
</table>

| Potential Application in ESC | Contaminant spatial distribution model | Prediction for contaminant or geologic phenomenon | Contaminant spatial distribution model | Prediction for contaminant or geologic phenomenon | Prediction for contaminant or geologic phenomenon | Prediction for contaminant or geologic phenomenon | Assess need for additional data, emphasis on geology |
| Data Needs | Remediation and sampling costs, cost of failure to remediate | Need enough data to develop a contour map | Can begin without data, uses data as it is collected | Need enough data to develop a contour map | Prior needs no data, updating requires data and spatial locations | Need enough data to develop a contour map | Alternative actions, failure modes, sample and failure costs |

| Uses Simulated Data | No | No | Yes | Yes | No | No | Yes |
| Handles ND Data Directly | No | No | No | Yes | No | No | No |
| Incorporates Cost | Yes | No | No | No | No | No | Yes |
| Requires a Covariance or Variogram Model | Yes | No | Yes | Yes | Yes | Yes | Usually |
3. ASSESSMENT AND DEVELOPMENT
OF GEOSTATISTICAL TOOLS FOR ESC:
FORMER MANUFACTURED GAS PLANT SITE

3.1 The Decision Making Process

Several of the four basic decisions identified in the DQO process, namely, 1) to determine whether the site poses a potential threat to human health or the environment, 2) to determine whether the concentration of contaminants exceeds applicable or relevant and appropriate requirements, 3) to determine the extent of contamination, and 4) a cleanup attainment decision to determine whether the final remediation levels or removal action levels have been achieved, are related to the issue of spatial distribution of contamination. The second decision requires specifying whether the ARARs are to be compared to the contaminant concentration for individual samples, an average concentration over portions of the site, an average over the entire site, or some combination of these criteria. The third decision suggests that the contaminant spatial distribution should be modeled to determine which areas of the site have higher local average contaminant concentrations, and which areas may be considered "clean".

Chapters 3 and 4 will focus on on-site sample location selection, rules to help decide when to stop sampling, and methods to estimate which portions of a site exceed ARARs. As each new piece of data is collected, it should be used to further develop the site conceptual model, to characterize the spatial distribution of contaminant concentrations, and may be used to decide if more data is needed, and where that data should be collected. While the conceptual model is a simplification of the actual site conditions (Sara, 1994), it should match reality as much as possible on the basis of realistic assumptions supported by data. This in turn will be used to make remedial decisions about the site.
3.2 Detailed Case Analysis

This Chapter provides an analysis of some data collected during the ESC demonstration at a former manufactured gas plant site in Marshalltown, Iowa. Particular emphasis is given to those data and data analysis applications which are not currently performed in an Ames Laboratory ESC project, and which may be useful in on-site decision making, including selection of subsequent sample locations and stopping rules, and post field work analysis to determine if a site needs remediation, and if so, which portions of the site need remediation.

3.2.1 FMGP Site in Marshalltown, IA, Phase 1

A remedial investigation (RI) had been completed prior to the Ames Laboratory ESC demonstration at the Marshalltown, Iowa, FMGP site (BVWST, 1992). In addition, the Ames Laboratory ESC report (Bevolo, Kjartanson, and Wonder, 1996) on the FMGP demonstration includes an analysis of all of the pertinent data that was collected during both the BVWST RI and the ESC demonstration. My intent here is to neither repeat that analysis nor provide an alternative analysis, but to demonstrate how several data analysis tools can be used to enhance the data interpretation, integration, and on-site decision making. The contaminants of interest at the FMGP site are 16 PAH compounds. The media of interest is soil. A plot of the site is shown in Figure 3.1.

The BVWST (1992) RI was performed by traditional site characterization methods whereby 16 soil borings and seven monitoring wells were used to determine the site geology and collect soil and groundwater samples for off-site analytical testing. Additionally, soil samples were collected in the crawl space beneath some of the on-site buildings. One of the first steps in an ESC program is to drill and log several boreholes to develop a general site stratigraphic column and to use for correlation/calibration of geophysical and direct push technologies. These borings are available from the BVWST RI.
3.2.1.1 Geology, Phase 1

Figure 3.2 shows a simplified soil boring log for monitoring well 3 (MW-3 from BVWST, 1992) showing six zones used in the analysis of soil PAH contamination at this site. It was known on the basis of 16 on-site BVWST logged soil borings and seven logged monitoring well borings on or near the site, that the general stratigraphy across the site was similar to that shown in Figure 3.2. Collectively, these 23 boring logs indicate that the lower cohesive unit slopes generally to the south, in the direction of the near surface groundwater flow, and may thin to zero along a bedrock ridge trending NW-SE through the NE corner of the site. This geologic information is useful in planning subsequent sampling locations as the coal-tar dense non-aqueous phase liquid (DNAPL) is expected to migrate vertically (and horizontally along preferential flow paths) through fractures in the upper cohesive soils and down through the granular soil unit to the top of the lower cohesive unit (LCU) where it will pool and/or move down-gradient along the LCU surface (Luthy et al., 1994).
Further geologic information was gathered during phase 1 by several geophysical surveys including ground penetrating radar (GPR), seismic reflection and refraction, electromagnetic offset logging (EOL), and borehole geophysical logging. The summary of this data by Bevolo, Kjartanson, and Wonder (1996) indicate that only the seismic refraction and borehole geophysical logging appeared to provide confident data. The seismic refraction data agreed reasonably well (usually within about two to three feet) with the BVWST borehole data with regards to defining the bedrock surface. While the borehole geophysical logs agreed with the BVWST logged boreholes, no new information beyond the interpreted interface between weathered and unweathered bedrock was gained. The relatively high wave velocity surface fill and noise from the railroad yard immediately south of the site negatively impacted the seismic surveys, while overhead power lines, buildings, metal fences and storage areas, and the relatively high electrical conductivity of the upper cohesive
unit (UCU) negatively impacted the GPR and EOL surveys. Additionally, not all of the geophysical contractors provided interpreted data to the Ames Lab ESC team, thus compounding the difficulty in attempting to use this data during the site investigation. Accordingly, no further use of the geophysical data is made here.

3.2.1.2 Analytical Screening Data, Phase 1 The analytical screening data collected includes three immunoassay (IMA) kits, a chemiluminescence (CL) analysis, and both passive and active soil gas measurements. The Millipore, Ensys and Quantix IMA kits gave a weighted average of the PAHs of interest. The Millipore and Ensys kits gave gravimetric PAH concentrations while the Quantix kit gave volumetric PAH concentrations. The CL method gave total PAH (less naphthalene) concentrations in ppm. Passive soil gas kits were returned to the vendor for analysis of naphthalene, anthracene/phenanthrene, and fluoranthene/pyrene by mass spectrometry and results are reported as ion counts. Active soil gas measurements were made on-site by gas chromatography/flame ionization detector and gas chromatography/photo ionization detector instruments with results for naphthalene reported in μg/L of vapor. Bevolo, Kjartanson, and Wonder (1996) report that the spatial distribution of PAH contamination determined by the screening technologies all compare fairly well.

The locations for the analytical screening measurements are shown in Figure 3.3. These 24 locations were selected to cover the expected region of soil PAH contamination with an approximately triangular spatial arrangement with about 50 feet between sample locations. Samples for the screening methods were collected at two depths: about 5 to 7 feet below ground surface (bgs), and about 9.5 to 15.5 feet bgs, called the upper and lower CL data, respectively. It was not possible to sample the deep zone at sample location ML027 due to some subsurface obstruction. Accordingly, there were 47 potential data values for each screening method.

Because these screening data are in generally good agreement I will not perform an analysis of each of them here. Ordinarily, the ESC program would not use this many screening methods at a
Figure 3.3. Plan view of the FMGP site with phase 1 screening data sample locations.
given site, but one of the objectives of this particular ESC demonstration was to compare a variety of screening methods. Of the IMA results, the Ensys and Quantix methods had one missing value while the Millipore method had seven missing values. The CL data were complete, with the exception that the ML026 deep value reported only "free DNAPL" which I gave a value of 50,000 ppm in the subsequent analysis.

A preliminary spatial analysis of the screening data will be useful to refine the site conceptual model, to estimate the spatial correlation structure, and to begin planning for the more intensive, phase 2 field work. Because the phase 2 soil samples will be analyzed on-site as they are collected, several days of data collection would normally be required to develop a data based spatial correlation model to improve the use of the spatial interpolators, kriging, Bayesian updating, and the method developed by Stein, as described above. Although the screening data are not acceptable for a risk analysis and may not be used in the final data analysis (depending upon what analyses the regulatory agencies will accept), an analysis of this data will aid our understanding of the spatial distribution of the PAH contaminants. Figure 3.4 shows the upper and lower CL data on a plan view of the FMGP site.

Figures 3.5 and 3.6 show the upper and lower CL data locations, respectively, with plots of the CL data versus the East-West and North-South directions and a dot plot illustrating a frequency distribution of the CL data. Note that about half of both the upper and lower CL data are non detect (ND) results with a quantification limit of 10 ppm. The exploratory data analysis plots in Figures 3.5 and 3.6 are modeled after similar plots by Cressie (1991). In both Figure 3.5 and 3.6, the CL data versus the easting indicate that the sample locations have captured the eastern edge of the PAH plume, but may not have captured the western edge of the PAH plume. Both figures also indicate that the major contaminant plume occurs near the north-central and center of the sampled area, with contaminant concentrations decreasing to the south. While there is insufficient data above the detection limit to make a confident determination, the dot plot suggests that even on the log scale, the data (above the MDL) appear to follow a more uniform than normal statistical distribution.
Figure 3.4. Plan view of the FMGP site with soil chemiluminescence data shown.
Figure 3.6. East-West and North-South marginal distributions and dot plot for the upper CL data.
Figure 3.6. East-West and North-South marginal distributions and dot plot for the lower CL data.
Semivariograms of the upper and lower base 10 logarithm CL data are shown in Figure 3.7. These semivariograms were calculated with the ND data from eastern-most sample locations ML003, ML007, and ML008 set aside. These ND data were set aside so as not to unduly influence the semivariogram estimate because they are separated from the contaminant plume by a series of ND data (ML010, ML004, ML006, ML006, and ML015) and clearly appear to be within an uncontaminated region of the FMGP site. The spherical semivariogram models were fit using a least sum of squared errors criteria implemented with the Solver routine in Microsoft Excel®. Note that while the lower CL data are more variable than the upper CL data (on the log scale) both show a range of influence of about 125 feet. This information may be useful during the on-site data analysis during ESC phase 2.

With a model of the semivariogram, we can use kriging, Bayesian updating (as described by equation 2.4), and the method developed by Stein to generate contour plots based on the logarithm of this data. Figure 3.8 shows contour plots of the logarithm of the CL data generated by ordinary kriging, Bayesian updating, and Stein's method for comparison. The Bayesian updating prediction was carried out using equations 2.4 under the assumption that the prior distribution had an unconditional mean equal to the median of the logarithm of the CL data with variance equal to the sill of the variogram (see Figure 3.7). From Figure 3.8 it is apparent that ordinary kriging and Bayesian updating produce nearly identical plots of the PAH distribution. Additionally, the method of Stein yields nearly identical contours as ordinary kriging and Bayesian updating in the region where the data are above the detection limit, however, Stein's method produces contours below the detection limit in the region where there are ND data. Figure 3.9 shows prediction standard deviations generated by ordinary kriging, Bayesian updating, and Stein's method for comparison. Again, from Figure 3.9 we see that ordinary kriging and Bayesian updating produce nearly identical results throughout the sample area. The prediction standard deviation plot generated by Stein's method is similar to the kriging and Bayesian updating plots in the region dominated by detect data, but is somewhat larger in the region dominated by ND data. This occurs because Stein's method
Figure 3.7. Isotropic semivariograms for the log(Chemiluminescence) data showing calculated points (+) and spherical model fit (solid line).
Figure 3.8. Contour plots of the upper log chemiluminescence for a) ordinary kriging, b) Bayesian updating, and c) Stein's method.
Figure 3.9. Contour plots of the upper log CL prediction standard deviation for a) ordinary kriging, b) Bayesian updating, and c) Stein's method.
"recognizes" the additional uncertainty due to the ND data, while kriging and Bayesian updating treat the ND data as though they are true observations.

The contour plots shown in Figure 3.8 can be used in the site contaminant conceptual model. To gain some appreciation for the uncertainty in these plots we can generate plots of the expected probability that a sample at a given location will exceed a given threshold. The Iowa Department of Natural Resources (IDNR) action level for total PAH (TPAH) in soils is 500 mg per kg of soil (500 mg/kg) and 100 mg/kg for carcinogenic PAHs. An analysis of PAH data contained in the IDNR Manufactured Gas Plant Data Management System (currently under development) from sites across Iowa reveals that naphthalene typically accounts for about 16 percent of the TPAH (based on 61 samples with detect measured values for all 16 PAH compounds). Because the CL data measures TPAH minus naphthalene, a comparison of the CL data with a threshold of 420 mg/kg (= 500(1 - 0.16) mg/kg) will provide some insight regarding contamination relative to the IDNR standard. Using the upper CL data, I have constructed three such plots using differing methods for comparison.

As mentioned previously (section 2.4.5), Stein's method will directly estimate the probability that the CL result will be less than or equal to the input threshold of 420 mg/kg (corresponding to TPAH ≤ 500 mg/kg); the probability of exceeding 420 mg/kg is one minus this. Under the assumption that the Z process follows a multivariate normal distribution with a spatial correlation characterized by a semivariogram or covariance function, ordinary kriging produces a prediction with a kriging error (true value minus predicted value) that follows a normal distribution with mean zero and a variance equal to the kriging prediction variance. Accordingly, under the joint distribution of \( \{Z(s_0), Z'\} \), where \( Z' = Z(s_1), \ldots, Z(s_n) \) are the observed values at locations \( s_i, i = 1, \ldots, n, \) we get the probability

\[
P(Z(s_0) \geq T | Z) = P \left( \frac{Z(s_0) - \hat{Z}(s_0)}{\sigma_k(s_0)} \geq \frac{T - \hat{Z}(s_0)}{\sigma_k(s_0)} \right) = 1 - \Phi \left( \frac{T - \hat{Z}(s_0)}{\sigma_k(s_0)} \right)
\]  

[3.1]
where \( \hat{Z}(s_o) \) and \( \sigma_k(s_o) \) are the ordinary kriging predicted value and prediction standard deviation at spatial location \( s_o \). \( T \) is a threshold value, and \( \Phi(\cdot) \) is the standard normal cumulative distribution function. This result can also be derived directly from the prediction interval for \( Z(s_o) \) developed by Cressie (1991, page 122).

The third probability plot considered here is based on indicator kriging whereby the data values are replaced by the indicator function \( I(s, T) \) at spatial location \( s \), defined by

\[
I(s, T) = \begin{cases} 
1, & \text{if } Z(s) \geq T \\ 
0, & \text{otherwise.} 
\end{cases} \tag{3.2}
\]

The indicator approach is described by Journel (1983) as a nonparametric method in that whatever the actual distribution of the variable of interest, the data are reduced to zeros and ones. The procedure is to perform ordinary kriging on the indicator values \( I(s_i, T) \) at sampled data locations \( s_i, i = 1, \ldots, n \). The predictor \( \hat{I}(s_o, T) \) developed from ordinary kriging gives an approximation to the probability \( P(Z(s_o) \geq T|I(s_1, T), \ldots, I(s_n, T)) \), (Cressie, 1991, page 282. Cressie actually uses an indicator function equal to one if \( Z(s) < T \), zero otherwise, and notes that ordinary kriging on the indicator values yields an approximation to \( P(Z(s_o) < T|\hat{I}(s_1, T), \ldots, \hat{I}(s_n, T)) \)). Note that the indicator values are invariant with respect to any monotonic transformation of the data and threshold value.

Ordinary kriging on the indicator values requires the development of an indicator semivariogram. Figure 3.10 shows the indicator semivariogram data based on equation 2.4 and spherical model for an isotropic indicator semivariogram for the indicator data developed from equation [3.2] with \( T = 420 \text{ mg/kg} \). Figure 3.11 shows the probability that the TPAH exceeds 500 mg/kg (corresponding to CL \( \geq 420 \text{ mg/kg} \)) estimated by each of the three methods discussed in the preceding paragraphs.

Probabilities between 0.2 and 0.8 are indicated by the shaded regions in Figure 3.11.

Examination of Figure 3.11a and 3.11b shows that the probabilities developed by Stein's method and ordinary kriging are nearly identical near the detect and ND data. The most notable difference between the Stein approach and ordinary kriging occurs near the extreme western edge of the site where there was no data collected; in this region the ordinary kriging approach gives greater
Figure 3.10. Isotropic semivariogram for upper CL indicator data (CL ≥ 420 mg/kg)

probabilities. The indicator results shown in Figure 3.11c are similar to Stein's method near the detect and ND data, but yield probabilities generally a little larger than the ordinary kriging approach in those regions where no data were collected. Which of these three methods is most believable depends in our belief in the representativeness of the CL data for the prediction of TPAH and on the degree to which the data actually follow a multivariate normal distribution. The indicator kriging approach may provide the better results in this case.

3.2.2 FMGP Site in Marshalltown, IA, Phase 2

The adaptive site selection method of David and Yoo (1993) assumes that sufficient data already exist to develop a "map" of the phenomenon of interest. This method does not incorporate information from other sources. Johnson's site selection method openly incorporates data from both the variable of interest and other soft information or hard data to choose subsequent sample locations, but is applicable only to variables, such as a contaminant concentration, for which it makes sense to model a probability of exceeding some threshold. The method developed by Englund and
Figure 3.11. Contour plots show estimated $P(\text{TPAH} > 500 \text{ mg/kg})$ based on the upper CL data (log scale). Sample locations shown as solid dots.
Heravi which incorporates sampling and remediation costs was also developed specifically to choose sampling locations for contaminants. Ultimately, how this data is used for regulatory decision making and risk analysis will depend on the regulatory authorities. While data other than analytical data for the variable of interest, such as soil sample visual or olfactory information, geophysical data, or screening data, might be used in the on-site sample location selection process, the assessment of when sufficient data have been collected should be made using only those data and procedures which are acceptable to the regulatory authorities.

3.2.2.1 Laser Induced Fluorescence and Geoprobe Electrical Conductivity, Phase 2

The SCAPS Laser Induced Fluorescence (LIF) system described by Koester et al. (1993), was used for both determination of soil stratigraphy and screening for PAH contamination. Post field work analysis of the LIF system (Stenback et al., 1994, see also Lieberman et al., 1993) indicates that coal tar derived contaminants have a strong fluorescence response to the 337 nm excitation light source used in the SCAPS LIF system. Additionally, the CPT soil behavior type classification was able to clearly distinguish between the major soil units and provided a greater level of stratigraphic detail than did the on-site continuously logged soil borings. Elevated LIF readings, indicating PAH contamination, were targeted as potential soil sample locations. Figure 3.12 shows output from a CPT push with LIF sensor probe from a location west of the former electric generation plant about 9 feet north of BVWST soil boring MW-4 (see Figure 3.1). Comparison of the soil boring log MW-4 and the CPT soil classification (Figure 3.12) shows generally good agreement with lateral offsets between soil unit contacts due to natural soil variations. Accordingly, soil unit contacts are relatively easy to pick from these plots. Post field work analysis of the LIF data indicate that this technology does not provide definitive contaminant concentration data, but is an excellent screening tool to locate PAH contamination in the subsurface (Stenback et al., 1995). LIF counts greater than 200 were found to be generally associated with total recoverable petroleum hydrocarbon (TRPH) concentrations greater than 500 mg/kg, LIF counts less than 100 were generally associated with
Figure 3.12. SCAPS CPT panel plot for push location ML-32 and adjacent BVWST soil boring MW-4 log.
TRPH less than 100 mg/kg, while LIF counts between 100 and 200 are associated with low to high levels of TRPH.

A Geoprobe electrical conductivity probe described by Christy, Christy, and Wittig (1994) was also used on-site to evaluate the effectiveness of this technology for both soil stratigraphy determination and potential for locating subsurface zones of PAH contamination. Post field work analysis of the Geoprobe electrical conductivity soundings indicate that the electrical conductivity trace versus depth was able to clearly delineate the soil strata (Stenback and Kjartanson, 1994). Because the soil electrical conductivity is a function of the soil type, porosity, water content, and pore fluid chemistry, it was not possible to determine the cause of a change in electrical conductivity without additional information. Nevertheless, we found that a dip in the electrical conductivity data in an otherwise relatively stable trace was frequently (about 75% of the time) associated with PAH contamination. Accordingly, dips in the conductivity trace provided some evidence of potential PAH contamination and thus were targeted as potential soil sample locations.

Figure 3.13 shows a Geoprobe soil electrical conductivity plot versus depth, together with a BVWST soil boring and CPT LIF trace taken about 10 feet away from a location just north of the storage shed (B-8 in Figure 3.1). As Figure 3.13 shows, the silty clay and silt soils are associated with elevated electrical conductivities (>80 mS/m) while the sandy soils tend to yield lower electrical conductivities (≤50 mS/m). Accordingly, soil unit contacts are relatively easy to pick from these plots. The LIF trace in Figure 3.13 shows evidence of PAH contamination directly overlying the LCU at about 35 feet below ground surface along with evidence of PAH contamination from about 3 to 23 feet below ground surface. Note that dips in the electrical conductivity trace within the sandy soil at 19 feet, 22-25 feet, and 35 feet generally correspond to elevated LIF counts with the elevated LIF counts occurring several feet higher in the soil profile possibly due to natural soil variations.

Because both the SCAPS LIF and the Geoprobe electrical conductivity gave indications of PAH contamination, subsequent soil sample locations for PAH analysis could be chosen on the basis of data from either of these tools together with the developing PAH contamination and/or geologic
Figure 3.13. Electrical conductivity, soil boring log, and LIF north of storage shed (Figure 3.1).
models. Because the SCAPS and Geoprobe electrical conductivity were used to gather information to develop both the soil stratigraphic model and the soil PAH contamination model, subsequent push locations may be determined by the application of the adaptive sample location methods described above.

3.2.2.2 Geology, Phase 2 Development of the site geologic model during phase 2 was based on the CPT and Geoprobe soil electrical conductivity. While the contacts between soil units (Figure 3.2) are gradational in some cases, they were relatively easy to pick off of the CPT and Geoprobe plots with an accuracy of about plus or minus one foot or less (see Figures 3.12 and 3.13).

As the geology, hydrogeology, and contaminant source impact the fate and transport of subsurface contamination, the need to define the geology of specific regions of the site is linked with the identification of areas of contamination and contaminant sources. In general, the goal was to define the soil strata underlying and bounding all regions of the site which showed evidence of PAH contamination. The surface topography of the LCU was of particular interest because this soil unit was expected to retard the downward migration of the coal tar DNAPL. To avoid opening a conduit in the LCU through which the coal tar DNAPL could migrate, all but several of the CPT and Geoprobe pushes were stopped within about two to three feet of the upper surface of the LCU. This was possible because the CPT and Geoprobe electrical conductivity data were available on a computer screen in real time as the push progressed so that the operator could see when the LCU had been penetrated and knew when to stop the push. In all cases where the LIF sensor reached the top surface of the LCU, the fluorescence intensity dropped to near zero indicating that the coal tar DNAPL was not migrating into the LCU.

I applied David and Yoo (1993) adaptive sample location selection technique to the selection of subsequent DPT locations to define the LCU surface topography to assess the method. The method first picks potential future sites from the area of interest which are farthest removed from the currently sampled sites, and then selects a site from among these candidates according to the minimum of maximum discrepancy functions as described in the preceding section 2.4.4. To
simulate this process, and be able to compare the results with the actual order in which the DPT sample locations were taken, I started with only the first seven sites (about one day of data) selected during the ESC demonstration. New sites were sequentially chosen according to the David and Yoo criteria from among the 66 sites actually selected in the ESC demonstration, but not yet selected in the simulation. Those sites not yet selected whose distance to nearest selected sites were within 95 percent of the maximum distance between sites already selected (i.e. this corresponds to $p = 0.95$ in equation 2.6) were chosen as potential next sample locations. There were between one and seven potential sites at each step of this process, with one, two, or three potential sites being most common. The next site was then chosen by the minimum maximum discrepancy described following equation 2.6 in section 2.4.4 above. The discrepancy function LCU maps were rectangular and covered most of the fenced area of the site.

I then applied the discrepancy function in equation 2.8 to the sample selection order according to the David and Yoo method and to the actual ESC demonstration order for comparison. Figure 3.14 shows these discrepancy functions. The discrepancy function for the David and Yoo selection order decreases relatively smoothly up to about 28 sample locations with no significant change beyond 28 sample locations. Because the discrepancy function compares the current map to a map generated with half of the current data, Figure 3.14 indicates that the major trends in the modeled LCU surface were reasonably well defined after about 14 DPT locations (one half of 28) using the David and Yoo sample order. The actual data selection order produced a discrepancy function which decreased rapidly up to about 14 sample locations, then increased and remained significantly above the David and Yoo selected sample location discrepancy from 17 to 33 sample locations, and finally remained relatively stable beyond 38 sample locations. Accordingly, the actual ESC LCU surface model was unstable until about 19 DPT results were available. This indicates that the David and Yoo method selected sample locations in an order which provided a rapid map convergence for the LCU surface elevation. Note however, that in the ESC demonstration, sample locations were selected to define PAH contamination as well as stratigraphy. In addition, several
closely spaced clusters of Geoprobe and CPT data were collected for the purpose of comparing these technologies. Figure 3.15 shows all 66 sample locations, along with the first 24 sample locations selected by both the David and Yoo method and the actual ESC demonstration. Comparison of Figures 3.15b and 3.15c indicates that the David and Yoo methodology tends to spread the sample locations out providing a more uniform coverage of the site.

While the discrepancy functions in Figure 3.14 indicate that the LCU surface was relatively well defined after about 14 (David and Yoo order) or 19 (actual order) DPT samples and that subsequent data added little to the refinement of the model, this information is not clear until the discrepancy function levels out at some "low" value. Even though the discrepancy plots level out beginning at about 28 samples (David and Yoo sample order) and 38 samples (actual order), in practice we might choose to observe additional samples until we are convinced that the surface is well defined. Use of a map discrepancy plot on-site would have indicated at about 40 to 50 sample locations that additional data was not adding much useful stratigraphic information so that the data collectors could have been discharged from this activity. However, the adequacy of sample

Figure 3.14. Discrepancy functions for LCU surface elevation modeling.
Figure 3.16. Map a) shows the 66 total LCU surface elevation sample locations. Maps b) and c) show the first 24 locations selected by b) David and Yoo method, c) ESC actual data. Points 2 and 3 lie to the east of the main map area.
coverage of the region of interest must also be considered. In the FMGP work, it was not possible to sample outside of the fenced property boundary (with the exception of sample locations 2 and 3 in Figure 3.15a collected along a dirt road leading into the FMGP site, or beneath the buildings and structures on the site. Accordingly, there is increased uncertainty in these areas.

As the investigation progresses, maps of various phenomenon can be plotted periodically. Maps of the LCU surface elevation were plotted on approximately a daily basis during the phase 2 FMGP investigation. Selected Hardy maps for days 2, 3, 5, and 10, corresponding to 14, 22, 36, and 66 sample locations, are plotted in Figure 3.16. Note that the general features of the LCU surface were defined relatively well after three days of sampling (compare day 3 with day 10 in Figure 3.16). Subsequent sampling after day three primarily identified localized depressions and elevated regions of the LCU surface. These plots indicate that at about the fifth day of sampling (36 sample locations), additional data was not adding much useful stratigraphic information so that the data collectors could have been discharged from this activity. Accordingly, these maps indicate convergence in agreement with the David and Yoo map discrepancy function shown in Figure 3.14. Note, however, that the distribution of samples selected during the ESC demonstration is not very uniform over the site at day five.

