Naval Facilities Engineering Command

ENGINEERING SERVICE CENTER
Port Hueneme, California 93043-4370

USER'S GUIDE
UG-2059-ENV

GUIDANCE FOR ENVIRONMENTAL BACKGROUND ANALYSIS VOLUME III: GROUNDWATER

Prepared by
Battelle Memorial Institute
Earth Tech, Inc.
NewFields, Inc.

April 2004

Approved for public release; distribution is unlimited.

Printed on recycled paper

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0811*

1. REPORT DATE ***(DD-MM-YYYY)*** April 2004	**2. REPORT TYPE** Final	**3. DATES COVERED** ***(From – To)*** April 2003 – April 2004

4. TITLE AND SUBTITLE **GUIDANCE FOR ENVIRONMENTAL BACKGROUND ANALYSIS – Volume III: Groundwater**	**5a. CONTRACT NUMBER** N47408-01-D-8207, Task Order No. 0059
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Battelle, Columbus, Ohio Earth Tech, Inc., Honolulu, HI NewFields, Inc., Atlanta, GA	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESSES Battelle, Columbus, Ohio Earth Tech, Inc., Honolulu, HI NewFields, Inc., Atlanta, GA	**8. PERFORMING ORGANIZATION REPORT NUMBER** **UG-2059-ENV**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Facilities Engineering Service Center 1100 23rd Avenue Port Hueneme, CA 93043-4370	**10. SPONSOR/MONITORS ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This guidance document provides instructions for characterizing background conditions at groundwater sites where past uses of the property have resulted in actual or suspected chemical releases. Background analysis is necessary to identify background chemicals, those derived from natural or anthropogenic sources not related to activities conducted at the site, and to estimate the chemical concentration ranges that represent background conditions at the site.

Step-by-step instructions for background analysis techniques or methods are presented for chemicals in groundwater. For many sites and target chemicals, background concentration ranges can be demonstrated by a “weight of evidence” process which consists of three steps: (1) Identify all actual or suspected on-site and upgradient contaminant sources, and evaluate the geologic and hydrogeologic characteristics of the site and background sampling areas, (2) Evaluate the geochemical characteristics of the groundwater and aquifer matrix, and (3) Construct post-plots and graphs to evaluate concentrations of the target COPCs with respect to space and time (Sections 2 and 3). After confirming that the background dataset is suitable for comparison to the site dataset, the project team can compare the maximum COPC concentrations detected at the investigation site to the estimated background range, and decide whether statistical comparative testing and the Comparative Method (Section 4) is required to determine whether differences between the site dataset and the background dataset are statistically significant. A case study (Section 5) illustrates the application of each method.

15. SUBJECT TERMS

Chemical concentration ranges, background, ambient concentrations, groundwater sites, cleanup efforts, environmental investigations, remedial investigation, chemical of potential concern

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			
U	U	U	U	203	**19b. TELEPHONE NUMBER** ***(include area code)***

The vendors and products, including the equipment and other materials identified in this report, are primarily for information purposes only. Although the authors may have used some of these vendors and products in the past, mention in this report does not constitute the authors' recommendation for using these vendors or products.

PREFACE

This guidance document supports and implements the September 2000 Navy background policy document, *Navy Interim Final Policy on the Use of Background Chemicals*. It is part of a series devoted to background analysis that provides instructions for characterizing background conditions at sites where past property uses have resulted in actual or suspected chemical releases. Each volume in this series is devoted to a particular medium. This volume focuses on analytical methods and procedures that can be used to identify background chemicals in the groundwater medium (whether from anthropogenic or natural sources), estimate the chemical concentration ranges that represent site-specific background conditions, and perform statistical comparative testing to determine whether differences between site data and background data are statistically significant. The background analysis techniques presented in this volume focus primarily on evaluation of groundwater metal concentration data. However, these techniques can also be applied to organic chemicals.

For the sake of completeness, each volume in the series opens with a summary of the September 2000 Navy background policy, and discussions of both state and federal regulatory requirements and guidance. Each volume then describes data review and assessment procedures, explains methods for background analysis and statistical comparative testing, and presents medium-specific case studies that illustrate application of the methods. Because each volume in this series is intended to serve as a stand-alone document, some identical or similar discussions occur across the volumes. The preface of each volume identifies the sections that are unique to that volume and the medium of concern.

The background analysis methods presented in Volumes I (soil) and II (sediment) rely on analysis of the spatial characteristics of the chemical data (Exploratory Data Analysis), the geochemical processes that result in strong correlations between certain chemicals in soil and sediment (the Geochemical Method), as well as statistical comparisons of site and background datasets (the Comparative Method). Exploratory data analysis of the spatial characteristics of a dataset (e.g., upgradient vs. downgradient data) is essential for groundwater background analysis. However, the Geochemical Method is not recommended for groundwater background analysis because geochemical relationships between COPCs and other chemical constituents in groundwater are extremely complex, and depend on a wide range of hydrogeological and geochemical parameters that can change quickly over time and space within the groundwater environment (e.g., hydraulic conductivity, pH, Eh, and many other characteristics of both the groundwater and aquifer matrix materials).

This volume presents a three-step “weight of evidence” approach to determine whether a prospective background dataset represents local groundwater background conditions, and therefore, is appropriate for comparison to site data. After confirming the appropriateness of the background dataset, the project team can compare the maximum COPC concentrations detected at the investigation site to the estimated background range in order to decide whether statistical comparative testing (the Comparative Method) is required. This volume both discusses site-wide comparisons to background datasets, and presents procedures for statistical comparison of individual measurements to background datasets. Individual comparisons are not recommended for soil and sediment background analyses, but are sometimes required during groundwater detection monitoring phases. The results of the Comparative Method are used to determine whether observed differences between site and the background measurements are statistically significant.

ACKNOWLEDGMENTS

This document was prepared by Battelle Memorial Institute, Earth Tech, Inc., and NewFields, Inc. for the Naval Facilities Engineering Command (NAVFAC). Numerous people from all three organizations contributed to this document including Tom Hanneman (Earth Tech) and Dr. Shahrokh Rouhani (NewFields) as primary authors, and Chris Zimmerman (Battelle) as the project manager.

NAVFAC's Risk Assessment Workgroup (RAW) and the RAW Background Subgroup made significant contributions to this document. In particular, NAVFAC's Dave Barclift, Chris Leadon, and Ruth Owens provided technical input, review, and direction throughout the project. Also, this guidance has gone through a peer review by independent experts in the field of environmental statistics.

Furthermore, this document incorporates significant information from two previous Naval Facilities Engineering Command (NAVFAC) documents*: Procedural Guidance for Statistically Analyzing Environmental Background Data* (DON 1998) and *Handbook for Statistical Analysis of Environmental Background Data* (DON, 1999a). Thus, authors and contributors to these two documents are noted. Although this guide is not intended to supersede these previous documents, it does serve as a compendium of the earlier guidance, provides additional instruction, and contains up-to-date information.

EXECUTIVE SUMMARY

This guidance document provides instructions for characterizing groundwater background conditions and comparing datasets representing groundwater impacted by an actual or potential chemical release to appropriate background datasets. Groundwater investigations and monitoring programs often require comparison of chemicals and chemical concentrations detected in groundwater samples from monitoring wells at an investigation site to corresponding data representing background conditions. Background analysis is therefore necessary to identify groundwater background chemicals—those derived from natural or anthropogenic sources not related to activities conducted at the site—and to estimate the chemical concentration ranges that represent site-specific groundwater background conditions.

According to the September 2000 Navy background policy document *Navy Interim Final Policy on the Use of Background Chemicals*, cleanup efforts at Navy sites should address only those risks associated with chemical concentrations that are elevated as a result of a site-related release. Cleanup efforts therefore must address only chemicals that have been released at the site—not background chemicals. Unacceptable risks may be associated with chemical concentrations within the background range. Although these risks are outside the scope of the Navy's Environmental Restoration Program, Navy policy requires restoration program personnel to convey information regarding all identified risks to stakeholders. The Navy background policy was developed to ensure compliance with federal and state laws and regulations, and is consistent with United States Environmental Protection Agency (U.S. EPA) technical guidance.

The background analysis techniques presented in this document are based on well-established statistical methods and geochemical relationships. The data analysis and statistical testing methods closely follow U.S. EPA's *Guidance for the Data Quality Objectives Process* and *Guidance for Data Quality Assessment: Practical Methods for Data Analysis.*

Groundwater background analysis is an integral component of most Navy groundwater site investigations and monitoring programs. Existing data that may be relevant to background should be reviewed and assessed during the initial phase of an investigation. The operational history, as well as geologic, hydrogeologic, and geochemical characteristics of the site and upgradient properties, should be evaluated to: (1) develop a list of chemicals of potential concern (COPCs), and (2) identify conditions that affect the spatial and temporal distributions of COPCs and background chemicals in groundwater. If additional or supplementary data are required, a sampling and analysis plan should be developed. The sampling and analysis program should be carefully designed and implemented to ensure that all data needed to evaluate groundwater background conditions are collected. After the additional data have been collected and incorporated into the dataset, the reported chemical concentrations should be compared to appropriate screening criteria to determine which chemicals should be carried forward for background analysis.

After identifying the COPCs that require background analysis and assembling all information and data needed for the analysis, a three-step "weight of evidence" process should be used to compile the evidence required to demonstrate that a prospective background dataset represents local groundwater background conditions, and is appropriate for comparison to site data: (1) Identify all actual or suspected on-site and upgradient contaminant sources, and evaluate the geologic and hydrogeologic characteristics of the site and background sampling areas. (2) Evaluate the geochemical characteristics of the groundwater and aquifer matrix. (3) Construct post-plots and graphs to evaluate concentrations of the target COPCs with respect to space and time. If no suitable locations for background monitoring wells (i.e., unimpacted regions of the aquifer) can be identified at or near the investigation site, the project team should expand the search for background data to other locations in the region, and repeat the evaluation process to identify a site with similar geologic, hydrogeologic, geochemical, and anthropogenic characteristics. Then, with stakeholder concurrence, the similar location can provide background groundwater data appropriate for comparison to the investigation site groundwater data.

The project team should compile the geologic, hydrogeologic, geochemical, and COPC concentration evidence, and present a summary of the evidence and background analysis conclusions for stakeholder review. If the choices of background wells and estimated background ranges are not acceptable to stakeholders, then the deficiencies should be identified, additional data should be collected if necessary, and the background analysis steps should be repeated as required.

After confirming that the background dataset represents local background conditions and is suitable for comparison to the site dataset, the maximum detected COPC concentrations should be screened against the background concentration ranges as a preliminary step. If the maximum concentrations are within the estimated background ranges, then comparative statistical testing is not required. If the maximum concentrations exceed the upper bound of the background range, but the exceedances are not high enough to be obviously inconsistent with the background population (based on site knowledge and professional judgment), then the Comparative Method should be implemented.

The Comparative Method uses various statistical testing approaches to determine whether differences between the site dataset and the background dataset are statistically significant. In general, the Comparative Method can be divided into two broad categories: (a) site-wide comparison for characterization and assessment purposes, and (b) individual comparison for detection monitoring purposes. Statistical procedures appropriate for characterization and assessment purposes are fundamentally different from those suitable for detection monitoring. If the comparative test results indicate that the detected differences are statistically significant (i.e., the site data or the individual monitoring data represent values with significantly higher COPC concentrations than the background population), then the data indicate that the site most likely has been impacted by a chemical release. Conversely, if the differences are not statistically significant, then the data indicate that COPC concentrations at the site are consistent with the background population.

CONTENTS

FIGURES

TABLES

BOXES

ACRONYMS AND ABBREVIATIONS

ACL	alternate concentration limit
AF	absolute frequency
amsl	above mean sea level
ANOVA	analysis of variance
ARAR	applicable or relevant and appropriate requirement
ASTM	American Society for Testing and Materials
bgs	below ground surface
Cal/EPA	California Environmental Protection Agency
CEC	cation exchange capacity
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
CF	cumulative frequency
CFR	Code of Federal Regulations
CNO	Chief of Naval Operations
COPC	chemical of potential concern
CSM	conceptual site model
CUSUM	cumulative sum
CV	coefficient of variation
CWA	Clean Water Act
DDT	dichlorodiphenyltrichloroethane
DEFT	Decision Error Feasibility Trials (software)
DL	detection limit
DNAPL	dense, nonaqueous-phase liquid
DoD	United States Department of Defense
DON	Department of the Navy
DOQ	digital orthophoto quadrangle
DQA	data quality assessment
DQO	data quality objective
DTSC	Department of Toxic Substances Control (of Cal/EPA)
EFA	Engineering Field Activity
EFD	Engineering Field Division
Eh	redox potential
FDEP	Florida Department of Environmental Protection
FDER	Florida Department of Environmental Regulation
GIS	Geographic Information System
GWMP	Groundwater Monitoring Plan
H_a	alternative hypothesis
H_o	null hypothesis
HDOH	Hawaii Department of Health

ID	identification
IR	Installation Restoration
IR CDQM	Installation Restoration Chemical Data Quality Manual
J	estimated concentration
K_{ow}	octanol-water partition coefficient
LNAPL	light, nonaqueous-phase liquid
MCL	maximum contaminant level
MDEQ	Michigan Department of Environmental Quality
meq/100g	milliequivalent per 100 grams
MERA	Michigan Environmental Response Act
mS/cm	millisiemens per centimeter
msl	mean sea level
mV	millivolt
NAPP	National Aerial Photography Program
NAVFAC	Naval Facilities Engineering Command
NCDC	National Climate Data Center
NCP	National Oil and Hazardous Substances Pollution Contingency Plan
ND	nondetect
NFA	no further action
NFESC	Naval Facilities Engineering Service Center
NGWA	National Groundwater Association
NOAA	National Oceanic and Atmospheric Administration
NQ	no qualifier/not qualified
NRC	Nuclear Regulatory Commission
NRCS	Natural Resources Conservation Service
NWS	National Weather Service
OEPA	Ohio Environmental Protection Agency
OERR	Office of Emergency and Remedial Response (of U.S. EPA)
OSWER	Office of Solid Waste and Emergency Response (of U.S. EPA)
PAH	polycyclic aromatic hydrocarbon
PCB	polychlorinated biphenyl
pcf	pound per cubic foot
ppb	part(s) per billion
ppm	part(s) per million
PVC	polyvinyl chloride
PZC	point of zero charge
QA	quality assurance
QAPP	quality assurance project plan
QC	quality control
RAGS	Risk Assessment Guidance for Superfund
RCRA	Resource Conservation and Recovery Act

RL	reporting limit
RPM	Remedial Project Manager
SAP	Sampling and Analysis Plan
SARA	Superfund Amendments and Reauthorization Act
SDWA	Safe Drinking Water Act
SOP	standard operating procedure
SWDIV	Southwest Division Naval Facilities Engineering Command
SWMU	solid waste management unit
TDS	total dissolved solids
TOC	total organic carbon
TSCA	Toxic Substances Control Act
U	nondetect
UJ	nondetect estimated
USC	United States Code
USDA	United States Department of Agriculture
U.S. EPA	United States Environmental Protection Agency
USGS	United States Geological Survey
UST	underground storage tank
VDEQ	Virginia Department of Environmental Quality
VSP	Visual Sample Plan
WRD	Water Resources Division (of USGS)
WRS	Wilcoxon Rank Sum test

GLOSSARY

α	Alpha is the tolerated probability of a Type I error in a hypothesis test.
β	Beta is the tolerated probability of a Type II error in a hypothesis test.
ε	In the Quantile test, epsilon is the proportion of a site in which chemicals are present at concentrations greater than background levels.
Alternative Hypothesis, H_a	The hypothesis that is accepted if the null hypothesis is rejected.
Ambient Chemical Concentrations	Total concentrations of both naturally occurring chemicals and anthropogenic chemicals not related to specific point sources or site releases.
Anion	A negatively charged chemical species. Sulfate (SO_4^{2-}) and chloride (Cl^-) are examples of anions.
Anthropogenic Background	Chemicals present in the environment due to human activities that are not related to specific point sources or site releases.
Aquifer	A water-bearing rock or sediment formation capable of yielding usable quantities of water, typically composed of unconsolidated materials such as sand and gravel, or consolidated bedrock such as limestone or fractured basalt.
Aquitard or Confining Layer	A relatively impermeable formation that restricts the vertical movement of water. The low permeability of an aquitard prevents the production of significant quantities of usable water from wells screened in the formation.
Background Area	See *Reference Area.*
Background Chemicals	Chemicals derived from natural or anthropogenic sources not associated with site-related chemical releases (i.e., sources not related to activities or operations conducted at the site).
Background Groundwater Conditions	The chemical characteristics of groundwater that has not been impacted by either a site-related or a nonsite-related chemical release. Background conditions may be affected by natural and anthropogenic chemical sources.
Basalt	A dark, fine-grained, extrusive (volcanic) igneous rock with a low silica content (40 to 50%), but rich in iron, magnesium, and calcium. Generally occurs in lava flows, but also as dikes. Basalt makes up most of the ocean floor and is the most abundant volcanic rock in the Earth's crust.
Baseline Groundwater Conditions	The chemical characteristics of groundwater not impacted by a site-related release. Groundwater can transport chemicals released at off-site areas to the site; therefore, baseline conditions can include the chemical characteristics of groundwater impacted by a source not related to the site. If groundwater is not impacted by off-site sources, background and baseline conditions will be the same.

Box and Whisker Plot	A graphic way of summarizing a set of data measured on an interval scale. Often used in exploratory data analysis, a box and whisker plot is a type of graph that shows the shape of the distribution, its central value, and variability. The picture produced consists of the most extreme values in the dataset (maximum and minimum values), the lower and upper quartiles, the median, and the mean.
Cation	A positively charged chemical species. Calcium (Ca^{2+}) and magnesium (Mg^{2+}) are examples of cations.
Censored Dataset	A dataset that contains one or more nondetects.
Clay Minerals	Finely crystalline, hydrous silicates formed from weathering of such silicate minerals as feldspar, pyroxene, and amphibole. Most common clay minerals belong to kaolinite, montmorillonite, and illite groups.
Comparative Method	Application of statistical tests for identifying COPCs. In this method, results collected from a nearby uncontaminated or “background” area are statistically compared to pooled or individual results from samples collected at a site of suspected contamination.
Chemical of Potential Concern (COPC)	A chemical that was released during past or current operations at a site, and which may be present at concentrations that pose potentially unacceptable risks to human health or the environment.
Darcy’s Law	A relationship that describes the rate of flow (Q) through a porous medium as directly proportional to the difference between hydraulic head at two points along the flowpath (dh) and the cross sectional area of the medium (A), and inversely proportional to the distance between the points (dl). Darcy’s law quantifies the rate of flow by adding a proportionality constant, the hydraulic conductivity (K), which is a property of the permeability of the porous medium, and the fluid density and viscosity. Darcy’s law can be expressed as: $$Q = \frac{-KA(dh)}{(dl)}$$
Data Quality Assessment (DQA) Process	A scientific and statistical data evaluation to determine if environmental investigation data are of the right type, location, quality, and quantity to support their intended use.
Data Quality Objective (DQO) Process	A series of planning steps based on the scientific method that are designed to ensure that the type, location, quantity, and quality of environmental data used in decision-making are appropriate for the intended application.
Degree of Freedom (f)	Describes the number of values in the final calculation of a statistic that are free to vary.
Dense, Nonaqueous-Phase Liquid (DNAPL)	An insoluble or low-solubility chemical with a density greater that that of water. DNAPLs sink to the bottom of an aquifer. Examples are chlorinated solvents such as trichloroethylene (TCE) and tetrachloroethylene (PCE or PERC).
Detection Limit	The minimum concentration of an analyte that can be measured within a given matrix and reported with a 99% confidence that the analyte concentration is greater than zero.
Distribution	The frequency (either relative or absolute) with which measurements in a dataset fall within specified classes. A graphic display of a distribution is referred to as a *histogram*.

Exploratory Data Analysis	A statistical and graphic procedure for examining data in order to describe the main distributional features of measured data.
Facility	See *Installation.*
Geostatistics	A class of statistical techniques for the analysis of spatially correlated data. In these analyses, variograms or related techniques are used to quantify and model the spatial correlation structure. Also includes various estimation techniques, such as kriging, that use spatial correlation models.
Histogram	A method of graphically displaying the characteristics of a distribution of items in a given population or sample. In a histogram, each measure is usually represented by a single block placed over the midpoint of the class interval into which the measure falls.
Hydraulic Conductivity (K)	In hydrogeology, the rate of flow of water through a unit cross section of an aquifer under a unit hydraulic head. Hydraulic conductivity is a measure of the permeability of a porous medium to water.
Hydraulic Gradient	In hydrogeology, the direction and magnitude of the slope of the water-table surface (unconfined aquifer) or potentiometric surface (confined aquifer). The magnitude of the hydraulic gradient is expressed as a ratio of the vertical change in hydraulic head per unit of horizontal distance.
Hydraulic Head	The height of a column of water, such as the elevation of the water level in a well. Hydraulic head is a measure of the potential energy associated with the elevation or pressure of a mass of water.
Hypothesis	An assumption about a property or characteristic of a set of data under study. The goal of statistical inference is to decide which of two complementary hypotheses is likely to be true. The *null hypothesis* (H_o) describes what is assumed to be the true state of nature; the *alternative hypothesis* (H_a) describes the complementary situation.
Inflection Point	A point on a curve where the direction of the curvature changes. For probability plots in background analysis, an inflection point that marks a distinct increase in slope typically represents the cutoff between two different populations present in the dataset. Examples of different populations include background versus impacted populations. The cutoff between these two populations represents the upper bound of the background concentration range.
Installation	The extent of a Navy property at which one or more activities have been or are being conducted. An installation may contain a number of sites, as well as parts or all of the investigated sediment basin. Also referred to as a *facility.*
Interquartile Range (IQR)	A measure of the spread of or dispersion within a dataset. The IQR is the difference between the 25th and 75th percentiles of the measured values of the sample. IQR is not affected by outliers.
Ion	A charged chemical species containing either more electrons than protons (anion) or fewer electrons than protons (cation).
Kriging	In geostatistics, a weighted-moving-average interpolation method in which the set of weights assigned to samples minimizes the estimation variance.

Light, Nonaqueous-Phase Liquid (LNAPL)	An insoluble or low-solubility chemical with a density less that that of water. LNAPLs float on the groundwater table. Examples are petroleum products such as gasoline.
Lognormal Distribution	A family of positive-valued, skewed distributions commonly used in environmental work.
Mean	A measure of the central tendency of a distribution, which is computed by averaging a dataset (totaling the various individual results and dividing by the number of results involved).
Median	A measure of the central tendency of a distribution, which is obtained by ranking the individual results in a dataset from smallest to largest and selecting the middle value. For an even number of results, the median is computed as the arithmetic average of two middle values.
Mode	A measure of the central tendency of a distribution, which is defined as the value in the population that occurs most frequently.
Naturally Occurring Background	Concentrations of naturally occurring chemicals in environments that have not been influenced by human activity.
Nondetects	Measurements reported by the analytical laboratory as below either the detection limit or the reporting limit.
Nonparametric Test	A statistical test that does not require any specific assumptions about the exact form of the underlying probability distributions of the investigated measures. Consequently, nonparametric tests generally are valid for a fairly broad class of distributions.
Normal (Gaussian) Distribution	A family of bell-shaped symmetrical distributions described by the mean and variance.
Null Hypothesis, H_o	The hypothesis that represents a theory that has been put forward, either because it is believed to be true or because it is to be used as a basis for argument, but has not been proved. The null hypothesis is assumed to be true, unless data and other evidence demonstrate otherwise with sufficient confidence.
Octanol-Water Partitioning Coefficient (K_{ow})	The ratio of the concentration of a chemical in octanol (an organic solvent used as a surrogate for natural organic matter) to its concentration in water at equilibrium and a specified temperature. Compounds with high K_{ow} values tend to adsorb readily to organic matter in soil or sediment.
Outlier	An unusually large or small measurement relative to the other measurements in the dataset, which is suspected to represent a population that is not the same as the population represented by the majority of the dataset.
Parametric Test	A test that requires specific assumptions about the exact form of the underlying probability distributions of the investigated measures. Parametric tests are not valid unless the underlying assumptions are met.
Partial Correlation	The correlation between two continuous variables that remains after the influence of one or more variables has been controlled or eliminated.

Permeability	In hydrogeology, the rate at which a fluid can flow through a porous medium. Permeability is a function only of the porous medium, i.e., it is independent of the fluid properties. Hydraulic conductivity is a function of the both the porous medium and the fluid.
pH	The negative logarithm of the hydrogen ion activity. The pH can range from 0 to 14, where 0 is the most acid, 7 is neutral, and 14 is the most basic or alkaline. Natural waters usually have a pH between 6.5 and 8.5.
Population	The entire collection of items that constitute the variable of interest.
Porosity	The ratio of the volume of void spaces in a porous medium to the total volume of the medium.
Power	The probability that a test will reject the null hypothesis, when the alternative hypothesis is true. Power is defined to be $1 - \beta$.
Probability Distribution	The rule for describing the probability measures associated with all the values of a random variable. For a discrete random variable, the probability distribution is described in terms of a probability mass function, which is a list of probabilities associated with each of the possible values of the discrete random variable. For continuous random variables, the probability distribution is described in terms of a probability mass function.
Pth Percentile	The specific value of a distribution that divides the set of measurements in such a way that the *P* percent of the measurements fall below (or are equal to) this value, and 100 – P percent of the measurements exceed this value.
Pth Quantile	The specific value of a distribution that divides the set of measurements in such a way that the proportion, *P*, of the measurements fall below (or are equal to) this value, and the proportion 1 – P of the measurements exceed this value.
Random Sample	A set of items that have been drawn from a population in such a way that each time an item was selected, every item in the population had an equal opportunity to appear in the sample. In environmental field investigations, random samples imply data that are collected in an unbiased, uncorrelated, and nonclustered manner.
Range	In descriptive statistics, the difference between the highest and lowest measured value. In geostatistics, the separation distance between any pair of measured values beyond which the pair are uncorrelated.
Reduction-Oxidation (Redox) Potential (Eh)	The energy gained in the transfer of 1 mole of electrons from an oxidant to hydrogen (H_2). The phase equilibrium (i.e., partitioning between the solid and aqueous phases) of many metals and other chemicals in the groundwater environment is strongly controlled by Eh.
Reference Area	An area where chemicals detected in groundwater are attributed to natural or anthropogenic background sources only. Also referred to as *background area.* Background or reference areas are often located upgradient of the impacted portions of a groundwater investigation site.
Regression	A set of techniques to characterize the manner in which one of the measures changes as the other measure changes.

Reporting Limit	The project-specific threshold limit established for a project for which, below a numerical value, the data are reported as nondetect (U) and presented as less than (<) a numerical value. This value is typically one to five times the detection limit, depending on the analytical method and matrix. The detection limit can vary considerably from sample to sample because of matrix effects. Ideally, the reporting limit will not change, and will be set high enough to account for matrix effects, yet low enough to meet project-specific DQOs.
Sediment	Any materials deposited at the bottom of water bodies, such as oceans, rivers, lakes, harbors, and storm drains.
Significance Level	In statistical hypothesis tests, the significance level is a fixed probability tolerated of wrongly rejecting the null hypothesis (H_o). It is the probability of a Type I error and is set by the investigator in relation to the consequences of such an error. Usually, the significance level is chosen to be 0.01, 0.05, or 0.10 (i.e., 1%, 5%, or 10%).
Site	A zone designated for investigation because of actual, suspected, or potential chemical releases. A site usually consists of both impacted and background areas. Site-specific field data are used to evaluate the extent of each area.
Skewness	A measure of asymmetry of the distribution of the sample data values.
Specific Conductivity	A measure of the ability of water to carry an electric current. Specific conductivity increases in direct proportion to the total concentration of ions in the water, and is therefore an indication of the concentration of total dissolved solids (TDS) in the solution.
Standard Deviation(s) (r^2)	A measure of dispersion of the distribution of the sample data values.
Standard Error	A measure of the variability (or precision) of a sample estimate, such as the computed arithmetic mean. Standard errors are needed to construct confidence intervals for the computed statistics.
Statistic	A measure of a statistical property of a population, computed based on sample results. An example of a statistic is the mean (i.e., average) of the measures in the sample.
Target Chemicals	Chemicals selected for background analysis. The list of target chemicals should include chemicals used at the site during its history, potential chemical breakdown products, potential site-specific background chemicals, chemicals required by regulatory criteria, and other parameters that may be useful for background analysis.
Target Population	The set of environmental space/time units within spatial and temporal boundaries for which a decision is needed on whether a chemical of interest is a COPC.
Total Dissolved Solids (TDS)	The total mass of solids dissolved in a water sample. Because TDS is directly proportional to specific conductivity, relative TDS concentrations can be determined in the field with an instrument that measures the conductivity of a sample.
Total Organic Carbon (TOC)	The total mass of organic carbon present in a water or soil sample.

Transformation to Linearity	A transformation of a response variable, or independent variable, or both, that produces an approximate linear relationship between the variables.
Type I Error	Falsely rejecting the null hypothesis when it is in fact true, and accepting the alternative hypothesis.
Type II Error	Falsely accepting the null hypothesis as being true.
Variance	A measure of dispersion of the distribution of a set of data values. The variance is the square of the standard deviation.
Variogram	A plot of the variance (one-half the mean squared difference) of paired sample measurements as a function of the distance (and optionally of the direction) between samples. Typically, all possible sample pairs are examined. Variograms provide a means of quantifying the commonly observed relationship that samples close together tend to have more similar values than samples far apart.

1. INTRODUCTION

This guidance document supports and implements Chief of Naval Operations (CNO) background policy by providing detailed instructions for evaluating background chemicals in groundwater. Background chemicals are derived from natural and anthropogenic sources not associated with site-related chemical releases (i.e., sources not related to site-specific activities or operations). Background analyses are essential for distinguishing between groundwater that has been impacted by a site-related chemical release and groundwater that has not been impacted by a site-related release.

In addition to contaminants associated with Navy and non-Navy chemical releases, naturally occurring background chemicals and anthropogenic background chemicals associated with nonpoint sources may occur in groundwater. The Navy is not responsible for cleanup of either natural or anthropogenic background chemicals; therefore, at most Navy groundwater investigation sites, background analysis will be necessary to define the nature and extent of site-related groundwater contamination and assess the need for cleanup.

1.1 Navy Policy and Guidance

The Navy has issued policy and guidance documents that address the role of background data in the Environmental Restoration Program, and present techniques for background analysis. Navy risk assessment policies specify requirements for the use of background data in human health and ecological risk assessments.

1.1.1 Navy Background Policy

The Navy *Interim Final Policy on the Use of Background Chemical Levels* was released by the CNO in September 2000 (DON, 2000b). The policy stresses the importance of eliminating background chemicals from the list of chemicals of potential concern (COPCs) carried through the risk assessment, and setting cleanup levels above the background range.

The policy specifically requires the following:

- ❑ Chemicals that may have been released at the site must be clearly identified to ensure that the Navy is focusing on remediating COPCs associated with the release.
- ❑ Chemicals detected at concentrations below the upper bound of the background range must be excluded from the full baseline risk assessment. All chemicals screened out because of background considerations must be discussed and documented in the risk characterization sections of the baseline risk assessment report.
- ❑ Cleanup levels must not be below the upper bounds of background ranges.

Conducting a background analysis early in the site investigation or groundwater monitoring process helps ensure that the Navy's cleanup responsibilities are clearly defined. The background analysis must be supported by adequate chemical, hydrogeologic, and geochemical information. Limited data (e.g., a few background groundwater samples from a single monitoring well) will not be sufficient to develop a defensible background analysis. The background analysis considers both naturally occurring and anthropogenic sources, as shown on Figure 1-1. The COPC selection process (which includes elimination of chemicals on the basis of the background analysis) should be discussed with regulators and conveyed to the community as early as possible. The methods used for background analysis must be scientifically based, technically defensible, and cost-effective.

As noted in the Navy background policy, in some cases unacceptable risks may be associated with chemical concentrations within the background range. Although this risk is outside the scope of the Navy's Environmental Restoration Program, restoration program personnel should convey the information to stakeholders. Chemicals screened

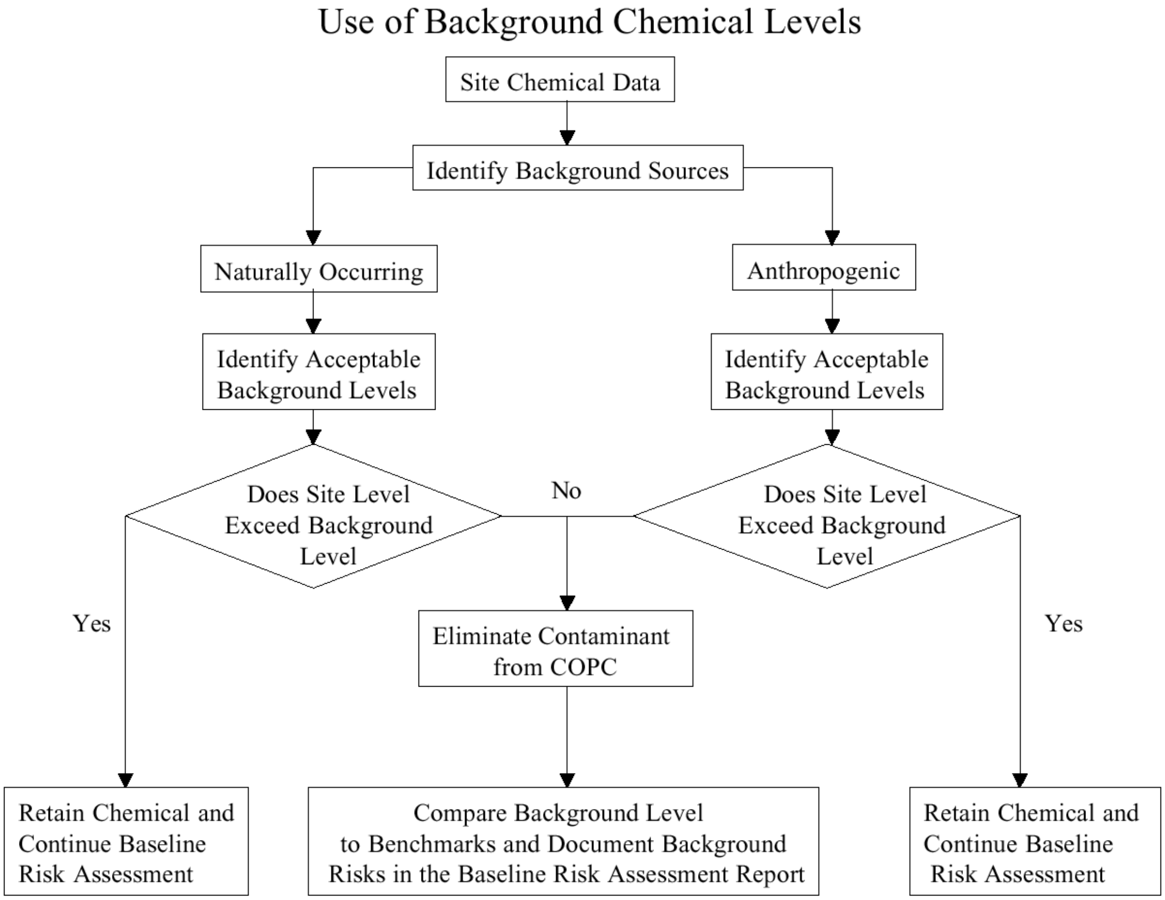


FIGURE 1-1. Navy policy on use of background chemical levels (DON, 2000b)

out because of background considerations should be evaluated against the appropriate risk-based screening criteria, and the results should be documented in the risk characterization sections of the baseline risk assessment report.

Cleanup efforts should be limited to chemicals associated with a site-related release that may pose unacceptable risks to human health or the environment. Background concentrations of metals in natural (unimpacted) groundwater can exceed regulatory and risk-based standards. However, the cleanup goal should never be set at a point below the upper bound of the background concentration range.

1.1.2 Navy Background Guidance

Although no Navy guidance specifically addresses analysis of groundwater background conditions, Navy Engineering Field Divisions/Activities (EFDs/EFAs) and the Naval Facilities Engineering Command (NAVFAC) have generated the following technical guidance documents for background data analysis:

- ***SWDIV/EFA West:*** *Procedural Guidance for Statistically Analyzing Environmental Background Data* (DON, 1998) and *Handbook for Statistical Analysis of Environmental Background Data* (DON, 1999a). These guidance documents provide detailed step-by-step instructions for graphical and statistical background analysis (i.e., statistical comparison of site data to reference area background data).

- ***NAVFAC:*** *Guidance for Environmental Background Analysis Volume I: Soil* (DON, 2002) and *Guidance for Environmental Background Analysis Volume II: Sediment* (DON, 2003). These documents provide detailed instructions for analysis of background conditions in soil and sediment. The guidance and procedures presented in these documents expand on principles described in the SWDIV/EFA West guidance document. Some of the guidance for analysis of groundwater background conditions presented in this document (Volume III of the series) builds on basic principles described in Volumes I and II.

1.1.3 Navy Risk Assessment Guidance

The Navy has issued the following policies to provide guidance on the use of background chemical concentrations in human health and ecological risk assessments:

- *Navy Policy for Conducting Human Health Risk Assessments Related to the Installation Restoration Program* (DON, 2001b)
- *Navy Policy for Conducting Ecological Risk Assessments* (DON, 1999b).

According to both policies, background chemicals should be screened out as early in the risk assessment process as possible. In human health risk assessments, background chemicals should be considered during the Tier 1A portion of the evaluation. In ecological risk assessments, background chemicals should not be evaluated until Step 3a of the Tier 2 baseline ecological risk assessment. The focus of subsequent risk calculations should be COPCs (i.e., chemicals detected at concentrations above the upper level of the background range).

1.2 Scope of the Groundwater Background Analysis Guidance Document

The purpose of this document is to provide detailed guidance on the following actions concerning background chemical analysis: (1) identifying background chemicals (i.e., chemicals from either natural or anthropogenic sources), (2) distinguishing between background chemicals and chemicals associated with a site-related release, (3) estimating the concentration ranges that represent background conditions, and (4) comparing chemical concentrations detected in groundwater at a suspected chemical release site to estimated background concentration ranges. For the purposes of this document, background conditions are defined as ambient chemical concentrations in groundwater that has not been impacted by either a site-related or a nonsite-related point source chemical release. Statistical procedures for comparison of site and background datasets often are required for groundwater monitoring; and, although these procedures are discussed within this document, the document does not provide guidance for implementing groundwater monitoring programs. The Naval Facilities Engineering Service Center (NFESC) has published an Interim Final *Guide to Optimal Groundwater Monitoring* (DON, 2000a). The objective of the *Optimal Groundwater Monitoring* document is to provide the information that Navy Remedial Project Managers (RPMs) and contractors need to design new groundwater monitoring programs or optimize existing programs in order to cost-effectively achieve monitoring objectives.

Many United States Environmental Protection Agency (U.S. EPA) and state guidance documents address procedures for groundwater monitoring and comparison between site groundwater data and fixed concentration standards, background data, or baseline (e.g., upgradient, side-gradient) data (see Section 1.4.2); however, like the *Guide to Optimal Groundwater Monitoring* (DON, 2000a), these documents do not focus on methods for identifying adequate background datasets and verifying that a dataset accurately represents background conditions. Therefore, this guidance document outlines procedures for evaluating chemical data, hydrogeologic characteristics, and geochemical parameters in order to estimate the upper bounds of background chemical concentration ranges and to identify and/or acquire datasets that adequately represent background conditions. This document also presents statistical techniques for comparing site data to background data in order to distinguish between groundwater that has been impacted by a site-related chemical release and groundwater that has not been impacted by a site-related release.

Resource Conservation and Recovery Act (RCRA) requirements define three distinct groundwater monitoring phases: detection monitoring, assessment (or compliance) monitoring, and corrective action monitoring. This document focuses primarily on groundwater background analysis and the role of background data in assessment monitoring (i.e., site-wide comparisons). However, an overview of the statistical testing approaches used for detection monitoring (i.e., individual comparisons) is presented in Section 4. The site-wide comparison techniques presented in this document also can be applied to corrective action monitoring plans (U.S. EPA, 1994b).

The focus of this guidance document is on background chemicals in groundwater; detailed discussions of the fate and transport of COPCs in groundwater are beyond the scope of the document. Similarly, this guidance does not address "baseline" conditions (as noted in Section 1.5.3, baseline conditions may be affected by chemicals associated with contaminated groundwater originating at upgradient properties) or co-mingled plumes of contaminated groundwater. Furthermore, the scope of this document includes analysis of background conditions that show a cyclical variation (e.g., seasonal trends); however, the document does not cover analysis of baseline or background conditions that show trends associated with significant changes in upgradient chemical characteristics (e.g., input of contaminants or significant increases in anthropogenic impacts).

Both natural processes (e.g., weathering and dissolution of naturally occurring metallic minerals) and anthropogenic processes (e.g., infiltration of chemicals including metals, pesticides, fertilizers, and other chemicals from urban and agricultural runoff) can result in elevated concentrations of various chemicals—including hazardous substances—in otherwise unimpacted groundwater. These background chemicals are derived from natural or nonpoint anthropogenic sources, and are not associated with site-related chemical releases. To evaluate the nature and extent of chemical contaminants, groundwater samples are analyzed for chemicals that may have been released during site activities. Because chemicals associated with a known or suspected Navy release also may be derived from background sources, background analysis should be conducted early in the site investigation process. This will ensure that only groundwater that has been impacted by a Navy chemical release is targeted for cleanup. Failure to distinguish between concentrations associated with a site-related chemical release and background levels may lead investigators to establish cleanup levels within the background range, resulting in unnecessary and costly remediation and potentially delaying property transfer and re-use. Furthermore, as discussed in Section 1.4, cleanup of chemicals present at concentrations within the background range is not consistent with established environmental regulations, policies, and guidelines.

The procedures presented in this guidance document will allow Navy environmental restoration personnel to technically and defensibly differentiate between chemicals and chemical concentrations associated with site-related releases and corresponding data representing background conditions. The techniques described in this document can be used to evaluate background levels of either organic or inorganic chemicals. Background chemicals occur naturally in all groundwaters and may be present at concentrations high enough to pose unacceptable risks to human and ecological receptors and exceed promulgated standards. Therefore, background concentrations must be characterized for most groundwater investigation sites in order to accurately evaluate the nature and extent of site-related contamination and assess the associated risks.

Techniques for statistical analysis of chemical data are integral parts of groundwater monitoring and background analysis. However, these techniques alone are not sufficient to understand and define the background conditions that exist at a particular site. An understanding of the geologic, geochemical, and hydrogeologic processes that control the occurrence and concentrations of naturally occurring chemicals in groundwater is also essential. This document pays special attention to the evaluation of physical and chemical processes that control metal concentrations in groundwater, including the geochemical characteristics of the groundwater environment.

This guidance document presents detailed instructions for background analysis and statistical comparative procedures, including:

- ***Section 2 (Data Review, Assessment, and Exploratory Analysis)*** presents procedures for evaluating the available groundwater data in order to identify the site-specific geologic, geochemical, and hydrogeologic processes that control the occurrence and concentrations of naturally occurring chemicals in groundwater.
- ***Section 3 (Identify Background Sampling Locations and Concentration Ranges)*** presents procedures for characterizing groundwater, identifying background wells, (i.e., wells that can be sampled to provide background data for comparison to data representing site groundwater), and estimating the background concentration range for each target COPC.
- ***Section 4 (Comparative Method)*** presents statistical testing methods for comparing site individual or pooled measurements to data that represent background conditions, in order to determine whether differences between the two populations are statistically significant. If the statistical tests indicate that COPC concentrations in the site groundwater are significantly higher than COPC concentrations in the background groundwater, then the data indicate that a site-related release has impacted groundwater.

1.3 Role of Background Analysis in the Environmental Restoration Program

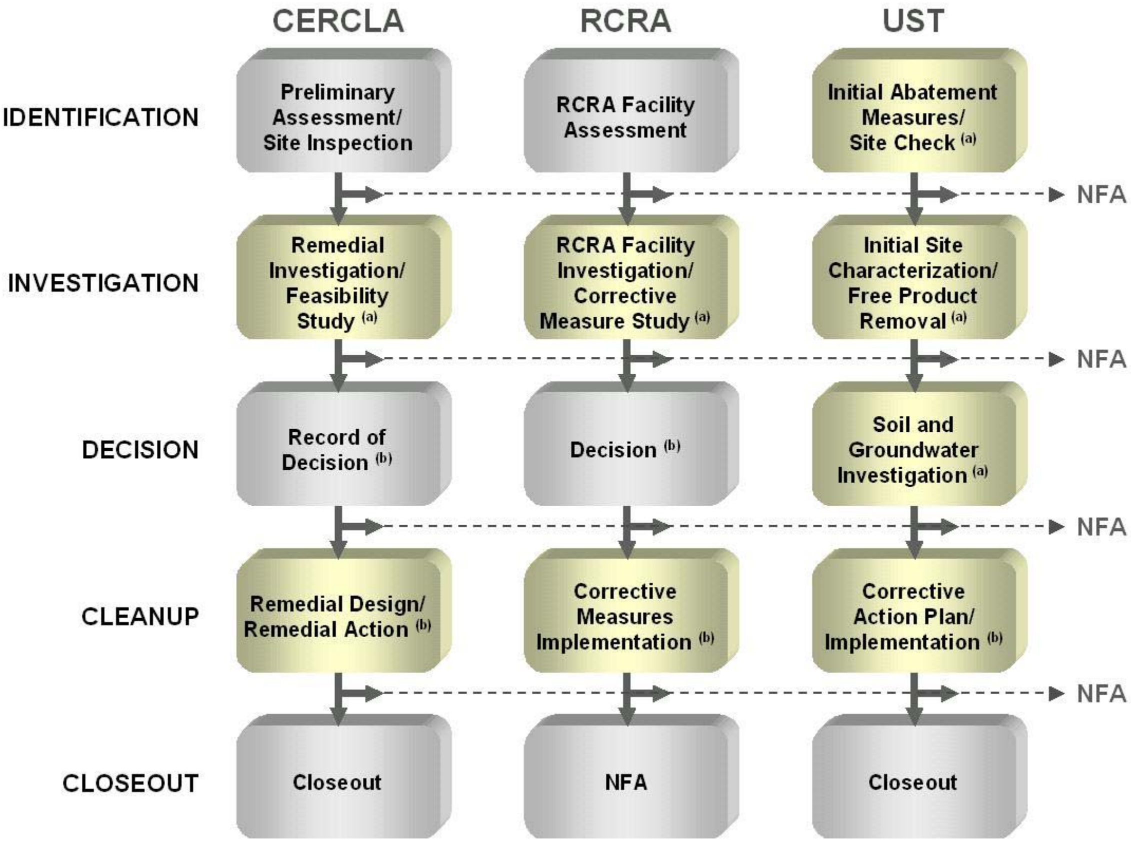
As noted in Section 1.4, federal law requires the Navy to protect human health and the environment and comply with applicable or relevant and appropriate requirements (ARARs) at all Navy chemical release sites. The Navy's responsibilities for site cleanup or other response actions necessary to comply with environmental laws and regulations cannot be defined until the nature and extent of contamination associated with a chemical release have been characterized. To evaluate the nature and extent of contamination, the Navy must distinguish between contamination associated with a chemical release and naturally occurring or anthropogenic background conditions; therefore, background analysis is an integral part of the environmental assessment, decision-making, and cleanup process.

Figure 1-2 illustrates the roles of background analysis in environmental investigation and restoration activities conducted under the three primary regulatory frameworks (CERCLA, RCRA, and the UST Program). Background analysis should be incorporated into the site identification, investigation, decision, and cleanup phases as necessary to achieve the following objectives:

- Identify chemicals associated with site-related releases (i.e., COPCs) for further consideration during the risk assessment process.
- Eliminate naturally occurring and anthropogenic chemicals that occur at concentrations within the background range from the list of target COPCs.
- Ensure documentation and discussion of chemicals eliminated from the list of target COPCs.
- Ensure adequate delineation of groundwater plumes with COPC concentrations above background levels.
- Ensure that the cleanup level established for each COPC is not below the upper bound of the site-specific background range.
- Obtain stakeholder concurrence with site characterization and remedial conclusions.

It should be noted that retention of a chemical as a COPC after comparison to background levels does not necessarily imply that the chemical will require cleanup. Decisions regarding cleanup should be made only after the baseline risk assessment and ARAR review are complete.

Background conditions should be systematically quantified to minimize uncertainties introduced by qualitative or semiquantitative background investigations. If background conditions are not understood, it may be impossible to determine whether a chemical release has occurred, to adequately define the nature and extent of contamination, or



Background Analysis Applications:

(a) Site Characterization:	Delineation of impacted media above background levels
Risk Assessment:	Determination of COPCs/computation of residual risks
(b) Cleanup:	Ensuring cleanup goals ≥ background levels

CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
RCRA	Resource Conservation and Recovery Act
UST	underground storage tank
NFA	no further action

FIGURE 1-2. Roles of background analysis within regulatory cleanup frameworks

to assess the residual contribution to risk made by the release. Background conditions must be characterized before the risk assessment process is completed so appropriate cleanup or other response actions can be recommended and implemented. Background analysis is essential for setting reasonable and attainable cleanup goals if cleanup is required. Project teams should never define cleanup goals within the background range established for the site.

The background analysis procedures described in this guidance document will reduce the probability of decision error, and therefore can help to reassure stakeholders and regulators that the Navy has made the correct decisions regarding response action for a particular site. In addition, by limiting the probability of decision error to acceptable levels, the Navy can avoid the unnecessary expense associated with cleaning up sites that do not require action to protect human health and the environment and thereby help ensure that funding is appropriately focused on contaminated sites. The project team should use the Data Quality Objectives (DQO) process (see Section 2.1) to reach agreement with stakeholders over acceptable decision error probability. Finally, background data are essential to provide scientifically defensible evidence to support a decision of no further action for a site; and, if cleanup is necessary, background data can aid in the selection of technically feasible and cost-effective remedial alternatives.

1.4 Statutory Requirements, Regulations, and Guidance

Numerous federal and state laws, regulations, and guidance pertain to background analysis and groundwater monitoring. Federal law requires the Navy to protect human health and the environment and comply with ARARs at all Navy chemical release sites. The Navy policy and guidance documents summarized above were developed to ensure compliance with laws and regulations that address background analysis and its role in the site assessment and cleanup process. The *Navy Installation Restoration Manual* (DON, 2001a) provides an overall synopsis of the environmental laws and regulations that define and affect the Navy Installation Restoration (IR) Program, and describes the procedures the Navy has developed to ensure compliance with these laws and regulations. As specified in the IR Manual, all actions at IR sites shall comply with the following:

- Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)
- Superfund Amendments and Reauthorization Act (SARA)
- National Oil and Hazardous Substances Pollution Contingency Plan (NCP).

Although compliance with U.S. EPA policy and guidance documents is not mandatory, the *IR Manual* also specifies that IR project teams shall reasonably interpret and apply U.S. EPA policy and guidance to make cleanup decisions and plan response actions.

1.4.1 Federal Laws and Regulations

Federal laws, including CERCLA/NCP, require protection of human health and the environment at sites where CERCLA hazardous substances have been released. To comply with CERCLA/NCP, the Navy must determine whether CERCLA hazardous substances have been released to the environment, assess the nature and extent of contamination, and evaluate associated risks to human health and the environment. To achieve these objectives, background analysis must be included as an integral component in the site assessment and cleanup process.

Federal Regulations Related to Background

U.S. EPA requirements acknowledge the importance of background analysis. In 42 USC §9604(a)(3)(A), CERCLA recognizes that remediation may be impractical or impossible if naturally occurring background levels are higher than regulatory criteria. According to this section:

> “The President shall not provide for a removal or remedial action under this section in response to a release or threat of a release of a naturally occurring substance in its unaltered form, or altered solely through naturally occurring processes or phenomena, from a location where it is naturally found.”

This provision clearly indicates that CERCLA does not require cleanup of background chemicals, even when they occur at concentrations that exceed federal, state, or local regulatory criteria.

RCRA requirements also recognize the importance of background analysis. RCRA gives the U.S. EPA authority to require cleanup of releases that impact environmental media within designated solid waste management units (SWMUs) (RCRA 3004[u]), and establishes requirements for groundwater monitoring.

Federal Regulatory Programs Requiring Groundwater Monitoring

Federal regulatory programs that may require groundwater monitoring (which frequently involves comparison of site data to background data) include the following:

- ***Resource Conservation and Recovery Act (RCRA).*** RCRA (Public Law 94-580

[1976]) regulates hazardous wastes under Subtitle C. Nonhazardous wastes (e.g., municipal solid wastes) are regulated under Subtitle D, and certain underground storage systems are regulated under Subtitle I. RCRA establishes specific requirements for groundwater monitoring under Subtitles C and D.

- ***Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA).*** CERCLA establishes no specific requirements for monitoring groundwater; however, groundwater monitoring may be required to meet overall CERCLA objectives: protection of human health and the environment, and compliance with ARARs.

- ***Toxic Substances Control Act (TSCA).*** TSCA (Public Law 94-469 [1976]) gives the U.S. EPA authority to identify and control chemicals that may pose unreasonable risks to human health or the environment during manufacture, distribution, processing, use, or disposal (particularly polychlorinated biphenyls [PCBs]). For example, initial groundwater monitoring is required to evaluate background groundwater conditions at chemical disposal sites.

- ***Clean Water Act (CWA).*** The CWA (Public Laws 92-500 [1972] and 95-217 [1977]) regulates discharges into surface waters. Because groundwater may discharge to surface water bodies, groundwater monitoring may be required to ensure compliance with the CWA.

- ***Safe Drinking Water Act (SDWA).*** The SDWA (Public Act 93-523 [1974]) provides for protection of drinking water supplies. The SDWA establishes chemical-specific criteria (MCLs) that specify the maximum chemical concentrations allowed in drinking water sources. MCLs may be used as groundwater protection standards for monitoring programs regulated under RCRA or CERCLA.

Overview of RCRA Groundwater Monitoring Requirements

Sites regulated under RCRA (e.g., municipal solid waste landfills, hazardous waste landfills, designated SWMUs, and certain underground storage systems) are subject to RCRA groundwater monitoring requirements. RCRA sections that specify requirements for groundwater monitoring include:

- ***40 CFR Part 258.*** Criteria for municipal solid waste landfills

- ***40 CFR Part 264.*** Regulations for owners and operators of permitted hazardous waste facilities

- ***40 CFR Part 265.*** Interim status standards for owners and operators of hazardous waste facilities

- ***40 CFR Part 267.*** Interim standards for owners and operators of new hazardous waste facilities

- ***40 CFR Part 270.*** Regulations for federally administered hazardous waste permit programs (Part B Permits)

- ***40 CFR Part 271.*** Requirements for authorization of state hazardous waste programs.

Compliance with RCRA groundwater monitoring requirements often involves comparing site groundwater data to background or baseline data. RCRA groundwater monitoring consists of three different phases: detection monitoring (i.e., individual comparisons), compliance or characterization and assessment monitoring (i.e., site-wide comparison), and corrective action. As noted above, this document focuses primarily on groundwater background analysis and the role of background data in characterization or assessment monitoring; however, an overview of the statistical testing approaches used for detection monitoring is presented in Section 4. The site-wide comparison techniques presented in this document also can be applied to reference-based corrective action monitoring plans (U.S. EPA, 1994b).

1.4.2 Guidance for Analysis of Groundwater Monitoring Data

Guidance for statistical analysis of groundwater monitoring data is available in numerous U.S. EPA, state, and private publications. These guidance documents typically describe preferred monitoring procedures and statistical comparison testing approaches that will comply with both RCRA and state-specific requirements, and generally focus on various statistical techniques for comparing site groundwater data to background or baseline (e.g., upgradient) groundwater data, or to fixed concentration limits (e.g., MCLs or alternate concentration limits [ACLs]). The guidance documents present various methods for interwell comparisons (i.e., comparison of concentrations detected in different wells to distinguish between impacted and unimpacted groundwater), and intrawell comparisons (i.e., comparative analysis of data representing a series of samples collected from a single well to identify statistically significant concentration trends that could indicate a chemical release). Box 1-1 lists several examples of guidance for groundwater monitoring data analysis.

1.4.3 U.S. EPA and State Background Guidance

The U.S. EPA and state technical guidance documents described in this section address issues related to background, including criteria for

BOX 1-1. Examples of guidance for groundwater monitoring data analysis

- ***U.S. EPA Office of Solid Waste:*** The interim final guidance document *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities* (U.S. EPA, 1989c) and its addendum (U.S. EPA, 1992d) provide guidance for statistical analysis of groundwater monitoring data at RCRA facilities, including comparison of site data to background, baseline, or fixed concentration limits, and intrawell comparisons.
- ***U.S. EPA Office of Policy, Planning, and Evaluation:*** *Methods for Evaluating Attainment of Cleanup Standards, Vol. 2: Ground Water* (U.S. EPA, 1992b) provides guidance for statistical analysis of groundwater monitoring data to verify attainment of groundwater cleanup standards and determine when remedial actions can be discontinued.
- ***American Society for Testing and Materials (ASTM):*** *D6312-98 Standard Guide for Developing Appropriate Statistical Approaches for Ground-Water Detection Monitoring Programs* (ASTM, 2003) describes several statistical approaches for analysis of groundwater monitoring data from hazardous and municipal solid waste disposal facilities. Objectives of the statistical testing are to detect a potential release from the facility as soon as possible while minimizing the probability of falsely concluding that the facility has impacted groundwater.
- ***Gibbons, R.D., 1994:*** The textbook *Statistical Methods for Groundwater Monitoring* (Gibbons, 1994) presents detailed discussions of detection monitoring to identify potential releases to groundwater, including statistical testing to compare individual site data to background, baseline, or fixed concentration limits, and intrawell comparisons. The text also addresses regulatory issues associated with statistical analysis of groundwater monitoring data, identifies computer software for implementing the statistical testing methods, and describes statistical methods (including several U.S. EPA-recommended methods) that should be avoided.
- ***Ohio Environmental Protection Agency (OEPA), Division of Drinking and Ground Waters:*** *Technical Guidance Manual for Hydrogeologic Investigations and Ground Water Monitoring* (OEPA, 1995) identifies recommended technical procedures for groundwater monitoring and statistical testing for comparison of groundwater monitoring data.
- ***Hawaii Department of Health (HDOH), Solid and Hazardous Waste Branch:*** *Landfill Groundwater Monitoring Guidance Document, Version 1.8* (HDOH, 2002) describes hydrogeologic and statistical methods for detection monitoring, assessment monitoring, and corrective action to meet RCRA 40 CFR Part 258 (Subtitle D) and State of Hawaii requirements for groundwater monitoring at solid waste facilities.
- ***Virginia Department of Environmental Quality (VDEQ):*** *Data Analysis Guidelines for Solid Waste Facilities* (VDEQ, 2003) provides guidance for statistical analysis of groundwater monitoring data collected at solid waste facilities, including methods for interwell and intrawell comparisons.

determining when a release has occurred, identifying cleanup requirements, and evaluating background concentration ranges. Although these guidance documents generally focus on background conditions in soil or sediment, some of the guidance also is applicable to groundwater background issues.

Examples of U.S. EPA background guidance are summarized in Box 1-2. Regional offices of the U.S. EPA have also issued technical guidance on background analysis; examples are summarized in Box 1-3. Also, several states have developed regulations and guidance that specifically address background data evaluation; representative examples are summarized in Box 1-4.

1.5 Key Definitions

Some of the keywords and concepts used in this guidance document are defined below; a more comprehensive list of terms is provided in the glossary section.

BOX 1-2. Examples of U.S. EPA background-related guidance

- ***U.S. EPA Office of Emergency and Remedial Response (OERR):*** OERR has published a number of guidance documents that describe how background levels should be considered under the CERCLA and RCRA programs. The OERR guidance manual *Risk Assessment Guidance for Superfund* (RAGS) stresses the importance of background analysis. Volume I, the *Human Health Evaluation Manual (Part A)* (U.S. EPA, 1989b), notes that “Background sampling is conducted to distinguish site-related contamination from naturally occurring or other nonsite-related levels of chemicals.” According to U.S. EPA’s *Guidance for Data Useability in Risk Assessment* (U.S. EPA, 1992a), one of the primary objectives of sampling and analysis programs for sites of suspected environmental contamination should be to determine “whether site concentrations are sufficiently different from background.” Similarly, U.S. EPA’s *Guidance for Conducting Remedial Investigations and Feasibility Studies Under CERCLA* (U.S. EPA, 1988) states that background sampling should be conducted to differentiate between chemical releases resulting from site operations and background conditions.
- ***U.S. EPA Office of Solid Waste:*** According to *RCRA Facility Investigation Guidance* (U.S. EPA, 1989a), “High variability in the chemical composition of soils makes determination of background levels for the constituents of concern essential. This is particularly important for quantification of toxic metals, because such metals commonly occur naturally in soil.”
- ***U.S. EPA Office of Solid Waste and Emergency Response (OSWER):*** OSWER has published an Engineering Forum Issue paper, *Determination of Background Concentrations of Inorganics in Soils and Sediments at Hazardous Waste Sites* (U.S. EPA, 1995), for educational use by its project managers. This publication provides a highly informative summary of the technical issues that should be considered to determine whether the concentrations of inorganic chemicals detected at a site are elevated relative to local background conditions. The first portion of the paper presents definitions and important factors influencing background concentrations. Issues addressed include selection of background sampling locations, considerations for selecting sampling procedures, and statistical analyses for determining if chemical concentrations at a background area and a waste site are significantly different. The paper notes in particular that it is not feasible to establish a single bright-line concentration value to define background for a particular chemical—background should instead be expressed as a concentration range determined by statistical analysis of the chemical data. The second portion of the paper is divided into two parts. Part A presents procedures for determining whether hazardous waste site-related activities have resulted in increased inorganic chemical concentrations in soils and sediments compared to background concentrations. These procedures are based on the approach employed by the State of Michigan Department of Environmental Quality (MDEQ, 1991, 1994). Part B presents approaches for determining background levels of inorganic chemicals at CERCLA sites and is a modification of a U.S. EPA issue paper addressing background (U.S. EPA, 1992c). Both sets of procedures are based on the comparative statistical approach to establishing background, which requires sampling of an off-site reference area.

BOX 1-3. Examples of U.S. EPA regional background guidance

- ***U.S. EPA Region 1:*** *Risk Updates (No. 5),* COPC Selection Process Update (U.S. EPA, 1999). In this bulletin, U.S. EPA Region 1 clarifies their intent to ensure that background chemicals are carried through the risk evaluation process. This regional guidance suggests that the relevance of background concentrations should be discussed in the risk characterization or uncertainty sections of the risk assessment.
- ***U.S. EPA Region 4:*** *Statistical Tests for Background Comparison at Hazardous Waste Sites* (Interim Draft Supplemental Guidance to RAGS) (U.S. EPA, 1998). This regional guidance document provides details of the statistical approaches that the Region 4 Office of Technical Services considers appropriate for comparing site chemical concentrations to background levels when selecting COPCs (see Section 4).
- ***U.S. EPA Region 8:*** *Evaluating and Identifying Contaminants of Concern for Human Health* (U.S. EPA, 1994a). This regional guidance document is intended to clarify the evaluation process for selecting COPCs for the human health baseline risk assessment process. In this bulletin, U.S. EPA Region 8 recommends the use of distributional tests (statistical tests used to determine if the central tendencies of two datasets are similar) in order to compare measured on-site datasets to background datasets (see Section 4).

BOX 1-4. Examples of state background guidance

- ***California Department of Toxic Substances Control (DTSC):*** DTSC provided a Final Policy entitled *Selecting Inorganic Constituents as Chemicals of Potential Concern at Risk Assessments at Hazardous Waste Sites and Permitted Facilities* (Cal/EPA, 1997). This policy describes the use of graphical techniques including probability plots, and describes various statistical tests used to distinguish background chemicals from site-related COPCs.
- ***Florida Department of Environmental Protection (FDEP; formerly Florida Department of Environmental Regulation):*** FDER published a technical document titled *A Guide to the Interpretation of Metal Concentrations in Estuarine Sediments* (FDER, 1988). This publication describes a very useful general approach to distinguishing between chemical concentrations associated with pollution and chemical concentrations that represent natural (background) conditions. In 1995, the FDEP released a special notice related to this publication describing the importance of using appropriate analytical methods for metals analyses (FDEP, 1995).
- ***Michigan Department of Environmental Quality (MDEQ):*** The Michigan Environmental Response Act or MERA (307 Protection Act, 1982) specifies that background conditions must be identified to determine whether chemical concentrations are elevated as a result of a chemical release. *Operational Memorandum No. 15* (MDEQ, 1993) to MERA established default background concentrations based on a comprehensive background survey completed in 1991.
- ***Pennsylvania Department of Environmental Protection:*** Title 25, §250.202 of the Pennsylvania Code sets requirements for establishing background chemical concentrations.

1.5.1 Site and Background Groundwater

The following terminology is used throughout this document:

- ❑ *Site* describes a zone designated for investigation due to actual, suspected, or potential chemical releases. A site usually includes both impacted and unimpacted areas. Site-specific field data are used to evaluate the extent of each area.
- ❑ *Background* groundwater conditions are defined as the chemical characteristics of groundwater that has not been impacted by

either a site-related or a nonsite-related chemical release. Background conditions may be affected by natural and anthropogenic chemical sources.

- *Installation* or *facility* describes the extent of a Navy property at which one or more activities have been or are being conducted. An installation may contain a number of sites, as well as both impacted and unimpacted (background) areas.

The concept of impacted versus unimpacted groundwater is depicted on Figure 1-3, which illustrates the contributions to total concentrations made by site-related and background chemicals.

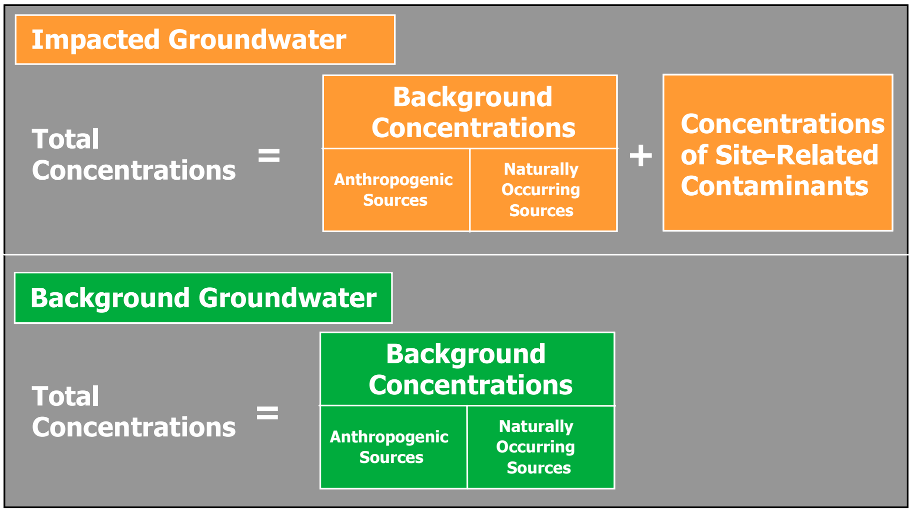
1.5.2 Background and Site-Related Chemicals

Background Chemicals

Chemicals derived from natural or anthropogenic sources not associated with site-related chemical releases (i.e., sources not related to activities or operations conducted at the site) are referred to as background chemicals. According to U.S. EPA (1989b), background chemicals fall into two categories:

- ***Naturally Occurring or Nonanthropogenic Chemicals:*** Chemicals present because of geochemical processes that have not been influenced by human activity. Naturally occurring organic and inorganic background chemicals in soil, sediment, and water are attributable to the natural geologic or hydrogeologic characteristics of the area. These chemicals have not been altered by human activity. Examples of natural sources of background chemicals include: minerals in soil and rock formations (metals), natural oil seeps (various hydrocarbons), and forest fires (polycyclic aromatic hydrocarbons [PAHs]).

- ***Anthropogenic Chemicals:*** Synthetic or natural substances that have been released to the environment as a result of human activities, but are not related to specific activities conducted at the site. These chemicals are usually ubiquitous near specific human activities that may impact very large areas. Anthropogenic background chemicals are generated by human activities, but are unrelated to specific point sources or site releases. U.S. EPA (1989b) cites the following sources of anthropogenic background chemicals: agricultural runoff, urban runoff, septic systems, air pollution, irrigation; agricultural and residential application of pesticides (e.g., arsenicals, DDT); industrial discharges, landfills, municipal sludge land application; and urban pollution (e.g., lead and PAHs from automobiles and combustion processes, salts used for road de-icing). Anthropogenic background chemicals typically are: (a) widely distributed in areas near specific human activities; (b) not related to site sources or releases; (c) not related to other point sources or releases; and (d) attributable to past or present legal applications or sources.



Source: Adapted from U.S. EPA (1995)

FIGURE 1-3. Concept of impacted versus unimpacted groundwater

Site-Related Chemicals

Site-related chemicals are chemicals released during past or current operations at an installation. In this document, such chemicals are referred to as COPCs if they may exist at concentrations that pose potentially unacceptable risks to human health or the environment. At some sites, elevated chemical concentrations may be the combined result of natural/anthropogenic (background) sources and a site-related chemical release.

Defining Background Conditions

Chemicals associated with background conditions and site-related releases, as well as chemicals that represent only background conditions, are routinely detected during sampling and analysis.

Unless background conditions are accounted for in one of the following two ways, project teams may unnecessarily remediate chemicals and groundwater in areas where no releases have occurred:

- ***Distinguishing COPCs from Background Chemicals.*** In this case, the project team must screen out background chemicals from the list of detected chemicals. Failure to make these distinctions could confound the investigation and remedial decisions. Chemicals detected at concentrations that do not exceed the upper bound of the background concentration range should be eliminated from consideration at the appropriate point in the risk assessment process (see Section 1.1). The remaining chemicals then are carried forward as COPCs for further evaluation during the risk assessment.

- ***Determining Background Levels of COPCs.*** If an individual chemical detected at a site is present due to both site-related and background sources, the project team will need to quantify the concentration range that represents background conditions. The upper bound of the background concentration range must be identified to (a) delineate the extent of a site-related chemical release; (b) calculate residual risks caused by a site-related release; and (c) determine the scope of required cleanup, should remediation become necessary.

1.5.3 Groundwater Background and Baseline Conditions

As noted in Section 1.2, this document focuses on identifying background chemicals (i.e., chemicals from either natural or anthropogenic sources), differentiating between background chemicals and chemicals associated with a site-related release, and estimating the concentration ranges that characterize background conditions. Although this document focuses on background conditions, baseline conditions can also be important considerations for groundwater monitoring programs.

- ***Background Groundwater Conditions.*** The chemical characteristics of groundwater not impacted by either a site-related chemical release or an off-site point source release. Background conditions can include the chemical characteristics of groundwater affected by both natural and anthropogenic sources.

- ***Baseline Groundwater Conditions.*** The chemical characteristics of groundwater not impacted by a site-related release. Groundwater can transport chemicals released at off-site areas to the site; therefore, baseline conditions can include the chemical characteristics of groundwater impacted by a source not related to the site. If groundwater is not impacted by off-site sources, background and baseline conditions will be the same.

If off-site chemical releases that could affect baseline groundwater conditions in the site vicinity are suspected, project teams should consult the groundwater monitoring literature for guidance (DON, 2000a).

1.6 Unique Aspects of Groundwater Background Evaluation

The unique aspects of the groundwater environment must be considered during background analysis. Due to the dynamic nature of the groundwater environment, methods and procedures for groundwater investigations and response actions are very different from those used for soil or sediment investigations and response actions.