3.2.2.3 Quantitative Chemical Analyses, Phase 2 The CL and other screening data results and site history information may provide guidance on where soil samples should be collected to develop more definitive maps of the PAH contamination. While the EarthVision software can provide three-dimensional views of the site geology and contaminant distribution models, it is often easier to utilize and interpret two-dimensional slices in the form of cross sections and plan view surface or contour maps. All of the statistical tools including kriging, Bayesian updating, Stein's method, and David and Yoo site selection and discrepancy functions, can be carried out in two or three-dimensions whereby the variable of interest is modeled as a function of two or three space coordinates. In stratified deposits where the nature of the spatial dependence may change from one soil type to the next, modeling each stratigraphic layer independently is warranted. Early in the ESC
Figure 3.16. Daily lower cohesive unit surface elevation (ft above msl) with data locations shown. Contour interval = 2 feet.
program the FMGP site was divided into the six zones which correspond to stratigraphic position as shown in Figure 3.2 and contamination within each zone was considered separately.

The quantitative chemical analysis data for PAH contamination was concentrated to the east of the heavily contaminated region, leaving too few samples with results above the method detection limit (MDL) to adequately characterize the spatial distribution of PAH contamination. In addition, some of the heavily contaminated soil samples were excluded from the quantitative PAH analysis, while other samples were diluted five, ten, or 100 times prior to analysis, thus raising the method detection limit (MDL) to five, ten, or 100 times the nominal MDL level of 1 mg/kg for each of the 16 PAH compounds. These problems result in some difficulties in the treatment of this data for characterizing the contaminant distribution. To gain some additional data for my purpose here, I have included data from the BVWST RI off-site analyses of soil samples for total PAH (TPAH). There are ten zone 1 (middle of the upper cohesive soil unit at a depth of about 10 feet) soil samples from the BVWST RI and 24 ESC samples with TPAH values. The BVWST and ESC samples from zone 1 are shown in Figure 3.17. The BVWST samples are identified with a "B" or "MW" identifier while the ESC samples have an "ML" identifier and were selected in numerical order corresponding to the number following the "ML". As shown in Figure 3.17, the BVWST MDL for TPAH is 1 mg/kg while the McLaren/Hart (ESC) MDL for TPAH was 16 mg/kg.

Figure 3.18 shows the 34 zone 1 TPAH values plotted versus the north-south and east-west directions. The TPAH versus east-west direction plot shows the contamination to be located within the center of the site with non-detect values to the east and west. The TPAH versus north-south direction plot shows that the contamination extends nearly the breadth of the site, but appears to taper off along both the north and south site boundaries. While there are too few detect data values for a confident assessment, the dot plot in Figure 3.18 indicates that the distribution of the logarithm of the TPAH values within the TPAH plume might be modeled as normally distributed.

To create a basis to compare the sample location selection methods proposed by David and Yoo (1993), Englund and Heravi (1994), and Johnson (1993 and 1996a), I simulated a TPAH
Figure 3.17. ESC (ML#) and BVWST (B-# or MW-#) TPAH for zone 1.
Figure 3.18. East-West and North-South marginal distributions and dot plot for the zone1 TPAH data (n=34).
contaminant plume that passes through all 34 of the measured TPAH values using the variogram estimated from the phase 1 upper CL data (see Figure 3.7). A TPAH soil concentration value was generated sequentially at each node of a grid with ten foot spacing covering the fenced area of the FMGP site using the sequential Gaussian simulation algorithm given by Deutsch and Journel (1992). The simulated values were truncated at 16 mg/kg to simulate the actual data that might have been observed. The resulting TPAH spatial distribution is shown in Figure 3.19. Other methods for conditional simulation are given by, for example, Cressie (1991), Davis (1987), Deutsch and Journel (1992), and Marcotte (1995).

The sample location selection method of David and Yoo (1993) requires that some data be available to form a basis from which to choose next sample locations. While the methods of Englund and Heravi (1994) and Johnson (1993 and 1996a) can be implemented without any data, they are best implemented with some data from which to form a reasonable basis to determine a next sample location. A reasonable set of initial sample locations includes several samples taken from the suspected source location, or within the suspected (or known) region of contamination, and several samples taken from the region that is expected to bound the contaminant plume. Samples taken from the suspected or known source or contaminated areas will confirm right away whether there is, or possibly is not, a contamination problem at the site (a positive identification of contamination is generally established either during or prior to ESC phase 1). Once it is established that contamination does exist at a level which requires further investigation, then several samples taken from regions expected to bound the plume will be useful to provide some initial indication of the magnitude of the problem. The BVWST data satisfied both of these conditions and were used to initiate sample location selection for comparison of the three sample location selection methods described here.

3.2.2.4 David and Yoo Method On the basis of the CL data results and the BVWST data, the potential sample region was defined as easting 550 feet to 850 feet and northing 380 feet to 600 feet, with sample locations being restricted to be within the FMGP site fenced area and outside of the
Figure 3.19. Simulated soil TPAH concentration (mg/kg) contours for zone 1.
Figure 3.20 shows the results of applying the David and Yoo algorithm to sequentially select 30 additional data locations from the simulated TPAH data (shown in Figure 3.19) beyond the ten initial BVWST data. A value of $p = 0.98$ in equation 2.6 was used. This resulted in between one and seven, with one, two and three being most common, candidate sample locations for additional sample location selection according to the David and Yoo criteria described in section 2.4.4. As each new sample location was selected, it was included, along with the measured value at that location (from the simulated surface in Figure 3.19), in the dataset. The David and Yoo method was repeatedly applied to select additional sample locations. As can be seen from Figure 3.20, the David and Yoo algorithm selects sample locations which provide a nearly uniform distribution over the region of interest.

I performed a second, simplified, simulation similar to the David and Yoo algorithm by simply choosing the next sample location as that site which was farthest removed from all previously selected sites. This is equivalent to letting $p = 1$ in equation 2.6; if two or more sites were equally far removed from the previously selected sites, as happened several times, I simply chose one at random. Figure 3.21 shows the discrepancy functions for the two cases ($p = 0.98$ and $p = 1$ in equation 2.6). For the David and Yoo algorithm I generated data up to sample number 61 in Figure 3.21, while for the modified method, I generated data up to sample number 70 (the discrepancy values are nearly identical for some portions of the plot). From Figure 3.21 it appears there may be little advantage gained by implementing the full David and Yoo procedure. Simply selecting the sample location with greatest distance to the nearest existing sample (maximum minimum distance) appears to provide adequate site coverage for the generation of the map. This is potentially useful in that selecting the location with the maximum minimum distance is considerably faster than implementing the full David and Yoo methodology.

3.2.2.5 Englund and Heravi Method. I applied the Englund and Heravi sample selection procedure using the cost structure and loss function used by Englund and Heravi, as shown in Figure 2.6. While the cost structure clearly has an impact on the actual costs, the sample locations selected
Figure 3.20. David and Yoo sample location selection for 30 additional zone 1 soil TPAH samples starting with the BVWST data. Numbers show the order in which the sample locations were selected.
on the basis of this cost structure are independent of the (positive) cost of remediating a remediation unit ($10,000 in Figure 2.6). In this example I used a square remediation unit ten feet on each side. The method uses block kriging to obtain the predicted average and prediction standard deviation of the average TPAH concentration in each remediation unit. I used the block kriging equations given by Isaaks and Srivastava (1989, pages 324 to 326) to carry out the block kriging. I assumed the TPAH values are lognormally distributed, and that for small blocks (ten by ten feet), the block averages are also approximately lognormally distributed in accord with the discussion by Joumel (1980, page 292). With the average and standard deviation and assumed normal distribution of the natural logarithm of TPAH, the expected loss for both false positive and false negative errors can be determined by numerical integration. In this setting a false positive error is made when the decision is to remediate a site because the estimated average contamination exceeds the action level when the true average contamination is actually below the action level. A false negative decision is made when the remediation unit is declared to have an average TPAH concentration below the action level when the true average concentration is actually above the action level.

Figure 3.21. Discrepancy functions for sampling for zone 1 TPAH contamination.
On the basis of the cost structure shown in Figure 2.6, for a potential false positive decision (predicted block average is greater than the action level), the expected loss is

\[ E(\text{Loss}) = \int_0^{\frac{z}{AL}} f(z)D(1 - \frac{z}{AL})dz \]  

[3.3]

where \( D \) is the cost of remediating a unit (\( D = $10,000 \) in Figure 2.6), \( z \) is the block average TPAH concentration, \( f(z) \) is the probability density function of the average TPAH concentration, and \( D(1 - \frac{z}{AL}) \) is the loss function. This loss function occurs because a decision is made to spend \( D \) dollars to remediate the block when in fact the TPAH value, \( z \), might be less than the action level, \( AL \). If \( z < AL \), then the actual cost is \( Dz/AL \), and hence the loss is \( D - Dz/AL = D(1 - z/AL) \). For a potential false negative decision (predicted block average is less than the action level), the expected loss is

\[ E(\text{Loss}) = \int_{\frac{z}{AL}}^{\infty} f(z)D(z/AL - 1)dz. \]  

[3.4]

In this case the loss is the difference between the actual cost, \( Dz/AL \), and the cost of a correct decision which is to remediate at a cost of \( D \) dollars for the block. Thus the loss is \( Dz/AL - D = D(z/AL - 1) \). Note that the average TPAH for a block of soil is some fixed number (ignoring changes over time) but we don't know what that number is. We only know the TPAH values at the sampled locations. It is the stochastic model that allows us to predict the average TPAH concentration and prediction standard deviation for each block, which, together with some assumption about the probability distribution of the error, allows us to construct a probability density function, \( f(z) \), for the true average TPAH for each block.

Under the assumption that the natural logarithm \( Y = \ln(Z) \) of the soil TPAH concentration, \( Z \), is normally distributed, we can perform kriging on the logarithm of the data and obtain normally distributed errors, where the error is the actual value minus the kriged value. Monotonic transformations of a random variable will preserve percentiles, but will not generally preserve moments, such as the mean, variance, or skewness. The mean and median (50th percentile) of a
normal distribution are identical. Because the kriged block average, \( \hat{y}_g(s_o) \), of the natural logarithm of TPAH is an estimate of the mean (and median) of a normally distributed value, the back transform, \( \exp(\hat{y}_g(s_o)) \), is an estimate of the median, but not the mean, of the actual (data scale) block average TPAH. To estimate the mean on the data scale, we can use the backtransformation given by Journel (1980).

Under the transformation \( Y = \ln(Z) \), the probability density function \( f(z) \) in [3.3] and [3.4] can be related to the normal density function \( g(y) \) as follows (Mendenhall, Scheaffer, and Wackerly, 1981):

\[
f(z) = g(y) \frac{dy}{dz} = g(y) \frac{1}{z}.
\]

Since \( \frac{dy}{dz} = \frac{1}{z} \), \( dz = zdy \), \( f(z)dz = (g(y)/z)(zdy) = g(y)dy \). Thus, we can evaluate [3.3] as follows:

\[
E(\text{Loss}) = \int_{-\infty}^{\ln(AL)} g(y)D(1 - \exp(y))/AL \, dy.
\]  

[3.6]

Similarly, [3.4] can be evaluated as

\[
E(\text{Loss}) = \int_{\ln(AL)}^{\infty} g(y)D(\exp(y)/AL - 1) \, dy.
\]  

[3.7]

Equations 3.6 and 3.7 are easy to integrate numerically because \( g(y) \) is a normal probability density function with mean \( \hat{y}_g(s_o) \) and variance equal to the block kriging variance.

At each step, the next best sample location is within that block which has the largest expected loss. Figure 3.22 shows the 30 sample locations selected using the ten BVWST data to start the sample location selection process. Note that this algorithm selects sample locations which attempt to define the action level, AL (compare Figure 3.22 with Figure 3.19). This occurs because regions with high TPAH values are nearly certain to yield block averages above the AL (assuming we
Figure 3.22. Englund sample location selection with AL = 500 mg/kg. Numbers show selection order for the first 30 samples with the initial ten BVWST data locations shown as open boxes.
have a reasonably representative dataset to work with) and have decision error rates with low expected losses, even if the uncertainty is relatively large. Likewise, regions with very low TPAH values are nearly certain to yield block averages below the AL with corresponding low decision error rates and low expected losses. But regions with predicted TPAH values near the AL may have a greater chance of being misclassified (false positive or false negative decision error) and have greater potential for a high expected loss. The strong tendency to cluster data near the AL level may occur with this simulation because the TPAH surface is relatively smooth. With a more erratic contaminant plume, the sample locations would likely be spread out to a greater extent.

The manner in which I implemented the procedure called for sample selection only at grid nodes spaced ten feet apart from Easting 550 feet to 850 feet, and northing 380 feet to 600 feet. One of the first 50 sample locations selected called for resampling a block that had already been sampled. In this case I selected a sample randomly located within that block and assigned it a value equal to the average of the four closest grid nodes. Englund and Heravi suggest that the next sample location be located at random within the block with the greatest expected loss at each sampling event.

We can also calculate the expected total cost as

\[ E(\text{Total Cost}) = (\text{Sample and Analytical Cost per Sample})(\text{Number of Samples}) + (\text{Remediation Cost per Unit})(\text{Number of Units with } Z_B > \text{AL}) + \sum_{\text{Units Not Remediated}} E(\text{Cost of Not Remediating Units with TPAH Contamination}) \quad [3.8] \]

where \( Z_B \) is the estimated block or unit average. The accuracy of the expected total cost calculation depends on how well the cost structure model fits reality. If the cost model fits reality well, then this calculation can be used to project remediation costs and might be useful to justify additional sampling or provide a stopping criteria. However, care must be taken in developing and applying such equations. For example, the IDNR action level used to initiate clean-up of PAH contaminated soils is 500 mg/kg TPAH, and 100 mg/kg CPAH (carcinogenic PAHs). However, once a decision to
initiate clean-up is made, the actual level to which in place soil may be left untreated may depend on factors such as economics, risk analysis, and whether the soil will impact groundwater quality or not (Golchin, 1996). Allowable TPAH concentrations for soil left untreated may range from near ND to 500 mg/kg, with additional requirements for CPAH levels. Adequately forecasting the amount of soil that might require treatment may require an accurate map of the plume beyond just that portion of the plume which is near the action level. Of course, the nature of the contaminant plume with depth must also be considered. This can be done by developing the TPAH (and CPAH) plume models within each stratigraphic zone in a manner similar to the analysis given here.

The expected total cost, $E(\text{Total Cost})$, for the cost structure shown in Figure 2.6, with the cost of sampling and analytical work taken as $1000 per sample, is shown in Figure 3.23. With successive sampling, the expected total cost generally declines primarily due to reduced prediction standard deviation effectively lowering the expected cost of failing to remediate contaminated units. The large reduction in expected total cost at sample number 18 (see Figure 3.23) is due to placement of sample 18 in the far northeast corner of the area considered for subsequent samples (see Figure 3.22). This is a result of the kriging process; kriging to a spatial location far removed from the data produces a prediction which is near the average of the data and has a large prediction standard deviation. Thus, prior to sample 18, block averages in the northeast corner were near the data average (337 mg/kg) with large standard deviations, resulting in large expected false negative losses. Sample 18 was a ND which essentially eliminated the northeast corner from further consideration for sampling by lowering both the predictions and prediction standard deviations in this region.

To the extent that this cost structure is meaningful for this FMGP, Figure 3.23 indicates that after about 43 samples, the expected total cost begins to level off at about $2.35 million. For this simulated data, there are 107 blocks with average TPAH > 500 mg/kg; at a unit remediation cost of $10,000, this amounts to $1.07 million of the total cost. The difference, $1.28 million, is largely associated with not remediating soils with TPAH contamination below 500 mg/kg because of the cost.
which is linearly proportional to the actual contaminant concentration; according to the cost structure these soils incur a cost of $D/AL ($10,000/500) times the TPAH concentration of the block. At a sampling cost of $1000 per sample, collecting 30 or even 50 samples amounts to less than 2.2 percent of the total cost. As a side note, the cost of not remediation low level contaminated soils (soils contaminated below the action level of 500 mg/kg) is based on the assumption that the danger posed by these soils is proportional to the actual concentration, which translates into twice as much is twice as costly (Englund and Heravi, 1994). To the extent that this is true, it supplies support for clean-up of soils with in-situ contamination well below the action level.

It is not my desire to promote this cost structure here; I only wish to illustrate the methodology. Careful consideration must be given to the development of a reasonable cost structure for a given application by considering items including different remediation methods, construction and operating costs, cost of future non-compliance, mobilization costs, sampling and analytical costs, remediation goals, regulatory concerns, and risk analysis and the expected cost to society. Some of these are unknown and may be difficult to estimate. In addition, some cost may be directly related to where the contamination is found, e.g. remediating soil under an existing building, such as the FMGP former purifier building, may incur additional expenses not associated with remediation of soils easily accessible from the surface.

3.2.2.6 Johnson's Method The method developed by Johnson (1993 and 1996a) starts by assigning prior beta probability distributions to each node of a grid over the site that define the probability, $\pi$ ($0 < \pi < 1$), that contamination exceeds some threshold, such as the action level. In other words, the unknown probability that the contaminant concentration exceeds the threshold is modeled as a random variable following a beta probability distribution with parameters $\alpha$ and $\beta$. The initial (prior) probability estimates can be made on the basis of "soft" data, such as geophysical data or site history describing the nature of the problem, or they can be constructed from "hard" analytical data. The parameters, $\alpha$ and $\beta$ ($\alpha > 0$ and $\beta > 0$), of the beta distribution are defined so that the expected value (mean) of the beta distribution at grid node $i$ is
Figure 3.23. Expected total cost (equation 3.8) for Englund and Heravi sample selection.

\[ E(\pi_j) = \frac{\alpha_i}{\alpha_i + \beta_i} \]  \hspace{1cm} [3.9]

and are chosen to reflect the degree of uncertainty according to the variance of the beta distribution

\[ \text{Var}(\pi_j) = \frac{\alpha_i \beta_i}{(\alpha_i + \beta_i)^2(\alpha_i + \beta_i + 1)} \]  \hspace{1cm} [3.10]

Small values of \(\alpha\) and \(\beta\) give a large variance (large uncertainty) and large values of \(\alpha\) and \(\beta\) give a small variance (small uncertainty). Once \(\alpha_i\) and \(\beta_i\) are defined for each node \(i\), they can be updated on the basis of indicator kriging on hard data as follows:

prior \(\alpha_i\) becomes posterior \(\alpha_i + X_i\) \hspace{1cm} [3.11]

prior \(\beta_i\) becomes posterior \(\beta_i + N_j - X_i\) \hspace{1cm} [3.12]

where

\[ N_j = \frac{2C_{so}}{\sigma_k^2} - 1 \text{ and } X_i = \hat{Z}(s)N_j, \]  \hspace{1cm} [3.13]
and $\hat{Z}(s)$ and $\sigma_k^2$ are the indicator prediction and variance, respectively, at node \( i \). At each new sampling event, which may consist of one or more samples, the existing prior parameters are updated to the posterior parameters according to [3.11] and [3.12]. These posterior parameters then become the prior parameters for the next sampling event.

Johnson suggests that large probabilities, say $E(\pi_i) > 0.8$ (or 0.9), indicate high likelihood that a classification of contamination being present is correct, and that low probabilities, say $E(\pi_i) < 0.1$, indicate high likelihood that a classification of contamination below the threshold is correct, while probabilities between these endpoints might be classified as state uncertain. Accordingly, one might sample locations that would be expected to maximize the number of decision points that would be classified as "contaminated", or to maximize the number of decision points that would be classified as "clean", or to minimize the number of decision points classified as state uncertain. In practice, classification of error rates should be based on the consequences of the error, with lower error rates given to the more severe type of error. However, in the spirit of managing uncertainty, we (the decision authority) must realize that we will rarely be able to achieve as low of error rates as desirable (we may rarely, if ever, achieve a zero error rate), and the lower we set error rates, the more data we must collect in order to achieve them.

I implemented this procedure using the indicator probability map for TPAH $> 500$ mg/kg developed from the CL data as shown in Figure 3.11c, i.e. using the indicator values at the 10 foot grid node spacing from which Figure 3.11c was developed as the prior probabilities, $E(\pi_i)$. I classified each grid node $i$ on a ten foot square grid as state uncertain if $0.2 < E(\pi_i) < 0.8$. At each node, prior parameter $\alpha_i$ was set at $\alpha_i = 0.01$, and prior parameter $\beta_i = \alpha_i/E(\pi_i) - \alpha_i$ was calculated so that $E(\pi_i) = \alpha_i/(\alpha_i + \beta_i)$ was equal to the prior kriged indicator values from Figure 3.11c. I used the ten zone 1 BVWST TPAH samples as an initial sampling event. Subsequent samples were selected sequentially at locations which were expected to minimize the number of nodes classified as state uncertain. This may be accomplished by examining each grid node, $i = 1,...,G$, in turn. While
considering grid node $i$, we assign the node a value of 1 if $E(n_j) \geq 0.5$, and 0 otherwise, and include this node in the existing indicator data based on the indicator values

$$I(s_j, T) = 1 \text{ if } Z(s_j) \geq T, \text{ and } I(s_j, T) = 0 \text{ otherwise,}$$

with $T = 500 \text{ mg/kg}$. Then we perform indicator kriging on this data, including the fabricated data at grid node $i$. At each node $j = 1, ..., G$, we determine $N_j$ and $X_j$ according to [3.13], and update the parameters $\alpha_j$ and $\beta_j$ according to [3.11] and [3.12]. Then determine, and store, the number of grid nodes classified as state uncertain. After doing this for each node $i$ we choose that node which gave the minimum number of nodes classified as state uncertain as the next sample point. This point is sampled and added to the existing data, and the process is repeated to select the next sample location. A flowchart of this process is given in Figure 3.24. Note that a relatively simple modification of the procedure, as shown in Figure 3.24, can be implemented to maximize the number of sites classified as either clean or contaminated.

The sample location selection process developed by Johnson can be somewhat slow for large grids (several thousand or more nodes) and large data sets (over one hundred observations). To speed the process, Johnson (1996b) has suggested several strategies including restricting the search area, searching over a coarse grid, such as every other grid node, using a nested grid search, or specifying a list of potential locations to choose from. In the simulation of the sample location selection that I performed for this work, I restricted the grid to Easting values of 580 feet to 800 feet, and northing values from 400 feet to 600 feet (compare with Figure 3.19) using a 10 foot grid node spacing. I further restricted the potential sample locations to be only those grid nodes which were classified as state uncertain according to the most current model at each step of the sample location selection process.

For comparison, I performed a somewhat similar analysis using only indicator kriging. Using the ten zone 1 BVWST TPAH data to start the process, I used indicator kriging with the indicator function as defined in [3.14] to estimate the probability that the TPAH concentration at each grid node is greater than or equal to 500 mg/kg. I selected the next sample location from within the state...
Define grid nodes, $s_i, i = 1, \ldots, G$, to cover the region of interest.
Define the threshold, $T$, usually the action level or clean-up level.
Define the State Uncertain criteria, e.g., $p_1 < E(s) < p_2$.
Define indicator values for existing "hard" data, $I(s, T) = 1$ if $Z(s) \geq T$, 0 otherwise.

Set $i = 1$

Fabricate indicator data at node $s_i$ as follows:
Assign 1 to grid node $s_i$ if $E(s_i) \geq 0.5$, 0 otherwise.

Concatenate fabricated data (last step) to "hard" indicator data.
Indicator krig to all grid nodes, $s_j, j = 1, \ldots, G$.
Use indicator prediction and prediction variance in equations 3.11 and 3.12 to get $N$ and $X$ according to equation 3.13.
Update $\alpha$, $\beta$, and $E(s_i)$ at each node.

Determine the number of nodes in the grid classified as state uncertain.
Store this number.
(Alternatively, store the maximum number of nodes classified as clean, or classified as contaminated)

Set $i = i + 1$

Is $i \leq G$?

Yes

No

Select next sample at the node that gives the minimum number of nodes classified as State Uncertain.
(Alternatively, select the node giving the maximum number of nodes classified as clean, or classified as contaminated)

Figure 3.24. Flowchart for Johnson sample location selection to minimize the state uncertain area.
uncertain region, defined as $0.2 < i(s,T) < 0.8$, at the grid node with the largest kriging variance, $\sigma_k^2$, where $i(s,T)$ is the indicator kriging predictor. The intent is to reduce the size of the state uncertain region by sampling where uncertainty, as measured by the kriging variance, is greatest. After each sampling event, indicator kriging was performed on the new dataset and a subsequent sample was selected from the new state uncertain region where the kriging variance was greatest. This procedure is much quicker than the site selection method developed by Johnson. I refer to this procedure as the uncertainty zone sampling method in the following text. An alternative, perhaps preferable, variation of this method is to use ordinary kriging (equation 3.1) or Stein's method to generate the probability plot to define the state uncertain region.

The sample locations for an additional 30 samples selected by both the Johnson and uncertainty zone sampling processes described in the preceding paragraphs are shown in Figure 3.25. From Figure 3.25a it is apparent that the sample locations selected by Johnson's procedure tend to concentrate near the threshold boundary (compare with Figure 3.19). From Figure 3.25b we see that the sample locations selected by the uncertainty zone sampling procedure tend to surround the region near the threshold boundary and provide a more uniform coverage of the contaminated soils than do the Johnson sample locations.

To further compare the Johnson and uncertainty zone sample location selections, the Johnson posterior probability, $P(\text{TPAH} \geq 500 \text{ mg/kg})$, and uncertainty zone indicator prediction for the probability, $P(\text{TPAH} \geq 500 \text{ mg/kg})$ are plotted in Figure 3.26. Note that the Johnson posterior and uncertainty zone probability plots in Figures 3.26a and 3.26b show two main differences. In the southwest corner of the contoured region, the uncertainty zone sampling probability contours indicate contamination is likely, while the Johnson method, which did not collect data in this region, does not. In this region, the uncertainty zone method matches the "true" plume better (see Figure 3.19). The second main difference occurs immediately east of the FMGP former purifier building where the uncertainty zone sampling probability contours lie largely beneath the building but the Johnson contours extend further to the east. This occurs because Johnson's data point 7 (see Figure 3.25a)
a) First 30 samples selected by Johnson's method.

b) First 30 samples selected by the uncertainty zone sampling method (see text).

Figure 3.25. Comparison of sample location selection by a) Johnson's method and b) uncertainty zone sampling. Numbers show the selection order. BVWST data shown as open boxes.
a) Johnson's method posterior Beta distribution mean, $P(TPAH>500 \text{ mg/kg})$.

b) Uncertainty zone sample selection $P(TPAH>500 \text{ mg/kg})$.

**Figure 3.26.** $P(TPAH>500 \text{ mg/kg})$ by a) posterior Beta distribution and b) uncertainty zone sampling.
gave a TPAH result above 500 mg/kg, while indicator data point 3, which lies to the west of Johnson's point 7 (see Figure 3.25b), gave a TPAH result below 500 mg/kg. Johnson's data matches the "true" plume better in this region (compare with Figure 3.19).

3.2.2.7 Sampling Method Comparisons To compare the performance of the David and Yoo, Englund and Heravi, Johnson, and uncertainty zone sample location selection methods, I computed the discrepancy functions for the sequential surfaces generated by comparing each method with the "true" TPAH surface illustrated in Figure 3.19. The map surfaces were generated using the MQ-B interpolation on the raw data as collected by each method. Because the TPAH MDL is 16 mg/kg, interpolated values less that 16 were set equal to 16. The resulting discrepancy plots are shown in Figure 3.27. The uncertainty zone sampling method appears to provide the best match, as judged by its low discrepancy values, most likely because this method provided the most uniform distribution of samples over the plume. While the David and Yoo method provides the most uniform sample distribution over the entire site, it places too many samples in uncontaminated areas resulting in the poorest match with the true surface, as judged by its large map discrepancy for most sample sizes shown in Figure 3.27. This problem with the David and Yoo method could be at least partially overcome by restricting the search area to include only that region believed to encompass the plume, thus providing a more uniform sample distribution over the plume. Because the Englund and Heravi method and the Johnson method both concentrate data near the action level, they do not provide sufficient data to produce a map closely matching the entire TPAH distribution.

To further compare the performance of the sample location selection methods, I performed sampling simulations to select 40 total samples for each method as discussed in the preceding paragraphs, however, with only two BVWST data values to start the sampling process. One of the two BVWST starting TPAH values lies within the plume with a value of 1868 mg/kg (greater than 500 mg/kg, the action level used in the sample selection process), and the other lies outside of the plume with a value of 1 mg/kg (less than 500 mg/kg). Samples were selected from the simulated TPAH
plume shown in Figure 3.19. The three Johnson method criteria were applied: sampling to maximize the number of nodes classified as clean, to maximize the number of nodes classified as contaminated, and to minimize the number of nodes classified as state uncertain. The England and Heravi method was applied in accordance with the cost structure in Figure 2.6 and using a cost structure with a false negative loss equal to double the value associated with Figure 2.6. Doubling the false negative loss represents an attempt to reduce uncertainty associated with remedial unit averages predicted to be below the action level but which might actually be above the action limit. Finally, the uncertainty zone sampling method using both indicator kriging and ordinary kriging to estimate the probability that TPAH exceeds 500 mg/kg was applied to generate sample locations. The sample location results are shown in Figures 3.28 and 3.29.

Discrepancy functions comparing the average absolute difference between the MQ-B map based on the sample locations generated by each method and the "true" TPAH surface (from Figure 3.19) are shown in Figure 3.30. Assuming the goal of the site characterization is to define the TPAH
a) Johnson method: maximize state clean criteria.

b) Johnson method: maximize state contaminated criteria.

c) Johnson method: minimize state uncertain criteria.

d) David and Yoo method.

Figure 3.28. Application of sample selection methods to the simulated TPAH surface starting with BVWST B-2 (1) and B-14 (2).
Figure 3.29. Application of sample selection methods to the simulated TPAH surface starting with BVWST B-2 (1) and B-14 (2).
plume, then for this simulation, Figure 3.30 indicates that the ordinary kriging based uncertainty zone sampling appears to provide generally good results over all sample numbers. The Johnson maximize state clean criteria yields data which do not give an MQ-B map that closely matches the simulated plume because many of the data locations are selected outside of the plume (see Figure 3.28a). Beyond sample number 31, the Johnson method based on the maximize state contaminated criteria provides the closest MQ-B map to the simulated surface; this is because the majority of these data are collected from within the plume, however, this leaves the plume edges less well defined (see Figure 3.28b). Doubling the England and Heravi false negative loss results in a poorer plume definition than the unaltered Englund and Heravi method. The Johnson minimize state uncertain, England and Heravi, David and Yoo, and indicator kriging based uncertainty zone sampling provide generally comparable results beyond about sample number 20.

To examine the effect of the prior on the Johnson sample location selection algorithm, I shifted the CL data prior used in the preceding analysis 20 feet to the west and then applied the Johnson algorithm. I started with the same BVWST points and selected an additional 38 samples using each of the three Johnson criteria. The resulting data locations are shown in Figure 3.31. The discrepancy functions for each of these datasets is shown in Figure 3.32, along with the map discrepancy for the ordinary kriging based uncertainty zone sampling method (also shown in Figure 3.30) for comparison. Examination of Figures 3.31 and 3.32 clearly show, for this example, that a bad choice of prior results in a poor selection of samples for the purpose of defining the contaminant plume. The maximize state clean and minimize state uncertain criteria perform poorly because they fail to detect the heavy contamination immediately west and south of the former purifier building (compare Figure 3.19 with Figures 3.31a and 3.31c). The maximize state contaminated criteria performs poorly because it fails to sample just east of the former purifier building leaving the east edge of the plume completely undefined (see Figure 3.31b). The uncertainty zone sampling method appears to provide good data upon which to model the contaminant plume.
Figure 3.30. Map discrepancy for sampling from the simulated TPAH plume with two initial BVWST sample locations.
Figure 3.31. Application of sample selection methods to the simulated TPAH surface starting with BVWST B-2 (1) and B-14 (2) using the CL prior shifted 20 feet to the west.
Figure 3.32. Map discrepancy for sampling from the simulated TPAH plume with two initial BVWST sample locations and the CL prior shifted 20 feet to the west.
The preceding analysis considered only a portion of the contaminated soils at this site. In practice, each of the zones must be considered. Commonly, once a sample location is selected, samples are taken from several or all of the depth zones of interest at that location. Accordingly, overlays of the two dimensional plots can be used to assess the spatial distribution versus depth. Alternatively, the three dimensional images available with EarthVision are ideal for this purpose. Each soil stratigraphic contact can be modeled in a manner similar to the LCU above. Isopach maps indicating soil unit thickness, and possible discontinuities, can be modeled wherever soil unit thickness data is available in sufficient quantity. The coal tar DNAPL might be expected to spread laterally somewhat as it migrates down into the subsurface soils, and would be expected to flow horizontally down the gradient of the LCU if a sufficient quantity of DNAPL is present. This type of soft information can be used to develop prior probability maps for the deeper stratigraphic zones to initiate a sampling scheme like that used by Johnson.

In addition to the measured PAH contamination, the CPT LIF and Geoprobe soil electrical conductivity data are other sources of information on which the placement of samples for PAH analysis can be based. These CPT and Geoprobe units were deployed to provide information on the geology and contamination. This information provides data to enhance the conceptual geologic model and suggest locations to sample for contamination. The geologic model itself might indicate a need to collect a sample at a certain location, independent of, or in conjunction with, the CPT or Geoprobe contaminant screening tools. Accordingly, the statistical techniques discussed in this section might at times play a dominant role in the selection of sample locations, while at other times the developing conceptual model and/or contaminant screening data will suggest areas where samples should be collected.

Statistical techniques can be utilized to insure a more uniform coverage of the site as a whole, and to focus on, and delineate, areas of contamination which will require remediation. A simple graphical way to show the maximum distance to the closest data points, i.e. the set \( T \) defined by David and Yoo in equation 2.6, is to plot circles centered on the data points of progressively
smaller radius as more and more data are collected. Such a plot is illustrated in Figure 3.33. This plot shows the area within a 50 foot radius of each sampled point in gray. The 50 foot radius was chosen to provide only a small region of potential sample sites (the unshaded area) in accord with p near one in equation 2.6 of the David and Yoo algorithm. Those areas not covered by gray are good candidate sample locations from the standpoint of producing a more uniform coverage of the site.

While the methods of Johnson and England and Heravi can define the boundary of a plume at some desired threshold, they might not provide sufficient data to clearly define the spatial distribution of the contaminants in regions with greater or lower contaminant level than the threshold used in the sample location selection algorithm. Since the actual cleanup levels may be considerably lower than the action level, it may be necessary to map the entire region of contaminated soils so that a proper assessment of applicable remediation technology and volume of soils needing treatment can be made. Accordingly, it is vital to understand from the project outset, all of the uses to which the data will be put so that a sufficient quantity of the right data is collected, i.e. the DQOs should be clearly defined prior to the collection of any data. Early sampling for contamination may focus on finding source areas, and areas of high contamination. Once it has been firmly established that soils at the site exceed the action level, the focus might shift to defining the plume edges at a level which is near the expected cleanup level. While the TPAH data illustrated in Figure 3.17 are poorly distributed for the purpose of defining the majority of the contaminant plume, they clearly indicate that there are soils in excess of the IDNR 500 mg/kg TPAH action level underlying the central portion of the FMGP.

Figure 3.34 shows an indicator semivariogram developed with T = 100 in equation 3.2 based on TPAH data within the apparent region of the expected plume from easting 500 feet to 800 feet. Based on this indicator semivariogram, a contour plot to estimate the probability that the soil TPAH concentrations exceed 100 mg/kg was developed using ordinary kriging on the indicator data. Figure 3.35 shows this plot overlaid on a plan view of the site together with gray areas indicating a 25 foot radius (one third of the semivariogram range of influence) from those points already sampled.
Figure 3.33. Plan of FMGP site complete 10 day sampled LCU locations shown. The gray area covers a 50 foot radius from the sampled locations.
Figures 3.34. Isotropic semivariogram for zone 1 TPAH indicator data (TPAH ≥ 100 mg/kg).

Probabilities less than 0.1 and greater than 0.8 are also indicated by grayed areas. Those areas not shaded in gray are good candidate locations for future samples. A figure such as Figure 3.35 can generally be produced faster than application of any of the Johnson, Englund and Heravi, or David and Yoo sample selection algorithms. The graphical technique illustrated in Figure 3.35, which I have called graphical uncertainty zone sampling, can provide an effective aid to sample location selection, particularly in conjunction with other information, such as LIF data, in the manner used in the ESC process, specifically, where people, and not computers, make the decisions. Further development of the graphical uncertainty zone sampling method is given in Chapter 4.

Modeling multiple zones for contamination and stratigraphy can be a time consuming task. Attempting to apply the slower statistical techniques, such as the methods of Johnson, David and Yoo, or England, to multiple regions of contamination, with multiple contaminants of concern, may be too time consuming to carry out effectively while on-site. In addition to the statistical routines themselves, the assumptions on which they are based must be considered. Assessing the statistical distribution of the population from which the data are drawn may be vital for a proper use of the sampling and modeling algorithms. We also need to consider the appropriateness of the choice of
Figure 3.36. ESC and BVWST TPAH for zone 1 with estimated probability that TPAH exceeds 100 mg/kg. Gray areas show 25 foot radius from sample locations and probabilities < 0.2 or > 0.8.
algorithms, i.e. is the remediation unit used in the Englund and Heravi algorithm appropriate for the problem at hand? Alternatively, the MB-Q method does not require any distributional assumptions or a spatial structure model (variogram or covariance), and will begin to show convergence to a spatial phenomena when the data are sufficiently dense that their spacing becomes less than the range of influence of the underlying phenomena. However, as seen previously, the David and Yoo algorithm may place too much data in regions where it is not needed.

We should consider the choice and development of the covariance model used; in the preceding analysis, the phase 1 CL data provided a basis for the variogram and a preliminary assessment of the statistical distribution. In some ESC projects, preliminary data may not be available so that a covariance model and distributional assumption may need to be made in the absence of data. In practice, as each 10 or 20 additional data values are obtained, variograms may be updated and distributional assumptions checked. During the initial few days on site, the data collected may be too few or too far apart to accurately model a variogram. To assess the nature of the variogram at low lags, several closely spaced sample locations are needed; a good time to collect such data is early in the program, such as during the DPT calibration program, so that an accurate variogram model may be developed to enhance the utility of the statistical methods used.

One way to represent three dimensional phenomena is to produce small profiles of a response versus depth overlain on a plan view map of the site. Figure 3.36 shows the LIF data plotted versus depth for each of the CPT LIF pushes overlain on a plan view of the FMGP site. Plots of this nature are useful during the phase 2 investigation to aid the development of a three dimensional picture of the site, and to aid the selection of future sample locations. Additionally, computer generated "three-D" plots displaying geologic or contaminant data can be generated and rotated to a variety of viewing angles to aid the interpretation of the data. However, computer generated "three-D" plots can be misleading if the user is not familiar with the algorithms used to generate the images and the practical limitations associated with these algorithms, particularly for extrapolations outside the range of the data.
Figure 3.36. CPT LIF fluorescence count data versus depth (depth is vertical, fluorescence truncated at 1000 counts).
The plot in Figure 3.36 shows all the LIF data collected at the site and thus, within the limitations of the small size of the individual LIF traces, represents a relatively complete way to illustrate this data. Figure 3.36 clearly indicates that subsurface contamination extends further to the northwest than indicated by the zone 1 TPAH data (see for example LIF traces L93 and L111 in Figure 3.36, compare with Figure 3.35). Figure 3.36 also shows the depth profile of LIF "hits" within the UCU, the top of the granular soil unit, and overlying the LCU surface at the base of the granular soil unit throughout the general area where contamination is indicated by indicator kriging for TPAH ≥ 100 as shown in Figure 3.35.

3.2.3 FMGP Site in Marshalltown, IA, Post Phase 2

Once the field data have been collected, some summary analysis of the data is made to update the conceptual model, determine optimal monitoring well locations, and provide input for a risk assessment and remedial design. Contour maps, surface maps, and cross sections are common methods of displaying geologic, hydrogeologic, and contaminant models. What may not be shown in some RI reports are plots which indicate the degree of uncertainty in the geologic, hydrogeologic, and contaminant models. Prediction standard deviation plots or probability maps can be useful for this purpose. Additionally, conditional simulations might be made whereby many possible surfaces, each passing through the data points and possessing the covariance observed in the data, might be useful. Contour maps together with prediction standard deviation plots, probability plots, or conditional simulations can be used to develop a range of likely surface models (potential deviations) for the development of a robust remedial resign which will perform under the expected variation of conditions. Alternatively, the decision analysis framework developed by Freeze et al. (1990) utilizes such conditional simulations to incorporate uncertainty and help choose between remedial designs, or suggest that more data of a certain type should be collected before a remedial design alternative is adopted.

During the post phase 2 data summarization, covariance and/or variogram models should be
updated, assumptions should be checked, and a variety of maps, cross-sections, and overlaid images should be developed to enhance the understanding of the conceptual model, and illustrate its uncertainties. For example, the LCU surface elevation plots illustrated in Figure 3.16 show an upward trend toward the northeast corner of the FMGP site. This trend could probably be accounted for with a linear model in both the east-west and north-south directions, i.e. write the LCU surface elevation $Z(s)$ as

$$Z(s) = \mu(s) + \delta(s), \quad [3.15]$$

where $\mu(s)$ is a linear function of the easting and northing values representing the trend, and $\delta(s)$ is a correlated error process which can be modeled by a variogram. Stenback et al. (1996) have modeled this surface using the median polish algorithm described by Cressie (1991). Details of the algorithm are given by Cressie (1991, pages 186-192) and are not repeated here. Basically, the median polish algorithm generates a coarse grid over the site, estimates a row effect, a column effect, and an overall term which are used to characterize the trend $\mu(s)$ in [3.15]. Residuals are defined as the observed data minus the median polish trend surface. These residuals are then modeled using ordinary kriging, a kriging surface of the residuals is generated, and the median polish trend surface is added back to the data to obtain the final model.

Using the 66 data points shown in Figure 3.15a, a coarse grid with nodes defined by the intersection of easting values 425, 540, 655, 770, and 887 feet and northing values 375, 457, 539, and 621 feet was defined, and the median polish algorithm was applied to this grid. The resulting median polish trend surface and kriged LCU residuals plus median polish surface are shown in Figure 3.37. The apparent southerly trending trough at easting 540 feet, south of northing 500 feet, occurs in a region where there is no data and is therefore due solely to the median polish surface.

Figure 3.38 shows a normal probability plot of the median polish residuals whereby the residuals plot nearly along straight line indicating that a normal distribution is a reasonable choice for the marginal distribution of residuals. An isotropic semivariogram of the median polish residuals is shown in Figure 3.39. A spherical model was fit to the semivariogram data for lags greater than 23
Figure 3.37. Kriged LCU residual plus median polish surface (a) and median polish trend surface (b).
Figure 3.38. Normal probability plot for the median polish residuals.

Figure 3.39. Isotropic semivariogram for the LCU median polish residuals.
feet. For lags less than 23 feet, a linear semivariogram that is continuous with the spherical model at 23 feet and passes near the semivariogram data point at 5 feet was developed.

I examined selected directional semivariograms (in the N-S, E-W, N45°E, and N45°W directions) to look for evidence of anisotropy. The directional semivariograms show considerably more scatter than the isotropic semivariogram and all but the N45°W semivariogram indicate that greater variation than was estimated for the isotropic model is possible. Nevertheless, I chose to use the isotropic semivariogram to model the LCU surface elevation.

Figure 3.40 shows the median polish kriged LCU surface elevation and prediction standard deviation. Note that the prediction standard deviation is about three to four feet over most of the sampled area. This indicates that a substantial amount of variability in the LCU surface has not been accounted for by this model. Under the assumption that the model error is normally distributed, the true surface at an unsampled point is expected to lie within about 2 standard deviations from the predicted value, with a nominal 95 percent confidence. Note that the prediction standard deviation increases with distance away from the body of data, and is quite large under the building in the southwest corner of the site. It is interesting to note that while the discrepancy function (Figure 3.14) and the daily maps (Figure 3.16) indicate that the map is converging toward the "true" surface, there is still a fairly large amount of uncertainty associated with point prediction. In other words, while the general trend of the surface has been relatively well defined, the small scale fluctuations account for prediction errors up to six to eight feet at unsampled locations away from the bulk of the data.

To further assess the model, I performed a cross-validation whereby each data point was removed from the dataset one at a time, and the remaining data were used to predict at the removed data location. In this way we generate a series of pairs of predicted versus observed data. If the model is performing well, the predicted versus observed data should lie approximately along a 45 degree line. Figure 3.41 shows the predicted versus observed LCU data based on the median polish kriging procedure. These data fall near the 45 degree line and indicate that the model is performing reasonably well. Additionally, the standardized residuals
Contours in feet above msl
Contour Interval = 2 feet

a) Median polish kriged LCU surface elevation.

Scale: feet
0 50 100 150 200

Figure 3.40. Median polish kriging LCU surface elevation (with data locations shown) a) and kriging prediction standard deviation b).

b) Kriging prediction standard deviation.
Figure 3.41. Cross validation predicted versus observed LCU surface elevation.

\[ z = \frac{\text{observed value} - \text{kriged value}}{\text{kriging prediction standard deviation}} \]  

should follow a standard normal distribution if the assumption about the normality of the median polish residual process is correct. Figure 3.42 shows a normal probability plot of the 66 standardized residuals from the cross-validation. The standardized residuals have a mean of 0.0 and standard deviation of 1.0, and appear to follow a normal distribution according to Figure 3.42.

An independent test of the model was also possible because there were 14 BVWST soil borings which intersected the LCU surface within the area were the ESC data was obtained but were not used to develop the median polish kriging model. I used this model to predict the LCU surface elevation at the BVWST soil boring locations, and then compared the predictions to the logged LCU surface elevations. The results of this are reported by Stenback et al. (1996) and indicate that the model performs relatively well, but shows a slight positive bias. More significantly, three of the standardized residuals from the predictions to the BVWST data locations were 3.94, 2.78 and 2.58, which are larger values than one would expect to see from a sample of 14 standard normal deviates. The remaining 11 standardized residuals ranged from -2 to 2, typical of standard normal deviates.
The implication is that there may be even more uncertainty in the model than is indicated by the prediction standard deviation as shown in Figure 3.40b.

Similar analyses can be performed on other variables, such as contaminant concentrations, to develop and test models which are used to display the data, make inferences to the population from which the data were drawn, and assess the level of uncertainty. Developing accurate models may be a time consuming task, but if any statistical inference is made on the basis of such models, it is essential that the assumptions be tested and documented so that an appropriate level of confidence can be attached and made clear to the decision authority.

To assess which portions of the site exceed ARARs, we may develop contour plots of the contaminant distribution, along with plots which indicate the probability of exceeding a specified threshold. For the zone 1 TPAH data, there are 33 measured values, including the BVWST data, of which 21 are non detects. Setting aside nondetect data locations ML036, ML038, ML042, ML070, ML073, and ML078 because they are separated from the plume area by a series of nondetect data

Figure 3.42. Normal probability plot for the median polish residuals.
gives the semivariogram data for the natural logarithm of the zone 1 TPAH data shown in Figure 3.43. Figure 3.43 shows spherical and Gaussian model semivariograms fit to the data using a minimum sum of squared errors implemented using Solver in Microsoft Excel. Both models fit the data equally well according to the sum of squared errors, however, there is clearly a difference between the models near the origin. While there is no data near the origin from which to choose one model over the other, there are several reasons why the Gaussian model may be the better choice. First, the ESC PAH (hence TPAH) measurements were made on gas chromatograph/mass spectrometry equipment and have a measurement error standard deviation which is nearly proportional to the magnitude of the value being measured, and the proportionality constant, k, varies from about 0.1 to 0.5 for the different PAH compounds (Stenback and Kjartanson, 1995). Assuming the measurement error is approximately normally distributed with zero mean and standard deviation equal to k times the true value, taking the natural log of a measured value yields a result with a measurement error standard deviation approximately equal to $k + k^2$ (for $0 \leq k \leq 0.5$). Accordingly, the nugget, or y-intercept for the semivariogram should be about 0.3 or so. Second, due to PAH concentration in fractures in the soil mass, fingering or stringer formation of the coal tar DNAPL during migration into the subsurface, and natural heterogeneity of the soil, there is good reason to expect a significant variation from sample to sample at the scale of the sample size (one inch diameter by two foot long soil cores). Accordingly, the Gaussian model with the nugget at 1.23 on the semivariogram axis is probably a better choice than the spherical model with a nugget of 0.15 (the units on the variogram axis are the square of the variable units, in this case (mg/kg)$^2$).

Because some of the data are below the TPAH method detection limit, the method of Stein is appropriate for the development of a contaminant distribution model. Because the BVWST MDL was 1 mg/kg TPAH and the phase 2 MDL was 16 mg/kg, and Stein's method is designed to treat only one detection limit, we could just assume that the MDL for all the zone 1 TPAH data is 16 mg/kg. Otherwise, we could perform ordinary kriging on the detect data as it is and use a separate model (e.g. ND region) for those regions dominated by ND data. Because the MDLs of 1 mg/kg and 16
Figure 3.43. Semivariogram for the natural log zone 1 TPAH data.

mg/kg are considerably lower than the action level of 500 mg/kg, it may not make much practical
difference which alternative we use.

A contour plot showing kriged contours for zone 1 TPAH contamination is shown in Figure
3.44a. This plot was developed on the natural log scale using the Gaussian semivariogram in Figure
3.43, and transformed back to the data scale (assuming the log scale TPAH data are normally
distributed) using the equation (Cressie, 1991)

$$Z(S_0) = \exp(W(S_0) + \alpha^2_k(s_0) / 2 - m) \quad [3.17]$$

where $\alpha^2_k(s_0)$ is the kriging prediction variance, $m$ is the Lagrange multiplier in equation 2.2, and
$W(S_0) = \ln(Z(s_0))$. Accordingly, we can write a model for the log scale as

$$W(s) = \ln(Z(s)) + \delta(s) \quad [3.18]$$

where $Z(s)$ is the observed value at spatial coordinate $s$ and $\delta(s)$ is a correlated error with correlation
described by the Gaussian semivariogram. The area with predicted TPAH in excess of 500 mg/kg is
shaded in Figure 3.44a. While we have only four measured TPAH values that exceed 500 mg/kg,
we find a substantial area with soils predicted to be above the action level. Note from Figure 3.44a
that it appears that some soil with TPAH exceeding 500 mg/kg exists under the buildings to the southwest of the main plume centered on the former purifier building. There also appears to be an increase in TPAH levels towards the corners of the plot away from the data locations; this is a peculiarity of the kriging process. Kriging to points far removed from data locations results in kriged values that tend toward the average of the data. This is one (but not the only) reason why these equations should not be used to extrapolate far outside the region where data was collected. A reasonable approach for this situation would be to krig up to the region dominated by nondetect values, and report only detection limit levels within the region where only nondetect data were observed. A measure of confidence in the contours shown in Figure 3.44a is given by the probability that TPAH exceeds 100 mg/kg, based on [3.1] is shown in Figure 3.44b. Note from Figure 3.44b that the probability of contaminated soils underlying the buildings to the southwest of the main plume, as well as soils near the corners of the contoured area, is low.

The large amount of phase 2 sampling on the east side of the plume has clearly defined this edge of the contaminated soils, however, the extent of contamination to the north, west and south is relatively uncertain due to the lack of samples in these regions. This is clearly illustrated in Figure 3.44b. Figure 3.44b shows the area surrounding the former purifier building to be contaminated above TPAH equal to 100 mg/kg with high probability (inside the probability = 0.9 contour). While there are several areas where the probability of contamination for selected samples within this area is 0.1 or less, a large portion of the site would be classified as state uncertain due to insufficient sampling on the basis of Figure 3.44b.

Bayesian updating to develop posterior prediction maps for the TPAH concentrations, on the basis of a constant mean and variance prior, could be performed in a manner identical to that used for the CL data as illustrated in Figures 3.8, 3.9, and 3.11. Alternatively, I used the CL posterior distribution as the mean and variance of the prior distribution, and then updated the mean and variance using the TPAH data. This was accomplished using equations 2.4 repeatedly, once to obtain the CL data posterior distribution shown in Figures 3.8 and 3.9, and then again, using the CL
a) Kriging predicted TPAH (mg/kg).

b) Estimated probability that TPAH exceeds 100 mg/kg.

Figure 4.44. Kriging results for zone 1 TPAH. a) Predicted TPAH (mg/kg), b) P(TPAH > 100 mg/kg).
posterior distribution as the new prior, to obtain the TPAH posterior distribution. I used the covariance function $C(h) = C(0) - \gamma(h)$, where $C(0)$ was taken to be the semivariogram sill (6.49, Gaussian model) in Figure 3.43, and $\gamma(h)$ is the Gaussian semivariogram in Figure 3.43. I used a measurement error standard deviation of 0.5 (on the log scale) for the CL data model to incorporate a relatively large amount of uncertainty associated with this screening data. This measurement error was incorporated into [2.4] by setting $E = 0.25I$, where $I$ is the identity matrix.

Figure 3.45 shows the results of the Bayesian updating. To estimate the mean on the data scale, I generated a series of 31 uniformly spaced values on the normal scale (i.e. the $\ln$(TPAH) scale) which cover the range from, say, the mean minus six standard deviations to the mean plus six standard deviations, and the associated probabilities that the normal random variable falls within mutually exclusive cells centered on each of these values (with the lower and upper cells extending to $-\infty$ and $+\infty$, respectively). An estimate of the data scale mean is then given by the sum of the exponentiated normal scale values times their respective probabilities:

$$\text{Data Scale Mean Estimate} = \hat{Z}(s_o) = \sum_i \exp(y_i)P(Y \in \text{cell } i). \quad [3.19]$$

where $y_i$ is a value centered in cell $i$, and $Y$ is a normally distributed random variable with mean $\hat{Y}(s_o)$ and standard deviation equal to the kriging prediction standard deviation.

Comparison of Figure 3.45 with the kriging results in Figure 3.44 shows that the Bayesian updating indicates a slightly larger area contaminated at the 500 mg/kg level and predicts lower TPAH levels near the edges of the contoured area. Also, the Bayesian updating estimation of the probability that sampled TPAH values would exceed 100 mg/kg (based on equation 3.1) shows closely spaced contours with a well defined plume edge, with the exception of the area beneath the large building in the southwest corner of the site. These differences are largely due to the influence of the CL prior data.

For comparison, I applied Stein's method to the zone 1 TPAH data using the ESC MDL of 16 mg/kg for all ND data. I assumed that the TPAH distribution is lognormally distributed, applied
a) Bayesian updating predicted TPAH (mg/kg).

b) Estimated probability that TPAH exceeds 100 mg/kg.

**Figure 4.45.** Bayesian updating results for zone 1 TPAH using the CL model as a prior. 
- a) Predicted TPAH (mg/kg),
- b) $P(\text{TPAH} > 100 \text{ mg/kg})$. 

---

[Diagram of Bayesian updating results for zone 1 TPAH using the CL model as a prior, showing predicted TPAH (mg/kg) and estimated probability that TPAH exceeds 100 mg/kg.]
Stein's method to the natural logarithm of the TPAH data, and back transformed the output to the TPAH data scale using a procedure similar to equation 3.19. I also used Stein's method to estimate the probability that the TPAH value at unsampled points is above 100 mg/kg (recall that Stein's method estimates the probability that the variable is less than user input thresholds directly; 1 minus this is the probability that the variable exceeds the threshold). The results of are illustrated in Figure 3.46. Comparison of Figures 3.44, 3.45, and 3.46 indicates that Stein's method did not produce as high of TPAH values under the former purifier building as did kriging or Bayesian updating, however. Stein's method produces a greater northern extent, but otherwise similar region, of soils with expected 500 mg/kg TPAH levels. Stein's probability map produces a sharper eastern and northwestern plume boundary than the kriging method, but is generally similar to the kriging method otherwise.

Similar maps can be constructed for other zones, and other contaminants (CPAH for example) of interest. There is generally no need to use kriging, Bayesian updating, and Stein's method for a particular analysis. I have done so here for comparative purposes. The situation should play a role in determining which methods are appropriate. If there are nondetect data, then Stein's method is appropriate. However, Stein's method is computer intensive and slow, so for a quick calculation, kriging or Bayesian updating might be preferable. If there are screening data, or some other prior information available that we want to incorporate into the analysis, then the Bayesian updating method is appropriate. Ordinary kriging does not require specification of a population mean (both Bayesian updating and Stein's method do) under the assumption of first order stationarity, and can be used when the variogram exists, but the covariance does not. Both the Bayesian updating and Stein's methods applied here require a covariance function.