1.6.1 Spatial Considerations

Chemicals can be transported with groundwater over long distances downgradient of their sources; groundwater flow direction is therefore a primary consideration when selecting locations for background groundwater sampling, or evaluating existing well networks. Areas upgradient of a site or potential contaminant source typically are considered the best locations for background groundwater monitoring wells. However, side-gradient or even downgradient locations also may be suitable for background wells. Chemical characteristics of groundwater—both background concentrations and COPC concentrations—can vary significantly with depth. Therefore, sampling depths, aquifer thickness, and chemical properties should be considered when evaluating background conditions. For example, light, nonaqueous-phase liquids (LNAPLs) will tend to float on the water table, and will therefore impact the upper portion of a groundwater body first, whereas dense, nonaqueous-phase liquids (DNAPLs) will sink toward the bottom of an aquifer, and therefore may impact the aquifer at greater depths. Hydrogeological and other characteristics that affect the spatial distribution of chemicals in the groundwater environment are discussed in more detail in Section 2.1.4.

1.6.2 Temporal Considerations

Variations in chemical characteristics of groundwater over time must also be considered for background groundwater analysis. Temporal variations are typically due to seasonal fluctuations in aquifer recharge rates. Semiannual or more frequent sampling over several years is required to accurately characterize seasonal trends. If these periodic trends are not accounted for during background analysis, the project team could erroneously identify a seasonal increase in background concentrations as evidence of a chemical release.

1.6.3 High Unit Costs of Groundwater Sampling

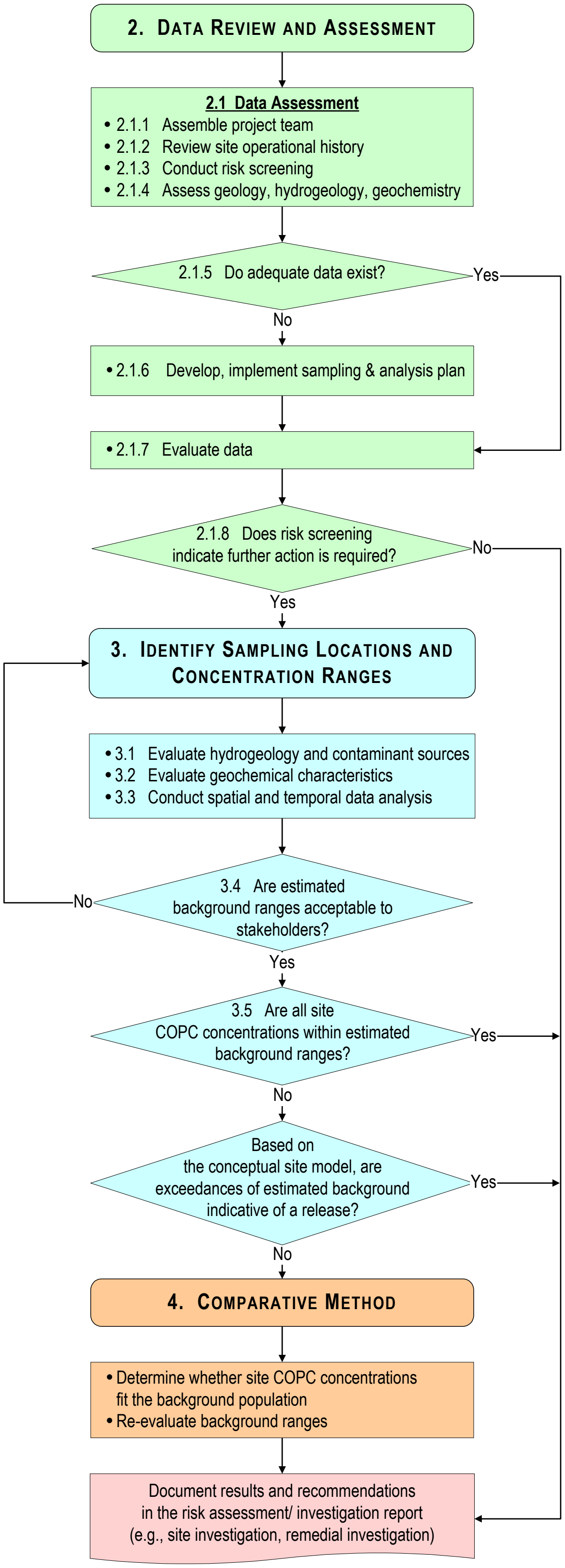
Groundwater sampling costs are usually far greater than soil or sediment sampling costs due to the expense associated with well installation and the time and equipment required to collect each groundwater sample. In addition, numerous sampling rounds may be required to establish the concentration pattern for each well. Budget constraints can therefore severely limit the amount of site and background data available for analysis. However, limited datasets can be used successfully to characterize background conditions by applying a “weight of evidence” approach, in which the chemical concentration data are augmented by additional lines of evidence based on hydrogeologic and geochemical knowledge.

1.7 Background Analysis Components Overview

The flowchart presented in Figure 1-4 illustrates the full sequence of background analysis and comparison of site data to background data for a groundwater investigation at a typical Navy installation. As indicated by the decision questions shown on the flowchart, certain steps can be bypassed depending on site conditions and project-specific requirements.

1.7.1 Groundwater Background Analysis

During the data review and assessment phase (Section 2), the data should be evaluated to identify data gaps, compare COPC concentrations to risk-based screening criteria, and determine the scope of supplementary sampling and analysis (if necessary). Background analysis requires representative datasets of adequate size to reliably characterize background conditions. Therefore, the adequacy of the site and background datasets should be addressed within the DQO and data quality assessment (DQA) framework. If the dataset contains insufficient data representing background conditions, additional sampling and analysis will be necessary to supplement the existing data. As discussed in Section 2.1.4, the project team must consider the hydrogeologic and geochemical characteristics of the site and potential background groundwater sampling locations. Once the dataset representing potentially impacted groundwater at the site (i.e., the suspected groundwater plume) is complete, representative exposure concentrations for the site groundwater can be compared to risk-based criteria. (Representative exposure concentrations for groundwater are



Notes:

Numeric prefixes correspond to groundwater guidance document section numbers.

COPC chemical of potential concern

FIGURE 1-4. Typical background data evaluation and comparative statistical testing for groundwater COPCs

usually conservatively defined as the maximum detected concentrations.) Chemicals with representative exposure concentrations below their risk-based criteria do not require background analysis and can be screened out. (Chemicals without risk-based screening criteria cannot be screened out at this stage.)

Chemicals not screened out after comparison to the appropriate criteria should be carried forward for background analysis by the "weight of evidence" methodology (Section 3). This methodology uses a three-step process to assemble the lines of evidence required to characterize groundwater background conditions:

- ❑ Step 1: Evaluate Geology, Hydrogeology, and Contaminant Sources
- ❑ Step 2: Evaluate Geochemical Conditions
- ❑ Step 3: Conduct Spatial and Temporal Groundwater Data Analysis

Objectives of these procedures are twofold: (1) to identify background monitoring wells, i.e., wells that can be sampled to provide background data for comparison to data representing potentially impacted site groundwater, and (2) to estimate the background concentration range for each target COPC.

After completing this process, investigators should compile the background analysis results and present conclusions for stakeholder review. If the choice of background wells and/or estimated background ranges are not acceptable to stakeholders, then the deficiencies should be identified, additional data should be collected if necessary, and the background analysis steps should be repeated as required.

After estimated background ranges acceptable to stakeholders have been established, the maximum detected concentrations of each target COPC should be compared to the estimated background ranges as a preliminary step. If the maximum concentrations of all COPCs are within the estimated background ranges, then no further background analysis or comparative statistical testing is required. If the maximum concentrations exceed the upper bound of the background range, but the exceedances are not high enough to be obviously inconsistent with the background population (based on site knowledge and professional judgment), then the statistical comparative methods described in Section 4 should be implemented. If the statistical comparison testing indicates that exceedances of an estimated background range are not statistically significant, then the background range estimate should be revised as necessary to reflect the maximum detected concentrations. (Note that this statement assumes that the site and background datasets are large enough to attain the desired test power. The desired power and adequacy of the datasets must be addressed within the DQO and DQA framework.)

However, if COPC concentrations in the site groundwater are considerably higher than the upper bound of the background range, and site knowledge and professional judgment indicate that the concentrations are consistent with a site-related release, then the time and resources required to implement the statistical comparison methods may not be justified. After completing this process, the results and conclusions should be documented in the investigation report (e.g., the Remedial Investigation Report) or appropriate section of the risk assessment report.

1.7.2 Statistical Comparative Methods

The statistical comparative methods (Section 4) are used to compare an individual measurement or the entire site-wide dataset representing chemical concentrations in potentially impacted groundwater to a dataset representing background conditions to determine whether differences between site and background measurements are statistically significant. Prior to implementing the statistical comparative tests (Section 4), the results of the background analysis (Section 3) should be used to separate the data for each COPC into two separate datasets, i.e., the site dataset and the background dataset. Use of adequately large representative datasets will ensure the reliability of statistical testing methods for site-wide comparisons. In the case of individual comparisons, however, verification resampling procedures are used to attain the desired reliability.

Stakeholder acceptance of the validity of the site and background datasets is particularly critical to ensure success of the statistical comparative methods. This acceptance can be attained by effectively communicating the background analysis results, including the geologic, hydrogeologic, and geochemical evidence used to identify and distinguish background data from site data. For the comparison to be valid, hydrogeologic and geochemical conditions at the background groundwater sampling locations must be similar to those in the site area; therefore, characteristics of potential background monitoring well locations should be thoroughly assessed in accordance with the guidance presented in Sections 2 and 3 before the data are used to represent background conditions for comparative statistical testing.

1.7.3 Background Analysis Documentation

Success of background analysis and comparative statistical testing is highly dependent on adequate and appropriate documentation of results and findings. The documentation should consist of at least the following components:

- Descriptions of the history and setting of the site and background groundwater sampling areas.
- Technically defensible rationale for selection of chemicals expected to occur in site groundwater because of natural and/or anthropogenic processes. This rationale should be based on professional judgment and a thorough assessment of regional and site-specific hydrogeologic, geochemical, and anthropogenic information.
- Descriptions of the site and background datasets.
- Rationale to demonstrate that the datasets are adequate for background analysis and statistical testing.
- Results of the background analysis and statistical testing, and documentation of each step in the process in sufficient detail to allow a comprehensive review.

Consistent with the Navy background policy (DON, 2000b), the background analysis documentation must address the following issues:

- Chemicals that may have been released at a site must be clearly identified to ensure that the Navy is focusing on remediating COPCs associated with the release.
- Chemicals detected at concentrations below the upper bound of the background range must not be included in the full baseline risk assessment. All chemicals screened out because of background considerations must be discussed and documented in the risk characterization section of the baseline risk assessment report.
- Cleanup levels must not be below the upper bound of background ranges.
- The methods presented in the background analysis report must be scientifically based, technically defensible, and cost-effective.

1.8 References

American Society for Testing and Materials (ASTM). 2003. *D6312-98 Standard Guide for Developing Appropriate Statistical Approaches for Ground-Water Detection Monitoring Programs.* West Conshohocken, PA.

California Environmental Protection Agency (Cal/EPA). 1997. *Selecting Inorganic Constituents as Chemicals of Potential Concern at Risk Assessments at Hazardous Waste Sites and Permitted Facilities. Final Policy.* Sacramento, CA: Cal/EPA Department of Toxic Substances Control (DTSC), Human and Ecological Risk Division. February.

Department of the Navy (DON). 1998. *Procedural Guidance for Statistically Analyzing Environmental Background Data.* San Diego, CA: DON SWDIV and EFA West. September.

Department of the Navy (DON). 1999a. *Handbook for Statistical Analysis of Environmental Background Data.* San Diego, CA: DON SWDIV and EFA West. July.

Department of the Navy (DON). 1999b. *Navy Policy for Conducting Ecological Risk Assessments.* Chief of Naval Operations Letter Ser. N453E/9U595355. Washington, DC. April 5.

Department of the Navy (DON). 2000a. Interim Final *Guide to Optimal Groundwater Monitoring*. Port Hueneme, CA: Naval Facilities Engineering Service Center. January.

Department of the Navy (DON). 2000b. *Navy Interim Final Policy on the Use of Background Chemical Levels.* Memo from Chief of Naval Operations (N45) (Rear Admiral L.C. Baucom, Director, Environmental Protection, Safety, and Occupational Health Division), to: Commander, Naval Facilities Engineering Command. 5090, Ser. N453E/OU595690. September 18.

Department of the Navy (DON). 2001a. *Department of the Navy Installation Restoration Manual, 2001 Update.* Draft. Alexandria, VA: Naval Facilities Engineering Command. August.

Department of the Navy (DON). 2001b. *Navy Policy for Conducting Human Health Risk Assessments Related to the Installation Restoration Program.* Chief of Naval Operations Letter Ser. N453E/1U595168. Washington, DC. February 12.

Department of the Navy (DON). 2002. *Guidance for Environmental Background Analysis, Volume I: Soil.* NFESC User's Guide UG-2049-ENV. Washington, DC: Naval Facilities Engineering Command. April.

Department of the Navy (DON). 2003. *Guidance for Environmental Background Analysis, Volume II: Sediment.* NFESC User's Guide UG-2054-ENV. Washington, DC: Naval Facilities Engineering Command. April.

Florida Department of Environmental Protection (FDEP). 1995. *Special Notice from the FDEP Contaminated Sediment Management Group.* Tallahassee, FL: FDEP Technical Review Section. April 5.

Florida Department of Environmental Regulation (FDER). 1988. *A Guide to the Interpretation of Metal Concentrations in Estuarine Sediments.* Eds. S.J. Schropp and H.L. Windom. Tallahassee, FL. April.

Gibbons, R.D. 1994. *Statistical Methods for Groundwater Monitoring.* New York: John Wiley & Sons.

Hawaii Department of Health (HDOH). 2002. *State of Hawaii Landfill Groundwater Monitoring Guidance Document.* Ver. 1.8. Solid and Hazardous Waste Branch. Honolulu, HI. September.

Michigan Department of Environmental Quality (MDEQ). 1991. *Michigan Background Soil Survey.* Formerly Michigan Department of Natural Resources. Prepared by Waste Management Division, Geotechnical Sampling and Support Unit. April.

Michigan Department of Environmental Quality (MDEQ). 1993. *MERA Operational Memorandum No. 15: Default Type A Cleanup Criteria.* From Alan J. Howard, Chief, MDEQ Environmental Response Division, to Environmental Response Division Staff. Lansing, MI. September 30.

Michigan Department of Environmental Quality (MDEQ). 1994. *Verification of Soil Remediation Guidance Document.* Revision 1. Lansing, MI.: DEQ Environmental Response and Waste Management Divisions. April.

Ohio Environmental Protection Agency (OEPA). 1995. *Technical Guidance Manual for Hydrogeologic Investigations and Ground Water Monitoring.* Columbus, OH: Division of Drinking and Ground Waters. February.

United States Environmental Protection Agency (U.S. EPA). 1988. *Guidance for Conducting Remedial Investigations and Feasibility Studies Under CERCLA.* Interim Final. EPA/540/G-89/004. Washington, DC: U.S. EPA Office of Emergency and Remedial Response. October.

United States Environmental Protection Agency (U.S. EPA). 1989a. *Interim Final RCRA Facility Investigation (RFI) Guidance, Vol. I of IV: Development of an RFI Work Plan and General Considerations for RCRA Facility Investigations.* EPA/530/SW-89-031. Washington, DC: U.S. EPA Office of Solid Waste, Waste Management Division. May.

United States Environmental Protection Agency (U.S. EPA). 1989b. *Risk Assessment Guidance for Superfund, Volume I, Human Health Evaluation Manual (Part A)*. Interim Final. EPA/540/1-89/002. Washington, DC: U.S. EPA Office of Emergency and Remedial Response. December.

United States Environmental Protection Agency (U.S. EPA). 1989c. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities.* Interim Final Guidance. EPA 530-SW-89-026. Washington, DC: U.S. EPA Office of Solid Waste, Waste Management Division. April.

United States Environmental Protection Agency (U.S. EPA). 1992a. *Guidance for Data Useability in Risk Assessment (Part A).* Final. Publication 9285.7-09A. PB92-963356. Washington, DC: U.S. EPA Office of Emergency and Remedial Response. April.

United States Environmental Protection Agency (U.S. EPA). 1992b. *Methods for Evaluating the Attainment of Cleanup Standards, Vol. 2: Ground Water.* EPA 230-R-92-014. Washington, DC: U.S. EPA Office of Policy, Planning, and Evaluation. July.

United States Environmental Protection Agency (U.S. EPA). 1992c. *Options for Addressing High Background Levels of Hazardous Substances at CERCLA Sites.* Draft issue paper. Washington, DC: U.S. EPA Office of Emergency and Remedial Response. June.

United States Environmental Protection Agency (U.S. EPA). 1992d. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities: Addendum to Interim Final Guidance.* Washington, DC: U.S. EPA Office of Solid Waste, Permits and State Programs Division. July.

United States Environmental Protection Agency (U.S. EPA). 1994a. *Evaluating and Identifying Contaminants of Concern for Human Health.* U.S. EPA Region 8 technical guidance. Prepared by U.S. EPA Region 8, Denver, CO.

United States Environmental Protection Agency (U.S. EPA). 1994b. *Statistical Methods for Evaluating the Attainment of Cleanup Standards, Vol. 3: Reference-Based Standards for Soils and Solid Media.* EPA 230-R-94-004. Washington, DC: U.S. EPA Office of Policy, Planning, and Evaluation. June.

United States Environmental Protection Agency (U.S. EPA). 1995. *Determination of Background Concentrations of Inorganics in Soils and Sediments at Hazardous Waste Sites.* U.S. EPA Engineering Forum Issue paper. EPA/540/S-96/500. Washington, DC: U.S. EPA Office of Solid Waste and Emergency Response. December.

United States Environmental Protection Agency (U.S. EPA). 1998. *Statistical Tests for Background Comparison at Hazardous Waste Sites.* Interim Draft Supplemental Guidance to RAGS. Prepared by U.S. EPA Region 4, Atlanta, GA.

United States Environmental Protection Agency (U.S. EPA). 1999. *Risk Updates, No. 5.* U.S. EPA Region 1, New England. Boston, MA. September.

Virginia Department of Environmental Quality (VDEQ). 2003. *Data Analysis Guidelines for Solid Waste Facilities.* Draft. VDEQ Office of Waste Programs. March.

2. DATA REVIEW, ASSESSMENT, AND EXPLORATORY ANALYSIS

The background analysis component of a groundwater investigation or monitoring program must be carefully planned and implemented to differentiate between background and site-related occurrences. Because some chemicals can pose risks to human health or the environment even at natural background levels, identification of these differences can facilitate decision-making and avoid potentially unnecessary remedial action. For example, some naturally occurring chemicals (e.g., arsenic) have characteristics that can complicate the remedy decision process:

- ❑ They have low risk-based criteria (i.e., their background levels may exceed risk-based criteria).
- ❑ They often are used in industrial, military, or commercial activities.
- ❑ They can occur naturally over wide concentration ranges.

Thorough data review is the key to effectively evaluating these characteristics and ensuring that background analyses are scientifically accurate and technically defensible for appropriate decision-making.

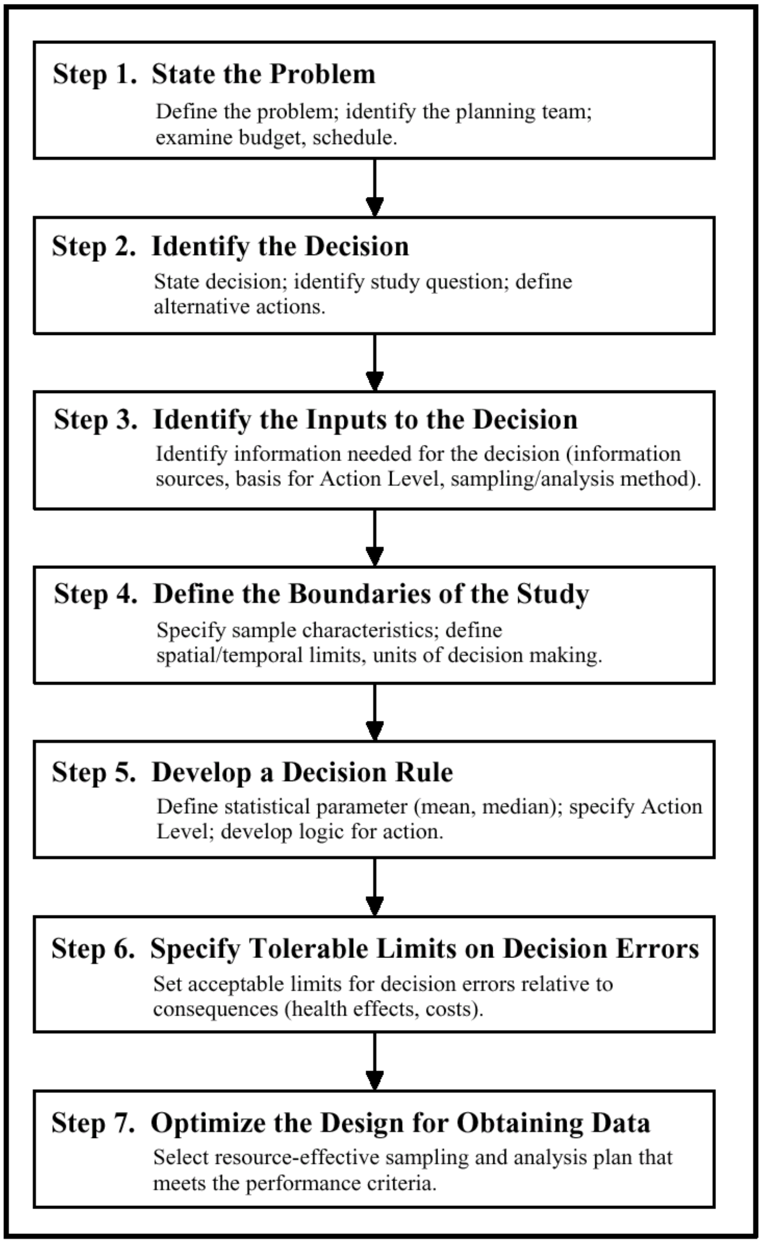
Groundwater sampling costs can be very substantial; therefore, review and assessment of existing data is particularly important for groundwater investigations. The procedures for data review and assessment described in this section should be implemented as part of any well-planned environmental site investigation, and are not intended to be repetitive. It is important to involve all stakeholders throughout the planning and execution phases of a background analysis. As noted in Section 1.3, evaluation of background conditions can greatly reduce the probability of decision error, and therefore can help to reassure stakeholders and regulators that the Navy has made the correct decisions regarding response actions, and has minimized the probability of decision errors.

2.1 Assessment of Local and Regional Physical and Chemical Characteristics

Background analysis should be based on a well-defined DQO decision-making framework. This framework is particularly useful when answering the following key questions:

- ❑ Is additional groundwater sampling and analysis necessary to provide the data required for background analysis?
- ❑ If additional sampling and analysis is necessary, what quality and quantity of samples and analyses are needed?

The available groundwater data must be reviewed and assessed to avoid the unnecessary effort and expense associated with collecting data not needed for the analysis. The data review and assessment process should be consistent with the most recent U.S. EPA *Guidance for the Data Quality Objectives Process* (U.S. EPA, 2000b). Specifically, the seven-step DQO planning process (Figure 2-1) should be used to determine the type, quantity, and quality of environmental data needed to support the decision-making process. Proper use of the DQO process will provide the scientific foundation for defensible decision-making by helping to assure that representative field samples are collected at appropriate locations and times, that appropriate techniques are used for graphic and statistical analysis of the resulting data, and that the graphic and statistical test results are properly interpreted. As part of the data review and assessment, the project team should verify that laboratory analytical data have been appropriately validated in accordance with the DQO requirements. When the DQO planning process is complete, supplementary groundwater samples, if necessary, should be collected according to the Sampling and Analysis Plan (SAP) or Groundwater Monitoring Plan (GWMP), as described in Section 2.1.6.



Step 1. State the Problem
Define the problem; identify the planning team; examine budget, schedule.

Step 2. Identify the Decision
State decision; identify study question; define alternative actions.

Step 3. Identify the Inputs to the Decision
Identify information needed for the decision (information sources, basis for Action Level, sampling/analysis method).

Step 4. Define the Boundaries of the Study
Specify sample characteristics; define spatial/temporal limits, units of decision making.

Step 5. Develop a Decision Rule
Define statistical parameter (mean, median); specify Action Level; develop logic for action.

Step 6. Specify Tolerable Limits on Decision Errors
Set acceptable limits for decision errors relative to consequences (health effects, costs).

Step 7. Optimize the Design for Obtaining Data
Select resource-effective sampling and analysis plan that meets the performance criteria.

FIGURE 2-1. Seven-step DQO planning process (Source: U.S. EPA, 2000b)

The data evaluation should be conducted according to the data quality assessment (DQA) process (U.S. EPA, 2000a), which consists of the following steps:

1. Review DQOs (output of each step of the DQO process) and sampling design
2. Conduct preliminary data review
3. Select the statistical test
4. Verify the assumptions
5. Draw conclusions from the data.

As part of the evaluation of available data, the project team must determine whether the data are sufficient to identify COPCs and then characterize the nature and extent of contamination at the site. As discussed in Section 2.1.3, the project team must also compare the data to risk-based screening criteria to determine whether it will be necessary to estimate background levels of chemicals identified as COPCs for the site. If the data are sufficient to characterize the nature and extent of contamination at the site and no COPCs are detected at concentrations above the screening criteria, then background analysis will not be required.

Groundwater background analysis requires information and data in addition to concentrations of the target chemicals in groundwater. For example, additional parameters such as pH and major ion concentrations also should be measured (see Section 2.1.4). These supplementary data then should be combined with the existing dataset.

2.1.1 Assemble Project Team

The first step in the data review and assessment phase of a background analysis is to assemble an appropriate project team. This team should include personnel with a broad range of expertise, such as:

- ❑ ***Geologists and Hydrogeologists.*** Geologists and hydrogeologists provide expertise in characterizing regional/site geologic and hydrogeologic characteristics, determining the adequacy of data coverage, and developing site and background area conceptual models.
- ❑ ***Statisticians.*** Statisticians provide expertise in designing background-specific sampling plans, selecting appropriate statistical tests and procedures, and interpreting the analytical results.
- ❑ ***Geochemists.*** Geochemists provide expertise in characterizing site-specific geochemical conditions, identifying parameters that must be investigated, developing appropriate sampling plans, identifying appropriate analytical methods, and evaluating site-specific elemental relationships.
- ❑ ***Toxicologists/Risk Assessors.*** Toxicologists/risk assessors provide expertise in characterizing site-specific risks to human health and the environment posed by background chemicals and COPCs, determining appropriate risk-based criteria, and documenting the results of the background analysis.

If it is not feasible to obtain input from experts in all of the disciplines listed above, at a minimum, personnel experienced in hydrogeology and statistics are essential to plan and implement a technically defensible groundwater background analysis strategy.

2.1.2 Review Site and Regional Operational History

A well-planned groundwater investigation requires information on the operational history of sites located above the groundwater body of concern, including their physical and chemical characteristics. This information is necessary to identify potential site-specific COPCs, to evaluate the need for further sampling, and, if necessary, to select appropriate sampling locations and analyses.

Operational History

Groundwater investigations and monitoring programs usually consider the operational history of a site and activities conducted upgradient of the site and potential background sampling locations by focusing on:

- ❑ Storage of hazardous substances or chemicals
- ❑ Transport, handling, or use of hazardous substances or chemicals
- ❑ Potential or known disposal or releases of hazardous substances or chemicals
- ❑ Potential or known disposal of municipal solid wastes
- ❑ Potential or known nonpoint sources of anthropogenic background chemicals (e.g., agricultural, commercial, industrial properties).

The primary sources of this information are the owners or operators of the site, and owners of upgradient properties. Previous environmental or regulatory investigations that have compiled information directly applicable to the evaluation of background conditions in the area will be of particular interest. Landowner records can provide key information concerning land use or activities at a site. Tax maps available from states or local municipalities show land ownership, and may therefore provide clues to the types of activities historically conducted in the area of concern. The following information regarding both the site and properties upgradient of the site should also be acquired if possible:

- History of industrial, agricultural, commercial, or residential uses
- Types of materials that were stored, handled, manufactured, or disposed of
- Locations of activities involving potential COPCs and potential releases
- Site-specific maps and "as-built" diagrams of current and former structures related to storage, handling, manufacture, or disposal of COPCs
- Aerial or land-based photographs that depict conditions at the site and surrounding area at various points in the past
- Locations of visible signs of potential COPC release (e.g., soil staining, discoloration, odor, and/or stressed vegetation)
- Data from monitoring and water supply wells in the site vicinity.

Other sources of information include: contaminant release incident report databases from federal, state, or local regulatory agencies; historical society records; title search agencies; master plans for larger facilities; tax maps, United States Geological Survey (USGS) maps, and other maps of the site and surrounding area; and employees or residents of neighboring properties. Several guidance documents present procedures for assessment of the operational history of an investigation site. The ASTM presents protocols for Phase I (ASTM, 2003e) and Phase II (ASTM, 2003f) site assessments.

Physical Setting

The physical characteristics of the site and surrounding area, such as topography, geography, and climate influence groundwater conditions, and therefore affect the fate and transport of both background chemicals and COPCs in groundwater. These characteristics should therefore be evaluated carefully during the data assessment phase.

Topographic and Geographic Information

Accurate maps of the investigation sites and the surrounding region are essential for background analysis. Two types of maps are required: a small-scale regional map for placing the site area in a regional context, and a detailed large-scale site map to plot prominent surface features, potential contaminant sources, impacted areas, and background areas.

Maps should provide topographic and geographic representations of several types of information needed for the investigation and background analysis:

- Terrain conditions, surface drainage patterns, and the types of rock and soil that occur within the watershed (to identify potential chemical sources and transport pathways)
- Locations and well construction data for existing monitoring wells and other groundwater sampling locations (to evaluate the spatial distribution of detected chemicals)
- Known or potential disposal or release locations in the site vicinity and upgradient area
- Other pertinent information, including property boundaries, right of ways, and utility corridors.

Depending on the location of the site, the following maps, charts, and aerial photographs may be available:

- ***USGS Quadrangle ("Quad") Maps.*** Quads are regional-scale maps that show regional topography, water bodies, landforms, streets, and general land use. Quads are available for most regions of the United States at different scales, most typically ranging from 1:24,000 (1 inch = 2,000 feet) to 1:250,000 (1 inch = 20,833 feet). Other maps at larger scales are available for limited areas. A list of available maps for the United States can be located at the USGS "Online Map Lists" Web page (mac.usgs.gov/maplists.html). The availability of large-scale maps will vary. Digital versions of many USGS quads are available in tiled format on the TerraServer website, a joint venture of the USGS and Microsoft Corporation (www.terraserver.microsoft.com).

- ❑ ***Photogrammetric Aerial Survey Maps (i.e., Aerial Photographs).*** Aerial photographs may show features not illustrated on USGS quads (e.g., stressed vegetation). Some photogrammetric surveys take photographs in pairs from slightly offset angles. When viewed stereoscopically, the photograph pairs produce three-dimensional images that clearly show topography and other site features. High-resolution 9 × 9 inch photographs are available from the National Aerial Photography Program (NAPP). NAPP is an interagency effort coordinated by the USGS to acquire new aerial photographic coverage of the conterminous United States every 5 to 7 years (approximately 90 percent coverage currently is available). The NAPP photographs are at a scale of 1:40,000; each covers the equivalent of one-quarter of a standard USGS 7.5-minute quad map. The NAPP images are also available in digital format as digital orthophoto quadrangles (DOQs), at a resolution of 1 meter per pixel. The DOQs are orthorectified so they can be read and measured as maps. Tiled versions of downsampled DOQs (linked to digitized USGS quad maps) can be freely viewed and downloaded from the TerraServer website (www.terraserver.microsoft.com).

- ❑ ***Historical/Archaeological Maps and Photographs.*** Historical and archaeological maps and photographs can identify land areas and water bodies that have been affected by natural processes or human activities.

Site-specific information may be difficult to obtain for some Naval facilities. As with many United States Department of Defense (DoD) facilities, aerial photography may not be allowed in areas or regions that are considered sensitive, and maps for sensitive areas may be classified. However, information and maps may be available for neighboring sites, or non-DoD areas.

A Geographic Information System (GIS) is useful for compiling and integrating groundwater site data. GIS is a highly effective tool for interpreting and co-analyzing different datasets that can be effectively analyzed only in a spatial format. For example, changes in the extent of contamination can be detected by overlaying locations of contamination detected during different site investigations. In addition, previously unidentified contaminant sources can be revealed by overlaying contaminant concentration data on historical aerial photographs. The older photographs may reveal relationships between clusters of detected contaminants and potential contaminant sources that no longer show up on USGS quads.

Climatological Information and Data

Aquifer recharge rates depend on precipitation rates, temperature conditions, and soil characteristics in the recharge area (i.e., the area in which groundwater is replenished by water that infiltrates and moves downward into the saturated zone). Concentrations of both background chemicals and COPCs may show seasonal trends associated with increases in recharge rates during periods of heavy rainfall or melting of snow packs. In many cases, increases in chemical concentrations are observed after periods of heavy precipitation due to increased leaching from vadose zone soils; however, decreased concentrations can also be associated with increased recharge rates. In addition to temperature and precipitation rates, data regarding evapotranspiration rates and wind conditions can also be useful. Climatological information and data are readily available from the following sources.

- ❑ ***The National Climate Data Center (NCDC).*** The NCDC (www.ncdc.noaa.gov) is the major repository for climate data in the United States. Time series of weather station data are available from the Web site, and maps of a particular variable can be plotted. A complete list of all surface data is available, including access to the raw daily data.

- ❑ ***National Oceanic and Atmospheric Administration (NOAA) Regional Climate***

Centers. There are six regional climate centers in the United States (www.wrcc.dri.edu/rcc.html). Each center provides climate information relevant to that particular part of the country. The regional climate centers are more likely to have esoteric information not available from the NCDC.

- ***State Climatologists.*** Almost all states in the United States employ a state climatologist, who has access to local climate records, especially those related to significant weather events. Some states have Web sites for easy access to local information and data.
- ***National Weather Service (NWS).*** The NWS (www.nws.noaa.gov) provides links to the Web pages of many of their local offices.

Regional and Site-Specific Geologic and Hydrogeologic Information

Naturally occurring chemicals in groundwater are attributable to the dissolution of minerals in rock and soil. Therefore, the distribution of these background chemicals in groundwater depends on the composition of the rock and soil that occur in the aquifer matrix and recharge area (see Section 2.1.4). Characteristics of the rock and soil formations that occur in the site and upgradient areas should be carefully evaluated. Accurate field geological observations and descriptions are essential to plan a technically defensible background analysis strategy.

A qualified geologist, soil scientist, or hydrogeologist should evaluate surface and subsurface soil and rock samples to assess the mineralogy of the aquifer matrix and upgradient recharge areas. In addition, geotechnical testing and chemical analysis of aquifer soil or rock samples are strongly recommended in order to quantify aquifer characteristics. These tests and analyses should be conducted according to ASTM geotechnical testing standards and U.S. EPA analytical methods (see Section 2.1.4).

Hydrogeologic information and data are essential for groundwater monitoring and background analysis. Data regarding hydrogeologic parameters (e.g., hydraulic gradient, hydraulic conductivity) are critical for identifying locations and depth intervals at which site-related groundwater contamination may be encountered, and to identify locations and depths likely to be free of groundwater contamination (i.e., background sampling locations).

Geologic and hydrogeologic maps, cross sections, and reports are available from the following sources:

- ***USGS and State Geological Survey Offices*** may be able to provide the geologic maps and cross sections necessary to assess the surface and subsurface distribution of rock and soil formations. USGS publications are available from libraries designated as federal document depositories and the U.S. Government Printing Office.
- ***The USGS Water Resources Division (WRD)*** is the leading source of hydrogeologic information in the United States. The WRD collects data representing stream flowrates, surface water quality, groundwater levels, and groundwater quality. The USGS maintains groundwater information Web pages (water.usgs.gov/ogw/index.html) containing a groundwater site inventory, groundwater level data, and groundwater quality data. The groundwater site inventory (waterdata.usgs.gov/nwis/inventory) contains more than 850,000 records of wells, springs, test holes, tunnels, drains, and excavations in the United States. Available site descriptive information includes well location information (i.e., latitude and longitude, well depth, site use, water use, and aquifer).
- ***Other Federal agencies*** that conduct hydrogeologic and related studies include the U.S. Army Corps of Engineers, Bureau of Land Management, Bureau of Reclamation, Soil Conservation Service, U.S. EPA, Nuclear Regulatory Agency, and Department of Energy.
- ***The United States Department of Agriculture (USDA), Natural Resources***

Conservation Service (NRCS), National Cooperative Soil Survey, and state soil offices can provide soil survey maps and reports indicating major soil associations, soil families, and soil series. Soil surveys can be useful to identify the geochemical characteristics of soils in the site and upgradient areas.

- ***The National Ground Water Association*** (www.ngwa.org) offers *Ground Water Online,* a bibliographic database of groundwater literature with more than 89,000 citations (available to NGWA members) (www.ngwa.org/gwonline/gwol.html). NGWA also operates the National Ground Water Information Center, the largest non-governmental clearinghouse on groundwater science and well technology in the world.

- ***Professional Journals*** that present articles on hydrogeologic research and groundwater include:

  Bulletin, International Association of Scientific Hydrology
  Ground Water
  Ground Water Monitoring and Remediation
  Journal American Water Works Association
  Journal of Applied Hydrogeology
  Journal of Contaminant Hydrology
  Journal of Environmental Forensics
  Journal of Hydrology
  Memoirs, International Association of Hydrogeologists
  Transactions, American Society of Civil Engineers
  Water Resources Bulletin
  Water Resources Research.

- ***Dissertations.*** Unpublished theses at the master's or Ph.D. level are often very useful sources of information on geology and groundwater. Many libraries have an index of "Dissertation Abstracts," that lists theses completed at accredited universities across the country. Libraries and universities located in the same geographical area as the investigation site are most likely to have pertinent theses. If the thesis is on file at a remote location, it often can be photocopied and mailed to the investigator. Theses are also available from dissertation services that can be accessed via the Internet.

A summary of potential sources of operational, physical, and geochemical information and data for assessing a groundwater investigation site and the surrounding area is presented in Box 2-1.

BOX 2-1. Assembling operational, physical, and geological information and data

Assessing the operational history, physical setting, and geological characteristics of a site and the surrounding area often involves assembling and reviewing a large amount of information and data.

Operational history. Property owners and operators of sites and facilities within the watershed are valuable sources of information. Results of previous environmental investigations at the site and neighboring or related sites are also of primary interest. In some locations, state and local governments have gathered information on past land uses, which may have included the disposal of hazardous wastes. For example, many states keep historical records of landfills and other sites that may pose threats to groundwater. Private industries also may have pertinent operational records; however, it may be difficult to obtain this information due to legal concerns.

Physical setting. Topographic and geographic information is acquired primarily from maps and aerial photographs. Potential sources include the USGS and the NAPP. Climatological information is available from the NCDC, NOAA, NWS, and state climatologists.

Geological and hydrogeological setting. Information and data including geologic and soil survey reports, maps, and cross sections are available from the USGS and state geological and soil survey offices. Geologic and hydrogeologic information is available from sources including the USGS, NGWA, and various professional journals. Unpublished master's and Ph.D. theses are available from state and private university libraries and dissertation services and can be valuable sources of geologic and hydrogeologic information.

Target Chemicals

The list of target chemicals should include chemicals used at the site during its history, potential chemical breakdown products, potential site-specific background chemicals, chemicals required by regulatory criteria, and other parameters that may be useful for background analysis (see Box 2-2). The list of target chemicals should be developed after evaluating the following information and data:

- Operational history of the investigation site and potential upgradient contaminant sources, particularly information regarding chemicals potentially used or stored at the sites
- Soil and groundwater chemical data and risk assessment results from previous investigations at the site and background sites
- Geologic, hydrogeologic, and anthropogenic characteristics of the site and background sites
- Regulatory requirements for groundwater monitoring at the site.

Identifying Background Groundwater Sampling Areas

Existing or prospective background monitoring well locations should be evaluated to identify the locations most likely to be free of groundwater contamination associated with both on-site and off-site sources. Areas upgradient of the actual, suspected, or potential contaminant sources at the investigation site are typically considered the best locations for background monitoring wells. However, depending on factors such as hydrodynamic dispersion and contaminant transport velocity (see Section 2.1.4), side-gradient or even downgradient locations may also be suitable for installation of background monitoring wells. Side-gradient locations may be suitable if they are outside the zone that could be impacted by lateral dispersion of the COPCs. Downgradient locations may be suitable if the distance from the source, time since the release, and contaminant transport velocities are such that the COPCs could not have reached the prospective downgradient monitoring wells. Groundwater modeling programs can be useful to assess the extent of groundwater impact, and select unimpacted locations for background monitoring wells (see Section 2.1.4).

BOX 2-2. Identifying target analytes

COPCs – Select based on Navy operational history, data from previous investigations, and regulatory requirements.

Examples:

Arsenic, chromium, and mercury (ship maintenance and building, aerial fallout, sewage effluent, fungicides)

Copper, lead, tin, and zinc (old paint, marine antifoulants)

Antimony, copper, and lead (firing ranges)

Arsenic and pesticides (wood preservatives, pesticide equipment rinsing, agricultural runoff)

Chromium (metal plating, alloys)

Lead (leaded gasoline, battery disposal)

Additional parameters useful for groundwater background analysis – Select based on geology, hydrogeology, and geochemistry (see Section 2.1.4).

Examples:

Cation concentrations (e.g., magnesium, calcium, sodium, potassium)

Anion concentrations (e.g., sulfate, chloride, carbonate, bicarbonate)

Salinity or total dissolved solids (TDS)

pH

Redox potential (Eh)

Whenever possible, background monitoring wells should be screened in the same aquifer formation and at the same depths as the on-site monitoring wells used to characterize the actual or suspected plume of contaminated groundwater at the investigation site. However, with stakeholder concurrence, it may be possible to select background monitoring well locations in aquifer formations that are spatially distinct from the site area. Such areas must have hydrogeologic, geochemical, and anthropogenic characteristics similar to those of the site area.

The process used to select appropriate locations for installing background monitoring wells must be technically defensible and acceptable to stakeholders, and should include the following steps:

- ***Identify pertinent operations and historical uses of the properties surrounding the investigation site.*** Current operations, past uses, and anthropogenic characteristics of locations upgradient of the potential background sampling locations should be examined to evaluate known or potential contaminant sources, locations of past releases, or locations of contaminant disposal.
- ***Compare physical, geologic, and hydrogeologic characteristics.*** The topographic, geographic, geologic, and hydrogeologic characteristics of potential background sampling areas should be compared to the corresponding characteristics of the investigation site and the actual or suspected groundwater contaminant plume area.
- ***Formulate a conceptual site model (CSM) representing the site and potential background sampling locations.*** A CSM is essential for identifying potential background groundwater sampling locations. Information on known or potential contaminant sources should be integrated with the physical, geologic, hydrogeologic, and geochemical characteristics of the site and surrounding area to develop a CSM. The CSM should address the following elements:
    - Locations of all on-site and off-site potential contaminant sources, both current and historic
    - Nature and extent of known soil and groundwater contamination
    - Geology (e.g., soil types, rock types)
    - Hydrogeology (e.g., depth to the water table, aquifer thickness, hydraulic gradient, hydraulic conductivity, transmissivity, flowpaths, locations of springs or seeps, existing production well locations)
    - Geochemical conditions (i.e., chemical characteristics of the aquifer matrix and groundwater)
    - Contaminant transport pathways
    - Existing monitoring well locations and groundwater data
    - Any other factors relevant to the understanding of the site.

In addition to its value for selecting background sampling locations, the CSM is an integral component of most groundwater site investigations and monitoring programs, and is very useful for communicating with stakeholders.

2.1.3 Conduct Risk Screening

After verifying that the analytical data have been collected and validated according to the DQO requirements, the detected chemical concentrations should be compared to appropriate screening criteria (e.g., U.S. EPA [2002] MCLs). Risk-based screening criteria should be identified in accordance with appropriate guidance, including U.S. EPA human health and ecological risk assessment guidance (1989, 1991, 1997a) and the Navy tiered ecological evaluation process (DON, 1999b). If the maximum detected concentrations of a chemical are equal to or less than its corresponding screening level, and the nature and extent of impacted groundwater have been adequately characterized, then no background analysis is necessary for that particular chemical.

For screening soil or sediment with maximum chemical concentrations above risk-based criteria, U.S. EPA RAGS protocols and the Navy risk assessment policy (DON, 1999b, 2001) allow use of representative exposure concentrations (e.g., the upper confidence limit of the mean concentration). However, because conditions under which receptors may be exposed to chemicals in groundwater are not the same as those for soil and sediment, representative exposure concentrations for groundwater are usually conservatively defined as the maximum detected concentrations.

2.1.4 Assess Geology, Hydrogeology, and Geochemistry

Information and data regarding the geologic, hydrogeologic, and geochemical characteristics of the investigation site, upgradient area, and potential background sites are essential for understanding the fate and transport of COPCs, selecting locations for background groundwater sampling, and providing the evidence needed to justify the choice of background well locations. A thorough evaluation of hydrogeology and geochemistry is particularly critical for groundwater monitoring and background analysis due to the high cost of monitoring well installation and groundwater sampling. The following sections present an overview of the primary geochemical and hydrogeological principles relevant to groundwater background analysis.

Geologic and hydrogeologic characteristics such as aquifer structure, porosity, hydraulic gradient, hydraulic conductivity, transmissivity, and hydrodynamic dispersion affect the fate and transport of both background chemicals and COPCs in groundwater. A thorough assessment of geologic and hydrogeologic information and data is therefore required to develop a conceptual site model, and identify appropriate locations for both site characterization and background monitoring wells.

The naturally occurring chemicals that occur in soils, sediments, and groundwater originate in the rocks that form the earth's crust. Therefore, the types and concentrations of background chemicals depend primarily on the composition of the parent rocks and their component minerals. Once rocks have been disintegrated by physical and chemical weathering, geochemical processes transport and redistribute the chemicals, particularly metals, in various geologic environments. A thorough understanding of the geochemistry of the local environment is essential to understand the origins of natural background chemicals in groundwater, and is particularly important to differentiate between background chemical conditions in groundwater and conditions that may be associated with a chemical release.

General Hydrogeologic Principals

Groundwater background analysis requires a comprehensive understanding of the basic hydrogeologic principals involved in groundwater recharge and flow, and the physical characteristics of groundwater systems.

Figure 2-2 illustrates the continuous system known as the hydrologic cycle that transfers water to and from the atmosphere, ocean, and land. Water enters the atmosphere by evaporation from the ocean, other bodies of water, and soil surfaces, and by transpiration through plants. These processes are collectively referred to as evapotranspiration. Water can also be transmitted directly from solid form to gaseous form through sublimation. As water vapor rises in the atmosphere, it cools with increased elevation and is transferred back to liquid form through the process of condensation. After accumulating in clouds, the liquid water falls to the earth as precipitation.

Water from precipitation infiltrates the ground surface and collects in surface reservoirs such as oceans, lakes, rivers, and glaciers. After percolating down through the unsaturated (vadose) zone, water collects in the saturated zone below the water table (the upper limit of the saturated zone). After entering the saturated zone, groundwater flows in a downslope direction until it is discharged to surface water bodies or at the ground surface.

It should be noted that both Figure 2-2 and the explanation of the hydrologic cycle presented above provide a simplified description of the process.

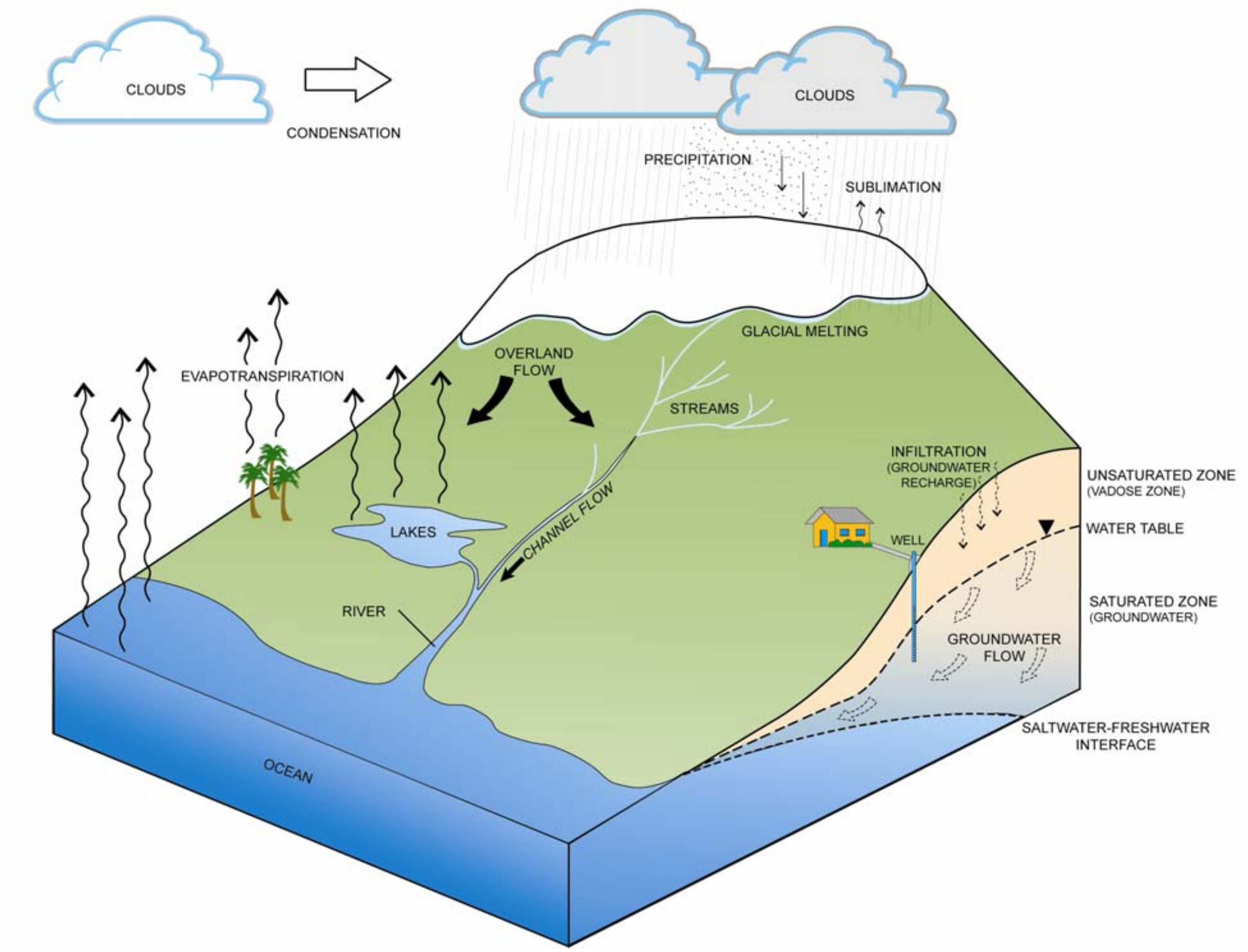


FIGURE 2-2. The hydrologic cycle

Specific conditions at a site are influenced by many natural and anthropogenic events.

Aquifers and Aquitards

Two primary classification terms have been used to describe the ability of a geologic formation to transmit groundwater: *aquifer* and *aquitard*. Aquifers are saturated, permeable geologic units that can transfer significant quantities of water under normal hydraulic conditions. Aquitards are geologic units that cannot transfer significant quantities of water under normal hydraulic conditions. The terms are qualitative to allow relative comparisons between geologic units, regardless of quantitative rates of flow. The term *aquiclude* is occasionally used to describe low permeability geologic units, but is seldom used in applied hydrogeology.

Aquifer Types

Aquifers are classified into three fundamental types: *unconfined, confined*, and *perched*. Figure 2-3 illustrates the three types of aquifers. The hydrogeology of a typical site may include a combination of multiple aquifer types.

Unconfined aquifers, also referred to as water-table or phreatic aquifers, exist in areas where all overlying geologic units have relatively high permeabilities. An unconfined aquifer is typically recharged by direct infiltration through the overlying unsaturated (vadose) zone.

Confined aquifers are overlain by relatively impermeable geologic units, and contain groundwater under pressure significantly greater than that of the atmosphere. Confined aquifers are recharged in

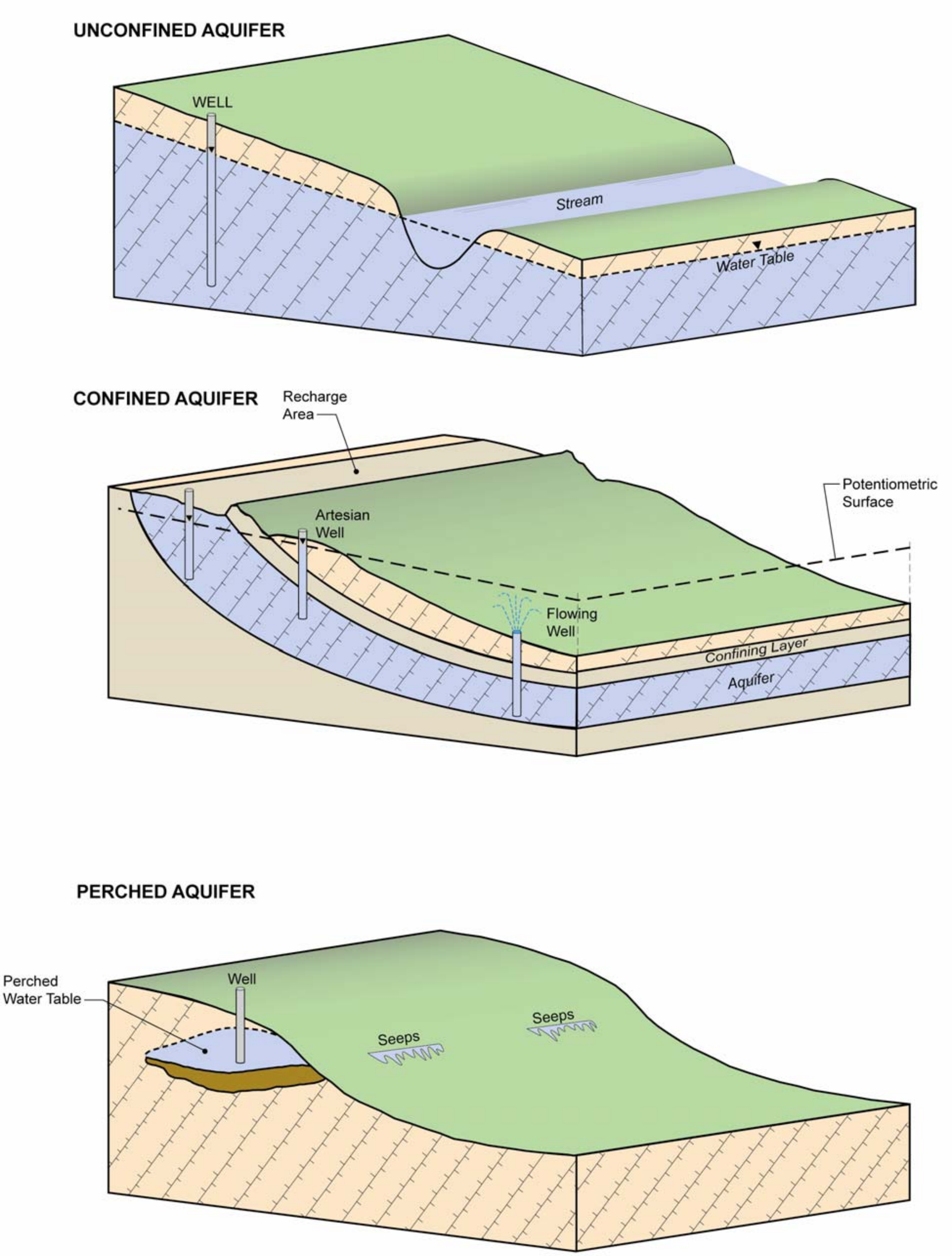


FIGURE 2-3. Aquifer types

areas where the formation crops out upgradient of the confined zone, or through leaks in the confining unit. When the pressure in a confined aquifer is sufficient to raise the potentiometric surface above the upper aquifer boundary, the aquifer is said to be artesian. If the water level in a well screened in the unit stabilizes at a level above the upper boundary of the confined aquifer, the well is known as an *artesian well*. If the water level rises above the level of the ground surface, the well is known as a *flowing artesian well*.

Perched aquifers are permeable geologic units underlain by impermeable formations that trap water in localized areas and prevent it from infiltrating downward. If the perched water table reaches the surface, water may be discharged in the form of seeps.

Groundwater Flow and Gradient

The magnitude of the hydraulic gradient (i.e., the vertical change in hydraulic head that occurs over a horizontal distance), in conjunction with the hydraulic conductivity, controls the rate of groundwater flow. The direction of the hydraulic gradient controls the direction of groundwater flow within the formation. Hydraulic head is the elevation at which the water level stabilizes after a well or piezometer has been allowed to equilibrate, and corresponds to the level of the potentiometric surface at the well location. A minimum of three groundwater monitoring wells or piezometers is required to establish the direction and magnitude of the hydraulic gradient. Figure 2-4 illustrates this relationship and the basis for Darcy's Law.

The principals of groundwater flow mechanics were investigated in experiments conducted by Henry Darcy in 1856. Darcy's work resulted in a law describing groundwater flow through a porous medium, known as Darcy's Law:

$$Q = \frac{-KA(dh)}{(dl)}$$

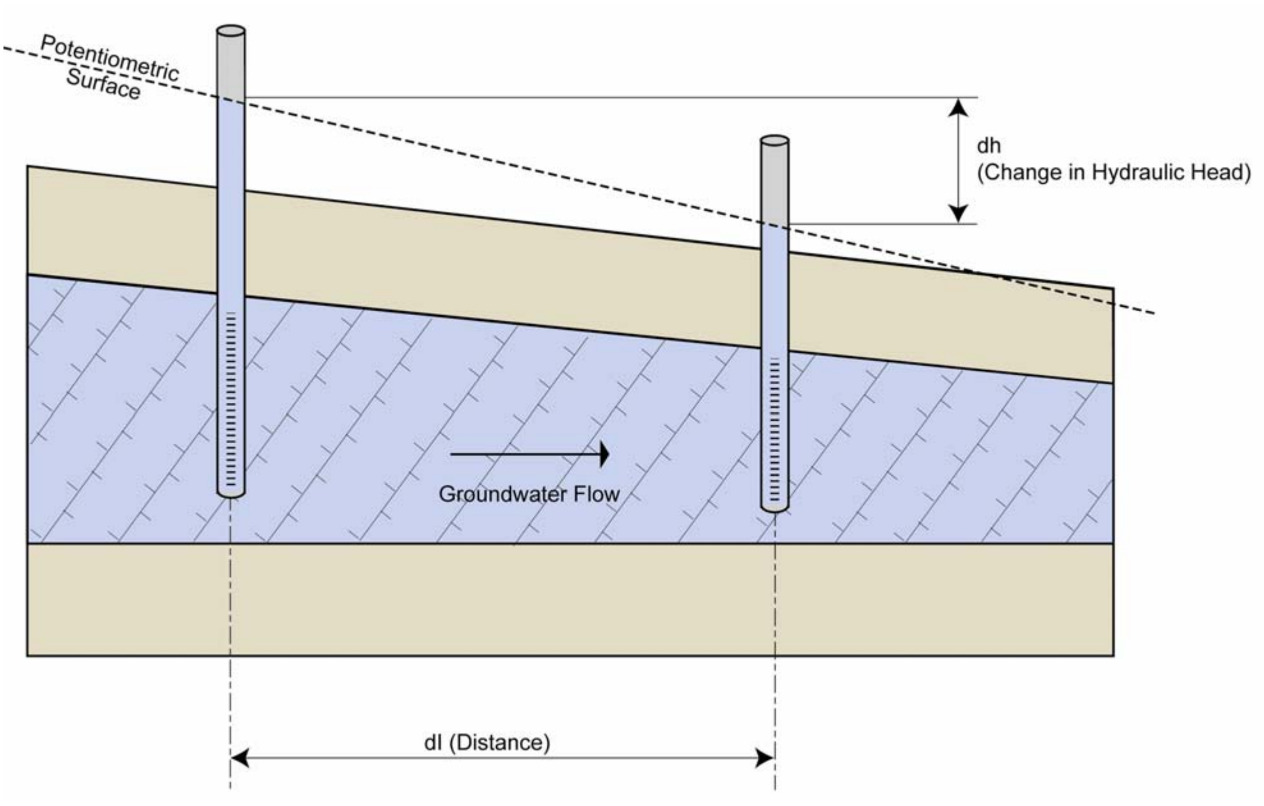


FIGURE 2-4. Hydraulic gradient

where Q is the volume of water passed in unit time, A is the cross-sectional area of the formation, dh is the change in hydraulic head, dl is the distance between reference points, and K is the hydraulic conductivity (described below). The ratio dh/dl is the hydraulic gradient.

Porosity

Groundwater moves through void spaces in subsurface rock, soil, or sediment. The ratio of the volume of the void spaces to the total volume of rock or sediment is called porosity, n. Porosity can be classified into two categories: primary porosity refers to voids created at the time of deposition (e.g., voids between sand particles); secondary porosity refers to voids developed after deposition (e.g., fractures or solution channels). Effective porosity, n_e, represents the total porosity available for fluid flow. As described below, n_e is a key parameter required to evaluate *seepage velocity* (i.e., the average linear velocity of groundwater flow between two points) (Fetter, 1994).

Hydraulic Conductivity, Seepage Velocity, and Transmissivity

The hydraulic conductivity of a porous medium, K, is a property of the medium that represents the quantity of water that will flow through a unit of cross sectional area per unit of time under a hydraulic gradient of 1. Hydraulic conductivity is quantified in terms of distance per unit of time. Coarse-grained porous media such as sand and gravel generally exhibit much higher K values than fine-grained media such as clay.

The seepage velocity (or average linear velocity, V) is the average velocity at which groundwater moves between two points (Fetter, 1994). Seepage velocity represents the velocity with which chemicals would be transported by groundwater in the absence of other factors such as hydrodynamic dispersion, and reversible sorption to the aquifer media (i.e., retardation). As indicated by the following modification of Darcy's Law, seepage velocity is inversely proportional to the effective porosity, n_e.

$$V = \frac{Q}{(n_e A)} = \frac{-K(dh)}{n_e(dl)}$$

Transmissivity, T, is a measure of the quantity of water that can be transmitted horizontally within the aquifer through an area representing a unit width and a depth equal to the full saturated thickness of the aquifer, under a hydraulic gradient of 1 (Fetter, 1994). Transmissivity is the product of the hydraulic conductivity, K, and the saturated thickness, b:

$$T = bK$$

Groundwater Modeling

Groundwater modeling can provide valuable information for designing groundwater monitoring programs and evaluating background conditions. Modeling can be useful to evaluate flow velocities, contaminant transport velocities, plume center-of-mass movement, plume spreading, and plume degradation. Groundwater contaminant modeling can range in complexity from simple analytical calculations, to semianalytical techniques, to multiphase stochastic numerical models that account for heterogeneous geology, hydrodynamic dispersion, contaminant mass-loss functions, and thermodynamic chemical equilibria.

Flow velocity (or transport-time) modeling can provide estimates of groundwater flow (seepage) velocity. Data required to estimate flow velocity include an estimate of the groundwater gradient, hydraulic conductivity, and effective porosity. Reversible sorption to aquifer matrix materials (e.g., organic matter and clays) and other factors such as chemical transformation and biological degradation, will retard the transport of nonconservative dissolved chemicals relative to the groundwater flow velocity. Plumes of different contaminants at the same site often move at different velocities, or a plume may separate into different constituents as the chemicals are retarded or degraded at different rates.

Numerous commercial groundwater modeling programs are available to predict contaminant fate and transport in the groundwater environment. Contaminant transport velocities, lateral and longitudinal spreading of a plume due to hydrodynamic dispersion, and attenuation due to biodegradation or other chemical reactions can all be approximated by fate and transport models. Under good calibration, advanced models can accurately predict contaminant plume movement in three dimensions. This information can be extremely useful to select locations for both site characterization monitoring wells and background monitoring wells, and to determine optimal sampling frequencies for a monitoring program.

Groundwater fate and transport modeling will be most successful if the site is geologically homogenous and adequately characterized with respect to chemical concentrations, hydrogeologic properties, and geochemical characteristics. In general, as the geologic complexity of a site increases, the cost of modeling increases, and modeling accuracy decreases. Modeling programs that combine groundwater flow with chemical fate and transport require a wide range of physical and chemical input parameters. The overall accuracy of the model will depend directly on the quality of the data used to characterize the input parameters. For example, the accuracy of modeling predictions is particularly sensitive to hydraulic conductivity, which can vary dramatically across different formations. Although there are numerous ways to characterize hydraulic conductivity, each has its own limitations (Freeze and Cherry, 1979; Fetter, 1994).

Natural and Anthropogenic Sources of Background Chemicals in Groundwater

Most of the solutes present in natural groundwater, particularly the metals, are derived from the dissolution of minerals in rock, soil, and sediment. The chemical characteristics of natural groundwater therefore are closely related to the composition of the minerals the water has contacted along the flowpath. However, because many geochemical parameters influence and control the distribution of chemicals between the solid and aqueous phase, the relationship between the mineralogy of aquifer solids and chemical concentrations in groundwater can be very complex. Relationships between the lithology and mineralogy of geological formations and the chemical characteristics of the associated groundwater, as well as details of the many geochemical factors that control concentrations of metals and other naturally occurring chemicals in groundwater, are described in the aqueous geochemistry literature. The literature can be particularly useful to identify the natural conditions that can lead to elevated concentrations of certain naturally occurring chemicals in groundwater. *Study and Interpretation of the Chemical Characteristics of Natural Water* (Hem, 1992) describes natural sources of metals and other chemical constituents commonly detected in groundwater, and the geochemical parameters that typically control their concentrations (e.g., solubility controls, pH effects, redox processes). *Geochemistry of Natural Waters* (Drever, 1997) and *Aqueous Environmental Chemistry* (Langmuir, 1997) also describe the processes that control the composition of groundwater and other natural waters. *Chemical Fate and Transport in the Environment* (Hemond and Fechner-Levy, 1999) provides a quantitative treatment of fate and transport processes in major environmental media. *Adsorption of Metals by Geomedia* (Jenne, 1998) features a wide range of reviews of the status of metal adsorption in geological media. *Contaminant Hydrogeology* (Fetter, 1999) presents a detailed account of the role of inorganic and organic chemicals in groundwater, mass transport in saturated media, transformation and attenuation of solutes, vadose zone monitoring, and remediation.

Anthropogenic background chemicals are associated with nonpoint sources. Point sources are confined, discreet, localized releases in areas typically an acre or less in size. An example of a point source is a one-time release from a UST. Conversely, nonpoint sources are dispersed or diffused over broad areas, and are typically active over long time periods. An example of a nonpoint source is prolonged pesticide use over a large agricultural area. Anthropogenic background chemicals are derived from agricultural, industrial, residential, and urban sources. Agricultural sources include pesticides, herbicides, fertilizers, and livestock waste. Industrial sources are extremely varied, and include storage, use, and transport of many different products and chemicals. Residential sources include septic tank systems and the storage or disposal of household chemicals. Urban sources are mainly associated with runoff of surface and transportation-related pollutants.

Geochemical and Geotechnical Characteristics

Groundwater geochemistry and the geotechnical characteristics of the aquifer control the concentrations and movement (i.e., fate and transport) of chemicals in the dissolved phase. Factors affecting the fate and transport of the chemicals identified as COPCs for a groundwater site should be evaluated to determine whether COPCs are likely to impact a proposed background monitoring well location, and to understand the factors that contribute to elevated concentrations of natural background chemicals in groundwater.

Many reactions and interactions between chemicals in the solid, liquid, and aqueous (i.e., dissolved) phases occur in the groundwater system. In general, if a chemical occurs in a particular phase at concentrations not in thermodynamic equilibrium with the other phases, then the environment will respond to restore equilibrium. For example, if the aquifer matrix contains relatively high concentrations of a naturally occurring or anthropogenic chemical that is soluble under the pH, Eh, and other geochemical conditions that exist within the aquifer, then the chemical will tend to partition from the solid phase into the aqueous phase until equilibrium is reached. Therefore, investigation of the natural geochemical characteristics of the aquifer and their effect on COPCs within the groundwater system is essential to understand both background conditions and the possible effects of a chemical release on groundwater. The geochemical and geotechnical parameters described in the following subsections are particularly important for understanding the fate and transport or chemicals in groundwater, and comparing the characteristics of different groundwater regimes.

Geochemical and Geotechnical Characteristics of the Aquifer Matrix

The geotechnical and geochemical parameters of the aquifer matrix materials control chemical concentrations and migration in the groundwater environment. Measurement of the geotechnical and geochemical parameters listed in Table 2-1 is recommended for subsurface soil samples (e.g., samples collected from monitoring well borings). Some of the most important parameters are described below.

Bulk density represents the weight of a soil sample divided by its volume. Bulk density is a useful property for comparing and classifying samples of soil, sediment, or rock collected during drilling for monitoring well installation.

Particle size distribution indicates the percentages of particles that fall within predefined size ranges (e.g., clay, silt, sand, and gravel). Because adsorption is a surface phenomenon, the rate and extent of adsorption increase as the surface area of the sorptive medium increases. Surface area increases as mineral grains are divided into smaller particles. Therefore, fine-grained soil and sediment particles, including both clay minerals and organic matter, have much greater sorption capacity than larger particles such as sand or gravel.

TABLE 2-1. Geotechnical and geochemical testing parameters, aquifer matrix materials

Parameter	Testing Method[a]	Method Source
Bulk Density	ASTM D2937	ASTM (2003b)
Particle Size Distribution	ASTM D422	ASTM (2003c)
Permeability	SW-846 9100	U.S. EPA (1997b)
Porosity	ASTM D854	ASTM (2003d)
Redox Potential (Eh)	ASTM D1498	ASTM (2003a)
Soil pH	SW-846 9045	U.S. EPA (1997b)
Total Organic Carbon (TOC)	SW-846 9060	U.S. EPA (1997b)
Cation Exchange Capacity (CEC)	SW-846 9081	U.S. EPA (1997b)

(a) Alternate test methods are also available.

Permeability, *k*, depends only on the properties of the porous medium (i.e., permeability is independent of the fluid type). Hydraulic conductivity is a function of the both the porous medium and the fluid. Permeability is related to hydraulic conductivity by the following equation:

$$K = \frac{k\rho g}{\mu}$$

where K is hydraulic conductivity, k is permeability, ρ is the density of the fluid, g is the acceleration due to gravity, and μ is the dynamic viscosity of the fluid. In general, permeability (and hydraulic conductivity) increases with increased grain size.

Porosity, n, is the ratio of the volume of voids within the porous medium to the total volume of the medium. As noted above, a measure of the effective porosity, $n_{e,}$ (i.e., the total porosity available for fluid flow), is necessary to calculate seepage velocity.

Redox, i.e., the reduction-oxidation potential (Eh) within the porous medium plays an important role in the geochemical processes that occur in groundwater. Redox reactions are defined as reactions that transfer electrons. The species receiving electrons is reduced, and the species donating electrons is oxidized. The phase equilibrium (i.e., partitioning between the solid and aqueous phases) of many metals and other chemicals in the groundwater environment is controlled by Eh. Redox reactions control the mobility of many inorganic compounds as well as biologically important elements such as nitrogen and sulfur.

pH controls equilibrium partitioning between the solid and aqueous phases for many metals and other chemicals in the groundwater environment. pH also affects the electric charge on the surface of fine-grained particles, particularly clays. At low pH, particles tend to have positively charged surfaces, whereas at high pH, negatively charged surfaces develop. Soils generally have pH values within the range of 4–8.5. High pH conditions tend to limit the concentrations and mobility of metals in groundwater by enhancing the tendency for metal ions (cations) to attach to clay particles. The point of zero charge (PZC) is the pH value at which a mineral surface effectively has no net charge. At pH values less than the PZC, a mineral surface has a net positive charge; at pH values greater than the PZC, net surface charge is negative. PZC pH values for the clay minerals kaolinite and montmorillonite, the iron oxide goethite, and quartz are listed in Table 2-2.