3.2.4 Summary of Statistical Applications at Marshalltown

3.2.4.1 Phase 1 Phase 1 screening data may be used to plan the sampling strategy and estimate initial variograms for use during the early stages of phase 2. CLP data (the BVWST data,
Figure 4.46. Stein's method results for zone 1 TPAH. a) Predicted TPAH (mg/kg), b) $P(\text{TPAH} > 100 \text{ mg/kg})$. 

a) Stein's method for predicted TPAH (mg/kg).

b) Estimated probability that TPAH exceeds 100 mg/kg.
for example) may be used to determine that there is contamination in sufficient concentrations to
determine whether the site poses a threat to human health or the environment, if this is not already
known, and to determine if the concentration of contaminants exceeds the ARARs. The DQO
process may be utilized to determine data needs so that a proper assessment of appropriate
statistical methods may be planned for use during phase 2.

3.2.4.2 Phase 2 Englund an Heravi's method or Johnson's method may be used to attempt
to accurately define a specified contaminant threshold, but may not provide sufficient data to
accurately model an entire contaminant plume. The David and Yoo method may be used provide a
uniform sample spacing which may be applied to the assessment of geologic contacts with good
success, and may be modified by restricting the search area to provide a uniform coverage
throughout the contaminated area. To successfully map an entire plume with reasonable accuracy
requires placement of samples throughout the contaminated region. Accordingly, the method of
David and Yoo (with restricted search area) or the graphical uncertainty zone sampling (see Figure
3.35 for example) can be used to select sample locations to effectively map a contaminant plume
with similar characteristics as the coal-tar plume. The approach used should depend on the data
needs (as identified in the DQOs).

The methods of Johnson, David and Yoo, and Englund and Heravi are computer intensive
but can be used on-site to select future sample locations. The graphical uncertainty zone sampling
(Figure 3.35) is a quick graphical alternative to aid sample location selection and may be used in
conjunction with screening data, such as electrical conductivity or LIF, to enhance data collection in
regions where essential information is lacking. Additionally, sampling until the plume is enclosed
within a low (0.1 or 0.2) probability of exceeding a desired threshold contour, such as the action level
or clean-up standard, could be used as an effective stopping criteria. More will be said about this
graphical uncertainty zone sampling technique in Chapter 4.

The discrepancy function may be used to indicate when sufficient data have been obtained
to produce an adequate map. Successive daily maps may be used to show convergence of a map
with additional sampling events, and for the FMGP site, such maps were as effective as the discrepancy function. The cost structure used in Engelund and Heravi's method may be used to indicate when additional sampling is adding little value to the expected total cost of the remediation. However, developing an accurate cost structure that models reality reasonably well may be difficult.

3.2.4.3 Post Phase 2 If an adequate covariance structure can be modeled from the data, then predictive maps, along with measures of uncertainty, such as prediction standard deviation plots and probability maps, may be produced. Such plots may be used by the decision authority to determine appropriate remedial actions, and may be used to develop robust remediation strategies that will perform well under the expected variation of site conditions. Alternatively, these uncertainties and modeled covariances for both the geologic and contaminant models may be used within the decision analysis framework developed by Freeze et al. (1990) assess and develop remedial actions, or suggest that additional data be collected.
4. ASSESSMENT AND DEVELOPMENT
OF GEOSTATISTICAL TOOLS FOR ESC:
SAVANNAH RIVER SITE OIL SEEPAGE BASIN

4.1 Detailed Case Analysis

This Chapter provides an analysis of some data collected during the ESC project at the D-Area Oil Seepage Basin (DOSB) at the USDOE Savannah River Site in South Carolina. Particular emphasis is given to those data and data analysis applications which are not currently performed in an Ames Laboratory ESC project, and which may be useful in on-site decision making, including selection of subsequent sample locations and stopping rules, and post field work analysis to determine if a site needs remediation, and if so, which portions of the site need remediation.

4.1.1 Savannah River Site, SC — Background

The Savannah River Site (SRS) ESC investigation program was a unit assessment RI activity carried out within the RCRA Facility Investigation/Remedial Investigation (RFI/RI) at the D-Area Oil Seepage Basin (DOSB), SRS, South Carolina. The phase 1 and phase 2 ESC work was carried out from May to October of 1995. The DOSB was constructed in 1952 as a series of at least three unlined eight foot deep trenches separated by berms to dispose of waste oils and other fluids not suitable for burning, (WSRC, 1995). Waste oils including machine cutting oil and transformer and other shop fluids were transported to the basin in drums, opened and dumped into the trenches. These waste oils were periodically burned along with general office and cafeteria waste. No historical evidence of overflow of the basin exists. This practice continued until 1973 when open burning at the SRS ceased. The basin received waste oils until 1975 when it was removed from service and backfilled with soil. An unknown number of 55 gallon drums, possibly containing waste oil, remained in the basin. The basin remains inactive and was covered by bushes and grasses and surrounded by trees during the ESC activities. A plan view of the DOSB is shown in Figure 4.1.
Carolina Bay (Wetlands)

Figure 4.1. Savannah River Site D-Area Oil Seepage Basin (DOSB).
Groundwater monitoring in the DOSB has indicated the presence of tetrachloroethylene (PCE), trichloroethylene (TCE), and vinyl chloride (VC) (WSRC, 1995). In 1988 a drum was punctured during a drilling activity and was found to contain a liquid with high concentrations of several organic compounds including TCE and PCE. Contaminants detected in soils collected from beneath the unit include PAHs, alkyl benzenes, phthalate, lead, chromium, and antimony. In 1993, a soil sampling event detected pesticides, polychlorinated biphenyls (PCBs), dioxins, furans, benzene, ethylbenzene, toluene, xylene, and naphthalene.

Covering the DOSB with soil has reduced the potential for exposure to humans and wildlife. Prior to the ESC investigation, risk assessment concerns were expected to consist of potential transport of unit-related contaminants to subsurface soils, groundwater, and offsite transport to surface water bodies. If in the future, institutional control of the area is lost and no further remedial action were taken, onsite residents could become potential receptors. Under the requirements of CERCLA, the DOSB was slated for additional assessment and perhaps, environmental remediation (WSRC, 1995). The assessment of the environmental media was accomplished using the ESC methodology.

The site is known to be underlain by a glauconitic marine clay referred to as the “green clay”. Prior to the ESC investigation the green clay was thought to exist at a depth of about 60 feet, however, little was known regarding the topographic nature of its surface, its thickness, or continuity. The green clay is overlain by alluvial sands, silts, and clays (WSRC, 1995, page 2-4). James and Freeze (1993) have studied radioactive waste at the SRS H-area about eight miles northeast of the D-area. They point out that the green clay acts as a strong confining layer, or aquitard, and while the green clay is likely to be largely continuous, a major source of discontinuity exists in the form of erosion by channeling which may have occurred during any of several periods of unconformity which have occurred since the green clay was deposited. The green clay is dissected along a number of present day stream channels at the SRS.
Four US Army Corps of Engineers SCAPS pushes were performed along the perimeter of
DOSB site during 1991 (Koester, et al., 1993). Three of the CPT pushes were made to a depth of 50
feet and one was pushed to 30 feet. The cone was equipped with the LIF sensor described for the
work at the FMGP in the preceding sections. The soil classifications (after Olsen) from these CPT
pushes indicate sandy soils overlying three to five feet of silty and clayey soils at a depth of about 5
to 15 feet. This sequence overlays eight to ten feet of sandy soils overlaying a three to eight foot
thick silt and clay unit at about the 25 to 30 foot depth range. This overlays interbedded sands, silts,
and clays. The LIF logs show elevated fluorescence intensities over the 10 to 20 foot range, with
some lesser “hits” near 30 and 45 feet bgs. These pushes occurred to the north, south, and west of
the trench area.

Selected soil gas measurements were made over the trench area in 1991. The soil gas data
for PCE, TCE, chloroform, ethane, Hg, propane, l-butane, N-butane, ethylene, and propylene all
indicated elevated regions and localized "hot" spots for these contaminants. Soil gas measurements
for VC, trans-1,2-dichloroethylene, 1,1,1-trichloroethane, and carbon tetrachloride soil gas
measurements were largely near or below the detection levels. PCE had the largest area of elevated
soil gas measurements and the soil gas data for PCE are shown in Figure 4.2. Figure 4.2 shows that
the elevated PCE soil gas measurements follow the trench area very closely, with the exception of
four slightly elevated measurements in the extreme northeast of the sampled area. These elevated
PCE measurements might be associated with a spill along the roadway, however, no record of this
exists to my knowledge. Extreme elevated TCE measurements are confined to the southwest trench
area while elevated chloroform and mercury measurements were found between the DOSB area and
the Carolina Bay. While soil gas has been used as a screening method to evaluate the extent of a
plume of groundwater contaminated with volatile organic compounds (Fetter, 1993), the PCE and
TCE soil gas measurements here all show NDs to the south and west of the DOSB area in the
direction of the near surface groundwater flow. As the phase 2 data show groundwater PCE and
TCE contamination trending southwest of the trenches in the direction of the groundwater flow, the
Figure 4.2. DOSB 1991 soil gas tetrachloroethylene (PCE) measurements.
soil gas PCE and TCE may be associated primarily with source material in the trenches and not with the groundwater plume. Apart from the PCE detected in the extreme northeast of the sampled area, the PCE soil gas data does not indicate the presence of any source areas outside of the trench area.

A soil sampling investigation in 1993 collected samples at 2-4, 6-8, and 12-14 foot depth intervals at 14 sample locations within the DOSB area. This investigation discovered oil stained soils underlying much of the trench area and analytical results indicating the presence of dioxins, furans, naphthalene, α-benzene hexachloride (α-BHC), β-BHC, α chlordane, γ chlordane, PAHs, DDE, DDT, PCBs, and metals (Al, Sb, As, Be, Cd, Ca, Cr, Co, Cu, Fe, Pb, Mg, Mn, Hg, Ni, Ag, V, and Zn).

4.1.2 SRS DOSB Phase 1 Applications

4.1.2.1 Geohydrology Three monitoring wells were installed in May 1983, and a fourth in June 1984 to characterize the geologic conditions and to monitor the groundwater level and groundwater quality of the basin. Locations of these wells, DOB-1 to DOB-4, are shown in Figure 4.3. These four polyvinyl-chloride (PVC) monitoring wells are screened from about 10 to 40 feet below ground surface (bgs). Quarterly data (1st quarter 1985 to 2nd quarter 1994) collected from these wells show water table depths varying between 4 to 14 feet bgs, however, potentiometric data is inconclusive with respect to groundwater flow direction (WSRC, 1995). In September 1994 six PVC water table piezometers were installed surrounding the DOSB (see DOB-5 to DOB-10 in Figure 4.3) with 20 foot screened zones extending from 3 to 23 feet bgs. Data from the six piezometers indicates that groundwater flows to the southwest.

4.1.2.2 Geology Phase 1 geophysical activities that occurred during the summer of 1995 included ground penetrating radar (GPR), EM38, EM31 and EM34 electrical conductivity measurements, and geophysical logging of two boreholes. The objectives of these activities were to map the aquifer stratigraphy (GPR, 40 MHz to map to a depth of about 60 to 70 feet, and 100 MHz to map to a depth of about 20 feet), including the green clay confining layer, and map soil conductivity for several depth ranges over the region (EM 38 to a depth of about 5 feet, EM 31 to
Figure 4.3. DOSB monitoring wells, piezometers, shallow groundwater flow direction, and rotosonic soil boring locations.
about a 20 foot depth, and EM34 to about a 50 foot depth). Two boreholes were drilled with the innovative rotosonic drilling method to obtain continuous soil cores from the ground surface to a depth of about 10 feet into the green clay. The rotosonic drill locations are referred to as DOL-1, drilled north (upgradient) of the DOSB area, and DOL-2, drilled to the southwest (downgradient) of the DOSB area; the locations of DOL-1 and DOL-2 are shown in Figure 4.3. PVC piezometers were installed in both DOL-1 and DOL-2 with a ten foot screen located at the top of the green clay. The piezometers were logged with natural gamma, gamma-gamma (density), neutron-neutron (porosity), and induction (conductivity) probes.

The EM measurements show elevated soil conductivity around the trenches with elevated conductivity values emanating from the trench area and extending to the west and southwest of the area in agreement with the direction of groundwater flow determined by the piezometers. These elevated conductivity values were associated with buried metallic and non-metallic materials and an inorganic groundwater plume. Downgradient wells DOB-2 and DOL-2 show elevated conductivities at a depth of about 25 feet that were interpreted as possible zones of inorganic contamination; however, phase 2 water quality measurements indicate that these elevated conductivities may be due to clay. The GPR data indicate uniform shallow soil conditions, disturbance in the trench areas, areas with local buried metallic and non-metallic targets, and large magnitude radar reflectors below the trenches and water table possibly associated with DNAPLs (Technos, Inc., draft report included with the 8/21/95 DOSB on-site technical meeting). The green clay was found at a depth of 47 feet in DOL-1 and 37 feet in DOL-2 (Technos, 1995). GPR results indicate the existence of a clay layer between 20 and 30 feet bgs that is continuous over much of the site. A thin layer of silts and cemented sands exists at a depth of about 10 feet over much of the site.

GPR results indicate that the green clay occurs at depths of about 35 to 45 feet, but may not be continuous over the site. No maps indicating depth to the green clay under the site, nor suspected areas where the green clay may or may not exist, were produced on the basis of the GPR data.
4.1.2.3 Analytical Data Phase 1 analytical data collected in the late spring and early summer of 1995 were measured by an off-site CLP laboratory using EPA SW-846 methods to obtain definitive type data. Data were collected at 14 sample locations. Groundwater samples were collected and analyzed for Cu, Fe, Mn, Ni, V, As, Cd, Cr, Pb, Ag, Ba, Sb, Be, PCE, TCE, VC, 4,4'-DDE, 4,4'DDT, α-BHC, β-BHC, α-chlordane, γ-chlordane, and dieldrin. On the basis of comparison of the phase 1 sampling results with background sample results (collected at six locations northeast of the DOSB area) and risk based concentrations, the phase 2 site specific contaminant list was developed and is listed in Table 4.1. The media of concern for the phase 2 ESC work was groundwater. Phase 1 data for PCE is shown in Figure 4.4.

<table>
<thead>
<tr>
<th>Contaminant Type</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metals</td>
<td>Sb, As, Be, Mn</td>
</tr>
<tr>
<td>Volatiles</td>
<td>PCE, TCE, Vinyl Chloride</td>
</tr>
<tr>
<td>Pesticides</td>
<td>DDE, DDT, α-BHC, β-BHC, α-Chlordane, γ-Chlordane, and Dieldrin</td>
</tr>
</tbody>
</table>

4.1.3 SRS DOSB Phase 2 Applications

The phase 2 ESC project work at the DOSB occurred during August and September 1995 and was a five week on-site investigation. Preliminary analyses of the phase 1 data were available, however, the full analyses of the geophysical data was not published until December 1995 (Technos, 1995).

4.1.3.1 Geology and Hydrogeology Further characterization of the site geology during phase 2 consisted of CPT testing by two CPT rigs: one operated by Applied Research Associates (ARA) and the other was a DOE SCAPS CPT truck. Both CPT rigs provided plots of cone rod sleeve stress, cone tip stress, sleeve to tip stress ratio corrected for overburden stress, and dynamic pore pressure generated during the push. During much of the investigation, these rigs did not provide
Figure 4.4. DOSB phase 1 groundwater tetrachloroethylene (PCE) off-site CLP measurements.
interpretations of the soil behavior type. The SCAPS rig performed 63 pore pressure dissipation tests at various depths and push locations where the dynamic pore pressure was elevated, indicating a low hydraulic conductivity soil. With the exception of a preliminary look at static pore pressures to determine vertical hydraulic gradients, no attempt was made to interpret the pore pressure dissipation data in the field. Post phase 2 analysis indicates that the data from the two rigs do not agree. Both CPT cone rods were 1.75 inch diameter rods with pore pressure sensor directly behind the cone tip.

A Geoprobe soil electrical conductivity probe was also used to help characterize the soil stratigraphy. However, because soil electrical conductivity is influenced by a variety of factors including the soil type, water content, and pore fluid chemistry, the electrical conductivity may respond to the presence of an inorganic plume, as is indicated by the phase 1 EM data, and thus a clear interpretation is difficult. The CPT and electrical conductivity probes were pushed adjacent to (within about 10 feet of) DOL-1 and DOL-2 to compare the response with the logged soil borings at those locations. There were no other soil borings at this site with detailed enough logs to provide useful correlations.

Water elevation was measured in the piezometers during the site investigation and it was determined that the depth to the water table was about seven feet throughout most of the DOSB area. For comparison with the dynamic pore pressures in the following paragraphs, a seven foot deep water table gives a hydrostatic pore pressure of about 23.0 psi at a depth of 60 feet.

Figure 4.5 shows the CPT and electrical conductivity output, along with nearby DOL-1 log for comparison. Figure 4.6 shows the CPT and electrical conductivity output, along with nearby DOL-2 log for comparison. In both Figures 4.5 and 4.6, the green clay depth as identified in the soil boring log is indicated by a horizontal line across the plots. The Olsen soil behavior classification shown in Figures 4.5 and 4.6 is after Olsen (1988). There is some similarity, but not complete agreement, of the CPT and electrical conductivity response to the supposed green clay between DOL-1 and DOL-2. Near DOL-1, the CPT dynamic pore pressure in the green clay is elevated above hydrostatic
Figure 4.5. S-001 CPT data, GP-001 data, and Technos DOL-1 log.
Figure 4.6. S-002 CPT data, GP-002, and Technos DOL-2 log.
pore pressure indicative of cohesive low permeability soils (Robertson, 1989), the tip and sleeve stress are both very high at the green clay surface and drop off with penetration into the green clay, and the Olsen soil classification indicates interbedded clays, silts and sands above the green clay with predominantly sand mixtures and sands within the green clay as identified in the DOL-1 log. Near DOL-2, the CPT dynamic pore pressure in the green clay is near zero or negative indicative of dilative overconsolidated low permeability soils (Robertson, 1989), the tip and sleeve stress are both high near the top of the green clay and drop off with penetration into the green clay, and the Olsen soil classification indicates sands above the green clay, silts and clays in the upper four feet of green clay, with sands interbedded with silts and clays in the lower green clay. The electrical conductivity log near DOL-1 shows a gradual increase with penetration into the green clay while the electrical conductivity log near DOL-2 shows a rapid increase at three feet into the green clay, followed by a decline at about 6 feet into the green clay, followed by a gradual increase with penetration deeper into the green clay.

In summary, the CPT and electrical conductivity response near DOL-1 and DOL-2 show the following characteristics with penetration into the green clay:

1. Neither of the electrical conductivity logs shows a characteristic response when entering the green clay, but both show a gradual increase with penetration into the green clay.

2. The dynamic pore pressure may be above, at, or below hydrostatic pore pressure (about 12 to 18.6 psi from 35 to 50 feet bgs) within the green clay and may not show any abrupt change upon entry into the green clay. CPT data near DOL-1 and DOL-2 both show pore pressures significantly above hydrostatic in the lower green clay noted as "hard clay" in the DOL-1 and DOL-2 logs.

3. The Olsen CPT soil classification may show sands or interbedded clays, silts, and sands both above and within the green clay. Sand mix and sand behavior type soils are common within the green clay.
4. While the tip and sleeve stress appear to be high and drop off upon entry into the green clay, such behavior is not unique (see Figure 4.6 at 50 to 55 feet for example) and may lead to erroneous identification of the green clay surface.

The previous factors lead to difficulty with on-site determination of the depth to the upper green clay surface. Picks of the depth to the green clay varied from person to person and from day to day as new data was obtained and reevaluated. However, the evidence suggests that the CPT tip and sleeve stress may be high just above the green clay and decrease with penetration into the green clay, and that both the electrical conductivity and pore pressure may gradually increase with penetration into the lower “hard clay” region of the green clay. On the basis of Figures 4.5 and 4.6, it seems clear that the nature of the green clay is variable but may be characterized by interbedded clays, silts and sands between about 35 to 50 feet bgs; this interpretation is consistent with the phase 1 geophysical data.

Figure 4.6 shows elevated electrical conductivity readings (>50 mS/m) near DOL-2 at 25 to 29 feet bgs and again at 40 to 42 feet bgs directly adjacent to clay/silt mix soil types according to the Olsen soil classification. Figure 4.5 shows no corresponding elevated electrical conductivity readings near DOL-1 adjacent to the interbedded clays from 31 to 46 feet bgs. It was proposed on-site during phase 2 (later reported by Technos, 1995) that the elevated electrical conductivity readings may be due to inorganic (metals) contaminants adsorbed to the clay minerals causing increased electrical conductivity in those clayey soils through which contaminated groundwater has passed. The ten foot unit detected by the phase 1 geophysical work is clear in both Figures 4.5 and 4.6. The 25 to 30 foot unit appears clearly as clayey soils with elevated pore pressures near DOL-2 (Figure 4.6) but shows only a faint response near DOL-1 (Figure 4.5).

While there was no clear identification of entry into the green clay from the DPT data, there was never an indication that the green clay was not present as indicated by interbedded Olsen CPT soil classification, elevated dynamic pore pressures, and elevated electrical conductivities. Figure 4.7 shows the CPT soil classification data as small scatterplots positioned at the push locations.
Push locations are indicated by either and "A" for ARA data or "S" for SCAPS data, followed by a number indicating the order in which the pushes were completed. In almost every CPT push, there is about 5 to 10 feet of sandy surface soils underlain by the "10 foot" layer of silts and clays. These soils overlay another sandy soil unit that is usually underlain by silty and clayey soils from about 25 to 30 feet bgs, called the "25 foot" unit. This in turn is underlain by interbedded clays, silts, and sands, often with about five feet of granular soils directly under the "25 foot" unit.

Figure 4.8 shows the dynamic pore pressure data plotted versus depth at the map locations where each push was made. Elevated pore pressures are common, but not universal, at the 10 foot and 25 foot levels. All pushes that were carried out to a depth near 60 feet show elevated pore pressures near the 45 to 55 foot range indicative of low hydraulic conductivity soils at that depth. In a similar fashion, the Geoprobe electrical conductivity data are shown in Figure 4.9. With almost every push there are elevated electrical conductivity readings beginning near the 30 to 40 foot range indicative of clayey soils. Note that three of the electrical conductivity traces, GP003, GP019, and GP035 (see Figure 4.9) showed anomalous responses with very low electrical conductivity readings with depth. CPT data S003 taken nine feet from GP003 shows a very typical response suggesting the anomaly is in the electrical conductivity trace and not the geology (S003 and GP003 are both shown in Figure 4.42 in the following section). There are no CPT pushes nearby GP019 or GP035 for comparison.

Given the access constraints imposed by trees at the site, the DPTs provide a relatively uniform distribution of push locations over the region expected to be influenced by contamination from the trenches. In a number of cases, both CPT rigs and the Geoprobe obtained push data from within about ten to 15 feet of one another for direct comparison. This provides the additional benefit of allowing assessment of the nature of variograms for closely spaced separation distances.

4.1.3.2 Analytical Data Groundwater samples for chemical constituents listed in Table 4.1 were collected by geoprobe groundwater sampling equipment. Samples were analyzed on-site in mobile laboratories using EPA SW-846 methods designed to produce definitive type data. Analytical
Figure 4.7. Olsen CPT Soil Behavior Classification
Figure 4.8. CPT Dynamic Pore Pressures at the SRS DOSB.
Figure 4.9. Geoprobe electrical conductivity versus depth at SRS DOSB (truncated at 80 mS/m).
results were generally available within several hours (or by the next day) of sample collection.

The sample location selection methods of Englund and Heravi, Johnson, and David and Yoo are all applicable in the DOSB ESC project setting. The method of Englund and Heravi was developed for use in a situation where a remediation unit can be defined and was applied in their paper (Englund and Heravi, 1994) to a soil contamination problem. To track groundwater contamination we might define a remediation unit as the volume of groundwater that could be captured and treated by a pumping well or containment system and use this as the remediation unit in the Englund and Heravi algorithm to adaptively determine the placement of subsequent samples.

As was done at Marshalltown, once a sample location had been selected, samples were generally collected at about five to seven depths, although some locations were sampled at only one depth. There were 14 site specific contaminants which were tested for and tracked during phase 2 (see Table 4.1). While ultimately each contaminant must be considered separately, to simplify the work one could consider contaminant groups (e.g. those with similar transport characteristics, such as a metals group or a volatiles group) simultaneously by considering a sample contaminated if the analytical result for any contaminant in the group is above its regulatory or risk based limit. If a sample is contaminated by this criteria, assign it a value of one, otherwise assign it a value of zero. The method developed by Johnson or the simplified indicator kriging/maximum minimum distance method (described for the preceding Marshalltown data) could be used with such grouped data to aid the selection of subsequent sample locations. Another way to simplify the problem is to first consider the lateral extent of contamination and then focus on the vertical spatial distribution. One approach might be to assign the largest measured value over selected sample depths at a given location to that location. This will (temporarily) reduce the three-dimensional problem to a two-dimensional problem. Alternatively, one could break the site into several zones, as was done for the FMGP site discussed in the preceding section. Such simplifications may be useful in the early stages until a three-dimension picture begins to form as more and more data are collected.
The geology and contaminant spatial distribution at the DOSB were complex and even on the last days on-site during phase 2 it was not entirely clear why contamination was found at some locations and not at others, and where to place subsequent samples. As an example, to apply Johnson's method as an aid to sample location selection, we could have developed a prior probability map describing were we expected to find PCE contamination exceeding the risk based concentration (RBC) of 1.1 ppb in groundwater. Considering the soil gas data in Figure 4.2, the phase 1 PCE data in Figure 4.4, and the fact the groundwater flow is from the northeast of the site and turns somewhat toward the south as it leaves the site, we could establish the prior probability map as shown in Figure 4.10. Johnson (as well as others, for example, Freeze et al., 1990) are willing to assume a variogram model, range, and sill in the absence of data, to begin the process. For example, on the basis of the soil gas PCE and phase 1 PCE data (Figures 4.2 and 4.4) we could assume that a spatial correlation exists up to about 100 feet. Accordingly, we might choose to model the indicator data for PCE > 1.1 ppb using an isotropic exponential semivariogram model with no nugget, a sill of 0.25, and a range of 100 feet at 95 percent of the sill. This translates into the following model:

\[ \gamma(h) = 0.25(1-\exp[-h/33.4]). \]  

where \( h \) is the separation distance.

Johnson's algorithm was applied to the indicator data based on the maximum PCE value observed over different sample depths at the first 28 phase 2 sample locations using the semivariogram model in [4.1], and the prior probability distribution shown in Figure 4.10. The results in the posterior beta distribution mean are shown in Figure 4.11. The state uncertain region, defined here by the probability statement \( 0.1 < P(\text{PCE} \geq \text{the PCE RBC}) < 0.9 \), is shown as the unshaded area in Figure 4.11. The dark shading (crosshatch) in Figure 4.11 shows a large region near the trench area, and several smaller regions south of the trenches, where the probability of contamination is expected to be high. On the basis of this criteria, Figure 4.11 shows that the PCE plume is not yet bounded and that additional samples are needed on all sides of the sampled area.
Figure 4.10. DOSB prior probabilities for tetrachloroethylene (PCE) above RBC.
Figure 4.11. Johnson's probability map based on [4.1] and the first 28 PCE data points.
As data is collected, semivariograms can, and should, be estimated and updated periodically, so that any statistically based maps will better reflect the actual conditions. Figure 4.12 shows indicator semivariograms for the maximum PCE at a sampled location after \( n = 28, 43, 58, \) and 73 of the total 73 phase 2 PCE data locations were sampled. The indicator values, \( l(PCE, s_j) \), where defined as

\[
l(PCE, s_j) = 1 \text{ if maximum } PCE \geq \text{RBC}, \quad l(PCE, s_j) = 0 \text{ otherwise,}
\]

where RBC is the risk based concentration above which the risk was determined to be excessive and \( s_j, i = 1,...,n \), are the sample locations. Note that with 28 sampled locations, the shape of the indicator semivariogram is beginning to show, and does not change much until all the data were obtained. When all the data were available (\( n = 73 \)), the semivariogram shows more variation at the lowest lag than indicated in the \( n = 28, 43, \) or 58 semivariograms. The exponential models shown in Figure 4.12 were generated using a least sum of squares criteria to estimate the nugget, sill, and range parameters; for the \( n = 73 \) case, the squared errors were weighted by the number of data pairs to downplay the large semivariogram value at a lag of 40 feet. All of these semivariogram models indicate a nugget near 0.1 (0.5 to 0.15), a sill near 0.3, and a range at 95 percent of the sill near 300 feet or greater. During the phase 2 field work, the assumed (e.g. equation 4.1) and/or calculated semivariograms would be updated periodically as additional data is collected, as shown in Figure 4.12.

To illustrate the evolution of the Johnson probability model and potential sample locations, I updated the Johnson prior in four successive stages utilizing the indicator semivariograms shown in Figure 4.12. Each stage uses the probability map and beta distribution parameters, \( \alpha \) and \( \beta \), from the prior stage, together with most recent estimate of the indicator semivariogram, to update the probability map. The first stage was shown in Figure 4.11 and was described previously. The second stage illustrated here is shown in Figure 4.13. Figure 4.13 was developed using the probability map shown in Figure 4.11 (along with its associated beta distribution parameters), the indicator semivariogram shown in Figure 4.12, \( n=28 \) case, and the first 43 PCE indicator values.
Figure 4.12. Isotropic exponential semivariograms for the PCE indicator in equation 4.2.

(defined by [4.2]). Figure 4.13 shows the same general areas with high probability of contamination that were seen in the prior stage (Figure 4.11), and also indicates that the eastern edge, and some of the northern edge of the plume is bounded by a region with low probability of contamination. The unshaded regions are state uncertain, according to the probability statement $0.1 < P(\text{PCE} \geq \text{PCE RBC}) < 0.9$, and represent regions where additional data is needed to close the boundaries of the plume (on the basis of the RBC taken as the threshold of interest).

Several points of caution regarding Figure 4.13 are in order. While it might appear that the plume has been bounded to the southwest according to the 0.1 probability contour, some of the PCE measurements in this region were above the detection limit, but below the RBC. Because the groundwater is believed to be moving toward the southwest, it is possible that contamination exists further to the southwest. Also, recall that there was PCE observed in the soil gas data near the
Figure 4.13. Johnson's probability map updated on the basis of Figure 4.11 and the first 43 PCE data points.
northeast corner of the DOSB (see Figure 4.2). While the map shown in Figure 4.13 indicates a low probability (< 0.2) of contamination in the northeast corner, there is no analytical data in this region to support a conclusion stating the area is clean. A probability of 0.95 of finding contamination in the northeast corner was incorporated into the initial prior probability map (see Figure 52) to account for PCE in the soil gas data observed there, and yet the two updating stages used to obtain Figure 4.13 have resulted in a low probability of PCE contamination there. The main point to make here is that all pertinent background information should be considered in the choice of subsequent sample locations. While use of the statistical sample location selection methods will almost certainly improve the quality of the data, reliance on any one method alone may lead to erroneous conclusions.

To follow up on the evolution of the Johnson probability model, the third stage shown in Figure 4.14 was developed using the beta distribution parameters associated with Figure 4.13, the indicator semivariogram shown in Figure 4.12 for the n=43 case, and the indicator values for the first 58 sample locations. Figure 4.14 shows basically the same area with a 0.9 or greater probability of PCE exceeding its RBC as shown in Figure 4.13. The inclusion of several samples within the northeast region of the DOSB has closed the 0.1 probability contour in this region (compare Figures 4.14 and 4.13), while the inclusion of several samples in the southwest corner of the mapped area has just opened the 0.1 contour southwest of the DOSB along the south road. A need for additional samples to the southeast between the south road and the DOSB is indicated on the basis of the state uncertain criteria.

Lastly, Figure 4.15 shows the probability map generated from the beta distribution parameters associated with Figure 4.14, the indicator semivariogram shown in Figure 4.12 for the n=73 case, and all 73 of the PCE indicator values obtained during phase 2. Comparison of Figure 4.15 with Figure 4.14 shows basically the same region with high probability of PCE contamination, while the region with low probability of contamination north of the site has expanded. The maps in
Figure 4.14. Johnson probability map updated with the first 58 PCE values.
Figure 4.15. Johnson probability based on all 73 phase 2 PCE data values.
Figure 4.15 and 4.14 both show the PCE plume trending to the southwest in the direction of the groundwater flow, as well as a need for additional samples south of the east edge of the DOSB to determine the extent of PCE contamination in that region. In addition, note that one of the four samples taken in the northeast corner of the DOSB showed a PCE concentration above the PCE RBC.

Figure 4.16 shows an indicator kriging plot for PCE developed from the indicator function in [4.2] and the indicator semivariogram shown in Figure 4.12 for the n = 73 case. Comparison of Figure 4.16 with Figure 4.15 shows some similarities and differences. The large range of influence (about 390 feet) and large nugget relative to the sill from the semivariogram in Figure 4.12 (n=73 case) results in a much smoother indicator kriging image than was given by the Johnson updating procedure. As with Figure 4.15, Figure 4.16 shows that the southeastern extent of the plume is not well defined indicating a need for additional samples to the southeast, and that the plume appears to be heading to the southwest off of the mapped area. There also appears to be some contamination in excess of the RBC near the northeast corner of the trench area, as was indicated by the 1991 soil gas data. The PCE data are illustrated in Figure 4.17.

A variogram for the natural logarithm of the maximum TCE over sample depths at each sample location is shown in Figure 4.18 for sample sizes of n = 57 and 72. To develop these plots, one large data value located within the trench area was set aside, and all nondetect values were assigned a value equal to one-half the MDL. The maximum TCE data over sample depths is shown in Figure 4.19. It was well into the phase 2 program before there were enough data to estimate the TCE semivariogram. Prior to about 55 sample locations, the TCE variograms indicated the TCE values appeared to be largely random with no observable spatial structure. Because England and Heravi’s method requires a variogram (or covariance model), this method could have been implemented as an aid to sample location selection using a semivariogram based on the practitioners best judgment, with the data-based variogram incorporated into the algorithm as it became available.
Figure 4.16. Indicator kriging for PCE above its RBC based on the variogram model in Figure 4.12, n = 73 case. Lightly shaded areas show low probability (<0.2), crosshatch shows high probability (>0.8).
Figure 4.17. SRS DOSB Phase 2 maximum PCE over sample depths. MDL = 0.38 ppb, RBC = 1.1 ppb.
Figure 4.18. Semivariograms for the TCE data.

A map showing the probability of TCE exceeding its RBC using the method of Stein is shown in Figure 4.20. Figure 4.20 was developed on the basis of the 57 observations used to generate the semivariogram in Figure 4.18, along with the corresponding semivariogram. Figure 4.20 indicates that the east, north, and west margins of the TCE plume are well defined and that the plume is trending off the map area to the southwest in the direction of the groundwater flow. One curious thing to note is the apparent band of low probability (about 0.2) of contamination exceeding the TCE RBC separating the plume southwest of the trench area from the plume in the extreme southwest of the map area. The pathway connecting these plumes is not apparent and was a source of confusion during the final few days of the phase 2 investigation.

Because each sample location was typically sampled at several depths, there was a large database from which to look for spatial correlations in the vertical direction by calculating variograms in the vertical direction. On the basis of all the data collected, there is no evidence of any significant vertical correlations in any of the volatiles or metals measured. This is likely a result of stratigraphic layering of silty and clayey soil seams and lenses with varying hydraulic conductivity resulting in predominantly localized vertical groundwater flow with minimal vertical mixing. An attempt to construct semivariograms based on maximum values over sample depths, as previously done for TCE and PCE, showed some spatial structure for manganese (Mn) only. In other words, with the exception of Mn and TCE, these data show no apparent spatial correlation. This may be
Figure 4.19. SRS DOSB Phase 2 maximum TCE over sample depths. MDL = 0.61 ppb, RBC = 1.6 ppb.
Figure 4.20. Probability that TCE exceeds its RBC by Stein's method based on 57 out of 73 total observations. Light shading shows low probability (<0.1), crosshatch shows high probability (>0.9).
predominantly due to the abundance of nondetect data. The vinyl chloride (VC) and antimony data are almost entirely nondetect while the log scale Mn data appear nearly normal and have only one nondetect value. The beryllium data was about two-thirds nondetect and the arsenic data was about 55 percent nondetect. The pesticide data were almost entirely nondetect, so it was not possible to look for spatial correlations in this data. Attempts to develop isotropic or anisotropic semivariograms for data within specified stratigraphic layers were unsuccessful with the DOSB data.

The map discrepancy function of David and Yoo (equation 2.7) is shown in Figure 4.21 for the maximum PCE and TCE data over sample depths discussed previously. Both the PCE and TCE discrepancy show an initial increase, followed by a decline up to about sample number 28. The PCE discrepancy shows little tendency to decline significantly beyond sample number 28, and generally shows little improvement overall. While there is a large decline from sample number 18 to 28 with the TCE data, there is little apparent tendency to stabilize until about sample number 50. Note that while these plots appear to have stabilized somewhat beyond sample number 50 or so, indicating some stability in the MQ-B surfaces, this does not necessarily mean that enough data has been

![Figure 4.21. Map discrepancy for PCE and TCE (natural log scale) based on equation 2.7.](image)
collected to fully characterize the contaminant plumes. There still remains the fact that the eastern edge of the PCE plume and the southwestern extent of the TCE and PCE plumes have not yet been well defined. In other words, the model surface may stabilize with respect to additional data, be they high or low, but we still need to check the data near the edges of the modeled region to see if we have captured the edge of the plume or not.

4.1.3.3 Sample Location Selection Method Comparisons To further assess and compare the David and Yoo, Englund and Heravi, Johnson, and uncertainty zone sampling strategies, I simulated a TCE plume that passes through each of the 73 phase 2 measured values (shown in Figure 4.19) using the sequential Gaussian simulation algorithm (Deutsch and Journel, 1992) based on the semivariogram shown in Figure 4.18 (n=72 case). The simulated TCE plume covers the site using a 25 foot grid node spacing and is illustrated in Figure 4.22. I then used the 14 phase 1 measured TCE values shown in Figure 4.23 as an initial dataset and applied each of the sample location selection algorithms to generate an additional 60 data values selected from the simulated TCE surface at the grid node locations. In each of the simulations discussed in this section, nondetect data were replaced by one half of the method detection limit.

The Johnson sample location method for TCE was carried out starting with the same prior probability map that was developed for PCE as shown in Figure 4.10. A spherical isotropic indicator semivariogram with a range of influence 100 feet, a sill of 0.25, and no nugget was assumed. Sample locations were selected using each of the three criteria suggested by Johnson: maximize the number of nodes classified as clean, maximize the number of nodes classified as contaminated, and minimize the number of nodes classified as state uncertain. Grid nodes were assumed to be "clean" if the probability that TCE exceeds its RBC (1.6 ppb) was less than 0.1, nodes with probability of contamination exceeding the RBC greater than 0.9 were assumed to be contaminated, and otherwise the node was declared to be state uncertain. The resulting sample locations are shown in Figure 4.24. Figure 4.24a, whereby the maximize the state clean criteria was used, shows sample locations selected largely north of the trench area where contamination is low or non-existent, and to the
Figure 4.22. Simulated SRS DOSB TCE (contours in ppb) surface and simulation sample area.
Carolina Bay (Wetlands)

Site North

Scale: feet

0 50 100 150 200

Wetland

Figure 4.23. DOSB phase 1 TCE values (ppb). The 0.5 values are nondetects.
southwest along the boundaries of the plume. Figure 4.24b, whereby the maximize the state contaminated criteria was used, shows sample locations selected largely within the plume but confined to the northern reaches of the plume southwest of the trenches, apparently being controlled by the TCE prior (identical to the PCE prior shown in Figure 4.10). Figure 4.24c, whereby the minimize the state uncertain criteria was used, shows sample locations selected roughly bounding the plume with low sample selection to the extreme southwest of the map area.

The David and Yoo sample selection was carried out using $p = 0.95$ in equation 2.6. The resulting sample locations are shown in Figure 4.24d. As expected, these samples are spread out in a relatively uniform manner throughout the sample area (as indicated in Figure 4.22).

The Englund and Heravi sample location selection algorithm was applied using the cost structure shown in Figure 2.6, as well as a modified cost structure based on Figure 2.6 with the false negative loss doubled. Doubling the false negative loss places emphasis on defining the remedial unit averages near the plume edges which have predicted concentrations below the action level (the RBC in this case) but might actually be above the action level. This is conservative toward protection of human health and the environment. Note that the relative loss from unit to unit is independent of the cost scale in Figure 2.6, thus, the sample location selection depends on the cost structure (i.e. linear, quadratic, etc.) but does not depend on the cost scale. In this simulation, a remedial unit is thought of as the volume of groundwater that might be captured by a pumping well for pump and treat remediation. Remedial units were defined as square blocks, 100 feet to a side, resulting in 57 remedial unit blocks covering the sample area. In practice, the optimal number and location of pumping wells is not known prior to the characterization of the site. Nevertheless, this (square) remedial unit size is probability smaller than the (near circular) zone of groundwater that would be captured by a pumping well and clearly includes many more remedial units (pumping wells) than would realistically be installed. However, this small remedial unit size should provide a better definition of the plume than would fewer, larger remedial units, many of which would likely be partially within and partially outside of the TCE plume. TCE values were converted to the log scale,
Figure 4.24. TCE sample location selection (■ shows 14 phase 1 initial samples). Range of influence = 100 feet.
and an isotropic covariance model was used with a range of influence of 100 feet and a TCE standard deviation of $\sqrt{3}$ ppb (see Figure 4.18, $n = 72$ case). Samples were selected at random from within the block with the greatest expected loss. The sample locations selected using this approach are shown in Figure 4.25. Comparison of Figure 4.25a and 4.25b indicates that doubling the false negative loss shows a slight tendency to place more samples outward from the margins of the plume.

The uncertainty zone sampling was implemented using both indicator kriging and ordinary kriging (equation 3.1) to estimate the probability that TCE exceeds the action level (on the log scale). Samples were collected from within the uncertain zone, defined by the probability $0.1 < P(\text{TCE} \geq \text{TCE RBC}) < 0.9$, where the kriging variance was greatest. The same indicator semivariogram as was used with Johnson's method (discussed previously in this section) was used with the indicator kriging uncertainty zone sampling. The same covariance used with the Englund and Heravi sampling (discussed previously in this section) was used with the ordinary kriging uncertainty zone sampling. The resulting sample locations are shown in Figure 4.25. Comparison of Figure 4.25c and 4.25d shows that both approaches provide a nearly uniform sampling density. Assuming the underlying assumptions are reasonably well met, the ordinary kriging approach is expected to provide better results because it utilizes the magnitude of the measured values as well as their spatial locations, while the indicator kriging method utilizes only zeros and ones together with the data locations. However, because the MDL (0.61 ppb) is close to the RBC (1.6 ppb), the ordinary kriging method may not perform well. Stein's method is better suited to data of this nature.

To investigate the effect of overestimating the range of influence on these sample location selection methods, each of the sampling methods discussed previously in this section was carried out using a range of influence of 200 feet (rather than 100 feet, as was used in the preceding discussion). The resulting data locations are shown in Figures 4.26 and 4.27. The David and Yoo method does not depend on a semivariogram or covariance model so is unchanged in this analysis (compare Figure 4.24d with Figure 4.26d). In general, it appears that increasing the range of
Figure 4.26. TCE sample location selection (● shows initial 14 phase 1 samples). Range of influence = 100 feet.
a. Johnson method - maximize state clean criteria.

b. Johnson method - maximize state contaminated criteria.

c. Johnson method - minimize state uncertain criteria.

d. David and Yoo method.

Figure 4.26. TCE sample location selection (● shows the 14 phase 1 initial samples). Range of influence = 200 feet.
Figure 4.27. TCE sample location selection (⊗ shows the 14 phase 1 initial samples). Range of influence = 200 feet.
influence provided a better distribution of sample locations relative to the TCE plume; compare Figures 4.24a and b with 4.26a and b, and Figures 4.25b and d with 4.27b and d, for example.

The effect of the lower state uncertain limit used with the ordinary kriging uncertainty zone sampling method was evaluated for probability levels of $p_1 = 0.1$, $0.2$, and $0.3$ for both the range of influence equal to 100 feet and 200 feet cases. The resulting sample locations are shown in Figures 4.28 and 4.29. Increasing both the range of influence and the lower uncertainty state probability level appear to focus the sampling on the plume with fewer surrounding samples; see for example Figure 4.28c and Figure 4.29b and c.

As a measure for comparison, the percent of grid nodes (25 foot, square node spacing) correctly classified as being less than, or greater than or equal to the TCE RBC for each dataset simulated in this section is listed in Table 4.2. The prediction calculations were made using ordinary kriging with the semivariogram parameters used to generate the simulated TCE plume. Table 4.2 indicates that the indicator and ordinary kriging based uncertainty zone sampling with $p_1 = 0.3$ yield the greatest percent of grid nodes correctly classified, while the same procedures with $p_1 = 0.1$ perform poorly, particularly when ordinary kriging is used. Using a larger value for $p_1$ (0.2 or 0.3, rather than 0.1) tends to force the sampling into the plume rather than away from the plume where the probability of contamination exceeding the RBC is high simply because of the low sampling density. In general, these methods seem to perform better when the larger range of influence is used.

4.1.3.4 Graphical Uncertainty Zone Sampling Method I propose a graphical approach to sample location selection based on sampling within the state uncertain region, as defined by the probability $p_1 < P(Z(s) \geq T) < p_2$, where $Z(s)$ is a contaminant concentration at location $s$ and $T$ is a threshold, such as an action limit or clean-up standard. The endpoints $p_1$ and $p_2$ are chosen to give an acceptable and achievable probability levels, e.g. $p_1 = 0.1$ and $p_2 = 0.9$ or so. The probability map may be constructed using ordinary kriging, indicator kriging, or the method of Stein. Samples are selected in such a way that they are not too close to existing samples, where the measure of
Figure 4.28. TCE sample location selection (● shows the 14 phase 1 initial samples). Range of influence = 100 ft.
Figure 4.29. TCE sample location selection (▧ shows the 14 phase 1 initial samples). Range of influence = 200 ft.
Table 4.2. Percent of TCE grid nodes correctly classified.

<table>
<thead>
<tr>
<th>Sampling Method</th>
<th>Percent of Nodes Correctly Classified as &lt; RBC or ≥ RBC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Range of Influence = 100 feet</strong></td>
<td></td>
</tr>
<tr>
<td>Indicator kriging uncertainty zone sampling, $p_1 = 0.3$</td>
<td>88.0</td>
</tr>
<tr>
<td>Ordinary kriging uncertainty zone sampling, $p_1 = 0.3$</td>
<td>87.5</td>
</tr>
<tr>
<td>Johnson maximize state contaminated</td>
<td>84.9</td>
</tr>
<tr>
<td>Englund and Heravi</td>
<td>84.6</td>
</tr>
<tr>
<td>Indicator kriging uncertainty zone sampling, $p_1 = 0.2$</td>
<td>84.5</td>
</tr>
<tr>
<td>Englund and Heravi, double false negative</td>
<td>82.6</td>
</tr>
<tr>
<td>Johnson maximize state clean</td>
<td>82.6</td>
</tr>
<tr>
<td>Johnson minimize state uncertain</td>
<td>82.5</td>
</tr>
<tr>
<td>Indicator kriging uncertainty zone sampling, $p_1 = 0.1$</td>
<td>82.5</td>
</tr>
<tr>
<td>David and Yoo *</td>
<td>82.2</td>
</tr>
<tr>
<td>Ordinary kriging uncertainty zone sampling, $p_1 = 0.2$</td>
<td>80.7</td>
</tr>
<tr>
<td>Ordinary kriging uncertainty zone sampling, $p_1 = 0.1$</td>
<td>80.7</td>
</tr>
</tbody>
</table>

**Range of Influence = 200 feet**

<table>
<thead>
<tr>
<th>Sampling Method</th>
<th>Percent of Nodes Correctly Classified as &lt; RBC or ≥ RBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indicator kriging uncertainty zone sampling, $p_1 = 0.3$</td>
<td>90.4</td>
</tr>
<tr>
<td>Ordinary kriging uncertainty zone sampling, $p_1 = 0.3$</td>
<td>88.9</td>
</tr>
<tr>
<td>Johnson maximize state contaminated</td>
<td>86.8</td>
</tr>
<tr>
<td>Indicator kriging uncertainty zone sampling, $p_1 = 0.2$</td>
<td>86.1</td>
</tr>
<tr>
<td>Indicator kriging uncertainty zone sampling, $p_1 = 0.1$</td>
<td>85.7</td>
</tr>
<tr>
<td>Englund and Heravi, double false negative</td>
<td>85.5</td>
</tr>
<tr>
<td>Englund and Heravi</td>
<td>85.4</td>
</tr>
<tr>
<td>Ordinary kriging uncertainty zone sampling, $p_1 = 0.2$</td>
<td>83.8</td>
</tr>
<tr>
<td>Johnson minimize state uncertain</td>
<td>82.5</td>
</tr>
<tr>
<td>David and Yoo *</td>
<td>82.2</td>
</tr>
<tr>
<td>Johnson maximize state clean</td>
<td>81.3</td>
</tr>
<tr>
<td>Ordinary kriging uncertainty zone sampling, $p_1 = 0.1$</td>
<td>80.5</td>
</tr>
</tbody>
</table>

* The David and Yoo method does not depend on a variogram or covariance model.

Closeness is about one quarter to one half the range of influence; to cover a larger area with fewer samples, a sample spacing of about the range of influence might be used. The basic approach is to start within the plume and work out until the plume is surrounded by a low probability contour, at which time the plume boundary is defined relatively well. Accordingly, this method also provides a convenient stopping rule. This method is ideally suited to phase 2 of an ESC project in that it can be quickly applied in the field, allows the user flexibility while providing guidance in the selection of future sample locations, and provides a frequently updated graphical representation of the contaminant conceptual model.
To demonstrate the method, I used the simulated TCE plume shown in Figure 4.22, and used the 14 phase 1 TCE values shown in Figure 4.23 as a starting dataset. Prior to data collection, we need to develop a semivariogram model. To begin, I assume a spherical semivariogram with a range of influence of 200 feet and a sill (TCE variance) of 1 ppb\(^2\) (about twice the phase 1 TCE variance). Ordinary kriging will be used to develop the probability maps. Figure 4.30 shows a probability map based on the 14 phase 1 TCE values using a threshold \(T\) equal to the TCE RBC of 1.6 ppb with probabilities greater than 0.9 or less than 0.2 shaded. Shaded circles with a radius of 67 feet (one third the assumed range of influence) are also shown. Unshaded regions represent good candidate sample locations. To better define the source area, five samples were selected within the shaded area near the trenches. Then, by starting within the plume and working out, I selected 15 additional sample locations, as illustrated in Figure 4.30. The phase 1 data along with the 20 samples just selected can be used to estimate a semivariogram, as shown in Figure 4.31, \(n = 34\) case. Note that samples can be selected one or several at a time using this method; I have selected 20 in this example to illustrate the process.

The 34 total TCE values and the estimated semivariogram shown in Figure 4.31, \(n = 34\) case, gives the probability map shown in Figure 4.32; areas with probabilities greater than 0.9 or less than 0.2 are shaded. The range of influence for the semivariogram (\(n = 34\)) in Figure 4.31 is about 150 feet. Accordingly, shaded circles with a radius of 50 feet (one third the range of influence) are shown centered at the sampled locations in Figure 4.32. Unshaded areas in Figure 4.32 represent good candidate sample locations from which 20 additional sample locations were selected, as shown in Figure 4.32. Pooling all the data together allows the semivariogram model to be updated, as shown in Figure 4.31, \(n = 54\) case.

The 54 total TCE values and the estimated semivariogram shown in Figure 4.31, \(n = 54\) case, gives the probability map shown in Figure 4.33; areas with probabilities greater than 0.9 or less than 0.2 are shaded. The range of influence for the semivariogram in Figure 4.31, \(n = 54\) case, is about 180 feet. Accordingly, shaded circles with a radius of 50 feet are shown centered at the
Figure 4.30. Probability that TCE exceeds 1.6 ppb based on the 14 phase 1 data - the shaded data point radius is 66 feet. Candidate sample locations are shown for 20 samples using the state uncertain spatial sampling method.
Figure 4.31. Semivariograms for the graphical uncertainty zone sampling simulation data.

sampled locations in Figure 4.33. Unshaded areas in Figure 4.33 represent good candidate sample locations from which 10 additional sample locations were selected, as shown in Figure 4.33. Pooling all the data together allows the semivariogram model to be updated, as shown in Figure 4.31, n = 64 case.

Repeating this process one more time gives the probability map shown in Figure 4.34. Note there is a probability contour of 0.3 surrounding the DOSB area in Figure 4.34. Our knowledge of the conditions by which the site became contaminated (dumping within the trenches) and the fact that the groundwater flow direction is to the southwest indicates little need to sample to the north, east, and northwest of the DOSB area, however, only by sampling within these areas can we be certain that they are contaminant free (even then, we can only be certain to within the sample measurement error and at the actual sample locations). An additional 10 sample locations are shown in Figure 4.34. This gives a total of 74 measurements from which the semivariogram shown in Figure 4.31, n = 74 case, was estimated.
Figure 4.32. Probability that TCE exceeds 1.6 ppb based on the phase 1 and sample 1 data (n=34) - shaded data point radius is 50 feet. Candidate sample locations are shown for 20 additional samples.
Candidate sample locations. 
Contour interval = 0.1 
Point size proportional to the square root of TCE

Figure 4.33. Probability that TCE exceeds 1.6 ppb (n=54) - shaded data point radius is 50 feet. Candidate sample locations are shown for 10 additional samples.
Figure 4.34. Probability that TCE exceeds 1.6 ppb (n=64) - shaded data point radius is 50 feet. Candidate sample locations shown for 10 additional samples.
Finally, because the detection limit (0.61 ppb) is near the RBC (1.6 ppb), I used Stein's method to generate the final probability map shown in Figure 4.35. Note that Stein's method can be used at each sampling event, however, the method may require several hours of computer time (e.g. overnight) and ordinary kriging may be preferable for quick field analyses. Figure 4.35 shows the entire TCE plume surrounded by a 0.1 probability contour. Comparing Figure 4.35 (based on 14 phase 1 and 60 phase 2 TCE values) with Figure 4.54 (based on the 73 TCE data values collected during phase 2) indicates that the plume boundaries are better defined with this simulation of uncertainty zone sampling data than with the actual ESC data. Figure 4.30 and Figures 4.32 through 4.35 show the development of a conceptual model, the data locations which support the model, and provide a means to assess whether additional data need to be collected, and if so, where.

Using ordinary kriging in an identical manner as described for the development of the data in Table 4.2, the final data from the graphical uncertainty zone sampling shown in Figure 4.35 yields 88.0 percent of the grid nodes correctly classified as being less than, or greater than or equal to the TCE RBC. Comparison with the results in Table 4.2 indicates that the graphical uncertainty zone sampling has provided good data for the characterization of the TCE plume.

4.1.4 SRS DOSB Post Phase 2 Applications

4.1.4.1 Geology Figures 4.37 to 4.48 show fence diagrams of CPT Olsen soil behavior classification, CPT dynamic pore pressure and geoprobe electrical conductivity. The locations of the fence diagrams is illustrated in Figure 4.36. The fence diagrams all show what I picked as the green clay. To do this, I assumed the green clay, being a marine deposit, is relatively horizontal and that the upper surface occurs within a region of interbedded or sandy soils at about 35 to 50 feet below ground surface. In general, there is often a slight, but abrupt, increase in the electrical conductivity, followed by a slow increase with depth below what I have classified as the green clay surface (e.g. see Figure 4.39). Often, but not always, there is an abrupt increase in the dynamic pore pressure upon entering the green clay. Examination of Figures 4.37 to 4.48 shows that it is not always clear
Figure 4.36. Probability that TCE exceeds 1.6 ppb (n=74) using Stein's algorithm. Probabilities less than 0.1 and greater than 0.9 are shaded.
Figure 4.36. SRS DOSB map showing locations of fence diagrams in Figures 4.37 to 4.48.
Figure 4.37. CPT and Geoprobe data south of the trench area.
CPT soil classification: <1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, after Olsen (1988).