TABLE 2-2. Point of zero charge pH values

Mineral	pH Values Corresponding to Point of Zero Charge
Montmorillonite	2–3
Kaolinite	4.6
Goethite	7.3–7.8
Quartz	2–3

Cation exchange capacity, a measure of the ability of the soil to exchange cations, is one of the most important properties governing the behavior and transport of metals in soil and groundwater (Kabata-Pendias, 2001). CEC controls the quantity of metallic cations that can be removed from solution by adsorption to soil particles. High CEC indicates a high capacity for adsorption of metals, and is typical of organic matter and clayey soils (Birkeland, 1999). Fine-grained soils, particularly clays, have large surface/volume ratios and tend to exhibit high CEC values. Lower CEC values (typical of sandy soils) indicate a low capacity for adsorption of metals.

Total organic carbon content is a measure of the concentration of organic matter in a sample. Because of the affinity of metals for organic matter, high TOC concentrations are associated with low dissolved-phase metal concentrations, and thus tend to limit the mobility of metals in groundwater (Deutsch, 1997). In addition to its role in cation exchange reactions, organic carbon can adsorb metals by forming chelate complexes (Alloway, 1990). Humic compounds with suitable reactive groups, such as hydroxyl, phenoxyl, and carboxyl also form complexes with metallic ions. Transport of many organic chemicals is retarded by reversible sorption to organic carbon in the aquifer matrix (Freeze and Cherry, 1979; Fetter, 1994). Relative partitioning of organic chemicals between the aqueous and nonaqueous phases can be estimated by K_{ow}, the ratio of partitioning between octanol (representing the organic carbon phase) and water. Compounds with high K_{ow} values tend to adsorb readily to organic matter in soil or sediment.

Geochemical Characteristics of Groundwater

The following groundwater geochemical parameters should be measured to compile the data needed to evaluate the behavior of the COPC chemicals in the groundwater environment and evaluate similarities and differences between groundwater samples:

- ☐ Temperature
- ☐ pH
- ☐ Redox potential (Eh)
- ☐ Dissolved oxygen
- ☐ Specific conductivity
- ☐ TDS
- ☐ Salinity
- ☐ Major ion concentrations (Na^+, Ca^{2+}, K^+, Mg^{2+}, SO_4^{2-}, Cl^-, CO_3^{2-}, HCO_3^-).

Parameters such as temperature, pH, Eh, and dissolved oxygen should be measured in the field immediately after groundwater samples are brought to the surface in order to minimize the effects of interaction with the atmosphere and biological reactions. Various instruments are available for measuring these parameters in the field. Instruments capable of simultaneously measuring multiple parameters and recording the data in digital form are particularly useful for groundwater sampling and monitoring programs.

Temperature can be a very useful parameter for distinguishing between samples collected from different water bodies. Temperature also affects biological activity, and is therefore an important parameter for evaluating the potential for biodegradation or biotransformation.

pH and **Eh** data are essential for evaluating equilibrium partitioning, and are very useful for distinguishing between samples collected from different water bodies. pH – Eh graphs show the equilibrium partitioning characteristics of minerals and ions in the groundwater system, and are therefore useful for predicting the phase in which a particular chemical will tend to exist in the groundwater body of concern. Freeze and Cherry (1979) and Fetter (1994) provide an introduction to pH – Eh diagrams.

Dissolved oxygen concentrations indicate whether aerobic or anaerobic conditions exist in groundwater, and therefore provide useful information to assess the potential for biodegradation or biotransformation of COPC chemicals.

Specific conductivity, TDS, and salinity are measures of the concentrations of dissolved solids in water, and are directly proportional to each other. Conductivity is a measure of the ability of a solution to carry an electric current and depends on the total concentration of ionized substances dissolved in water. Specific conductivity, TDS, and salinity data are useful for making quick comparisons between groundwater samples.

Major ion concentration data provide very useful information for evaluating the similarities and differences between groundwaters. As groundwater moves through the saturated zone, TDS and major ion concentrations increase due to dissolution of minerals along the flowpath. The cations Na^+, Ca^{2+}, K^+, Mg^{2+} and the anions SO_4^{2-}, Cl^-, CO_3^{2-}, and HCO_3^- account for more than 90 percent of the total dissolved species in most groundwaters. A common method of plotting relative abundance of the major ions is the trilinear diagram. This plotting technique allows numerous groundwater analyses to be plotted on a single diagram, and provides a quick and concise means of visually comparing chemical similarities and differences among groups of groundwater analyses, as well as identifying mixing trends between different groundwaters. Further information on the trilinear plotting diagram is presented in Section 3, and in Freeze and Cherry (1979) and Fetter (1994).

It should be noted that groundwater geochemical parameters such as TDS, major ion, and trace element concentrations often show significant spatial variability in freshwater-saltwater transition zones along coastlines, and near inland surface water bodies such as lakes and rivers. Box 2-3 describes some issues that should be considered when selecting background monitoring well locations and evaluating groundwater data to characterize background conditions in these transitional environments.

Geochemistry of COPCs in the Groundwater Environment

Knowledge of the chemistry and geochemistry of the chemicals identified as COPCs for a groundwater site, particularly metals, is essential for evaluating groundwater background conditions. The geochemical parameters described above should be measured in both in soil and groundwater to provide the data needed to fully assess groundwater background conditions. The geochemical literature can be a valuable source of information regarding the behavior of metals and other chemicals in the groundwater environment. Literature sources include Langmuir (1997); Drever (1997); Hemond and Fechner-Levy (1999); Jenne (1998); Deutsch (1997); and Fetter (1999).

For example, when geochemical conditions favor partitioning into the solid phase, metals tend to be immobile in the environment. However, when geochemical conditions change due to natural or anthropogenic causes, metals can partition from the solid phase into the aqueous phase and be transported with groundwater. Weathering is an example of a natural process that can change geochemical conditions and mobilize metals. In chemical weathering, chemicals are altered to soluble forms that are readily transported with groundwater. For example, lead carbonate ($PbCO_3$) in soil reacts with carbonic acid (H_2CO_3) from infiltrated rainwater to release Pb^{2+} into solution.

Metals exist in nature primarily in positive valence (or oxidation) states, and many form stable compounds in more than one oxidation state. Thermodynamically, virtually all metals in the elemental form are unstable with respect to redox reactions in environments where they are exposed to air and

BOX 2-3. Evaluating background conditions in transitional environments

Groundwater geochemical parameters, including concentrations of both major constituents and trace elements, can show considerable spatial variability in transitional environments. As noted above, groundwater geochemical conditions at background monitoring well locations must be similar to those in the investigation site area. In addition, under some conditions, a geochemical trend associated with a transition zone could be confused with the concentration gradient associated with a site related chemical release. Therefore, if the investigation or background site is located in a transitional environment, the project team should carefully consider geochemical trends associated with the transition zone when selecting background monitoring well locations and evaluating groundwater data to characterize background conditions.

- **Transition from freshwater to saltwater.** Groundwater concentrations of the major ions and trace elements associated with seawater can increase significantly with distance toward the coastline in the transition zone between freshwater and saltwater. Diffusion causes mixing of saltwater and freshwater in the transition zone; therefore, the extent of the transition zone depends on the hydrodynamic dispersion characteristics of the aquifer formation (Freeze and Cherry, 1979). The location of the transition zone can change due to seasonal recharge variations. Tidal effects cause diurnal variations. Decreases in the level of the water table (unconfined aquifer) or potentiometric surface (confined aquifer) caused by pumping significant volumes of groundwater from wells can allow seawater intrusion and cause the transition zone to migrate inland.
- **Recharge from surface water bodies.** The effect of surface water bodies such as lakes, ponds, and rivers on the groundwater system should be evaluated to understand natural geochemical trends within the groundwater body of concern. Large permanent lakes are usually discharge areas for regional groundwater systems; however, groundwater bodies can also be recharged by leakage through the beds of lakes, ponds, and rivers (Freeze and Cherry, 1979). Groundwater recharge from surface water can affect chemical concentration trends at a site. For example, dilution associated with recharge from surface water may cause local decreases in the concentrations of dissolved constituents.

water. Many metals have more than one potentially stable positive oxidation state. These different oxidation states can have dramatically different chemical properties. For example, almost 4 g/L of ferrous iron (Fe^{2+}) can dissolve in distilled water maintained at pH 7.0. However, if the water is exposed to the air and the iron is oxidized to Fe^{3+}, essentially all the iron will precipitate, reducing the dissolved iron concentration by more than eight orders of magnitude. The oxidation or valence state of a metal can also affect its toxicity. For instance, As^{3+} is considerably more toxic to both aquatic organisms and humans than As^{5+}.

Metal dissolution, precipitation, and adsorption are strongly related to redox and pH conditions. Eh – pH diagrams can be useful to evaluate equilibrium partitioning in the groundwater environment, and determine whether the COPC metals will tend to exist in the solid or aqueous phases. An introduction to Eh and pH diagrams is presented in Fetter (1994) and Freeze and Cherry (1979). The project team should also consult the geochemical literature to obtain the specific data and information necessary to evaluate the effects of different Eh – pH conditions on equilibrium partitioning of the COPCs identified for the particular groundwater site.

2.1.5 Determine Whether Adequate Site/Background Groundwater Data Exist

Once quantitative and qualitative groundwater information and data have been compiled and reviewed, the project team must determine whether the existing data are adequate. If the project team concludes that additional data are needed to proceed with the background analysis, a SAP or GWMP should be developed within the framework of the DQO process. The existing data should be reviewed as follows to identify data gaps and assess the adequacy of the dataset for background analysis:

- ☐ ***Review number of measurements.*** The background analysis and comparative statistical methods (Section 4) presented in this guidance document typically require chemical data representing at least 10 groundwater monitoring wells located in both impacted and unimpacted zones, and a series of semiannual or quarterly sampling events. If the available dataset represents only a few samples collected from a small number of wells and a limited number of sampling events, the background analysis methods presented in the following sections may not yield reliable results. As discussed in Box 2-4, experience at Navy sites demonstrates that background analysis based on small datasets should be avoided if possible; however, it may be possible to adequately characterize background conditions if the numerical data are augmented with hydrogeological and geochemical evidence to demonstrate that data collected at certain locations most likely represent background conditions.

- ☐ ***Review sampling locations and depths.*** Locations of the monitoring wells with respect to potential on-site and off-site contaminant sources should be reviewed to ensure that the dataset is likely to include data from wells and sampling depth intervals not impacted by COPCs. Hydrogeological information and data, such as groundwater flow direction, flow velocity, and aquifer thickness, must be included in this review. COPC characteristics should also be evaluated; for example, LNAPLs will tend to impact primarily the upper portion of a groundwater body first, whereas DNAPLs will sink toward the bottom of an aquifer.

- ☐ ***Review sampling times.*** Data representing a series of distinct sampling events are required to evaluate seasonal trends in concentrations of both background chemicals and COPCs. At least two rounds of sampling (i.e., wet and dry season sampling events) are required. However, semiannual or quarterly sampling over several years is usually necessary to fully characterize seasonal trends. If data representing seasonal variations are insufficient, the project team could erroneously identify a seasonal increase in background concentrations as evidence of a chemical release.

- ☐ ***Review target chemicals and geochemical parameters.*** The dataset should include

BOX 2-4. Analysis based on small datasets

In many cases, the available quantitative data (i.e., concentration measurements) adequately represent groundwater conditions within the actual or suspected contaminant plume, but contain a very limited number of "background" measurements collected from a few wells located upgradient of the suspected source area, or in another area believed to be nonimpacted. In these cases, the quantitative dataset by itself will be too small to reliably characterize background conditions.

If necessary, it may be possible to generate a rough estimate of the background concentration range for a chemical by evaluating a small dataset. In such cases, the background analysis must include a thorough review of site history to accurately identify contaminant source locations, and the hydrogeological and geochemical evidence should be thoroughly evaluated to augment the numerical dataset. Although these procedures can provide evidence to justify selection of background wells and may yield defensible background range estimates for the target COPCs, background range estimates based solely on a small dataset may be questioned as unreliable.

Past experience at Navy sites has demonstrated that relying solely on small background datasets to estimate background concentration ranges can lead to the following problems:

- **Unreliable Statistical Analyses:** When the background dataset is small, it is difficult or impossible to make statistically robust and reliable estimates of background concentration ranges; furthermore, statistical comparison of the background dataset to the dataset representing the potentially impacted population will be problematic and prone to large errors.
- **Unexplainable Large Concentration Ranges:** Concentration values within a small background dataset can be highly variable and cover a very wide range. These situations tend to confuse the analysis and lead to the erroneous labeling of true background concentrations as statistical outliers. This may raise questions as to whether the higher concentrations represent background conditions or indicate the presence of contamination.
- **Unproductive Negotiations:** Limited background datasets are often questioned as not representative of actual conditions, leading to complicated negotiations, qualitative discussions, and speculative conjectures. In such cases, the background dataset may become the primary subject of remedial discussions rather than a useful tool for decision-making.

groundwater data representing the site-specific COPCs, potential products of chemical reactions or breakdown of COPCs (if any), potential site-specific background chemicals, chemicals required by regulatory criteria (i.e., groundwater monitoring requirements), and geochemical parameters that may be useful for background analysis.

❑ ***Review soil sampling data.*** Data representing chemical concentrations in soil samples collected in potentially impacted areas at the site and in unimpacted areas in the site vicinity can be useful to associate chemicals detected in groundwater samples with potential sources of COPCs and background chemicals.

❑ ***Review sampling and analysis methods.*** The sampling and analysis methods used to generate the data should be reviewed and compared to appropriate standard operating procedures (SOPs) to ensure that approved procedures have been followed for monitoring well installation, well development, groundwater sampling, sample analysis, and data validation. SOPs are available in NAVFAC and EFD/EFA-specific procedures manuals.

2.1.6 Develop and Implement Sampling and Analysis Plan or Groundwater Monitoring Plan

The sampling and analysis program should be carefully designed and implemented to fill the data gaps identified by the evaluation described in Section 2.1.5, and should ensure that the data needed to evaluate background conditions and meet groundwater monitoring requirements are collected. A GWMP is required for a long-term groundwater monitoring program, whereas a groundwater characterization investigation requires a SAP. Guidance for preparing SAPs and GWMPs is available in

Lesnik and Crumbling (2001) and *Guide to Optimal Groundwater Monitoring* (DON, 2000), respectively. Careful planning is particularly important for groundwater investigations and monitoring due to the high cost of drilling, well installation, groundwater sampling, and sample analysis. The project team should consult the groundwater monitoring literature (e.g., Nielsen, 1991) and review applicable Navy SOPs (e.g., NAVFAC and EFD/EFA-specific procedures manuals) to ensure that the procedures specified in the SAP or GWMP meet applicable requirements. As additional data needs become evident (e.g., if it is necessary to increase the number of monitoring wells or collect data from off-site wells), the SAP or GWMP should be amended or an addendum should be prepared.

The project team must seek concurrence of stakeholders (e.g., U.S. EPA and state regulators). A draft version of the SAP or GWMP should be prepared and submitted for review. The SAP or GWMP should be implemented only after concurrence by all stakeholders. Finally, the project team must ensure compliance with all requirements and procedures established by the SAP or GWMP during the field sampling and laboratory analytical programs.

The SAP or GWMP should specify the DQOs and the following procedures that will be used to collect the data, and should address factors that must be considered when new data are combined with a previously existing dataset (see Section 2.1.7).

- ☐ ***Number of measurements.*** The SAP or GWMP should identify the number of additional measurements required to meet the DQOs. As noted in Section 2.1.5, the background analysis and comparative statistical methods (Section 4) presented in this guidance document typically require chemical data representing at least 10 groundwater monitoring wells located in both impacted and unimpacted zones, and a series of semi-annual or quarterly sampling events. The SAP or GWMP should provide for installation of new monitoring wells or additional sampling events as required to augment the existing dataset.
- ☐ ***Sampling locations and depths.*** In many cases, most of the existing groundwater data will represent samples collected from monitoring wells installed at and downgradient of on-site contaminant sources. Therefore, it is often necessary to augment the dataset by installing new monitoring wells in locations likely to yield samples of unimpacted groundwater. Hydrogeological information and data should be evaluated with respect to the locations of known and potential contaminant sources to select locations and screen depth intervals for the new wells.
- ☐ ***Sampling times.*** Seasonal variations in recharge rates should be evaluated to select sampling times appropriate for collecting the groundwater data needed to characterize seasonal concentration trends. At least two rounds of sampling (i.e., wet and dry season sampling events) are required to evaluate seasonal trends. However, depending on climatic conditions in the recharge area, semi-annual or quarterly sampling over several years may be required. Regulatory sampling frequency requirements also should be reviewed to select appropriate sampling intervals.
- ☐ ***Groundwater data.*** The groundwater samples should be analyzed for site-specific COPCs, potential products of chemical reactions or breakdown of COPCs (if any), potential site-specific background chemicals, chemicals required by regulatory criteria (i.e., groundwater monitoring requirements), and other data that may be useful for background analysis (see Section 2.1.4).
- ☐ ***Aquifer matrix characteristics.*** The characteristics of the aquifer matrix play a very important role in the distribution of background chemicals in the groundwater environment. Therefore, soil or rock samples should be collected from the monitoring well borings to acquire the data needed to evaluate aquifer matrix characteristics. The soil or rock samples should be examined by a qualified geologist, soil scientist, or engineer, and analyzed for site-specific COPCs and potential background chemicals. Soil samples

also should be submitted for testing to evaluate the geotechnical and chemical parameters listed in Table 2-1.

- ***Monitoring well installation.*** Monitoring wells should be installed by a licensed well driller under the supervision of a qualified geologist or hydrogeologist. Soil or rock samples should be collected from each boring to assess subsurface lithology. The geologist or hydrogeologist should determine whether lithologic sampling should be continuous or restricted to specified intervals, depending on the anticipated subsurface complexity, and then should adjust sampling intervals as necessary based on field observations. Samples should also be collected from the boring for chemical analysis to evaluate the nature and extent of soil contamination (if necessary). The first indications of groundwater in the boring should be recorded, and the monitoring well should be screened across the static groundwater level as necessary to account for seasonal water table fluctuations. Primary monitoring wells are typically constructed of 2-inch- or 4-inch-diameter, Schedule 40 polyvinyl chloride (PVC). The wells should be completed with a filtration sand pack around the well screen, a bentonite seal atop the sand, and the remaining annular space grouted to the surface. After installation and development, the monitoring well locations should be surveyed by a licensed land surveyor to establish location coordinates and top-of-casing elevations. SOPs for well installation and development are available in NAVFAC and EFD/EFA-specific procedures manuals.

- ***Sampling protocol.*** Procedures for drilling, monitoring well installation, well development, and groundwater sampling must be carefully selected and implemented to ensure that the groundwater data are representative of the groundwater body of concern. Proper procedures are essential to avoid cross-contamination, ensure that the well is screened across the appropriate depth interval, and limit the turbidity of groundwater samples. Groundwater samples should be filtered in the field for analysis of dissolved chemical concentrations; however, analysis of total chemical concentrations (i.e., the sum of dissolved- and solid-phase chemical concentrations) will require unfiltered samples. The project team should identify and implement appropriate procedures after consulting the groundwater monitoring literature (e.g., Nielsen, 1991) and reviewing the applicable SOPs (e.g., NAVFAC and EFD/EFA-specific procedures manuals).

- ***Laboratory analysis.*** Methods for laboratory analysis must be selected and implemented in accordance with the DQOs and applicable Navy SOPs. The SAP or GWMP should specify laboratory reporting limits (RLs) for all analytes. The RL is typically one to five times the detection limit (DL), depending on the analytical method and matrix. The DL can vary considerably from sample to sample because of matrix effects. Ideally, the RL will not change, and will be set high enough to account for matrix effects, yet low enough to meet project-specific DQOs. RLs must be low enough to allow comparison to risk-based screening criteria, and, to ensure that data quality is adequate for background analysis, should be set at levels well below groundwater background ranges cited in the literature. Because certain chemicals can pose unacceptable risks to ecological receptors at low concentrations, analytical methods with exceptionally low RLs, such as NOAA National Status and Trends methods, may be required. In addition, the SAP or GWMP should focus on the collection of only those analytical data necessary to evaluate contamination, estimate background ranges, or achieve other site-specific DQOs—unnecessary data should not be collected. The SAP or GWMP also must establish quality assurance (QA) and quality control (QC) procedures for the field sampling and laboratory analytical programs (for example, the SAP or GWMP should specify requirements for field QC sampling).

2.1.7 Evaluate Data

After the analytical data are received from the laboratory, the data should be evaluated to ensure they are of the right type, quality, and quantity for the intended background analysis method. The U.S. EPA guidance manual *Guidance for Data Quality Assessment* (2000a) provides general guidance for data quality assessment, and describes how it fits into the DQO process. Data validation is a key component in the data evaluation process.

Validate Data

Data should be validated for quality in accordance with Appendix H, Data Validation, of the Navy *Installation Restoration Chemical Data Quality Manual (IR CDQM)* (DON, 1999a). The data validation procedures assign qualifiers to the data that give the end users a qualitative measure of data usability. If no qualifier (NQ) is assigned to an analytical measurement, all QC criteria associated with the measurement were within acceptance criteria and the chemical was quantified at a concentration above the laboratory RL. Otherwise, measurements are typically assigned one the qualifiers listed in Table 2-3.

Combine Datasets

Background analysis often involves combining two or more separate datasets, such as previously existing groundwater data and supplemental data acquired during additional sampling. Combining two or more datasets to form a larger dataset may improve the ability of the analysis to differentiate between background and COPCs. Pooling the data will increase the number of data points available for background analysis and improve the reliability of the results. However, an inappropriate combination of datasets can have the opposite effect.

Before datasets are combined, the spatial and temporal boundaries within which a chemical is evaluated must be carefully defined. From a statistical point of view, such spatial and temporal areas are considered target populations. If multiple datasets are to be combined, they must all be representative of the same target population. Datasets that represent groundwater with dissimilar hydrogeological and geochemical characteristics are not suitable for combination. For example, data representing the chemical characteristics of a deep confined freshwater aquifer should not be combined with data representing a brackish near-surface caprock groundwater body.

Ideally, the datasets being considered for pooling should be obtained using the same sampling design and analytical methods. For example, certain classes of organic compounds (e.g., PAHs and PCBs) can be quantified either as individual compounds or as functional groups. Before organic chemical datasets are combined, investigators should verify that the same procedures were used to calculate total concentrations for each dataset. It also is important to verify that measurements in all the datasets being considered for pooling have similar quality characteristics. For example, the RLs and measurement biases should be sufficiently low, and an adequate number of blank and duplicate samples should be taken to check for the magnitude of bias and variability. (Requirements for blank and duplicate sampling frequencies should be documented in the SAP or GWMP.) Furthermore, to ensure consistency, similar sampling and handling procedures should be used for

TABLE 2-3. Data qualifiers

Qualifier	Definition	Explanation
J	Estimated concentration	The analyte was positively identified, and the associated concentration value is an estimated quantity. “J” data are biased, but provide definitive analyte identification, and are usually reliable.
U	Nondetect	The sample was analyzed for the chemical, but the chemical was not detected at a concentration above the sample quantitation limit (either the RL or the DL).
UJ	Nondetect estimated	The sample was analyzed for the chemical, but the chemical was not detected at a concentration above the sample quantitation limit (either the RL or the DL). The associated value is an estimate and may be inaccurate or imprecise.
N	Tentatively identified	The analysis indicates the presence of an analyte that has been “tentatively identified.” The associated value represents the approximate concentration
R	Rejected	The analytical results are rejected due to serious deficiencies in the ability to analyze the sample and meet quality control criteria. The presence or absence of the analyte cannot be verified.

all the samples used to generate the datasets to be pooled. In addition to assessing the hydrogeological and geochemical characteristics of the groundwater, graphic and statistical methods can be helpful to confirm that the datasets represent the same target population (Box 2-5).

2.1.8 Screen Data and Identify COPCs for Background Analysis

After the new data have been validated according to the DQO requirements and combined with the previously existing data, the maximum detected chemical concentrations should be compared to appropriate risk-based screening criteria (as described in Section 2.1.3) to identify COPCs requiring background analysis.

2.2 Overview of Groundwater Background Analysis Methodology

After identifying the COPCs that require background analysis, the project team should implement the background analysis methodology presented in Section 3 of this document. The objectives of these procedures are: (1) to identify background monitoring wells, (i.e., wells that can be sampled to provide background data), (2) to estimate the background concentration range for each target COPC, and (3) to ensure that the background dataset is adequate and appropriate for statistical comparison to the dataset representing potentially contaminated site groundwater. These procedures follow a three-step "weight of evidence" process to assemble the lines of evidence required to characterize groundwater background conditions.

Step 1: Evaluate Geology, Hydrogeology, and Contaminant Sources. The first step in the process involves evaluating local and regional hydrogeology, and assessing the locations and characteristics of actual and suspected contaminant sources (both on-site and off-site) to prepare a conceptual site model for background analysis. The conceptual site model can then be used to compile the hydrogeological and contaminant source information and data to aid in identifying locations suitable for background monitoring wells.

Step 2: Evaluate Geochemical Conditions. In this step, the project team should evaluate geochemical data to confirm that groundwater data collected from existing or prospective background monitoring wells are appropriate for use as background data. If the geochemical characteristics of the prospective background groundwater are not similar to those of the site groundwater, then the data should not be used to represent local background conditions. Groundwater geochemistry can be characterized by evaluating parameters such as pH, salinity, and concentrations of the major dissolved components (i.e., the dominant cations and anions). The geochemistry of the rock or soil that forms the aquifer matrix should also be evaluated during this step. This will provide further evidence to confirm that the background data are suitable for comparison to site data, and may allow investigators to identify potential sources of the groundwater background chemicals.

BOX 2-5. Statistical methods for comparing datasets

Graphic Methods: Histograms, boxplots, and probability plots of individual datasets are useful to assess the similarity between datasets.

Two Datasets: The Comparative Method (Section 4) can be used to evaluate the difference between mean or median concentrations. Differences in the variance of measurements for the two datasets that have a normal distribution (with possibly different means) could be tested using the F test described in U.S. EPA (2000a, p. 4-33) and Conover (1998). The Squared Ranks Test of variances (Conover, 1998, p. 300) may be used to test for equality of variances. This test may be applied regardless of the shape of the data distributions.

Multiple Datasets: The Kruskal-Wallis test (Gilbert, 1987, p. 250; Conover, 1998, p. 288) may be used to assess differences among their median concentrations. Equality of variances of more than two datasets can be tested as described by Conover (1998, p. 303). Both of these tests may be applied regardless of the shape of the data distributions.

Step 3: Conduct Spatial and Temporal Groundwater Data Analysis. In this step, concentrations of the target COPCs are plotted on graphs and maps to compare concentrations with respect to both space and time. This step will allow investigators to identify the wells that can be used to represent background conditions and estimate the associated background concentration ranges.

After completing this process, investigators should compile and summarize the hydrogeologic, geochemical, and target COPC concentration evidence, and present the background analysis conclusions for stakeholder review. If the choice of background wells and estimated background ranges are not acceptable to stakeholders, then the deficiencies should be identified, additional data should be collected if necessary, and the background analysis steps should be repeated as required. If the background analysis conclusions are acceptable, then the project team should decide whether the statistical comparative methods described in Section 4 of this document are necessary.

If the maximum detected concentrations of all target COPCs in the site groundwater are within the estimated background ranges, then statistical comparison is not required. If the maximum concentrations exceed the upper bound of the background range, but the exceedances are not high enough to be obviously inconsistent with the background population, then the statistical comparison methods described in Section 4 should be implemented. If the statistical comparison shows that exceedances of the initially estimated background range are not statistically significant (i.e., the on-site chemical concentrations are consistent with the background concentrations), then the background range estimate should be revised upward to include the on-site concentration values. (Note that this statement assumes that the site and background datasets are large enough to attain the desired test power; the desired power and adequacy of the datasets must be addressed within the DQO and DQA framework.) However, if COPC concentrations in the site groundwater are considerably higher than the upper bound of the background range, and site knowledge and professional judgment indicate that the concentrations are consistent with a site-related release, then the time and money required to implement the statistical comparative methods may not be justified. After completing this process, the results and conclusions should be documented in the investigation report (e.g. the Remedial Investigation Report) or appropriate section of the risk assessment report.

2.3 Overview of Statistical Comparative Testing

The statistical comparative testing methods are designed to compare data representing background conditions against individual or pooled data representing potentially impacted groundwater. The reliability of site-wide comparisons is ensured by the use of datasets large enough to adequately represent both populations. In the case of individual comparisons, verification resampling procedures are used to attain the desired reliability. Prior to implementing the statistical comparative tests, the results of the background analysis methods presented in Section 3 should be used to separate the data for each COPC into two separate datasets, i.e., the site dataset (representing potentially impacted groundwater) and the background dataset.

Comparative statistical tests are categorized as parametric or nonparametric. Parametric tests are based on specific distributional assumptions, (e.g., normality of mean concentrations), whereas nonparametric tests require no such assumptions. Each test is designed to assess specific aspects of the investigated data. For example, certain tests are designed to evaluate the similarity of extreme site and background concentrations, whereas others are designed to assess central tendencies (median or mean) of the observed concentrations.

To ensure the success of the comparative statistical methods, stakeholder acceptance of the validity of the site and background datasets is particularly critical. This acceptance can be attained by effectively communicating the background analysis results, including the hydrogeological and geochemical evidence used to identify and distinguish potential background data from site data.

2.4 Cost-Benefit Analysis

Characterization of groundwater background conditions and statistical comparative testing require significant effort, cost, and time. The project team will need to evaluate hydrogeologic, geochemical, and contaminant source characteristics; identify background well locations; estimate background concentration ranges; ensure that the background dataset is adequate for statistical comparison; and (if necessary) implement the statistical comparative tests. In addition, significant stakeholder interaction typically is required for successful completion of projects involving background analysis and statistical comparison. Decision-makers should assess whether the potential benefits of background analysis and statistical comparative testing justify the associated costs. In general, the benefits of background analysis are most likely to justify the associated costs when groundwater is not impacted or only marginally impacted by site-related releases. In addition, background analysis can be very beneficial for sites located in areas that have been exposed to long-term anthropogenic (not site-related) chemical sources.

2.5 Regulatory and Stakeholder Acceptance

The methods used for background analysis and statistical comparison must be acceptable to the regulators and other stakeholders involved in the project. The most favorable conditions are: (a) the stakeholders promote the use of the methodology; (b) the stakeholders have accepted the methodology for similar sites; (c) the stakeholders do not insist on the use of alternative, overly conservative procedures; and (d) the stakeholders are willing to accept the decision process prior to sampling and analysis. Examples of U.S. EPA and state technical guidance on background analysis procedures are presented in Section 1.

2.6 References

Alloway, B.J. 1990. *Heavy Metals in Soils.* Glasgow: Blackie and Son.

American Society for Testing and Materials (ASTM). 2003a. *D1498-00 Standard Practice for Oxidation-Reduction Potential of Water.* West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 2003b. *D2937-00 Standard Test Method for Density of Soil in Place by the Drive-Cylinder Method.* West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 2003c. *D422-63(1998) Standard Test Method for Particle-Size Analysis of Soils.* West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 2003d. *D854-02 Standard Test Methods for Specific Gravity of Soil Solids by Water Pycnometer.* West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 2003e. *E1527-00 Standard Practice for Environmental Site Assessments: Phase I Environmental Site Assessment Process*. West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 2003f. *E1903-97 Standard Guide for Environmental Site Assessments: Phase II Environmental Site Assessment Process*. West Conshohocken, PA.

Birkeland, P.W. 1999. *Soils and Geomorphology.* 3rd ed. New York: Oxford University Press.

Conover, W.J. 1998. *Practical Nonparametric Statistics*. 3rd ed. Wiley Series in Probability and Statistics, Applied Probability and Statistics Section. New York: Wiley & Sons.

Department of the Navy (DON). 1999a. *Navy Installation Restoration Chemical Data Quality Manual (IR CDQM).* NFESC Special Report SP-2056-ENV. Washington, DC: Naval Facilities Engineering Command, Naval Facilities Engineering Service Center. September.

Department of the Navy (DON). 1999b. *Navy Policy for Conducting Ecological Risk Assessments.* Chief of Naval Operations Letter Ser. N453E/9U595355. Washington, DC. April 5.

Department of the Navy (DON). 2000. Interim Final *Guide to Optimal Groundwater Monitoring*. Port Hueneme, CA: Naval Facilities Engineering Service Center. January.

Department of the Navy (DON). 2001. *Navy Policy for Conducting Human Health Risk Assessments Related to the Installation Restoration Program.* Chief of Naval Operations Letter Ser. N453E/1U595168. Washington, DC. February 12.

Deutsch, W.J. 1997. *Groundwater Geochemistry: Fundamentals and Applications to Contamination.* Boca Raton, FL: CRC Press.

Drever, J.I. 1997. *Geochemistry of Natural Waters: Surface and Groundwater Environments.* 3rd ed. Upper Saddle River, NJ: Prentice Hall.

Fetter, C.W. 1994. *Applied Hydrogeology*. 3rd ed. Upper Saddle River, NJ: Prentice-Hall.

Fetter, C.W. 1999. *Contaminant Hydrogeology.* Upper Saddle River, NJ: Prentice-Hall.

Freeze, R.A., and J.A. Cherry. 1979. *Groundwater*. Englewood Cliffs, NJ: Prentice-Hall.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring.* New York: Van Nostrand Reinhold.

Hem, J.D. 1992. *Study and Interpretation of the Chemical Characteristics of Natural Water*. 3rd ed. U.S. Geological Survey Water-Supply Paper 2254. Alexandria, VA: USGS.

Hemond, H.F., and E.J. Fechner-Levy. 1999. *Chemical Fate and Transport in the Environment.* 2nd ed. San Diego, CA: Academic Press.

Jenne, E.A., ed. 1998. *Adsorption of Metals by Geomedia.* San Diego: Academic Press.

Kabata-Pendias, A. 2001. *Trace Elements in Soils and Plants.* 3rd ed. Boca Raton, FL: CRC Press.

Langmuir, D. 1997. *Aqueous Environmental Chemistry.* Englewood Cliffs, NJ: Prentice-Hall.

Lesnik, B. and D. Crumbling. 2001. *Guidelines for Preparing SAPs Using Systematic Planning and PBMS.* Environmental Testing & Analysis, Vol.10, No.1. January/February. pp. 26-40. Available at: http://cluin.org/download/char/etasaparticle.pdf.

Nielsen, D.M. 1991. *Practical Handbook of Groundwater Monitoring.* Chelsea, MI: Lewis Publishers.

United States Environmental Protection Agency (U.S. EPA). 1989. *Risk Assessment Guidance for Superfund, Volume I, Human Health Evaluation Manual (Part A)*. Interim Final. EPA/540/1-89/002. Washington, DC: U.S. EPA Office of Emergency and Remedial Response. December.

United States Environmental Protection Agency (U.S. EPA). 1991. *Risk Assessment Guidance for Superfund, Volume I—Human Health Evaluation Manual (Part B, Development of Risk-Based Preliminary Remediation Goals).* Interim. EPA/540/R-92/003. Washington, DC: U.S. EPA Office of Emergency and Remedial Response. December.

United States Environmental Protection Agency (U.S. EPA). 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessment.* Interim final. EPA/540/R-97/006. Washington, DC: U.S. EPA Office of Solid Waste and Emergency Response. June.

United States Environmental Protection Agency (U.S. EPA). 1997b. *Test Methods for Evaluating Solid Waste, Physical/Chemical Methods, SW-846.* 3rd ed., Final Update IIIA. Washington, DC: U.S. EPA Office of Solid Waste.

United States Environmental Protection Agency (U.S. EPA). 2000a. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis.* EPA QA/G-9, QA 00 Update. EPA/600/R-96/084. Washington, DC: U.S. EPA Office of Environmental Information. July.

United States Environmental Protection Agency (U.S. EPA). 2000b. *Guidance for the Data Quality Objectives Process.* EPA QA/G-4. EPA/600/R-96/055. Washington, DC: U.S. EPA Office of Environmental Information. August.

United States Environmental Protection Agency (U.S. EPA). 2002. *Current Drinking Water Standards: National Primary and Secondary Drinking Water Regulations.* EPA 816-F-02-013. Washington, DC: U.S. EPA Office of Ground Water and Drinking Water. July.

3. IDENTIFY BACKGROUND SAMPLING LOCATIONS AND CONCENTRATION RANGES

This section presents a methodology for identifying a background dataset for COPCs that require background analysis. The background dataset can be used both to estimate background concentration ranges, and for comparison to site concentrations. The objectives of the methodology are as follows:

- ❑ Identify the locations of existing or prospective background wells (i.e., wells that can be sampled to provide background data).
- ❑ Estimate the background concentration range for each target COPC.
- ❑ Ensure that the background dataset is adequate and appropriate for statistical comparison to the site dataset.

The background analysis methodology uses the following three-step "weight of evidence" process to provide the evidence required to characterize groundwater background conditions:

1. Evaluate the geologic and hydrologic characteristics of the site and background sampling areas, and the locations and characteristics of actual and suspected contaminant sources.
2. Evaluate the geochemical characteristics of the groundwater and aquifer matrix.
3. Construct graphs and maps to evaluate concentrations of the target COPCs with respect to space and time.

A case study illustrating the application of the methodology described in this section to a hypothetical investigation site is presented in Section 5.

3.1 Geology, Hydrogeology, and Contaminant Source Evaluation

The first steps for identifying a background dataset for COPCs are to evaluate the local and regional geology and hydrogeology, assess the locations and characteristics of actual and suspected contaminant sources (both on site and off site), and prepare a CSM for background analysis. The CSM is used to compile hydrogeological and contaminant source information and data in order to allow investigators to identify existing wells or prospective monitoring well locations for sampling to characterize background conditions (i.e., locations likely to be unimpacted by on-site or off-site contaminant sources). The conclusions reached after completing this step should justify the selection of background well locations and sampling depths, and thus provide the first lines of evidence for the background evaluation.

If no suitable locations for background monitoring wells (i.e., unimpacted regions of the aquifer) can be identified at or near the investigation site, the project team should expand the search for background data to other locations in the region, and repeat the evaluation process to identify a site with similar geologic, hydrogeologic, geochemical, and anthropogenic characteristics. Then, with stakeholder concurrence, the similar location can provide background groundwater data appropriate for comparison to the on-site groundwater data. A detailed CSM of the prospective background area is essential to demonstrate that the groundwater conditions are similar to conditions at the site. The CSM for the prospective background site should also address the hydrogeologic and contaminant source parameters outlined below.

3.1.1 Evaluate Geology and Hydrogeology

Geologic and hydrogeologic parameters that should be evaluated to develop the CSM include:

- ❑ ***Geologic cross section.*** To assess geologic formation, soil type, and groundwater regime (the required number of cross sections will depend on site conditions and subsurface complexity).

- ❑ ***Aquifer type (unconfined, confined, perched) and geologic structure.*** To evaluate groundwater flow within the aquifer, between different groundwater bodies, and potential barriers to groundwater flow (e.g., confining layers, aquitards).

- ❑ ***Aquifer thickness, depth to groundwater, magnitude of seasonal and tidal water table elevation variations.*** To determine appropriate depths and lengths of monitoring well screened intervals.

- ❑ ***Direction of the hydraulic gradient.*** To assess the direction of groundwater flow and chemical transport.

- ❑ ***Magnitudes of the hydraulic gradient, hydraulic conductivity, and porosity.*** To assess groundwater seepage velocity.

- ❑ ***Aquifer transmissivity.*** To measure the amount of water that can be delivered to a well through a unit width given a unit hydraulic gradient.

- ❑ ***Estimated hydrodynamic dispersion.*** To assess potential for COPCs to spread out in directions both parallel and perpendicular to the groundwater flow direction.

- ❑ ***Characteristics of the aquifer matrix.*** To assess the effect of the aquifer matrix materials on the transport of chemicals in groundwater. For example, reversible adsorption to clays and organic carbon may retard chemical transport velocities.

- ❑ ***Potential for natural attenuation due to chemical and biological processes.*** To assess potential for reduction in COPC concentrations with both time and distance from the contaminant source.

- ❑ ***Locations of nearby streams, rivers, lakes, ponds, springs, seeps, private wells, and production wells.*** To describe potential discharge areas and groundwater/surface water interactions.

3.1.2 Identify Contaminant Source Locations and Characteristics

Contaminant source and COPC parameters that should be evaluated to prepare the CSM include:

- ❑ ***Locations of all actual and suspected contaminant sources, both on site and off site.*** To identify areas where groundwater is likely to be impacted, and assess potentially unimpacted areas for installation of background monitoring wells.

- ❑ ***Nature and extent of the contaminant source.*** To evaluate release of COPCs from the source media. For example, COPCs may reach the groundwater directly (e.g., infiltration of free product), or COPCs may be adsorbed to soil particles and then released by leaching and transported downward toward the water table. Also, it should be determined whether the primary contaminant source is still active (e.g., ongoing release or leak from storage tanks).

- ❑ ***COPC characteristics affecting transport.*** To evaluate mobility of COPCs in the groundwater environment. Important chemical characteristics include solubility, affinity for organic carbon, chemical partitioning to solids, and the effects of pH and redox conditions.

- ❑ ***Depths of contaminant sources with respect to the ground surface and the water table.*** To evaluate potential for transport of COPCs downward through the vadose zone toward the water table.

3.1.3 Develop Conceptual Site Model for Background Analysis

As shown in Figure 3-1, the CSM should include a plan view (illustrating surface features, the hydraulic gradient, well locations, etc.), one or more cross-sectional views (illustrating the geology of the vadose and saturated zones), a table of supplemental information and data, and a summary of other relevant information.

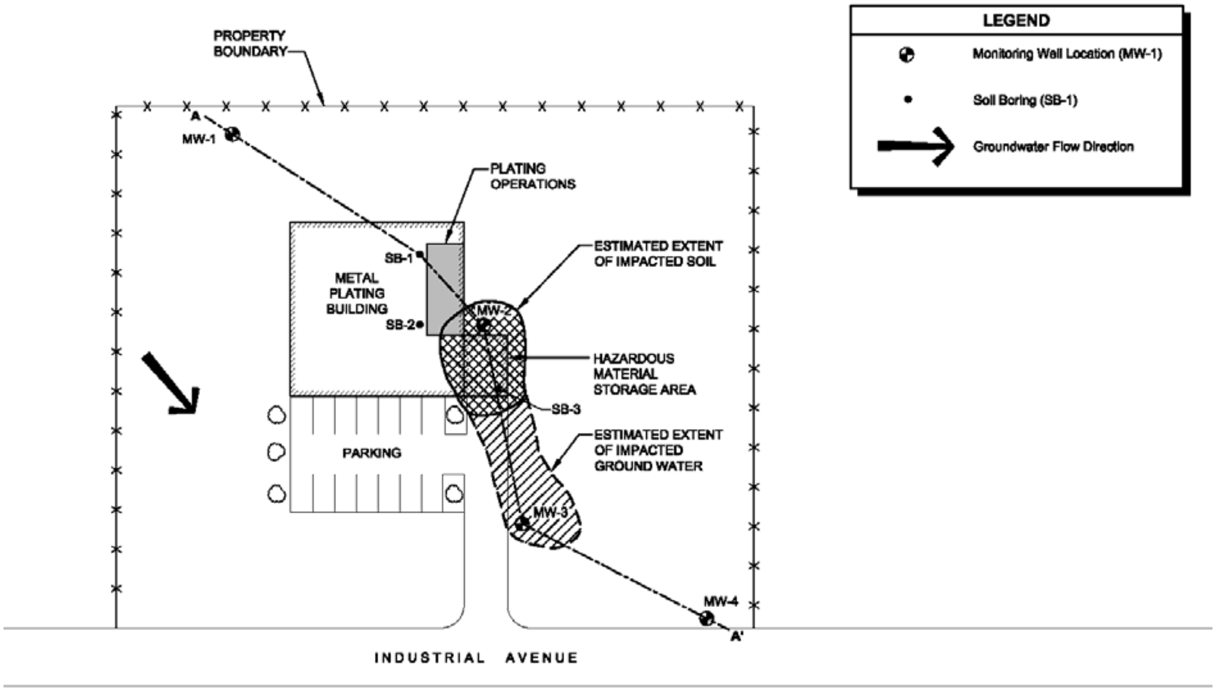


FIGURE 3-1a. Example conceptual site model for background analysis, plan view

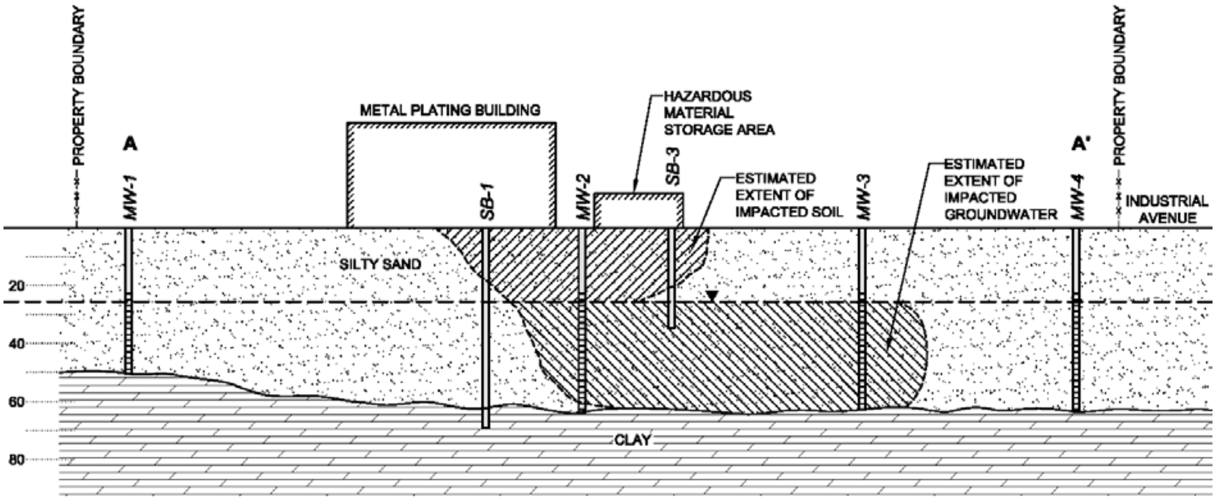


FIGURE 3-1b. Example conceptual site model for background analysis, cross sectional view

The CSM plan view should depict the following items:

- Locations of current site features (e.g., buildings, roads)
- Locations of relevant historical site features
- Locations of known and suspected contaminant sources
- Nature and extent of known surface and subsurface soil contamination
- Nature and extent of known groundwater contamination

- Surface water bodies
- Locations of monitoring wells and any other wells on the property and in the surrounding area
- Locations of subsurface utilities or other natural or artificial subsurface structures that could provide preferential pathways for contaminant migration with groundwater
- Water table elevation contours or groundwater flow directions
- Line showing trend of CSM cross section.

The CSM cross section should depict the following items:

- Rock and soil formations
- Saturated formations
- Water table and potentiometric surface elevations
- Monitoring well locations, depths, and screened intervals
- Groundwater flow directions
- Surface water bodies
- Preferential pathways for contaminant migration with groundwater.

Additional relevant data and information can be presented in tabular or narrative format, and should include:

- Descriptions of the soil and rock formations
- Hydraulic gradient, conductivity, porosity, and estimated groundwater seepage velocity for each saturated formation (Section 2.1.4)
- Geochemical and geotechnical parameters of the aquifer that affect chemical mobility, including, soil characteristics such as pH, TOC concentrations, and CEC (Section 2.1.4)
- Fate and transport properties of COPCs, including the chemical characteristics of COPCs that could significantly affect their partitioning between aqueous and non-aqueous phases and mobility in groundwater (e.g., ionic charge for inorganic chemicals, K_{ow} for organic chemicals) (Section 2.1.4).
- Utility of the groundwater (e.g., used for drinking water, irrigation)
- Narrative of site history.

3.2 Geochemical Evaluation

Data and information representing the geochemical characteristics of the groundwater and aquifer matrix are very useful to complement and support the evidence assembled during the hydrogeologic and contaminant source evaluation, and confirm that the background dataset is appropriate for comparison to site data. If the geochemical characteristics of the prospective background groundwater site are not similar to those of the investigation site, then the data should not be used for comparison to site data. Evaluation of groundwater geochemical characteristics can be used to avoid potential problems associated with collecting samples from an area that is apparently upgradient of the impacted area, but which is actually part of a separate groundwater regime. For example, aquitards, confining layers, and other geologic structures can separate groundwater bodies with very different geochemical characteristics.

If background sampling locations within the same aquifer or same local area as the actual or suspected contaminant plume cannot be located, then other prospective background locations must be identified. For this purpose, geochemical information and data should be used to confirm that groundwater conditions in the prospective background location are similar to site groundwater conditions. After this has been confirmed, the prospective background data can be used for comparison to site data.

3.2.1 Aquifer Matrix Geochemical Characteristics

The geochemical characteristics of the aquifer matrix materials should be evaluated to assess the behavior and transport of COPCs and background chemicals in both the vadose and saturated zones, and provide evidence to confirm that the background groundwater data are suitable for comparison to site groundwater data. In addition, soil sampling data (if available) should be reviewed to identify potential sources of groundwater background chemicals and COPCs in the local environment. As noted in Section 2.1.4, the aquifer matrix geochemical parameters that should be characterized include Eh, pH, CEC, and TOC.

3.2.2 Groundwater Geochemistry

As described in Section 2.1.4, groundwater geochemistry can be characterized by evaluating general chemical parameters and concentrations of the major dissolved components (i.e., the major cations and anions).

Evaluate Groundwater General Chemical Parameters

General chemical parameters (including pH, Eh, TDS concentrations, salinity, specific conductivity, dissolved oxygen concentrations, and temperature) should be quantified, tabulated, and plotted graphically to allow comparison among samples and groups of samples. Similarities and differences can be observed by inspection of the tables or graphs. Statistical comparison tests (Section 4.2) can be performed if necessary to confirm that differences between the samples representing two groups are statistically significant, i.e., the data indicate that the samples are most likely drawn from separate populations (e.g., two different groundwater bodies).

Evaluate Groundwater Major Ion Concentration Data

Concentrations of the major ions should be quantified to compare groundwater quality among samples. More than 90 percent of the dissolved solids in groundwater are composed of the following major ions:

- ❑ ***Cations:*** Sodium (Na^+), calcium (Ca^{2+}), potassium (K^+), magnesium (Mg^{2+})
- ❑ ***Anions:*** Sulfate (SO_4^{2-}), chloride (Cl^-), bicarbonate (HCO_3^-), and carbonate (CO_3^{2-})

Piper trilinear diagrams (Figure 3-2) are particularly useful to evaluate similarities and differences between samples or groups of samples (Piper, 1944). Other graphical formats, such as the Stiff diagram, also can be used to compare major ion characteristics among groundwater samples; however, these diagrams may not be convenient for graphical analysis of a large number of samples (Freeze and Cherry, 1979).

The Piper trilinear diagram plots each major ion concentration as a percentage of the sum of all the major ion concentrations (in milliequivalents) on two base triangles. The data points in the two base triangles are projected onto the center grid. Similarities and differences between groundwater samples can be identified by inspecting the diagrams. Data points that cluster together on the diagram represent groundwater samples with similar proportions of major ion concentrations, and therefore indicate that the samples are likely to be drawn from the same, or very similar, groundwater regimes. For example, the Piper trilinear diagram presented in Figure 3-2 indicates that proportions of the major ion concentrations in groundwater samples from MW-1, MW-2, and MW-3 are very similar to each other (within the chloride type field), and distinct from the MW-4, MW-5, MW-6, MW-7, and MW-8 samples (within the bicarbonate type field).

To construct a Piper trilinear diagram, concentrations of the major ions are converted to milliequivalents, and the percentage of the total number of milliequivalents is calculated and plotted for each ion. Details of the procedures for conversion to milliequivalents and construction of Piper trilinear diagrams can be found in Fetter (1994) and Freeze and Cherry (1979).

Statistical comparison tests (Section 4.2) can be performed if necessary to confirm that the major ion data indicate the samples are most likely drawn from separate populations. Note that the analysis of any mixture of waters A and B will plot on a straight line between points A and B in the Piper diagram plotting field, where points A and B represent the analyses of each of the two components (if the ions do not react chemically as a result of mixing) (Hem, 1992).

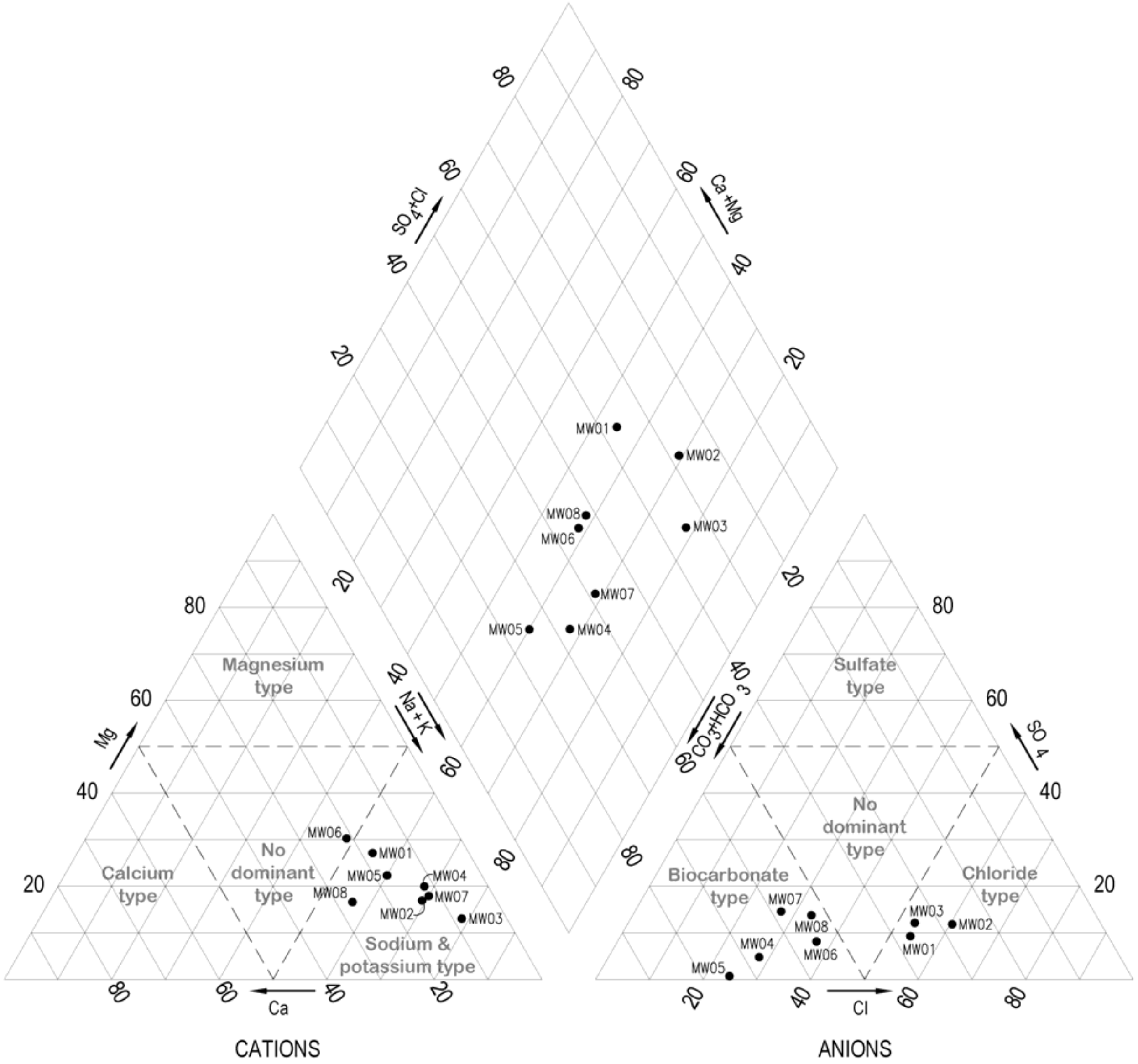


FIGURE 3-2. Example of Piper diagram

3.3 Spatial and Temporal Groundwater Data Analysis

Spatial and temporal groundwater data analysis involves the use of post-plots and graphs to evaluate concentration trends for the target COPCs with respect to space and time. By evaluating the post-plots and graphs in conjunction with the hydrogeologic, geochemical, and contaminant source data, investigators can distinguish between concentrations likely to represent background levels and concentrations that may be associated with a chemical release. If the dataset is large enough, probability plots of target COPC concentration data may also be useful to distinguish between populations likely to represent background conditions and populations that may represent contamination.

3.3.1 Post-Plots

Visual inspection of spatial plots can enable investigators to distinguish between impacted and unimpacted regions of the aquifer. Anomalies identified by inspection of the plot should be evaluated with respect to the hydrogeologic, geochemical, and contaminant source data to determine whether they could be associated with a chemical release, or could represent background conditions. For example, anomalously high COPC concentrations in groundwater samples collected downgradient of a suspected contaminant source location are most likely associated with impacted groundwater, whereas relatively low concentrations at locations upgradient or crossgradient of the source are likely to represent background conditions. A post-plot is a map of the site that has groundwater concentration data associated with each monitoring well location posted on it in order to assess potential anomalies in the spatial distribution of each target COPC. Post-plots that display clear spatial concentration trends, such as the example plot shown in Figure 3-3, can provide very strong evidence to distinguish between impacted and background groundwater. Apparent trends may be associated with hydrogeologic conditions such as the presence of an aquitard or other hydraulic discontinuity. Therefore, the hydrogeologic and geochemical data must be evaluated before making any conclusions, or assuming that the low-concentration data can be used to represent background conditions. Also, post-plots that display little or no spatial concentration trends may indicate that all concentrations represent background conditions. However, depending on the distribution of monitoring wells with respect to the contaminant source location, the lack of a trend could also indicate that all concentrations represent impacted groundwater, and that additional sampling (e.g., upgradient of suspected contaminant sources) is needed to delineate the contaminant plume.

3.3.2 Spatial/Temporal Concentration Plots

A spatial/temporal concentration plot displays the groundwater data distribution with respect to space and time, and thus shows both spatial relationships and the effect of seasonal variations on groundwater COPC concentrations. Separate spatial/temporal concentration plots can be constructed for the site and background datasets, or a single plot based on the combined dataset (i.e., both site and background data) can be constructed.

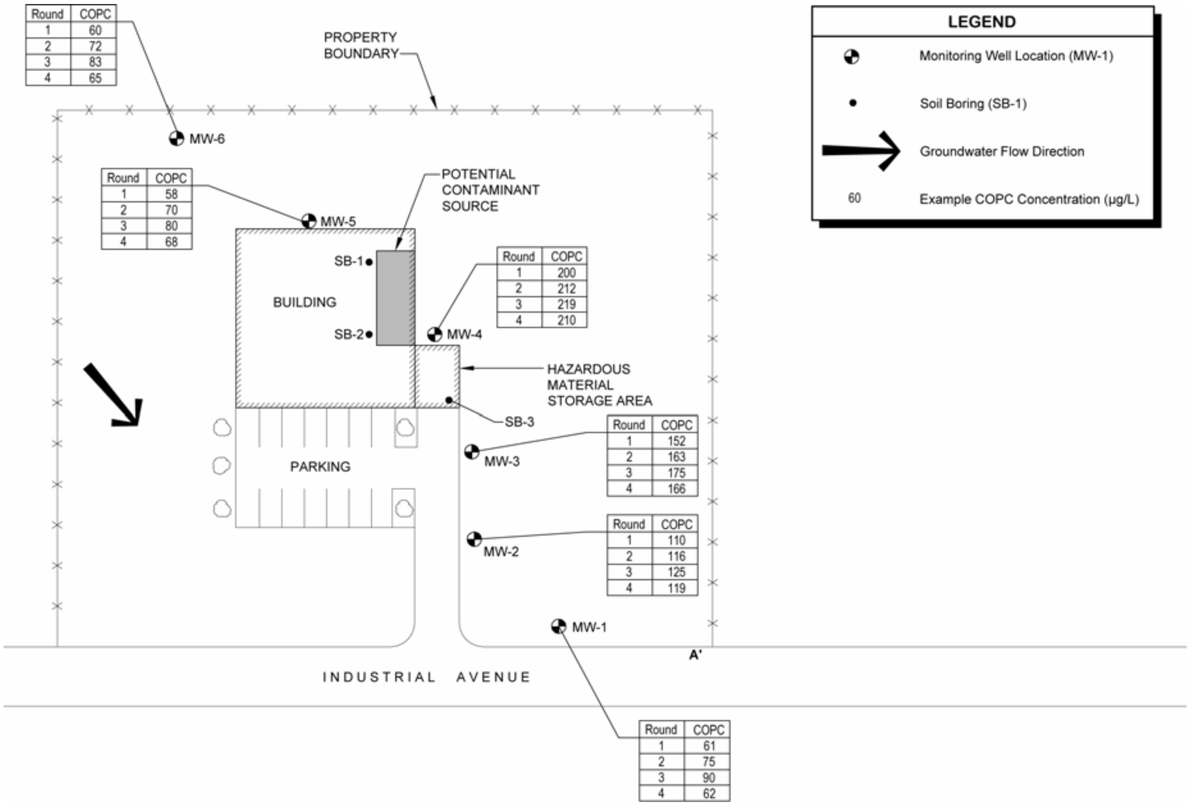


FIGURE 3-3. Example post-plot

Figure 3-4 presents an example of a spatial/temporal concentration plot. (Note that other graphical formats could also be adapted according to the user's preference to display the data required for spatial and temporal analysis.) This plot (and the corresponding data presented in Table 3-1) shows a clear spatial relationship, which suggests that MW-1, MW-5, and MW-6 (which exhibit consistently low concentrations) are outside of the impacted zone, and that MW-4 (which exhibits the highest concentrations) is located within or just downgradient of the source area. Concentrations then decrease in MW-2 and MW-3. Again, it must be noted that the apparent relationship may be associated with hydrogeologic conditions such as the presence of an aquitard or other hydraulic discontinuity; therefore, the hydrogeologic and geochemical data must be evaluated before making any conclusions, or assuming that the low-concentration data can be used to represent background conditions. If the MW-1, MW-5, and MW-6 groundwater data can be shown to represent the same groundwater body as the other well data, and these wells are not likely to be impacted by other COPC sources, then the associated data probably represent background conditions.

Inspection of the temporal trend indicates that the highest concentrations correspond to the winter (rainy season) sampling event. The increased concentrations detected at MW-1, MW-5, and MW-6 during the winter sampling event are likely to be associated with increased leaching of background chemicals from soil. Therefore, in this case, the evidence indicates that the highest concentration detected among the MW-1, MW-5, and MW-6 samples (90 µg/L) represents a reasonable estimate of the upper bound of the background concentration range. Evidence from the CSM should be presented to support this conclusion. As discussed in the following subsection, an intrawell comparison can be useful to confirm that the temporal variations are not evidence of contamination.

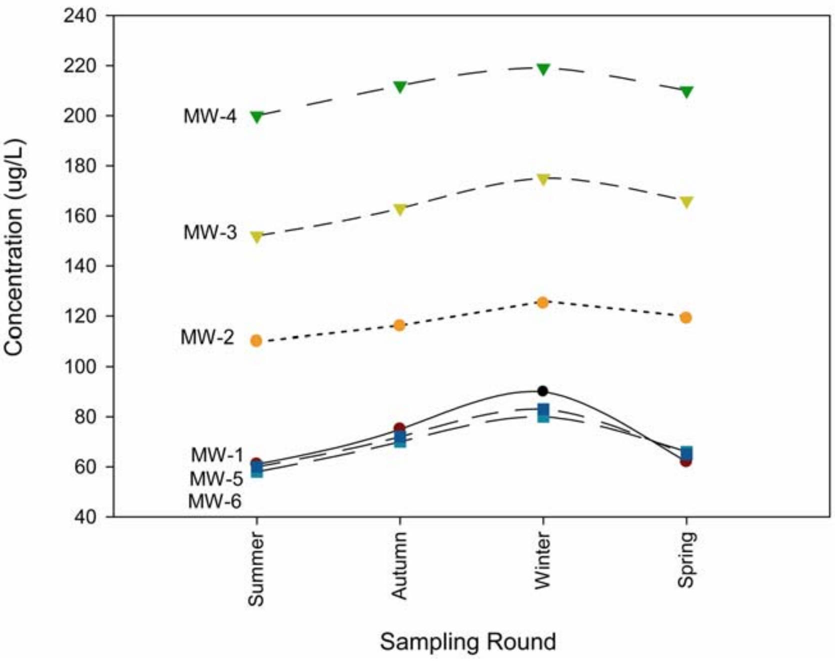


FIGURE 3-4. Example spatial/temporal concentration plot

3.3.3 Intrawell Comparisons

If sufficient data are available, intrawell comparisons can be performed to assess whether temporal concentration trends indicate that a particular monitoring well has been impacted by a site-related chemical release. Intrawell comparison refers to analysis of COPC concentration data collected from a single monitoring well over a series of sampling

TABLE 3-1. Example COPC concentration data

Sampling Event	MW-1	MW-2	MW-3	MW-4	MW-5	MW-6
Summer	61	110	152	200	58	60
Autumn	75	116	163	212	70	72
Winter	90	125	175	219	80	83
Spring	62	119	166	210	66	65

Note: All concentrations in µg/L.

events. One method, the combined Shewart-cumulative sum (CUSUM) control chart procedure (Gibbons et al., 1992; Gibbons and Discerning Systems, 1994), provides a statistical and visual tool capable of detecting both sudden and gradual changes in groundwater chemistry among samples collected from a single well. Statistically significant increases in concentrations over baseline conditions may be evidence that a chemical release has started to impact groundwater at a particular location.

The combined Shewart-CUSUM control chart procedure requires a minimum number of independent background concentration data points to provide a reliable estimate of the mean and standard deviation of background concentrations for the target COPC. Given the expected low variability of background concentrations, eight measurements (corresponding to eight independent sampling events) are usually adequate. Examples are provided in Gibbons (1994, page 161). Once these background data have been collected from the monitoring well of concern and plotted on the control chart, subsequent sampling results are statistically compared to the estimated control limit, in terms of both their absolute magnitude and cumulative sum.

The combined Shewart-CUSUM control chart procedure assumes the data are independent (i.e., sample independence) and normally distributed. The assumption of sample (data) independence requires that sufficient time has elapsed between sampling events (sampling interval) (Gibbons and Discerning Systems, 1994). A qualified hydrogeologist or groundwater scientist should evaluate site-specific hydrogeological factors (e.g., groundwater seepage velocity and contaminant transport velocities) to determine an optimal sampling interval that will ensure sample independence. If the analysis relies on data from previous sampling events, the sampling intervals should be evaluated to confirm that the samples can be considered independent. The assumption of normality is typically not a great concern because the data usually can be transformed adequately. Procedures for constructing and evaluating combined Shewart-CUSUM control charts are presented in Gibbons et al. (1992).

3.3.4 Probability Plots

A probability plot is a graph of concentration values plotted against their cumulative probabilities. Probability plots can be used to estimate background concentration ranges for target COPCs by identifying outliers and differentiating between separate populations within the dataset (e.g., a population that represents background conditions, and a population that represents contamination). However, if the dataset is small or the data are dependent (common occurrences in groundwater background analysis), it might not be possible to use a probability plot in order to distinguish between populations.

Probability plots can be constructed by plotting the data against quantiles of the hypothesized distribution on standard graph paper as described in Box 3-1, or by plotting cumulative probabilities (as percentages) on special probability paper as described in Box 3-2. However, probability plots for background analysis are usually plotted with the aid of a computer and statistical software program (e.g., U.S. EPA DataQUEST software [U.S. EPA, 1997]). Procedures for constructing a probability plot when the dataset contains multiple nondetect values are described in Box 3-3. Nondetect values are usually replaced with surrogate values. However, if the dataset contains a large number of nondetects, surrogate replacements may yield misleading results. In these cases, review by a statistician is recommended.

Use of Probability Plots to Evaluate Population Distributions

Probability plots are also useful for identifying the population distribution that best fits the data. If a background population is normally distributed, the data will plot along a straight line on a linear-scale normal probability plot. If a population is lognormally distributed, the logarithms or natural logarithms of the data will plot along a straight line on a linear-scale normal probability plot. Box 3-4 describes the use of probability plots to evaluate population distributions. Boxes 3-1 and 3-2 provide examples of the procedures used to construct a probability plot when the null hypothesis is that the data are normally distributed. These procedures,

BOX 3-1. Directions for constructing a probability plot (from U.S. EPA, 2000)

Let $x_1, x_2, \ldots, x_n$ represent the *n* data points. A normal probability plot of the data can be constructed as follows.

STEP 1: Order all the *n* data from smallest to largest and denote the ordered *distinct* (different) data values by $x_{(1)}, x_{(2)}, \ldots, x_{(n')}$, where n' may be less than n. For each distinct data value, compute the absolute frequency, AF_i. The absolute frequency is the number of times each distinct value occurs. If a data value occurs only once, the absolute frequency for that value is 1. If a data value occurs more than once, count the number of times the distinct value occurs. For example, consider the dataset 1, 2, 3, 3, for which $n = 4$ and $n' = 3$. The absolute frequency of value 1 is 1, i.e., $AF_1 = 1$. The absolute frequency of value 2 is 1, i.e., $AF_2 = 1$. But the absolute frequency of value 3 is 2, i.e., $AF_3 = 2$, as 3 appears two times in the dataset.

STEP 2: Compute the cumulative frequency (CF), for each of the n' distinct data values. The CF_i is the number of data points that are less than or equal to $x_{(i)}$, that is, $CF_i = \sum_{j=1}^{i} AF_j$. Using the data given in Step 1, the CF for value 1 is 1, the CF for value 2 is 2 (i.e., 1+1), and the CF for value 3 is 4 (i.e., 1+1+2).

STEP 3: Compute $Y_i = \frac{CF_i}{(n+1)}$ for each distinct data value.

STEP 4: Determine from the standard normal distribution (Table B-1) the quantile associated with each value of Y_i. Denote the quantile of the *i*th distinct data value by Z_i.

STEP 5: Plot the pairs (x_i, Z_i). If the plot of these points is well fit by a straight line, the data most likely fit a normal distribution. Otherwise, the data may be better fit by another distribution.

or a statistical software program, can also be used to test the null hypothesis that data are lognormally distributed, by plotting logarithms or natural logarithms of the data instead of the untransformed data against a linear scale. Alternatively, it is possible to test the null hypothesis that data are lognormally distributed by plotting the untransformed data against a log-scale.

Use of Probability Plots to Evaluate Background Conditions

The presence of multiple populations in a dataset results in a segmented probability plot. Therefore, probability plots can be useful for assessing whether measurements represent different populations. The combined dataset should be used to construct a probability plot. This may allow investigators to differentiate between data representing site groundwater and data likely to represent background conditions.

A continuous straight-line plot with no large gaps indicates that the data fit the hypothesized distribution, and represent a single population—most likely a naturally occurring population. Significant deviations or data gaps indicate that more than one population exists at the site, suggesting that contamination may be present. An inflection point or discontinuity in a probability plot may indicate the threshold separating two different populations in the dataset. Ambient (local background) conditions are usually conservatively defined as the range of concentrations associated with the low-concentration segment of the population. If concentrations in the upper range depart from the line or trend shown on the probability plot (i.e., a distinct increase in slope or discontinuity occurs), then the upper range of the concentration distribution represents a separate population. In this case, the lower-range concentrations are likely to represent background conditions, whereas the upper-range concentrations are likely to represent contamination. Figure 3-5 illustrates a probability plot with an inflection point at 10 µg/L, indicating that the upper bound of the background concentration range may be approximately 10 µg/L. This observation should be confirmed by evaluating the hydrogeological, geochemical, and contaminant source data (Section 3.1), as well as the spatial and temporal distribution of the data (Section 3.2). Details regarding use of probability plots to distinguish between populations representing background conditions and populations representing potential contamination are presented in Box 3-5.

BOX 3-2. Example: Constructing a probability plot by graphing cumulative percentages on probability plotting paper

Consider the following n = 14 data points that have been ordered from smallest to largest: 5, 6, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11, 11.5, 12, 13. To test the hypothesis that the data are normally distributed, construct a normal probability plot.

STEP 1: Because there are no duplicate values in the dataset, the AF of each value is 1.

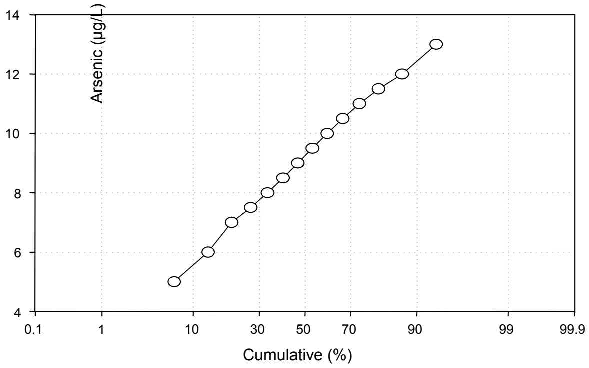
STEP 2: Compute the cumulative frequency for each data value as shown in the table below.

STEP 3: The cumulative percentages $Y_i = 100[CF_i / (n+1)]$ for each of the 14 distinct data values are shown in the last column of the table. The cumulative percentage associated with an individual data value is the probability (expressed as a percentage) that a randomly selected value from the dataset will be less than or equal to that individual data value.

STEP 4: Plot the n=14 pairs of X_i, Y_i on probability plotting paper.

It appears the plot is approximately linear; therefore, the data are assumed to be normally distributed.

i	Individual X_i	Absolute Frequency AF_i	Cumulative Frequency CF_i	Cumulative Percentage Y_i
1	5	1	1	6.7
2	6	1	2	13.3
3	7	1	3	20.0
4	7.5	1	4	26.7
5	8	1	5	33.3
6	8.5	1	6	40.0
7	9	1	7	46.7
8	9.5	1	8	53.3
9	10	1	9	60.0
10	10.5	1	10	66.7
11	11	1	11	73.3
12	11.5	1	12	80.0
13	12	1	13	86.7
14	13	1	14	93.3



BOX 3-3. Use of multiple nondetects in probability plots

Nondetect values usually are replaced with surrogate values; however, if the dataset contains large number of nondetects, surrogate replacements may yield misleading results. In these cases, review by a statistician is recommended. Possible alternatives for surrogate replacements are:

- Replace nondetects by one-half of the DL for each nondetect, or
- Replace nondetects with a dummy value at or below the lowest detected value.

Alternatively, Akritas et al. (1994, p. 227) and Michael and Schucany (1986, p. 476, equation 11.8) have developed statistical procedures for constructing probability plots when multiple nondetects are present. However, because these methods are somewhat complicated and have not been evaluated for constructing probability plots for background analysis, the assistance of a statistician is recommended.

3.4 Compile Lines of Evidence

The project team should compile and summarize the evidence provided by evaluating the hydrogeologic and contaminant source characteristics, geochemical characteristics, and the spatial and temporal COPC concentration distributions. The evidence should then be used to identify the monitoring wells that represent background conditions, and estimate the upper bound of the background concentration range for each target COPC.

BOX 3-4. Use of probability plots to evaluate population distributions

Probability plots often are used to visually evaluate the null hypothesis that the data are well fit (modeled) by a specified distribution. Frequently, the null hypothesis is that the dataset has either a normal or lognormal distribution; however, other distributions such as the Weibull and Gamma distributions (Gilbert, 1987, p. 157) sometimes are used. If the graph of plotted points in a probability plot appears linear with little scatter or deviation about the line, the results indicate that the data are well fit by the hypothesized distribution being tested. If the hypothesized distribution is the normal distribution, the data values are not transformed and are plotted on a linear scale (y-axis). If the hypothesized distribution is the lognormal distribution, the procedures are the same, except the logarithms or natural logarithms of the data are plotted, or a log-scale is used. A lognormally distributed population plotted on a log scale will yield a straight probability curve, as will a normally distributed population plotted on a linear scale. However, when a normally distributed population is plotted on a log scale, the curve will appear convex when viewed from above. Conversely, when a lognormally distributed population is plotted on a linear scale, the curve will appear concave when viewed from above.

Figure A is a probability plot constructed to test the null hypothesis that the data have a normal distribution. Note that the x-axis represents cumulative percentages for the standard normal distribution.

If a probability plot does not exhibit a linear pattern for the hypothesized distribution, the characteristics of the curve may indicate that the data fit another type of distribution. Three typical distribution characteristics will cause probability plots to deviate from a straight line: asymmetry (skewness), outliers, and heavy tails of the distribution. (Helsel and Hirsch [1992, pp. 30-33] describe these three conditions in detail.) If a probability plot is constructed on a linear scale to test the null hypothesis that the data are normally distributed, but the dataset is actually skewed to the right, the normal probability plot will be concave when viewed from above. If the dataset is skewed to the left, the graph will be convex when viewed from above when plotted on a linear scale.

The plotted points in Figure A form a concave curve, indicating that the dataset is skewed to the right. Because lognormal distributions are right-skewed, it is logical to test the hypothesis that the dataset is well fit by a lognormal distribution. Figure B shows a probability plot of the natural logarithms of the data. The plotted line is well fit by a straight line; therefore, it may be tentatively accepted that the data are lognormally distributed. However, this result can be checked by the Shapiro-Wilk W test discussed in Appendix A.1.

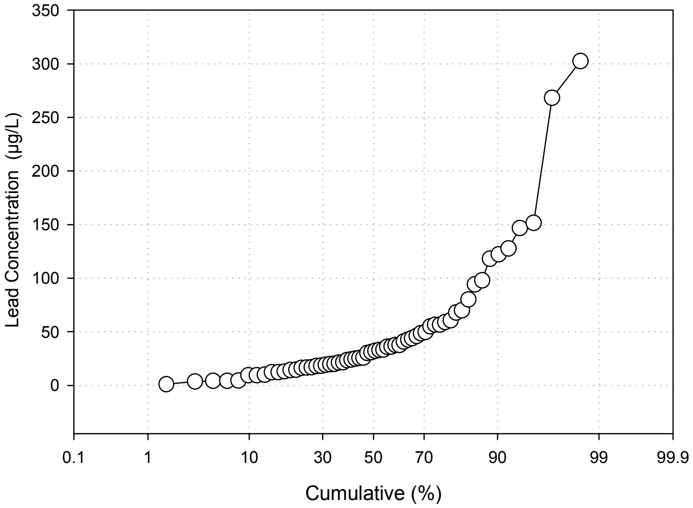


Figure A

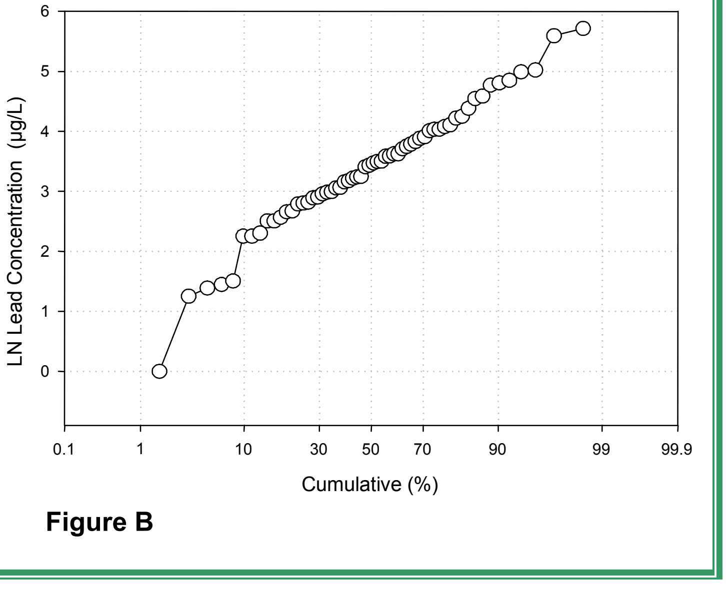


Figure B

3.5 Determine Acceptability of Background Wells and Concentration Ranges

After completing the background analysis process outlined in Sections 3.1 through 3.4, the project team must decide whether the analysis has produced technically defensible and reliable estimates of the background concentration ranges for each target COPC, and present the conclusions for stakeholder review. If the choice of background wells and estimated background ranges are not acceptable to stakeholders, then the deficiencies should be identified, additional data should be collected if necessary, and the background analysis steps should be repeated as required to address stakeholder concerns. The following criteria should be considered:

- The identified background ranges must be derived according to technically defensible procedures, and must be supported by adequate data.
- The identified background ranges must be consistent with the known physical and chemical characteristics of the groundwater environment (e.g., geology, hydrogeology, and geochemistry).
- The identified background ranges must be acceptable to stakeholders as representative of ambient chemical concentrations (i.e., total concentrations of both naturally occurring chemicals and anthropogenic chemicals not related to specific point sources or site releases).

If the above conditions are met, then the background analysis is completed for the target chemical. The estimated background concentration range should be documented for use in subsequent groundwater monitoring or risk evaluations, and the background dataset can be used for statistical comparison to concentrations representing site groundwater.

If the above conditions are not met for a target COPC, and background conditions must be characterized to comply with groundwater monitoring requirements or complete a risk assessment, then the background analysis should be repeated to correct the deficiencies. It may be necessary to collect additional data before proceeding with the analysis. If the inconclusive results are associated with a high percentage of U or UJ data points in the dataset, it may be necessary to reanalyze samples by a method that will yield lower RLs.

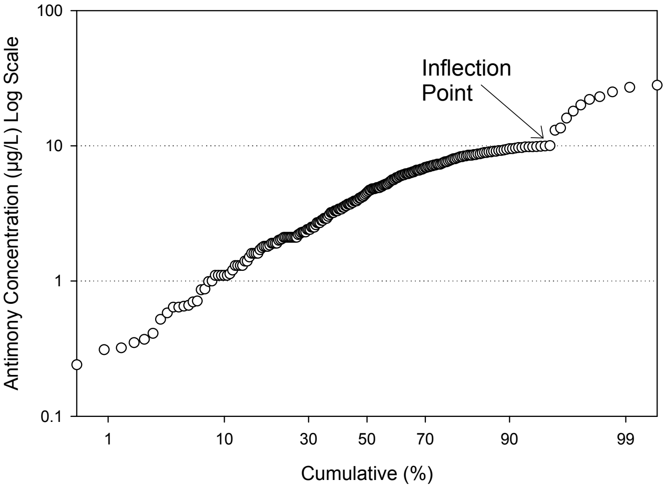


FIGURE 3-5. Example probability plot

3.6 Screen Site Data against Background Data and Evaluate the Need for Statistical Comparative Analysis

After estimated background ranges acceptable to stakeholders have been established, the maximum concentrations of each COPC should be screened against the estimated background ranges. If the maximum concentrations of all COPCs are within the estimated background ranges, then no further background analysis or statistical comparison is required. If the exposure concentrations exceed the upper bound of the background range, but the exceedances are not high enough to be obviously inconsistent with the background population, then the statistical comparative methods described in Section 4 should be implemented. If the statistical comparative testing reliably indicates that exceedances of an estimated background range are not statistically significant, then the background

BOX 3-5. Use of probability plots to identify background ranges

Groundwater datasets often contain both impacted and background measurements. The presence of these multiple populations in a dataset results in a segmented probability plot. Therefore, probability plots can be used to assess whether the measurements should be separated into different populations. An abrupt change in slope (inflection point) in a probability plot may signify the delimiter value separating two different populations in the investigated dataset. Singh et al. (1994) present a procedure for identifying inflection points as a means to evaluate background ranges.

The probability plotting method for background analysis typically involves one of the following cases:

- ***Single Populations:*** A nonsegmented probability plot with no inflection points indicates a single population, e.g., a background population. Note that the gradual curves obtained when a normally distributed population is plotted against a log-scale, and when a lognormally distributed population is plotted against a linear scale, do not contain inflection points.
- ***Background Delimiters:*** Segmented probability plots or probability plots with inflection points suggest the existence of multiple subpopulations, including possible outliers. In many cases, an inflection point at which the slope increases, or a break between a segment with a gradual slope followed by a segment with a steeper slope, represents the delimiter between the lowest concentration subpopulation (i.e., the background population) and a higher concentration subpopulation (i.e., potential contamination). This is a conservative approach, because the background range is viewed as a single population, represented by the lowest subpopulation. However, when multiple inflection points are evident on a probability plot, the upper bound of the background range may be considerably higher than the value associated with the lowest inflection point.
- ***Multiple Inflection Points:*** The background range may be composed of multiple natural or anthropogenic subpopulations. In these situations, the datasets will yield segmented probability plots with multiple inflection points, and the lowest inflection point will not represent the upper bound of the background range.
- ***Nondelimiting Inflection Points:*** Not all inflection points can be considered background delimiters. Specifically, if the subpopulation above an infection point forms a segment with a more gradual slope than the lower subpopulation, then the inflection point should not be considered a background delimiter. In this case, both segments can be treated as part of the same population.

Probability plots offer a simple way of graphically describing data and determining background ranges. However, to avoid any misleading conclusions, interpretation of these plots should always be supported by other accompanying analyses, as discussed in Sections 3.1 and 3.2. The *Handbook for Statistical Analysis of Environmental Background Data* (DON, 1999, Section 2.5.4) provides further discussion of the potential limitations of probability plots. For further discussion, including typical outcomes of probability plot analyses, readers are referred to Helsel and Hirsch (1992).

range estimate should be revised as necessary to reflect the maximum detected concentrations. (Note that this statement assumes that the site and background datasets are large enough to attain the desired test power. The desired power and adequacy of the datasets must be addressed within the DQO and DQA framework.) However, if COPC concentrations in site groundwater are considerably higher than the upper bound of the background range, and site knowledge and professional judgment indicate that the concentrations are consistent with a site-related release, then the time and resources required to implement the statistical comparative methods may not be justified.

3.7 References

Akritas, M.G., T.F. Ruscitti, and G.P. Patil. 1994. Statistical Analysis of Censored Environmental Data. In *Handbook of Statistics, Volume 12, Environmental Statistics*, eds. G.P. Patil and C.R. Rao. New York: North-Holland.

Department of the Navy (DON). 1999. *Handbook for Statistical Analysis of Environmental Background Data.* San Diego, CA: DON SWDIV and EFA West. July.

Fetter, C. W. 1994. *Applied Hydrogeology*. 3rd ed. Upper Saddle River, N.J.: Prentice-Hall.

Freeze, R.A., and J.A. Cherry. 1979. *Groundwater*. Englewood Cliffs, NJ: Prentice-Hall.

Gibbons, R.D. 1994. *Statistical Methods for Groundwater Monitoring.* New York: Wiley.

Gibbons R.D., N.E. Grams, F.H. Jarke, and K.P. Stoub. 1992. Practical Quantitation Limits. *Chemometrics and Intelligent Laboratory Systems* 12:225–235.

Gibbons, R.D., and Discerning Systems. 1994. *DUMPStat: An Artificially Intelligent Computer Program for Statistical Analysis of Groundwater Monitoring Data under Subtitle C and Subtitle D Regulations.* Chicago, IL: Robert D. Gibbons LTD (unpublished).

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring.* New York: Van Nostrand Reinhold.

Helsel, D.R., and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. New York: Elsevier.

Hem, J.D. 1992. *Study and Interpretation of the Chemical Characteristics of Natural Water*. 3rd ed. U.S. Geological Survey Water-Supply Paper 2254. Alexandria, VA: USGS.

Michael, J.R., and W.R. Schucany. 1986. Analysis of Data from Censored Samples. Pp. 461–496 in *Goodness-of-Fit Techniques*, eds. R.B. D'Agostino and M.A. Stephens. New York: Marcel Dekker.

Piper, A.M. 1944. A Graphic Procedure in the Geochemical Interpretation of Water Analyses. *Transactions, American Geophysical Union* 25:914–23.

Singh, A., A.K. Singh, and G. Flatman. 1994. Estimation of Background Levels of Contaminants. *Mathematical Geology* 26(3).

United States Environmental Protection Agency (U.S. EPA). 1997. *Data Quality Evaluation Statistical Toolbox (DataQUEST), User's Guide, EPA QA/G-9D, QA96 Version.* EPA/600/R-96/085. Washington, DC: U.S. EPA Office of Research and Development. December.

United States Environmental Protection Agency (U.S. EPA). 2000. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis.* EPA QA/G-9, QA 00 Update. EPA/600/R-96/084. Washington, DC: U.S. EPA Office of Environmental Information. July.

4. COMPARATIVE METHOD

4.1 Overview

Groundwater investigations and monitoring can require a variety of comparisons involving background data. The Comparative Method consists of statistical procedures for comparing chemical concentrations in site groundwater to chemical concentrations representing background levels. The appropriate type and scope of the statistical procedure depends on the objective of the intended comparison.

In general, the objectives of comparative statistical testing for groundwater can be divided into two broad categories: (a) site-wide comparison (i.e., characterization and assessment), and (b) individual comparison (i.e., detection monitoring). Statistical procedures appropriate for characterization and assessment of groundwater are fundamentally different from those suitable for detection monitoring. In addition, the outcomes of these procedures are quite different. This section identifies the comparative statistical test methods that are appropriate to achieve each objective, and presents guidance for implementing each test method.

4.1.1 Definition and Purpose of Site-Wide Comparisons

Site-wide comparison methods are appropriate for use in groundwater characterization and assessment investigations. In such investigations, available site data (i.e., measured concentrations in monitoring wells located downgradient of potential sources) are collectively compared to those measured at background wells in order to determine whether they belong to the same population. These procedures recognize that background concentrations represent a range of values, and not a single concentration. Thus, observed variability in the site data, including few elevated values, may well represent the expected variation in the background population.

Thus, for site-wide comparisons, statistical hypothesis tests are used to compare the site dataset to the background dataset. The results of these tests allow investigators to identify chemicals that occur at significantly higher concentrations in site groundwater than in background groundwater; these chemicals are declared COPCs. Uncertainty in this decision can be caused by factors such as a limited amount of groundwater data (as a result of inevitable resource constraints). The statistical testing methods account for uncertainty by quantifying the probability of making false conclusions based on the data. Similar methods are also applied in reference-based evaluations of post-cleanup conditions (U.S. EPA, 1994).

Site-wide comparisons are discussed in Section 4.2. The key questions addressed in Section 4.2 are:

- ❑ What statistical procedures or tests should be used to determine if a chemical is a COPC?
- ❑ What testing approaches should be avoided to reduce the probability of falsely concluding that a chemical is a COPC?
- ❑ How is the minimum number of measurements needed for the selected statistical test determined?
- ❑ How is the selected statistical test performed?

4.1.2 Definition and Purpose of Individual Comparisons

Individual comparison methods are specifically designed for groundwater detection monitoring efforts. In such investigations, a certain number of site monitoring wells (usually denoted as compliance wells) are subject to periodic sampling. At the conclusion of each sampling round, individual COPC measurements are simultaneously compared to the dataset representing background concentrations. These comparisons are intended to detect the release of COPCs in a timely manner. Therefore, unlike the site-wide comparison method, individual measurements are the focus of comparison rather than site-wide data.

Individual comparisons have a high probability of error, i.e., they falsely identify individual measurements as above-background exceedances, even when the site and background data belong to the same population. To address this deficiency and provide a reasonable approach to meet regulatory detection requirements, a series of statistical procedures involving verification resampling has been developed. The outcome of these statistical procedures determines whether a release has occurred (i.e., whether a COPC has been detected at a concentration which is indicated by the exceedance of a COPC above its background level). Such an exceedance, if verified, triggers further investigations to assess the extent of the detected release.

Individual comparisons are discussed in Section 4.3. The key questions addressed in Section 4.3 are:

❑ What statistical procedures or tests should be used to determine if an individual COPC measurement in a compliance well exceeds its background level?

❑ What testing approaches should be used to reduce the probability of falsely concluding that the background range for detecting a COPC has been exceeded?

❑ How is the minimum number of measurements needed for the selected statistical test determined?

❑ How is the selected statistical test performed?

4.1.3 Limitations of Recommended Statistical Procedures

All of the statistical tests discussed and recommended in Sections 4.2 and 4.3 are based on a number of assumptions that limit their applicability. Prior to any statistical comparisons, validity of these assumptions must be assessed. When these assumptions are not applicable, alternative procedures must be pursued. The main assumptions and their limiting factors are presented as follows:

❑ ***Site-Wide versus Individual Comparisons:*** The statistical tests discussed in Section 4.2 are specifically designed for collective comparison of site and background datasets. These tests are not applicable to cases of individual sample comparisons: Individual comparisons are addressed by different statistical procedures that are discussed in Section 4.3. Care must be taken to avoid any misapplication of the recommended tests.

❑ ***Independent versus Correlated Data:*** The groundwater samples should be collected far enough apart in space and time to minimize any spatial or temporal correlations among measured concentrations. In many instances, however, the available groundwater data will exhibit spatial and temporal correlations. Repeated sampling within a short time from the same well can result in strong temporal correlations. These correlations reduce the information value of the collected data. To avoid such information losses, Gibbons (1994) recommends that groundwater samples from the same well be taken at least three months apart. Spatial correlation also occurs where monitoring wells are located close together. In these situations, geostatistical procedures (Section 4.4) should be used to analyze the data and to assess their variability.

❑ ***Temporal Stationarity versus Variability:*** The statistical tests recommended in this section are based on the assumption that groundwater samples collected at different times are representative of a temporally stationary population (i.e., a population whose distribution is independent of time). This assumption applies at sites where background concentrations are expected not to vary with respect to sampling time. In instances where background concentrations display temporal variations, such as seasonal cycles or long-term trends, the nature of such trends should be determined prior to any statistical analysis. Various parametric and nonparametric techniques for trend analysis are discussed in Gibbons and Coleman (2001, Chapter 16) and Helsel and Hirsch (1992, Chapter 12). In cases of seasonal cycles, U.S. EPA (1992, page 5-10) proposes a simple method to compute long-term, season-specific mean concentrations, which are then used to “de-trend” the data prior to statistical testing.

- ***Spatial Stationarity versus Variability:*** The statistical tests recommended in this section are based on the assumption that groundwater samples collected from different wells are representative of spatially stationary populations, i.e., populations whose distributions are independent of space. This assumption is applicable to sites where background concentrations are expected to be invariant with respect to sampling location. The spatial variability of background concentrations can be assessed by analysis of variance (ANOVA) tests (Section 4.4). If the presence of a significant spatial variability among background wells is confirmed, intrawell procedures (Section 3.3.3) should be pursued.

- ***Unbiased versus Biased Coverage:*** The statistical tests recommended in this section are expected to perform best if based on the assumption that the available site and background datasets provide unbiased coverage of the investigated aquifer in space and time. In most cases, however, the available data are derived from monitoring wells that are clustered downgradient of suspected contamination sources. Furthermore, certain wells of interest may have been sampled far more frequently than others. In cases of large space-time clustering, alternative methods including ANOVA and geostatistics (Section 4.4) should be used.

- ***Test Reliability:*** The statistical tests recommended in this section are based on the assumption that adequate site and background datasets can be generated in order to attain the desired test confidence and power. If site and background measurements have already been collected and the budget does not allow for additional sampling, the desired confidence or power may not be attainable. In other words, the available site/background data by themselves may not contain enough information for the selected test to make a confident decision. In such cases it is highly recommended to include other reliable information, such as expert knowledge about hydrogeology and physical properties of the aquifer, as well as site history and operations, in order to enhance the reliability of the ensuing decisions.

4.2 Site-Wide Comparisons

Site-wide comparisons are used to determine whether the site and groundwater datasets are derived from the same statistical population. For this purpose, monitoring wells are divided into two groups, denoted as site and background wells. The locations of these wells are identified by the procedures described in Section 3. Site wells are located immediately adjacent and/or downgradient of the potential release sources, whereas background wells are usually situated upgradient, side-gradient, or distant downgradient of the site, where the potential for site-related impacts is negligible. If background sampling locations within the same aquifer or same local area as the actual or suspected contaminant plume cannot be located, then hydrogeological and geochemical information and data can be used to demonstrate that groundwater conditions in another area (usually a nearby area) are similar to site groundwater conditions, and that the background data are therefore appropriate for comparison to site data.

The statistical tests discussed in this section assume that spatial and temporal variations of background chemicals are negligible. Therefore, each site or background sample is viewed as equally representative of its respective population. The grouped site and background data are then compared to assess whether a detected chemical should be considered as a COPC or as a background chemical.

In addition to the assumptions and limitations discussed in Section 4.1.3, site-wide comparisons require that the site and background datasets are representative of hydrogeologically, geochemically, and anthropogenically similar aquifers. Furthermore, for optimum results, the two datasets should be approximately the same size and cover the same or similar time periods. If adequate background datasets do not exist, then appropriate reference areas, including upgradient, downgradient, and/or distant downgradient areas, must be identified for well installation and sampling. This process must consider the site aquifer system, especially the groundwater flow direction, and must be conducted in accordance with DQO principles. The scope of the background groundwater analyses also must be adequately comprehensive in order to demonstrate the geochemical, hydrogeologic, and anthropogenic similarity of the site and background groundwater samples. Concurrence of various stakeholders on the appropriateness and representativeness of the background dataset is necessary for successful application of site-wide comparisons.

4.2.1 Common Site-Wide Comparative Statistical Tests

An initial, tentative selection of the most appropriate statistical test(s) should be made during the DQO planning process. This selection should be based on: (1) the number of samples required for the various tests to achieve the specified DQO performance goals, (2) the particular distribution (normal or lognormal) expected of the data to be collected, (3) the likely spatial pattern of groundwater contamination, and (4) information in published statistical papers that demonstrate the performance of the candidate tests for various data distributions and contamination scenarios. After all data have been collected and exploratory data analyses have been conducted as discussed in Sections 2 and 3, a final selection of the statistical test(s) can be made.

To aid the user in selecting the most appropriate statistical test(s), the assumptions, advantages and disadvantages of each test discussed in this chapter are provided in Table 4-1. In this regard, selection of the optimal site-wide comparison test depends in part on the spatial pattern of groundwater contamination which is expressed as a hypothesis, such as:

- ❑ ***Site-wide impact:*** All site measurements from the targeted aquifer are likely to be impacted, and thus, their distribution is simply shifted to higher values than the distribution of background measurements. This scenario is encountered if the groundwater contaminant plume extends across the entire site. In this case, the difference between the two datasets can be assessed by comparing their mean or median concentrations.
- ❑ ***Localized or hotspot impact:*** Only a small portion of the distribution of site measurements is expected to be higher than the distribution of background measurements. This scenario is encountered if the groundwater contaminant plume underlies only portions of the site. In this case, the difference between the two datasets can be assessed by comparing their highest concentrations.

For the case of a site-wide impact, the Wilcoxon Rank Sum (WRS) test, the Gehan test, the two-sample t test, and the Satterthwaite two-sample t test are preferred. The Slippage test, Quantile test, and two-sample test of proportions are best suited to identify chemicals that have elevated concentrations in only a small portion of the site (i.e., localized or hotspot impact). If portions of the site aquifer system can be segregated into distinct zones and/or layers, the WRS, Gehan, or two-sample t test can be used to compare data from each zones or depth to the background dataset. The main features of these tests are discussed in Table 4-1 and subsequent parts of this section.

4.2.2 Statistical Testing Approaches Not Recommended for Site-Wide Comparisons

This section describes two methods for comparing site and background groundwater data that are not recommended to identify COPCs. These methods are not acceptable because, as shown in following paragraphs, the probability of error can be very high.

Comparing Maximum Site and Maximum Background Measurements

One approach to test whether a chemical is a COPC is to compare the maximum measurement among site data with the maximum background measurement, using the following decision rule:

TABLE 4-1. Assumptions and advantages/disadvantages of statistical tests to determine whether site concentrations are larger than background concentrations

Test Statistic	Objectives/Assumptions	Advantages	Disadvantages
Slippage Test	• Objective is to test for differences in the right tail (largest values) of the site and background concentration distributions. • More nondetects are allowed than for other tests considered in this section. • At least one detected (quantified) background measurement is present and it is larger than the largest nondetect value. • No assumptions are required with regard to the shape of site and background data concentration distributions.	• Very simple to conduct the test. • No distribution assumptions are necessary. • Many nondetects are permitted. • Can be used in conjunction (in tandem) with tests that focus on the detecting differences in the mean or median.	• May require a large number of measurements to have adequate power to detect differences in site and background concentrations.
Quantile Test	• Objective is to test for differences in the right tail (largest values) of site and background concentration distributions. • Below-detection values are not among the largest *r* data values in the pooled set of site and background data. • No assumptions are required with regard to the shape of the site and background data concentration distributions.	• Relatively simple to conduct the test. • No distribution assumptions are necessary. • Can have more power to detect differences in the right tail of site and background distributions than tests like the WRS, Gehan, or two-sample t tests that focus on the mean or median. • Can be used in conjunction (in tandem) with tests that focus on detecting differences in the mean or median.	• May require a large number of measurements to have adequate power to detect differences in site and background concentrations. • Test may be inconclusive if nondetects are present among the largest *r* data values.
Wilcoxon Rank Sum Test	• Objective is to test for differences in the medians of the site and background populations. • Only one RL (all nondetects have the same value), which is less than the smallest detected datum. • No more than 40% of both the site and background datasets are nondetects. • The site and background data concentration distributions have the same shape (variance).	• Nonparametric—i.e., no distribution assumptions necessary (however, the test is based on the assumption that the variance of the site distribution is the same as the variance of the background distribution). • In general, the test has more power to detect shift in site median than the two-sample t tests when the site and background data distributions are asymmetric (skewed to the right, to high concentrations). • Can be used in conjunction (in tandem) with Slippage and Quantile tests so that differences in the right tails of the site and background distributions, as well as differences in medians, can be detected. • Although manual calculations are labor intensive, this test can easily be performed using a number of readily available statistical software packages.	• Relatively more complex to compute by hand. • Too many nondetects prevent use of the test.

TABLE 4-1. (cont'd)

Test Statistic	Objectives/Assumptions	Advantages	Disadvantages
Gehan Test	• Objective is to test for differences in the medians of the site and background populations. • Nondetects do *not* have the same value (multiple RLs exist). • The censoring mechanism that generated the nondetects is the same for the site and background populations. • No assumptions are required with regard to the shape of the site and background data concentration distributions.	• Can be used when multiple RLs are present. • Same advantages as for the WRS test.	• Relatively complicated to compute by hand. • The performance of the test is not known as well as that of the WRS test. • Must assume the same censoring mechanisms apply to the site and background data.
Two-Sample t Test	• Objective is to test for differences in the means of the site and background populations. • Both site and background mean concentrations are normally distributed. • Below-detection-values have no significant impact on computed means (e.g., less than 15% of measurements are below detection). • The site and background data istributions have the same shape (variance).	• Most powerful test for detecting a shift in the site mean from the background mean, if the site and background data are normally distributed. • Certain transformations may be able to normalize the data. These transformations can be readily performed using statistical software.	• The test requires a statistical evaluation of the assumption of equal total variances for the site and background populations. • In general, the power will be less than that of the WRS test, if the data are not normally distributed. • Normal distribution assumption often is violated. • Outliers can affect the test results. • Not well suited for datasets that contain nondetects.
Satterthwaite Two-Sample t Test	• Objective is to test for differences in the means of the site and background populations. • Both site and background data have a normal distribution. • No nondetects are present. • Site and background data distributions are expected or known to have different shapes (i.e., unequal variances).	• Test can be used when the site and background distributions have unequal variances.	• The test is relatively complicated to compute by hand. • Same disadvantages as for the two-sample t test.
Two-Sample Test of Proportions	• Objective is to test for differences in the proportions of the site and background data above a given cutoff level. • Test may be used when more than 50% of the site or background datasets are nondetects. • No assumptions are required with regard to the shape of the site and background data concentration distributions.	• No distribution assumptions are necessary. • Relatively simple test to perform. • Can be used when many nondetects are present.	• A test based on proportions may not be what is really needed (e.g., it may be more appropriate to test for differences in means).

❑ If the maximum site measurement exceeds the maximum background measurement, then declare the chemical a COPC; otherwise declare the chemical not a COPC.

As discussed in O'Brien and Gilbert (1997), if the site and background datasets have the same concentration distribution and the same number of data points, the probability is 50% that the maximum measurement occurs in the site dataset and 50% that it occurs in the background dataset. Thus, the chance is 50% that the chemical will be declared to be a COPC, when in fact the chemical occurs at background levels in the site. The probability of erroneously declaring a chemical to be a COPC increases if the site dataset has more data points that the background dataset. In fact, if the site dataset has n measurements and the background dataset has m measurements, the probability of an incorrect decision is $p = n/(n + m)$. For example, if $n = 20$ and $m = 10$, then $p = 20/30 = 2/3$. In this case, the probability that this testing approach would erroneously identify a chemical as a COPC is 67%.

TABLE 4-2. Probabilities that one or more of *n* site measurements will exceed the 95th percentile of the background distribution if the site and background distributions are identical

n	$1 - (0.95)^n$
1	0.05
2	0.10
5	0.23
8	0.34
10	0.40
12	0.46
21	0.67
64	0.96

The above decision rule is not acceptable because: (1) the probability of incorrectly identifying a chemical as a COPC can be very high, and (2) correct determination of whether or not a chemical is a COPC is critically dependent on which area—the site or the background area—is represented by the most measurements. (Procedures for background analysis should not be confused with U.S. EPA or Navy risk assessment protocols: although maximum concentrations are typically compared to risk-based screening criteria in the initial phase of the risk assessment process, they should not be used to test whether a chemical is a COPC with respect to background.)

Comparing the Maximum Site Measurement to a Background Threshold

Another decision rule that might be used to decide if a chemical at the site is a COPC is:

- ❑ If one or more site measurements exceed the 95th percentile of the background data, declare the chemical a COPC; otherwise declare the chemical not a COPC.

Suppose the site and background distributions are identical and, thus, the chemical is not a COPC. Then, if the above decision rule is used, it can be shown that the probability that one or more of n site measurements will exceed the 95th percentile is equal to $1 - (0.95)^n$, where 0.95 is the probability that any randomly drawn (representative) site measurement is less than the 95th percentile of the background distribution. The expression $1 - (0.95)^n$ takes on the values shown in Table 4-2 for various values of n.

For example, if the background and site distributions are identical and $n = 21$, the probability that one or more of the site measurements will exceed the 95th percentile of the background distribution is 0.67. In other words, there is a 67% chance of falsely identifying a chemical as a COPC. If more extensive sampling is conducted at the site, for example, if $n = 64$, the probability of falsely concluding that the chemical is a COPC is 96%!

Other threshold values include the 90th or 99th percentiles, the background mean, two times the background mean, or an upper confidence limit on the background mean (U.S. EPA, 1998b). Regardless of which threshold value is selected, it will correspond to some percentile (perhaps unknown) of the background distribution. Therefore, no matter which threshold value is used, if site measurements are individually compared to the threshold value, the basic problem of excessive decision errors remains—only the specific probability of making an erroneous decision changes.

Based on the above discussion, it is recommended any background threshold or maximum comparison:

- ❑ Should only be considered as a means to suggest the need for additional investigation of whether or not a chemical is a COPC; and
- ❑ Should never be the only test applied to determine if a chemical is a COPC.

4.2.3 Recommended Site-Wide Comparative Statistical Tests

Unlike the background threshold and maximum comparisons discussed in Section 4.2.2, comparative statistical tests do not have the problems of elevated false decision error rates. Following are some general words of advice about using the Comparative Method to decide which chemicals are COPCs:

- ***Not Suitable for Detection Monitoring:*** The site-wide comparisons are mainly aimed at determining whether a given chemical is a COPC or a background chemical. These methods rely on the collective site and background datasets, and thus, are not suitable for individual COPC detection purposes. Procedures suitable for individual comparisons are discussed in Section 4.3.

- ***Assumption Verification and Limitations*:** The assumptions that underlie the site-wide comparative statistical hypothesis tests should always be reviewed. These reviews must be supported by the results of the exploratory data analyses presented in Sections 2 and 3. As discussed in Section 4.1.2, there are limitations that may require the use of alternative methods depending on the observed site and background conditions.

- ***Datasets*:** Background datasets should be comparable to the site data. Ideally, both datasets are unbiased and representative of geochemically, hydrogeologically, and anthropogenically similar aquifers. Furthermore, for optimum results, the two datasets should be nearly the same size and correspond to the same (or a similar) time period.

- ***Preferred Site-Wide Tests:***

  Use the nonparametric Slippage test (Section 4.2.4) as a quick method to decide which chemicals are COPCs by comparing extreme measurements from the site to the maximum background measurements. This test is expected to perform well if an impacted groundwater zone exists and it is limited to only a portion of the site (i.e., the localized or hotspot impact hypothesis).

  Use the nonparametric Quantile test (Section 4.2.5) if an important criterion for deciding which chemicals are COPCs is whether the extreme concentrations in the site are higher than the extreme background concentrations. Similar to the Slippage test, the Quantile test is expected to perform well if an impacted groundwater zone exists and it is limited to only a portion of the site.

  Use the nonparametric Wilcoxon Rank Sum test (Section 4.2.6) to compare median site measurements to median background measurements and thus decide which chemicals are COPCs. The WRS test is expected to perform well if an impact has occurred and it covers the site more or less uniformly (i.e., the site-wide impact hypothesis).

  Use the nonparametric Gehan test (Section 4.2.7) instead of the WRS test if the background or site datasets contain multiple RLs. Similar to the WRS test, the Gehan test is expected to perform well if an impact has occurred and it covers the site more or less uniformly.

  Use the two-sample t test (Section 4.2.8) if the mean concentrations of the background and site datasets can be assumed to be normally distributed with about the same variance, and if very few or no nondetects are present. This test is expected to perform well if an impact has occurred and it covers the site more or less uniformly.

  Use the Satterthwaite two-sample t test (Section 4.2.9) if the mean concentrations of the background and site datasets are assumed to be normally distributed with different variances, and if very few nondetects are present. This test is expected to perform well if an impact has occurred and it covers the site more or less uniformly.

  Use the nonparametric two-sample test of proportions (Section 4.2.10) if more than 50% of the background or site measurements are nondetects. This test focuses primarily on the portion of measurements in excess of a given cutoff value.

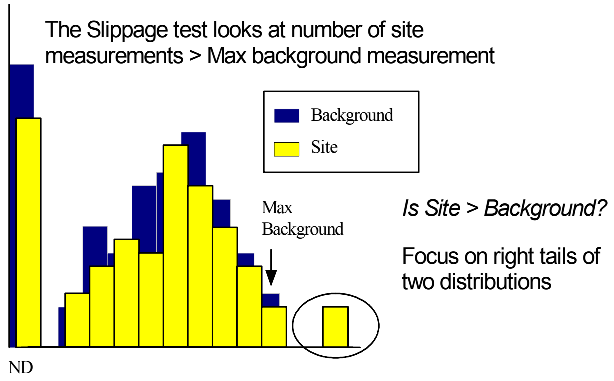
- ***Most Common Test Combinations*:** Often, the Slippage and WRS tests are used in tandem for site-wide comparisons. First, these two tests are nonparametric, and thus allow for the occurrence of more nondetects without any specific statistical distribution assumptions. Second, the combined use of these two tests addresses both the localized and site-wide impact hypotheses.

- ***Direct Statistician Involvement:*** Consult an experienced environmental statistician whenever disputes regarding the most appropriate statistical testing methods arise. In some cases, the characteristics of the site and background datasets may violate the assumptions required for the proposed statistical test methods. In these cases it may be necessary to implement the advanced methods described in Section 4.4, including geostatistical procedures to accommodate spatially correlated data. Under such conditions, involvement of an experienced statistician or geostatistician is essential.

4.2.4 Slippage Test

Groundwater Contamination Scenario

Site history information indicates that operations may have released a chemical into the site groundwater. The particular chemical of interest also is known to occur naturally and/or anthropogenically in groundwater. The decision question is: are concentrations of this chemical within the site greater than those in the background area? If so, the chemical will be considered a COPC. Knowledge of site operations suggests that if releases of the chemical did occur, the chemical is likely to have impacted only a portion of the site aquifer (i.e., localized impact).



Role of the Data Quality Objectives Process

The stakeholders and regulators used the DQO planning process to agree:

- On the methods that should be used to collect, handle, prepare, and measure the chemicals in groundwater samples.
- On the fact that a large number of measurements are or will be nondetects.
- On the usage of larger measurements among the site and background data as the primary criteria to decide whether a chemical is a COPC.
- On the values of the design parameters used to determine the required minimum number of site and background groundwater measurements (see the following subsection entitled "Guidance on Implementing the Slippage Test").

The Slippage test is appropriate for this scenario because it uses only the largest few data values and does not require any assumptions about the underlying distributions of the site and background measurements. The assumptions behind the Slippage test are summarized in Table 4-1.

Advantages and Disadvantages

- The Slippage test involves counting the number of site measurements that exceed the largest background datum and comparing that count with a critical value from a special table (see Box 4-1). Therefore, the Slippage test is extremely easy to conduct.
- The Slippage test considers only the largest background measurement and the largest measurements among the site data. Therefore, it is critical to verify that these elevated values do not represent mistakes or errors made during sample collection, handling, measurement, or data processing. Statistical tests for outliers (Appendix A.3) can be used for this purpose. If the extreme values prove to be outliers due to mistakes or errors, they should be eliminated. In particular, it is

BOX 4-1. Procedure for conducting the Slippage test

1. Specify the tolerable probability, α (also known as test significance), that the Slippage test will incorrectly declare that the site concentrations tend to be higher than the background concentrations, i.e., the probability that the chemical will be incorrectly identified as a COPC. α can be set only at 0.01 or 0.05, because critical values for the test are available only for these two α values (Step 7 below). Note: When both the Slippage test and the WRS test are conducted, the α level of the combined tests will be approximately the sum of the α levels selected for each test.
2. Specify the values of ε (the proportion of a site within which chemical concentrations are substantially greater than background levels) and of the power $(1 - \beta)$ the stakeholders and regulators have decided are important for the Slippage test.
3. Determine the approximate number of required measurements from Table 4-3.
4. Make sure that at least n samples from the site and m samples from the background area are available. If necessary, collect additional data, and analyze each sample for the chemical of interest. Some of the measurements may be nondetects.
5. Determine the value of the largest *detected* background measurement. In making this determination, ignore all nondetects that may be present in the background dataset.
6. Count the number, K, of detected site measurements that are larger than the largest detected background measurement. In making this determination, ignore all nondetects in the site dataset.
7. If α was set at approximately 0.01, determine the critical value K_c from Table B-2. If α was set at approximately 0.05, determine K_c from Table B-3. Note that the value of K_c depends on n and m ($n = m$).
8. If K is larger than the critical value K_c, declare that the site concentrations for the chemical of interest tend to be larger than the background concentrations for that chemical, i.e., the chemical is a COPC.

highly recommended to examine the suspiciously lager values, even if the outlier test does not qualify these values as outliers.

- In general, the Slippage test is designed to address the localized or hotspot impact hypothesis. If the exact nature of groundwater contamination is unknown, other tests, such as the WRS test, should be used in conjunction with the Slippage test to assess both the site-wide impact hypothesis and the hotspot impact hypothesis.
- The Slippage test can be viewed as a quick test to assess the COPC-status of a chemical.
- Table 4-1 provides a summary of the advantages and disadvantages of the Slippage test.

Guidance on Implementing the Slippage Test

The Slippage test procedure is shown in Box 4-1. Boxes 4-2 and 4-3 provide two examples of its use. The first step in implementing the Slippage test is to determine the minimum number of site and background measurements, n and m, respectively, required for the test to have adequate power to correctly declare that the chemical of interest is a COPC. The required values of n and m depend not only on the required power, but also on the following design parameters:

- The proportion, ε, of the site data that has concentrations greater than background.
- The magnitude of the difference between site area concentrations and background area concentrations.
- The tolerable probability, α (or test significance), that the Slippage test will declare the chemical to be a COPC when in fact it is not a COPC.
- The underlying distributions (for example, normal or lognormal) of the site and background concentration measurements.

BOX 4-2. Example 1 of the Slippage test

1. Suppose $\alpha = 0.01$ is selected.
2. Suppose $\varepsilon = 0.50$ and a desired power of 0.80 are selected.
3. The approximate minimum number of measurements needed is $n = m = 10$ (from Table 4-3).
4. Suppose the following representative measurements of the chemical of interest are obtained (listed in order from smallest to largest):

   Background Data: 23, 36, 37, 37, 44, 57, 60, 61, 61, 79
   Site Data: 15, 15, 20, 29, 30, 39, 60, 89, 90, 100
5. The value of the largest background measurement is 79.
6. $K = 3$ detected site measurements are larger than 79.
7. Using Table B-2 with $n = m = 10$, we find the critical value K_c is 6.
8. Therefore, the Slippage test declares that the evidence is insufficient to declare that the chemical is a COPC because $K = 3$ is not larger than $K_c = 6$.
9. However, do *not* conclude that the chemical is *not* a COPC. Instead, also conduct the WRS test (Section 4.2.6) on these data.

BOX 4-3. Example 2 of the Slippage test

1. Suppose $\alpha = 0.05$ is selected.
2. Suppose $\varepsilon = 0.30$ and a desired power of 0.80 are selected.
3. The approximate minimum number of measurements needed is $n = m = 15$ (from Table 4-3).
4. Suppose the following 30 representative measurements of the chemical of interest are obtained (listed in order from smallest to largest):

   Background Data: <3, <3, <4, <7, <7, <8, 8, 15, <16, <16, <17, <17, <22, <24, 25
   Site Data: <5, <10, 11, 13, <22, 23, <24, <36, <40, 70, 89, <100, 115, 200, <300
5. The value of the largest detected background measurement is 25.
6. $K = 4$ detected site measurements are larger than 25. Note that nondetects with elevated RL are ignored.
7. Using Table B-3 with $n = m = 15$, we find the critical value K_c is 4.
8. Therefore, the Slippage test declares that the evidence is insufficient to declare that the chemical is a COPC because $K = 4$ is not larger than $K_c = 4$.
9. Normally, the WRS test also would be performed to complement the results of the Slippage test. However, the datasets contain so many nondetects the WRS test cannot be computed (see Section 4.2.3). The Gehan test (Section 4.2.7) should be used in place of the WRS test.

Little information is present in the scientific literature concerning the best values of n and m for use in the Slippage test. However, Gilbert and Simpson (1990) provide enough information to construct a table for this guidance document (Table 4-3) that provides the approximate minimum number of measurements, n and m (for $n = m$) that should be used in the Slippage test to achieve a power (probability) of approximately 0.80 and 0.90 for various values of ε. These results are for the case where the value selected for α is between 0.025 and 0.05. Additional information on the power of the Slippage test is given in Gilbert and Simpson (1990, Figure 3). If a value of α smaller than 0.025 is selected, the number of samples in Table 4-3 would have to be increased

TABLE 4-3. Minimum number of measurements (*n* and *m, n = m*) required by the Slippage test to achieve a power of approximately 0.80 or 0.90 when a proportion, ε, of the site has concentrations substantially larger than background concentrations

	Number of Required Measurements (n and m)	
	Power	
Selected Proportion	0.80	0.90
ε = 0.10	60	75
ε = 0.15	40	50
ε = 0.20	30	35
ε = 0.25	25	30
ε = 0.30	15	25
ε = 0.35	15	20
ε = 0.40	15	20
ε = 0.45	10	15
ε = 0.50	10	10
ε = 0.60	10	10

Adapted from: Gilbert and Simpson (1990, Table 1 and Figure 3).

for the Slippage test to retain a power of 0.80 or 0.90. If a value of α larger than 0.05 is selected, the number of samples in Table 4-3 could be decreased somewhat and the Slippage test would still have a power of 0.80 or 0.90.

4.2.5 Quantile Test

Groundwater Contamination Scenario

The groundwater contamination scenario described for the Slippage test also applies to the Quantile test. The decision question is: are concentrations of the target chemical within the site greater than the background concentration values? If the target chemical concentrations exceed background, the chemical should be declared a COPC. The knowledge of site operations indicates that the extent of any potential impact is localized. This situation suggests that the Quantile test is appropriate.

Role of the Data Quality Objectives Process

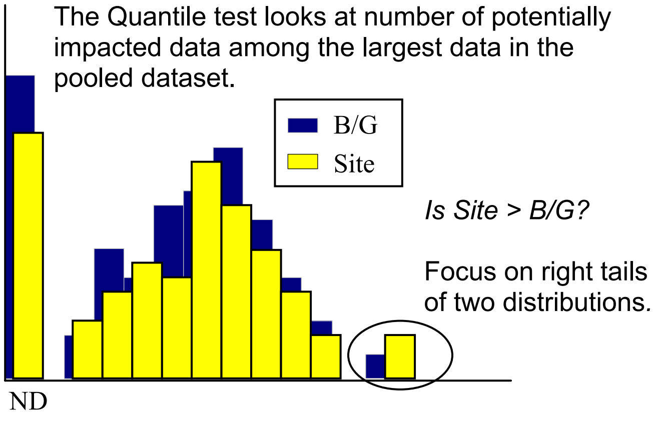
The stakeholders and regulators used the DQO planning process to agree:

- On the methods that should be used to collect, handle, prepare, and measure the chemicals in the groundwater samples.
- On the fact that a large number of measurements are or will be nondetects.
- On the values of the design parameters used to determine the required minimum number of site and background groundwater measurements (see the following subsection entitled “Guidance on Implementing the Quantile Test”).

The Quantile test is appropriate for this scenario because (1) it is a valid test regardless of the underlying distributions of the site and background datasets, and (2) the test looks for differences in the right tails of the site and background concentration distributions. The assumptions behind using the Quantile test are summarized in Table 4-1.

Advantages and Disadvantages

- The Quantile test is closely related to the Slippage test. It consists of looking at the largest *r* measurements in the pooled site and background datasets, and counting the number of those measurements that are from the site. If *k* or more of the *r* measurements are site measurements, the Quantile test declares the chemical to be a COPC.



The Quantile test looks at number of potentially impacted data among the largest data in the pooled dataset.

Is Site > B/G?

Focus on right tails of two distributions.

- Any number of nondetects is permitted in the site and background datasets, as long as all nondetects are smaller than the smallest of the *r* largest detected measurements in the pooled dataset.
- Use of the Quantile test does not require knowledge of the underlying concentration distribution of the chemical of interest.
- The Quantile test focuses on comparing the right tails of the site and background distributions rather than comparing the medians or means of the two distributions. This implies that the Quantile test is appropriate to address the localized impact hypothesis. If the nature of the impact is unknown, the Quantile test should be used in tandem with the WRS test. The WRS test addresses the site-wide impact hypothesis by looking for differences in the medians.
- Table 4-1 provides a summary of the advantages and disadvantages of the Quantile test.

Guidance on Implementing the Quantile Test

The Quantile test procedure is shown in Box 4-4, and Boxes 4-5 and 4-6 provide two examples of its use. As with other tests discussed in this document, the first step in implementing the Quantile test is to determine the minimum number of site and background measurements, *n* and *m*, respectively. The minimum number of measurements depends on the desired power of the test to correctly declare a chemical a COPC. As with the Slippage test, the required values of *n* and *m* also depend on the following:

- The proportion, ε, of the site that has concentrations greater than the background.
- The magnitude of the difference between site concentrations and background concentrations.
- The tolerable probability, α (or test significance), that the Quantile test will declare

BOX 4-4. Procedure for conducting the Quantile test

1. Select the tolerable probability, α (also known as test significance), that the Quantile test will incorrectly declare that the site concentrations tend to be larger than the background concentrations. The probability α may be set at 0.01, 0.025, 0.05, or 0.10. Note: When both the Quantile test and the WRS test are conducted, the α level of the combined tests will be approximately the sum of the α levels selected for each test.
2. Specify the values of ε and of the power ($1 - \beta = 0.80$ or 0.90) desired for the test.
3. Use the values of ε and power specified in Table 4-4 to approximate the required number site and background measurements. Table 4-5 may be used if it is important to detect when a proportion ε of site concentrations is only slightly larger than background.
4. Make sure at least *n* samples from the site and *m* samples from the background area are available. If necessary, collect additional samples and analyze each sample for the chemical of interest. Some of the measurements may be nondetects.
5. List the pooled site and background measurements from smallest to largest. The total number of pooled measurements is $n + m$.
6. Using the values of *n* and *m*, use Table B-4, B-5, B-6, or B-7 (depending on whether α was selected to be 0.01, 0.025, 0.05, or 0.10, respectively) to find the values of *r* and *k* needed to conduct the Quantile test.
7. Determine from the ordered list of pooled site and background measurements whether or not *k* or more of the largest *detected r* measurements are site measurements. (Note: ignore any nondetects when determining the largest detected *r* measurements). If so, the Quantile test indicates that the chemical is a COPC. If not, the test indicates that the data are insufficient for the Quantile test to conclude that the chemical is a COPC, and the WRS test should be conducted.

BOX 4-5. Example 1 of the Quantile test

1. Suppose $\alpha = 0.05$ is selected.
2. Suppose $\varepsilon = 0.50$ is selected, and a power of 0.80 is needed to determine whether site concentrations are distinctly higher than background concentrations.
3. For these values of α, ε, and power, Table 4-4 indicates that a minimum of $n = m = 10$ measurements are required for the Quantile test.
4. Suppose the 20 measurements are as follows (the same data were used to illustrate the Slippage test in Box 4-2):

   Background Data: 23, 36, 37, 37, 44, 57, 60, 61, 61, 79
   Site Data: 15, 15, 20, 29, 30, 39, 60, 89, 90, 100

5. The 20 pooled and ordered background and site data are (*P* and *B* indicate site and background, respectively):

   P P P B P P B B B P B B P B B B B P P P
   15, 15, 20, 23, 29, 30, 36, 37, 37, 39, 44, 57, 60, 60, 61, 61, 79, 89, 90, 100

6. As $\alpha = 0.05$ was selected in Step 1, Table B-6 indicates that for $n = m = 10$ that $r = k = 4$.
7. Among the largest $r = 4$ measurements in the pooled measurements (79, 89, 90, and 100), 3 are from the site area. Therefore, because $3 < k$= 4, the Quantile test indicates the data are insufficient to conclude the chemical is a COPC.

BOX 4-6. Example 2 of the Quantile test

1. Suppose $\alpha = 0.01$ is selected.
2. Suppose $\varepsilon = 0.50$ and a power of 0.80 is needed to determine whether site concentrations are distinctly higher than background concentrations.
3. For these values of α, ε, and power, Table 4-4 indicates that $n = m = 15$ measurements are required for the Quantile test.
4. Suppose the data are as follows:

   Background Data: <3, <3, <4, <7, <7, <8, 8, 15, <16, <16, <17, <17, 22, <24, <25
   Site Data: <5, <10, 11, 13, <22, 23, <24, <36, <40, 70, 89, 100, 115, 200, 300

5. The 30 pooled and ordered background and site data are:

   B B B P B B B B P P P B B B B B P B P B P B
   <3,<3, <4, <5, <7, <7, <8, 8, <10 11, 13, 15, <16, <16, <17, <17, <22, 22, 23, <24, <24, <25,

   P P P P P P P P
   <36,<40,70, 89, 100, 115, 200, 300

6. As $\alpha = 0.01$ was selected in Step 1, Table B-4 indicates that for $n = m = 15$ that $r = k = 6$.
7. Among the largest $r = 6$ detected measurements (70, 89, 100, 115, 200, 300), all 6 are from the site dataset. Therefore, because k (i.e., 6) of the highest r (i.e., 6) measurements are from the site area, the Quantile test indicates that the chemical is a COPC.

that the chemical is a COPC when in fact it is not a COPC.

- The underlying distribution (e.g., normal or lognormal) of the site and background concentration measurements.

U.S. EPA (1994, Tables A.2, A.3, A.4, and A.5) provides information on the values of *n* and *m* required for the Quantile test to achieve a prescribed power to correctly declare that a chemical is a COPC. A portion of this information is summarized in Tables 4-4 and 4-5. These tables list the approximate number of measurements needed for the Quantile test to have a power of approximately 0.80 or 0.90 to correctly declare a chemical a COPC, for cases where the tolerable probability, α, of incorrectly declaring a chemical to be a COPC is 0.01, 0.025, 0.05, or 0.10. Stakeholders and regulators should agree on an acceptable value of α.

The numbers of measurements listed in Tables 4-4 and 4-5 were obtained by assuming that the data are normally distributed. If it is suspected that

TABLE 4-4. Minimum number of measurements (*n* and *m*, *n* = *m*) required by the Quantile test to achieve a power of approximately 0.80 or 0.90 when a proportion, ε, of the site concentrations are distinctly higher than background concentrations[(a)]

α	0.01		0.025		0.05		0.10	
Power	**0.80**	**0.90**	**0.80**	**0.90**	**0.80**	**0.90**	**0.80**	**0.90**
ε = 0.10	>100	>100	100	>100	80	100	55	70
ε = 0.20	55	60	40	40	35	40	25	35
ε = 0.30	25	30	20	25	20	20	15	15
ε = 0.40	20	25	15	20	15	15	10	15
ε = 0.50	15	20	15	15	10	10	10	10
ε = 0.60	10	15	10	10	10	10	10	10
ε = 0.70	10	10	10	10	10	10	10	10
ε = 0.80	10	10	10	10	10	10	10	10
ε = 0.90	10	10	10	10	10	10	10	10
ε = 1.0	10	10	10	10	10	10	10	10

(a) *n* = *m* were obtained for the case where the normal site concentration distribution is shifted to the right of the normal background concentration distribution by the amount Δ/s = 4 (U.S. EPA, 1994, Tables A.2, A.3, A.4, and A.5). α is the tolerable probability (selected by stakeholders and regulators) that the Quantile test will incorrectly declare that the chemical is a COPC. Δ/s is the amount (in units of standard deviation) that the distribution of 100ε% of the measurements from the site is shifted to the right of the distribution in the background area.

TABLE 4-5. Minimum number of measurements (*n* and *m*, *n* = *m*) required by the Quantile test to achieve a power of approximately 0.80 or 0.90 when a proportion, ε, of the site concentrations are somewhat higher than background concentrations[(a)]

α	0.01		0.025		0.05		0.10	
Power	**0.80**	**0.90**	**0.80**	**0.90**	**0.80**	**0.90**	**0.80**	**0.90**
ε = 0.10	>100	>100	>100	>100	>100	>100	>100	>100
ε = 0.20	>100	>100	>100	>100	>100	>100	>100	>100
ε = 0.30	>100	>100	>100	>100	>100	>100	>100	>100
ε = 0.40	>100	>100	>100	>100	>100	>100	>100	>100
ε = 0.50	>100	>100	>100	>100	>100	>100	>100	>100
ε = 0.60	>100	>100	>100	>100	>100	>100	>100	>100
ε = 0.70	>100	>100	100	>100	75	>100	70	>100
ε = 0.80	>100	>100	75	>100	60	>100	50	>100
ε = 0.90	>100	>100	60	100	50	100	40	100
ε = 1.0	>100	>100	50	75	50	75	30	75

(a) *n* = *m* were obtained for the case where the normal site concentration distribution is shifted to the right of the normal background concentration distribution by the amount Δ/s = 1 (U.S. EPA, 1994, Tables A.2, A.3, A.4, and A.5). α is the tolerable probability (selected by stakeholders and regulators) that the Quantile test will incorrectly declare that the chemical is a COPC. Δ/s is the amount (in units of standard deviation) that the distribution of 100ε% of the measurements from the site is shifted to the right of the distribution in the background area.

measurements are skewed to the right and perhaps have a lognormal rather than a normal distribution, the number of samples should probably be increased somewhat to achieve the 0.80 and 0.90 power levels. Furthermore, the measurement numbers listed in Table 4-4 are for cases in which approximately 85% of site concentrations are higher than background concentrations (i.e., distinctly higher site concentrations). The measurement numbers listed in Table 4-5 are for cases in which only approximately 5% of site concentrations are higher than background concentrations (i.e., somewhat higher site concentrations). The numbers of measurements listed in Table 4-5 are larger than in Table 4-4 because more measurements are necessary to achieve the same power when the concentration differences are small. The Quantile test can be computed using the software Environmental Stat for S-PLUS (see Appendix C.3).

4.2.6 Wilcoxon Rank Sum Test

Groundwater Contamination Scenario

The approximate extent of potential groundwater contamination has been determined by the stakeholders and regulators during the DQO process. Site knowledge indicates that if contamination exists, it is most likely uniformly distributed throughout the site rather than occurring as hotspots.

Role of the Data Quality Objectives Process

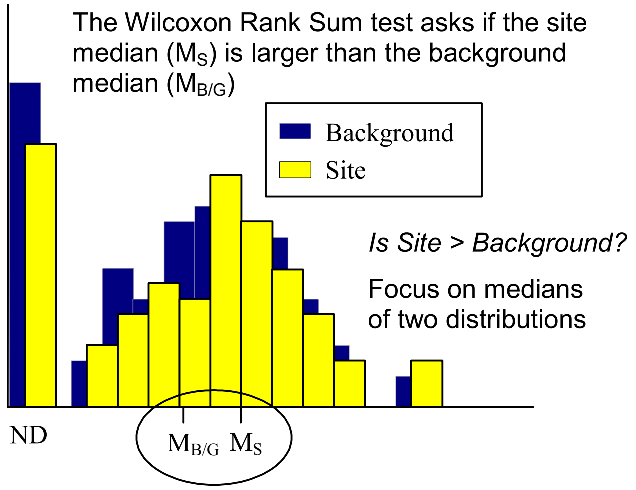
The stakeholders and regulators used the DQO planning process to agree:

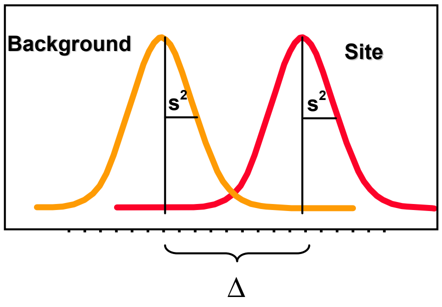
- On the methods that should be used to collect, handle, prepare, and measure the chemicals in groundwater samples.
- On the fact that no more than 40% of the measurements are or will be nondetects.
- On the values of design parameters used to determine the minimum number of site and background measurements (see the subsection entitled "Guidance on Implementing the WRS Test" that follows).

The WRS test is appropriate for this scenario because (1) it is a valid and reliable test regardless of the underlying distribution of the investigated data, and (2) the test determines if the difference between median concentrations of site and background data is statistically significant. The assumptions behind the WRS test are summarized in Table 4-1.

Advantages and Disadvantages

- Similar to the Slippage and Quantile tests, the WRS test is nonparametric. These tests do not require any specific assumptions about the exact form of the underlying data distributions.
- The power of the WRS test (for correctly detecting when the median site concentration is greater than the median background concentration) is known from theory and practice to be as high or higher than other statistical tests that evaluate differences in averages.





Is the site data distribution shifted to the right of the background data distribution by a significant amount Δ?

- The WRS test is based on the assumption that the underlying distribution of the site dataset has the same shape (variance) as the distribution of the background dataset. This assumption implies that the two distributions are the same, except that the site data distribution may be shifted to higher concentrations than the background data distribution. The assumption of equal variances should be evaluated using descriptive statistics and graphic plots of the site and background data.
- If nondetects occur, all of them must have the same RL, and that RL must be less than the smallest reported concentration. If multiple RLs are present throughout the set of measurements, then the Gehan test should be used instead of the WRS test.
- The WRS test should not be used if more than 40% of the site or background datasets are nondetects.
- The WRS test does not place great importance (weight) on the larger site and background measurements.
- The WRS test should be used in conjunction with the Slippage or Quantile Test so that both site-wide and localized contamination can be detected with adequate probability.

The software Environmental Stats for S-PLUS (see Appendix C.3) can be used to compute the WRS test. Table 4-1 provides a summary of the advantages and disadvantages of the Quantile test.

Guidance on Implementing the WRS Test

Box 4-7 describes the steps necessary to perform the WRS test when $n < 20$ and $m < 20$, and Box 4-8 provides an example of that procedure. Box 4-9 describes how to conduct the WRS test when $n \geq 20$ and $m \geq 20$, and Box 4-10 provides an example of that procedure. To implement the WRS test, determine the minimum number of site and background groundwater measurements, denoted by n and m, respectively. A formula for computing n and m is given in U.S. EPA (1994, Equation 6.3). This sample-size formula requires the following inputs:

- The tolerable probability, α, that the WRS test will incorrectly declare that the chemical is a COPC. Often, α is set at a value in the range of 0.01 to 0.10.
- The power (probability) the WRS test should have to correctly declare that the chemical is a COPC when that is in fact the case.
- The amount Δ/s (in units of standard deviation, s) by which the site median concentration exceeds the background median concentration that must be detected with the required power.
- The proportion of the total number of site and background groundwater samples that will be collected in the background area. If this proportion is 0.50, then $n = m$.

When $n = m$ is desired (the ideal case), a formula for determining the number of site and background measurements is given in MARSSIM (NRC et al., 1997, Equation 5-1, p. 5-28). However, rather than use the formulas in U.S. EPA (1994) or NRC et al. (1997), it is simpler to select n and m from Table 4-6 if it is desired to have $n = m$. The values of $n = m$ listed in Table 4-6 were obtained using Equation 5-1 in NRC et al. (1997) and then increasing that value by 20% to account for uncertainties and the likelihood that missing or unusable measurements will occur. The free computer software code Visual Sample Plan (VSP), developed by the U.S. Department of Energy's Pacific Northwest National Laboratory, also can be used to compute the number of samples for conducting a WRS test (see Appendix C.3).

4.2.7 Gehan Test

Groundwater Contamination Scenario

The groundwater contamination scenario developed by the stakeholders and regulators during the DQO process is the same as for the WRS test in Section 4.2.6: If contamination has occurred, it is most likely uniformly distributed throughout the site. However, in this case, the datasets are expected to contain multiple nondetects with different RLs.

BOX 4-7. Procedure for conducting the WRS test when the number of site and background measurements is small (*n*<20 and *m*<20)

1. Specify the tolerable probability, α (also known as test significance), that the WRS test will incorrectly declare that the site concentrations tend to be larger than the background concentrations, i.e., the probability that the chemical will be incorrectly identified as a COPC. Note: When both the WRS and Quantile tests are conducted, the α level of the combined tests will be approximately the sum of the α levels selected for each test.
2. Specify the value of Δ/s and of power, where Δ/s is the magnitude of the difference in median site and background concentrations that must be detected by the WRS test with the specified power. The notation Δ/s indicates the shift is expressed in units of standard deviation (s) of the underlying background site concentration distributions for the chemical of interest.
3. Use the specified values of α, Δ/s, and power in Table 4-6 to determine the number of site and background measurements needed when it is desired to have n equal to m. If equal n and m values are not desired, use Equation 6.3 in U.S. EPA (1994) and increase that value by 20% to guard against missing or unusable measurements.
4. If measurements are available from past sampling efforts, verify that their number is at least as large as the number indicated in Table 4-6. Collect additional samples, if necessary to achieve the required number of samples.
5. List and rank the pooled set of $n + m$ site and background measurements from smallest to largest, keeping track of which measurements came from the site and which came from the background area. Assign the rank of 1 to the smallest value among the pooled data, the rank of 2 to the second smallest value among the pooled data, and so forth.

   If a few measurements are tied (identical in value) assign the average of the ranks that would otherwise be assigned to the tied observations. If several measurement values are tied, average the ranks separately for each of those measurement values.

   If a few nondetects occur (say, <10%), and if all such values are less than the smallest detected measurement in the pooled dataset, handle the nondetects as tied at an arbitrary value less than the smallest detected measurement. Assign the average of the ranks that would otherwise be assigned to these tied nondetects (the same procedure as for tied detected measurements).

   If between 10% and 40% of the pooled dataset are nondetects, and all are less than the smallest detected measurement, use the WRS test procedure in Box 4-9, even if n and m are less than 20. Note: The procedure in Box 4-9 is for the case where m and n are both 20 or larger. That procedure provides only an approximate test if it is used when n and m are both smaller than 20. In that case, the test should not be used to decide whether the chemical is a COPC until additional information is obtained by collecting more samples and analyzing by a method with a lower RL.
6. Calculate the sum of the ranks of the site measurements. Denote this sum by R, then calculate W as follows:

   $$W = R - n(n + 1) / 2$$
7. Use the values of n and m and α to enter Table B-8 to find the critical value w_α, where α has been specified in Step 3 above. Table B-8 can be used only if α has been chosen to be 0.05 or 0.10.

   If $W > nm - w_\alpha$ the WRS test indicates the site concentration distribution is shifted to the right of the background concentration distribution, i.e., that the chemical is a COPC.
8. If the WRS test declares that the chemical is *not* a COPC, this conclusion may indicate (1) the chemical is indeed not a COPC, or (2) the assumptions that underlie the WRS test are not valid for the site and background measurements, or (3) an insufficient number of measurements (n and m) were obtained for the WRS test to detect the difference that actually exists in site and background concentration distributions.

BOX 4-7. (cont'd)

The possibility that the causes in items (2) or (3) may have resulted in the WRS test declaring that the chemical is not a COPC should be evaluated. Review the DQO planning process records to make sure the actual number of samples (n and m) corresponds to the number of samples determined necessary to detect a difference (shift) between the site and background medians of magnitude Δ/s. For case 3, the shift in the concentration distribution may in fact be smaller than the shift selected by the stakeholders, in which case no additional measurements are needed.

Also, update the estimated number of site and background measurements needed by using a software package such as Visual Sample Plan (Hassig et al., 2002). Collect additional samples if needed.

BOX 4-8. Example of the WRS test when the number of site and background measurements is small (*n*<20 and *m*<20)

Suppose it is necessary to determine if a chemical detected in groundwater at a Navy site is a COPC.

1. Suppose α was specified to be 0.05.
2. Suppose Δ/s and the power were specified to be 1.5 and 0.95, respectively. That is, the stakeholders and regulators specified that if the median of the site concentration distribution is greater than the median of the background distribution by the amount $\Delta/s = 1.5$, then enough measurements should be obtained so that the WRS test has a power of 0.95 of detecting that fact.
3. Using these values of α, Δ/s, and power to enter Table 4-6, we find that $n = m = 18$ measurements are needed for the WRS test.
4. Then, 18 samples from both the site and the background areas were collected using a suitable probability-based sampling design (for example, simple random sampling or sampling at the nodes of a square or triangular grid) and each sample was analyzed for the chemical of interest. Suppose the measurements were:

Background Data:	22,	32,	9,	12,	3,	7,	11,	2,	9,	11,	13,	16,	20,	25,	<1,	<1,	17,	21
Site Data:	24,	33,	5,	9,	36,	<1,	10,	50,	9,	19,	15,	10,	28,	9,	3,	15,	4,	19

5. Next, the data are pooled together and listed from smallest to largest. The ranks of the site data are determined (the site and background data and ranks are denoted by P and B, respectively):

	B	B	P	B	P	B	P	P	B	P	P	P	B	B	P	P	B	B
Data:	<1	<1	<1	2	3	3	4	5	7	9	9	9	9	9	10	10	11	11
Rank:	2	2	2	4	5.5	5.5	7	8	9	12	12	12	12	12	15.5	15.5	17.5	17.5

	B	B	P	P	B	B	P	P	B	B	B	P	B	P	B	P	P	P
Data:	12	13	15	15	16	17	19	19	20	21	22	24	25	28	32	33	36	50
Rank:	19	20	21.5	21.5	23	24	25.5	25.5	27	28	29	30	31	32	33	34	35	36

6. Sum the ranks of the site measurements to obtain $R = 2 + 5.5 + 7 + \ldots + 34 + 35 + 36 = 350.5$. Therefore,

$$W = R - n(n+1) / 2 = 350.5 - 18(19) / 2 = 179.5$$

7. Enter Table B-8 with $\alpha = 0.05$ and $n = m = 18$ to obtain $w_{0.05} = 110$.

   Computing $nm - w_\alpha = 18 \times 18 - 110 = 214$. Therefore, $W < nm - w_\alpha$ (i.e., $179.5 < 214$). The WRS has indicated the evidence is insufficient to declare the chemical is a COPC.