Figure 4.38. CPT and Geoprobe data south of the trench area.
CPT soil classification: 1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, Olsen (1988).

Figure 4.39. CPT and Geoprobe data south of the trench area.
CPT soil classification: 1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, after Olsen (1988).

Figure 4.40. CPT and Geoprobe data south of the trench area.
CPT soil classification: <1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, after Olsen (1988).

Figure 4.41. CPT and Geoprobe data south of the trench area.

Figure 4.42. CPT and Geoprobe data southeast of the trench area.
Figure 4.43. CPT and Geoprobe data through the trench area.
CPT soil classification: <1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, after Olsen (1988).

Figure 4.44. CPT and Geoprobe data through the trench area.
CPT soil classification: <1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, after Olsen (1988).

Figure 4.46. CPT and Geoprobe data north of the trench area.
CPT soil classification: <1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, after Olsen (1988).

Figure 4.46. CPT and Geoprobe data along the road north of the trench area.
Figure 4.47. CPT and Geoprobe data north of the trench area.

CPT soil classification: <1 = Clays, 1-2 = Silt Mix, 2-3 = Sand Mix, 3-4 = Sands, 4-5 = Sand and Gravel, after Olsen (1988).
Figure 4.48. Geoprobe electrical conductivity data west of the trench area.
when the green clay has been penetrated. In addition, Figures 4.37 to 4.48 show regions where the ten foot unit is clearly present, and other regions where it is only weakly identified (see Figure 4.41 for example). There appears to be two distinct 25-30 foot units present. One gives a very definite elevated electrical conductivity peak, a definite elevation of dynamic pore pressures, and is generally associated with clayey soils (see Figure 4.40 for example); this unit is not present at all locations (see Figure 4.39 and 4.44 for example). I marked this abrupt 25-30 foot unit as "25" in Figures 4.37 to 4.48. The other 25-30 foot unit gives a more subtle, usually thicker, region of elevated electrical conductivities, little or no elevation in dynamic pore pressure, and is characterized by silt mix and sand mix soils types (compare from right to left in Figure 4.41 for example). Figures 4.37 to 4.48 show 69 of the 109 DPT results.

From a similar analysis of the remaining DPT results I was able to generate sufficient green clay surface elevation data to develop the semivariogram shown in Figure 4.49. Figure 4.50 shows the green clay surface elevation data and shows no strong evidence of a trend in the east-west or north-south directions. The dot plot in Figure 4.50 shows the green clay surface elevation to be approximately normally distributed. Ordinary kriging of the green clay surface elevation data using the semivariogram shown in Figure 4.49 and a one foot measurement error variance yields the green clay surface map shown in Figure 4.51 and the prediction standard deviation map shown in Figure 4.52.

An RFI/RA report (WSRC, 1996) has been written for the SRS DOSB. The WSRC (1996) report shows a contour plot of the green clay surface that varies from zero to about ten feet higher in elevation than the plot in Figure 4.51. The WSRC report does not provide any estimate of the prediction standard deviation. The WSRC green clay surface elevation was generally selected to be at the top of the interbedded soils directly below the 25' unit. I chose green clay surface elevations often within the interbedded soils where there was evidence of increasing electrical conductivity and/or increasing pore pressures indicative of low permeability soils. My attempt was to be consistent with the correlations relating the DPT results to the DOL-1 and DOL-2 soil boring logs.
The prediction standard deviation plot in Figure 4.52 fails to show all of the uncertainty associated with the determinations of the green clay surface. This is because of my assumption that the data used to perform the kriging are accurate to within a one foot measurement error variance. In fact, there appears to be a gradational boundary between the green clay and the overlying soils; differences as great as ten feet between Figure 4.51 and the WSRC (1996) green clay surface indicate that the actual uncertainty may be several times greater than indicated by the prediction error standard deviations shown in Figure 4.52. The apparent strong spatial correlation at low lags (less than about 50 feet) exhibited by the semivariogram in Figure 4.49 may be an artifact of my attempt to pick the upper green clay surface elevation with some continuity between adjacent DPT data. Thus, Figure 4.52 could result in a misleading conclusion of greater precision than actually exists with respect to the green clay surface elevation.

4.1.4.2 Contaminant Spatial Distribution Figure 4.53 shows a map of the predicted maximum TCE over sample depths using the Stein algorithm based on the semivariogram for n=72 shown in Figure 4.18. Figure 4.53 clearly shows the lateral extent of the TCE plume which emanates
Figure 4.50. SRS DOSB green clay surface elevation data locations and distribution (feet above msl); 91 observations.
Figure 4.61. Green clay surface elevation (feet above msl).
Figure 4.52. Green clay surface elevation prediction standard deviation (feet).
Site North Carolina Bay (Wetlands)

Figure 4.53. TCE prediction (ppb) by Stein's method. Light shading shows area predicted to be less than the RBC of 1.6.
from the trench area and follows the groundwater flow to the southwest of the mapped area. This map provides a firm basis for the placement of monitoring wells.

From the same application of the Stein algorithm used to generate Figure 4.53, I generated a probability map showing the probability that a TCE value at an unsampled location would exceed the TCE RBC. This probability map is shown in Figure 4.54. Note that Figure 4.54 shows similar features as the Stein probability map shown in Figure 4.20, but is based on all the data, whereas the map in Figure 4.20 was based on 15 fewer data values. This indicates that the last 15 or so TCE samples did not provide much additional information – this conclusion is in good agreement with the map discrepancy for TCE shown in Figure 4.21. Figure 4.20 showed that the TCE plume was quite well defined by the complete closure of the 0.1 probability contour surrounding the east, north, and west sides of the plume. Figure 4.54 also shows that the TCE plume is quite well defined. Recall, however, that the southeastern edge of the PCE plume was not yet well defined even when all the data collected were considered (see Figure 4.15 for example).

To assess the TCE distribution with depth, I constructed two cross sections running down the main portion of the plume from north to south. Figure 4.55 shows the locations of these cross sections, with the environmental sample locations denoted by an "E" followed by the number indicating the sequence in which the sample was collected. Figures 4.56 and 4.57 show these cross sections with the 10 foot, 25 foot, and green clay soils units shown, as determined from nearby DPT results. The dashed lines in Figures 4.56 and 4.57 separate the detect from nondetect data. As shown in Figures 4.56 and 4.57, there appears to be a downward trend associated with the TCE plume as the plume moves toward the southwest in the direction of the groundwater flow.

Note the relatively large TCE concentrations in Figure 4.56 directly below the trench area associated with sample location E003. This may be associated with residual DNAPL from stringers which may have migrated through the soil from a free product source in the trench at some time in the past.
Figure 4.64. Probability that TCE exceeds the TCE RBC by Stein’s method. Areas with low probability (< 0.1) and high probability (> 0.8) are shaded.
Figure 4.65. Cross section lines A-A' and B-B'. "E" locations are TCE sample locations. Samples were also collected in monitoring wells DOL-1 and DOL-2.
Figure 4.66. Simplified cross section showing major soil units and measured groundwater TCE concentrations (ppb). Dashed lines separate detect from nondetect (ND) results. Groundwater flows from A' to A.
**Figure 4.57.** Simplified cross section showing major soil units and measured groundwater TCE concentrations (ppb). Dashed lines separate detect from nondetect (ND) results. Groundwater flows from B’ to B.
Similar maps and cross sections can be constructed to develop the geologic and contaminant site conceptual models for other areas of the site, and for the other contaminants. Cross sections such as shown in Figures 4.56 and 4.57 provide a basis for optimal placement of screen intervals for monitoring wells.

4.1.4.3 Hydrogeology As mentioned previously, water elevations in the piezometers surrounding the DOSB indicated a water table depth of about 7.3 feet throughout most of the area during the phase 2 investigation. To assess the possibility of determining the depth to the water table on the basis of the CPT pore pressure data (called CPTU data below), I compared those portions of the dynamic pore pressure data that appear to be approximately hydrostatic versus depth. Robertson et al. (1992) state that excess dynamic pore pressure will dissipate as fast as they are generated in clean medium to coarse grained sands. At the SRS many of the sandy soils contain significant amounts of clays and silts, however, some portions of the pore pressure data taken in sandy soils appear to lie approximately along a line equal to the hydrostatic increase in pore pressure with depth. Figure 4.58 shows an example for CPTU push S003. The depth intercept and slope of this line will give an estimate of the water table depth and the pore pressure increase with depth, respectively. The increase in pore pressure with depth gives a direct measure of the vertical hydraulic gradient.

While a graphical approach, as illustrated in Figure 4.58, could be used to obtain these estimates, a simple linear regression will provide estimates as well as standard errors of the estimates of both the water table and increase in pore pressure with depth. Linear regression of pore pressure on depth will give the slope directly, but the depth intercept must be obtained by rearranging the regression equation: \( y = mx + b \), where \( y \) = the pore pressure, \( x \) = depth, \( m \) is the slope and \( b \) is the pore pressure intercept. We may solve for the water table depth, \( x_o \), at the intercept \( y = 0 \), as

\[
x_o = -\frac{b}{m}.
\]  

[4.3]
Intercept gives an estimate of the water table depth.

Slope gives an estimate of pore pressure increase with depth.

Olsen Soil Classification

Discard this data

Sandy soils with stable pore pressure
Use this data

Discard this data

Discard outlying data

Sandy soils Stable region Use this data

Figure 4.68. Pore pressure data for CPT S003.
An estimate of the standard error of $x_0$ is given by

$$\text{standard error of } x_0 = \frac{\text{standard error of } b}{m}. \quad [4.4]$$

The standard error of the slope, $m$, and the y intercept, $b$, are given directly by the linear regression. Equation 4.4 is approximate, but will be sufficient for the practical application given here.

There were only six SCAPS and three ARA pushes with pore pressures stable enough within selected sandy regions to apply linear regression to estimate the water table depth and pore pressure increase with depth. The results of the regressions (and application of equations 4.3 and 4.4) for these nine cases are given in Table 4.3. While there is no direct measurement of water table depth at the push locations for comparison, the estimated water table depths near ten feet at S001, S013, and S024 are about 2 to 2.5 feet greater than expected on the basis of the surrounding piezometers. Additionally, CPT pushes S001 and A001 were performed within about four feet of each other on nearly level ground; a water table change of $10.3 - 6.6 = 3.7$ feet over four feet lateral distance is highly unlikely. When subtracted from the ground surface elevation, the CPT derived water table elevations in Table 4.3 vary over a six foot range from 139.6 to 145.6 feet above msl while the piezometer data vary over a 1.5 foot range from 141.1 to 142.6 feet above msl over approximately the same region. For these reasons it appears as though these CPT derived water table depths may be accurate only to within about two feet or so. Note that this is a considerably greater uncertainty than is indicated by the standard errors listed in Table 4.3.

The three ARA estimates of pore pressure increase with depth are significantly less than the six SCAPS estimates (see Table 4.3). CPT pushes S001 and A001 were within about four feet laterally of one another and should give a similar response. Both A007 and S003 were pushed near the southeast corner of the site about 40 feet apart and both S007 and A015 were pushed in the northeast corner of the site about 150 feet apart. Accordingly, the disagreement between the ARA and SCAPS pore pressure increase with depth does not appear to be related to the geology. The most likely explanation is that the difference may be related to the calibration of the two different piezocone instruments. Had we performed this analysis on-site, we may have detected the problem.
<table>
<thead>
<tr>
<th>CPT Push ID</th>
<th>Water Table Depth (ft) *</th>
<th>Pore Pressure Increase with Depth (ft water/ft depth) *</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCAPS S001</td>
<td>10.3 (0.4)</td>
<td>1.23 (0.02)</td>
</tr>
<tr>
<td>SCAPS S003</td>
<td>7.4 (0.3)</td>
<td>1.24 (0.01)</td>
</tr>
<tr>
<td>SCAPS S004</td>
<td>7.7 (0.5)</td>
<td>1.35 (0.02)</td>
</tr>
<tr>
<td>SCAPS S007</td>
<td>7.1 (0.6)</td>
<td>1.15 (0.03)</td>
</tr>
<tr>
<td>SCAPS S013</td>
<td>10.0 (0.8)</td>
<td>1.32 (0.02)</td>
</tr>
<tr>
<td>SCAPS S024</td>
<td>9.7 (1.0)</td>
<td>1.27 (0.04)</td>
</tr>
<tr>
<td>ARA A001</td>
<td>6.6 (0.6)</td>
<td>1.02 (0.02)</td>
</tr>
<tr>
<td>ARA A007</td>
<td>7.7 (0.3)</td>
<td>1.05 (0.01)</td>
</tr>
<tr>
<td>ARA A015</td>
<td>8.3 (0.4)</td>
<td>0.98 (0.02)</td>
</tr>
</tbody>
</table>

* two times the standard error is given in parentheses

In time to resolve it through a calibration check on each CPT rig. While upward hydraulic gradients have been observed across some of the deeper aquitards at other areas within the Savannah River Site (WSRC, 1995), it seems unlikely that an upward hydraulic gradient as large as any indicated by the SCARS data would be observed within these near surface soils above the green clay. Thus, the ARA data are probably more representative of the actual site conditions.

Pore pressure dissipation tests were carried out by the SCAPS rig at various depths within soils that gave a large positive pore pressure response during a push. A pore pressure dissipation test is carried out by stopping the advancement of the push rod, and measuring the pore pressure with time while the excess pore pressure dissipates. In most cases, these tests were not carried out until complete dissipation of excess pore pressures, but they were carried out long enough to enable estimation of the hydraulic conductivity and coefficient of consolidation of the surrounding soil. A pore pressure dissipation test cannot be conducted in a (coarse grained) soil where excess pore pressures either don't develop or dissipate within a few seconds.

The pore pressure dissipation tests were used to estimate the hydraulic conductivity and coefficient of consolidation using the methodology given by Robertson and Campanella (1989) and
Robertson et al. (1992). While there are many factors which complicate the theoretical solutions, such as soil anisotropy, soil layering, nonlinearity, soil macrofabric, soil disturbance, clogging of the porous filter element, the CPTU dissipation test can provide an economic and useful means of approximating consolidation properties (Robertson et al., 1992). Figure 4.59a shows the CPTU dissipation test data for CPT push S001 at a depth of 58 feet. Figure 4.59b illustrates the square root of time method to obtain the initial pore pressure at time zero when the test is started. Note the redistribution of pore pressures whereby an initial increase in pore pressure is followed by a decrease with time in Figure 4.59b. In stiff, heavily overconsolidated soils, the pore-pressure gradient around a cone tip can be extremely large; it has been suggested that this may be responsible for the initial redistribution of pore pressure, however, poor saturation of the porous element can also cause a similar response (Robertson et al., 1992).

Figure 3 in the Robertson et al. (1992, page 543) paper may be used to estimate the horizontal coefficient of consolidation, $c_h$, on the basis of $t_{50}$, the time to reach 50 percent dissipation of the total excess pore pressure. Their Figure 3 gives a range of values corresponding to a rigidity index ranging from 50 to 500. Equations relating $t_{50}$ [minutes] to $c_h$ [cm$^2$/minute] in the Robertson et al. Figure 3 are:

$$c_h = 1.5 \times 10^{(\log(20)-\log(t_{50}))}$$  \hspace{1cm} [4.5]

for a rigidity index of 500, and

$$c_h = 1.5 \times 10^{(\log(6)-\log(t_{50}))}$$  \hspace{1cm} [4.6]

for a rigidity index of 50. Equations 4.5 and 4.6 are for a porous filter located just behind the cone tip (as with the SCAPS pore pressure sensor), and a 15 cm$^2$ (1.75 in diameter rod) cone.

In the absence of laboratory data for the rigidity index, we can use equations 4.5 and 4.6 to estimate a range of values for $c_h$. Because all the CPTU tests were carried out by the SCAPS rig, I used a static pore pressure, $u_o$, of 1.15 times the geostatic pore pressure to accommodate the results in Table 4.3. Whenever there was appreciable initial pore pressure redistribution I used the square root of time method to estimate the initial pore pressure, $u_i$, otherwise I simply used the
Hydrostatic pore pressure assuming a 7.3 foot deep water table = 22.0 psi.

Figure 4.69. a) S001 pore pressure dissipation test data at 58 feet bgs.
   b) Square root of time method to obtain the pore pressure at time zero.
maximum pore pressure to obtain $t_{50}$. The estimation of $t_{50}$ is illustrated in Figure 4.60 for the S001
data taken at a depth of 58 feet. In Figure 4.50, $u_i$ is the measured pore pressure at time $t$, and the
normalized excess pore pressure at time $t$, $(u_i - u_{eq})/\left(|u_i - u_{eq}|\right)$, is plotted along the vertical axis.

An estimate of the hydraulic conductivity is given by

$$k_h = c_h m_h \gamma_w,$$

[4.7]

where $m_h$ is the volumetric compressibility and $\gamma_w$ is the unit weight of water. An estimate of $m_v$ may
be obtained from $m_v = 1/(\alpha q_c)$, where $\alpha$ is value adapted from work by Sangerat and may be found
in Table 4.3, page 106, of Robertson and Campanella (1989), and $q_c$ is the measured CPT tip stress
in bars. Table 4.3 in Robertson and Campanella gives a range of values for $\alpha$ for each soil type
and $q_c$ listed. Using the procedure described in the previous paragraphs, I calculated several
estimates of $c_h$ using equations 4.5 and 2.24. Using these two estimates of $c_h$ together with the
lower and upper values of $\alpha$ from Table 4.3 of Robertson and Campanella, I calculated two
estimates of the hydraulic conductivity based on equation 4.7, using the combination of $\alpha$'s and $c_h$'s
which give the largest and smallest values for $k_h$.

Table 4.4 shows the CPT ID, CPTU dissipation test depth, Olsen soil type, initial pore
pressure $u_i$, equilibrium pore pressure $u_{eq}$, measured tip stress $q_c$, $t_{50}$, $\alpha$, and estimated $c_h$ and $k_h$ for
all 63 of the CPTU tests that were conducted during phase 2 at the SRS DOSB. The rows in Table
4.4 are sorted by CPTU test depth. Figure 4.61 shows push locations and the maximum of the two
estimated hydraulic conductivity values for CPTU dissipation tests conducted in the green clay as
identified in Figure 4.51. The minimum (over sample depths) hydraulic conductivity is plotted in
Figure 4.61 in those cases where several tests were conducted at one push location. The hydraulic
conductivity values plotted in Figure 4.61 are consistent with silts and low permeability silty sands,
but are just above the values for unweathered marine clays, as reported by Freeze and Cherry
(1979, Table 2.2).
Figure 4.60. Graphical method to determine the time to reach 50 percent of the total pore pressure dissipation (S001 data at a depth of 58 ft).
Figure 4.61. Estimated hydraulic conductivity (10^-6 cm/s) from the CPTU dissipation testing within the green clay.
228

Table 4.4. Pore pressure dissipation test results.
Depth Soil Type
CRT
Qc
"/
"o
(psi) (psi) (bar)
Push
(feet) (Olsen)
0.9 51.8
S006
8.6 day
220.1
108.0
6.8 19.3
S004
20.7 day
103.6
9.7 62.6
S027
22.3 day
9.0 18.9
S023
2Z8 silt mix
110.2
8.3 29.2
S005
24.0 silt mix
151.2
8.3 30.2
S013
128.3
24.0 silt mix
8.8 16.3
30.7
S024
24.7 day
66.2
9.0 30.5
25.0 silt mix
S022
154.0
9.8 22.3
SOOS
27.1 silt mix
9.8 38.4
203.3
S022
27.2 day
95.6
11.3 19.7
SOI 9
29.6 day
8010
30.0 sand mix
26.5
11.3 13.5
73.0 12.3 13.2
8012
31.8 silt mix
8006
32.3 sand mix
44.0
13.4 23.1
13.7 25.7
8007
32.9 sand mix
42.6
8026
34.4 day
31.2 13.3 61.0
14.3 45.5
8021
36.0 sand mix
112.5
8020
36.2 sand mix
71.3
13.9 58.0
8007
36.6 silt mix
63.5
14.8 16.0
8023
36.9 silt mix
218.1
14.8 55.8
S006
39.1 silt mix
68.1
15.8 19.5
16.9
6.5
SOU
39.5 day
29.3
90.6
16.3 32.0
8014
39.9 silt mix
8017
40.1 silt mix
80.1
17.1 19.7
9.3 91.7
8020
41.1 sand mix
96.7
8015
48.9
9.3
43.1 sand mix
20.4
9.8
8006
44.8 silt mix
109.3
18.8
8017
45.0 silt mix
46.6
18.8
9.3
8007
47.6 silt mix
90.6
20.3 11.9
8008
130.6
19.5 42.1
48.3 sand mix
8014
49.4 sand mix
145.6 20.8 50.0
8008
49.6 sand mix 314.8
21.3 40.4
8011
49.9 silt mix
271.2 21.3 42.6
8017
50.1 silt mix
150.4
21.3 17.8
S004
252.9 21.3 33.1
50.2 silt mix
8010
50.5 silt mix
178.9
15.8 34.7
S006
51.0 silt mix
154.0 21.8 18.7
8007
51.0 silt mix
117.7
21.8 144
S009
51.1 silt mix
296.6
21.8 40.4
S013
54.1 silt mix
154.8
23.3 21.3
8002
57.8 sand mix
177.9 25.3 36.2
8001
58.0 sand
99.9
25.3 77.4
8003
60.0 sand
30.6
26.3 59.6
8005
168.5
60.0 sand mix
26.3 47.2

%0
(min)
2.39
4.66
1.11
0.38
2.54
1.04
0.29
0.26
0.49
1.21
0.58
0.09
2.44
0.18
0.16
0.16
7.34
0.21
0.38
0.76
0.19
0.19
0.34
0.23
1.18
1.73
1.66
0.18
0.94
1.48
0.39
19.33
2.38
22.91
0.84
0.49
1.78
3.31
22.74
0.71
0.58
0.68
0.04
1.09

a
1 -2.5
2 -5
1 -2.5
1 -3
3 -6
3 -6
2 -5
3 -6
3 -6
1 -2.5
2 -5
1 -3
1 -3
3 -6
3 -6
1 -2.5
3 -6
3 -6
1 -3
3 -6
1 -3
3 -8
3 -6
1 -3
3 -6
1 -3
1 -3
1 -3
1 -3
3 -6
3 -6
3 -6
3 -6
1 -3
3 -6
3 -6
1 -3
1 -3
3 -6
3 -6
3 -6
3 -6
3 -6
3 -6

(crti^/min)
3.8
13
1.9
6.4
8.1
27
79
24
3.5
12
8.7
29
100
31
35
120
18
61
25
7.4
16
52
330
100
3.7
12
50
170
56
190
56
190
1.2
4.1
43
140
79
24
39
12
160
47
160
47
88
26
39
130
7.6
25
5.2
17
5.4
18
50
170
9.6
32
20
6.1
23
77
0.47
1.6
3.8
13
0.39
1.3
36
11
18
61
5.1
17
2.7
9.1
0.4
1.3
13
42
52
16
13
44
750
230
8.3
28

kh
(10^ cm/sec)
0.48
4
0.33
2.7
0.85
7.1
6.8
68
0.33
2.2
0.78
5.2
6.2
52
3.1
21
15
2.2
1.3
11
2.6
21
400
40
1.5
15
5.9
39
6.0
40
6.0
50
0.073
0.49
13
2.0
8.1
81
0.58
3.8
13
130
15
130
2.3
15
110
11
0.23
1.5
3.1
31
3.0
30
29
290
4.4
44
0.39
2.6
1.3
8.4
0.031
0.21
0.24
1.6
0.12
1.2
0.88
5.9
9.6
1.4
1.5
15
1.0
10
0.027
0.18
1.6
11
1.2
7.8
3.1
0.47
10
69
0.48
3.2


Table 4.4. (continued)

<table>
<thead>
<tr>
<th>CPT Push</th>
<th>Depth (feet)</th>
<th>Soil Type</th>
<th>$u_l$ (psi)</th>
<th>$u_o$ (psi)</th>
<th>$q_c$ (bar)</th>
<th>$t_{SO}$ (min)</th>
<th>$\alpha$</th>
<th>$c_h$ (cm$^2$/min)</th>
<th>$k_h$ ($10^{-6}$ cm/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S008</td>
<td>60.0</td>
<td>sand mix</td>
<td>298.2</td>
<td>26.3</td>
<td>32.2</td>
<td>1.31</td>
<td>3 - 6</td>
<td>6.9</td>
<td>23</td>
</tr>
<tr>
<td>S009</td>
<td>60.0</td>
<td>sand</td>
<td>160.0</td>
<td>26.3</td>
<td>113.2</td>
<td>1.94</td>
<td>3 - 6</td>
<td>4.6</td>
<td>15</td>
</tr>
<tr>
<td>S010</td>
<td>60.0</td>
<td>sand</td>
<td>250.6</td>
<td>26.3</td>
<td>65.4</td>
<td>0.64</td>
<td>3 - 6</td>
<td>14</td>
<td>47</td>
</tr>
<tr>
<td>S011</td>
<td>60.0</td>
<td>sand mix</td>
<td>168.4</td>
<td>26.3</td>
<td>42.1</td>
<td>1.24</td>
<td>3 - 6</td>
<td>7.3</td>
<td>24</td>
</tr>
<tr>
<td>S014</td>
<td>60.0</td>
<td>sand mix</td>
<td>110.7</td>
<td>26.3</td>
<td>22.8</td>
<td>0.89</td>
<td>3 - 6</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>S017</td>
<td>60.0</td>
<td>sand mix</td>
<td>161.4</td>
<td>26.3</td>
<td>49.4</td>
<td>1.13</td>
<td>3 - 6</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>S020</td>
<td>60.0</td>
<td>sand</td>
<td>237.8</td>
<td>26.3</td>
<td>75.5</td>
<td>0.34</td>
<td>3 - 6</td>
<td>26</td>
<td>88</td>
</tr>
<tr>
<td>S021</td>
<td>60.0</td>
<td>sand mix</td>
<td>47.8</td>
<td>26.3</td>
<td>6.7</td>
<td>2.34</td>
<td>1 - 3</td>
<td>3.8</td>
<td>13</td>
</tr>
<tr>
<td>S023</td>
<td>60.0</td>
<td>sand mix</td>
<td>219.0</td>
<td>26.3</td>
<td>28.8</td>
<td>1.54</td>
<td>3 - 6</td>
<td>5.8</td>
<td>19</td>
</tr>
<tr>
<td>S025</td>
<td>60.0</td>
<td>sand mix</td>
<td>220.0</td>
<td>26.3</td>
<td>30.7</td>
<td>15.00</td>
<td>3 - 6</td>
<td>0.6</td>
<td>2</td>
</tr>
<tr>
<td>S027</td>
<td>60.0</td>
<td>sand mix</td>
<td>396.9</td>
<td>26.3</td>
<td>36.1</td>
<td>3.49</td>
<td>3 - 6</td>
<td>2.6</td>
<td>8.6</td>
</tr>
<tr>
<td>S002</td>
<td>60.1</td>
<td>sand</td>
<td>60.3</td>
<td>26.3</td>
<td>63.1</td>
<td>0.58</td>
<td>3 - 6</td>
<td>16</td>
<td>52</td>
</tr>
<tr>
<td>S006</td>
<td>60.1</td>
<td>sand</td>
<td>198.3</td>
<td>26.3</td>
<td>76.0</td>
<td>0.61</td>
<td>3 - 6</td>
<td>15</td>
<td>49</td>
</tr>
<tr>
<td>S007</td>
<td>60.1</td>
<td>clay</td>
<td>60.2</td>
<td>26.3</td>
<td>4.2</td>
<td>3.94</td>
<td>3 - 6</td>
<td>2.3</td>
<td>7.6</td>
</tr>
<tr>
<td>S012</td>
<td>60.1</td>
<td>sand</td>
<td>434.0</td>
<td>26.3</td>
<td>53.6</td>
<td>1.54</td>
<td>3 - 6</td>
<td>5.8</td>
<td>19</td>
</tr>
<tr>
<td>S015</td>
<td>60.1</td>
<td>sand</td>
<td>154.5</td>
<td>26.3</td>
<td>70.2</td>
<td>1.03</td>
<td>3 - 6</td>
<td>8.7</td>
<td>29</td>
</tr>
<tr>
<td>S018</td>
<td>60.1</td>
<td>sand mix</td>
<td>335.1</td>
<td>26.3</td>
<td>41.5</td>
<td>3.64</td>
<td>3 - 6</td>
<td>2.5</td>
<td>8.2</td>
</tr>
<tr>
<td>S019</td>
<td>60.1</td>
<td>sand mix</td>
<td>210.2</td>
<td>26.3</td>
<td>29.2</td>
<td>0.64</td>
<td>3 - 6</td>
<td>14</td>
<td>47</td>
</tr>
<tr>
<td>S024</td>
<td>60.9</td>
<td>sand</td>
<td>162.9</td>
<td>26.8</td>
<td>66.5</td>
<td>1.08</td>
<td>3 - 6</td>
<td>8.3</td>
<td>28</td>
</tr>
</tbody>
</table>

4.1.5 Summary of SRS DOSB Statistical Applications

The abundance of nondetect data is good, of course, from an environmental standpoint, but nondetect data makes it difficult to estimate spatial structure and generate contaminant concentration maps, prediction standard deviation maps, and probability maps. Of course, contaminant concentration maps can be generated with spline functions, inverse distance weighting schemes, or other methods, but contouring such data implies a belief in some underlying continuity. Without describing the underlying structure via a variogram or covariance function, we cannot develop prediction variance or probability maps.