   As the WRS did not declare that the chemical is a COPC, the DQO process notes are reviewed to make sure the number of measurements specified to meet the α, Δ/s, and power requirements were indeed obtained. Also, to update the estimated number of site and background measurements needed, use a software package such as VSP (Hassig et al., 2002). If the number of samples computed using that equation exceeds the number used in the WRS test, collect the indicated number of new site and background samples.

BOX 4-9. Procedure for conducting the WRS test when the number of site and background measurements is large ($n \geq 20$ and $m \geq 20$)

1. Specify the tolerable probability, α (also know as test significance), that the WRS test will incorrectly declare that the site concentrations tend to be larger than the background concentrations, i.e., the probability that the chemical will be incorrectly identified as a COPC. Note: When both the WRS test and Quantile test are conducted, the α level of the combined tests will be approximately the sum of the α levels selected for each test.
2. Specify the value of Δ/s and of power, where Δ/s is the magnitude of the difference in median site and background concentrations that must be detected by the WRS test with the specified power. The notation Δ/s indicates the shift is expressed in units of standard deviation (s) of the underlying background and site concentration distributions for the chemical of interest.
3. Use the specified values of α, Δ/s, and power in Table 4-6 to determine the number of site and background measurements needed when it is desired to have n equal m. If equal n and m values are not desired, use Equation 6.3 in U.S. EPA (1994) and increase that value by 20% to guard against missing or unusable measurements.
4. If measurements are available from past sampling efforts, verify that their number is at least as large as the number indicated in Table 4-6. Collect additional samples, if necessary to achieve the required number of samples.
5. List and rank the pooled set of $n + m$ site and background measurements from smallest to largest, keeping track of which measurements came from the site and which came from the background area. Assign the rank of 1 to the smallest value among the pooled data, the rank of 2 to the second smallest value among the pooled data, and so forth.

   If <40% of the measurements in the pooled dataset are tied (identical in value) assign the average of the ranks that would otherwise be assigned to the tied observations. If several measurement values are tied, average the ranks separately for each of those measurement values.

   If <40% of the pooled dataset are nondetects, and if all such values are less than the smallest detected measurement in the pooled dataset, handle those nondetects as being tied at an arbitrary value less than the smallest detected measurement. Assign the average of the ranks that would otherwise be assigned to this group of tied values (the same procedure as for detected measurements that are tied). Note: The total number of tied detected measurements and tied nondetects should not exceed 40% of the total number of measurements.

   If more than 40% of the pooled data are nondetects, then do not use the WRS test. The Gehan test should be used instead (Section 4.2.7).
6. Calculate the sum of the ranks of the site measurements. Denote this sum by R.
7. Calculate

   $$w_{1-\alpha} = n(n+1) / 4 + z_{1-\alpha} [n (n + 1) (2n + 1)/24]^{1/2}$$

   where $z_{1-\alpha}$ is the $100(1-\alpha)$ percentile of the standard normal distribution, which is tabulated in Table B-1. For example, if $\alpha = 0.05$, then $z_{1-\alpha} = z_{0.95} = 1.645$ from Table B-1.
8. The WRS test declares that the chemical is a COPC if $R > w_{1-\alpha}$.

BOX 4-10. Example of the WRS test when the number of site and background measurements is large ($n \geq 20$ and $m \geq 20$)

1. Suppose α is specified to be 0.01.
2. Suppose Δ/*s* and the power were specified to be 1.8 and 0.95, respectively. That is, the stakeholders and regulators specified that if the median of the site concentration distribution is Δ/*s* = 1.8 units greater than the median of the background distribution, then enough measurements should be obtained so the WRS test has a power of 0.95 of detecting that fact.
3. Using these values of α, Δ/*s*, and power to enter Table 4-6, we find that $n = m = 20$ measurements are needed for the WRS test, where *n* and *m* are the number of site and background measurements, respectively.
4. Then 20 samples from both the site and the background areas were collected using a suitable probability-based sampling strategy, for example, simple random sampling. Suppose the measurements were (listed in increasing magnitude):

   Background Data:<10, <10, <10, <10, 12, 15, 15, 18, 22, 26, 27, 29, 29, 29, 55, 60, 77, 90, 101, 150

   Site Data: <10, <10, <10, 25, 27, 27, 36, 36, 99, 101, 103, 140, 145, 150, 180, 190, 199, 200, 250, 300
5. Next, the data are pooled together and listed from smallest to largest. Then the ranks of the site data are determined (the site and background data and ranks are denoted by *P* and *B,* respectively).

	B	B	B	B	P	P	P	B	B	B	B	B	P	B	B	P	P	B	B	B	P
Data:	<10	<10	<10	<10	<10	<10	<10	12	15	15	18	22	25	26	27	27	27	29	29	29	36
Rank:	4	4	4	4	4	4	4	8	9.5	9.5	11	12	13	14	16	16	16	19	19	19	21

	P	B	B	B	B	P	B	P	P	P	P	B	P	P	P	P	P	P	P
Data:	36	55	60	77	90	99	101	101	103	140	145	150	150	180	190	199	200	250	300
Rank:	22	23	24	25	26	27	28.5	28.5	30	31	32	33	34	35	36	37	38	39	40

6. The sum of the ranks of the site data is R = 4 + 4 + 4 + 13 + 16 + … + 39 + 40 = 507.5.
7. Also, $w_{0.99} = n(n+1) / 4 + z_{0.99} [n (n + 1) (2n + 1)/24]^{1/2}$

   $= 20(21) / 4 + 2.33[20(21)(41)/24]^{1/2}$

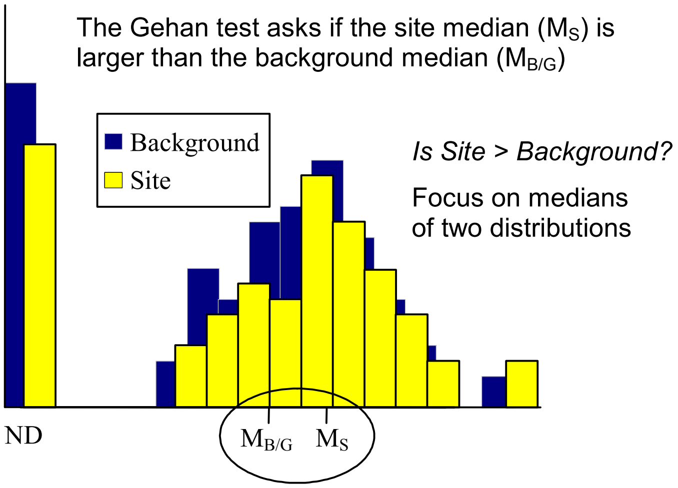
   $= 167.4$

   where $z_{0.99}$ = 2.33 is the 99th percentile of the standard normal distribution, that is found in Table B-1.
8. Because $R > w_{0.99}$, that is, 507.5 > 167.4, the WRS test determines the chemical to be a COPC.

Role of the Data Quality Objectives Process

The stakeholders and regulators used the DQO planning process to agree:

- On the methods that should be used to collect, handle, prepare, and measure the chemicals in the groundwater samples.
- On the fact that it is unlikely that more than 40% of the measurements will be nondetects.
- On the fact that the nondetects may have different RLs.


The Gehan test asks if the site median (M_S) is larger than the background median ($M_{B/G}$)

Is Site > Background?

Focus on medians of two distributions

TABLE 4-6. Number of site and background measurements (*n* and *m*, *n* = *m*) required by the WRS test to achieve a desired power[(a)]

Δ/s	α= 0.01					α= 0.025					α= 0.05					α= 0.10					α=0.25				
	Power					Power					Power					Power					Power				
	0.99	**0.975**	**0.95**	**0.90**	**0.75**	**0.99**	**0.975**	**0.95**	**0.90**	**0.75**	**0.99**	**0.975**	**0.95**	**0.90**	**0.75**	**0.99**	**0.975**	**0.95**	**0.90**	**0.75**	**0.99**	**0.025**	**0.05**	**0.10**	**0.25**
0.1	5452	4627	3972	3278	2268	4827	3870	3273	2846	1748	3972	3273	2726	2157	1355	3278	2846	2157	1655	964	2268	1748	1355	964	459
0.2	1370	1163	998	824	570	1163	973	823	665	440	998	823	685	542	341	824	685	542	416	243	570	440	341	243	116
0.3	614	521	448	370	256	521	436	369	298	197	448	369	307	243	153	370	298	243	187	109	256	197	153	109	52
0.4	350	297	255	211	148	297	248	210	170	112	255	210	175	139	87	211	170	139	106	62	146	112	87	62	30
0.5	227	193	166	137	95	193	162	137	111	73	166	137	114	90	57	137	111	90	69	41	95	73	57	41	20
0.6	161	137	117	97	67	137	114	97	76	52	117	97	81	64	40	97	78	64	19	29	67	52	40	29	14
0.7	121	103	88	73	51	103	86	73	59	39	88	73	61	48	30	73	59	48	37	22	51	39	30	22	11
0.8	96	81	69	57	40	81	68	57	46	31	69	57	48	38	24	57	46	38	29	17	40	31	24	17	8
0.9	77	66	58	47	32	65	55	46	38	25	56	48	39	31	20	47	38	31	24	14	32	25	20	14	7
1.0	64	55	47	39	27	55	46	39	32	21	47	39	32	26	16	39	32	25	20	12	27	21	16	12	6
1.1	55	47	40	33	23	47	39	33	27	18	40	33	28	22	14	33	27	22	17	10	23	18	14	10	5
1.2	48	41	35	29	20	41	34	29	24	16	35	29	24	19	12	29	24	19	15	9	20	16	12	9	4
1.3	43	36	31	26	18	36	30	26	21	14	31	26	22	17	11	26	21	17	13	8	18	14	11	8	4
1.4	38	32	28	23	16	32	27	23	19	13	28	23	19	15	10	23	19	15	12	7	16	13	10	7	4
1.5	35	30	25	21	15	30	25	21	17	11	25	21	18	14	9	21	17	14	11	7	15	11	9	7	3
1.6	32	27	23	19	14	27	23	19	16	11	23	19	16	13	8	19	16	13	10	6	14	11	8	6	3
1.7	30	25	22	18	13	25	21	18	15	10	22	18	15	12	8	18	15	12	9	6	13	10	8	6	3
1.8	28	24	20	17	12	24	20	17	14	9	20	17	14	11	7	17	14	11	9	5	12	9	7	5	3
1.9	26	22	19	15	11	22	19	16	13	9	19	16	13	11	7	16	13	11	8	5	11	9	7	5	3
2.0	25	21	18	15	11	21	18	15	12	8	18	15	13	10	7	15	12	10	8	5	11	8	7	5	3
2.25	22	19	16	14	10	19	16	14	11	8	16	14	11	9	6	14	11	9	7	4	10	8	6	4	2
2.5	21	18	15	13	9	18	15	13	10	7	15	13	11	9	6	13	10	9	7	4	9	7	6	4	2
2.75	20	17	15	12	9	17	14	12	10	7	15	12	10	8	5	12	10	8	6	4	9	7	5	4	2
4.0	19	16	14	12	8	16	14	12	10	6	14	12	10	8	5	12	10	8	6	4	8	6	5	4	2
4.5	18	16	13	11	8	16	13	11	9	6	13	11	9	8	5	11	9	8	6	4	8	6	5	4	2
4.0	18	15	13	11	8	15	13	11	9	6	13	11	9	7	5	11	9	7	6	4	8	6	5	4	2

(a) Power is the probability the WRS test correctly declares that the chemical is a COPC.

Adapted from: NRC et al. (1997, Table 5.3).

- On the values of the design parameters used to determine the minimum number of site and background groundwater measurements (see the sub-section entitled "Guidance on Implementing the Gehan Test" that follows).

The Gehan test (Palachek et al., 1994) is appropriate for this scenario because the site and background datasets were likely to contain multiple nondetects with different RLs. This test is appropriate for the site-wide impact hypothesis. The assumptions behind the Gehan test are summarized in Table 4-1.

Advantages and Disadvantages

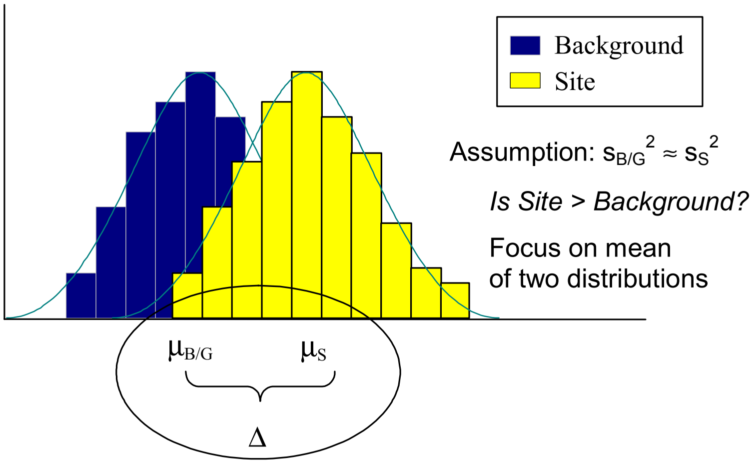
- The Gehan test can be used when the background or site datasets contain multiple nondetects with different RLs.
- If the RLs are different for the site and background datasets, then the test results may be an indication of this analytical difference rather than an indication that the chemical is a COPC.
- The Gehan test should be used in conjunction with the Slippage or Quantile Test so that both site-wide and localized contamination hypotheses can be evaluated.

Table 4-1 summarizes the advantages and disadvantages of the Gehan test.

Guidance on Implementing the Gehan Test

The Gehan test procedure for $n \geq 10$ and $m \geq 10$ is presented in Box 4-11. An example of the test is presented in Box 4-12. If $n < 10$ or $m < 10$, the procedure in Box 4-13 may be used to conduct the Gehan test. The minimum number of site and background measurements required to conduct the Gehan test may be approximated using the method described for the WRS test in Section 4.2.6.

The two-sample t test asks if the mean of the site distribution (μ_S) is greater than the mean of the background distribution ($\mu_{B/G}$)



4.2.8 Two-Sample t Test

Groundwater Contamination Scenario

The groundwater contamination scenario developed by the stakeholders and regulators during the DQO process is the same as for the WRS and Gehan tests (i.e. if contamination has occurred), it has most likely uniformly impacted the site.

Role of the Data Quality Objectives Process

The stakeholders and regulators used the DQO planning process to agree:

- On the methods that will be used to collect, handle, prepare, and measure the groundwater samples.
- On the fact that it is likely that very few nondetects will be reported by the laboratory.
- On the values of the design parameters used to determine the minimum number of site and background sediment measurements (see the subsection entitled "Guidance on Implementing the Two-Sample t Test").
- That normally distributed mean concentrations are expected. For large datasets (n and m greater than 30), the normality of the mean can be viewed as valid based on the Central Limit Theory.

BOX 4-11. Procedure for conducting the Gehan test when $n \geq 10$ and $m \geq 10$

1. Specify the tolerable probability, α (also known as test significance), that the Gehan test will incorrectly declare that the site median is larger than the background median, i.e., the probability that the chemical will be incorrectly identified as a COPC.
2. Specify the value of Δ/s and the power, where Δ/s is the magnitude of the difference in median site and background concentrations that must be detected by the Gehan test with the specified power. The notation Δ/s indicates the shift is expressed in units of standard deviation (s) of the underlying background and site concentration distributions for the chemical of interest.
3. Use the specified values of α, Δ/s, and the power in Table 4-6 to determine the number of site and background measurements needed when it is desired to have n equal to m. If it is not desired to have n equal to m, use Equation 6.3 in U.S. EPA (1994) and increase that value by 20% to guard against missing or unusable measurements.
4. If measurements are available from past sampling efforts, verify that their number is at least as large as the number indicated in Table 4-6. Collect additional samples if necessary to achieve the required number of samples.
5. List the combined m background and n site measurements, including the nondetects, from smallest to largest, where the total number of combined samples is $N = m + n$. The below-detection symbol (<) is ignored when listing the N data from smallest to largest.
6. Determine the N ranks, $R_1, R_2, \ldots, R_N$, for the N ordered data values using the method described in the example given in Box 4-12.
7. Compute the N scores, $a(R_1), a(R_2), \ldots, a(R_N)$ using the formula $a(R_i) = 2R_i - N - 1$, where i is successively set equal to 1, 2, …, N.
8. Compute the Gehan statistic, G, as follows:

$$G = \frac{\sum_{i=1}^{N} h_i a(R_i)}{\left\{ mn \sum_{i=1}^{N} [a(R_i)]^2 / [N(N-1)] \right\}^{1/2}} \tag{1}$$

where h_i = 1 if the ith datum is from the site population
= 0 if the ith datum is from the background population
$N = n + m$
$a(R_i) = 2R_i - N - 1$, as indicated above.

9. The Gehan test declares that the chemical is a COPC if $G \geq Z_{1-\alpha}$, where $Z_{1-\alpha}$ is the $100(1-\alpha)$th percentile of the standard normal distribution, which is obtained from Table B-1. Otherwise, the test declares that the evidence is not strong enough to conclude that the chemical is a COPC.

- That measurements from the site are expected to have approximately the same variance as the background measurements. If a statistical test, e.g., an F test described in Conover (1998) and U.S. EPA (2000, p. 4-33), indicates the site and background measurements may not have the same variance, but both mean concentrations appear to be normally distributed, then the Satterthwaite two-sample t test (Section 4.2.9) should be used to test for differences in the site and background means.

The two-sample t test is appropriate for this scenario because the assumptions of normality of mean, equal variances for background site data, and the general absence of nondetects are or will be valid. However, once the measurements are obtained, these assumptions should be evaluated by observation and statistical testing. The assumptions

BOX 4-12. Example of the Gehan test

1. Suppose α was specified to be 0.05.
2. Suppose Δ/*s* and the power were specified to be 2.0 and 0.90, respectively. That is, the stakeholders and regulators specified that if the median of the site concentration distribution is greater than the median background distribution by the amount Δ/*s* = 2.0, enough measurements should be obtained so the Gehan test has a power of 0.90 of detecting that fact.
3. Using the specified values of Δ/*s* and power in Table 4-6, we find that n = m = 10 measurements are needed to conduct the Gehan test.
4. The 10 samples from the site and the background area were collected using a suitable probability-based sampling design (for example, simple random sampling or sampling at the nodes of a square or triangular grid) and each sample was analyzed for the chemical of interest. Suppose the measurements are:

Background:	1	<4	5	7	<12	15	18	<21	<25	27
Site:	2	<4	8	17	20	25	34	<35	40	43

5, 6 and 7. Use the following procedure to determine the *N* = 20 ranks $R_1, R_2, \ldots, R_{20}$ and the 20 scores $a(R_i)$. Refer to Table 1 and the bullet list below as you go through the steps.

Table 1. Calculations to Determine the Ranks, R_i, and the Scores, $a(R_i)$

Data	h_i	Index_i	d_i	e_i	R_i	$a(R_i)$
1	0	0	1	0	4	–13
2	1	0	2	0	5	–11
<4	1	1	2	1	4.5	–12
<4	0	1	2	2	4.5	–12
5	0	0	3	2	9	–7
7	0	0	4	2	8	–5
8	1	0	5	2	9	–3
<12	0	1	5	3	6	–9
15	0	0	6	3	10.5	0
17	1	0	7	3	11.5	2
18	0	0	8	3	12.5	4
20	1	0	9	3	14.5	6
<21	0	1	9	4	8	–5
<25	0	1	9	5	8	–5
25	1	0	9	5	15.5	10
27	0	0	10	5	16.5	12
34	1	0	12	5	17.5	14
<35	1	1	12	6	9.5	–2
40	1	0	13	6	19	17
43	1	0	14	6	20	19

- List the combined *m* background and *n* site measurements, including the nondetects, from smallest to largest, as illustrated in column 1 of Table 1. Ignore the below-detection symbol (<) when listing the *N* data from smallest to largest.
- Place a 0 or 1 in the second column of Table 1 (the column with heading h_i) using the following rule:

  h_i = 0 if the *i*th measurement is from the background dataset
  = 1 if the *i*th measurement is from the site dataset

- Place a 0 or 1 in the 3rd column of Table 1 (the column with heading Index_i) using the following rule:

  Index_i = 0 if the *i*th measurement is a detect
  = 1 if the *i*th measurement is a nondetect value

- When moving down the data in column 1, determine the values of parameters d and e (columns 4 and 5 in Table 1) using the following rules:
  - If the first datum in column 1 is a detect, i.e., if Index_i = 0, then set d = 1 and e = 0 in the first row of Table 1.
  - If the first datum in column 1 is a nondetect value, i.e., if Index_i = 1, then set d = 0 and e = 1 in the first row of Table 1.
  - For each successive row (rows 2 through n = 20), increase *d* by 1 whenever the datum in column 1 in that row is a detect, i.e., whenever Index = 0.
  - For each successive row, increase *e* by 1 whenever the datum in column 1 in that row is a nondetect value, i.e., when Index = 1.

BOX 4-12. (cont'd)

- Let T denote the total number of nondetects in the pooled background and site datasets. For the previous data there are $T = 6$ nondetects. Compute the rank of the ith datum (i.e., of the datum in the ith row in the previous table) as follows:
    - $R_i = d_i + (T + e_i)/2$ if the datum in column 1 of the ith row is a detect, i.e., if $h_i = 0$ for the ith row.
    - $R_i = (T + 1 + d_I)/2$ if the datum in column 1 of ith row is a nondetect value, i.e., if $h_i = 1$ for the ith row.
- Compute the $n = 20$ scores, $a(R_1), a(R_2), \ldots, a(R_{20})$, using the formula

  $$a(R_i) = 2R_i - N - 1$$

  for successive values of i = 1, 2, …, 20.

8. Compute the Gehan statistic, G:

$$G = \frac{(-11)+(-12)+(-3)+2+6+10+14+(-2)+17+19}{\{10*10[(-13)^2+(-11)^2+(-12)^2+\ldots+(-2)^2+(17)^2+(19)^2]/20*19\}^{1/2}}$$

$$= 40 / [(100*1942) / (20*19)]^{1/2}$$
$$= 40 / 22.606$$
$$= 1.77$$

9. In Step 1 above we specified that $\alpha = 0.05$. When $\alpha = 0.05$, Table B-1 yields $Z_{1-\alpha} = Z_{0.95} = 1.645$. As $G > 1.645$, i.e., $1.77 > 1.645$, the Gehan test declares that the chemical is a COPC.

BOX 4-13. Procedure for conducting the Gehan test when *n*<10 and *m*<10

1. Generate on a computer all possible orderings of the combined n site and background measurements. Denote the *number* of possible orderings by M.
2. Compute the G statistic (Equation 1 in Box 4-11) for each of these orderings to generate an empirical distribution (histogram) of the M values of G.
3. Determine the $100(1 - \alpha)$th percentile of the empirical distribution of G generated by Step 2 as follows (from Gilbert, 1987, p. 141) where α is the tolerable probability that the test procedure described in this box will incorrectly declare that the chemical is a COPC:
    - Order the M values of G from smallest to largest.
    - Compute $k = (1 - \alpha)(M + 1)$
    - If k is an integer, the $(1 - \alpha)$th percentile is the kth largest value of the ordered M values of G.
    - If k is not an integer, determine the value of k', where k' is the largest integer less than k. Compute the $(1 - \alpha)$th percentile by linear interpolation between the k'th and $(k' + 1)$th largest values of G.
4. If the value of G computed, *using the ordering actually observed for the collected background and site data,* equals or exceeds the $100(1-\alpha)$th percentile obtained above, conclude that the chemical is a COPC.

behind the two-sample t test are summarized in Table 4-1.

Advantages and Disadvantages

- ❑ To implement the two-sample t test, all nondetects must be replaced with a surrogate value, such as one-half the DL. However, if a large number of nondetects exist (e.g., more than 15% of measurements) the test results cannot be viewed as reliable.
- ❑ The WRS or Gehan tests should be used in place of the two-sample t test if the tests for normality indicate the data are not normally distributed and if n and m are too small (less

than 30) for the estimated means to be normally distributed. However, if *n* and *m* are small, the power of the WRS and Gehan tests may not be adequate.

- ❑ If large numbers of nondetects exist (up to 40% of data), nonparametric tests, such as the WRS or Gehan tests, should be used. The two-sample test of proportions (Section 4.2.10) is appropriate if more than 40% of the data are nondetects.
- ❑ If the exact nature of the spatial extent of the groundwater contamination is not known, the two-sample t test should be accompanied by the Quantile or Slippage tests. These latter tests address the localized (hotspot) impact hypothesis.
- ❑ If the mean concentrations of site and background measurements are normally distributed, but their variances are different, then the Satterthwaite two-sample t test should be used instead of the two-sample t test. The difference between variances can be assessed by statistical tests, such as the F test described in Conover (1998, p. 300) and U.S. EPA (2000). Note, however, that real data are never completely normal, and the power of these common parametric tests is reduced when the data distributions depart significantly from normality.
- ❑ The two-sample t test is the most common statistical test for assessing the differences between mean concentrations of two datasets. Most statistical software packages can compute the two-sample t test.

Table 4-1 summarizes the advantages and disadvantages of the two-sample t test.

Guidance on Implementing the Two-Sample t Test

The minimum number of site (*n*) and background (*m*) measurements required for the two-sample t test should be approximated using the procedure outlined in Box 4-14. Software such as VSP (Hassig et al., 2002) and Decision Error Feasibility Trials (DEFT) (U.S. EPA, 2001) provide options for rapid computation of *n* and *m* values. An example of the evaluation of Equation 1 in Box 4-14 is given in Box 4-15. After *n* and *m* have been determined, the samples collected, and measurements reported by the laboratory, summary statistics should be computed for both the site and background datasets. In particular, the computed sample variance of the site measurements should be compared with the computed sample variance of the background measurements to determine if they are approximately equal, a required assumption of the two-sample t test. A procedure (an F test) for testing whether the two sample variances are equal is provided in Conover (1998) and U.S. EPA (2000, p. 4-33). This procedure is commonly available in statistical software packages.

If some measurements appear to be unusually large compared to the remainder of the measurements in the dataset, a test for outliers should be conducted (see Appendix A.3). Outliers should be identified, investigated to determine if they represent mistakes or errors, and, if necessary, discarded. If the site and background datasets are small (*n* and *m* less than 30) they should then be tested for normality using probability plots and normality tests.

After the assumptions of equal variances (and normality for small datasets) have been shown to be reasonable, the two-sample t test can be conducted. The test procedure is presented in Box 4-16; an example of the procedure is presented in Box 4-17. Boxes 4-18 and 4-19 present methods and examples of calculating some of the statistical parameters presented in Box 4-16.

4.2.9 Satterthwaite Two-Sample t Test

Groundwater Contamination Scenario

The groundwater contamination scenario for the Satterthwaite two-sample t test is the same as for the two-sample t test (Section 4.2.8), i.e. if contamination exists, it is probably uniformly impacted the site. However, the Satterthwaite two-sample t test was selected (instead of the two-sample t test) because it can be used when measurements from the site are not expected to have approximately the same variance as the background measurements.

BOX 4-14. Procedure for calculating the number of site and background measurements required to conduct the two-sample t test

The formula for calculating the number of site (n) and background (m) measurements required to conduct the two-sample t test is:

$$n = m \approx \frac{2s^2(Z_{1-\alpha} + Z_{1-\beta})^2}{(\mu_s - \mu_b)^2} + 0.25*(Z_{1-\alpha})^2 \tag{1}$$

where: s^2 = expected variance of the measurements at both site and background areas (ideally, the value of s^2 used should be approximated using measurements previously obtained from the site and background areas, or obtained in a special pilot study at the site and background areas)

α = the tolerable probability that the two-sample t test will incorrectly declare that the chemical is a COPC (α is usually specified to be a small value such as 0.01, 0.025, 0.05 or 0.10)

$1 - \beta$ = the required power (probability) that the two-sample t test will declare that the chemical is a COPC when that is indeed the case (β is usually specified to be ≥ 0.80)

$\mu_s - \mu_b$ = true site mean (μ_s) minus the true background mean (μ_b); i.e., the difference in the *true* (unknown) means of the site and background areas that the stakeholders and regulators have agreed needs to be detected by the two-sample t test with power (probability) equal to $1 - \beta$.

$Z_{1-\alpha}$ = the $100(1 - \alpha)$ percentile of the standard normal distribution, which is found in Table B-1 (for example, if $\alpha = 0.05$, Table B-1 indicates $Z_{1-0.05} = Z_{0.95} = 1.645$)

$Z_{1-\beta}$ = the $100(1 - \beta)$ percentile of the standard normal distribution, which is found in Table B-1 (for example, if $1 - \beta = 0.80$, then we find from Table B-1 that $Z_{0.80} = 0.84$)

Stakeholders and regulators should determine the appropriate values of the parameters in Equation 1 in this box during the DQO planning process.

BOX 4-15. Example of the procedure for calculating the number of site and background measurements required to conduct the two-sample t test

Suppose the values of the parameters in Equation 1 in Box 4-14 were specified by the stakeholders and regulators as follows:

$s^2 = 7.5$
$\alpha = 0.025$
$1 - \beta = 0.80$
$\mu_s - \mu_b = 4$

Table B-1 indicates that $Z_{1-\alpha} = Z_{0.975} = 1.96$ and $Z_{1-\beta} = Z_{0.80} = 0.84$

Therefore, Equation 1 is:

$$\begin{aligned} n &= m \approx 2*7.5*(1.96 + 0.84)^2 / 4^2 + 0.25*(1.96)^2 \\ &= 8.31 \text{ or } 9 \end{aligned}$$

Therefore, nine site and nine background measurements are required for the two-sample t test to attain the performance specified (by the values of α and $1 - \beta$) to detect a difference in true means of size $\mu_s - \mu_b = 4$ when the variance of the data at the site and background areas is $s^2 = 7.5$.

The reader may want to try other values of s^2 and $\mu_s - \mu_b$ to see how $n = m$ change for the specific values of α and $1 - \beta$ given above.

BOX 4-16. Procedure for conducting the two-sample t test

1. Stakeholders and regulators used the DQO process to select values of s^2, α, 1 – β, and $\mu_s - \mu_b$, and the procedure in Box 4-14 (as illustrated in Box 4-15) to determine the number of site (*n*) and background (*m*) measurements. VSP (Hassig et al., 2002) and DEFT (U.S. EPA, 2001) software also can be used to compute *n* and *m*.
2. Collect the samples and obtain the *n* and *m* site and background measurements.
3. Suppose
   - the *n* site measurements are denoted by $x_1, x_2, \ldots, x_n$
   - the *m* background measurements are denoted by $y_1, y_2, \ldots, y_m$
4. Compute the two-sample t test statistic, denoted by *T*:

$$T = \frac{\bar{x} - \bar{y}}{\{(n+m)[n-1)s_x^2 + (m-1)s_y^2]/[nm(n+m-2)]\}^{1/2}}$$

where $\bar{x}$ = the arithmetic mean of the *n* site measurements
$\bar{y}$ = the arithmetic mean of the *m* background measurements
s_x^2 = the sample variance of the *n* site measurements (the formula for computing this variable is given in Box 4-18)
s_y^2 = the sample variance of the *m* background measurements (see Box 4-18)

5. The two-sample t test declares:
   - that the chemical is a COPC if $T \geq t_{1-\alpha,\ n+m-2}$
   - that insufficient evidence exists to conclude that the chemical is a COPC if $T < t_{1-a,\ n+m-2}$

   where $t_{1-\alpha,\ n+m-2}$ is the 100(1 – α) percentile of the *t* distribution that has $n + m - 2$ degrees of freedom. The value of $t_{1-\alpha,\ n+m-2}$ is determined from Table B-9 by using that table with the values of 1 – α and $n + m - 2$. Note that the value of α was specified in Step 1, as part of the process for determining the required number of site and background measurements.

If the two-sample t test does not declare that the chemical is a COPC, it may indicate: (1) the chemical is indeed not a COPC, or (2) the assumptions that underlie the t test are not valid for the site and background measurements, or (3) an insufficient number of measurements (*n* and *m*) were obtained for the t test to be able to detect the difference in site and background concentration distributions that actually exists.

The possibility that the test did not declare the chemical to be a COPC due to items (2) or (3) should be evaluated.

- First, review the DQO planning process records to make sure the number of samples (*n* and *m*) agrees with what was determined necessary to detect the difference between site and background means that was considered important.
- Second, review the computations that were conducted to test for normality and equality before the t test was calculated. Verify that the tests were done correctly using the appropriate data. Redo the tests if necessary.
- Third, the shift in the site concentration distribution may in fact be smaller than the shift selected by the stakeholders as being important to detect, in which case no additional measurements are needed. However, as the true difference in means is unknown, update the estimated number of site and background measurements needed to detect the critical (important) shift in the site mean by calculating the variance of the site and background measurements (s_x^2 and s_y^2, respectively), and use the larger of these two estimated variances in Equation 1 of Box 4-14. If this new value, denoted by *n′*, is larger than either the number of site or background measurements obtained or used in the t test, collect additional samples so *n′* site and *n′* background measurements are available. Then redo the t test.

BOX 4-17. Example of the two-sample t test

1. Suppose the values of the parameters in Equation 1 of Box 4-14 were specified by the stakeholders and regulators to be $s^2 = 7.5$, $\alpha = 0.025$, $1 - \beta = 0.80$, and $\mu_s - \mu_b = 4$. In Box 4-15 it was shown that $n = m = 10$ for these parameter values.
2. The n = m measurements were obtained.
3. Suppose the values were as follows:

   Site Data (x): 90, 77, 81, 210, 92, 130, 110, 120, 140, 84
   Background Data (y): 23, 15, 78, 26, 90, 99, 87, 34, 17, 10

   No potential outliers are apparent in either dataset. Therefore, tests for outliers do not appear necessary. Each dataset should tested for normality (Appendix A.1). The reader is encouraged to conduct these tests. Suppose the tests indicate the data can be assumed to be normally distributed.
4. The following calculations were conducted:

   $\bar{x} = 113.4$
   $\bar{y} = 47.9$
   $s_x^2 = 1623.82$
   $s_y^2 = 1287.21$

$$T = \frac{\bar{x}-\bar{y}}{\left\{(n+m)\,[(n\text{-}1)\,s_x^2+(m\text{-}1)\,s_y^2]\,/\,[nm\,(n+m\text{-}2)]\right\}^{1/2}}$$

$$= \frac{113.4-47.9}{\{(10+10)\ [9*1623.82+9*1287.21]/[10*10\ (10+10-2)]\}^{1/2}}$$

$$= \frac{65.5}{17.06}$$

$$= 3.84$$

5. The value of $t_{1-\alpha,\, n+m-2}$, i.e., of $t_{0.975,\, 18}$ is found from Table B-9 to be 2.101. Therefore, as $T > 2.101$, i.e., $3.84 > 2.101$, the two-sample t test declares that the chemical is a COPC.

Role of the Data Quality Objectives Process

The stakeholders and regulators concluded that the site and background datasets meet all the conditions for use of the two-sample t-test (Section 4.2.8) with one exception: based on prior data and statistical tests, or expert knowledge, the site and background datasets are expected to have different variances.

Advantages and Disadvantages

- The Satterthwaite two-sample t test can be applied to datasets with unequal variances. A statistical procedure for testing whether two sample variances are equal is provided in Conover (1998) and U.S. EPA (2000, p. 4-33).
- Like the two-sample t test, the Satterthwaite two-sample t test assumes that the mean concentrations are normally distributed. This assumption is valid for large datasets (*n* and *m* greater than 30); however, for small datasets, tests for normality of the measurements should be conducted to assure the validity of this assumption.
- To implement the Satterthwaite two-sample t test, all nondetects must be replaced with a surrogate value, such as one-half the DL. However, if a large number of nondetects

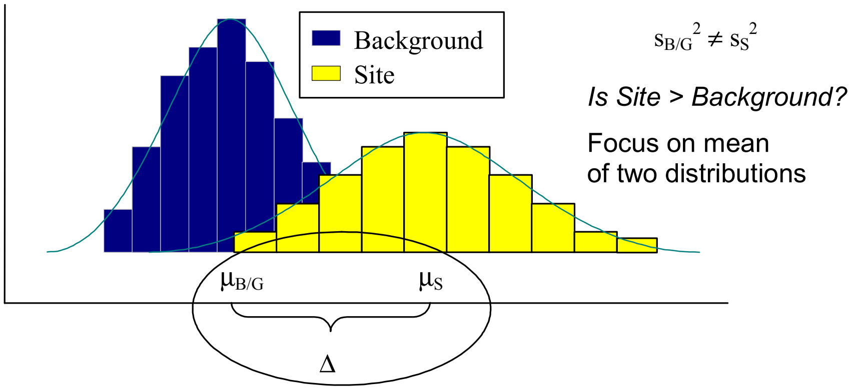
BOX 4-18. Descriptive summary statistics for datasets with no nondetects

Descriptive Statistics	Definitions and Computation
Arithmetic Mean ($\bar{x}$)	$\bar{x} = (x_1 + x_2 + \ldots + x_n) / n$
Median (when *n* is an odd integer)	The middle value of the *n* measurements after they are arranged in order of magnitude from smallest to largest.
Median (when *n* is an even integer)	The arithmetic average of the middle two of the ordered measurements.
*p*th Sample Percentile	The value (not necessarily an observed measurement) that is greater than or equal to p% of the values in the dataset and less than or equal to (1 – p)% of the data values, where $0 < p < 1$. Compute $k = p(n + 1)$, where n is the number of measurements. If *k* is an integer, the *p*th percentile is the *k*th largest measurement in the ordered dataset. If *k* is not an integer, the *p*th percentile is obtained by linear interpolation between the two measurements in the ordered dataset that are closest to *k*.
Range	The difference between the maximum and minimum measurements.
Interquartile Range	The 75th sample percentile minus the 25th sample percentile.
Sample Standard Deviation (s)	A measure of dispersion (spread or variation) of the n measurements in a dataset that is computed as follows: $s = \{[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2] / (n - 1)\}^{1/2}$
Sample Variance (s^2)	The sample variance is the square of the sample standard deviation, i.e., Sample Variance = s^2.
Coefficient of Variation (CV)	The CV is a measure of relative standard deviation that is computed as follows: $CV = s / \bar{x}$.

BOX 4-19. Examples of descriptive summary statistics for datasets with no nondetects

Descriptive Statistics	Example Calculations
Arithmetic Mean ($\bar{x}$)	Suppose there are five data, say 50, 34, 52, 62, 60. Then the arithmetic mean is: $\bar{x} = (50 + 34 + 52 + 62 + 60) / 5 = 51.6$
Median (when n is an odd integer)	For the five data (after being ordered from smallest to largest) 34, 50, 52, 60, 62, the median is 52.
Median (when n is an even integer)	Suppose there are six data, which when ordered from smallest to largest are 0.1, 0.89, 2.0, 3.01, 3.02, 4.0. Then the median is (2.0 + 3.01) / 2 = 2.50.
*p*th Sample Percentile	Suppose the dataset (after being ordered) is 34, 50, 52, 60, 62, and we want to estimate the 60th percentile, i.e., p = 0.6. Now, k = 0.6 (5 + 1) = 3.6. Because *k* is not an integer, we linearly interpolate between the 3rd and 4th largest measurements, i.e., the 0.60 sample percentile is 52 + 0.6 (60 – 52) = 56.8.
Range	For the dataset 50, 34, 52, 62, 60, the range is 62 – 34 = 28.
Interquartile Range	The 75th sample percentile of the (ordered) dataset 34, 50, 52, 60, 62 is 60 + 0.5(62 – 60) = 61. The 25th sample percentile is 34 + 0.5(50 – 34) = 42. Therefore, the interquartile range is 61 – 42 = 19.
Sample Standard Deviation (s)	The sample standard deviation of the dataset 50, 34, 52, 62, 60 is: $s = \{ [(50 - 51.6)^2 + (34 - 51.6)^2 + (52 - 51.6)^2 + (62 - 51.6)^2 + (60 - 51.6)^2] / 4\}^{1/2} = 11.08$
Sample Variance (s^2)	The sample variance of the dataset 50, 34, 52, 62, 60 is the square of the sample standard deviation, i.e., variance = $(11.08)^2 = 122.77$.
Coefficient of Variation (CV)	The CV for the dataset 50, 34, 52, 62, 60 is CV = 11.08 / 51.6 = 0.21.

The Satterthwaite two-sample t test asks if the mean of the site distribution (μ_S) is greater than the mean of the background distribution ($\mu_{B/G}$)



exist (e.g., more than 15% of measurements) the test results cannot be considered reliable.

- If a large number of nondetects exist, other tests should be used to address the site-wide impact hypothesis. For example, the Gehan test should be used if multiple RLs are present. The two-sample test of proportions (Section 4.2.10) is appropriate if more than 40% of the data are nondetects.
- If the exact nature of the spatial extent of the groundwater contamination is unknown, the Satterthwaite two-sample t test should be accompanied by the Quantile or Slippage tests. These tests address the localized (i.e., hotspot) impact hypothesis.
- If the datasets are small (*n* and *m* less than 30), and the tests for normality indicate the measurements are not normally distributed, a nonparametric test such as the Gehan test should be used in place of the two-sample t test.

Table 4-1 summarizes the advantages and disadvantages of the Satterthwaite two-sample t test.

Guidance on Implementing the Satterthwaite Two-Sample t Test

Ideally, the same number of measurements should be obtained for both the site and background areas. The required number of measurements (*n*) should be approximated by the same procedure used for the two-sample t test (see Box 4-14). The larger of the site or background variance should be used for s^2 in Equation 1 of Box 4-14.

When the *n* measurements have been obtained, the normality of mean concentrations can be assumed to be reasonable if *n* exceeds 30. For smaller datasets, the assumption of normality should be evaluated by appropriate testing methods (Appendix A.1). The Satterthwaite two-sample t test procedure is described in Box 4-20. An example of this procedure is presented in Box 4-21.

4.2.10 Two-Sample Test of Proportions

Groundwater Contamination Scenario

Consider the case where the site is expected to contain hotspot contamination, i.e., localized impact hypothesis. In addition, a large number of measurements from both the site and background areas are expected to be nondetects. Due to the large number of nondetects, no specific assumptions can be made about the statistical distribution of the sampling data.

Role of the Data Quality Objectives Process

Given the anticipated large number of nondetects, it is difficult to conduct a valid statistical test on differences between mean or median concentrations of the site and background datasets. Therefore, the stakeholders and regulators decided to test the site and background data to identify the dataset that has the largest proportion of concentrations greater that a specified cut-off concentration, *C*. After determining the magnitude of the cut-off concentration limit, *C*, the DQO planning team agreed:

- That the null and alternative hypotheses to be tested are

$$H_o: \quad P_s \le P_b$$
$$H_a: \quad P_s > P_b$$

BOX 4-20. Procedure for conducting the Satterthwaite two-sample t test

1. Use the DQO process to select values of α, β, $\mu_s - \mu_b$ and the larger of the site and background variances (s^2). Then use the procedure in Box 4-14, as illustrated in Box 4-15 to determine the number of measurements for both the site and the background areas.
2. Collect the samples and obtain the site and n background measurements.
3. Suppose
   - the n site measurements are denoted by $x_1, x_2, \ldots, x_n$
   - the n background measurements are denoted by $y_1, y_2, \ldots, y_n$
4. Compute the Satterthwaite two-sample t test statistic, denoted by T_s:

$$T_s = \frac{\bar{x} - \bar{y}}{\left(s_x^2/n + s_y^2/n\right)^{1/2}}$$

where $\bar{x}$ = the arithmetic mean of the n site measurements
$\bar{y}$ = the arithmetic mean of the n background measurements
s_x^2 = the sample variance of the n site measurements (the formula for computing this variable is given in Box 4-18)
s_y^2 = the sample variance of the n background measurements (see Box 4-18).

5. Compute the approximate degrees of freedom, f, as follows:

$$f = \frac{\left(s_x^2/n + s_y^2/n\right)^2}{\left(s_x^2/n\right)^2/(n-1) + \left(s_y^2/n\right)^2/(n-1)}$$

Note: the Satterthwaite t-test can be computed when the number of site and background measurements are not equal. In that case, n_x and n_y would replace n in these equations, as appropriate.

6. The Satterthwaite two-sample t test declares that:
   - the chemical is a COPC if $T_s \geq t_{1-\alpha,\,f}$
   - insufficient evidence exists to conclude that the chemical is a COPC if $T_s < t_{1-\alpha,\,f}$

where $t_{1-\alpha,\,f}$ is the 100(1 − α) percentile of the t distribution that has f degrees of freedom. The value of $t_{1-\alpha,\,f}$ is determined from Table B-9 by entering 1 − α and f. If f is not an integer, linear interpolation may be used to determine $t_{1-a,\,f}$ from Table B-9.

If the two-sample t test does not declare that the chemical is a COPC, it may indicate (1) the chemical is indeed not a COPC, or (2) the assumptions that underlie the t test are not valid for site and background measurements, or (3) an insufficient number of measurements were obtained for the Satterthwaite t test to detect the difference in site and background concentration distributions that actually exists.

The possibility that the test did not declare the chemical to be a COPC due to items (2) or (3) should be evaluated.

1. First, review the DQO planning process records to make sure the number of samples agrees with what was determined necessary to detect the difference between site and background means that was considered important.
2. Second, review the computations that were conducted to test for normality and equality before the t test was calculated. Verify that the tests were done correctly using the appropriate data. Redo the tests if necessary.
3. Third, the shift in the concentration distribution may, in fact, be smaller than the shift selected by the stakeholders, in which case no additional measurements are needed. However, as the true difference in means is unknown, update the estimated number of site and background measurements needed by calculating the variance of the site and background measurements (s_x^2 and s_y^2, respectively), and using the larger of these two estimated variances in Equation 1 of Box 4-14. If this new value, denoted by n', is larger than the number of site and background measurements obtained and used in the t test, then collect additional samples to obtain n' site and n' background measurements. Then redo the Satterthwaite t test.

BOX 4-21. Example of the Satterthwaite two-sample t test

1. Suppose a preliminary study was conducted to estimate the variance of the background and site measurements, and the variance of the site data was found to be significantly larger than the background data variance. Suppose the larger of the two estimated variances was 15. Therefore, 15 were selected as the s^2 value to use in Equation 1 in Box 4-14. (If very few site and background measurements were obtained in the preliminary study, say less than 10 for each, the value for s^2 may be increased by 20% or so to guard against not taking enough measurements.) Also, suppose the values of the other parameters in Equation 1 in Box 4-14 were specified by the stakeholders and regulators during the DQO process to be $\alpha = 0.10$, $1 - \beta = 0.90$ and $\mu_s - \mu_b = 4$. For these parameter values, the reader may verify that Equation 1 in Box 4-14 gives the value $n = 14.1$, rounded down to $n = 14$.

2. Therefore, $n = 14$ site and $m = 14$ background samples were collected and measured using the methods specified during the DQO process and as documented in the Quality Assurance Project Plan (QAPP).

3. Suppose the measurements are as follows:

   Site Data (x): 7.2, 4.3 10.9, 11.5, 2.0, 6.4, 12.1, 2.2, 0.5, 0.9, 1.1, 2.0, 5.1, 10.5
   Background Data (y): 8.1, 14.2, 5.0, 2.5, 7.2, 4.9, 10.8, 1.1, 8.5, 11.3, 9.2, 2.7, 4.1, 9.1

4. No potential outliers appear to be present in either dataset. Therefore, tests for outliers do not appear to be necessary. Each dataset should be evaluated graphically and with a formal statistical test (Appendix A.1) to verify that the normality assumption is reasonable for each dataset. The reader may verify that the normality assumption appears to be reasonable for both datasets.

5. Next, the following calculations are conducted:

$$\bar{x} = 5.41$$

$$\bar{y} = 6.84$$

$$s_x^2 = 18.708$$

$$s_y^2 = 14.316$$

$$T_s = \frac{\bar{x} - \bar{y}}{\left(s_x^2 / n + s_y^2 / n\right)^{1/2}}$$

$$= \frac{5.41 - 6.84}{(18.708/14 + 14.316/14)^{1/2}}$$

$$= \frac{-1.43}{1.536}$$

$$= -0.931$$

$$f = \frac{(18.708/14 + 14.316/14)^2}{(18.709/14)^2 / 13 + (14.316/14)^2 / 13}$$

$$= \frac{5.564}{0.1374 + 0.08043}$$

$$= 25.54 \text{ degrees of freedom}$$

6. Using linear interpolation between $t_{0.90,\,25} = 1.316$ and $t_{0.90,\,26} = 1.315$ in Table B-9, we find that $t_{0.90,\,25.54} = 1.3155$. Therefore, as $T_s < 1.3155$, i.e., as $-0.931 < 1.3155$, the Satterthwaite two-sample *t* test does *not* declare that the chemical is a COPC. Indeed, the estimated mean of the site measurements is less than the estimated mean of the background measurements.

BOX 4-21. (cont'd)

7. As the test did not declare that the chemical is a COPC, the DQO process records and QAPP should be reviewed to confirm that all requirements for collecting the type, quantity, and quality of measurements were correctly followed. Next, evaluate whether the number of measurements used in the test ($n = 14$) was too small to achieve the allowable α and β decision error rates specified during the DQO process (see Step 1 in this example) and recorded in the QAPP. To do so, either compute Equation 1 in Box 4-14 by hand, or use VSP (Hassig et al., 2002) or DEFT (U.S. EPA, 2001) software using the larger of the estimated site and background variances, i.e., using $s^2 = 18.7$, as computed in Step 5. From Equation 1, $n = 16.1$, which is rounded up to 17, when $s^2 = 18.7$, $\alpha = 0.10$, $1 - \beta = 0.90$ and $\mu_s - \mu_b = 4$. Therefore, three additional samples should be collected and measured in each area (the background and site areas). Simple random sampling should be used to determine the new sampling locations in the field. Also, the collection and measurement protocols for obtaining the new data should be exactly the same as for the original data (specified in the QAPP). Then the Satterthwaite two-sample t test should be recomputed using the new background and site datasets, each of which consists of 14 old and 3 new measurements. Before conducting the Satterthwaite t test, graphic and statistical tests for normality should be conducted on the new datasets ($n = 17$) to verify that the normality assumption is still reasonable.

where P_s and P_b are the true proportions of the site and background measurements, respectively, that exceed C.

- ❑ On the methods that will be used to collect, handle, prepare, and measure the groundwater samples.
- ❑ On the values of the design parameters used to determine the minimum number of site and background groundwater measurements (see the subsection entitled "Guidance on Implementing the Two-Sample Test of Proportions").

The assumptions behind the two-sample test of proportions are summarized in Table 4-1.

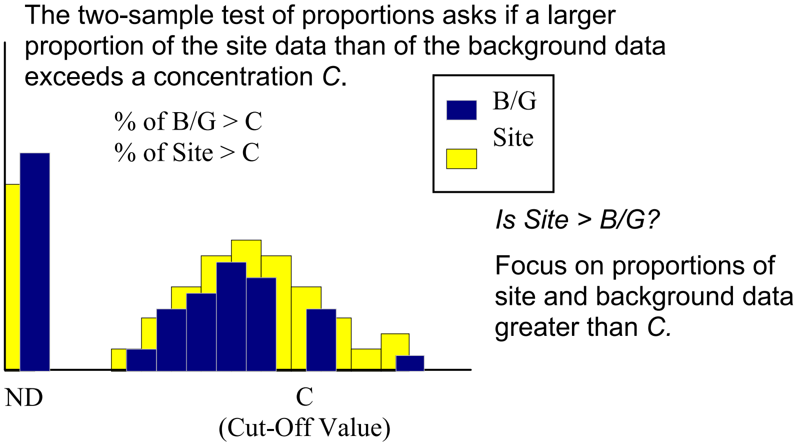
Advantages and Disadvantages

- ❑ The two-sample test of proportions may be conducted regardless of the underlying distribution of measurements, i.e., the test is nonparametric.
- ❑ The two-sample test of proportions is most suitable in cases where the site and background datasets contain a large number of nondetects.
- ❑ The two-sample test of proportions simply focuses on the proportions of measurements that exceed a given cut-off value.

Table 4-1 summarizes the advantages and disadvantages of the two-sample test of proportions.

Guidance on Implementing the Two-Sample Test of Proportions

The minimum number of site (n) and background (m) measurements required to conduct the two-sample test of proportions should be approximated by the procedure outlined in Box 4-22, which is available in VSP (Hassig et al., 2002) or DEFT (U.S. EPA, 2001) software. An example of the procedure is given in Box 4-23.



BOX 4-22. Procedure for calculating the number of site and background measurements required for the two-sample test of proportions

The formula for calculating the number of site (n) and background (m) measurements required for the two-sample test of proportions is as follows (from U.S. EPA, 2000):

$$n = m = \frac{2(Z_{1-\alpha} + Z_{1-\beta})^2 \, \overline{P}\,(1 - \overline{P})}{D^2} \qquad (1)$$

where: $\overline{P}$ = $(P_s + P_b) / 2$
P_s = the proportion of the true site distribution of potential measurements that exceeds C
P_b = the proportion of the true background distribution of potential measurements that exceeds C
α = the tolerable probability that the two-sample test of proportions will incorrectly reject H_o, i.e., the probability that the chemical will be incorrectly identified as a COPC (α is usually specified to be a small value such as 0.01, 0.025, 0.05, or 0.10)
$1 - \beta$ = the power (probability) required that the two-sample test of proportions will declare that the chemical is a COPC when that is indeed the case (β is usually specified to be ≥ 0.80)
D = the difference in the true (unknown) proportions of the site and background distributions that exceed the constant C, that must be detected with probability $1 - \beta$. That is, the stakeholders and regulators have agreed that the difference D needs to be detected by the two-sample test of proportions with a power (probability) equal to $1 - \beta$.
$Z_{1-\alpha}$ = the $100(1 - \alpha)$ percentile of the standard normal distribution, that is tabulated in Table B-5 (for example, if $\alpha = 0.05$, then Table B-1 indicates that $Z_{1-0.05} = Z_{0.95} = 1.645$)
$Z_{1-\beta}$ = the $100(1 - \beta)$ percentile of the standard normal distribution, that is tabulated in Table B-1 (for example, if $1 - \beta = 0.80$, we find from Table B-5 that $Z_{0.80} = 0.84$)

Stakeholders and regulators should determine the appropriate values of the parameters in Equation 1 during the DQO planning process.

BOX 4-23. Example of the procedure for calculating the number of site and background measurements required for the two-sample test of proportions

Suppose the values of the parameters in Equation (1) of Box 4-22 were specified by the stakeholders and regulators as follows:

D = 0.20
α = 0.025
β = 0.20
$Z_{1-\alpha}$ = $Z_{0.975} = 1.96$ and $Z_{1-\beta} = Z_{0.80} = 0.84$ (from Table B-1).

Because P_s and P_b are true values and are therefore unknown, estimates of these true proportions must be supplied by a preliminary sampling study at the site background areas. This study must be conducted according to the same sampling and analysis protocol that will be used in the main study. Suppose a preliminary study based on 20 background samples and 20 site samples yields P_s and P_b estimates of 0.30 and 0.15, respectively. Therefore, $\overline{P} = (0.30 + 0.15) / 2 = 0.225$. Therefore, equation (1) in Box 4-22 is:

$$n = m = 2(1.96 + 0.84)^2 \, 0.225 \, (1 - 0.225) / 0.20^2 = 68.35$$

68.35 are rounded up to 69. Therefore, 69 background area samples and 69 site samples are needed. Because the 20 site and 20 background samples have already been collected, handled, and measured by the methods required for the full study, only 49 new site and 49 new background measurements need to be collected.

After the data evaluation process (Section 2.1.7) has been completed (i.e., once it has been determined that the data contain no errors, and have been collected, handled, and measured according to the DQO specifications), and the assumptions required for the two-sample test of proportions have been shown to be reasonable, then the two-sample test of proportions may be conducted. The test procedure is presented in Box 4-24. An example is provided in Box 4-25.

4.3 Individual Comparisons

The procedures described in Section 4.2 are appropriate for site-wide comparisons. In many instances, however, such site-wide pooling of the data is not appropriate. For example, at landfills subject to RCRA requirements, periodic monitoring is mandated as a means for timely detection and control of any release (U.S. EPA, 1992). Such monitoring efforts involve the comparison of individual measurements of COPCs at each compliance well to their corresponding background levels, and these individual comparisons are fundamentally different from site-wide comparisons. Some of these differences are:

- ❑ Individual comparisons are not concerned about the determination of the COPC-status of investigated chemicals. Instead, these comparisons are concerned about previously identified COPCs, whose above-background exceedances, even in a single compliance well, may be an indication of a release.
- ❑ Individual comparisons generally assume that both background and compliance wells are initially not impacted. The subsequent sampling and comparisons are thus intended to be a safeguard against any potential future releases, and thus cannot rely on the pooled site data from the past. Instead, they focus on data representing the most recent samples from each individual compliance well.
- ❑ Individual comparisons are not one-time tests; instead, they are intended to be implemented upon completion of each sampling round. As a result they are referred to as "partially sequential," i.e., fixed background data but sequential future sampling (Davis and McNichols, 1999). In contrast, site-wide comparisons are one-time tests to determine the COPC-status of investigated chemicals.

As discussed in Section 4.2.2, individual comparisons are bound to yield a high probability of falsely declaring above-background exceedances. As noted in Gibbons (1994), with a 95% confidence for an individual comparison, 10 repeated comparisons yield a site-wide confidence of $(0.95)^{10} = 0.60$. Such a confidence corresponds to a significance of $1 - 0.60 = 0.40$, or a 40% chance of falsely declaring a background exceedance, even when the observed site and background concentrations are from the same population. The site-wide confidence deteriorates rapidly as the number of monitoring wells and COPCs increase. This issue is quite problematic, because erroneous identification of background or groundwater protection standard exceedances leads to costly and unnecessary site assessments and characterizations.

The individual comparisons recommended in this section recognize the problem by providing procedures to meet the regulatory detection requirements, while controlling the level of ensuing errors. These procedures attempt to control the level of site-wide errors by selecting the appropriate individual error levels. For example, in the case of ten individual comparisons, the site-wide significance of 5% (or a 95% site-wide confidence) can be attained if individual significances are reduced to 0.5%. Such an individual significance yields a site-wide confidence of $(0.995)^{10} = 0.95$ or 95%. Under such an arrangement, the site-wide chance of falsely declaring an exceedance is reduced to 5%. This reduction, however, is attained by relaxing the bounds of individual comparisons, which results in a lower test power (i.e., the chances of missing true above-background exceedances have increased). To create a reasonable balance between the two types of error, the recommended procedures rely on the following:

- ❑ **Controlling the Number of Comparisons (*k*):** Attaining the desired site-wide confidence can lead to unacceptably low test powers, especially when the numbers of comparisons are high. Therefore, every attempt must be made to maintain *k* at a reasonable level. *k* is a function of the

BOX 4-24. Procedure for conducting the two-sample test of proportions (from U.S. EPA, 2000)

1. Stakeholders and regulators use the DQO process to select values of α, β, D, and C (recall that C is the concentration limit of interest).
2. Conduct a preliminary sampling and analysis study at the background and site areas to obtain estimates of the true proportions P_s and P_b of the site and background populations that exceed C. Then use the procedure in Box 4-22 to determine n and m, the necessary number of site and background measurements.
3. Collect, handle, and analyze the n and m samples, as specified in the sampling and analysis plan and the QAPP.
4. Suppose

   - n site measurements are denoted by $x_1, x_2, \ldots, x_n$
   - m background measurements are denoted by $y_1, y_2, \ldots, y_m$

   Note: In this document it is recommended that $n = m$. However, the following formulas are for the more general case where the number of site measurements, n, and the number of background measurements, m, are not equal.
5. Let k_s and k_b be the number of site and background measurements, respectively, that exceed C.
6. Compute $p_s = k_s/n$, which is the estimated proportion of the true distribution of site measurements that exceed C.
7. Compute $p_b = k_b/m$, which is the estimated proportion of the true distribution of background measurements that exceed C.
8. Compute

   $$p = (k_s + k_b) / (n + m)$$
9. Compute k_s, k_b, $n(1 - p_s)$, and $m(1 - p_b)$. If all of these quantities are greater than or equal to 5, continue with step 10. If not, seek assistance from a statistician, because the computations for the test are more complicated when these quantities are less than 5.
10. Compute the test statistic:

    $$Z_s = (p_s - p_b) / [p(1 - p)(1/n + 1/m)]^{1/2}$$
11. Use Table B-1 to find $Z_{1-\alpha}$
12. If $Z_s \geq Z_{1-\alpha}$ the test has declared that $P_s > P_b$, i.e., that the true proportion of the site measurements greater than the concentration value C is greater than the true proportion of the background measurements greater than C.

    If $Z_p < Z_{1-\alpha}$, then the data do not provide enough evidence to conclude that $P_s > P_b$. In that case, go to step 13.
13. Suppose the test declares that the data do not provide enough evidence to conclude that $P_s > P_b$. This conclusion may indicate (1) the chemical is not a COPC, or (2) the assumptions that underlie the test are not valid for the potentially impacted and background measurements, or (3) an insufficient number of measurements (n and m) were obtained for the test to detect the difference D that actually exists. The possibility that the test did not declare that $P_s > P_b$ due to items (2) or (3) should be evaluated. Review the DQO planning process records to make sure the number of measurements (n and m) agrees with what was determined necessary to detect the specified difference D. Use Equation 1 in Box 4-22 to recalculate the number of measurements required for the test. This can be done using VSP (Hassig et al., 2002) or DEFT (U.S. EPA, 2001) software. Those computations should be done using the estimates p_s and p_b in place of P_s and P_b, respectively. If the new n value is greater than what was used to compute the test statistic, collect and analyze the additional samples necessary to achieve the new n value and redo the test.

BOX 4-25. Example of the two-sample test of proportions

1. Suppose the stakeholders and regulators specified that $\alpha = 0.025$, $\beta = 0.20$, D = 0.20 and $C = 1$ ppb for the chemical of interest.
2. Also suppose that a preliminary study was conducted at the site and background area to obtain estimates of the true proportions P_s and P_b. Suppose these estimates were 0.30 and 0.15, respectively. Then, as illustrated in Box 4-23, 69 measurements are needed from the site area and from the background area.
3. A total of 138 measurements are obtained. Suppose $k_b = 19$ of the 69 background measurements were greater than *C,* i.e., greater than 1 ppb. Furthermore, suppose that $k_p = 24$ of the site measurements were greater than *C.* Therefore,

   p_b = 19/69 = 0.275
   p_s = 24/69 = 0.347
   $p = (k_p + k_b) / (n + m) = (19 + 24) / (69 + 69) = 0.3116$
4. Also,

   $mp_b = 69(0.275) = 19$
   $np_s = 69(0.347) = 24$
   $m(1 - p_b) = 69(1–0.275) = 50$
   $n(1 - p_s = 69(1–0.347) = 45$

   All the above numbers are greater than 5. Therefore, we continue on with the test as described in Box 4-24.
5. The test statistic is computed as follows:

   $$\begin{aligned} Z_p &= (p_s - p_b) / [p(1 - p)(1/n + 1/m)]^{1/2} \\ &= (0.347 - 0.275) / [0.3116(1 - 0.3116)(1/69 + 1/69)]^{1/2} \\ &= 0.072 / [0.2145*(0.014493 + 0.014493)]^{1/2} \\ &= 0.072 / 0.0789 \\ &= 0.913 \end{aligned}$$
6. From Table B-1 we find that $Z_{1-\alpha} = Z_{0.975} = 1.96$
7. As $Z_p < 1.96$, i.e., $0.913 < 1.96$, the data do not provide sufficient information for the test to reject H_o and declare that the chemical is a COPC.
8. Equation 1 in Box 4-22 is then re-evaluated to determine whether the test did not reach a statistically significant conclusion because the datasets have fewer measurements than are required to achieve the power of $1 - \beta = 0.80$ when D = 0.20 (at the $1 - \alpha = 0.975$ confidence level). We obtain:

   $$n = m = 2(1.96 + 0.84)^2\, 0.3116(1 - 0.3116) / 0.2^2 = 84.09$$

   This indicates that 85 site measurements (n) and 85 background measurements (m) are needed; therefore, the number of measurements ($n = m = 69$) is insufficient.

In conclusion, based on the data, the true difference D is estimated to be 0.347 – 0.275 = 0.072. However, the two-sample test of proportions was not able to declare that this difference is large enough to conclude that $P_s > P_b$.

numbers of compliance wells (n_{wells}) and targeted COPCs (n_{COPCs}). If COPCs are independent of each other, then $k = n_{wells} \times n_{COPCs}$. In detection monitoring problems, however, all COPCs are associated with the same source, and thus are highly correlated. Under such conditions, k approaches n_{wells}.

Timely detection of releases is the main objective of these monitoring efforts, so compliance wells should be limited to those located immediately downgradient of potential sources. Furthermore, the list of targeted COPCs should be limited to those chemicals that are most persistent and least sorptive, i.e., COPCs whose breakthrough can be viewed as the earliest indicators of a release. Such a combination of compliance wells and indictor COPCs provides a reliable basis for timely detection of future releases, while controlling the balance between the test confidence and power.

- **Verification Resampling Plan:** The reliability of observed individual exceedances can be further ensured through verification resampling. Such sampling efforts provide a number of benefits. First, resampling can verify a true exceedance in a timely manner. Second, erroneous exceedances and unnecessary assessments can be avoided. Third, reasonable site-wide confidence can be attained while preserving the test power, i.e., both types of error (false positives and negatives) can be maintained at low levels. The regulatory community has recognized the benefits of verification resampling. Davis and McNichols (1999) cite the California Plan (*Barclay's Code of California Regulations*, 1999), in which an exceedance is verified by three or more additional resamples. Under the California Plan, if any of the three or more resamples exceeds the background prediction limit, the exceedance is verified. U.S. EPA (1998a) suggests a "modified California" plan, according to which, an exceedance is verified, if at least two of the three resamples are above the background prediction limit. Other verification plans include the so-called "1-of-*m*" plan, according to which an exceedance is verified if all of the subsequent $m - 1$ resamples are above background (Davis and McNichols, 1999). Gibbons (1994) demonstrates that the "1-of-*m*" plans provide more powerful test with higher site-wide confidence, when compared to similar procedures under the California Plan.

- **Independent Resamples:** The resampled data from the same compliance well are assumed to be independent of each other, i.e., temporally uncorrelated. As a result, resampling cannot be performed immediately after the original compliance samples. For this reason, Gibbons (1994) suggests an interval of at least three months between resampling events to ensure their independence. In aquifers with low transmissivity, longer intervals may be necessary to obtain uncorrelated resampling data.

4.3.1 Common Individual Comparative Statistical Tests

An initial, tentative selection of the most appropriate individual comparative test(s) should be made during the DQO planning process. This selection should consider: (1) the number of compliance wells and COPCs that need to be targeted for detection monitoring, (2) the site-wide confidence as specified by DQO performance goals, (3) verification resampling procedures, (4) the particular distribution (normal or lognormal) of the background/prerelease data, and (5) information in published statistical papers that demonstrate the performance of the candidate tests for various data distributions and contamination scenarios (e.g., Gibbons, 1994; Davis and McNichols, 1999). After exploratory analyses of available background and prerelease data as discussed in Sections 2 and 3, a final selection of the statistical test(s) can be made.

The recommended tests should be applied by focusing on the most critical compliance wells and indicator COPCs, as discussed above. Large number of comparisons would lead to high probability of error. In general, use the parametric prediction limit test (Section 4.3.4) if the available background and prerelease data or their transformations (e.g. logarithms) indicate a tendency toward normal distribution. The normality of the investigated data must be assessed using probability plots (Section 3.3.4) and/or an appropriate normality test (Appendix A.1). If the investigated data or their transformations proved to be not normally distributed or if they contain more than 15% nondetects, the nonparametric prediction limit test (Section 4.3.5) should be pursued.

4.3.2 Statistical Testing Approaches Not Recommended for Individual Comparisons

As noted in Section 4.2.3, site-wide comparative statistical procedures are unsuitable for individual comparisons. Gibbons (1994, page 258) provides further discussion about tests and procedures that should be avoided during individual comparisons. Approaches that are not recommended include two main classes of tests:

- Individual comparisons devoid of verification resampling are not recommended. As discussed in Sections 4.2.2 and 4.3, individual comparisons with no verification scheme yield elevated chances of false exceedances. Given the costly consequences of erroneous exceedances, every effort must be made to avoid individual comparisons without resampling plans.

- Statistical comparison tests that require data pooling, including site-wide comparisons (Section 4.2) and ANOVA (Section 4.4), are not recommended for individual comparisons because they are generally nonsequential. Furthermore, implementation of verification resampling on a large-scale, such as resampling the entire site, is time-consuming and may lead to significant delays in confirming true exceedances.

4.3.3 Recommended Individual Comparative Statistical Tests

Recommended individual comparisons provide tools that control the problem of elevated false positives rates by (1) pursuing individual confidence levels consistent with the desired site-wide confidence, (2) focusing on the critical compliance wells and indicator COPCs in order to limit the number of simultaneous individual comparisons, and (3) implementing verification resampling procedures. These tests are discussed in this section. Following are some general words of advice about using individual comparisons to decide whether a release has occurred:

- ***Not Suitable for Site-Wide Comparisons:*** Individual comparisons are intended for detecting the occurrence of a release. These methods do not rely on the collective site and background datasets, and thus, are not suitable for characterization or assessment purposes. Site-wide comparisons are discussed in Section 4.2.

- ***Assumption Verification and Limitations*****:** The assumptions that underlie individual comparisons should always be reviewed. As discussed in Section 4.1.2, limitations may require the use of alternative methods depending on the observed background and prerelease conditions.

- ***Datasets*****:** An adequate background and/or prerelease dataset should be available. Ideally, these datasets are unbiased and representative of the aquifers of interest. Limited background data can yield unreliable test results.

- ***Preferred Site-Wide Tests:***

    - Use the parametric prediction limit test (Section 4.3.4) if the background/prerelease site data or their transformation (e.g., logarithms) display a tendency to normal distribution. Normality of the investigated data or their transformations should be tested using probability plots (Section 3.3.4) and/or normality tests (Appendix A.1).

    - Use the nonparametric prediction limit test (Section 4.3.5) if the background/prerelease data or their transformations do not display any normal tendency, or if they contain more than 15% nondetects.

- ***Direct Statistician Involvement:*** Consult an experienced environmental statistician whenever disputes regarding the most appropriate statistical testing methods arise. In some cases, the characteristics of the site and background datasets may violate the assumptions required for the proposed statistical test methods. In these cases it may be necessary to implement the advanced methods described in Section 4.4, including geostatistical procedures to accommodate spatially correlated data. Under such conditions, involvement of an experienced statistician or geostatistician is essential.

4.3.4 Parametric Prediction Limit Test

Groundwater Release Scenario

The site includes a waste disposal landfill, which is subject to a detection monitoring requirements. Failure of the landfill containment system could release a range of COPCs into the aquifer. COPCs with higher solubility and lower sorptivity are considered the most rapid indicators of a release. However, these indicator COPCs are also known to occur as natural and/or anthropogenic background chemicals in groundwater. Exceedances of background levels in the downgradient site compliance monitoring wells, if verified, will trigger a site assessment investigation.

Role of the Data Quality Objectives Process

The stakeholders and regulators used the DQO planning process to agree:

- ❑ On the methods that should be used to collect, handle, prepare, and measure the chemicals in groundwater samples.
- ❑ On the choice of selected indicator COPC(s).
- ❑ On the number and location of site compliance wells.
- ❑ On the frequency of detection monitoring.
- ❑ On the resampling plan to verify an above-background exceedance.
- ❑ On the fact the background and prerelease data representing the indicator COPC or their transformations (e.g., logarithms) are normally distributed.
- ❑ On the value of the site-wide confidence level and its corresponding individual confidence level (see the following subsection, "Guidance on Implementing the Parametric Prediction Limit Test").

The parametric prediction limit is appropriate for this scenario because it is designed for normally distributed data and incorporates the selected verification resampling plan.

Advantages and Disadvantages

- ❑ The parametric prediction limit test allows direct incorporation of the desired site-wide and individual confidence levels based on the number of expected comparisons.
- ❑ The parametric prediction limit test includes the verification resampling. In this section, examples of the "1-of-*m*" plans are presented. Alternative verification resampling plans are described by Gibbons (1994).
- ❑ The parametric prediction limit test requires the normality of the background/prerelease data or their transformations (e.g., logarithms). Prediction limits based on the Poisson distribution are also discussed in Gibbons (1994, page 79). If no specific distribution can be assumed, such as in the case of a large number of nondetects among the background and/or prerelease measurements, then the nonparametric prediction limit test (Section 4.3.5) should be pursued.

Guidance on Implementing the Parametric Prediction Limit Test

The procedure for conducting the parametric prediction limit test is shown in Box 4-26. Box 4-27 provides an example of its use. Given the parametric nature of this approach, normality of the available background/prerelease concentrations of the indicator COPCs or their transformations (e.g., logarithms) should be tested by appropriate graphical and/or statistical tests (Section 3 and Appendix A.1). Upon confirmation of the normality of the background data, the parametric prediction limit test can be implemented.

The objective of the parametric prediction limit test is to compute the background prediction limit of each indicator COPC. This limit is then used for comparison of individual sampling results from various compliance wells. The background prediction limit is defined as $\bar{x} + Ks$, where $\bar{x}$ and s are the mean and standard deviation of background/prerelease concentrations of the indicator COPC, and K is the prediction limit factor.

The background prediction limit factor, K, depends on a number of factors, including: (a) the

BOX 4-26. Procedure for conducting the parametric prediction limit test

1. Stakeholders and regulators use the DQO process to select values of site-wide α, number of compliance wells, number of indicator COPCs, and the verification resampling scheme.
2. Count the available background and site presample data (n). Compute the background/presample mean ($\overline{x}$) and standard deviation (s) of all indicator COPCs concentrations.
3. Determine the prediction limit factor K. Tables 4-7 and 4-8 provide K values for site-wide α of 0.05 for 1 and 10 indictor COPCs, respectively, assuming a "1-of-2" verification resampling plan. K values for other combinations of α, number of indictor COPCs, and verification sampling plans are provided by Gibbons (1994, pages 23-31).
4. Compute the background prediction limit as $\overline{x} + Ks$.
5. Compare every single future indicator COPC measurement to its corresponding computed background prediction limit.
6. If an exceedance is observed in a compliance well, resample the well allowing at least a three-month interval between resampling events. Verify the exceedance according to the selected plan.
7. If the exceedance is verified, initiate site assessment.

BOX 4-27. Example of the parametric prediction limit test

1. Suppose the stakeholders and regulators specified that site-wide α = 0.05, 10 compliance wells, a single indicator COPC (previously determined to be the most mobile COPC), and a "1-of-2" verification resampling plan.
2. There are 20 background/prerelease measurements of the indicator COPC, with a mean of 10 ppb and a standard deviation of 5 ppb.
3. K is determined using Table 4-7, yielding a value of 1.67.
4. The background prediction limit is computed as 10 + 1.67(5) = 18.35 ppb.
5. Future indicator COPC measurements are then compared to 18.35.
6. An initial exceedance of 20 ppb is observed in a compliance well; however, the subsequent resample taken three months later is only 11 ppb. Therefore, the original exceedance is deemed as unverified.

selected site-wide significance, α (site-wide confidence = 1 – α); (b) number of background/prerelease measurements, n; (c) the selected resampling verification plan; and (d) the number of comparisons, k. Table 4-7 provides K values for individual comparisons of a single indicator COPC at a site-wide significance of 0.05 or a 95% confidence with a "1-in-2" resampling plan (i.e., an exceedance is verified when the resampling result exceeds the background prediction limit). Table 4-8 lists K values for the same conditions when 10 indicator COPCs are evaluated. Comparisons of these two tables indicate that a large set of indicator COPCs, especially when the number of background data is small, can lead to very large K values. Such large K values can lead to low test power (i.e., a high probability of missing true exceedances). Therefore, every effort must be made to limit the number of indicator COPCs. For parametric K values under a variety of verification plans and comparisons, readers are referred to Gibbons (1994, pages 23-31).

4.3.5 Nonparametric Prediction Limit Test

Groundwater Release Scenario

Similar to the previous case (Section 4.3.4), the site includes a waste disposal landfill subject to

TABLE 4-7. Simultaneous normal prediction limit factors (K) (site-wide α= 0.05/one indicator COPC/"1-of-2" verification resampling plan)

n (Background Sample Size)	Number of Compliance Wells									
	5	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**
4	2.47	3.00	3.31	3.52	3.69	3.82	3.94	4.03	4.12	4.19
8	1.72	2.03	2.21	2.33	2.43	2.51	2.57	2.63	2.67	2.72
12	1.56	1.82	1.97	2.07	2.15	2.21	2.27	2.31	2.36	2.39
16	1.48	1.72	1.86	1.95	2.03	2.08	2.13	2.18	2.13	2.25
20	1.44	1.67	1.8	1.89	1.95	2.01	2.06	2.1	2.08	2.16
24	1.41	1.63	1.76	1.84	1.91	1.96	2.01	2.05	2.04	2.11
28	1.39	1.61	1.73	1.81	1.88	1.93	1.97	2.01	2.02	2.07
32	1.38	1.59	1.71	1.79	1.85	1.90	1.95	1.98	1.99	2.04
36	1.37	1.58	1.69	1.77	1.84	1.89	1.93	1.96	1.98	2.02
40	1.36	1.56	1.68	1.76	1.82	1.87	1.91	1.95	1.96	2.00
44	1.35	1.56	1.67	1.75	1.81	1.86	1.9	1.93	1.95	1.99
48	1.34	1.55	1.66	1.74	1.8	1.85	1.89	1.92	1.94	1.98
52	1.34	1.54	1.66	1.73	1.79	1.84	1.88	1.91	1.94	1.97
56	1.33	1.54	1.65	1.73	1.79	1.83	1.87	1.91	1.93	1.96
60	1.33	1.53	1.64	1.72	1.78	1.83	1.87	1.90	1.92	1.95
64	1.33	1.52	1.64	1.72	1.77	1.82	1.86	1.89	1.92	1.95
68	1.32	1.52	1.64	1.71	1.77	1.82	1.85	1.89	1.91	1.94
72	1.32	1.52	1.63	1.71	1.77	1.81	1.85	1.88	1.91	1.94
76	1.32	1.52	1.63	1.71	1.76	1.81	1.85	1.88	1.90	1.93
80	1.32	1.52	1.63	1.71	1.76	1.80	1.84	1.88	1.90	1.93
84	1.32	1.51	1.62	1.71	1.76	1.80	1.84	1.87	1.90	1.93
88	1.31	1.51	1.62	1.71	1.75	1.80	1.84	1.87	1.89	1.92
92	1.31	1.51	1.62	1.69	1.75	1.80	1.83	1.87	1.89	1.92
96	1.31	1.51	1.62	1.69	1.75	1.79	1.83	1.86	1.89	1.92
100	1.31	1.51	1.62	1.69	1.75	1.79	1.83	1.86	1.89	1.91

Adapted from: Gibbons (1994, page 23).

TABLE 4-8. Simultaneous normal prediction limit factors (K) (site-wide α= 0.05/10 indicator COPC/"1-of-3" verification resampling plan)

n (Background Sample Size)	Number of Compliance Wells									
	5	10	15	20	25	30	35	40	45	50
4	5.86	6.96	7.60	8.06	8.42	8.71	8.95	9.15	9.34	9.50
8	2.98	3.36	3.59	3.75	3.87	3.97	4.06	4.14	4.20	4.26
12	2.51	2.78	2.94	3.06	3.15	3.22	3.28	3.34	3.38	3.43
16	2.31	2.55	2.69	2.78	2.86	2.92	2.97	3.02	3.06	3.09
20	2.21	2.42	2.55	2.64	2.70	2.76	2.81	2.85	2.88	2.91
24	2.14	2.35	2.46	2.55	2.61	2.66	2.70	2.74	2.77	2.80
28	2.10	2.29	2.40	2.48	2.54	2.59	2.63	2.67	2.70	2.73
32	2.07	2.25	2.36	2.44	2.50	2.54	2.58	2.62	2.65	2.67
36	2.04	2.23	2.33	2.40	2.46	2.50	2.54	2.58	2.60	2.63
40	2.02	2.20	2.30	2.38	2.43	2.48	2.51	2.54	2.57	2.60
44	2.01	2.18	2.28	2.35	2.41	2.45	2.49	2.52	2.55	2.57
48	1.99	2.17	2.27	2.34	2.39	2.43	2.47	2.50	2.53	2.55
52	1.98	2.16	2.25	2.32	2.37	2.42	2.45	2.48	2.51	2.53
56	1.98	2.15	2.24	2.31	2.36	2.40	2.44	2.47	2.49	2.52
60	1.97	2.14	2.23	2.30	2.35	2.39	2.42	2.45	2.48	2.50
64	1.96	2.13	2.22	2.29	2.34	2.38	2.41	2.44	2.47	2.49
68	1.95	2.12	2.21	2.28	2.33	2.37	2.40	2.43	2.46	2.48
72	1.95	2.11	2.21	2.27	2.32	2.36	2.40	2.43	2.45	2.47
76	1.94	2.11	2.20	2.27	2.32	2.36	2.39	2.42	2.44	2.47
80	1.94	2.10	2.20	2.26	2.31	2.35	2.38	2.41	2.44	2.46
84	1.94	2.10	2.19	2.26	2.30	2.34	2.38	2.41	2.43	2.45
88	1.93	2.09	2.19	2.25	2.30	2.34	2.37	2.40	2.42	2.45
92	1.93	2.09	2.18	2.25	2.29	2.33	2.37	2.39	2.42	2.44
96	1.93	2.09	2.18	2.24	2.29	2.33	2.36	2.39	2.41	2.44
100	1.92	2.08	2.18	2.24	2.29	2.33	2.36	2.39	2.41	2.43

Adapted from: Gibbons (1994, page 26).

detection monitoring requirements. Detection of above-background concentrations of the indicator COPCs may indicate failure of the landfill containment system. Background exceedances, if verified, trigger a site assessment investigation. Unlike the previous case, however, the available background/prerelease data of indictor COPCs include a large number of nondetects. Therefore, normality of the data or their transformation cannot be assumed.

Role of the Data Quality Objectives Process

The stakeholders and regulators used the DQO planning process to agree:

- On the methods that should be used to collect, handle, prepare, and measure the chemicals in groundwater samples.
- On the choice of selected indicator COPC(s).
- On the number and location of site compliance wells.
- On the frequency of the detection monitoring.
- On the resampling plan to verify an above-background exceedance.
- On the fact the type of the statistical distribution of background and prerelease data of the indicator COPCs or their transformations (e.g., logarithms) cannot be presumed.
- On the desired site-wide confidence level (see the following subsection, "Guidance on Implementing the Nonparametric Prediction Limit Test").

The nonparametric prediction limit test is appropriate for this scenario because it is does not require any specific distributional assumption and incorporates the selected verification resampling plan.

Advantages and Disadvantages

- The nonparametric prediction limit test allows the use of background datasets containing large number of nondetects. The test is applicable to any case where no parametric assumption can be made about indicator COPCs.
- Davis and McNichols (1999) demonstrate that the recommended power of the nonparametric limit test is comparable to similar parametric tests, such as those described in Section 4.3.4.
- The recommended nonparametric prediction limit test incorporates verification resampling. In this section, examples of the original and modified California Plans (Section 4.3) are presented (*Barclay's Code of California Regulations*, 1999). Results associated with a wide range of alternative verification resampling plans are provided by Gibbons (1994, pages 52-75) and Davis and McNichols (1999).
- The nonparametric prediction limit for background is defined as the largest value out of *n* background/presample measurements. Under this definition, the confidence associated with the prediction limit is a function of *n*, which cannot be adjusted except through more background sampling (Gibbons, 1994, page 33). Therefore, in cases of a large number of comparisons, the only approach for improving site-wide confidence is through additional background measurements.

Guidance on Implementing the Nonparametric Prediction Limit Test

The procedure for conducting the nonparametric prediction limit test is shown in Box 4-28. Box 4-29 provides an example of its use. The background prediction limit is simply defined as the largest background/prerelease measurement. The resulting site-wide significance depends on: (1) number of background/prerelease measurements, *n*, and (2) the number of comparisons, *k*. Tables 4-9 and 4-10 provide site-wide significance levels under the California Plan (verified exceedance when any of three resampling results exceed) and the modified California Plan (verified exceedance when at least two of three resampling results exceed), respectively.

BOX 4-28. Procedure for conducting the nonparametric prediction limit test

1. Stakeholders and regulators use the DQO process to select values of the target site-wide significance (α), number of compliance wells, number of indicator COPCs, and the verification resampling scheme.
2. Identify the number of background measurements (n) to attain the targeted site-wide significance (α), based on the number of comparisons (number of compliance wells x number of indictor COPCs), and the selected verification resampling plan. Tables 4-9 and 4-10 provide α values based on the original and modified California Plans. Test confidence ($1-\alpha$) values based on other verification resampling plans are provided by Gibbons (1994, pages 52-75).
3. If less than n background/presample measurements are available, then collect additional data from background or nonimpacted site wells and analyze for the indicator COPCs.
4. Identify the highest concentration of indicator COPCs among background/presample data as their corresponding nonparametric background prediction limit.
5. Compare every single future indicator COPC measurement to its corresponding computed background prediction limit.
6. If an exceedance is observed in a compliance well, resample the well allowing at least a three-month interval between resampling events. Verify the exceedance according to the selected plan.
7. If the exceedance is verified, initiate site assessment.

BOX 4-29. Example of the nonparametric prediction limit test

1. Suppose the stakeholders and regulators specified that site-wide $\alpha = 0.03$, four compliance wells, a single indicator COPC (previously determined to be the fastest traveling COPC), and the modified California verification resampling plan (follow an exceedance with three resamples; verify an exceedance if at least two of the resampling results exceed the prediction limit).
2. Using Table 4-10, the minimum number of background/presample data should be approximated. For this purpose, along the column associated with the selected number of comparisons (i.e., four), identify the cell containing the nearest significance to the selected α (0.03). In this example, this significance is 0.03072, which is associated with a minimum of 10 background samples. Only eight background/prerelease measurements are available; therefore, two additional samples from background and nonimpacted wells are collected.
3. The highest concentration of the indicator COPC among the 10 background/prerelease samples is 18 ppb; this is defined as the nonparametric background prediction limit.
4. Future indicator COPC measurements are then compared to 18 ppb.
5. An initial exceedance of 20 ppb is observed in a compliance well; however, results of three subsequent resampling events at three-month intervals are 15, 21, and 11 ppb, respectively. Therefore, the original exceedance is deemed as unverified.

Note that the modified California Plan provides higher site-wide confidence when compared to the original California Plan. Also note that reasonable site-wide confidence levels require a large number of background samples when large numbers of comparisons are planned. In certain cases, generating such large background datasets within a reasonable period of time may not be economically feasible.

4.4 Advanced Statistical Procedures

As noted in Section 4.1, the recommended site-wide and individual comparative procedures are based on certain assumptions that limit their applicability. These methods are not designed for spatially or temporally correlated, nonstationary, or clustered data. Analyses of such datasets require more advanced techniques. These techniques include:

TABLE 4-9. Smallest simultaneous site-wide significance (α) of nonparametric prediction limit tests (California Plan with three resamples)

n (Background Sample Size	Number of Comparisons					
	2	4	7	10	20	40
10	0.05082	0.09240	0.14313	0.18429	0.28204	0.39510
20	0.01588	0.03063	0.05096	0.06948	0.12151	0.19761
30	0.00762	0.01496	0.02548	0.03547	0.06557	0.11494
50	0.00292	0.00579	0.01002	0.01416	0.02733	0.51220
75	0.00134	0.00267	0.00464	0.00659	0.01296	0.25050
100	0.00076	0.00153	0.00266	0.00379	0.00750	0.01470
150	0.00034	0.00069	0.00120	0.00172	0.00342	0.00677
200	0.00020	0.00039	0.00068	0.00098	0.00194	0.00387

Adapted from: Davis and McNichols (1999).

TABLE 4-10. Smallest simultaneous site-wide significance (α) of nonparametric prediction limit tests (modified California Plan with three resamples)

n (Background Sample Size	Number of Comparisons					
	2	4	7	10	20	40
10	0.04638	0.03072	0.04942	0.06565	0.10839	0.16647
20	0.00298	0.00587	0.01120	0.13980	0.26100	0.46620
30	0.00101	0.00201	0.00348	0.00492	0.00954	0.18050
50	0.00024	0.00049	0.00085	0.00121	0.00239	0.00470
75	0.00008	0.00015	0.00027	0.00038	0.00076	0.00150
100	0.00030	0.00070	0.00012	0.00016	0.00033	0.00066
150	0.00010	0.00002	0.00004	0.00005	0.00001	0.00020
200	0.00000	0.00001	0.00002	0.00002	0.00004	0.00009

Adapted from: Davis and McNichols (1999).

- **Analysis of Variance or ANOVA:** In many instances, available data are either spatially variable (e.g., mean background concentrations vary from well to well) or spatially clustered (e.g., some wells are sampled significantly more frequently than others). Under such cases, the recommended site-wide comparisons (Section 4.2) are not appropriate. Instead, ANOVA tests (McBean and Rovers, 1998) should be considered. (Note that although ANOVA tests are not recommended for individual comparisons, they can be useful for dealing with spatially variable or clustered data.) ANOVA tests allow simultaneous comparisons of multiple sets of data. The Kruskal-Wallis test (Box 2-5) is an example of a nonparametric ANOVA test. ANOVA tests can be used to determine the significance of spatial variability among background wells. For this purpose, data from each background well are grouped as a separate dataset. The well-specific datasets then can be compared using an appropriate ANOVA test. Another example involves cases of spatially clustered data. Under such conditions, site-wide comparisons become unduly influenced by measurements from frequently sampled wells. The effects of such clustering can be reduced if data from each well are treated as a separate dataset. Such well-specific datasets can then be compared by an appropriate ANOVA test.

- **Geostatistics:** In many cases, background data are spatially correlated. Such correlations limit the application of the recommended site-wide and individual comparisons (Section 4.1). Under such instances, geostatistical methods (Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989) can be used. These methods incorporate the spatial correlation of the data into the estimation process. If background data display a strong spatial correlation, the concentration of indicator COPCs can be estimated at site well locations. Comparison of these estimated values to actual measured concentrations at various site wells then can be performed in order to determine the likelihood of an exceedance.

- ❑ **Statistical Classification Methods:** The recommended site-wide and individual comparisons (Sections 4.2 and 4.3, respectively) are basically univariate techniques in which measurements of each chemical or COPC are considered separately. In many instances, however, background chemicals may display particular compositional patterns different from those of site-related COPCs. Under such cases, multivariate statistical classification methods can be used to distinguish background data from site data. Among these multivariate techniques are principal component analysis (Pielou, 1984), cluster analysis (Harman, 1970), and finite mixture distributions (Everitt and Hand, 1981), which create empirical groupings of the investigated data based on the similarity of their compositional patterns. Other methods such as discriminate function analysis (Hand, 1981; Gibbons, 1994) group investigated sample results according to known patterns associated with background and various site-related releases.

Applications of these and other advanced statistical techniques require direct involvement and advice of an experienced environmental statistician or geostatistician.

4.5 References

Barclay's California Code of Regulations. 1991. Title 22. Section 66264.97 (concerning hazardous waste facilities) and Title 23, Section 2550.7(e)(8) (concerning solid waste facilities). San Francisco: Barclay's Law Publications.

Conover, W.J. 1998. *Practical Nonparametric Statistics*. 3rd ed. Wiley Series in Probability and Statistics, Applied Probability and Statistics Section. New York: Wiley & Sons.

Davis, C.B., and R.J. McNichols. 1999. Simultaneous Nonparametric Prediction Limits. *Technometrics*. Vol. 41, No. 2, 90-112. May.

Everitt, B.S., and D. J. Hand. 1981. *Finite Mixture Distributions*. London: Chapman and Hall.

Gibbons, R.D. 1994. *Statistical Methods for Groundwater Monitoring.* New York: Wiley.

Gibbons, R.D., and D.E. Coleman. 2001. *Statistical Methods for Detection and Quantification of Environmental Contamination*. New York: Wiley.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring.* New York: Van Nostrand Reinhold.

Gilbert, R.O., and J.C. Simpson. 1990. Statistical Sampling and Analysis Issues and Needs for Testing Attainment of Background-Based Cleanup Standards at Superfund Sites. Pp. 1–16 in *Proceedings of the Workshop on Superfund Hazardous Waste: Statistical Issues in Characterizing a Site: Protocols, Tools, and Research Needs,* eds. H. Lacayo, R.J. Nadeau, G.P. Patil, and L. Zaragoza. University Park, PA: Pennsylvania State University, Center for Statistical Ecology and Environmental Statistics.

Hand, D.J. 1981. *Discrimination and Classification*. New York: Wiley.

Harman, H. 1970. *Modern Factor Analysis*. Chicago: University of Chicago Press.

Hassig, N.L., R.F. O'Brien, J.E. Wilson, B.A. Pulsipher, R.O. Gilbert, C.A. McKinstry, D.K. Carlson, and D.J.Bates. 2002. *Visual Sample Plan, Version 2.0, User's Guide.* Pacific Northwest National Laboratory. PNNL-14002. September.

Helsel, D.R., and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. New York: Elsevier.

Isaaks, E.H., and R.M. Srivastava. 1989. *An Introduction to Applied Geostatistics.* New York: Oxford University Press.

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

McBean, E.A., and F. Rovers. 1998. *Statistical Procedures for Analysis of Environmental Monitoring Data and Risk Assessment.* Vol. 3. Prentice Hall Environmental Series. Pearson Education. July.

Nuclear Regulatory Commission, U.S. Department of Energy, U.S. Environmental Protection Agency, and Department of Defense (NRC et al.). 1997. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*. Final report. NUREG-1575; EPA 402/R-97-016. December. http://www.epa.gov/radiation/marssim/.

O'Brien, R.F., and R.O. Gilbert. 1997. Comparing Sites to Background, Part 1. *Environmental Testing & Analysis* (September/October): 12-14.

Palachek, A.D., D.R. Weier, T.R. Gatliffe, D.M. Splett, and D.K. Sullivan. 1994. *Statistical Methodology for Determining Contaminants of Concern by Comparison of Background and Site Data with Applications to Operable Unit 2.* Internal report. SA-94-010. Golden, CO: EG&G Rocky Flats Inc., Rocky Flats Plant, Statistical Applications.

Pielou, E.C. 1984. *The Interpretation of Ecological Data: A Primer on Classification and Ordination*. New York: Wiley.

United States Environmental Protection Agency (U.S. EPA). 1992. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities: Addendum to Interim Final Guidance*. Washington, DC: Office of Solid Waste. http://www.epa.gov/swertio1/chartext_edu.htm#stats.

United States Environmental Protection Agency (U.S. EPA). 1994. *Statistical Methods for Evaluating the Attainment of Cleanup Standards, Vol. 3: Reference-Based Standards for Soils and Solid Media*. EPA 230-R-94-004. Washington, DC: Office of Policy, Planning, and Evaluation. http://www.epa.gov/swertio1/chartext_edu.htm#stats.

United States Environmental Protection Agency (U.S. EPA). 1998a. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities: Unified Guidance*. Draft. Washington, DC: Office of Solid Waste. http://www.epa.gov/swertio1/chartext_edu.htm#stats.

United States Environmental Protection Agency (U.S. EPA). 1998b. *Statistical Tests for Background Comparison at Hazardous Waste Sites.* Supplemental Guidance to RAGS: Region 4 Bulletins-Addition #1. Atlanta, GA. November. http://www.epa.gov/swertio1/chartext_edu.htm#stats.

United States Environmental Protection Agency (U.S. EPA). 2000. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis. EPA QA/G-9, QA 00 Update.* EPA/600/R-96/084. Office of Environmental Information. July. http://www.epa.gov/quality1/qa_docs.html.

United States Environmental Protection Agency (U.S. EPA). 2001. *Data Quality Objectives Decision Error Feasibility Trials Software (DEFT)-User's Guide.* EPA QA/G-4D. EPA/240.B-01/007. http://www.epa.gov/quality.

5. CASE STUDY

The case study presented in this section demonstrates the application of the methods described in Sections 3 and 4 to a groundwater investigation conducted at a hypothetical Navy chemical release site. Background analysis and statistical comparison were necessary to assess whether the hypothetical site groundwater is impacted by COPCs released to soil at the site.

Products containing various metals were stored at the investigation site, and pipelines formerly transported leaded gasoline and other petroleum products across the site area. Arsenic, copper, lead, and zinc were detected in surface soil samples collected at the site, and in groundwater samples collected from monitoring wells south of the impacted surface soil area. All copper and zinc concentrations detected in groundwater were below the screening criteria identified for the site, whereas arsenic and lead concentrations detected in groundwater samples collected in one area of the site exceeded the screening criteria. Soil in the region is known to contain relatively high concentrations of naturally occurring arsenic and lead; therefore, some or all of the arsenic or lead detected in groundwater could be attributable to natural sources.

Note: the case study is intended only to demonstrate application of the general methodology. The presence or absence of a specific technique in the case study is not intended to reflect upon the utility of that technique. Furthermore, the datasets used for the case study are not intended as examples of the minimum or optimum quantity of data needed for background analysis or statistical comparative testing.

5.1 Identifying Background Sampling Locations and Concentration Ranges

Figure 5-1 shows the location of the hypothetical investigation site and a nearby site selected for evaluation as a potential background site. The investigation site (Figure 5-2) consists of two parcels of Navy property separated by a major highway: Area 1 on the north, and Area 2 on the south. The investigation and potential background sites are located near an ocean inlet used as a harbor by the Navy.

5.1.1 Investigation Site Evaluation

COPC (arsenic and lead) concentrations above the screening criteria identified for the site were detected only in groundwater samples collected from the Area 2 monitoring wells. Therefore, background analysis is necessary to assess whether the Area 2 groundwater is impacted by a site-related chemical release. The Area 1 monitoring wells were initially considered as potential background wells because they appeared to be located upgradient of the Area 2 wells. However, a comprehensive evaluation of the geology, hydrogeology, and geochemistry of the investigation site area was necessary to determine whether the Area 1 wells could be considered background wells. The information and data collected during the evaluation were integrated to develop a CSM for the site. As will be shown later in the case study, the CSM indicated that groundwater conditions in Area 1 are not representative of background conditions at Area 2.

Contaminant Sources

As shown on Figure 5-2, two potential sources of metals contamination were identified at the investigation site: (1) A former Hazardous Materials Storage Facility located near the center of Area 1, and (2) a pair of pipelines that formerly carried leaded gasoline and other petroleum products along the north boundary of Area 2 (one pipeline is now inactive and the other has been removed). Arsenic, copper, lead, and zinc were detected in surface soil samples collected in the former Hazardous Materials Storage Facility area. No soil sampling data are available for the pipeline area. Arsenic and lead were detected at concentrations above screening criteria in groundwater samples collected from some of the Area 2 monitoring

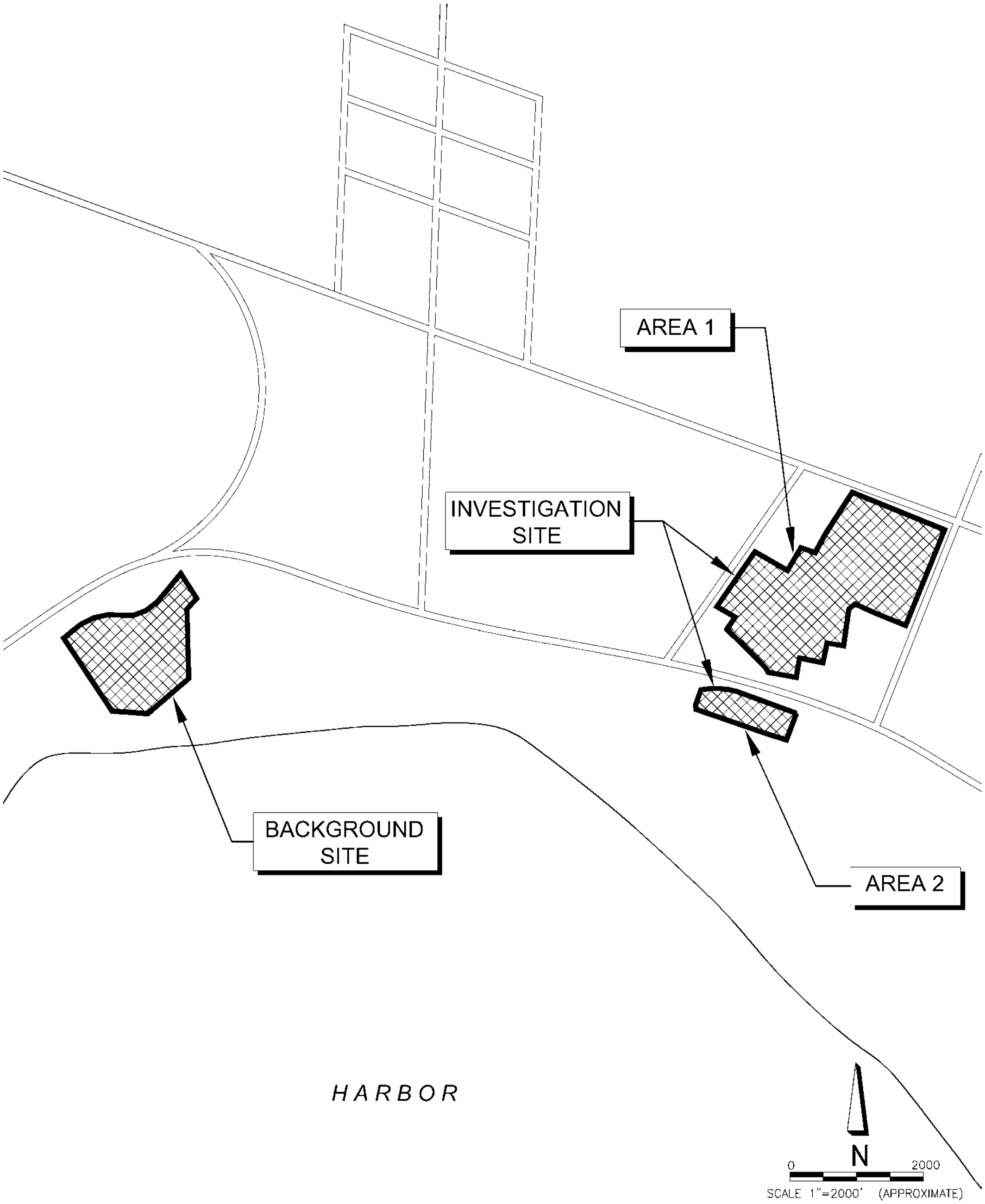


FIGURE 5-1. Location map, investigation and background sites

wells. The elevated concentrations of these two metals could be attributable to the site-related contaminant sources, or may be associated with background sources. Although anthropogenic nonpoint sources of various chemicals may exist at the residential and commercial properties surrounding the site, no chemical releases were identified in the off-site area.

Geology

The investigation and potential background sites are located on an extensive lowland coastal plain formed by interbedded alluvial and marine sedimentary deposits, including sand, gravel, and low permeability beds of clay and silt known collectively as caprock. The caprock formations overlie

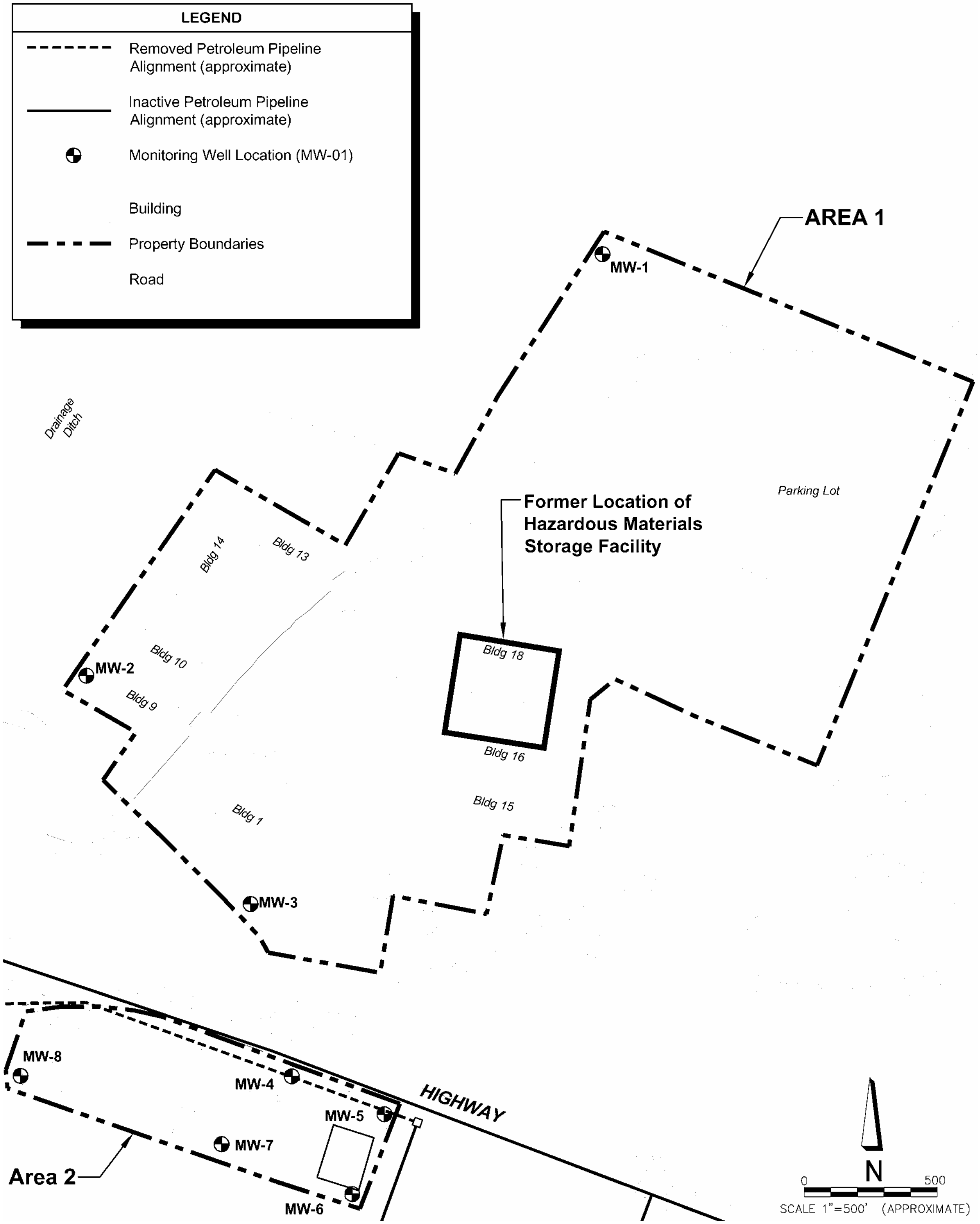


FIGURE 5-2. Site layout and monitoring well locations, investigation site

a thick sequence of basalt flows that extends to great depths throughout the region.

The cross section illustrated in Figure 5-3 is based on observations made during drilling for monitoring well installation at the investigation site and the results of other investigations at both Navy and non-Navy sites in the area. The cross section illustrates the differences between the geology and hydrogeology of Areas 1 and 2. Area 1 is located above a layer of silty clay that directly overlies the basalt bedrock. The silty clay layer dips steeply in the transition zone between Area 1 and Area 2, and is overlain in Area 2 by a wedge of saturated caprock sediments.

Hydrogeology

The entire region is underlain by an extensive basal (i.e., low-level) groundwater aquifer within the fractured basalt bedrock. Lenses and wedges of caprock sediments overlie the basalt bedrock in many areas in the region. The caprock strata are often saturated, and typically are isolated hydraulically from the basal groundwater by low-permeability clay layers that overlie the basalt. The caprock water-bearing zone occurs within interbedded silts, clayey silts, and silty clays, and in local lenses of sand and gravel. As shown in Figure 5-3, the basal aquifer can be either unconfined (as in Area 1) or confined (as in Area 2). These hydrogeologic conditions are very common in the region, and result in two distinct types of groundwater at the investigation site and surrounding region:

- Basal groundwater within the fractured basalt bedrock aquifer (as in Area 1)
- Near-surface caprock groundwater (as in Area 2).

The basal groundwater originates as rainwater falling in higher drainage basins and infiltrating downward to the basalt bedrock. Fresh water of the basal aquifer floats on and displaces salt water,

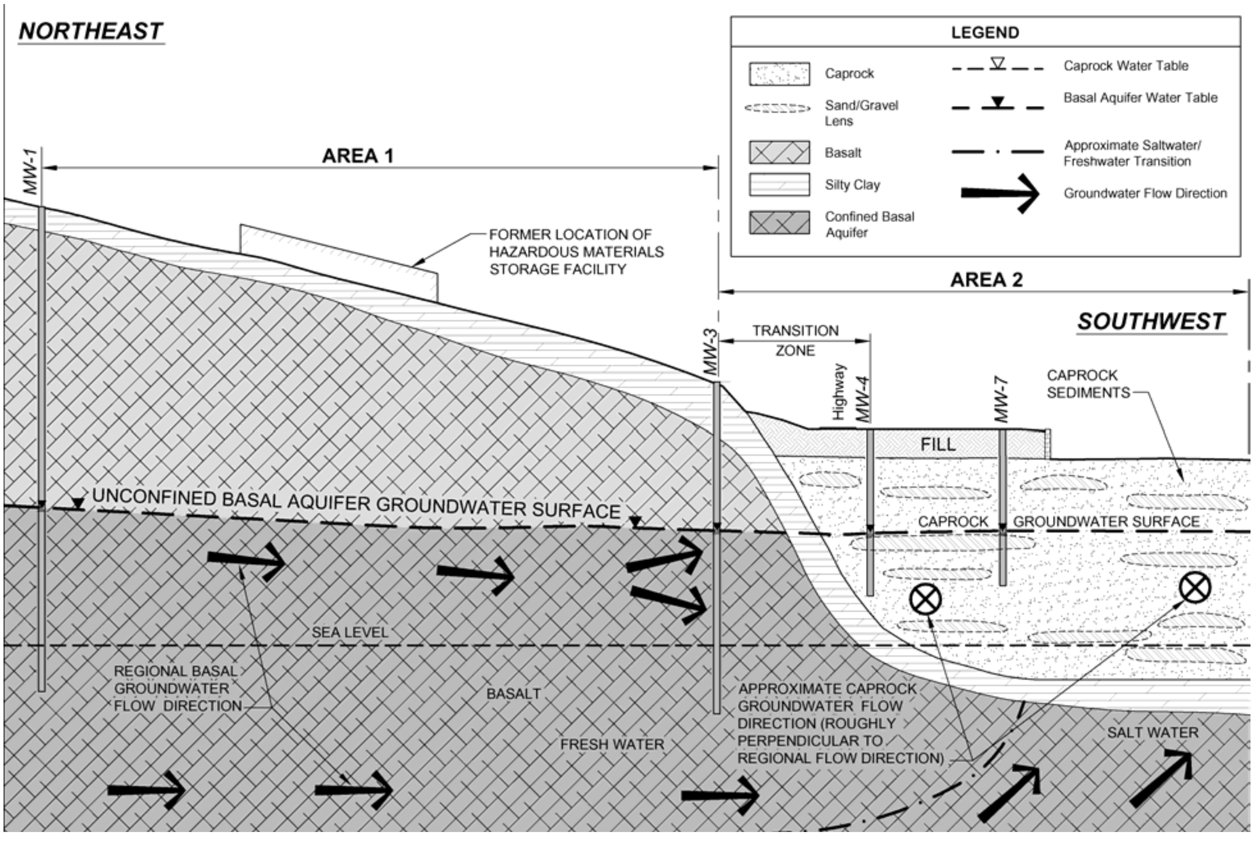


FIGURE 5-3. Geologic cross section, investigation site

which saturates the basalt bedrock at depth. The basal groundwater generally migrates seaward through the highly permeable, fractured basalt, and typically flows beneath relatively impermeable sedimentary confining layers as it approaches the shoreline.

A contour map showing the configuration of the water tables in Areas 1 and 2 at the investigation site is presented in Figure 5-4. The hydrogeological data indicate that the basal groundwater in Area 1 and the caprock groundwater in Area 2 exhibit significant differences:

- ❑ ***Area 1.*** Monitoring wells MW-1, MW-2, and MW-3 penetrate the bedrock basalt, and demonstrate water levels and a gentle hydraulic gradient consistent with regional unconfined basal groundwater conditions.

- ❑ ***Area 2.*** Monitoring wells MW-4 through MW-8 penetrate caprock sediments, and demonstrate water levels and a steeper hydraulic gradient inconsistent with the pattern observed in Area 1.

Groundwater levels measured during drilling and after well installation are summarized below. At Area 1, groundwater was initially encountered at depths of 118.9, 86.8, and 63.0 feet below ground surface (bgs) in wells MW-1, MW-2, and MW-3, respectively. These depths correspond to water table elevations of 13.31, 13.62, and 14.10 feet above mean sea level (amsl). As indicated in Table 5-1, differences between groundwater levels observed during drilling and levels observed after well installation were minimal. These minor changes in groundwater levels suggest the groundwater at Area 1 is part of an unconfined aquifer.

Unlike Area 1, where the groundwater occurs in fractured basalts with very high hydraulic conductivity, groundwater at Area 2 occurs in clay-dominated caprock sediments with low measured conductivity.

At Area 2, groundwater was initially encountered at depths of 29.50, 33.10, 29.10, 24.75, and 19.31 feet bgs in wells MW-4, MW-5, MW-6, MW-7, and MW-8, respectively. As shown in Table 5-2, these depths correspond to elevations of 11.97, 13.23, 13.28, 13.66, and 11.17 feet amsl. After the MW-4 and MW-8 borings were drilled and the water levels were allowed to equilibrate following well installation and development, the water levels rose 1.32 feet and 2.04 feet, respectively. The rise in head observed in MW-4 and MW-8 suggests that the caprock groundwater in Area 2 may be confined in some areas by clay lenses with very low hydraulic conductivity.

Water table elevations recorded in all wells screened in the basal aquifer beneath Area 1 decrease systematically and evenly in a seaward direction. Conversely, water table elevations measured in the Area 2 wells, which are screened in the caprock groundwater, show more variability and do not conform to the smooth seaward trend observed in Area 1. In Area 2, groundwater flows in directions both opposite and perpendicular to that of the basal groundwater in Area 1 (Figure 5-4). The differences in groundwater levels between Area 1 and Area 2 support the conclusion that a low-permeability formation (the silty clay layer) forms a hydraulic barrier between the two areas.

Geotechnical and Geochemical Evaluation

Data representing the geotechnical and geochemical characteristics of the investigation site aquifer matrix materials, and the geochemical characteristics of the groundwater were evaluated to provide additional evidence to determine whether the Area 1 groundwater data are appropriate for comparison to the Area 2 groundwater data.

Investigation Site Aquifer Matrix Materials

Seven samples of the aquifer matrix material were collected from the Area 1 and Area 2 monitoring well borings for geotechnical and geochemical testing. The test results are summarized in Table 5-3.

The geotechnical and geochemical parameters show clear distinctions between the caprock and basal aquifer matrix materials, and therefore indicate that the two water-bearing formations are part of separate groundwater systems. For example, the hydraulic conductivities measured for the Area 2 samples are several orders of magnitude lower than the conductivities measured for the Area 1 samples. The TOC concentrations measured for

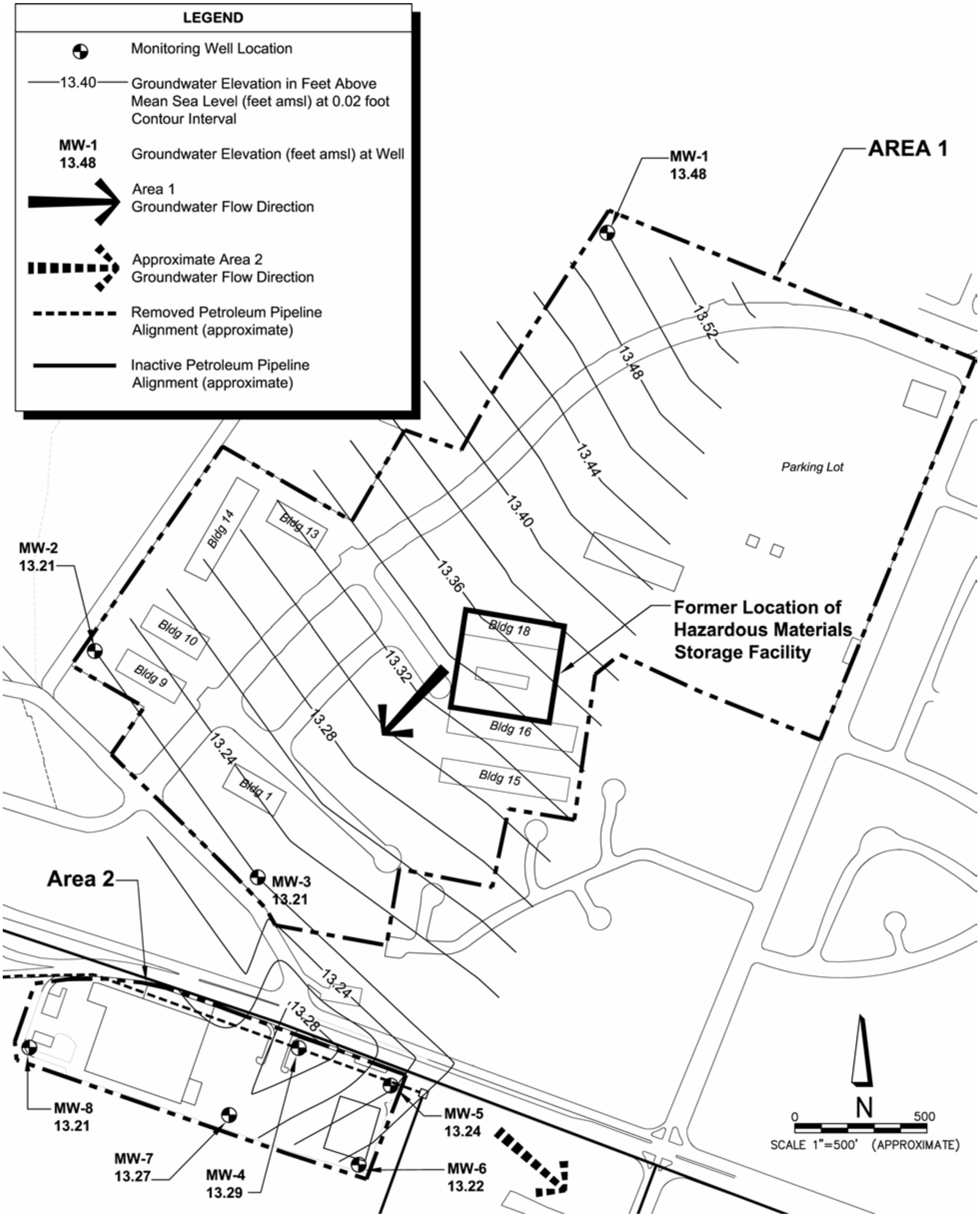


FIGURE 5-4. Groundwater contour map, investigation site

TABLE 5-1. Groundwater elevations, investigation site Area 1

Well ID	Groundwater Level Observed during Drilling		Groundwater Level Measured after Well Installation	
	Depth (feet bgs) (a)	**Elevation (feet amsl)**	**Depth (feet bgs) (a)**	**Elevation (feet amsl)**
MW-1	118.90	13.31	118.73	13.48
MW-2	86.80	13.62	87.21	13.21
MW-3	63.00	14.10	63.89	13.21

(a) Measured from top of inner well casing.

TABLE 5-2. Groundwater elevations, investigation site Area 2

Well ID	Groundwater Level Observed during Drilling		Groundwater Level Measured after Well Installation	
	Depth (feet bgs) (a)	**Elevation (feet amsl)**	**Depth (feet bgs) (a)**	**Elevation (feet amsl)**
MW-4	29.50	11.97	28.18	13.29
MW-5	33.10	13.23	33.09	13.24
MW-6	29.10	13.28	29.16	13.22
MW-7	24.75	13.66	25.14	13.27
MW-8	19.31	11.17	17.27	13.21

(a) Measured from top of inner well casing.

TABLE 5-3. Aquifer matrix geotechnical and geochemical test results, investigation site

Boring ID(a)	Moisture Content by ASTM D2216 (%)	Bulk Density by ASTM D2937 (pcf)	Total Porosity(b)	Specific Gravity by ASTM D854	Saturated Hydraulic Conduc-tivity by ASTM D5084 (cm/sec)	TOC by U.S. EPA Method 415.2 (mg/kg)	CEC by U.S. EPA Method 9081 (meq/100g)	Redox by ASTM D1498 (mV)	Soil pH by ASTM D2976
					Area 1 (Basal Aquifer)				
MW-2GTS28.0	42.7	80.1	0.563	2.94	6.1E–03	<140	42.6	255	7.53
MW-1GTS15.0	31.4	80.6	0.559	2.93	4.8E–02	225	27.7	264	7.42
MW-3GTS11.5	29.1	79.7	0.561	2.91	2.3E–02	217	40.8	272	7.43
					Area 2 (Caprock Groundwater)				
MW-6GTS16.5	36.4	87.0	0.518	2.89	2.7E–09	1,170	68.5	385	8.23
MW-5GTS21.5	40.0	83.6	0.538	2.90	3.1E–08	1,598	84.8	391	8.16
MW-8GTS11.5	34.3	88.1	0.510	2.88	6.1E–08	1,350	62.8	480	8.39
MW-8GTS26.5	39.8	83.1	0.521	2.78	2.3E–08	1,452	55.9	452	8.29

(a) Boring ID indicates well location and geotechnical soil sampling depth in feet.
(b) Porosity is calculated from density and specific gravity.
ID = identification.
meq/100g = milliequivalent of sodium per 100 grams.
mV = millivolt.
pcf = pounds per cubic foot.

the Area 2 samples are significantly higher than those measured for the Area 1 samples. CEC and pH values are also considerably higher in the Area 2 samples.

Investigation Site Groundwater

Groundwater samples collected during four rounds of monitoring at the investigation site were analyzed for the major ions: Na$^+$, Ca^{2+}, K$^+$, Mg^{2+}, SO$_4^{2-}$, Cl$^-$, and HCO$_3^-$. TDS, specific conductivity, salinity, and temperature were also measured. The groundwater geochemistry data corresponding to the first two sampling rounds are presented in Table 5-4.

The investigation site groundwater geochemistry data indicate significantly different chemical compositions between Area 1 and Area 2. Average concentrations of all the major ions are consistently higher in the Area 2 (caprock) groundwater than in the Area 1 (basal) groundwater. The major ion data were plotted on Piper trilinear diagrams to systematically evaluate the differences in ionic compositions among the groundwater samples. Piper trilinear diagrams representing the major ion concentrations detected in the investigation site groundwater are presented in Figures 5-5 and 5-6. As shown in the diagrams, the Area 1 groundwater (MW-1, MW-2 and MW-3) is classified as chloride type, whereas the Area 2 groundwater (MW-4, MW-5, MW-6, MW-7, and MW-8) is classified as bicarbonate type. The TDS, specific conductivity, salinity, and temperature data also show clear differences between the groundwater bodies.

Geochemistry of the COPCs in the Groundwater Environment

Groundwater beneath Area 1 exists exclusively in basaltic formations, whereas the caprock groundwater beneath Area 2 exists within alluvial and marine sedimentary deposits near the coastline. Compared to the basal groundwater, the caprock groundwater has high concentrations of many chemicals, including the major ions and trace

TABLE 5-4. Major ion concentrations in groundwater, investigation site

Sample ID	Ca^{2+} (mg/L)	Mg^{2+} (mg/L)	Na$^+$ (mg/L)	K$^+$ (mg/L)	Cl$^-$ (mg/L)	HCO$_3^-$ (mg/L)	SO$_4^{2-}$ (mg/L)	TDS (mg/L)	Specific Conductivity (mS/cm)	Salinity (%)	Temperature (°C)
Round One											
Area 1 (Basal Groundwater)											
MW-1	11.5	10.9	46.7	3.01	53	74	10	220	0.360	0.01	25.3
MW-2	11.6	11.2	92.4	3.45	100	91	18	310	0.566	0.02	24.4
MW-3	13.0	10.2	123	4.18	134	133	34	380	0.687	0.02	25.4
Area 2 (Caprock Groundwater)											
MW-4	26.7	25.4	162	9.86	94	375	47	600	0.940	0.04	27.9
MW-5	33.3	25.4	133	8.89	88	438	3	550	1.00	0.04	26.2
MW-6	30.9	25.8	159	7.26	150	468	58	600	1.02	0.04	27.1
MW-7	17.9	23.0	186	9.53	150	499	61	640	1.07	0.04	27.0
MW-8	113	20.0	184	9.65	200	500	56	1,100	1.75	0.07	25.7
Round Two											
Area 1 (Basal Groundwater)											
MW-1	10.4	9.46	34.7	2.35	52	60	12	210	0.376	0.01	23.5
MW-2	11.3	9.71	69.1	2.58	94	74	24	320	0.609	0.02	24.8
MW-3	7.29	7.57	81.8	3.01	134	110	30	350	0.682	0.02	24.8
Area 2 (Caprock Groundwater)											
MW-4	23.5	23.7	149	6.12	86	350	20	550	0.940	0.04	27.3
MW-5	34.6	27.8	135	5.33	76	390	1	540	1.07	0.04	27.7
MW-6	82.6	72.1	211	7.82	480	438	81	1,200	2.17	0.1	26.9
MW-7	24.4	18.5	137	5.87	130	470	61	610	1.02	0.04	27.9
MW-8	114	23	177	9.41	170	486	67	800	1.33	0.06	26.4

mS/cm = millisiemens per centimeter.

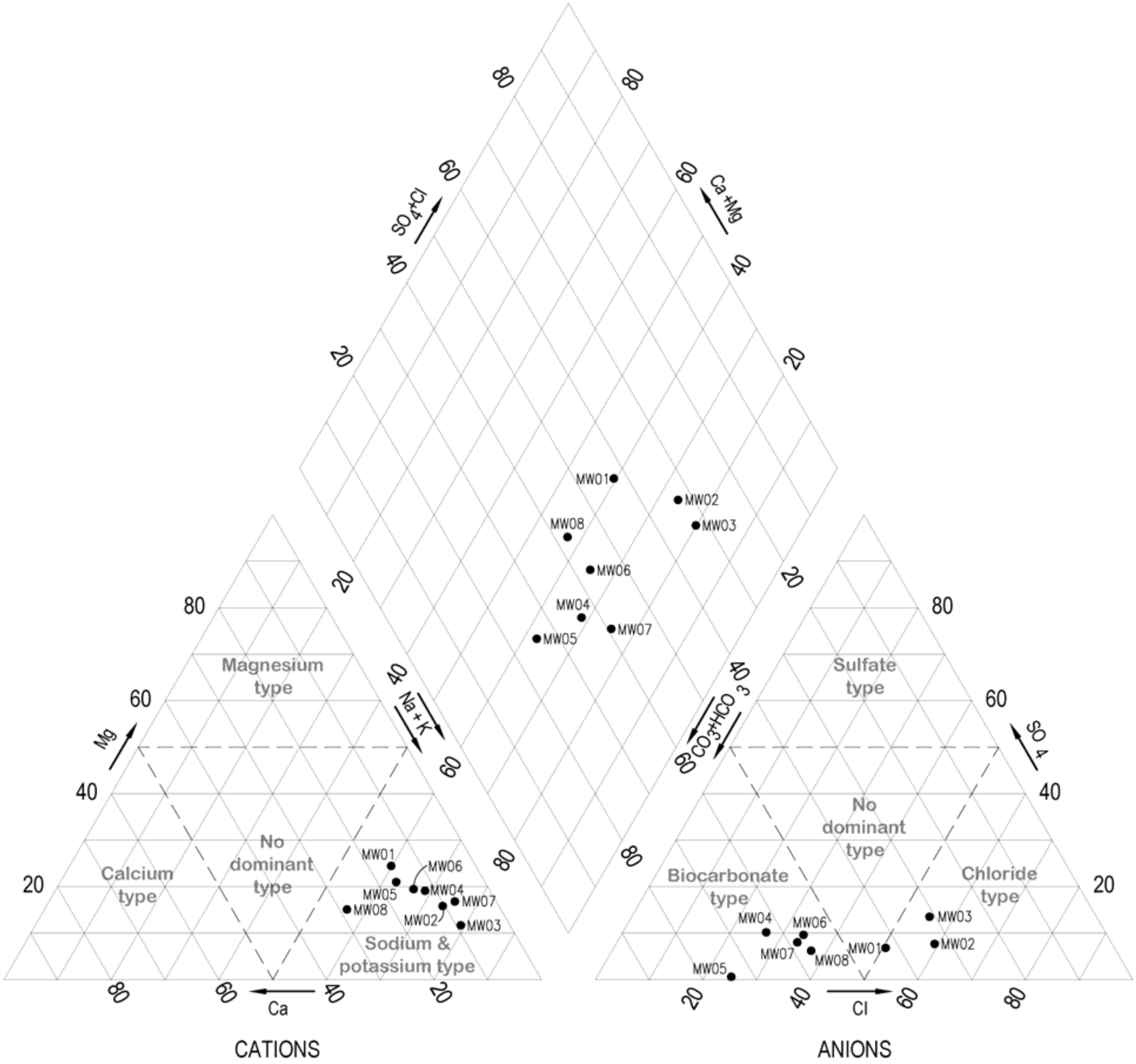


FIGURE 5-5. Piper trilinear diagram Round 1, investigation site

metals. The caprock sediments typically have trace metal concentrations orders of magnitude higher than the basalts. In addition, the hydraulic conductivity of the caprock formation is much lower than the conductivity of the fractured basalts of the basal aquifer, resulting in much longer groundwater residence time within the caprock. This prolonged contact between the groundwater and the aquifer matrix, combined with the higher trace metal concentrations, results in higher average concentrations of all dissolved species, including the major ions and trace metals. Concentrations of the COPC metals, arsenic and lead, detected in groundwater samples collected from three Area 1 wells and five Area 2 wells over four sampling events, are presented in Table 5-5. The data indicate that concentrations of the COPC metals in the Area 1 basal groundwater are well below concentrations in the Area 2 caprock groundwater.

Arsenic. The phase equilibrium (i.e., partitioning between the solid and aqueous phases) of arsenic in the groundwater environment is strongly controlled by Eh and pH. Under the Eh and pH ranges typical in soils, sediments, and groundwater, the As^{5+} and As^{3+} valence states predominate. Under

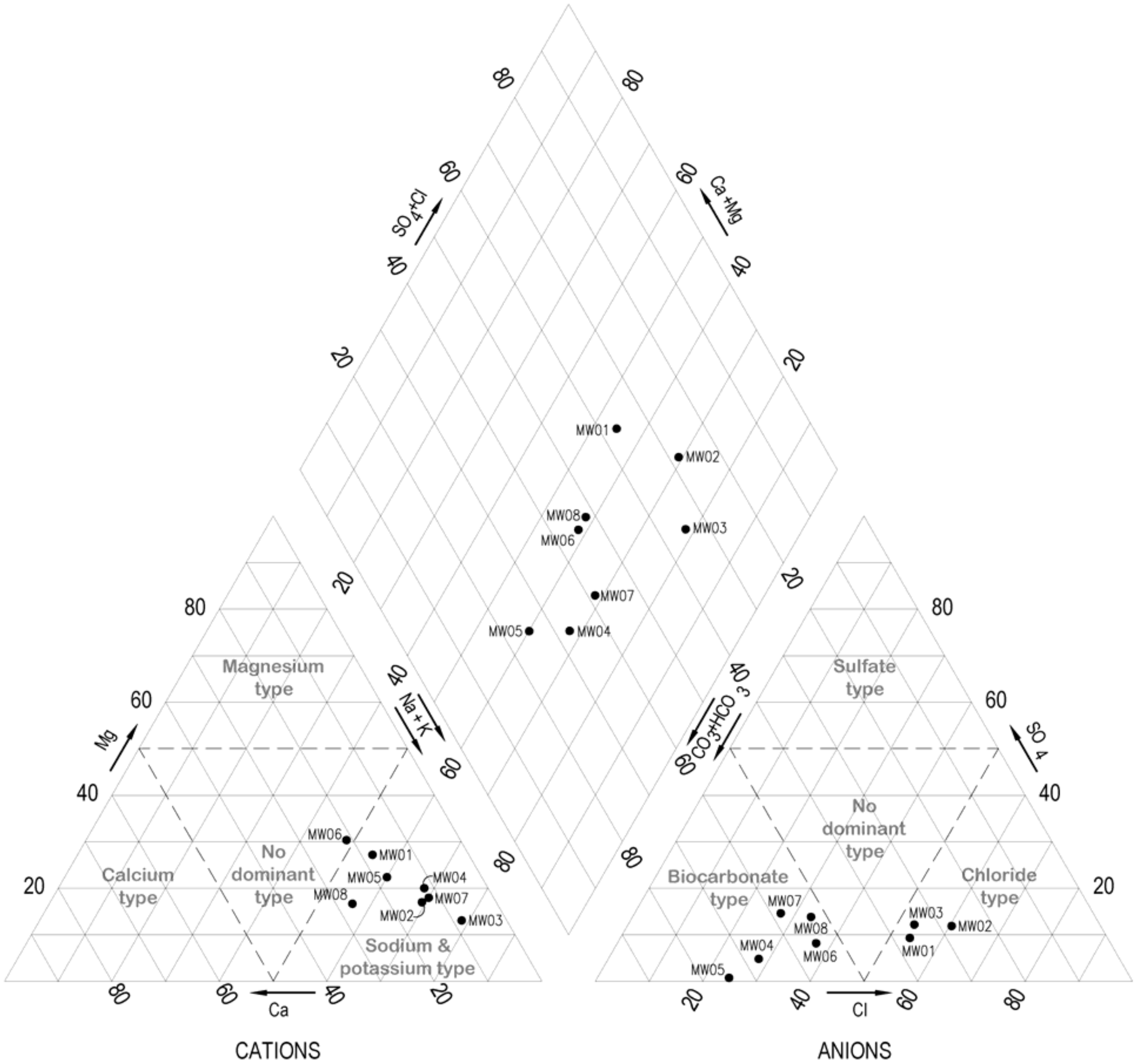


FIGURE 5-6. Piper trilinear diagram Round 2, investigation site

oxidizing conditions (high Eh), arsenic remains in the As^{5+} valence state and has limited solubility (Masscheleyn et al., 1991), whereas under reducing conditions (low Eh), the As^{3+} valence state predominates. As^{3+} is about 40 times more soluble than As^{5+} (U.S. EPA, 1984). Cercelius et al. (1986) also studied the behavior of arsenic, and found that, under typical environmental conditions, As^{5+} (as the $H_2AsO_4^-$ complex) predominates, whereas in low Eh environments, the more soluble As^{3+} (as the complex $H_3AsO_3^0$) is dominant. Although Eh is the most significant parameter with respect to arsenic solubility and mobility, pH also has some effect on arsenic mobility: low pH tends to favor conversion of As^{5+} to As^{3+} under higher Eh conditions (Masscheleyn et al., 1991).

Arsenic concentrations detected in the investigation site groundwater over four rounds of sampling are shown on Figure 5-7. The relatively high arsenic concentrations detected in the Area 2 groundwater are most likely attributable to localized reducing conditions in the saturated caprock sediments. As shown in Table 5-3, samples of the caprock formation had measured hydraulic conductivity values as low as 2.7×10^{-9} cm/sec, indicating that highly impermeable zones exist within the formation. These impermeable zones restrict the movement of groundwater and the oxygen and oxidized species (e.g., nitrate, sulfate, ferric iron, carbon dioxide) dissolved in the groundwater, and therefore form localized reducing environments. Under these conditions, naturally occurring arsenic associated with the sediments will tend to partition from the solid phase into the aqueous phase.

Lead. Lead is usually a relatively immobile element in the groundwater environment due to its low solubility under typical Eh and pH conditions, and its tendency to sorb to aquifer matrix materials. Naturally occurring lead in most soils and sediments is associated primarily with clay minerals, manganese oxides, iron and aluminum hydroxides, and organic matter. Particles of these materials provide a large surface area to sorb lead; therefore, lead mobility and leachability in media with high percentages of these components tend to be low (Xinde Cao et al., 2003). Soils with high CEC also readily adsorb high concentrations of lead. In addition, lead is much more soluble under low pH conditions than at neutral or high pH (Xinde Cao et al., 2003). High pH favors precipitation of lead as hydroxide, and promotes the formation of very stable lead-organic complexes.

The caprock sediments contain significant quantities of clay, and have relatively high CEC and TOC concentrations. These characteristics, combined with the elevated pH values measured in the Area 2 soil samples (Table 5-3), indicate that naturally occurring lead in the caprock sediments will tend to remain in the solid phase. Lead concentrations detected in the investigation site groundwater over four rounds of sampling are shown on Figure 5-7. The relatively high lead concentrations detected in groundwater samples from MW-4 and MW-5 in Area 2 could be evidence of a release of leaded gasoline from the pipelines that formerly conveyed petroleum products along the northern boundary of Area 2. The highest dissolved lead concentrations were detected in groundwater samples from MW-5, the only well in which free-phase petroleum product has been observed.

TABLE 5-5. Arsenic and lead concentrations (µg/L) in groundwater, investigation site

		Round			
Well ID	**COPC**	**1**	**2**	**3**	**4**
Area 1					
MW-1	As	0.76 U	0.76 UJ	0.9 UJ	0.73 UJ
	Pb	0.1 U	0.1 U	0.1 U	0.1 U
MW-2	As	0.9 U	0.18 J	0.175 J	0.17 J
	Pb	0.1 U	0.1 U	0.1 U	0.1 U
MW-3	As	1.3 UJ	1.3 UJ	0.95 UJ	1.2 UJ
	Pb	0.1 U	0.1 U	0.1 U	0.1 U
Area 2					
MW-4	As	3.5	3.2	3.1	3.15
	Pb	2.2	2	2.1	1.8
MW-4D	As	3.4	3.24	2.94	2.90
	Pb	2.4	2.3	2.2	1.8
MW-5	As	5.4	5.6	5.3	3.2
	Pb	22	18	32	33.1
MW-6	As	1.1	1.5	4.8	1.5
	Pb	0.08 U	0.09 U	0.09 U	0.15 UJ
MW-7	As	5.4	4.7	4.6	4.7
	Pb	0.22 UJ	0.22 UJ	0.45	0.22 UJ
MW-8	As	5.1	4.1	1.9	5.4
	Pb	0.67	0.48	0.59	0.45

As = arsenic.
Pb = lead.
D = duplicate sample.
J = estimated concentration.
U = nondetect.
UJ = nondetect estimated.

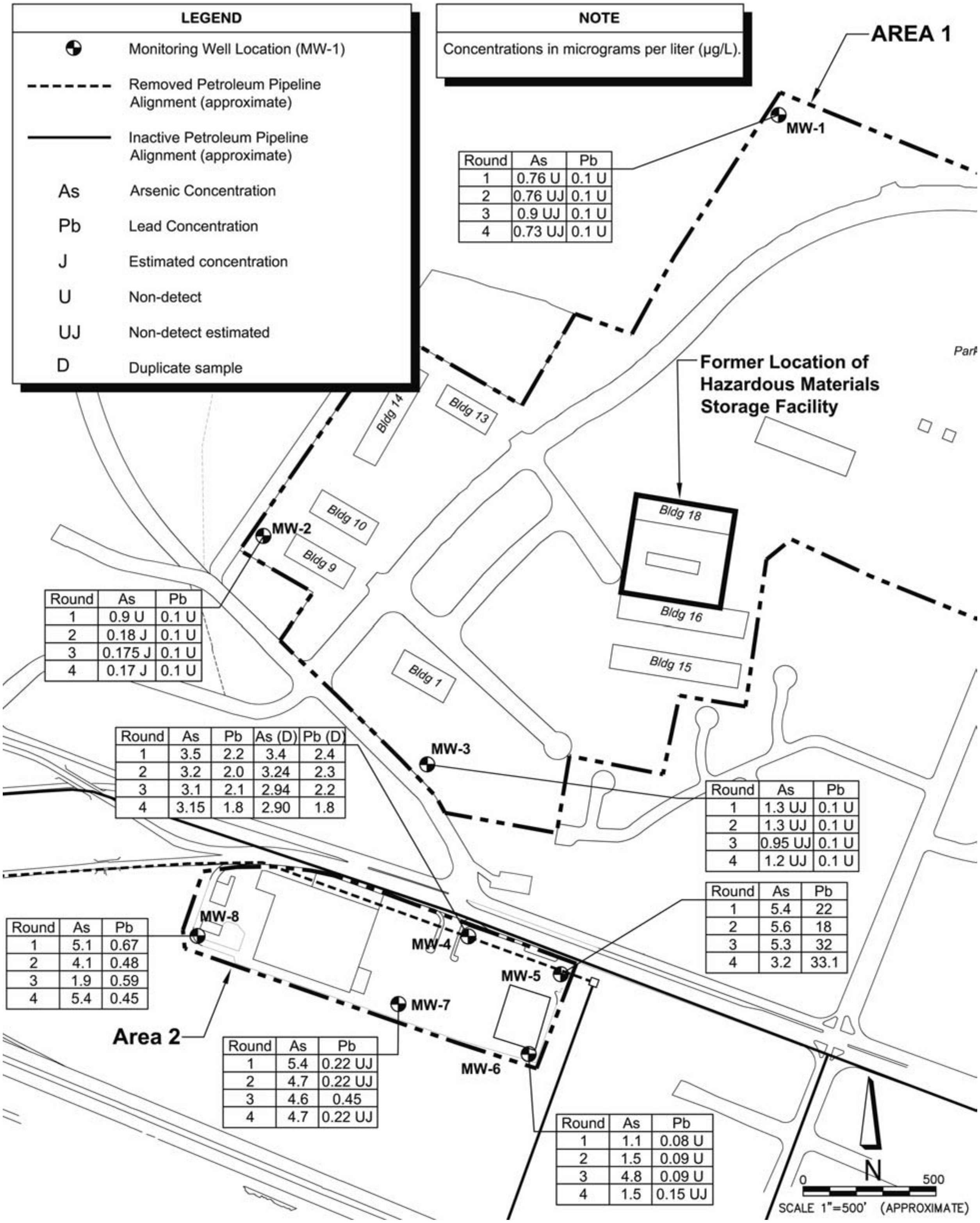
Investigation Site Conclusions

Results of the geologic, hydrogeologic, and geochemical evaluation clearly indicate that the Area 1 and Area 2 groundwater bodies represent two different groundwater regimes. Therefore, the Area 1 groundwater data should not be used to define background conditions with respect to the Area 2 groundwater. Additional data are needed to develop a background dataset appropriate for comparison to the Area 2 groundwater dataset.