The geology at the SRS consists of a variety of interbedded, often discontinuous, soil types. While the DPTs were able to characterize the soil deposits at a given location, it was often difficult to find continuity between adjacent pushes. In addition, the two calibrating boreholes, DOL-1 and DOL-2, appear to give partially conflicting information, making identification of the green clay aquitard
difficult. It appears as though the green clay surface is not well defined and may best be characterized as a gradational contact within interbedded sands, silts, and clays with hydraulic conductivities indicative of silty fine grained sands. Two to four additional deep boreholes would have provided useful information for calibrating the DPTs — either to clarify, or confirm, the ambiguity.

The ESC phase 2 work ceased because the planned activities were largely accomplished and the time and money allotted to conduct the field work had run out. While several of the statistical tools discussed previously can be used to show that the site was reasonably well characterized, they also show that the characterization could have been improved, possibly with about 20 percent less contaminant data than was actually collected. The discrepancy function for the PCE and TCE data shown in Figure 4.21 indicates that the TCE plume is relatively well defined with about 50 of the 73 total sample locations, but the PCE plume remains less well defined. Figure 4.20 shows that TCE plume was relatively well defined and bounded (within the pre-arranged "site" boundaries) after about 80 percent of the data had been collected.

Figure 4.13 shows that, with only about 60 percent of the total PCE data collected, probability maps constructed with Johnson's updating procedure indicate that additional data southeast of the trench area is needed to define the edge of the PCE plume in that region, while the north and west edge of the plume is relatively well defined. The need for data southeast of the trench area is also clearly shown in subsequent Figures 4.14 and 4.15. Recall that knowledge of the soil gas data suggested a need for samples near the northeast corner of the DOSB area and groundwater flow direction suggested a need to place samples further to the southwest. Both the TCE and PCE data clearly show the plume extending off the southwest portion of the phase 2 work area in the direction of the groundwater flow (e.g., see Figures 4.15 and 4.53). Plots such as Figures 4.53 to 4.57 can be used to determine optimal placement and screen intervals for monitoring wells.

It seems clear that on-site use of Johnson's method, Stein's method, and indicator kriging would have helped to identify areas where additional samples might have been collected. Ordinary
Kriging of the upper green clay surface was carried out during the phase 2 investigation, however, the large uncertainty in what was called green clay on the basis of the DPTs rendered the green clay surface plot questionable. Nevertheless, the green clay surface plot generated on-site was used to help determine sample depths for the groundwater samples. Also, indicator kriging was used on-site, as illustrated in Figure 4.16 for the PCE data, to identify areas where additional samples should be collected using the Johnson "state uncertain" criteria. During the last week on-site, the indicator kriging identified the area southwest of the trenches as needing additional samples, and samples were placed there, however, even more sampling in this region was needed to bound the PCE plume.
5. ASSESSMENT OF METHODOLOGY: ADVANTAGES AND LIMITATIONS

In this chapter I will briefly review some of the investigation methods and tools which are often used in a site characterization. My focus will be the advantages and limitations of these methods and how they might best be incorporated into the ESC process. As these tools are well described elsewhere, I will give only brief descriptions of them here.

5.1 Geophysical Tools

Geophysical methods include a range of surface and borehole measurement techniques for investigation of the subsurface. Benson (1993) provides a concise review of the commonly used methods which have been proven effective for waste site assessments. Surface geophysical methods include ground penetrating radar (GPR), electromagnetics (EM), resistivity, seismic refraction and reflection, micro-gravity, metal detection, and magnetics. Downhole geophysical methods include natural gamma, gamma-gamma, neutron, induction, resistivity, single point resistance, spontaneous potential, temperature, fluid conductivity, flow, and caliper. Many geophysical measurements can be made relatively quickly allowing a much greater sampling density than can be achieved by the collection of individual samples and selected boreholes. Geophysical measurements generally measure a greater volume of material than soil or borehole samples, and coupled with a greater sampling density, provide a greater likelihood of detecting anomalous conditions. Surface geophysical methods provide non-destructive, in-situ, measurements of physical, electrical, or geochemical properties of the natural or contaminated soil or rock. Which method, if any, is applicable to a given problem depends on the information sought and the existing site conditions. While the borehole geophysics are themselves non-destructive, they require the existence of a borehole in which to make the measurements.
These geophysical tools have proved successful for initial site surveys if site conditions are amenable, but require that experienced personnel design and interpret the surveys. A major inhibiting factor to the general use of geophysical techniques has historically been choice of inappropriate techniques due to not understanding the limitations, however, selection of the proper techniques implemented by trained personnel can increase the level of understanding of a site without incurring the large costs often associated with installation of soil borings and monitoring wells (Preslo and Stoner, 1991).

5.1.1 Surface Geophysical Methods

The following geophysical method summaries are taken primarily from Benson (1993), Technos (1992, volume 1), and Technos (1993).

**GPR**  GPR radiates high frequency electromagnetic waves into the subsurface. Energy reflected back to the surface is received by an antenna, the signal is recorded, and a cross sectional picture of shallow conditions may be produced. GPR reflections occur whenever there is a change in the dielectric constant or electrical conductivity between two materials. The "depth" to a reflector is measured as a two-way travel time for the electromagnetic wave; to convert this to a depth requires either some calibration measurement (such as a logged borehole) to determine the velocities of the electromagnetic wave through the site specific materials, or assumptions about these velocities. GPR may be used to detect areas of disturbed soil, metallic or non-metallic buried drums or tanks, and has some use in mapping contaminant plumes.

**Electromagnetic Methods**  Two types of electromagnetic methods are in use: time-domain and frequency-domain. Both induce electric currents in the ground by electromagnetic induction whereby the subsurface conductivity, or its reciprocal, resistivity, is measured. With the frequency-domain method a transmitter continually radiates energy into the ground, while a receiver measures changes in the magnitude of the currents induced within the ground. With the time-domain system, the transmission of energy into the ground is cycled on and off, while a receiver measures changes
in the induced currents as a function of time. Electromagnetic instruments do not require electrical contact with the ground, making rapid measurement possible. Both profiling and sounding may be performed. Profiling refers to mapping lateral variations in subsurface conductivity at a given measurement depth. Sounding refers to determining the vertical changes in electrical conductivity for correlation with rock or soil types.

**Resistivity** The resistivity method measures subsurface electrical resistivity by passing an electric current into the ground from a pair of surface electrodes. The voltage in the ground surface due to the current is measured by a second pair of electrodes. Greater electrode spacing gives a greater depth measurement. The apparent resistivity of the subsurface materials is calculated based on the geometry of the electrode array, the applied current, and the measured voltage. The resistivity method can be used for both profiling and sounding, as described in the previous paragraph.

**Seismic Refraction and Seismic Reflection** Seismic refraction and reflection may be used to determine the depth to bedrock, the depth to the water table, or assess the continuity of geologic strata, faults, and buried channels. Seismic waves are transmitted into the subsurface where they are refracted or reflected when they pass from one material to another which has a different seismic velocity. An array of geophones on the surface measures the travel times of the seismic waves from the source to the geophones. For most refraction work, the first arriving compression waves are used, while for reflection work, the latter arriving compression waves are used. Seismic refraction will not detect a lower velocity layer underlying a higher velocity layer and may not detect a thin layer, however, if a sufficient velocity contrast between layers exists, up to three of four soil and rock layers may be determined. The depth to a layer where seismic refraction or reflection occurs is measured as the time required for the seismic wave to travel from the source to the geophone. This must be converted to a depth by correlation with known information (such as a borehole), or some assumptions, regarding the seismic velocities of the subsurface materials.
A seismic reflection survey measures two-way travel times for seismic waves reflected from subsurface materials and is capable of much deeper investigations with less seismic source energy. The method is commonly applied to depths of greater than 15 to 30 meters. While a high frequency energy source improves vertical resolution, the method is limited by the ability to transmit high frequency energy into soil and rock, particularly if there is loose soil near the surface.

**Micro-Gravity** Micro-gravity measures changes in the earth's gravitational field caused by changes in the density of soil and rock. The gravimeter is a very sensitive instrument designed to measure extremely small changes in the gravitational field. The gravimeter must be thermostatically controlled and may be affected by ground noise, wind, and temperature. To compensate for instrument drift, measurements must be made at a base station each hour or so. Earth tides and changes in elevation also affect the instrument. Gravity data may be presented as a profile or as a contour map. Micro-gravity is used to detect geologic anomalies such as bedrock channels, fractures, and cavities.

**Metal Detection** Metal detection may be used to locate buried cables, pipes, drums, property stakes, and other metal debris, including delineating the boundaries of trenches containing metal debris. The method can detect both ferrous and non-ferrous, e.g. Cu and Al, metal. Metal detectors have a short range whereby an object the size of a coin may be detected at about 0.5 to 1 meter, and objects such as 55 gallon drums may be detected up to about one to three meters. Metal detectors may be affected by nearby metal fences, pipes, vehicles, buildings, and possibly by changing soil conditions.

**Magnetometry** A magnetometer measures the intensity of the earth's magnetic field and may be used to map regional magnetic conditions. The common use at waste sites is to detect buried drums, tanks, and pipes. The magnetometer responds only to ferrous metals and will not detect non-ferrous metals because ferrous metals create a local variation in the earth's magnetic field while non-ferrous metals do not. Either total of gradient magnetic measurements are commonly made. Total field measurements respond to the total surrounding magnetic field, which may be
affected by natural and cultural magnetic noise. Gradient measurements are made by a gradiometer which consists of two magnetic sensors separated vertically or horizontally by 1 to 2 meters. Gradient measurements are insensitive to natural changes and minimize most cultural effects. The gradiometer responds to the local gradient, it is better able to locate a small target, such as a buried drum, however, it is less sensitive than a total field instrument. Under ideal conditions, a single drum can be detected at depths up to six meters with a total field magnetometer, and up to about three meters with a gradiometer. Massive piles of drums may be detected up to about 15 meters with a total field magnetometer, and about seven meters with a gradiometer.

**Thermal** Thermal infrared sensors may be used to measure ground temperature variations to locate fractures, caves and seeps. Measurements are easy to make by either hand carried or vehicle mounted instruments which do not require any ground contact. Thermal sensing ground probes may be installed for time series measurements or to obtain thermal gradient data. Thermal measurements are sensitive to diurnal and seasonal temperature changes.

Table 5.1 lists uses, advantages, and limitations for the surface geophysical methods commonly used in waste site characterizations.

### 5.1.2 Borehole Geophysical Methods

Several borehole logging methods are available for determining the characteristics of soil, rock, or fluid along the length of a boring or monitoring well. These downhole measurements can sometimes be correlated with known geologic information from one boring, and then used to identify and correlate with geologic strata from another boring without soil or rock samples. All of these methods may be used in uncased boreholes, however, some may not be applicable in boreholes cased with either PVC or steel or both.

The following summary of borehole geophysical methods is taken from Benson (1993) and Technos (1992, volume 2).
<table>
<thead>
<tr>
<th>Geophysical Method</th>
<th>Use</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPR</td>
<td>Profiling and mapping to about 30 m depth</td>
<td>Highest resolution of any surface geophysical method</td>
<td>Penetration limited by soil conditions, less for wet or conductive soils</td>
</tr>
<tr>
<td></td>
<td>Detect water table in coarse grained soils</td>
<td>Preliminary field analysis is possible</td>
<td>Interface must have sufficient contrast to show in radar profile</td>
</tr>
<tr>
<td></td>
<td>Detect metallic and non-metallic buried objects such as pipes and drums</td>
<td>Data may be obtained at a rapid speed</td>
<td>Overhead reflections from trees and power lines may show on the record</td>
</tr>
<tr>
<td></td>
<td>Cross sections of natural and anomalous geologic conditions</td>
<td>Measurements are easy to make</td>
<td>Two-way travel time depth scale must be calibrated or assumed</td>
</tr>
<tr>
<td></td>
<td>Some use in mapping contaminant plumes</td>
<td>May be performed through asphalt and concrete under some conditions</td>
<td>Penetration depth is site specific</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Does not require intrusive ground contact</td>
<td>Lower frequency gives greater penetration depth with lower resolution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Measures a smaller volume than resistivity for a given depth and provides better lateral resolution</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires less space for soundings than resistivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Very rapid measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No intrusive ground contact needed (but may come into contact with the ground)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Station measurements up to about 70 m depth</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May make measurements over time to monitor plume dynamics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires less space for soundings than resistivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Very rapid measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No intrusive ground contact needed (but may come into contact with the ground)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Station measurements up to about 70 m depth</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May make measurements over time to monitor plume dynamics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires less space for soundings than resistivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Very rapid measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No intrusive ground contact needed (but may come into contact with the ground)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Station measurements up to about 70 m depth</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May make measurements over time to monitor plume dynamics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires less space for soundings than resistivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Very rapid measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No intrusive ground contact needed (but may come into contact with the ground)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Station measurements up to about 70 m depth</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May make measurements over time to monitor plume dynamics</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires less space for soundings than resistivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Very rapid measurements</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No intrusive ground contact needed (but may come into contact with the ground)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Station measurements up to about 70 m depth</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May make measurements over time to monitor plume dynamics</td>
<td></td>
</tr>
<tr>
<td>Geophysical Method</td>
<td>Use</td>
<td>Advantages</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Resistivity</strong></td>
<td>Measure subsurface resistivity</td>
<td>No depth limit -- depends on electrode geometry,</td>
<td>Requires good ground contact and long electrode arrays</td>
</tr>
<tr>
<td></td>
<td>Sounding or profiling</td>
<td>commonly less than 300 m</td>
<td>Integrates a large volume of subsurface materials</td>
</tr>
<tr>
<td></td>
<td>Mapping of 2 to 3 layers</td>
<td>Provides better vertical resolution than frequency domain EM</td>
<td>Affected by cultural features, e.g. metal, buildings, power lines, vehicles</td>
</tr>
<tr>
<td></td>
<td>Mapping hydrogeologic anomalies</td>
<td>May be less sensitive to cultural interferences than EM</td>
<td>Deep soundings require a large electrode array</td>
</tr>
<tr>
<td></td>
<td>Measure depth and thickness of geologic strata</td>
<td></td>
<td>Cannot be used over asphalt or concrete</td>
</tr>
<tr>
<td></td>
<td>Locate buried wastes</td>
<td></td>
<td>Less effective at very low resistivity values (use EM)</td>
</tr>
<tr>
<td><strong>Magnetometry</strong></td>
<td>Detect buried ferrous metal objects</td>
<td>May detect a single drum up to 6 m deep</td>
<td>Detects ferrous metals only</td>
</tr>
<tr>
<td></td>
<td></td>
<td>May detect massive piles of drums up to 15 m deep</td>
<td>May be affected by magnetic minerals in the soil, steel debris, pipes, fences, buildings, vehicles, and changes in the earth's magnetic field</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Measurements are easy to make</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No intrusive ground contact</td>
<td></td>
</tr>
<tr>
<td><strong>Metal Detection</strong></td>
<td>Locating buried metal objects</td>
<td>Will detect ferrous and non-ferrous metal</td>
<td>Short detection range -- coin size objects from 0.5 to 1 m depth, drums up to 3 m depth</td>
</tr>
<tr>
<td></td>
<td>Detect buried drums, define boundaries of trenches containing metal debris</td>
<td>No intrusive ground contact</td>
<td>Affected by nearby metal objects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provides better spatial definition than EM</td>
<td>May be affected by changing soil conditions</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Geophysical Method</th>
<th>Use</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Seismic Refraction</strong></td>
<td>Determine depth to bedrock, depth to water table, continuity of strata, locate faults Determine bedrock surface Characterize rock type and degree of weathering by seismic velocity Generally applied to shallow (&lt; 100 m) investigations Depth and thickness of geologic strata, structure and anomalous conditions Measure in-situ elastic moduli of soil and rock</td>
<td>Up to 3 to 4 layers of soils and rock can be resolved Can provide a depth under each geophone Simple source of seismic energy can often be used (i.e. a sledge hammer)</td>
<td>Seismic velocity must increase with deeper layers to detect them Seismic wave travel time must be calibrated on-site to obtain depths Cannot detect thin layers Deep surveys require a great energy source Sensitive to ground vibrations from vehicles, wind, waves Long survey lines, about 4 to 5 times the measurement depth Requires intrusive ground contact Deep measurements may require an explosive energy source</td>
</tr>
<tr>
<td><strong>Seismic Reflection</strong></td>
<td>Profiling and mapping soil and rock Measure depth and thickness of geologic strata, structural, and anomalous conditions Measurements from about 15 m to 1000 m or more</td>
<td>Deeper investigation with less energy than seismic refraction Can provide relatively detailed geologic cross sections to depths of about 200 m Shorter survey line than seismic refraction</td>
<td>Sensitive to ground vibrations, e.g. vehicles, trains Two-way travel time depth scale must be calibrated Ability to transmit energy into soil or rock, loose surface soil limits resolution Intrusive ground contact May require extensive data processing</td>
</tr>
<tr>
<td><strong>Micro-Gravity</strong></td>
<td>Detect and map shallow localized geologic anomalies such as bedrock channels, fractures, and cavities</td>
<td>Data may be presented as contours or as a profile No depth limit, commonly used to 300 m or so May characterize conditions in geologic or cultural environments where other methods would fail Can be used inside buildings</td>
<td>Very delicate instrument, subject to shock damage Affected by temperature variations, wind, ground vibrations, earth tides, and changes in elevation Slow, extensive data reduction Need a base station for instrument drift correction Slow and tedious field operation</td>
</tr>
</tbody>
</table>
Natural Gamma The natural gamma log measures the natural gamma radiation emitted by soils and rocks. Dominant gamma emitting radioisotopes are potassium 40 and daughter products of the uranium and thorium decay series. The natural gamma log is used primarily to identify high clay content soils and shales which are typically high in potassium. The log may be taken in a PVC or steel cased hole, under either saturated or unsaturated conditions. Potassium feldspar sands may also be detected by the natural gamma log (Ault, Logan, and Madaj, 1994).

Gamma-Gamma (Density log) The gamma-gamma log may be used in cased or uncased holes above or below the water table. The probe has a gamma radiation source and detector which provides a response in counts per second. The count rate is used to indicate relative density, and can be calibrated to give the actual density of the in-situ material. Borehole diameter variations and well construction factors can affect this log.

Neutron-Neutron (Porosity log) The neutron-neutron probe has a neutron source and a neutron detector which provides a response in counts per second. The count rate is inversely proportional to the water content and may be calibrated to provide the actual moisture content or porosity of the in-situ material. The log will not differentiate between water bound by clays and shales from free water. The log may be used in open or cased holes above and below the water table.

Induction log The induction log is an electromagnetic induction tool for measuring the electrical conductivity of soil and rock in open or PVC cased holes above or below the water table. This log may be used for identification of lithology and inorganic contamination. The probe does not require contact with the formation. Because this log has a relatively large investigation radius, it is almost totally insensitive to borehole and construction effects. Because the electrical conductivity of the soil or rock is a function of the material, porosity, permeability, and pore fluids, a particular response may be caused by any of several possible subsurface conditions. Accordingly, without additional information it may not be possible to correlate a response with a definite subsurface material or condition.
**Resistivity log**  The resistivity log measures the apparent resistivity of rock and soil in the borehole. Because resistivity is the reciprocal of conductivity, the resistivity log responds to the same properties as the induction log. However, the resistivity log requires contact with the borehole wall, and can only be run in an uncased hole filled with water or drilling fluid. Variations in the electrode geometry give rise to different levels of resolution and radius of measured material adjacent to the borehole.

**Resistance log**  The resistance log, also called single-point resistance, measures the electrical resistance of the materials between a downhole electrode and a surface electrode. It can only be run in uncased holes with the downhole probe in the saturated zone. The primary uses are geologic correlation and identification of fractures. The log may be strongly affected by the conductivity of the borehole fluid.

**Spontaneous Potential log**  The spontaneous potential (SP) log measures the natural potential (in millivolts) developed between the borehole fluid and the surrounding materials. The SP log is subject to noise from the electrodes, hydrogeologic conditions, and the borehole fluids. The SP log must be run in an uncased hole under saturated conditions. The measurements are qualitative and may be used to characterize lithology, provide an indication of fluid flow, and provide information on the geochemical oxidation-reduction conditions.

Several additional logs are available. The temperature log, usually run in cased holes within the saturated zone, may be used to indicate fluid flow. A variety of devices to measure fluid flow within uncased boreholes within the saturated zone exist. A fluid conductivity log measures the specific conductance of the borehole fluids and may be used to detect some types of contamination.

### 5.2 Direct Push Technology

Direct push technology (DPT) consists of devices to push or drive steel rods into the ground and may deploy a variety of sensors and soil, soil vapor, and groundwater sampling devices. The
two basic direct push tools are the cone penetrometer and percussion probing. Advantages of DPTs over drilling include no drilling waste, safer working conditions, rapid data collection, and lower costs.

5.2.1 Percussion Probing

A geoprobe is a vehicle-mounted, hydraulically-powered, soil probing machine that utilizes static force and percussion to advance small diameter sampling tools into subsurface soils to collect soil core, soil gas, or groundwater samples, or measure electrical conductivity. Geoprobe® is also a registered trademark of Kejr Engineering, Inc., of Salina, Kansas (Kejr Engineering, 1995). The percussion probing tools may be pushed to depths of 70 feet or more through soil materials, but will not penetrate sound bedrock, hard subsurface materials such as cobbles, or coarse gravels. Because of the percussion aspect of the geoprobe, it is possible that the geoprobe could penetrate cemented sands which might stop the advancement of a CPT rod. As described in the preceding chapter, a soil electrical conductivity sensor has been developed for the geoprobe. The geoprobe is an excellent platform for the development of additional subsurface sensors such as a piezocone, an LIF system, and possibly tip stress and sleeve stress sensors to obtain additional subsurface information, and broaden the utility of this device. However, the percussion driving aspect of the probe will almost certainly cause some engineering challenges with the development of new sensors.

A major advantage of the geoprobe is its relatively small size. Being a truck mounted device, it is easily maneuverable wherever a pickup truck can go. The geoprobe can also acquire data relatively fast. At the SRS DOSB, about four or five pushes were made per day to depths ranging from about 25 to 60 feet, with 60 feet being more common. Disadvantages of the geoprobe include the lack of sensors available and the reluctance of some regulatory authorities to accept soil and groundwater samples from this technology. Geoprobe rods may need to be decontaminated by hand scrubbing in a water bath upon removal from the subsurface. The system is typically operated by two people, but may be operated by one person if necessary.
5.2.2 Cone Penetration Testing

Cone penetration rigs use hydraulic rams to push steel rods into subsurface soils using the weight of the truck, or ground anchors, as a counter force. Accordingly, CPT trucks are often large, 20 tons or more, resulting in reduced maneuverability on sites with natural or cultural obstructions, or rough terrain. The CPT will not penetrate sound bedrock, hard subsurface obstructions (such as cobbles), coarse gravel, very dense sands, and coarse gravel. The rods may be equipped with a variety of sensors, including tip stress, sleeve stress, a resistivity probe, pore pressure sensor, and any of several LIF systems. Additionally, several types of soil and groundwater sampling devices are available for CPT rigs. Most CPT rigs are equipped to grout the push rod hole and steam clean the push rods upon removal from the subsurface.

A major advantage of the CPT is its ability to classify the soil on the basis of the tip stress and sleeve stress (some classifications also use the dynamic pore pressure) according to several empirically developed soil behavior types. The CPT rigs at the SRS DOSB each pushed typically two 60 foot holes per day, with some holes stopping at about 25 to 30 feet due to cemented sands at that depth. At the Marshalltown FMGP site, the CPT generally pushed four or five holes per day to depths ranging from about 25 to 40 feet. These rigs generally require two or three persons to operate. A disadvantage is the large size causing reduced maneuverability on some sites.

5.3 Rotosonic Drilling

The rotosonic drilling tool used at the SRS is considered an innovative technology and is briefly discussed here because it has several desirable features for hazardous waste site characterization (Barrow, 1996). Rotosonic drilling (also called rotasonic, sonicor, vibratory, or resonant sonic drilling) utilizes a drill head that imparts high frequency, high force vibrations into a rotating steel pipe to advance the pipe into the subsurface for the collection of minimally disturbed soils cores. The drilling may proceed through almost any formation, from soils, to soils with cobbles or gravel, to rock. The method requires no mud, water, air, or other circulating medium for
penetration, is quite fast, and yields minimal drill cuttings in the process. The soil core barrel is advanced in stages, and is withdrawn in five or ten foot sections to retrieve and place the soil core into clear plastic bags. Virtually the entire core may be retrieved for further inspection.

While several types of soil disturbance have been observed (Stephan, 1995), soil cores retrieved by this method are generally very adequate for soil classification and sampling for contaminant analysis, but may show too much disturbance for laboratory hydraulic conductivity testing. The disturbance is due to the fluidization of some soils as a response to the vibratory action of the drilling operation. At the SRS DOSB, borings DOL-1 (65 feet deep) and DOL-2 (55 feet deep) were drilled with the rotosonic method with nearly continuous core retrieval (a two foot section was lost) with well screen and casing set in nine hours. Because drilling spoils are minimal, the operation is fast, and the entire soil or rock core may be retrieved and contained in clear plastic bags, this technology is very promising for future incorporation into environmental site investigations.

5.4 Application of Exploration Tools in ESC

The general sequence of application of these techniques in an ESC program begins with a geophysical survey to cover a large portion, or possible all, of the site to identify anomalous conditions and aid development of the initial site conceptual model. Several boreholes will be drilled and logged for calibration of the geophysics and subsequent DPT tools. Some environmental sampling may be done near the expected source area and to obtain background samples and samples from selected locations within the site. A percussion probing tool with soil electrical conductivity sensor may be utilized to further correlate with the geophysical results, and begin to further refine the geologic and contaminant models. The CPT will follow the electrical conductivity probe as the CPT gives the most definitive assessment of the soil conditions, and can be used to aid the interpretation of the electrical conductivity measurements to assess whether the electrical conductivity is responding to contaminants or geologic conditions. Finally, the DPT tools will be
utilized to begin sampling in accord with the expected transport behavior of the site contaminants as based on the developing geologic and contaminant models.