In addition, the COPC concentration data, geochemical characteristics of the caprock and COPCs, and locations of the suspected contaminant sources suggest that, although the arsenic detected in the Area 2 groundwater is most likely associated with natural sources, some of the lead detected in the Area 2 groundwater could be associated with a release from the petroleum product pipelines.

5.1.2 Background Site Evaluation

A nearby site with characteristics likely to be similar to those of the investigation site was selected as



LEGEND	
⊕	Monitoring Well Location (MW-1)
- - - - - -	Removed Petroleum Pipeline Alignment (approximate)
———	Inactive Petroleum Pipeline Alignment (approximate)
As	Arsenic Concentration
Pb	Lead Concentration
J	Estimated concentration
U	Non-detect
UJ	Non-detect estimated
D	Duplicate sample

NOTE
Concentrations in micrograms per liter (µg/L).

MW-1

Round	As	Pb
1	0.76 U	0.1 U
2	0.76 UJ	0.1 U
3	0.9 UJ	0.1 U
4	0.73 UJ	0.1 U

MW-2

Round	As	Pb
1	0.9 U	0.1 U
2	0.18 J	0.1 U
3	0.175 J	0.1 U
4	0.17 J	0.1 U

MW-4

Round	As	Pb	As (D)	Pb (D)
1	3.5	2.2	3.4	2.4
2	3.2	2.0	3.24	2.3
3	3.1	2.1	2.94	2.2
4	3.15	1.8	2.90	1.8

MW-3

Round	As	Pb
1	1.3 UJ	0.1 U
2	1.3 UJ	0.1 U
3	0.95 UJ	0.1 U
4	1.2 UJ	0.1 U

MW-5

Round	As	Pb
1	5.4	22
2	5.6	18
3	5.3	32
4	3.2	33.1

MW-8

Round	As	Pb
1	5.1	0.67
2	4.1	0.48
3	1.9	0.59
4	5.4	0.45

MW-7

Round	As	Pb
1	5.4	0.22 UJ
2	4.7	0.22 UJ
3	4.6	0.45
4	4.7	0.22 UJ

MW-6

Round	As	Pb
1	1.1	0.08 U
2	1.5	0.09 U
3	4.8	0.09 U
4	1.5	0.15 UJ

FIGURE 5-7. Arsenic and lead concentrations in groundwater, investigation site

a potential background site. The site location is shown on Figure 5-1; the site layout and monitoring well locations are shown on Figure 5-8. Potential contaminant sources, geology, hydrogeology, and geochemistry were evaluated to determine whether groundwater data from the potential background site are appropriate for comparison to data from Area 2 of the investigation site.

Contaminant Sources

Although diesel fuel was formerly stored in Tanks S-26 and S-27, there is no evidence of a chemical release that could impact groundwater at the site. Anthropogenic nonpoint sources of various chemicals may exist at the residential and commercial properties surrounding the site; however, no chemical releases were identified in the off-site area.

Geology

The geology of the background site (shown on the geological cross section illustrated in Figure 5-9) is very similar to that of the investigation site; however, unlike the investigation site, the caprock formation is continuous across the entire background site. The caprock decreases in thickness as the ground surface slopes down toward the harbor shoreline. Impermeable clay and silty clay layers at the base of the caprock overlie the basalt bedrock. Driller's logs from artesian wells in the agricultural area downgradient of the site indicate that the confined basal aquifer was encountered at a depth of approximately 45–58 feet bgs (40–50 feet below msl). This is consistent with the inferred depth to basal groundwater in Area 2 at the investigation site.

Hydrogeology

The hydrogeology of the background site is also similar to that of the investigation site. However, the unconfined near-surface caprock water-bearing zone is continuous throughout the site area. As in Area 2 of the investigation site, the caprock groundwater occurs within interbedded silts, clayey silts, and silty clays, and in local lenses of sand and gravel, whereas the basal aquifer exists within the fractured basalts that underlie the sediments. Also as in Area 2 of the investigation site, impermeable clay and silty clay layers at the base of the caprock sequence form confining layers over the basal aquifer. The potentiometric surface of the confined basal aquifer dips gently toward the shoreline. As shown in Figure 5-9, the elevation of the unconfined caprock water table is less than that of the basal aquifer potentiometric surface. The elevation of the potentiometric surface in the background site area is approximately 16 feet amsl, whereas the elevation of the caprock water table near the center of the site is approximately 12 feet amsl. The difference in hydraulic head between the caprock and basal groundwater bodies is evidence of an impermeable hydraulic barrier, and reflects a difference in elevation between the two recharge areas. The confined basal aquifer is recharged by heavy rainfall in the higher elevations of the watershed, whereas the caprock water-bearing zone is recharged primarily by local sources including rainfall and irrigation water.

Table 5-6 lists the caprock groundwater levels observed during drilling and the water levels measured after well installation and development. The water table contour map presented in Figure 5-10 depicts the caprock water table surface at the background site. The caprock water table is 80 to 90 feet bgs near the northern boundary of the site, and becomes closer to the surface as the topography slopes down toward the shoreline. Caprock water table elevations range from approximately 14 feet amsl in the northern portion of the site to about 8 feet amsl near the southern boundary of the site. Although the water level data collected in Area 2 of the investigation site are limited to a relatively small area, the hydraulic gradient in the saturated background site caprock appears to be similar to that observed in the Area 2 caprock.

Geochemical and Geotechnical Evaluation

Data representing the geotechnical and geochemical characteristics of the background site aquifer matrix materials, as well as data on the geochemical characteristics of the groundwater, were evaluated to provide additional evidence to determine whether the background site groundwater data are appropriate for comparison to the Area 2 groundwater data.

Background Site Aquifer Matrix Materials

Six samples of the aquifer matrix material were collected from the monitoring well borings that

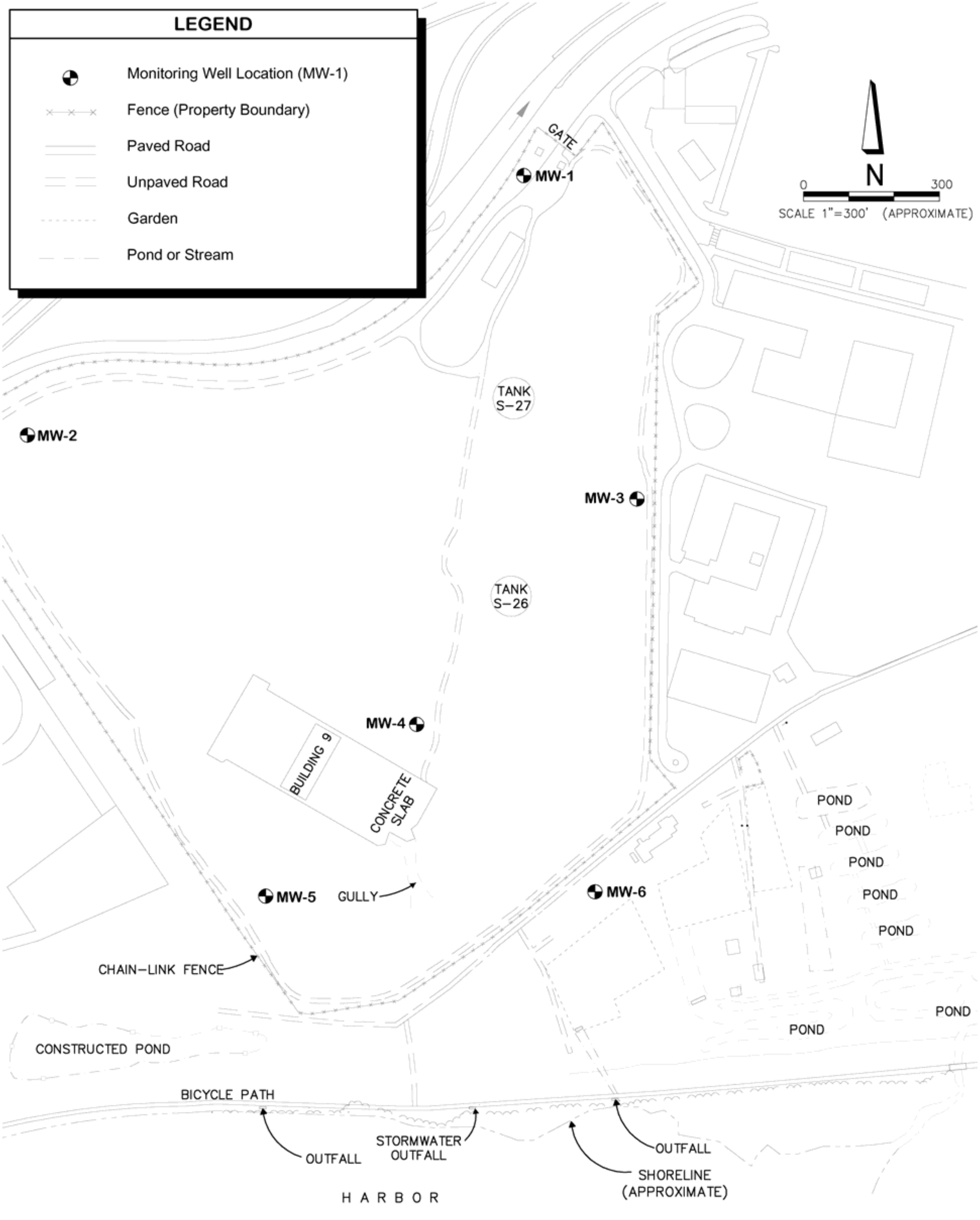


FIGURE 5-8. Site layout and monitoring well locations, background site

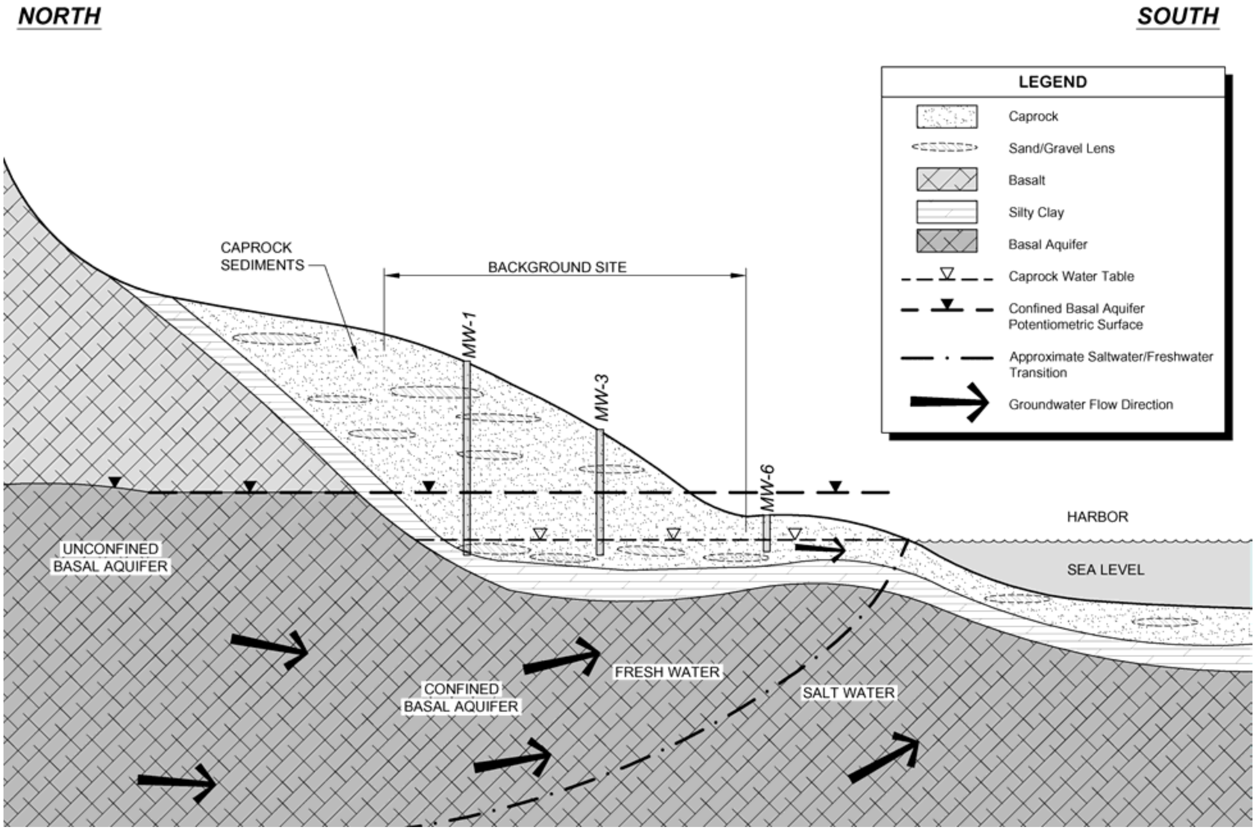


FIGURE 5-9. Geologic cross section, background site

TABLE 5-6. Caprock groundwater elevations, background site

	Groundwater Level Observed during Drilling		Groundwater Level Measured after Well Installation	
Well ID	Depth (feet bgs) [(a)]	Elevation (feet amsl)	Depth (feet bgs) [(a)]	Elevation (feet amsl)
MW-1	113.00	13.74	112.26	13.00
MW-2	98.54	14.42	97.21	13.09
MW-3	60.36	13.05	59.28	11.97
MW-4	84.69	11.80	83.71	10.82
MW-5	63.82	10.23	62.59	9.00
MW-6	29.23	8.81	27.99	7.57

(a) Measured from top of inner well casing.

penetrate the caprock formation at the background site for geotechnical and geochemical testing. The test results are summarized in Table 5-7.

The background site geotechnical and geochemical parameters (Table 5-7) are very similar to those of Area 2 of the investigation site (Table 5-3), and indicate that the background site caprock and Area 2 caprock are most likely part of the same geologic formation.

Background Site Groundwater

Groundwater samples collected at the background site were analyzed for the major ions: Na^+, Ca^{2+}, K^+, Mg^{2+}, SO_4^{2-}, Cl^-, and HCO_3^-. TDS, specific

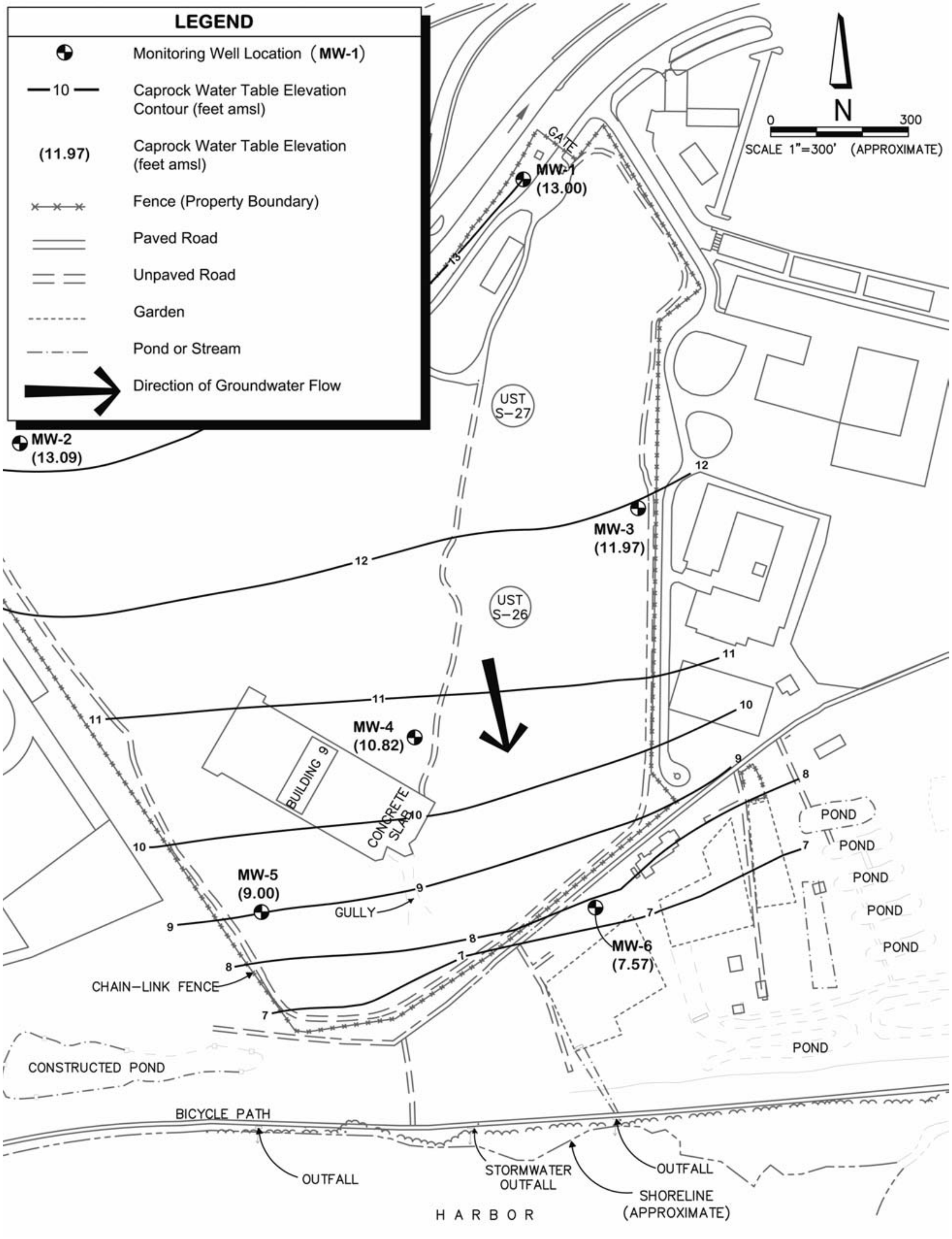


FIGURE 5-10. Groundwater contour map, background site

TABLE 5-7. Aquifer matrix geotechnical and geochemical test results, background site

Boring ID[a]	Moisture Content by ASTM D2216 (%)	Bulk Density by ASTM D2937 (pcf)	Total Porosity[b]	Specific Gravity by ASTM D854	Saturated Hydraulic Conductivity by ASTM D5084 (cm/sec)	TOC by U.S. EPA Method 415.2 (mg/kg)	CEC by U.S. EPA Method 9081 (meq/100g)	Redox by ASTM D1498 (mV)	Soil pH by ASTM D2976
MW-1GTS58.0	38.1	86.2	0.524	2.90	2.6E–09	1,180	82.3	393	8.26
MW-2GTS55.0	39.6	83.6	0.530	2.85	3.1E–08	1,570	59.1	386	8.29
MW-3GTS48.5	38.9	83.6	0.528	2.84	2.8E–08	1,510	82.3	402	8.35
MW-4GTS28.0	34.6	88.1	0.510	2.88	4.1E–08	1,350	68.8	386	8.35
MW-5GTS18.0	35.4	83.3	0.523	2.80	3.2E–08	1,410	54.3	450	8.28
MW-6GTS17.5	37.3	83.4	0.519	2.78	2.5E–08	1,350	55.6	450	8.32

(a) Boring ID indicates well location and geotechnical soil sampling depth in feet.
(b) Porosity is calculated from density and specific gravity.

conductivity, salinity, and temperature also were measured. The groundwater geochemistry data collected during the first round of background site groundwater sampling are presented in Table 5-8.

The Piper trilinear diagram presented in Figure 5-11 is based on the major ion concentration data associated with Area 2 of the investigation site and the background site, presented in Tables 5-4 and 5-8, respectively. The diagram illustrates the close similarities between major ion concentrations detected in the two caprock groundwater bodies. TDS, specific conductivity, salinity, and temperature values were also very similar. The geochemistry data therefore support the conclusion that the background site caprock groundwater and Area 2 caprock groundwater are most likely part of the same groundwater system.

Arsenic and Lead Concentrations in Background Site Groundwater

The arsenic and lead concentrations detected in groundwater samples collected from six wells at the background site over eight sampling events are presented in Table 5-9.

Arsenic. Arsenic concentrations detected in the background site groundwater are presented on the site map shown in Figure 5-12. A spatial/temporal concentration plot for arsenic is presented in Figure 5-13.

Lead. Lead concentrations detected in the background site groundwater are presented on the site map shown in Figure 5-12. A spatial/temporal concentration plot for lead is presented in Figure 5-14.

TABLE 5-8. Major ion concentrations in groundwater, background site

Sample ID	Ca^{2+} (mg/L)	Mg^{2+} (mg/L)	Na^{+} (mg/L)	K^{+} (mg/L)	Cl^{-} (mg/L)	HCO_3^{-} (mg/L)	SO_4^{2-} (mg/L)	TDS (mg/L)	Specific Conductivity (mS/cm)	Salinity (%)	Temperature (°C)
Round One											
Background Site (Basal Groundwater)											
MW-1	22	23.1	125	8.23	86	485.6	34	650	0.89	0.05	26.3
MW-2	25.2	24.2	159	9.64	94	423.5	47	760	1.21	0.045	26.4
MW-3	25.9	24.3	127	7.44	88	445.6	45	780	1.26	0.05	26.6
MW-4	32	25.6	202	8.19	89	458.6	58	620	0.915	0.048	26.8
MW-5	21	25.1	164	7.89	95	456	61	580	1.02	0.043	26.7
MW-6	27	21.3	185	9.66	85	478	56	800	1.33	0.06	26.4

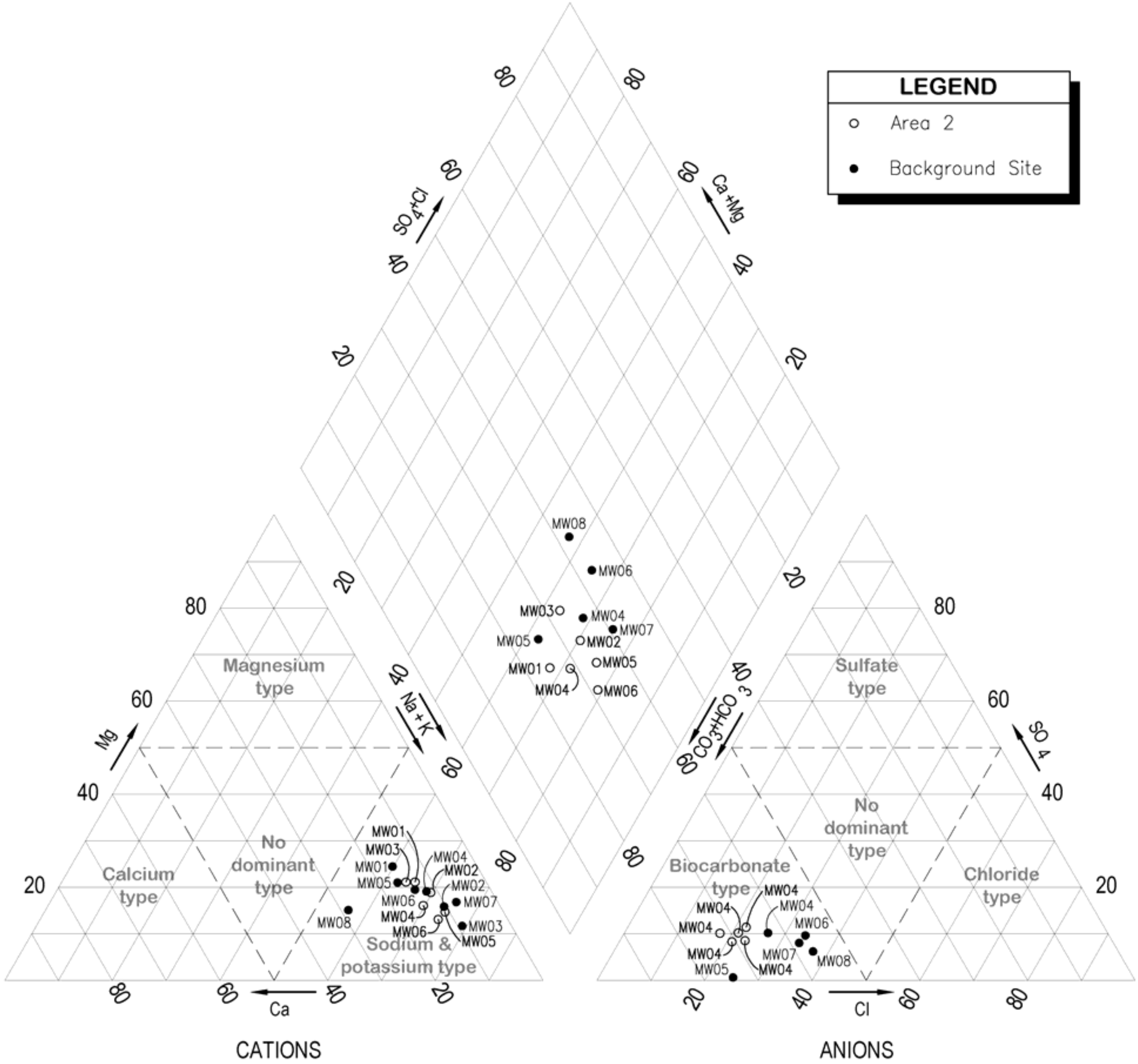


FIGURE 5-11. Piper trilinear diagram, background site and investigation site Area 2

Background Site Conclusions

Results of the geological, hydrogeological, and geochemical evaluation indicate that the caprock groundwater at the background site occurs within the same formation as groundwater at Area 2 of the investigation site, and that the hydrogeologic and geochemical characteristics of caprock groundwater at the two sites are very similar. In addition, there is no evidence of a chemical release at the background site. Arsenic and lead concentrations in the background groundwater are relatively constant within each individual well, and systematic variations of arsenic and lead are strongly correlated to dry and rainy seasons. These observations suggest that arsenic and lead are in steady-state equilibrium in the groundwater system. Under equilibrium conditions, naturally occurring arsenic and lead are mobilized from the sediments into the groundwater to maintain chemical equilibrium under the geochemical conditions existing in groundwater system. The evaluation presented above, combined with inspection of the arsenic and lead data, indicates that all arsenic and lead concentrations detected in groundwater at the background

TABLE 5-9. Arsenic and lead concentrations (µg/L) in groundwater, background site

		Round							
Well	**COPC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
MW-1	As	5.5	5.6	4.8	5.4	4.5	5.5	4.8	5.3
	Pb	0.56	0.9	0.7	1.1	0.9	0.9	0.8	0.9
MW-2	As	4.5	4.8	3.2	4.8	3.5	4.7	4.2	4.6
	Pb	1.2	1.4	1.1	1.2	0.98	1.2	1.1	1.15
MW-3	As	2.7	3.2	2.3	3.2	2.7	3.2	2.5	2.8
	Pb	0.18 UJ	0.23	0.14 UJ	0.29	0.15 J	0.21	0.15 UJ	0.24
MW-4	As	2.1	2.85	3.2	3.1	2.1	2.9	3.2	3.7
	Pb	0.09 U	0.12 UJ	0.09 U	0.14 UJ	0.13	0.15 UJ	0.08 U	0.12
MW-5	As	4.2	5.3	4.8	5.4	4.3	5.6	4.3	5.4
	Pb	0.78	0.9	0.7	0.86	0.75	0.9	0.65	0.9
MW-6	As	1.1	2.2	1.1	2.1	1.2	1.5	1.2	1.5
	Pb	0.75	1.05	0.65	0.8	0.65	0.78	0.67	0.71

J = estimated concentration.
U = nondetect.
UJ = nondetect estimated.

site represent natural background conditions. Therefore, the maximum detected arsenic concentration (5.6 µg/L), and the maximum detected lead concentration (1.4 µg/L), represent technically defensible estimates of the upper bounds of the groundwater background concentration ranges.

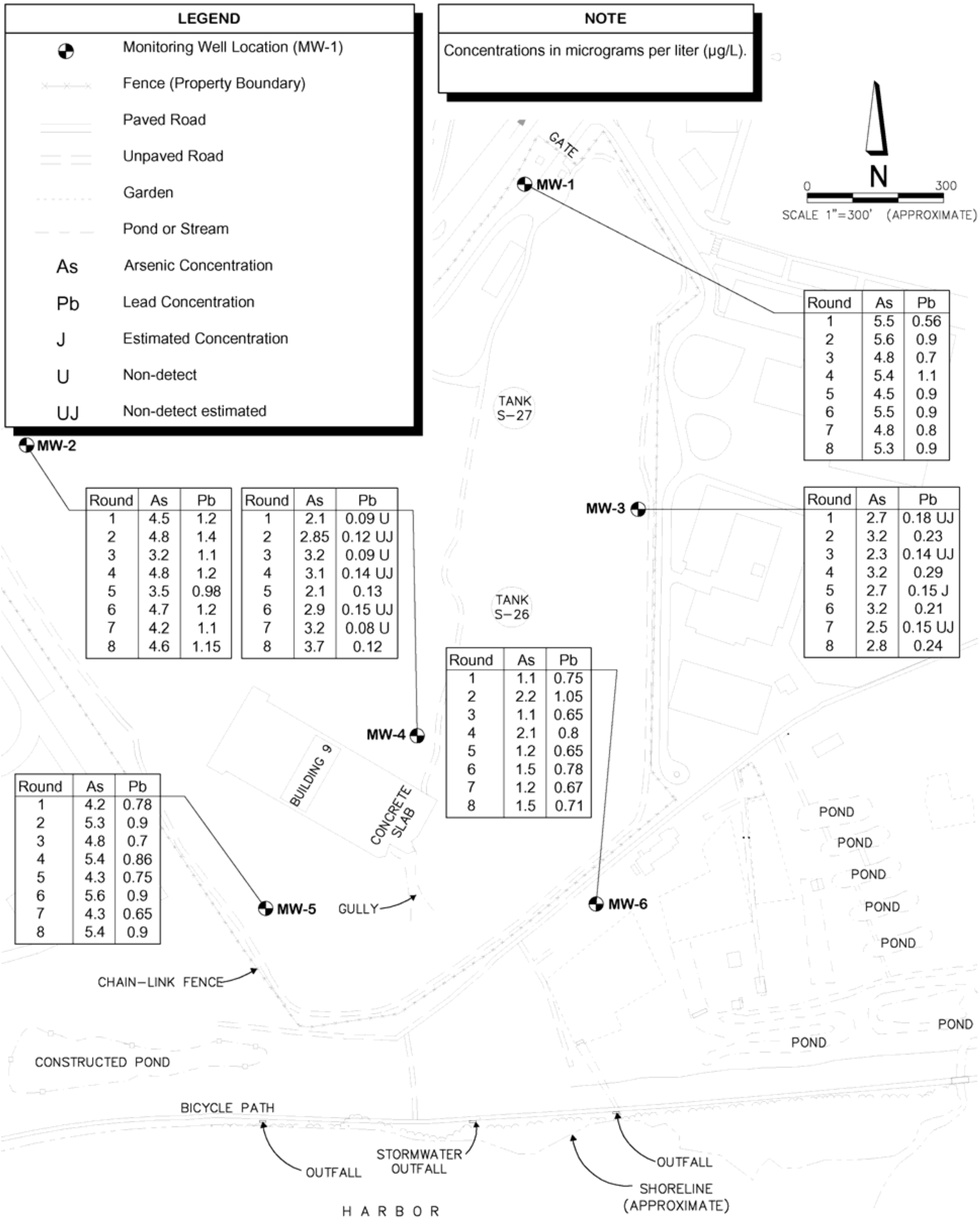
5.2 Arsenic and Lead Concentrations, Area 2 vs. Background Site

COPC concentrations detected in groundwater samples collected from monitoring wells in Area 2 of the investigation site were compared to the concentrations detected at the background site. The Area 2 (caprock) groundwater data were not compared to the Area 1 (basal) groundwater data because the geologic, hydrogeologic, and geochemical evaluation demonstrated that the two groundwater bodies are part of separate groundwater systems with very different geochemical characteristics.

As noted in Section 5.1.1, the COPC concentration data, geochemical characteristics of the caprock and COPC metals, and locations of the suspected contaminant sources suggest that the arsenic detected in Area 2 groundwater is most likely associated with natural sources, whereas some of the lead may be associated with a chemical release.

The maximum arsenic concentration detected in the Area 2 groundwater, 5.6 µg/L, is identical to the maximum concentration detected at the background site. Therefore, in this case, the analysis indicates that arsenic is not a COPC, and that 5.6 µg/L represents a technically defensible estimate of the upper bound of the groundwater background concentration range. However, as noted in Section 4.2.2, in some cases, simple numerical comparison of maximum concentrations can lead to erroneous conclusions. For example, if the maximum arsenic concentrations detected in the Area 2 groundwater were somewhat higher than the maximum concentration detected at the background site, statistical comparative testing would be necessary to confirm that the exceedance is statistically significant before arsenic is declared a COPC. Although in this case the Comparative Method is not necessary to confirm that arsenic is not a COPC, Section 5.3 illustrates the procedures for statistical comparative testing of the arsenic data.

Lead concentrations above the estimated upper bound of the background range, 1.4 µg/L, were detected in groundwater samples collected from two Area 2 monitoring wells, MW-4 and MW-5. The maximum lead concentrations, 18 to 33 µg/L, were detected in the MW-5 samples, and are above the estimated upper bound of the background range. However, statistical comparative analysis is required to determine whether the background exceedances in Area 2 are widespread or only limited to the MW-4 and MW-5 areas.



LEGEND	
◓	Monitoring Well Location (MW-1)
—×—×—	Fence (Property Boundary)
——	Paved Road
— —	Unpaved Road
· · · ·	Garden
— — —	Pond or Stream
As	Arsenic Concentration
Pb	Lead Concentration
J	Estimated Concentration
U	Non-detect
UJ	Non-detect estimated

NOTE

Concentrations in micrograms per liter (µg/L).

Round	As	Pb
1	5.5	0.56
2	5.6	0.9
3	4.8	0.7
4	5.4	1.1
5	4.5	0.9
6	5.5	0.9
7	4.8	0.8
8	5.3	0.9

Round	As	Pb
1	4.5	1.2
2	4.8	1.4
3	3.2	1.1
4	4.8	1.2
5	3.5	0.98
6	4.7	1.2
7	4.2	1.1
8	4.6	1.15

Round	As	Pb
1	2.1	0.09 U
2	2.85	0.12 UJ
3	3.2	0.09 U
4	3.1	0.14 UJ
5	2.1	0.13
6	2.9	0.15 UJ
7	3.2	0.08 U
8	3.7	0.12

Round	As	Pb
1	2.7	0.18 UJ
2	3.2	0.23
3	2.3	0.14 UJ
4	3.2	0.29
5	2.7	0.15 J
6	3.2	0.21
7	2.5	0.15 UJ
8	2.8	0.24

Round	As	Pb
1	1.1	0.75
2	2.2	1.05
3	1.1	0.65
4	2.1	0.8
5	1.2	0.65
6	1.5	0.78
7	1.2	0.67
8	1.5	0.71

Round	As	Pb
1	4.2	0.78
2	5.3	0.9
3	4.8	0.7
4	5.4	0.86
5	4.3	0.75
6	5.6	0.9
7	4.3	0.65
8	5.4	0.9

FIGURE 5-12. Arsenic and lead concentrations in groundwater, background site

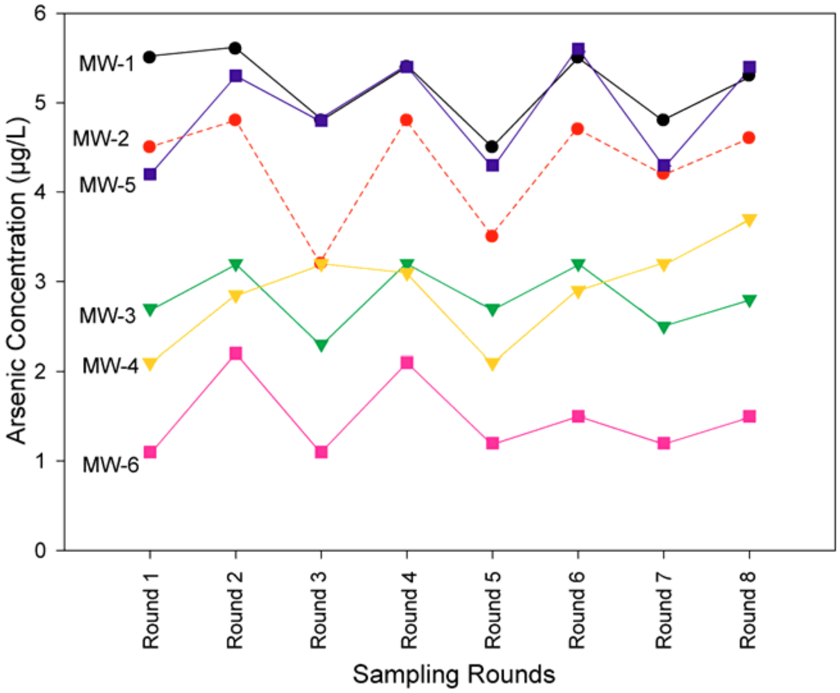


FIGURE 5-13. Spatial/temporal concentration plot, arsenic concentrations in groundwater, background site

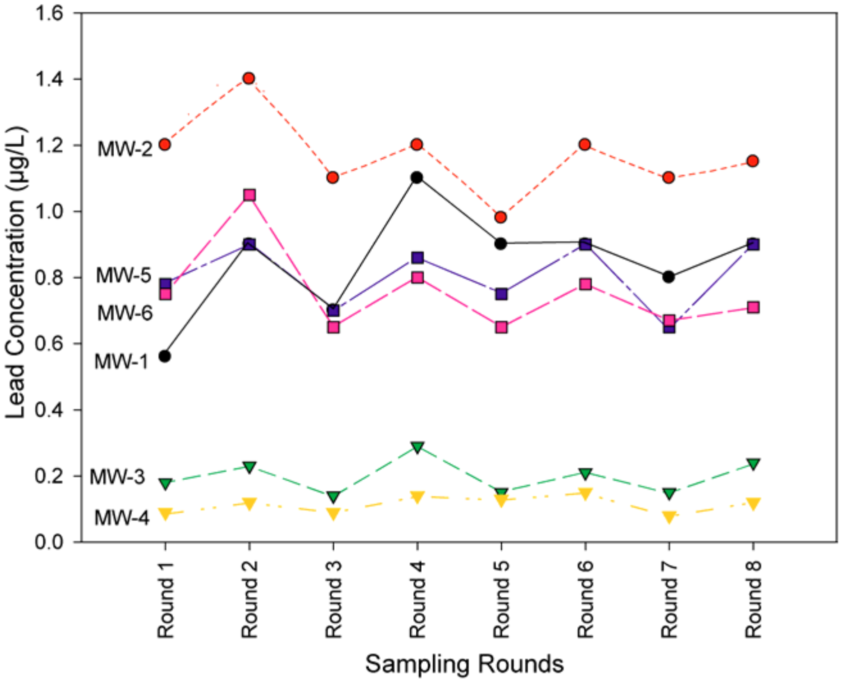


FIGURE 5-14. Spatial/temporal concentration plot, lead concentrations in groundwater, background site

5.3 Comparative Method Analysis

The Comparative Method was used to confirm the findings of the exploratory data and weight-of-evidence analyses presented above for the COPC metals (arsenic and lead). For this purpose, arsenic and lead concentrations from the background monitoring wells were considered "background data," whereas the Area 2 arsenic and lead concentrations were considered "site data." Figure 5-1 shows the relative locations of the background and investigation sites. The investigated samples were collected far apart in time and space, and thus were viewed as independent data. The COPC data associated with Area 2 and the background site are listed in Tables 5-5 and 5-9, respectively. Table 5-10 lists the descriptive statistics of the investigated data.

Table 5-10 also includes the results of the Shapiro-Wilk W test. As discussed in Section A.1.1, U.S. EPA (2000) recommends this test for evaluating the normality of datasets with less than 50 measurements. As listed in Table 5-10, normality tests are not conducted for the excluded Area 2 datasets (datasets not including MW-4 and MW-5 measurements). For such small datasets, usually with less than 20 independent measurements, results of any normality test are unreliable.

Previous analyses indicated that arsenic is a background chemical; however, lead measurements in samples from MW-4 and MW-5 are indicative of potential localized impacts. The results presented in Table 5-10 indicate that when data from wells MW-4 and MW-5 are excluded from the analysis, there is a significant impact on the lead statistics, which further supports the conclusion that lead measurements are elevated at MW-4 and MW-5.

5.3.1 Comparative Method Application

In order to determine whether arsenic and lead concentrations in Area 2 are statistically different from those in the background area, the Area 2 arsenic and lead data (i.e., site data) were statistically compared to the background arsenic and lead data.

Potential contamination in Area 2 either could cover the entire area (i.e., site-wide impact), or could be limited to small zones within the site (i.e., localized, or hotspot, impact). Given the general nonnormal nature of the datasets and the large number of nondetects among lead measurements, two types of nonparametric tests were selected:

- ❑ The Slippage test, to assess the differences among elevated concentrations in Area 2 and the background area; and
- ❑ The WRS test, to assess the difference between median concentrations of the two datasets.

Both these tests are nonparametric and do not require any specific distributional assumptions.

TABLE 5-10. Descriptive statistics

Parameter	Area 2 Data		Area 2 Data (Excluding MW-4 and MW-5 Data)		Background Data	
	Arsenic (µg/L)	Lead (µg/L)	Arsenic (µg/L)	Lead (µg/L)	Arsenic (µg/L)	Lead (µg/L)
Number of Samples	24	24	12	12	48	48
Mean	3.74	5.23	3.73	0.31	3.60	0.65
Standard Deviation	1.37	10.02	1.69	0.21	1.40	0.39
Minimum	1.1	0.08 U	1.1	0.08 U	1.1	0.08
Maximum	5.6	33.1	5.4	0.67	5.6	1.4
Median	3.45	1.24	4.65	0.22	3.35	0.73
Shapiro-Wilk Normality Test	0.067>0.05 (normal)	0.000<0.05 (not normal)	–	–	0.011<0.05 (not normal)	0.001<0.05 (not normal)
Percent Nondetects	0%	29%	0%	58%	0%	21%

5.3.2 Slippage Test Results

Results of the Slippage test ($\alpha = 0.05$), shown in Table 5-11, indicate an unusually large number of elevated lead concentrations in Area 2 compared to the maximum background area concentrations. This implies that site lead data display localized (i.e., hotspot) impact. Also, when the Slippage test

TABLE 5-11. Results of Slippage test ($\alpha = 0.05$)

Arsenic			Lead		
Area 2	**Background**	**All Data**	**Area 2**	**Background**	**All Data**
5.60	5.60	K = 0	33.10	1.40	K = 12
5.40	5.60	$K_c = 3$	32.00	1.20	$K_c = 3$
5.40	5.50	$K < K_c$	22.00	1.20	$K > K_c$
5.40	5.50	Arsenic is not a COPC	18.00	1.20	Lead is a COPC
5.30	5.40		2.40	1.15	All data except MW-4/MW-5 (highlighted) data
5.10	5.40		2.30	1.10	K = 0
4.80	5.40		2.20	1.10	$K_c = 2$
4.70	5.30		2.20	1.10	$K < K_c$
4.70	5.30		2.10	1.05	Lead is not a COPC
4.60	4.80		2.00	0.98	
4.10	4.80		1.80	0.90	
3.50	4.80		1.80	0.90	
3.40	4.80		0.67	0.90	
3.24	4.80		0.59	0.90	
3.20	4.70		0.48	0.90	
3.20	4.60		0.45	0.90	
3.15	4.50		0.45	0.90	
3.10	4.50		0.22	0.86	
2.94	4.30		0.22	0.80	
2.90	4.30		0.22	0.80	
1.90	4.20		0.15	0.78	
1.50	4.20		0.09	0.78	
1.50	3.70		0.09	0.75	
1.10	3.50		0.08	0.75	
	3.20			0.71	
	3.20			0.70	
	3.20			0.70	
	3.20			0.67	
	3.20			0.65	
	3.20			0.65	
	3.10			0.65	
	2.90			0.56	
	2.85			0.29	
	2.80			0.24	
	2.70			0.23	
	2.70			0.21	
	2.50			0.18	
	2.30			0.15	
	2.20			0.15	
	2.10			0.15	
	2.10			0.14	
	2.10			0.14	
	1.50			0.13	
	1.50			0.12	
	1.20			0.12	
	1.20			0.09	
	1.10			0.09	
	1.10			0.08	

Note: Highlighted cells correspond to lead data from Area 2 MW-4 and MW-5.
Test results indicate that data within the highlighted cells do not fit the background population.

is run and lead data from MW-4 and MW-5 are excluded, the results of the test are reversed, which further confirms the presence of localized impacts at these wells.

The Slippage test also indicated that elevated arsenic concentrations in Area 2 do not differ significantly from the maximum background area arsenic concentrations. These results confirm the earlier findings, which indicated that arsenic occurs only as a background chemical.

5.3.3 WRS Test Results

Results of the WRS test, shown in Table 5-12, results indicate that the median arsenic concentration in Area 2 groundwater is not significantly different from the median arsenic concentration in the background site groundwater. This result confirms the earlier findings, which indicated that arsenic occurs only as a background chemical.

The WRS test result for lead indicates that lead is not a COPC, which means that the observed lead concentrations at Area 2 do not display a site-wide impact. Instead, as detected by the Slippage test, the elevated concentrations of lead in MW-4 and MW-5 are localized. Under such a condition, the background data along with measurements from the unimpacted portion of Area 2 can be used to determine the background range for lead. This range extends from nondetect to 1.4 µg/L.

TABLE 5-12. Results of WRS test

Area	Number of Measurements	Mean Rank	Sum of Ranks
Arsenic			
Area 2	24	38	916
Background	48	36	1712
Test Statistics			
Wilcoxon W	1,712.00		
Significance	0.6322	>0.05	<95% significance
Conclusion: Arsenic is not a COPC.			
Lead			
Area 2	24	42	1015
Background	48	34	1614
Test Statistics			
Wilcoxon W	1,613.50		
Significance	0.0978	>0.05	<95% significance
Conclusion: Lead is not a (site-wide) COPC.			

5.4 Summary of Case Study Results

Results of the case study analysis are summarized as follows:

- All arsenic concentrations detected in the Area 2 groundwater samples are within the background range. The estimated background arsenic concentration range extends from 1.1 to 5.6 µg/L.

- Lead is a groundwater COPC, and the above-background concentrations detected in Area 2 represent one or more localized hotspots.

  The extent of lead contamination in the Area 2 groundwater is most likely limited to the vicinity of MW-4 and MW-5. The estimated background lead concentration range extends from nondetect to 1.4 µg/L.

5.5 References

Cercelius, E.A., N.S. Bloom, C.E. Cowan, and E.A. Jeane. 1986. *Speciation of Selenium and Arsenic in Natural Waters and Sediments*. Arsenic Speciation, EPRI Project 20202-2. Sequim, WA: Battelle Pacific Northwest Laboratory.

Masscheleyn, P.H., R.D. Delaune, and W.H. Patrick, Jr. 1991. Effect of Redox Potential and pH on Arsenic Speciation and Solubility in a Contaminated Soil. *Environmental Science and Technology* 25:1414–1419.

United States Environmental Protection Agency (U.S. EPA). 1984. *Health Assessment Document for Inorganic Arsenic*. U.S. Environmental Protection Agency. EPA-600/8-83-021F. Washington, DC: U.S. EPA Office of Health and Environmental Assessment, Office of Research and Development. March.

United States Environmental Protection Agency (U.S. EPA). 2000. *Guidance for Data Quality*

Assessment: Practical Methods for Data Analysis. EPA QA/G-9, QA 00 Update. EPA/600/R-96/084. Washington, DC: Office of Environmental Information. July. http://www.epa.gov/quality1/qa_docs.html.

Xinde Cao, L.Q. Ma, M. Chen, D.W. Hardison, Jr., and W.G. Harris. 2003. Weathering of Lead Bullets and Their Environmental Effects at Outdoor Shooting Ranges. *J. Environ. Qual.* 32:526–534.

APPENDIX A: SUPPLEMENTARY STATISTICAL TESTS

A.1 Statistical Tests of Normality

A.1.1 Shapiro-Wilk W Test

The Shapiro-Wilk W test is highly recommended for testing whether data have a normal distribution. It also may be used to test for a lognormal distribution, if the data are first transformed by computing the natural logarithm of each datum. The W test is recommended in several U.S. EPA guidance documents (U.S. EPA, 1992a, 2000) and in many statistical texts (Gilbert, 1987; Conover, 1998). It is available in many software packages including GRITS/STAT (U.S. EPA, 1992b) and DataQUEST (U.S. EPA, 1997). The W test has been shown to have more power than other tests to detect when data are not from a normal or lognormal distribution. The W test should be conducted in conjunction with constructing normal and lognormal probability plots in order to visually confirm whether the normal or lognormal distribution is an acceptable fit to the data. The W test:

- ❑ Requires the use of a table of coefficients (Table B-10) and critical values (Table B-11).
- ❑ Can only be conducted if the number of samples is less than or equal to 50 because the table of critical values (Table B-11) does not extend beyond n = 50. For larger datasets, either the D'Agostino test can be performed or the Shapiro-Wilk W test can be used as it is implemented in statistical software packages such as SPSS (see Appendix C.3).
- ❑ Is easily conducted using DataQUEST if $n \leq 50$.
- ❑ Is not designed to process datasets with non-detects.
- ❑ May not have sufficient power to detect non-normality if the underlying distribution is only slightly different than the normal distribution or if the number of data in the dataset is too small.

The computations needed to conduct the W test are provided in Box A-1 along with an example.

A.1.2 D'Agostino Test

The D'Agostino test (D'Agostino, 1971) may be used to test if data are from a normal distribution. This test can be used when the number of samples exceeds 50. D'Agostino (1971) showed the performance of the test compares favorably with other tests. The same assumptions and their verification for applying the W test also apply to this test. The D'Agostino test:

- ❑ Cannot be conducted if $n < 50$ or $n > 1,000$
- ❑ Requires the use of a special table of critical values (Table B-12)
- ❑ Is not designed to process datasets with non-detects
- ❑ May not have large power to detect non-normality if the underlying distribution is only slightly different than the normal distribution or if the number of data in the dataset is small.

The computations necessary to conduct the test are provided in Box A-2 along with an example.

A.1.3 Other Tests

In addition to the Shapiro-Wilk W and D'Agostino tests, other statistical procedures can be used to test hypotheses about which probability distribution best fits a dataset. These tests are commonly called "goodness-of-fit tests." A thorough summary of the scientific literature on this topic is provided in D'Agostino and Stephens (1986). U.S. EPA (2000) provides descriptions of several tests, most of which can be conducted using the DataQUEST software (U.S. EPA, 1997).

U.S. EPA (2000) recommends the use of the Shapiro-Wilk W test if the number of samples is less than 50; otherwise, either the Filliben statistic or the studentized range test is recommended. The Filliben test (Filliben, 1975; U.S. EPA, 1997) is closely related to the Shapiro-Wilk W test. The studentized range test also is recommended except when the data appear to be lognormally distributed. The studentized range test, as illustrated in U.S. EPA (2000), is simpler to compute than the Shapiro-Wilk W test and critical values needed for the test are available for sample sizes up to 1,000.

If several goodness-of-fit tests are applied to the same dataset, the test results may differ. Under such conditions, the results of the most appropriate tests with respect to the investigated dataset, in conjunction with probability plots, should be considered as the basis for subsequent decisions.

A.2 Descriptive Summary Statistics for Datasets with Large Numbers of Nondetects

Descriptive summary statistics of datasets with large numbers of nondetects can be computed using the Cohen method or calculating a rimmed mean or a Winsorized mean and standard deviation. These methods are defined and their assumptions, advantages, and disadvantages are listed in Box A-3. Examples of computing the median, trimmed mean, the Winsorized mean and standard deviation are illustrated in Box A-4. The Cohen method for computing the mean and standard deviation of a normally distributed set of data that contains nondetects (i.e., a censored dataset) is explained and illustrated in Box A-5.

Cautionary Note

If more than 50% of the measurements in the dataset are nondetects, the loss of information is too great for descriptive statistics to provide much insight into the underlying distribution of measurements. The only descriptive statistics that might be possible to compute are *p*th percentiles for values of *p* that are greater than the proportion of nondetects present in the sample and when no nondetects are greater than the k(n+1)th largest datum, where *k* is defined in Box 4-18 of the main document.

U.S. EPA (2000) cautions that no general procedures exist for the statistical analyses of censored datasets that can be used in all applications of statistical analysis. For this reason, U.S. EPA guidelines should be implemented cautiously. U.S. EPA (2000) also suggests the data analyst should consult a statistician for the most appropriate way to statistically evaluate or analyze a dataset that contains nondetects.

Akritas et al. (1994, pp. 221-242) provide a review of the statistical literature that deals with the statistical analysis of censored environmental datasets. Further review is provided by Helsel and Hirsch (1992) and Singh and Nocerino (2002).

A.3 Statistical Tests for Outliers

This section provides detailed description of various statistical tests for determining outliers. The assumptions, advantages, and disadvantages of each test are provided in Table A-1. The procedures for conducting the Dixon extreme value test, the Discordance test, and the Walsh test, with an example for each, are provided in Boxes A-6, A-7, and A-8, respectively. The Rosner test is described in Box A-9 and illustrated in Box A-10. It should be noted that in background analysis, outlier measurements are deleted, only if they are determined to be the results of field, laboratory, or processing errors.

A.4 References

Akritas, M.G., T.F. Ruscitti, and G.P. Patil. 1994. Statistical Analysis of Censored Environmental Data. In *Handbook of Statistics, Volume 12, Environmental Statistics*, eds. G.P. Patil and C.R. Rao. New York: North-Holland.

Conover, W.J. 1998. *Practical Nonparametric Statistics*. 3rd ed. Wiley Series in Probability and Statistics, Applied Probability and Statistics Section. New York: Wiley & Sons.

D'Agostino, R.B. 1971. An Omnibus Test of Normality for Moderate and Large Size Samples. *Biometrika* 58:341-348.

D'Agostino, R.B., and M.A. Stephens. 1986. *Goodness-of-Fit Techniques*. New York: Marcel Dekker.

Filliben, J.J. 1975. The Probability Plot Correlation Coefficient Test for Normality. *Technometrics* 17:111-117.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring.* New York: Van Nostrand Reinhold.

Helsel, D.R., and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. New York: Elsevier.

Singh, A., and J. Nocerino. 2002. Robust Estimation of Mean and Variance Using Environmental Data Sets with Below Detection Limit Observations. *Chemometrics and Intelligent Laboratory Systems 60.* pp. 69-86.

United States Environmental Protection Agency (U.S. EPA). 1992a. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities – Addendum to Interim Final Guidance.* Washington, DC: Office of Solid Waste, Permits and State Programs Division. June.

United States Environmental Protection Agency (U.S. EPA). 1992b. *User Documentation: GRITS/STAT, v4.2.* EPA/625/11-91/002. Cincinnati: Center for Environmental Research Information. http://www.epa.gov/esd/databases/grits-stat/abstract.htm.

United States Environmental Protection Agency (U.S. EPA). 1997. *Data Quality Evaluation Statistical Toolbox (DataQUEST), User's Guide, EPA QA/G-9D, QA96 Version.* EPA/600/R-96/085. Washington, DC: Office of Research and Development. December. http://www.epa.gov/quality1/qa_docs.html.

United States Environmental Protection Agency (U.S. EPA). 2000. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis. EPA QA/G-9, QA 00 Update.* EPA/600/R-96/084. Washington, DC: Office of Environmental Information. July. http://www.epa.gov/quality1/qa_docs.html.

BOX A-1. Shapiro-Wilk W test procedure

Select the significance level, α, desired for the test, where $0 < \alpha < 0.5$. That is, select the probability, α, that can be tolerated of the W test declaring that the measurements in the dataset are not from a normal distribution when in fact they are from a normal distribution.

- Compute the arithmetic mean of the *n* data: $\bar{x} = (x_1 + x_2 + \ldots + x_n) / n$
- Compute the denominator *d* of the W test statistic using the *n* data and $\bar{x}$:

$$d = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2$$

- Order the *n* data from smallest to largest. Denote these "sample order statistics" by $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$, where $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$.
- Compute *k*, where $k = n/2$ if *n* is an even integer and $k = (n - 1)/2$ if *n* is an odd integer
- Turn to Table B-10 to obtain the coefficients $a_1, a_2, \ldots, a_k$ for the value of *n*.
- Compute the W test statistic

$$W = \{ a_1(x_{(n)} - x_{(1)}) + a_2(x_{(n-1)} - x_{(2)}) + \ldots + a_k(x_{(n-k+1)} - x_{(k)})\}^2 / d$$

- Conclude that the dataset is not normally distributed if the value of *W* is less than the critical value given in Table B-11 for the selected significance level α.

Example:

- Suppose $\alpha = 0.05$
- Suppose there are $n = 10$ measurements in the dataset:

1.20, 0.13, 1.69, 1.05, 1.12, 0.45, 2.06, 0.60, 0.76, 1.37

- The arithmetic mean of these data is

$$\bar{x} = (1.2 + 0.13 + 1.69 + 1.05 + 1.12 + 0.45 + 2.06 + 0.60 + 0.76 + 1.37) / 10$$
$$= 1.04$$

- The denominator *d* of the W test statistic using the *n* data and $\bar{x}$ is:

$$d = (1.2 - 1.04)^2 + (0.13 - 1.04)^2 + \ldots + (1.37 - 1.04)^2 = 3.05$$

- Order the $n = 10$ measurements from smallest to largest to obtain:

0.13, 0.45, 0.60, 0.76, 1.05, 1.12, 1.20, 1.37, 1.69, 2.06

- Compute $k = n/2 = 10/2 = 5$ because *n* is an even integer.
- In Table B-10, the $k = 5$ coefficients are

$$a_1 = 0.5739,\ a_2 = 0.3291,\ a_3 = 0.2141,\ a_4 = 0.1224,\ a_5 = 0.0399$$

- Therefore, the computed W statistic is:

$$W = \{0.5739(2.06 - 0.13) + 0.3291(1.69 - 0.45) + 0.2141(1.37 - 0.60) + 0.1224(1.20 - 0.76) + 0.0399(1.12 - 1.05)\}^2 / 3.05$$
$$= 0.989$$

The critical value from Table B-11 for $n = 10$ and $\alpha = 0.05$ is 0.842. Therefore, because 0.989 is not less than 0.842, the measurements appear to be normally distributed. The data do not provide convincing evidence that the distribution of the measurements is not normal.

BOX A-2. Procedure for conducting the D'Agostino test

- Select the significance level, α, desired for the test, where $0 < \alpha < 0.5$.
- Compute $s = \{[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2] / n\}^{1/2}$
- Order the *n* data from smallest to largest. Denote these sample order statistics by

$$x_{(1)}, x_{(2)}, \ldots, x_{(n),} \text{ where } x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$$

- Compute $D = \{[1 - 0.5(n+1)]x_{(1)} + [2 - 0.5(n+1)]x_{(2)} + \ldots + [n - 0.5(n+1)]x_{(n)} \} / n^2s$
- Compute $Y = (D - 0.282094) / (0.02998598 / n^{1/2})$
- Conclude the data are not from a normal distribution, if Y is less than the critical value $Y_{\alpha/2}$ or greater than the critical value $Y_{1-\alpha/2}$, that are found in Table B-12 for each value of *n*.

Example (from Gilbert, 1987, p. 161):

- Suppose $\alpha = 0.05$
- Suppose n = 115 and the computed value of *s* is

$$\{[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2] / 115\}^{1/2} = 0.4978$$

- Then the value of n^2s, the denominator of *D*, is $(115)^2(0.4978) = 6{,}583$
- As $0.5(n+1) = 0.5(116) = 58$, and using the sample order statistics $x_{[i]}$, the numerator of *D* equals

$$\{[1-58]x_{(1)} + [2-58]x_{(2)} + \ldots + [115 - 58]x_{(115)} \} = 1{,}833.3$$

- Therefore, $D = 1833.3 / 6583 = 0.2785$
- Therefore, $Y = (0.2785 - 0.282094) / (0.02998798 / 115^{1/2}) = -1.29$
- Using Table B-12, linear interpolation indicates that $Y_{0.025} = -2.522$ and $Y_{0.975} = 1.339$.
- Because −1.29 is not less than −2.522 and not larger than 1.339, it cannot be concluded that the measurements are not normally distributed.

BOX A-3. Descriptive statistics when 15% to 50% of the dataset are nondetects

Method	Assumptions	Advantages	Disadvantages
Median (when *n* is an odd or an even integer): Determine the median in the usual way as illustrated in Box 4-18 of the main document.	• The largest nondetect is less than the median of the entire dataset (detects + nondetects); i.e., there are no nondetects in the upper 50% of the measurements.	• A simple procedure.	• The median cannot be determined, if the assumption is not true.
100p% Trimmed Mean: Determine the percentage (100p%) of measurements below the DL. Discard the largest *np* measurements and the smallest *np* measurements. Compute the arithmetic mean on the n(1 − 2p) remaining measurements.	• All nondetects have the same DL. • All detects are larger than the DL. • The number of nondetects is no more than *np.* • The underlying distribution of measurements is symmetric (not skewed). • $0 < p < 0.50$.	• Trimmed mean is not affected by outliers that have been trimmed from the dataset.	• Cannot be used if the assumptions are not true.

BOX A-3. (cont'd)

Method	Assumptions	Advantages	Disadvantages
Winsorized Mean ($\bar{x}_w$): If n' nondetects are in the lower tail of a dataset with n measurements (including nondetects). • Replace the n' nondetects by the next *largest* detected datum. • Also replace the n' largest measurements by the next *smallest* measurement. • Obtain the Winsorized Mean, $\bar{x}_w$, by computing the arithmetic mean of the resulting set of n measurements.	• All nondetects have the same DL. • All detects are larger than the DL. • The underlying distribution of the measurements is symmetric (not skewed).	• Winsorized mean is not affected by outliers that are among the largest measurements.	• Cannot be used if the assumptions are not true.
Winsorized Standard Deviation (s_w) Suppose n' nondetects are in the lower tail of a dataset with n measurements (detects plus nondetects). • Replace the n' nondetects by the next *largest* detected datum. • Also replace the n' largest measurements by the next *smallest* measurement. • Compute the standard deviation, *s,* of the new set of n measurements. • Compute $s_w = [s(n-1)]/(v-1)$ where $v = n - 2n'$ is the number of measurements not replaced during the Winsorization process.	• All nondetects have the same DL. • All detects are greater than the DL. • The underlying distribution of the measurements is symmetric (not skewed). • The quantity v must be greater than 1.	• If the measurements are normally distributed, then confidence intervals for the mean can be computed using the method in Gilbert (1987, p. 180).	• Cannot be used if the assumptions are not true.
Cohen Method for Mean and Standard Deviation (see Box A-5).	• All nondetects have the same DL. • The underlying distribution of the measurements is normal. • Measurements obtained are representative of the underlying normal distribution.	• Has good performance if the underlying assumptions are valid and if the number of samples is sufficiently large.	• The assumptions must be valid.
*p*th Sample Percentile The *p*th sample percentile is computed as described in Box 4-18 of the main document.	• All nondetects have the same DL. • All detects are greater than the DL. • The computed value of k (see Box 4-18 of the main document) must be larger than the number of nondetects plus 1.	• Provides an estimate of the value that is exceeded by 100(1 – p)% of the underlying population.	• Cannot be computed when the assumption on k is not valid.

Sources: Gilbert (1987); U.S. EPA (2000).

BOX A-4. Examples of computing the median, trimmed mean, and Winsorized mean and standard deviation using a dataset that contains nondetects

The following examples use this dataset of 12 measurements (after being ordered from smallest to largest): <0.15, <0.15, <0.15, 0.18, 0.25, 0.26, 0.27, 0.36, 0.50, 0.62, 0.63, 0.79. Note three nondetects are in this dataset, but each one has the same RL, 0.15. If multiple RLs are present, consult a statistician for the best way to summarize the data.

Median

The median of the dataset is (0.26 + 0.27) / 2 = 0.265. Note the nondetects do not have any impact on computing the median because fewer than half of the data were nondetects.

100p% Trimmed Mean

The percentage of nondetect measurements is 100(3/12) = 25%. Therefore, set p = 0.25 and compute the 25% trimmed mean (25% of n is 3). Discard the smallest 0.25(12) = 3 and largest 3 measurements, i.e., discard the three nondetects and the measurements 0.62, 0.63, 0.79. Compute the arithmetic mean on the remaining six measurements: Trimmed Mean = (0.18 + 0.25 + 0.26 + 0.27 + 0.36 + 0.50) / 6 = 0.30. This estimate is valid, if the underlying distribution of measurements is symmetric. If the distribution is not symmetric, this trimmed mean is a biased estimate.

Winsorized Mean

Replace the three nondetects by the next largest detected datum, which is 0.18. Replace the three largest measurements by the next smallest measurement, which is 0.50. Compute the arithmetic mean of the new set of 12 data: 0.18, 0.18, 0.18, 0.18, 0.25, 0.26, 0.27, 0.36, 0.50, 0.50, 0.50, 0.50.

$$\bar{x}_w = (0.18 + 0.18 + 0.18 + 0.18 + 0.25 + 0.26 + 0.27 + 0.36 + 0.50 + 0.50 + 0.50 + 0.50) / 12 = 0.32$$

This estimate is valid if the underlying distribution of measurements is symmetric. If the distribution is not symmetric, this Winsorized mean is a biased estimate.

Winsorized Standard Deviation

Replace the three nondetects by the next largest detected datum, which is 0.18. Replace the three largest measurements by the next smallest measurement, which is 0.50. Compute the standard deviation, s, of the new set of 12 data:

$$s = [(0.18 - 0.32)^2 + (0.18 - 0.32)^2 + (0.18 - 0.32)^2 + (0.18 - 0.32)^2 + (0.25 - 0.32)^2 + (0.26 - 0.32)^2 + (0.27 - 0.32)^2 + (0.36 - 0.32)^2 + (0.50 - 0.32)^2 + (0.50 - 0.32)^2 + (0.50 - 0.32)^2 + (0.50 - 0.32)^2] / 11$$

$$= 0.1416$$

Compute $v = n - 2n' = 12 - 2(3) = 6$

Compute the Winsorized Standard Deviation:

$$s_w = [s(n - 1)]/(v - 1) = [0.1416(11)] / 5 = 0.31$$

This estimate is valid if the underlying distribution of measurements is symmetric. If the distribution is not symmetric, this Winsorized standard deviation is a biased estimate.

BOX A-5. Cohen method for computing the mean and variance of a censored dataset (U.S. EPA, 2000; Gilbert, 1987, p. 182)

- Let the single nondetect limit be denoted by ND. Let $x_1, x_2, \ldots, x_n$ denote the *n* measurements in the dataset, including those that are less than ND. Let *k* be the number out of *n* that are greater than the ND.
- Compute $h = (n - k)/n$, which is the fraction of the *n* measurements that are below the ND.
- Compute the arithmetic mean of the *k* measurements that exceed the ND as follows:

$$\overline{x}_c = (x_1 + x_2 + \ldots + x_k) / k$$

where $x_1, x_2, \ldots,$ and x_k are all the measurements >ND.

- Compute the following statistic using the *k* measurements that exceed the ND:

$$s_c^2 = [(x_1 - \overline{x}_c)^2 + (x_2 - \overline{x}_c)^2 + \ldots + (x_k - \overline{x}_c)^2] / k$$

- Compute $G = s_c^2 / (\overline{x}_c - ND)^2$
- Obtain the value of λ from Table B-13 for values of h and γ. Use linear interpolation in the table if necessary.
- Compute the Cohen mean and variance as follows:

$$\text{Cohen mean} = \overline{x}_c - \lambda(\overline{x}_c - ND)$$

$$\text{Cohen variance} = s_c^2 + \lambda(\overline{x}_c - ND)^2$$

- Cohen standard deviation is the square root of Cohen variance.

Example:

- n = 25 measurements of a chemical in sediment samples were obtained. The nondetect limit was equal to 36. Five measurements were reported as <36. The data obtained were:

<36, <36, <36, <36, <36, 49, 49, 59, 61, 62, 62, 65, 65, 65, 70, 72, 80, 80, 99, 99, 104, 110, 140 142, 144

- Compute $h = (25 - 20)/25 = 0.20$ = fraction of the 25 measurements that are below the ND.
- Compute the arithmetic mean of the 20 measurements that exceed the ND:

$$\overline{x}_c = (49 + 49 + 59 + \ldots + 142 + 144) = 83.85$$

- Compute $s_c^2 = [(49 - 83.85)^2 + (49 - 83.85)^2 + (59 - 83.85)^2 + \ldots + (142 - 83.85)^2 + (144 - 83.85)^2] / 20$

$$= 882.63$$

- Compute $G = 882.63 / (83.85 - 36)^2 = 0.385$.
- Using Table B-13, linear interpolation between $\gamma = 0.35$ and $\gamma = 0.40$ for $h = 0.20$ indicates that $\lambda = 0.291$.
- Therefore, Cohen mean and variance are:

$$\text{Cohen mean} = 83.85 - 0.291(83.85 - 36) = 69.9$$

$$\text{Cohen variance} = 882.63 + 0.291(83.85 - 36)^2 = 1{,}548.9$$

- Cohen standard deviation = $(1{,}548.9)^{1/2} = 39.4$

TABLE A-1. Assumptions, advantages, and disadvantages of outlier tests

Statistical Test	Assumptions	Advantages	Disadvantages
Dixon Test	• $n \leq 25$ • Measurements are representative of the underlying population. • The measurements without the suspect outlier are normally distributed; otherwise, see a statistician. • Test can be used to test for either one suspect large outlier or one suspect small outlier. The latter case is not considered here as it is not of interest for identifying COPCs.	• Simple to compute by hand. • The test is available in the DataQUEST software (U.S. EPA, 1997).	• Test should be used for only one suspected outlier. Use the Rosner test if multiple suspected outliers are present. • Must conduct a test for normality on the dataset after deleting the suspect outlier and before using the Dixon test.
Discordance Test	• $3 < n \leq 50$ • Measurements are representative of underlying population. • The measurements without the suspected outlier are normally distributed; otherwise, see a statistician. • Test can be used to test that the largest measurement is a suspected outlier or the smallest measurement is a suspected outlier. The latter case is not considered here as it is not of interest for identifying COPCs.	• Simple to compute by hand. • The test is available in the DataQUEST software (U.S. EPA, 1997).	• Test can be used for only one suspected outlier. Use the Rosner test if there are multiple suspected outliers. • Must conduct a test for normality on the dataset after deleting the suspect outlier and before using the Discordance test.
Rosner Test	• $n \geq 25$ • Measurements are representative of underlying population. • The measurements without the suspected outliers are normally distributed; otherwise, see a statistician.	• Can test for up to 10 outliers. • The test is available in the DataQUEST software (U.S. EPA, 1997).	• Must conduct a test for normality after deleting the suspected outliers and before using the Rosner test. • Computations are more complex than for the Dixon test or the Discordance test.
Walsh Test	• $n > 60$ • Measurements are representative of the underlying population. • Test can be used to test that the largest *r* measurements or the smallest *r* measurements are suspected outliers. The latter case (discussed in U.S. EPA, 2000) is not considered here as it is not of interest for identifying COPCs.	• Can test for 1 or more outliers. • The measurements need not be normally distributed. • Need not conduct a test for normality before using the test. • The test is available in the DataQUEST software (U.S. EPA, 1997).	• Must have $n > 60$ to conduct the test • The test can only be performed for the $\alpha = 0.05$ and 0.10 significance levels, and the α level used depends on *n*: the $\alpha = 0.05$ level can only be used if $n > 220$ and the $\alpha = 0.10$ level can only be used if $60 < n \leq 220$. • Test calculations are more complex than for the Dixon test or the Discordance test. • The number of identified suspected outliers, *r*, are accepted or rejected as a group rather than one at a time.