5.5 Summary Remarks

There are a wide variety of geophysical tools available, each with its own strengths and weaknesses. A proper choice of tools must involve an expert in the geophysics arena. One needs to seriously consider the objectives of the investigation. Out of this might come a list of potential tools, be they geophysical, direct push technologies, or more traditional methods. Then the site geologic and cultural conditions must be evaluated to determine the likelihood of success of an applied technology. For example, the conclusions in the Marshalltown FMGP ESC report (Bevolo, Kjartanson, and Wonder, 1996) indicate that the GPR data added little useful information to the characterization effort, the seismic reflection data gave no depth indications, and the electromagnetic offset logging did not appear to correlate with any observed features at the site. While the borehole geophysical logging gave confident data, it did not provide any new information, and the seismic refraction data provided confident bedrock data which agreed with the borings at the site. Many of the problems with interpretation of the geophysical data at the FMGP were associated with cultural features, on-site traffic, the rail yard south of the site, buildings, fences, powerlines, etc., and surface fill materials over the site.

The pre-ESC magnetic and GPR surveys at the SRS DOSB identified the trench boundaries and areas of disturbed soil. The ESC phase 1 GPR and EM data were used to more accurately define the trench boundaries and disturbed soil area. The ESC phase 1 EM data indicated an organic plume source in the trench area and trending to the west and south in accord with the groundwater flow. This information was useful for planning the phase 2 investigation. The borehole geophysics at DOL-1 and DOL-2 assisted with correlation of technologies and stratigraphy. The Technos association of elevated geoprobe conductivity readings with inorganic contaminants was well laid out in their December 1995 report (Technos, 1995). The suggestion was made during
phase 2 by Technos that a relation between high electrical conductivity and inorganic contamination may exist, and this was useful to the investigation.

While the phase 1 geophysical data may be useful to aid the development of the site conceptual model when integrated with the phase 2 data, one goal of the phase 1 geophysical work is to locate anomalous areas for further investigation during phase 2. Accordingly, sufficient time must be allowed between phase 1 and phase 2 for a complete assessment of the phase 1 data so that maximum use of this information can be made during the ESC phase 2 investigation. Because the geophysical data often does not stand on its own, but serves to point out anomalous conditions to be explored by more invasive methods, the phase 1 geophysical data is most useful to the phase 2 investigation if it is fully interpreted, reported, and integrated prior to the phase 2 investigation.

Post phase 1, pre phase 2 maps showing where the green clay was clearly seen in the geophysical data, and where it appeared to be missing or unclear, would have been useful prior to phase 2. Technos (1995) did provide a contour map of the intermediate aquitard (25' unit) as Figure 38 in their report (not shown here) which clearly demonstrated the continuity of the aquitard. While the metals contamination was more widespread south of the trench area than was indicated by the region of elevated electrical conductivities identified by Technos, the Technos geophysical data clearly identified an inorganic contaminant plume trending to the southwest of the DOSB study area.

The major benefit of the geophysical data is its widespread coverage of the site, and its ability to indicate anomalous conditions. In order for the proper data to be collected during phase 2 to interpret the geophysical data, maps or cross-sections showing anomalous conditions should be constructed prior to the phase 2 field work so that the phase 1 information can be incorporated into the sampling plan as appropriate. To obtain a better assessment and integration of the phase 1 data requires not only sufficient time between phase 1 and phase 2, but also requires a sufficient number of professionals with expertise in geology, geophysics, geochemistry, civil engineering, statistics, geohydrology, management, and so forth, to properly evaluate the data, develop the site conceptual model, and plan the phase 2 activities.
6. RECOMMENDED APPROACH TO ESC

In this chapter I provide a summary of the statistical methods described in Chapters 3 and 4 and recommend an approach to utilization of these methods in the ESC process. Basically, the DQO process should drive the field work, so that the site specific questions that need to be addressed are developed in order that sufficient data of the right quality is collected to provide the maximum amount of information at a reasonable cost to support defensible decision making. This goal is not new, but has always been an underlying force in any good statistical design for data collection. The DQOs formalize the process of designing the data collection work plan so that the pertinent questions can be answered with an appropriate level of confidence.

The level of confidence in an interpretation of spatial data may depend on the amount of data, and the spatial arrangement of the data. Because the spatial characteristics of each new problem are different, the amount of data required to achieve a given level of confidence will vary from site to site. Due to cost and time constraints and inherent geological variability, there will always be some uncertainty. A proper statistical analysis, however, can quantify that uncertainty. While some statistical methods incorporate costs directly, each method to determine placement of subsequent samples examined here indirectly attempts to minimize costs and maximize information by the proper placement of samples on the basis of the currently available data.

Statistical methods may be used as guidance tools in conjunction with geophysical and DPT tools, depending on the needs of the study. Statistical methods can be used to help insure that the final data collected will meet the criteria for the intended use of the data, i.e. risk assessment. Environmental sample location for definitive data may be based on screening type data and/or on the current understanding of the groundwater flow system and geology as they pertain to the movement and transport of subsurface contaminants. I do not propose that statistical tools for sample location selection be used exclusively, but rather that they be used in conjunction with other available information, and in particular, that they be used to insure that the data objectives set forth
at the inception of the project are met as well as can be expected within the limitations of allotted
time and money.

6.1. Summary of Adaptive Site Selection Methods

This section provides a brief summary of the adaptive site selection algorithms considered in
Chapter 3. The summaries include some practical considerations regarding their use in an ESC
program.

6.1.1 The David and Yoo Method

This method is applicable to any phenomenon for which an iso-contour plot can be
constructed. The algorithm first selects a subset of potential future sample locations that are farthest
removed from the currently available data, i.e. locations with maximum minimum distance to
previously sampled locations. The Hardy method, which is simply kriging with a linear variogram
equal to the spatial distance between data, is used for spatial prediction. A future sample location is
selected which, together with some earlier subset of the data, would have given a map close to the
current map. The result is a very nearly uniform distribution of sample locations within the specified
sample area. While there are instances where a concentration of data within an area is needed to
clarify anomalous conditions, define contaminant plumes, or to assess spatial correlation, a relatively
uniform distribution of data locations over the area of interest generally provides a good database for
the assessment of site conditions.

The method may be used to provide a good initial coverage of the site as a whole, but may
need to modified by restricting the search area to coincide with, for example, contaminated regions
or anomalous areas, so that data is not collected in areas which provide little or no useful
information. The method is computer intensive but is usable in the field as a viable aid to sample
location selection.
6.1.2 The Englund and Heravi Method

This method incorporates sampling and analytical costs, together with costs associated with remediation and failure to remediate, into an adaptive sample location selection method. The method is based on an assumption that a decision rule exists whereby a block, or remediation unit, is remediated if its estimated average contaminant concentration exceeds an action level, otherwise the block is not remediated. A remediation unit could be the minimum volume of TPAH contaminated soil that might be excavated for, say, treatment by thermal desorption. The method can be used whenever a block average contaminant concentration for a remediation unit can be defined. The method uses kriging to predict block averages and obtain prediction variances to assess the expected cost associated with failing to remediate when in fact we should, and the cost associated with remediating when the actual contaminant concentration is below the action level. A total cost including sampling, analysis, and remediation may be calculated. However, developing a proper cost function that adequately represents reality may be difficult, primarily due to uncertainties associated with future non-compliance and indirect costs associated with risk to human health and the environment.

The method places samples in areas where the expected loss is greatest. These generally turn out to be near the expected action level boundary of the contaminant plume, but also depend to a large extent on the current sample configuration, predicted block averages, and prediction variances. The method can be used to select future sample locations individually, or in groups of, say, five or ten at a time. The algorithm is computer intensive but is usable in the field as an aid to sample location selection.

6.1.3 The Johnson Method

This method uses a beta distribution to model the probability that the contaminant concentration at an unsampled location exceeds a threshold. A prior probability map, and associated beta distribution parameters, that shows the probability with which contamination is expected to
exceed an action level must be developed for each node of a grid covering the site. The prior probability map can be constructed on the basis of existing hard or soft data, such as phase 1 analytical and geophysical data. Heuristically based equations modeled after Bayesian updating of the beta distribution, and based on indicator kriging of hard data, are used to update the beta distribution parameters following each sampling event. The method presumes we can define probability levels for which we are confident deciding an area is "clean", "contaminated", or "state uncertain." The state uncertain criteria is defined according to the probability \( p_1 < P(Z(s) \geq T) < p_2 \), where \( Z(s) \) is a contaminant concentration at spatial location \( s \), and \( T \) is a threshold, such as an action level. The endpoints need to be specified; Johnson suggests using \( p_1 = 0.1 \) and \( p_2 = 0.9 \). A grid node is classified as clean if the probability \( P(Z(s) \geq T) \) is less than \( p_1 \), and is classified as contaminated if \( P(Z(s) \geq T) \) is greater than \( p_2 \). One of three criteria may be used to select a future sample location: 1) maximize the area declared clean, 2) maximize the area declared contaminated, or 3) minimize the area declared state uncertain. If enough data are collected, these three methods may eventually converge to the "correct" plume, however, they differ in the early stages of sample collection. Criteria 1 generally attempts to locate samples outside the expected plume and gradually work in, criteria 2 locates samples inside the expected plume and gradually works out, and criteria 3 locates samples near the expected plume boundary where uncertainty is high.

The Johnson sample location selection appears to be sensitive to the choice of prior probabilities reflecting expected contamination at grid nodes covering the site. This technique can be used to select a single future sample location, or a group of future sample locations. The method is computer intensive but is usable in the field as an aid to sample location selection.

6.1.4 Uncertainty Zone Sampling Method

I propose a quick and simple graphical sample location method that proceeds as follows: choose potential sample locations from within a state uncertain region as defined by Johnson, but as determined on the basis of a kriging or Stein map, which are separated from existing data locations
by about one fourth to one times the variogram range of influence. If insufficient data to estimate the variogram exists, then initial sampling at about one fourth to one third times the expected variogram range of influence from previously sampled locations will provide data from which the variogram may be estimated. Subsequent samples may be spaced at about one half to one times the variogram range of influence, if desired, to cover a larger area with fewer samples. The uncertainty zone sampling method provides a graphical conceptual model that by design suggests good candidate future sample locations while allowing the field worker to "hand" select sample locations in accord with the needs of the ESC program, and provides a graphical assessment of when sufficient data have been collected to characterize the contaminant plume. The method fits very well in the ESC program whereby people, and not computers, make the decisions.

An alternative, computer based sampling method is to select a future sample location from within the state uncertain region where the kriging prediction variance is greatest. To avoid excessive sampling in uncontaminated areas, this method may perform best if a high lower state uncertain probability level is used, e.g. \( p_1 = 0.2 \) to 0.3.

This graphical method provides the ESC site manager with a choice of potentially good candidate sample locations to choose from. In accordance with the dynamic work plan, the site manager may now incorporate other criteria or information, which were not part of the statistical algorithm to generate a sample location, into the decision where to place future samples. The graphical uncertainty zone sampling method can be done on a computer within about one minute or so whereby a map showing good candidate sites may be generated. Such a map is illustrated in Figure 6.1 which shows an indicator kriging map for PCE at the SRS with probabilities less than 0.1 or greater than 0.8 and the area within a 50 foot radius from previously sampled points as the shaded area. Unshaded areas in Figure 6.1 are good candidate sample locations. The area southeast of the trenches near the 0.5 contour clearly represents an area where additional sampling is needed.
Figure 6.1. Graphical uncertainty zone sampling method to aid on-site sample location selection. Unshaded areas are good candidate sample locations.
6.2 Stopping Rules

The following section lists several stopping rules with some practical considerations regarding their use. Pitfalls and additional considerations are provided.

6.2.1 Discrepancy Functions and Map Convergence

The map discrepancy function developed by David and Yoo may be used to indicate when sufficient information exists to adequately map a phenomenon. The discrepancy function is a measure of the volume (or average distance) between two maps. David and Yoo construct these maps using the Hardy method, however, the discrepancy function may also be applied to maps generated by one of the methods which incorporates spatial dependence. Indication that we have collected enough data to adequately characterize the phenomenon of interest is given by consistently low successive discrepancy function values. Simply producing periodic plots of the variable of interest, such as the FMGP lower cohesive unit maps shown in Figure 6.2 (Figure 3.15 repeated), may indicate map convergence just as readily as the discrepancy function.

It is possible for the discrepancy function to indicate map convergence even when there is insufficient data collected to adequately define the variable of interest. For example, suppose we sample one half of a contaminant plume very well. While we now may know a great deal about the sampled half of the plume, we may know nothing about the other half of the plume. For this scenario, a discrepancy function might indicate map convergence, yet we only have information regarding one half of the area of interest. Accordingly, we need to consider the spatial distribution of the data in relation to the information we need to obtain from the map and cannot rely on the discrepancy function alone to indicate when to stop sampling. This can be accomplished by overlying a plot of sample locations on a contour map to be sure, for example, that contaminant plume edges are bounded with sufficiently low concentration values — if not, then additional data may be needed.
Figure 6.2. Daily lower cohesive unit surface elevation (ft above msl) with data locations shown. Contour interval = 2 feet.
6.2.2 England and Heravi's Total Cost Function

Whether or not the method of England and Heravi is used to select sample locations, one can look at a plot of the total expected cost as a function of the sample number (e.g. see Figure 3.23). This plot may be erratic initially, but is expected to eventually level out and then begin increasing as the plume(s) area becomes well defined. Once the costs associated with uncertainty about the plume location and magnitude of contaminant concentrations level out, the cost of additional sampling and analysis will begin to cause the total cost function to increase. At this point there is cost incentive to cease sampling activities, or at least consider the cost/benefit of additional samples.

6.2.3 Graphical Uncertainty Zone Sampling

The graphical uncertainty zone sampling method provides a method to assess whether additional sampling is needed by noting whether the contaminant plume is bounded by a sufficiently low probability contour. One could choose to sample until a sufficiently low probability (such as 0.1) that contaminant concentrations exceed the threshold at unsampled locations is achieved. The method not only shows if and when such a probability level is achieved, but also shows where additional samples need to be placed to achieve a low probability contour surrounding the contaminant plume.

6.3 Recommendations for ESC

This section provides recommendations for the selection and use of statistical methods and characterization tools in the ESC process. The recommendations are listed separately for phase 1, phase 2, and post phase 2 ESC activities. While each site may have its own unique aspects, there are some general conclusions which can be drawn. Some practical considerations are given.
6.3.1 ESC Phase 1

6.3.1.1 Geophysics, Geology, and Hydrogeology  On the basis of the DQO determined needs, the advantages and limitations of the potential geophysical methods, and the known or anticipated site conditions, an expert in geophysical methods should be consulted to design appropriate geophysical surveys, if any, to be performed. The results of these surveys should be interpreted by experts and presented in reports to the ESC team in time for them to incorporate the information into the phase 2 work plan.

Consideration of existing information will determine the need for additional soil borings, monitoring wells or piezometers to be installed at the site. Sufficiently detailed well logs should exist to allow correlations for the direct push technologies. If sufficient logs don't exist, then installation of several deep borings at the site must be carried out. The number of borings needed may depend on the complexity of the site geology. At the FMGP in Marshalltown, the geology was relatively simple and several borings were sufficient for DPT correlations. At the SRS DOSB, the alluvial stratigraphy overlying the green clay was more difficult to correlate with the DPT data. Several additional deep borings would have been very helpful, either to clarify the correlations, or to confirm the ambiguity.

6.3.1.2 Contamination  The DQOs should drive the process. If background samples are needed, then it is best to spread them out as much as possible to better represent the natural variation in background conditions. This may be accomplished by defining the region upgradient from the site which is not expected to be affected by contaminants originating from the site. One could then choose a selection of random locations to take the background samples. Alternatively, rather than rely on chance to provide a good coverage of the background area, using a regular grid, or the David and Yoo methodology, will provide a more uniform coverage of the background sample region. If the background region has an irregular shape, then using the maximum minimum distance to previously selected sample locations, beginning with an initial random location, may provide the easiest way to obtain a relatively uniform sampling of the background region.
To provide a reasonable data set from which to develop the initial site conceptual model and begin planning the phase 2 sampling plan, several samples should be selected from the region(s) known or considered to be source areas. This should clearly establish whether contamination at the site exceeds ARARs and should provide a basis for development of the site specific contaminant list for the phase 2 work. The number of sample locations might be chosen during the DQO process. If there is any prior information which might indicate a range of spatial correlation, then this could be used to define a sample spacing to adequately cover the expected source area. Alternatively, selecting five to ten percent of the expected total number of phase 2 samples is a reasonable rule of thumb. If possible, these data should be tested with the same level of QA/QC, accuracy, and precision as will be used for the phase 2 data so that the phase 1 data may be incorporated into the definitive dataset right from the beginning. This will allow the adaptive sample selection rules to be based on hard data right from the start of the phase 2 work.

Additional samples should be selected from across the entire site in an effort to indicate the extent of the contamination. Again the same level of data quality as will be used for the phase 2 work is preferable to allow confident future use of the data during phase 2. About five to 20 percent of the expected total number of samples should be sufficient for this purpose. Either a sample grid or the David and Yoo method should be used to insure a good coverage of the site. The maximum minimum distance approach will work around both the source location data and irregular shaped site boundaries in a straightforward manner and is recommended for this purpose. The maximum minimum distance approach may be carried out graphically by plotting shaded circles around the existing data locations in such a way that only a small portion of the site remains uncovered; this uncovered portion of the site defines the candidate sample locations. While this procedure may be extended quite naturally to three dimensions, it may be best to treat three dimensional contamination using a series of layers defined on the basis of stratigraphy. This layered approach is reasonable under the assumption that the behavior of contaminant transport within a layer may be relatively
consistent, but may vary across layers, particularly for layers which are expected to confine either the contaminant or the flow of groundwater through the subsurface.

6.3.2 ESC Phase 2

6.3.2.1 Direct Push Technologies DQOs will drive the process and determine what type of data and tools are necessary to carry out the investigation. Direct push methods have become standard ESC tools for characterization of the geology, as well as for soil and groundwater sampling for contaminant data. If a CPT is to be used to characterize the soil stratigraphy, then either the CPT output, or the ESC data analysis program, should be able to develop a soil classification. A pore pressure sensor may be used to perform pore pressure dissipation tests to estimate hydraulic conductivity in fine grained soils. Geoprobe electrical conductivity probe may be used to characterize soil stratigraphy. If one or more key stratigraphic layers are to be mapped, then the David and Yoo algorithm may be used to deploy both the geoprobe and/or CPT. The David and Yoo algorithm may be applied to the elevation, depth, or thickness of a soil unit. While sample locations may be chosen to explore anomalies detected by geophysical methods, or simply to resolve ambiguous or conflicting data in critical regions, application of the David and Yoo algorithm is still valid and recommended as a means of providing a more uniform sample distribution.

In the early stage of the phase 2 work, the direct push technologies should be pushed adjacent to several logged boreholes for site specific calibration. This serves a dual purpose of providing several closely spaced data clusters that may be used to explore the nature of the variogram for small lag distances. Additionally, it has been traditional for the Ames Laboratory ESC program to create several CPT/geoprobe data clusters throughout the site for direct comparison of these methods. As this practice provides additional data for development of the variogram, and the variogram is needed for a proper kriging analysis, it is recommended that such data clusters be developed during the first few days, to enhance subsequent data analysis.
6.3.2.2 Contamination  While knowledge of the contaminant source area, the direction of groundwater flow, soil type, and the transport characteristics of the contaminant might suggest potential sample locations for environmental samples, any of the preceding adaptive sample location selection methods may be used to help insure that the objectives of the site characterization are met. How effectively these tools can be applied depends largely on the level of expertise of the person applying them, and also on how well the objectives of the data collection program and remedial goals are understood. For example, consider the Marshalltown FMGP PAH contamination. While the action level was 500 mg/kg TPAH in soil, the actual clean up level might be considerably lower than this. Accordingly, this must be taken into consideration when setting the action level threshold that is used in the Englund and Heravi, Johnson, and uncertainty zone sampling methods.

For contaminant sampling, I recommend that the graphical uncertainty zone sampling, Englund and Heravi, or the Johnson method be used to choose additional sample locations. If Johnson's method is used, care must be taken to insure that a poor choice of prior does not negatively impact the sampling design. Choice of which method to use might depend on the available software, and whether or not the ESC team or stakeholders want to build expected costs into the sample selection program. Englund and Heravi's method requires a block kriging program, numerical integration, and a cost structure. The uncertainty zone sampling method requires the ability to produce a probability map (indicator or ordinary kriging, or Stein's method), and a method to overlay a contour map onto a site plan with the data locations shown as user defined circles of arbitrary size. The EarthVision® software currently used by the Ames Laboratory ESC team has the capability to perform the kriging calculations and map overlays to implement the uncertainty zone sampling method in the field. Software is available to perform the calculations required for Johnson's method (Johnson, 1996).

Proper use of the spatial statistical methods (for all but the David and Yoo method) requires development of a spatial correlation model. This requires adequate data collected at close enough spacings to allow the estimation of a variogram or covariance function. Until such data is available,
one must simply choose a certain model form and model parameters on the basis of whatever data is available, knowledge of the process under study, and knowledge of the statistical method to be applied. Accordingly, some closely spaced data collected early in the program may be of great use in this regard. The phase 1 data may also be utilized to help determine an appropriate model. The David and Yoo algorithm can be implemented without a spatial dependence model, and may be modified to provide very adequate data by restricting the search area for potential next sample locations to certain regions, such as within the region believed to contain the plume.

Probability maps showing the probability of a contaminant concentration exceeding a threshold may be of great use in both deciding where future samples should be collected, and whether enough data have been collected to adequately characterize the site. Such maps should always consider the spatial distribution of the data locations (by simply plotting them on the map) so that it is clear where the map has adequate data support, and where data is lacking, so that unwarranted extrapolation beyond the data mass may be avoided. Probability maps may be developed from indicator or ordinary kriging, direct application of Johnson's method, or by Stein's method. Stein's method may be used during the phase 2 investigation, but due to potential processing time, it may be desirable to run the program overnight. Note that distributional assumptions must be checked to verify the appropriateness of ordinary kriging and Stein's method for this application.

Proper utilization of any of these methods may require a substantial portion of a person's time during the site investigation. This time will be spent building models of spatial dependence, checking distributional assumptions, and modeling individual contaminants or spatial phenomenon. With each additional contaminant, and each geologic layer that is considered, the problem becomes more and more time consuming. Data may come in faster than it can be properly interpreted and analyzed. Simplifications, such as those mentioned in the preceding SRS analysis in Chapter 4 might be considered to better manage the data while in the field.
Any of the stopping rules discussed in section 6.2 may be used to indicate when sufficient data have been collected. These methods generally indicate when additional data begins to add little to the understanding of the site conditions. Appropriate use of the site selection and probabilistic assessment maps can also be used to indicate directly when sufficient sampling has occurred, e.g. we may choose to sample until the plume is enclosed within a 0.1 probability contour indicating a low probability of exceeding an action level.

6.3.3 Post Phase 2 Recommendations

Ordinary kriging, Bayesian updating, or Stein's method may be used to model the data. Each of these methods requires a spatial dependence model. Block kriging may be used to assess average concentrations, along with prediction variances, for remediation units or exposure units. Stein's method is particularly appropriate when there are non-detect data (which may always be the case), however, this method is computer intensive and may require up to several hours or more, depending on the computer speed, number of grid nodes, and the number of observations, for the development of a contaminant model. In addition to estimating actual concentrations, Stein's method can be used to directly estimate the probability that the contaminant exceeds user defined thresholds, and as many thresholds as desired may be incorporated into one simulation.

If spatial correlations can be discovered in the data, then predictive models which incorporate this spatial dependence can be utilized for prediction and to assess levels of certainty via the prediction variance and/or development of maps showing various probabilities of interest. These spatial correlations may be used to assess a range of conditions which might be expected to be encountered so that effective and robust remedial designs can be made. Other standard characterization methods, such as cross sections, should be used as much as possible to extract the maximum amount of information from the data.
6.4 Recommendations for Further Work

This work does not solve all of the site characterization problems, but it does provide some useful techniques which advance the current state of practice in Expedited Site Characterization. While nondetect data may always be present in environmental data, the method of Stein will remain useful to modeling such data. Development of a method to model data with multiple detection limits would provide a valuable enhancement to interpretation of multiple detection limit data.

I have provided some geostatistical methods to model both geologic data and contaminant data separately, and we have used our knowledge of the geology and its effect on contaminant transport together with the developing contaminant model to suggest future sample locations for contaminant data. However, development of one model incorporating both geologic and contaminant data, together with contaminant migration behavior, to suggest future sample locations would be a valuable research area.

I have made several comparisons of the Johnson, Englund and Heravi, David and Yoo, and uncertainty zone sampling methods. Further comparisons to explore the effects of a poor prior with Johnson's method and alternative cost structures with the Englund and Heravi method would be useful. While these methods appear to work well with a moderate overestimation of the range of influence, further investigation of the sensitivity of these sample methods to the variogram or covariance model parameters would be useful. Further testing of these sampling methods with both computer generated data, and especially actual field data, would be valuable.

6.5 Concluding Remarks

This thesis has provided a comparison of three adaptive sample location selection algorithms, and the development of a graphical uncertainty zone sampling method, and has demonstrated how each of these methods can be used to enhance an Expedited Site Characterization program. Stopping rules, based on the David and Yoo discrepancy function and the Englund and Heravi total expected cost method, and the graphical uncertainty zone sampling
method, to help decide when sufficient data have been collected are given, and additional considerations beyond the straight application of these methods are discussed. Comparison with the time series-like daily maps already used in the ESC process show that the daily map surfaces may indicate convergence just as well as the map discrepancy function. The uncertainty zone sampling method provides a graphical site conceptual model of the contamination, as well as providing good candidate sample locations, and suggesting when enough data have been collected to define the location and extent of a contaminant plume.

Data collected in an ESC program is generally not compatible with the hypothesis testing framework described in the USDOE DQO process. Uncertainty will always exist in environmental waste site characterization and effective management of this uncertainty is vital to proper decision making. Statistical spatial analysis tools are given in this work which enable the quantification of uncertainty to aid the decision making process.

While the statistical methods described here are not a panacea, they can be used to effectively enhance the standard site characterization methodologies, and may provide several measures of uncertainty. These uncertainty measures include prediction variances for spatial statistical models applied to either geologic data or contaminant data, and probabilistic measures applied to contaminant spatial distributions. The preceding analysis of data collected during the ESC program at the SRS DOSB indicate that a slightly better level of environmental characterization could have been achieved with about 80 percent of the sampling that was actually carried out. The analysis of the FMGP lower cohesive unit (LCU) surface elevation indicate that this surface could have been defined adequately with about 75% of the data that was actually collected.

An assessment of geophysical methods and direct push technologies is given with emphasis on how these methods may best be utilized in an ESC project. The DQOs, together with information regarding the site history, the nature of the current problem, and expected or known site conditions, will determine which site characterization methods are utilized. While each site has its own unique
aspects, some general conclusions may be drawn which are potentially useful to many site investigations.

Spatial statistical methods may provide input into the decision making process in several ways. Predictive models, together with a prediction variance, may be developed. These can be used to identify optimal locations for the placement of monitoring wells. Probabilistic models may be developed which indicate where contamination is likely to exist, and where it likely absent or uncertain. These may be used by the appropriate decision authority to decide on remedial actions, or suggest that uncertainty is unacceptably large and additional data is needed. The method developed by Stein for prediction with non-detect data is shown to be potentially useful within the ESC process.

The focus of this work rests with application of statistical tools to enhance the USDOE ESC process. However, a review of the ESC process, and several similar site characterization processes is also given. Other processes considered are the M³ process, FAST, SACM, and SAFER. While the M³ and FAST processes are site characterization methods, SACM and SAFER have a broader focus by considering the entire environmental restoration process. The ESC process is shown to be compatible with the SACM and SAFER processes.
REFERENCES


References


Golchin, J., 1996, MGP and Radioactive Sites Program, Iowa Department of Natural Resources, Environmental Protection Division, Des Moines, Iowa, Personal communication, September 1996.


USDOE, 1996, Managing Human Health and Ecological Risk Assessments Under CERCLA and RCRA, A DOE Workshop, developed by: The Office of Environmental Policy and Assistance, RCRA/CERCLA Division (EH-413), Gaithersburg, Maryland, March 5-7, 1996.