BOX A-6. Procedure for conducting the Dixon extreme value outlier test (U.S. EPA, 2000)

- Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the *n* measurements in the dataset after they have been listed in order from smallest to largest. The parentheses around the subscripts indicate the measurements are ordered from smallest to largest.
- $x_{(n)}$ (the largest measurement) is suspected of being an outlier.
- Perform test for normality on $x_{(1)}$ through $x_{(n-1)}$.
- Specify the tolerable decision error rate, α (significance level), desired for the test. α may only be set equal to 0.01, 0.05, or 0.10 for the Dixon test.
- Compute C

$$\begin{aligned} C &= [x_{(n)} - x_{(n-1)}] / [x_{(n)} - x_{(1)}] \text{ if } 3 \le n \le 7 \\ &= [x_{(n)} - x_{(n-1)}] / [x_{(n)} - x_{(2)}] \text{ if } 8 \le n \le 10 \\ &= [x_{(n)} - x_{(n-2)}] / [x_{(n)} - x_{(2)}] \text{ if } 11 \le n \le 13 \\ &= [x_{(n)} - x_{(n-2)}] / [x_{(n)} - x_{(3)}] \text{ if } 14 \le n \le 25 \end{aligned}$$

If C exceeds the critical value in Table B-14 for the specified *n* and α, then declare that $x_{(n)}$ is an outlier and should be investigated further.

Example: Suppose the ordered dataset is 34, 50, 52, 60, 62. We wish to test if 62 is an outlier from an assumed normal distribution for the n = 5 data. Perform a test for normality on the data 34, 50, 52, 60. Any test for normality will have little ability to detect nonnormality on the basis of only four data values. Suppose α is selected to be 0.05, i.e., there should be no more than a 5% chance that the test will incorrectly declare the largest observed measurement to be an outlier. Compute C = (62 – 60)/(62 – 34) = 0.071. Determine the test critical value from Table B-14. The critical value is 0.642 when n = 5 and α = 0.05. As 0.071 is less than 0.642, the data do not indicate the measurement 62 is an outlier from an assumed normally distribution.

BOX A-7. Procedure for conducting the Discordance outlier test (U.S. EPA, 2000)

- Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the *n* measurements in the dataset after they have been listed in order from smallest to largest.
- $x_{(n)}$ (the largest measurement) is suspected of being an outlier.
- Specify the tolerable decision error rate, α (significance level) desired for the test. α may be specified to be 0.01 or 0.05 for the Discordance outlier test.
- Compute the sample arithmetic mean, $\bar{x}$, and the sample standard deviation, *s*.
- Compute $D = [x_{(n)} - \bar{x}] / s$
- If D exceeds the critical value from Table B-15 for the specified *n* and α, $x_{(n)}$ is an outlier and should be further investigated.

Example: Suppose the ordered dataset is 34, 50, 52, 60, 62. We wish to test if 62 is an outlier from an assumed normal distribution for the data. Suppose α is selected to be 0.05. Using the n = 5 data, we compute $\bar{x} = 51.6$ and s = 11.08. Therefore, D = (62 – 51.6) / 11.08 = 0.939. The critical value from Table B-15 for n = 5 and α = 0.05 is 1.672. As 0.939 is less than 1.672, the data do not indicate the measurement 62 is an outlier from an assumed normally distribution.

BOX A-8. Procedure for conducting the Walsh outlier test (U.S. EPA, 2000)

- Let $x_{(1)}$, $x_{(2)}$, …, $x_{(n)}$ denote *n* measurements in the dataset after they have been listed in order from smallest to largest. Do not apply the test if $n < 60$. If $60 < n \leq 220$, then use $\alpha = 0.10$. If $n > 220$, then use $\alpha = 0.05$.
- Identify the number of possible outliers, *r,* where *r* can equal 1.
- Compute: $c = [(2n)^{1/2}]$, $k = r + c$, $b^2 = 1/\alpha$,

  $$a = (1 + b\{(c - b^2)/(c - 1)\}^{1/2}) / (c - b^2 - 1)$$

  where [ ] indicates rounding the value to the largest possible integer (i.e., 3.24 becomes 4).
- The Walsh test declares that the *r* largest measurements are outliers (with a α level of significance) if

  $$x_{(n+1-r)} - (1 + a)x_{(n-r)} + ax_{(n+1-k)} > 0$$

Example: Suppose $n = 70$ and that $r = 3$ largest measurements are suspected outliers. The significance level $\alpha =$ 0.10 must be used because $60 < n \leq 220$. That is, we must accept a probability of 0.10 that the test will incorrectly declare that the three largest measurements are outliers.

- Compute
  $c = [(2 \times 70)^{1/2}] = 12$
  $k = 3 + 12 = 15$
  $b^2 = 1 / 0.10 = 10$
  $a = 1 + 3.162\{(12 - 10) / (12 - 1)\}^{1/2}\} / (12 - 10 - 1) = 2.348$
- $x_{(n+1-r)} = x_{(70+1-3)} = x_{(68)}$ is the 68th largest measurement (two measurements are larger)
  $x_{(n-r)} = x_{(70-3)} = x_{(67)}$ is the 67th largest measurement
  $x_{(n+1-k)} = x_{(70+1-15)} = x_{(56)}$ is the 56th largest measurement
- Order the 70 measurements from smallest to largest. Suppose $x_{(68)} = 83$, $x_{(67)} = 81$, and $x_{(56)} = 20$.
- Compute $x_{(n+1-r)} - (1+a)x_{(n-r)} + ax_{(n+1-k)} = 83 - (1+2.348)81 + 2.348(20) = -141.22$ which is smaller than 0. Therefore, the Walsh test indicates that the three largest measurements are not outliers.

BOX A-9. Procedure for conducting the Rosner outlier test (U.S. EPA, 2000)

STEP 1:

- Select the desired significance level α, i.e., the tolerable probability that the Rosner test will falsely declare that outliers are present.
- Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ denote *n* measurements in the dataset after they have been listed in order from smallest to largest, where $n \geq 25$.
- Identify the maximum number of possible outliers, denoted by *r*.

STEP 2:

- Set i = 0 and use the following formulas to compute the sample arithmetic mean, labeled $\overline{x}^{(0)}$, and $s^{(0)}$ using all *n* measurements. Determine the measurement that is farthest from $\overline{x}^{(0)}$ and label it $y^{(0)}$.

$$\overline{x}^{(i)} = (x_1 + x_2 + \ldots + x_{n-i}) / (n - i)$$

$$s^{(i)} = \{[(x_1 - \overline{x}^{(i)})^2 + (x_2 - \overline{x}^{(i)})^2 + \ldots + (x_{n-i} - \overline{x}^{(i)})^2] / (n - i)\}^{1/2}$$

- Delete $y^{(0)}$ from the dataset of *n* measurements and compute (using i = 1 in the above formulas) the sample arithmetic mean, labeled $\overline{x}^{(1)}$, and $s^{(1)}$ on the remaining n – 1 measurements. Determine the measurement that is farthest from $\overline{x}^{(1)}$ and label it $y^{(1)}$.
- Delete $y^{(1)}$ from the dataset and compute (using i = 2 in the above formulas) the sample arithmetic mean, labeled $\overline{x}^{(2)}$, and $s^{(2)}$ on the remaining n – 2 measurements.
- Continue using this process until the *r* largest measurements have been deleted from the dataset.
- The values of $\overline{x}^{(0)}$, $\overline{x}^{(1)}$, …, $s^{(0)}$, $s^{(1)}$, … are computed using the above formulas.

STEP 3:

- To test if there are *r* outliers in the dataset, compute

$$R_r = [\,|\,y^{(r-1)} - \overline{x}^{(r-1)}\,|\,] / s^{(r-1)}$$

- Determine the critical value λ_r from Table B-16 for the values of *n, r,* and α.
- If R_r exceeds λ_r, conclude *r* outliers are in the dataset.
- If not, test if r – 1 outliers are present. Compute

$$R_{r-1} = [\,|\,y^{(r-2)} - \overline{x}^{(r-2)}\,|\,] / s^{(r-2)}$$

- Determine the critical value λ_{r-1} from Table B-16 for the values of *n,* r – 1, and α.
- If R_{r-1} exceeds λ_{r-1}, conclude r – 1 outliers are in the dataset.
- Continue on in this way until either it is determined that there are a certain number of outliers are present or that no outliers exist at all.

BOX A-10. Example of the Rosner outlier test

STEP 1: Consider the following 32 data points (in ppm) listed in order from smallest to largest: 2.07, 40.55, 84.15, 88.41, 98.84, 100.54, 115.37, 121.19, 122.08, 125.84, 129.47, 131.90, 149.06, 163.89, 166.77, 171.91, 178.23, 181.64, 185.47, 187.64, 193.73, 199.74, 209.43, 213.29, 223.14, 225.12, 232.72, 233.21, 239.97, 251.12, 275.36, and 395.67.

A normal probability plot of the data identified four potential outliers: 2.07, 40.55, 275.36, and 395.67. Moreover, a normal probability plot of the dataset after excluding the four suspect outliers provided no evidence that the data are not normally distributed.

STEP 2: First use the formulas in Step 2 of Box A-9 to compute $\bar{x}^{(0)}$ and $s^{(0)}$ using the entire dataset. Using subtraction, it was found that 395.67 was the farthest data point from $\bar{x}^{(0)}$, so $y^{(0)} = 395.67$. Then 395.67 was deleted from the dataset and $\bar{x}^{(1)}$ and $s^{(1)}$ are computed on the remaining data. Using subtraction, it was found that 2.07 was the farthest value from $\bar{x}^{(1)}$, so $y^{(1)} = 2.07$. This value then was dropped from the data and the process was repeated to determine $\bar{x}^{(2)}$, $s^{(2)}$, $y^{(2)}$ and $\bar{x}^{(3)}$, $s^{(3)}$, $y^{(3)}$. These values are summarized in the table.

i	$\bar{x}^{(i)}$	$s^{(i)}$	$y^{(i)}$
0	169.92	73.95	395.67
1	162.64	62.83	2.07
2	167.99	56.49	40.55
3	172.39	52.18	275.36

STEP 3: To apply the Rosner test, first test if four outliers are present. Compute

$$R_4 = |y^{(3)} - \bar{x}^{(3)}| \,/\, s^{(3)} = |275.36 - 172.39| \,/\, 52.18 = 1.97$$

Suppose we want to conduct the test at the $\alpha = 0.05$ level, i.e., we can tolerate a 5% chance of the Rosner test falsely declaring four outliers. In Table B-16, $\lambda_4 = 2.89$ when $n = 32$, $r = 4$ and $\alpha = 0.05$. As $R_4 = 1.97$ is less than 2.89, it is concluded that four outliers are not present. Therefore, test if three outliers are present. Compute

$$R_3 = |y^{(2)} - \bar{x}^{(2)}| \,/\, s^{(2)} = |40.55 - 167.99| \,/\, 56.49 = 2.26$$

In Table B-16 $\lambda_3 = 2.91$ when $n = 32$, $r = 3$ and $\alpha = 0.05$. Because $R_4 = 2.26$ is less than 2.91, it is concluded that three outliers are not present. Therefore, test if two outliers are present. Compute

$$R_2 = |y^{(1)} - \bar{x}^{(1)}| \,/\, s^{(1)} = |2.07 - 162.64| \,/\, 62.83 = 2.56$$

In Table B-16, $\lambda_2 = 2.92$ for $n = 32$, $r = 2$ and $\alpha = 0.05$. As $R_2 = 2.56$ is less than 2.92, it is concluded that two outliers are not present in the dataset. Therefore, test if one outlier is present. Compute

$$R_1 = |y^{(0)} - \bar{x}^{(0)}| \,/\, s^{(0)} = |395.67 - 169.92| \,/\, 73.95 = 3.05$$

In Table B-16 $\lambda_1 = 2.94$ for $n = 32$, $r = 1$ and $\alpha = 0.05$. Because $R_1 = 3.05$ is greater than 2.94, then it is concluded at the $\alpha = 0.05$ significance level that one outlier is present in the dataset. Therefore, the measurement 395.67 is considered to be a statistical outlier. It will be further investigated to determine if the measurement is an error or a valid data value.

APPENDIX B: COMPARATIVE METHOD REFERENCE TABLES

TABLE B-1. Cumulative standard normal distribution (values of the probability ϕ corresponding to the value z_ϕ of a standard normal random variable)

z_ϕ	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5674	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

TABLE B-2. Critical values (K_c) for the Slippage test for $\alpha = 0.01$

Number of Background Measurements, m	Number of Site Measurements, n																								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
2	/	/	/	/	/	/	/	/	/	/	/	/	13	14	15	16	17	18	19	20	21	22	23	23	24
3	/	/	/	/	/	/	7	8	9	10	11	11	12	13	14	15	15	16	17	18	18	19	20	21	22
4	/	/	/	/	5	6	7	8	8	9	10	10	11	12	12	13	14	14	15	16	16	17	18	19	19
5	/	/	/	4	5	6	6	7	8	8	9	9	10	11	11	12	12	13	14	14	15	15	16	17	17
6	/	/	/	4	5	5	6	6	7	8	8	9	9	10	10	11	11	12	12	13	14	14	15	15	16
7	/	/	3	4	5	5	6	6	7	7	8	8	9	9	10	10	10	11	11	12	12	13	13	14	14
8	/	/	3	4	4	5	5	6	6	7	7	8	8	8	9	9	10	10	11	11	12	12	12	13	13
9	/	/	3	4	4	5	5	5	6	6	7	7	8	8	8	9	9	10	10	10	11	11	12	12	12
10	/	/	3	4	4	4	5	5	6	6	6	7	7	7	8	8	9	9	9	10	10	11	11	11	12
11	/	/	3	3	4	4	5	5	5	6	6	6	7	7	7	8	8	9	9	9	10	10	10	11	11
12	/	/	3	3	4	4	4	5	5	5	6	6	6	7	7	7	8	8	8	9	9	9	10	10	10
13	/	2	3	3	4	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8	9	9	9	10	10
14	/	2	3	3	4	4	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8	9	9	9	9
15	/	2	3	3	3	4	4	4	5	5	5	5	6	6	6	7	7	7	7	8	8	8	9	9	9
16	/	2	3	3	3	4	4	4	4	5	5	5	6	6	6	6	7	7	7	7	8	8	8	8	9
17	/	2	3	3	3	4	4	4	4	5	6	6	6	6	6	6	6	7	7	7	7	8	8	8	8
18	/	2	3	3	3	3	4	4	4	5	5	5	5	5	6	6	6	6	7	7	7	7	8	8	8
19	/	2	3	3	3	3	4	4	4	4	5	5	5	6	6	6	6	6	6	7	7	7	7	8	8
20	/	2	3	3	3	3	4	4	4	4	5	5	5	5	5	6	6	6	6	7	7	7	7	7	8
21	/	2	3	3	3	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	7	7	7	7	7
22	/	2	3	3	3	3	3	4	4	4	4	5	5	5	5	5	6	6	6	6	6	7	7	7	7
23	/	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	6	6	6	6	6	7	7	7
24	/	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	6	6	6	6	6	6	7	7
25	/	2	2	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	6	6	6	6	6	6	7
26	/	2	2	3	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6	6
27	/	2	2	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	6	6	6	6	6	6
28	/	2	2	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6
29	/	2	2	3	3	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	5	6	6	6	6
30	/	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6
31	/	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	5	6	6	6
32	/	2	2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5	5	6	6
33	/	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	6	6	6
34	/	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5	6	6
35	/	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5
36	/	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5
37	/	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5	5
38	/	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5	5	5
39	/	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5	5	5
40	/	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5	5
41	/	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5	5
42	/	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5	5
43	/	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	5	5	5	5
44	/	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5
45	/	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	5	5	5
46	/	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	5	5	5
47	/	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	5	5
48	/	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	5	5
49	/	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	5
50	/	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	5

TABLE B-2. Critical values (K_c) for the Slippage test for $\alpha = 0.01$ (continued)

	Number of Site Measurements, n																								
Number of Background Measurements, m	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
2	25	26	27	28	29	30	31	32	32	33	34	35	36	37	38	39	40	41	41	42	43	44	45	46	47
3	22	23	24	25	26	26	27	28	29	30	30	31	32	33	33	34	35	36	37	37	38	39	40	40	41
4	20	21	21	22	23	23	24	25	25	26	27	27	28	29	30	30	31	32	32	33	34	34	35	35	36
5	18	18	19	20	20	21	22	22	23	23	24	25	25	26	26	27	28	28	29	29	30	31	31	32	32
6	16	17	17	18	18	19	19	20	21	21	22	22	23	23	24	24	25	25	26	26	27	28	28	29	29
7	15	15	16	16	17	17	18	18	19	19	20	20	21	21	22	22	23	23	24	24	25	25	26	26	26
8	14	14	15	15	16	16	16	17	17	18	18	19	19	19	20	20	21	21	22	22	23	23	23	24	24
9	13	13	14	14	14	15	15	16	16	16	17	17	18	18	18	19	19	20	20	20	21	21	22	22	22
10	12	12	13	13	13	14	14	15	15	15	16	16	16	17	17	18	18	18	19	19	19	20	20	21	21
11	11	12	12	12	13	13	13	14	14	14	15	15	15	16	16	16	17	17	18	18	18	19	19	19	20
12	11	11	11	12	12	12	13	13	13	14	14	14	15	15	15	16	16	16	17	17	17	17	18	18	18
13	10	10	11	11	11	12	12	12	13	13	13	14	14	14	14	15	15	15	16	16	16	17	17	17	17
14	10	10	10	11	11	11	11	12	12	12	13	13	13	13	14	14	14	15	15	15	15	16	16	16	17
15	9	10	10	10	10	11	11	11	11	12	12	12	13	13	13	13	14	14	14	14	15	15	15	15	16
16	9	9	9	10	10	10	10	11	11	11	11	12	12	12	13	13	13	13	14	14	14	14	15	15	15
17	9	9	9	9	10	10	10	10	11	11	11	11	12	12	12	12	12	13	13	13	13	14	14	14	14
18	8	9	9	9	9	9	10	10	10	10	11	11	11	11	12	12	12	12	12	13	13	13	13	14	14
19	8	8	8	9	9	9	9	10	10	10	10	10	11	11	11	11	12	12	12	12	13	13	13	13	13
20	8	8	8	8	9	9	9	9	9	10	10	10	10	11	11	11	11	11	12	12	12	12	12	13	13
21	8	8	8	8	8	9	9	9	9	9	10	10	10	10	10	11	11	11	11	11	12	12	12	12	12
22	7	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10	10	11	11	11	11	11	12	12	12
23	7	7	8	8	8	8	8	8	9	9	9	9	9	10	10	10	10	10	11	11	11	11	11	11	12
24	7	7	7	8	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10	10	11	11	11	11	11
25	7	7	7	7	8	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10	10	10	11	11	11
26	7	7	7	7	7	7	8	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10	10	10	11
27	7	7	7	7	7	7	7	8	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10	10	10
28	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10
29	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	9	9	9	9	9	9	9	10	10	10
30	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9	9	9	10
31	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9	9	9
32	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9	9
33	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9
34	6	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	9	9	9
35	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	9	9
36	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	9
37	5	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8
38	5	5	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8
39	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8	8
40	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	8	8	8	8
41	5	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	8	8	8
42	5	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7	8	8
43	5	5	5	5	5	5	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7	8
44	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7
45	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7
46	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7
47	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7
48	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7
49	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7
50	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7

TABLE B-3. Critical values (K_c) for the Slippage test for $\alpha = 0.05$

Number of Site Measurements, n

Number of Background Measurements, m

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	20	21	22	23	24	25
2	/	/	/	/	5	6	7	8	9	9	10	11	12	13	13	14	15	16	16	17	18	19	20	20	21
3	/	/	/	4	5	5	6	7	7	8	9	9	10	11	11	12	12	13	14	14	15	16	16	17	18
4	/	/	3	4	4	5	5	6	6	7	8	8	9	9	10	10	11	11	12	12	13	13	14	14	15
5	/	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	9	10	10	11	11	12	12	13	13
6	/	2	3	3	4	4	4	5	5	6	6	6	7	7	8	8	8	9	9	10	10	10	11	11	12
7	/	2	3	3	3	4	4	5	5	5	6	6	6	7	7	7	8	8	8	9	9	9	10	10	11
8	/	2	3	3	3	4	4	4	5	5	5	6	6	6	6	7	7	7	8	8	8	9	9	9	10
9	/	2	2	3	3	3	4	4	4	5	5	5	5	6	6	6	7	7	7	7	8	8	8	9	9
10	/	2	3	3	3	3	4	4	4	4	5	5	5	5	6	6	6	6	7	7	7	8	8	8	8
11	/	2	2	3	3	3	3	4	4	4	4	5	5	5	5	6	6	6	6	7	7	7	7	8	8
12	/	2	2	3	3	3	3	3	4	4	4	4	5	5	5	5	6	6	6	6	6	7	7	7	7
13	/	2	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	7	7	7
14	/	2	2	2	3	3	3	3	3	4	4	4	4	4	5	5	5	5	5	6	6	6	6	6	7
15	/	2	2	2	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	6	6	6	6	6
16	/	2	2	2	2	3	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6
17	/	2	2	2	2	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	6	6	6
18	/	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	5	6
19	/	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	5
20	1	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5	5
21	1	2	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5
22	1	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5	5
23	1	2	2	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5
24	1	2	2	2	2	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5
25	1	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5
26	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5
27	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4
28	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4
29	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4
30	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4
31	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4
32	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4
33	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4
34	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4
35	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4
36	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4
37	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4
38	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4
39	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	4	4
40	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	4
41	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	4
42	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3
43	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3
44	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3
45	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3
46	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3
47	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3
48	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3
49	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3
50	1	1	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3

TABLE B-3. Critical values (K_c) for the Slippage test for $\alpha = 0.05$ (continued)

Number of Site Measurements, n

Number of Background Measurements, m

SITE	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1	26	27	28	29	30	31	32	33	34	35	36	37	38	39	39	40	41	42	43	44	45	46	47	48	49
2	22	23	23	24	25	26	26	27	28	29	30	30	31	32	33	33	34	35	36	37	37	38	39	40	40
3	18	19	19	20	21	21	22	23	23	24	24	25	26	26	27	28	28	29	30	30	31	31	32	33	33
4	15	16	17	17	18	18	19	19	20	20	21	21	22	22	23	23	24	24	25	26	26	27	27	28	28
5	14	14	14	15	15	16	16	17	17	18	18	18	19	19	20	20	21	21	22	22	23	23	23	24	24
6	12	12	13	13	14	14	14	15	15	16	16	16	17	17	18	18	18	19	19	20	20	20	21	21	21
7	11	11	12	12	12	13	13	13	14	14	14	15	15	15	16	16	16	17	17	18	18	18	19	19	19
8	10	10	11	11	11	12	12	12	13	13	13	13	14	14	14	15	15	15	16	16	16	17	17	17	17
9	9	9	10	10	10	11	11	11	11	12	12	12	13	13	13	13	14	14	14	15	15	16	16	16	16
10	9	9	9	9	10	10	10	10	11	11	11	11	12	12	12	12	13	13	13	14	14	14	14	15	15
11	8	8	9	9	9	9	9	10	10	10	10	11	11	11	11	12	12	12	12	13	13	13	13	14	14
12	8	8	8	8	8	9	9	9	9	10	10	10	10	10	11	11	11	11	12	12	12	12	12	13	13
13	7	7	8	8	8	8	8	9	9	9	9	9	10	10	10	10	11	11	11	11	11	12	12	12	12
14	7	7	7	7	8	8	8	8	8	9	9	9	9	9	10	10	10	10	10	11	11	11	11	11	12
15	7	7	7	7	7	7	8	8	8	8	9	9	9	9	9	9	9	10	10	10	10	10	11	11	11
16	6	6	7	7	7	7	7	7	8	8	8	8	8	9	9	9	9	9	9	10	10	10	10	10	10
17	6	6	6	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9	9	10	10	10
18	6	6	6	6	6	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9	9	10
19	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	9
20	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9	9	9
21	5	5	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	9
22	5	5	5	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8	8
23	5	5	5	5	5	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8	8	8	8
24	5	5	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	8	8
25	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	8
26	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7
27	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7
28	4	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7
29	4	4	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	7	7	7
30	4	4	4	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	7	7
31	4	4	4	4	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	6
32	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6
33	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6
34	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6
35	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6
36	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6
37	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	6	6	6
38	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	6	6
39	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	6
40	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5
41	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5
42	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5
43	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5
44	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5
45	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5
46	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5
47	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5
48	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5
49	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5
50	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5

TABLE B-4. Values of *r, k,* and α for the Quantile test when α is approximately equal to 0.01

Number of Site Measurements, n

Number of Reference (Background) Measurements, m

	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
5	***r, k*** α		11,1 0.008	13,13 0.015	16,16 0.014	19,19 0.013	22,22 0.013	25,25 0.013	28,28 0.012											
10		6,6 0.005	7,7 0.013	9,9 0.012	11,11 0.011	13,13 0.010	14,14 0.014	16,16 0.013	18,18 0.012	19,19 0.015	21,21 0.014	23,23 0.013	25,25 0.012	26,26 0.015	28,28 0.014	30,30 0.013				
15	3.3 0.009	7,6 0.007	6,6 0.008	7,7 0.012	8,8 0.014	10,10 0.009	11,11 0.011	12,12 0.013	13,13 0.014	15,15 0.011	16,16 0.012	17,17 0.013	18,18 0.014	19,19 0.015	21,21 0.012	22,22 0.013	23,23 0.014	24,24 0.015	26,26 0.013	27,27 0.013
20	6,4 0.005	4,4 0.008	5,5 0.009	6,6 0.010	7,7 0.011	8,8 0.011	9,9 0.011	10,10 0.011	11,11 0.011	12,12 0.011	13,13 0.011	14,14 0.012	15,15 0.012	16,16 0.012	17,17 0.012	18,18 0.012	19,19 0.012	19,19 0.015	20,20 0.015	21,21 0.015
25	4,3 0.009	7,5 0.012	4,4 0.015	5,5 0.013	6,6 0.011	7,7 0.010	8,8 0.009	9,9 0.009	9,9 0.014	10,10 0.012	11,11 0.011	12,12 0.011	12,12 0.015	13,13 0.014	14,14 0.013	15,15 0.012	16,16 0.011	16,16 0.014	17,17 0.014	18,18 0.013
30	4,3 0.006	3,3 0.012	4,4 0.009	5,5 0.007	6,6 0.006	6,6 0.012	7,7 0.010	8,8 0.008	8,8 0.013	9,9 0.011	10,10 0.009	10,10 0.013	11,11 0.011	12,12 0.010	12,12 0.013	13,13 0.012	14,14 0.011	14,14 0.014	15,15 0.012	15,15 0.015
35	2,2 0.013	3,3 0.008	4,4 0.006	4,4 0.014	5,5 0.010	6,6 0.007	6,6 0.012	7,7 0.009	7,7 0.014	8,8 0.011	9,9 0.009	9,9 0.013	10,10 0.010	10,10 0.014	11,11 0.011	11,11 0.015	12,12 0.012	13,13 0.011	13,13 0.013	14,14 0.012
40	2,2 0.008	3,3 0.008	7,5 0.013	4,4 0.007	5,5 0.006	5,5 0.012	6,6 0.008	6,6 0.013	7,7 0.009	7,7 0.013	8,8 0.010	8,8 0.014	9,9 0.011	9,9 0.014	10,10 0.011	10,10 0.014	11,11 0.012	11,11 0.014	12,12 0.012	12,12 0.014
45	2,2 0.008	6,4 0.008	3,3 0.013	4,4 0.007	4,4 0.014	5,5 0.008	5,5 0.014	6,6 0.009	6,6 0.013	7,7 0.009	7,7 0.013	8,8 0.009	8,8 0.012	9,9 0.009	9,9 0.012	10,10 0.009	10,10 0.012	10,10 0.015	11,11 0.012	11,11 0.014
50		4,3 0.013	3,3 0.010	4,4 0.005	4,4 0.010	5,5 0.006	5,5 0.010	5,5 0.015	6,6 0.009	6,6 0.013	7,7 0.009	7,7 0.012	8,8 0.009	8,8 0.011	8,8 0.014	9,9 0.011	9,9 0.013	10,10 0.010	10,10 0.012	10,10 0.015
55		4,3 0.010	3,3 0.008	7,5 0.013	4,4 0.008	4,4 0.014	5,5 0.007	5,5 0.011	6,6 0.007	6,6 0.010	6,6 0.014	7,7 0.009	7,7 0.012	8,8 0.008	8,8 0.010	8,8 0.013	9,9 0.009	9,9 0.012	9,9 0.014	10,10 0.011
60		4,3 0.008	3,3 0.007	3,3 0.014	4,4 0.006	4,4 0.011	5,5 0.006	5,5 0.009	5,5 0.013	6,6 0.007	6,6 0.010	6,6 0.014	7,7 0.009	7,7 0.011	7,7 0.014	8,8 0.010	8,8 0.012	8,8 0.015	9,9 0.010	9,9 0.013
65		4,3 0.007	3,3 0.006	3,3 0.012	6,5 0.006	4,4 0.009	4,4 0.013	5,5 0.007	5,5 0.010	5,5 0.014	6,6 0.008	6,6 0.011	6,6 0.014	7,7 0.009	7,7 0.011	7,7 0.014	8,8 0.009	8,8 0.011	8,8 0.014	9,9 0.010
70		2,2 0.014	6,4 0.008	3,3 0.010	7,5 0.013	4,4 0.007	4,4 0.011	5,5 0.005	5,5 0.008	5,5 0.011	5,5 0.015	6,6 0.008	6,6 0.011	6,6 0.014	7,7 0.009	7,7 0.011	7,7 0.013	8,8 0.009	8,8 0.011	8,8 0.013
75		2,2 0.013	4,3 0.014	3,3 0.008	3,3 0.014	4,4 0.006	4,4 0.009	4,4 0.013	5,5 0.006	5,5 0.009	5,5 0.012	6,6 0.007	6,6 0.009	6,6 0.011	6,6 0.014	7,7 0.009	7,7 0.011	7,7 0.013	8,8 0.008	8,8 0.010
80		2,2 0.011	4,3 0.012	3,3 0.007	3,3 0.012	6,5 0.006	4,4 0.008	4,4 0.011	5,5 0.005	5,5 0.007	5,5 0.010	5,5 0.013	6,6 0.007	6,6 0.009	6,6 0.012	6,6 0.014	7,7 0.008	7,7 0.010	7,7 0.013	7,7 0.015
85		2,2 0.010	4,3 0.010	3,3 0.006	3,3 0.011	7,5 0.013	4,4 0.006	4,4 0.009	4,4 0.013	5,5 0.006	5,5 0.008	5,5 0.011	5,5 0.014	6,6 0.008	6,6 0.010	6,6 0.012	6,6 0.014	7,7 0.008	7,7 0.010	7,7 0.012
90			4,3 0.009	3,3 0.005	3,3 0.009	3,3 0.014	4,4 0.005	4,4 0.008	4,4 0.011	5,5 0.005	5,5 0.007	5,5 0.009	5,5 0.012	5,5 0.015	6,6 0.008	6,6 0.010	6,6 0.012	6,6 0.014	7,7 0.008	7,7 0.010
95			4,3 0.008	6.4 0.008	3.3 0.008	3.3 0.013	6.5 0.005	4.4 0.007	4.4 0.010	4.4 0.013	5.5 0.006	5.5 0.008	5.5 0.010	5.5 0.013	6.6 0.007	6.6 0.008	6.6 0.010	6.6 0.012	6.6 0.014	7.7 0.008
100			4,3 0.007	4,3 0.014	3,3 0.007	3,3 0.011	7,5 0.013	4,4 0.006	4,4 0.008	4,4 0.011	4,4 0.015	5,5 0.007	5,5 0.009	5,5 0.011	5,5 0.013	6,6 0.007	6,6 0.009	6,6 0.010	6,6 0.012	6,6 0.014

TABLE B-5. Values of *r, k,* and α for the Quantile test when α is approximately equal to 0.025

Number of Site Measurements, n

Number of Reference (Background) Measurements, m

	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
5	***r, k*** α		9,9 0.030	12,12 0.024	15,15 0.021	17,17 0.026	20,20 0.024	22,22 0.028	25,25 0.025											
10		7,6 0.029	6,6 0.028	8,8 0.022	9,9 0.029	11,11 0.024	12,12 0.029	14,14 0.025	15,15 0.029	17,17 0.025	18,18 0.029	20,20 0.026	21,21 0.029	23,23 0.026	24,24 0.029	26,26 0.026	27,27 0.029			
15	11,5 0.030	6,5 0.023	5,5 0.021	6,6 0.024	7,7 0.026	8,8 0.027	9,9 0.028	10,10 0.029	11,11 0.030	13,13 0.022	14,14 0.023	15,15 0.023	16,16 0.024	17,17 0.025	18,18 0.025	19,19 0.026	21,21 0.021	21,21 0.027	22,22 0.027	23,23 0.027
20	8,4 0.023	3,3 0.030	4,4 0.026	5,5 0.024	6,6 0.022	7,7 0.020	12,11 0.021	13,12 0.024	9,9 0.028	10,10 0.026	11,11 0.024	12,12 0.023	13,13 0.022	13,13 0.029	14,14 0.027	15,15 0.026	16,16 0.025	17,17 0.024	17,17 0.029	18,18 0.028
25	2,2 0.023	8,5 0.027	6,5 0.021	7,6 0.023	5,5 0.025	6,6 0.020	10,9 0.026	7,7 0.027	8,8 0.023	13,12 0.027	9,9 0.027	10,10 0.024	11,11 0.022	11,11 0.028	12,12 0.025	13,13 0.023	13,13 0.028	14,14 0.025	15,15 0.023	15,15 0.028
30	6,3 0.026	6,4 0.026	9,6 0.026	4,4 0.021	7,6 0.029	5,5 0.026	9,8 0.024	6,6 0.029	7,7 0.023	12,11 0.021	8,8 0.025	9,9 0.021	9,9 0.027	10,10 0.023	10,10 0.029	11,11 0.025	11,11 0.030	12,12 0.026	13,13 0.023	13,13 0.027
35	7,3 0.030	4,3 0.030	3,3 0.023	6,5 0.020	4,4 0.026	10,8 0.022	5,5 0.027	9,8 0.024	6,6 0.027	7,7 0.020	7,7 0.027	8,8 0.021	8,8 0.027	9,9 0.022	9,9 0.027	10,10 0.022	10,10 0.027	11,11 0.022	11,11 0.027	12,12 0.023
40	3,2 0.029	4,3 0.022	8,5 0.028	11,7 0.025	6,5 0.028	4,4 0.030	10,8 0.026	5,5 0.027	9,8 0.023	6,6 0.026	10,9 0.028	7,7 0.024	12,11 0.020	8,8 0.023	8,8 0.029	9,9 0.022	9,9 0.027	10,10 0.021	10,10 0.026	11,11 0.021
45	3,2 0.023	8,4 0.029	6,4 0.030	3,3 0.026	8,6 0.021	4,4 0.023	7,6 0.025	5,5 0.020	5,5 0.028	9,8 0.023	6,6 0.024	10,9 0.026	7,7 0.022	7,7 0.027	8,8 0.020	8,8 0.025	8,8 0.030	9,9 0.023	9,9 0.027	10,10 0.021
50		2,2 0.025	6,4 0.022	3,3 0.021	11,7 0.027	6,5 0.026	4,4 0.026	7,6 0.028	5,5 0.021	5,5 0.028	9,8 0.022	6,6 0.023	6,6 0.029	7,7 0.020	7,7 0.025	12,11 0.020	8,8 0.022	8,8 0.026	13,12 0.027	9,9 0.023
55		2,2 0.022	4,3 0.029	8,5 0.028	3,3 0.028	8,6 0.021	4,4 0.020	4,4 0.029	10,8 0.021	5,5 0.022	5,5 0.028	9,8 0.022	6,6 0.023	6,6 0.028	10,9 0.029	7,7 0.023	7,7 0.027	12,11 0.023	8,8 0.023	8,8 0.027
60		14,5 0.022	4,3 0.024	8,5 0.021	3,3 0.023	11,7 0.029	6,5 0.024	4,4 0.023	7,6 0.023	10,8 0.024	5,5 0.023	5,5 0.029	9,8 0.022	6,6 0.022	6,6 0.027	10,9 0.027	7,7 0.021	7,7 0.025	7,7 0.030	8,8 0.021
65		6,3 0.028	7,4 0.021	6,4 0.025	10,6 0.025	3,3 0.029	8,6 0.021	6,5 0.029	4,4 0.026	7,6 0.026	10,8 0.026	5,5 0.023	5,5 0.029	9,8 0.022	6,6 0.021	6,6 0.026	10,9 0.026	7,7 0.020	7,7 0.024	7,7 0.028
70		6,3 0.024	2,2 0.029	6,4 0.021	8,5 0.028	3,3 0.025	13,8 0.026	6,5 0.023	4,4 0.022	4,4 0.028	7,6 0.028	10,8 0.027	5,5 0.024	5,5 0.029	9,8 0.022	6,6 0.021	6,6 0.025	6,6 0.029	10,9 0.030	7,7 0.022
75		11,4 0.022	2,2 0.026	4,3 0.028	8,5 0.022	3,3 0.022	9,6 0.028	8,6 0.021	6,5 0.027	4,4 0.024	7,6 0.023	7,6 0.030	10,8 0.029	5,5 0.024	5,5 0.029	9,8 0.021	6,6 0.021	6,6 0.024	6,6 0.028	10,9 0.028
80		7,3 0.028	2,2 0.024	4,3 0.024	6,4 0.028	10,6 0.024	3,3 0.027	13,8 0.027	6,5 0.023	4,4 0.020	4,4 0.026	7,6 0.024	10,8 0.023	5,5 0.020	5,5 0.025	5,5 0.029	9,8 0.021	6,6 0.020	6,6 0.024	6,6 0.027
85		3,2 0.029	2,2 0.021	4,3 0.021	6,4 0.023	8,5 0.028	3,3 0.023	9,6 0.030	8,6 0.020	6,5 0.026	4,4 0.022	4,4 0.028	7,6 0.026	10,8 0.024	5,5 0.021	5,5 0.025	5,5 0.029	9,8 0.021	6,6 0.020	6,6 0.023
90			5,3 0.020	11,5 0.027	9,5 0.023	8,5 0.023	3,3 0.021	3,3 0.028	13,8 0.028	6,5 0.022	6,5 0.029	4,4 0.024	4,4 0.029	7,6 0.028	10,8 0.026	5,5 0.022	5,5 0.025	5,5 0.030	9,8 0.021	9,8 0.025
95			10,4 0.029	2,2 0.029	4,3 0.028	6,4 0.029	10,6 0.023	3,3 0.025	11,7 0.026	8,6 0.020	6,5 0.025	4,4 0.021	4,4 0.026	7,6 0.024	7,6 0.029	10,8 0.027	5,5 0.022	5,5 0.026	5,5 0.030	9,8 0.021
100			6,3 0.029	2,2 0.027	4,3 0.025	6,4 0.025	8,5 0.028	3,3 0.022	3,3 0.029	13,8 0.028	6,5 0.022	6,5 0.028	4,4 0.023	4,4 0.027	7,6 0.025	10,8 0.022	10,8 0.028	5,5 0.022	5,5 0.026	5,5 0.030

TABLE B-6. Values of *r, k,* and α for the Quantile test when α is approximately equal to 0.05

Number of Reference (Background) Measurements, m	Number of Site Measurements, n																			
	5	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**	**100**
5	***r, k*** α		8,8 0.051	10,10 0.057	13,13 0.043	15,15 0.048	17,17 0.051	19,19 0.054	21,21 0.056											
10		4,4 0.043	5,5 0.057	14,12 0.045	8,8 0.046	9,9 0.052	10,10 0.058	12,12 0.046	13,13 0.050	14,14 0.054	15,15 0.057	17,17 0.049	18,18 0.052	19,19 0.055	20,20 0.057	21,21 0.059	23,23 0.053			
15	2,2 0.053	3,3 0.052	4,4 0.050	5,5 0.048	6,6 0.046	7,7 0.045	8,8 0.044	9,9 0.043	9,9 0.060	10,10 0.057	11,11 0.055	12,12 0.054	13,13 0.052	14,14 0.051	15,15 0.050	16,16 0.049	16,16 0.058	17,17 0.057	18,18 0.056	19,19 0.055
20	9,4 0.040	8,5 0.056	6,5 0.040	4,4 0.053	5,5 0.043	9,8 0.052	6,6 0.056	7,7 0.048	8,8 0.043	8,8 0.057	9,9 0.051	10,10 0.046	10,10 0.057	11,11 0.052	12,12 0.048	12,12 0.057	13,13 0.053	14,14 0.049	14,14 0.057	15,15 0.054
25	6,3 0.041	6,4 0.043	3,3 0.046	6,5 0.052	4,4 0.055	5,5 0.041	5,5 0.059	6,6 0.046	11,10 0.042	7,7 0.050	8,8 0.042	8,8 0.053	9,9 0.045	9,9 0.055	10,10 0.048	11,11 0.042	11,11 0.050	11,11 0.058	12,12 0.052	12,12 0.060
30	3,2 0.047	2,2 0.058	10,6 0.052	3,3 0.058	11,8 0.045	4,4 0.056	8,7 0.045	5,5 0.054	6,6 0.040	6,6 0.053	7,7 0.041	7,7 0.052	8,8 0.042	8,8 0.051	9,9 0.042	9,9 0.050	9,9 0.059	10,10 0.049	10,10 0.057	11,11 0.049
35	8,3 0.046	2,2 0.045	6,4 0.058	3,3 0.043	6,5 0.041	4,4 0.040	4,4 0.057	8,7 0.043	5,5 0.051	9,8 0.052	6,6 0.047	6,6 0.058	7,7 0.043	7,7 0.053	8,8 0.041	8,8 0.049	8,8 0.057	9,9 0.046	9,9 0.053	10,10 0.044
40	4,2 0.055	5,3 0.048	4,3 0.057	10,6 0.059	3,3 0.053	6,5 0.048	4,4 0.043	4,4 0.058	8,7 0.042	5,5 0.048	9,8 0.047	6,6 0.042	6,6 0.051	11,10 0.042	7,7 0.045	7,7 0.053	8,8 0.041	8,8 0.048	8,8 0.055	9,9 0.043
45	4,2 0.045	9,4 0.047	2,2 0.059	8,5 0.052	3,3 0.042	8,6 0.041	6,5 0.054	4,4 0.045	4,4 0.058	8,7 0.041	5,5 0.046	5,5 0.057	9,8 0.056	6,6 0.047	6,6 0.055	11,10 0.046	7,7 0.047	7,7 0.054	8,8 0.041	8,8 0.047
50		6,3 0.052	2,2 0.050	6,4 0.051	12,7 0.050	3,3 0.049	8,6 0.049	6,5 0.059	4,4 0.047	4,4 0.059	8,7 0.041	5,5 0.045	5,5 0.054	9,8 0.051	6,6 0.043	6,6 0.050	6,6 0.058	7,7 0.042	7,7 0.048	7,7 0.054
55		3,2 0.059	2,2 0.043	4,3 0.056	8,5 0.058	3,3 0.041	5,4 0.041	6,5 0.046	9,7 0.042	4,4 0.048	4,4 0.059	8,7 0.040	5,5 0.043	5,5 0.052	9,8 0.048	6,6 0.040	6,6 0.047	6,6 0.054	11,10 0.043	7,7 0.043
60		3,2 0.052	5,3 0.052	4,3 0.046	6,4 0.059	3,3 0.035	3,3 0.047	8,6 0.043	6,5 0.051	9,7 0.046	4,4 0.049	4,4 0.059	13,10 0.052	5,5 0.042	5,5 0.050	5,5 0.058	9,8 0.054	6,6 0.044	6,6 0.050	6,6 0.056
65		3,2 0.045	5,3 0.043	2,2 0.053	6,4 0.048	10,6 0.050	3,3 0.040	3,3 0.053	6,5 0.041	6,5 0.055	4,4 0.042	4,4 0.050	4,4 0.060	13,10 0.052	5,5 0.041	5,5 0.048	5,5 0.055	9,8 0.051	6,6 0.041	6,6 0.047
70		8,3 0.057	9,4 0.048	2,2 0.047	4,3 0.055	8,5 0.050	5,4 0.041	3,3 0.046	3,3 0.057	6,5 0.045	6,5 0.058	4,4 0.043	4,4 0.051	4,4 0.060	13,10 0.051	5,5 0.041	5,5 0.047	5,5 0.054	9,8 0.048	9,8 0.057
75		8,3 0.049	6,3 0.056	2,2 0.043	4,3 0.047	6,4 0.054	10,6 0.053	3,3 0.040	3,3 0.051	8,6 0.044	6,5 0.049	9,7 0.041	4,4 0.044	4,4 0.052	5,5 0.060	13,10 0.051	8,7 0.047	5,5 0.046	5,5 0.052	5,5 0.058
80		4,2 0.059	6,3 0.048	5,3 0.053	2,2 0.055	6,4 0.046	8,5 0.055	5,4 0.042	3,3 0.045	3,3 0.055	6,5 0.041	6,5 0.052	9,7 0.043	4,4 0.045	4,4 0.053	7,6 0.058	13,10 0.051	8,7 0.046	5,5 0.045	5,5 0.051
85		4,2 0.054	3,2 0.058	5,3 0.047	2,2 0.050	4,3 0.054	4,3 0.048	10,6 0.056	5,4 0.049	3,3 0.049	3,3 0.059	6,5 0.044	6,5 0.055	9,7 0.046	4,4 0.046	4,4 0.053	7,6 0.059	10,8 0.060	8,7 0.045	5,5 0.044
90			3,2 0.053	5,3 0.041	2,2 0.046	6,4 0.059	6,4 0.051	8,5 0.058	5,4 0.042	3,3 0.044	3,3 0.053	8,6 0.045	6,5 0.047	6,5 0.058	4,4 0.041	4,4 0.047	4,4 0.054	7,6 0.059	10,8 0.060	8,7 0.045
95			3,2 0.048	9,4 0.048	2,2 0.042	2,2 0.056	4,3 0.059	8,5 0.050	10,6 0.058	5,4 0.048	3,3 0.048	3,3 0.056	6,5 0.041	6,5 0.050	9,7 0.040	4,4 0.042	4,4 0.048	4,4 0.054	7,6 0.059	10,8 0.059
100			3,2 0.044	6,3 0.057	5,3 0.054	2,2 0.052	4,3 0.053	6,4 0.056	10,6 0.049	5,4 0.043	3,3 0.043	3,3 0.051	3,3 0.059	6,5 0.044	6,5 0.053	9,7 0.042	4,4 0.043	4,4 0.049	4,4 0.055	7,6 0.059

TABLE B-7. Values of *r, k,* and α for the Quantile test when α is approximately equal to 0.10

Number of Site Measurements

Number of Reference (Background) Measurements, m

	5	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**	**100**
5	***r, k*** α		7,7 0.083	8,8 0.116	10,10 0.109	12,12 0.104	14,14 0.100	15,15 0.117	17,17 0.112											
10		3,3 0.105	4,4 0.108	5,5 0.109	6,6 0.109	7,7 0.109	8,8 0.109	9,9 0.109	10,10 0.109	11,11 0.109	12,12 0.109	13,13 0.109	14,14 0.109	15,15 0.109	16,16 0.109	17,17 0.109	18,18 0.109			
15	9,4 0.098	10,6 0.106	3,3 0.112	4,4 0.093	5,5 0.081	5,5 0.117	6,6 0.102	7,7 0.092	7,7 0.118	8,8 0.106	9,9 0.098	9,9 0.118	10,10 0.109	11,11 0.101	11,11 0.118	12,12 0.110	13,13 0.104	13,13 0.118	14,14 0.111	15,15 0.106
20	3,2 0.091	2,2 0.103	5,4 0.093	3,3 0.115	4,4 0.085	4,4 0.119	5,5 0.093	10,9 0.084	6,6 0.099	7,7 0.083	7,7 0.102	8,8 0.088	8,8 0.105	9,9 0.092	9,9 0.107	10,10 0.095	10,10 0.108	11,11 0.098	11,11 0.110	12,12 0.100
25	4,2 0.119	7,4 0.084	8,5 0.112	3,3 0.080	3,3 0.117	4,4 0.080	4,4 0.107	8,7 0.108	5,5 0.101	10,9 0.088	6,6 0.096	6,6 0.114	7,7 0.093	7,7 0.108	8,8 0.091	8,8 0.104	8,8 0.117	9,9 0.100	9,9 0.112	10,10 0.098
30	4,2 0.089	5,3 0.089	2,2 0.106	14,8 0.111	3,3 0.088	3,3 0.119	9,7 0.116	4,4 0.100	8,7 0.093	5,5 0.088	5,5 0.106	6,6 0.080	6,6 0.095	6,6 0.110	7,7 0.087	7,7 0.100	7,7 0.113	8,8 0.092	8,8 0.103	8,8 0.115
35	5,2 0.109	3,2 0.119	2,2 0.086	6,4 0.119	5,4 0.091	3,3 0.093	3,3 0.120	9,7 0.112	4,4 0.094	4,4 0.114	8,7 0.107	5,5 0.094	5,5 0.110	6,6 0.081	6,6 0.094	6,6 0.107	6,6 0.120	7,7 0.094	7,7 0.105	7,7 0.116
40	5,2 0.087	3,2 0.098	5,3 0.119	2,2 0.107	12,7 0.109	5,4 0.102	3,3 0.097	6,5 0.100	9,7 0.109	4,4 0.090	4,4 0.107	8,7 0.097	5,5 0.086	5,5 0.099	5,5 0.112	6,6 0.082	6,6 0.093	6,6 0.104	6,6 0.116	7,7 0.089
45	6,2 0.103	3,2 0.082	5,3 0.094	2,2 0.091	6,4 0.115	7,5 0.086	5,4 0.112	3,3 0.100	6,5 0.101	9,7 0.107	4,4 0.087	4,4 0.102	4,4 0.117	8,7 0.107	5,5 0.091	5,5 0.103	5,5 0.115	6,6 0.083	6,6 0.093	6,6 0.103
50		7,3 0.083	9,4 0.115	7,4 0.097	2,2 0.108	10,6 0.112	5,4 0.090	3,3 0.084	3,3 0.103	6,5 0.102	9,7 0.105	4,4 0.084	4,4 0.098	4,4 0.112	8,7 0.099	5,5 0.084	5,5 0.095	5,5 0.105	5,5 0.116	6,6 0.083
55		4,2 0.109	3,2 0.114	5,3 0.114	2,2 0.095	6,4 0.112	14,8 0.111	5,4 0.098	3,3 0.088	3,3 0.105	6,5 0.103	9,7 0.104	4,4 0.082	4,4 0.095	4,4 0.107	4,4 0.120	8,7 0.107	5,5 0.088	5,5 0.098	5,5 0.108
60		4,2 0.095	3,2 0.100	5,3 0.097	2,2 0.084	2,2 0.109	8,5 0.119	5,4 0.082	5,4 0.105	3,3 0.091	3,3 0.106	6,5 0.103	9,7 0.102	4,4 0.081	4,4 0.092	4,4 0.103	4,4 0.115	8,7 0.100	5,5 0.083	5,5 0.092
65		4,2 0.084	3,2 0.089	5,3 0.082	7,4 0.090	2,2 0.097	6,4 0.110	12,7 0.113	5,4 0.089	5,4 0.111	3,3 0.093	3,3 0.108	6,5 0.104	9,7 0.101	7,6 0.084	4,4 0.090	4,4 0.100	4,4 0.110	8,7 0.094	8,7 0.107
70		5,2 0.115	7,3 0.101	9,4 0.106	5,3 0.112	2,2 0.088	2,2 0.109	8,5 0.114	7,5 0.081	5,4 0.096	3,3 0.083	3,3 0.096	3,3 0.109	6,5 0.104	9,7 0.101	7,6 0.082	4,4 0.088	4,4 0.097	4,4 0.107	4,4 0.117
75		5,2 0.103	7,3 0.088	3,2 0.111	5,3 0.098	7,4 0.101	2,2 0.099	2,2 0.119	10,6 0.117	5,4 0.083	5,4 0.102	3,3 0.085	3,3 0.098	3,3 0.110	6,5 0.105	9,7 0.100	7,6 0.081	4,4 0.086	4,4 0.095	4,4 0.104
80		5,2 0.093	4,2 0.116	3,2 0.101	5,3 0.086	7,4 0.086	2,2 0.091	2,2 0.109	8,5 0.110	14,8 0.110	5,4 0.089	5,4 0.107	3,3 0.088	3,3 0.099	3,3 0.111	6,5 0.105	6,5 0.120	9,7 0.116	4,4 0.084	4,4 0.093
85		4,2 0.106	4,2 0.106	3,2 0.092	9,4 0.117	5,3 0.111	2,2 0.083	2,2 0.101	2,2 0.118	10,6 0.112	7,5 0.084	5,4 0.094	5,4 0.111	3,3 0.090	3,3 0.101	3,3 0.112	6,5 0.105	6,5 0.119	9,7 0.114	4,4 0.083
90			4,2 0.097	3,2 0.085	3,2 0.119	5,3 0.099	7,4 0.095	2,2 0.093	2,2 0.109	8,5 0.108	12,7 0.114	5,4 0.083	5,4 0.099	3,3 0.082	3,3 0.092	3,3 0.102	3,3 0.112	6,5 0.105	6,5 0.119	9,7 0.113
95			4,2 0.089	7,3 0.100	3,2 0.110	5,3 0.089	7,4 0.084	2,2 0.086	2,2 0.102	2,2 0.117	10,6 0.108	14,8 0.117	5,4 0.088	5,4 0.103	3,3 0.084	3,3 0.094	3,3 0.103	3,3 0.113	6,5 0.106	6,5 0.118
100			4,2 0.082	7,3 0.090	3,2 0.102	5,3 0.080	5,3 0.109	2,2 0.080	2,2 0.095	2,2 0.110	6,4 0.118	12,7 0.109	7,5 0.086	5,4 0.093	5,4 0.108	3,3 0.086	3,3 0.095	3,3 0.104	3,3 0.114	6,5 0.106

TABLE B-8. Critical values (ω_α) for the WRS test (n = the number of site measurements; m = the number of background measurements)

n	α	m																		
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	0.05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	4	5	5	5
	0.10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
3	0.05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	0.10	1	2	2	3	4	5	6	6	7	8 1	9	10	11	11	12	13	14	15	16
4	0.05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	0.10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23
5	0.05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	0.10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
6	0.05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	0.10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
7	0.05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	0.10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
8	0.05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	48
	0.10	3	6	8	11	14	17	20	23	25	28	31	34	37	40	43	46	49	52	55
9	0.05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	0.10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
10	0.05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	0.10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
11	0.05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	0.10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
12	0.05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	0.10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87
13	0.05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	8
	0.10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
14	0.05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	0.10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
15	0.05	4	8	13	19	24	29	34	40	45	51	56	62	67	73	78	84	89	95	101
	0.10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
16	0.05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	0.10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120
17	0.05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	0.10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
18	0.05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	0.10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
19	0.05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	0.10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
20	0.05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	0.10	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

TABLE B-9. Critical values for the two-sample t test

Degrees of Freedom	1 – α								
	.70	.75	.80	.85	.90	.95	.975	.99	.995
1	0.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	0.617	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.536	0.691	0.866	1.074	1.34	1.753	2.131	2.602	2.947
16	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704
60	0.527	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660
120	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617
∞	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576

TABLE B-10. Coefficients a_k for the Shapiro-Wilk W test for normality

k\n	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5868	0.5739
2	—	0.0000	0.1677	0.2413	0.28D6	0.3031	0.3164	0.3244	0.3291
3	—	—	—	0.0000	0.0875	0.1401	0.1743	0.1976	0.2141
4	—	—	—	—	—	0.0000	0.0561	0.0947	0.1224
5	—	—	—	—	—	—	—	0.0000	0.0399

k\n	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0,2540	0.2553	0.2561	0.2565
4	0.1429	0.1506	0.1707	0.1802	0.1876	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0 0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	—	—	0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8	—	—	—	—	0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9	—	—	—	—	—	—		0.0163	0.0303	0.0422
10	—	—	—	—	—	—	—	—	0.0000	0.0140

k\n	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043	0.3018	0.2992	0.2968	0.2944
3	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533	0.2522	0.2510	0.2499	0.2487
4	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151	0.2152	0.2151	0.2150	0.2148
5	0.1736	0.1764	0.1787	0.1007	0.1822	0.1836	0.1840	0.1857	0.1864	0.1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	0.1092	0.1150	0.1201	0.1245	0.1263	0.1316	0.1346	0.1372	0.1395	0.1415
8	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089	0.1128	0.1162	0.1192	0.1219
9	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876	0.0923	0.0965	0.1002	0.1036
10	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672	0.0728	0.0778	0.0822	0.0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	—	—	0.0000	0.0107	0.0200	0.0284	0.0358	0.0424	0.0483	0.0537
13	—	—	—	—	0.0000	0.0094	0.0178	0.0253	0.0320	0.0381
14	—	—	—	—	—	—	0.0000	0.0084	0.0159	0.0227
15	—	—	—	—	—	—	—	—	0.0000	0.0076

TABLE B-11. Critical values for the Shapiro-Wilk W test for normality

n	$W_{0.01}$	$W_{0.02}$	$W_{0.05}$	$W_{0.10}$	$W_{0.50}$
3	0.753	0.756	0.767	0.789	0.859
4	0.687	0.707	0.748	0.792	0.935
5	0.686	0.715	0.762	0.806	0.927
6	0.713	0.743	0.788	0.826	0.927
7	0.730	0.760	0.803	0.838	0.928
8	0.749	0.778	0.818	0.851	0.932
9	0.764	0.791	0.829	0.859	0.935
10	0.781	0.806	0.842	0.869	0.938
11	0.792	0.817	0.850	0.876	0.940
12	0.805	0.828	0.859	0.883	0.943
13	0.814	0.837	0.866	0.889	0.945
14	0.825	0.846	0.874	0.895	0.947
15	0.835	0.855	0.881	0.901	0.950
16	0.844	0.863	0.887	0.906	0.952
17	0.851	0.869	0.892	0.910	0.954
18	0.858	0.874	0.897	0.914	0.956
19	0.863	0.879	0.901	0.917	0.957
20	0.868	0.886	0.905	0.920	0.969
21	0.873	0.884	0.908	0.923	0.960
22	0.878	0.892	0.911	0.926	0.961
23	0.881	0.895	0.914	0.928	0.962
24	0.884	0.898	0.916	0.930	0.963
25	0.886	0.901	0.918	0.931	0.964
26	0.891	0.904	0.920	0.933	0.965
27	0.894	0.906	0.923	0.935	0.965
28	0.896	0.908	0.924	0.936	0.966
29	0.898	0.910	0.926	0.937	0.966
30	0.900	0.912	0.927	0.939	0.967
31	0.902	0.914	0.929	0.940	0.967
32	0.904	0.915	0.930	0.941	0.968
33	0.906	0.917	0.931	0.942	0.968
34	0.908	0.919	0.933	0.943	0.969
35	0.910	0.920	0.934	0.944	0.969
36	0.912	0.922	0.935	0.945	0.970
37	0.914	0.924	0.936	0.946	0.970
38	0.916	0.925	0.938	0.947	0.971
39	0.917	0.927	0.939	0.948	0.971
40	0.919	0.928	0.940	0.949	0.972
41	0.920	0.929	0.941	0.950	0.972
42	0.922	0.930	0.942	0.951	0.972
43	0.923	0.932	0.943	0.951	0.973
44	0.924	0.933	0.944	0.952	0.973
45	0.926	0.934	0.945	0.953	0.973
46	0.927	0.935	0.945	0.953	0.974
47	0.928	0.936	0.946	0.954	0.974
48	0.929	0.937	0.947	0.954	0.974
49	0.929	0.937	0.947	0.955	0.974
50	0.930	0.938	0.947	0.955	0.974

TABLE B-12. Critical values for the D'Agostino test for normality (values of Y such that 100p% of the distribution of Y is less than Y_p)

n	$Y_{0.005}$	$Y_{0.01}$	$Y_{0.025}$	$Y_{0.05}$	$Y_{0.10}$	$Y_{0.90}$	$Y_{0.95}$	$Y_{0.975}$	$Y_{0.99}$	$Y_{0.995}$
50	–3.949	–3.442	–2.757	–2.220	–1.661	0.759	0.923	1.038	1.140	1.192
60	–3.846	–3.360	–2.699	–2.179	–1.634	0.807	0.986	1.115	1.236	1.301
70	–3.762	–3.293	–2.652	–2.146	–1.612	0.844	1.036	1.176	1.312	1.388
80	–3.693	–3.237	–2.613	–2.118	–1.594	0.874	1.076	1.226	1.374	1.459
90	–3.635	–3.100	–2.580	–2.095	–1.579	0.899	1.109	1.268	1.426	1.518
100	–3.584	–3.150	–2.552	–2.075	–1.566	0.920	1.137	1.303	1.470	1.569
150	–3.409	–3.009	–2.452	–2.004	–1.520	0.990	1.233	1.423	1.623	1.746
200	–3.302	–2.922	–2.391	–1.960	–1.491	1.032	1.290	1.496	1.715	1.853
250	–3.227	–2.861	–2.348	–1.926	–1.471	1.060	1.328	1.545	1.779	1.927
300	–3.172	–2.816	–2.316	–1.906	–1.456	1.080	1.357	1.528	1.826	1.983
350	–3.129	–2.781	–2.291	–1.888	–1.444	1.096	1.379	1.610	1.863	2.026
400	–3.094	–2.753	–2.270	–1.873	–1.434	1.108	1.396	1.633	1.893	2.061
450	–3.064	–2.729	–2.253	–1.861	–1.426	1.119	1.411	1.652	1.918	2.090
500	–3.040	–2.709	–2.239	–1.850	–1.419	1.127	1.423	1.668	1.938	2.114
550	–3.019	–2.691	–2.226	–1.841	–1.413	1.135	1.434	1.682	1.957	2.136
600	–3.000	–2.676	–2.215	–1.833	–1.408	1.141	1.443	1.694	1.972	2.154
650	–2.984	–2.663	–2.206	–1.826	–1.403	1.147	1.451	1.704	1.986	2.171
700	–2.969	–2.651	–2.197	–1.820	–1.399	1.152	1.458	1.714	1.999	2.185
750	–2.956	–2.640	–2.189	–1.814	–1.395	1.157	1.465	1.722	2.010	2.199
800	–2.944	–2.630	–2.182	–1.809	–1.392	1.161	1.471	1.730	2.020	2.211
850	–2.933	–2.621	–2.176	–1.804	–1.389	1.165	1.476	1.737	2.029	2.221
900	–2.923	–2.613	–2.170	–1.800	–1.386	1.168	1.481	1.743	2.037	2.231
950	–2.914	–2.605	–2.164	–1.796	–1.383	1.171	1.485	1.749	2.045	2.241
1000	–2.906	–2.599	–2.159	–1.792	–1.381	1.174	1.489	1.754	2.052	2.249

TABLE B-13. Values of the parameter λ for the Cohen estimates of the mean and variance of normally distributed datasets that contain nondetects

	h											
γ	**.01**	**.02**	**.03**	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**	**.10**	**.15**	**.20**
0.00	.010100	.020400	.030902	.041583	.052507	.063625	.074953	.08649	.09824	.11020	.17342	.24268
0.05	.010551	.021294	.032225	.043350	.054670	.066159	.077909	.08983	.10197	.11431	.17925	.25033
0.10	.010950	.022082	.033398	.044902	.056596	.068483	.080563	.09285	.10534	.11804	.18479	.25741
0.15	.011310	.022798	.034466	.046318	.058356	.070586	.083009	.09563	.10845	.12148	.18985	.26405
0.20	.011642	.023459	.035453	.047829	.059990	.072539	.085280	.09822	.11135	.12469	.19460	.27031
0.25	.011952	.024076	.036377	.048858	.061522	.074372	.087413	.10065	.11408	.12772	.19910	.2762
0.30	.012243	.024658	.037249	.050018	.062969	.076106	.089433	.10295	.11667	.13059	.20338	.2819
0.35	.012520	.025211	.038077	.051120	.064345	.077736	.091355	.10515	.11914	.13333	.20747	.2873
0.40	.012784	.025738	.038866	.052173	.065660	.079332	.093193	.10725	.12150	.13595	.21129	.2925
0.45	.013036	.026243	.039624	.053182	.066921	.080845	.094958	.10926	.12377	.13847	.21517	.2976
0.50	.013279	.026728	.040352	.054153	.068135	.082301	.096657	.11121	.12595	.14090	.21882	.3025
0.55	.013513	.027196	.041054	.055089	.069306	.083708	.098298	.11208	.12806	.14325	.22225	.3072
0.60	.013739	.027849	.041733	.055995	.070439	.085068	.099887	.11490	.13011	.14552	.22578	.3118
0.65	.013958	.028087	.042391	.056874	.071538	.086388	.10143	.11666	.13209	.14773	.22910	.3163
0.70	.014171	.028513	.043030	.057726	.072505	.087670	.10292	.11837	.13402	.14987	.23234	.3206
0.75	.014378	.029927	.043652	.058556	.073643	.088917	.10438	.12004	.13590	.15196	.23550	.32489
0.80	.014579	.029330	.044258	.059364	.074655	.090133	.10580	.12167	.13775	.15400	.23858	.32903
0.85	.014773	.029723	.044848	.060153	.075642	.091319	.10719	.12225	.13952	.15599	.24158	.33307
0.90	.014967	.030107	.045425	.060923	.075606	.092477	.10854	.12480	.14126	.15793	.24452	.33703
0.95	.015154	.030483	.045989	.061676	.077549	.093611	.10987	.12632	.14297	.15983	.24740	.34091
1.00	.015338	.030850	.046540	.062413	.078471	.094720	.11116	.12780	.14465	.16170	.25022	.34471

	h											
γ	**.25**	**.30**	**.35**	**.40**	**.45**	**.50**	**.55**	**.60**	**.65**	**.70**	**.80**	**.90**
0.00	.31862	.4021	.4941	.5961	.7096	.8388	.9808	1.145	1.336	1.561	2.176	3.283
0.05	.32793	.4130	.5066	.6101	.7252	.8540	.9994	1.166	1.358	1.585	2.203	3.314
0.10	.33662	.4233	.5184	.6234	.7400	.8703	1.017	1.185	1.379	1.608	2.229	3.345
0.15	.34480	.4330	.5296	.6361	.7542	.8860	1.035	1.204	1.400	1.630	2.255	3.376
0.20	.35255	.4422	.5403	.6483	.7673	.9012	1.051	1.222	1.419	1.651	2.280	3.405
0.25	.35993	.4510	.5506	.6600	.7810	.9158	1.067	1.240	1.439	1.672	2.305	3.435
0.30	.36700	.4595	.5604	.6713	.7937	.9300	1.083	1.257	1.457	1.693	2.329	3.464
0.35	.37379	.4676	.5699	.6821	.8060	.9437	1.098	1.274	1.475	1.713	2.353	3.492
0.40	.38033	.4735	.5791	.6927	.8179	.9570	1.113	1.290	1.494	1.732	2.376	3.520
0.45	.38665	.4831	.5880	.7029	.8295	.9700	1.127	1.306	1.511	1.751	2.399	3.547
0.50	.39276	.4904	.5967	.7129	.8408	.9826	1.141	1.321	1.528	1.770	2.421	3.575
0.55	.39679	.4976	.6061	.7225	.8517	.9950	1.155	1.337	1.545	1.788	2.443	3.601
0.60	.40447	.5045	.6133	.7320	.8625	1.007	1.169	1.351	1.561	1.806	2.465	3.628
0.65	.41008	.5114	.6213	.7412	.8729	1.019	1.182	1.368	1.577	1.824	2.486	3.654
0.70	.41555	.5180	.6291	.7502	.8832	1.030	1.195	1.380	1.593	1.841	2.507	3.679
0.75	.42090	.5245	.6367	.7590	.8932	1.042	1.207	1.394	1.608	1.851	2.528	3.705
0.80	.42612	.5308	.6441	.7676	.9031	1.053	1.220	1.408	1.624	1.875	2.548	3.730
0.85	.43122	.5370	.6515	.7781	.9127	1.064	1.232	1.422	1.639	1.892	2.568	3.754
0.90	.43622	.5430	.6586	.7844	.9222	1.074	1.244	1.435	1.653	1.908	2.588	3.779
0.95	.44112	.5490	.6656	.7925	.9314	1.085	1.255	1.448	1.668	1.924	2.607	3.803
1.00	.44592	.5548	.6724	.8005	.9406	1.095	1.287	1.461	1.882	1.940	2.626	3.827

TABLE B-14. Critical values for the Dixon extreme value test for outliers

n	Level of Significance α		
	0.10	**0.05**	**0.01**
3	0.886	0.941	0.988
4	0.679	0.765	0.889
5	0.557	0.642	0.780
6	0.482	0.560	0.698
7	0.434	0.507	0.637
8	0.479	0.554	0.683
9	0.441	0.512	0.635
10	0.409	0.477	0.597
11	0.517	0.576	0.679
12	0.490	0.546	0.642
13	0.467	0.521	0.615
14	0.492	0.546	0.641
15	0.472	0.525	0.616
16	0.454	0.507	0.595
17	0.438	0.490	0.577
18	0.424	0.475	0.561
19	0.412	0.462	0.547
20	0.401	0.450	0.535
21	0.391	0.440	0.524
22	0.382	0.430	0.514
23	0.374	0.421	0.505
24	0.367	0.413	0.497
25	0.360	0.406	0.489

TABLE B-15. Critical values for the Discordance test for outliers

n	Level of Significance	
	0.01	**0.05**
3	1.155	1.153
4	1.492	1.463
5	1.749	1.672
6	1.944	1.822
7	2.097	1.938
8	2.221	2.032
9	2.323	2.110
10	2.410	2.176
11	2.485	2.234
12	2.550	2.285
13	2.607	2.331
14	2.659	2.371
15	2.705	2.409
16	2.747	2.443
17	2.785	2.475
18	2.821	2.504
19	2.854	2.532
20	2.884	2.557
21	2.912	2.580
22	2.939	2.603
23	2.963	2.624
24	2.987	2.644
25	3.009	2.663
26	3.029	2.681
27	3.049	2.698
28	3.068	2.714
29	3.085	2.730
30	3.103	2.745
31	3.119	2.759
32	3.135	2.773
33	3.150	2.786
34	3.164	2.799
35	3.178	2.811
36	3.191	2.823
37	3.204	2.835
38	3.216	2.846
39	3.228	2.857
40	3.240	2.866
41	3.251	2.877
42	3.261	2.887
43	3.271	2.896
44	3.282	2.905
45	3.292	2.914
46	3.302	2.923
47	3.310	2.931
48	3.319	2.940
49	3.329	2.948
50	3.336	2.956

TABLE B-16. Approximate critical values for the Rosner test for outliers

n	r	α	
		0.05	**0.01**
25	1	2.82	3.14
	2	2.80	3.11
	3	2.78	3.09
	4	2.76	3.06
	5	2.73	3.03
	10	2.59	2.85
26	1	2.84	3.16
	2	2.82	3.14
	3	2.80	3.11
	4	2.78	3.09
	5	2.76	3.06
	10	2.62	2.89
27	1	2.86	3.18
	2	2.84	3.16
	3	2.82	3.14
	4	2.80	3.11
	5	2.78	3.09
	10	2.65	2.93
28	1	2.88	3.20
	2	2.86	3.18
	3	2.84	3.16
	4	2.82	3.14
	5	2.80	3.11
	10	2.68	2.97
29	1	2.89	3.22
	2	2.88	3.20
	3	2.86	3.18
	4	2.84	3.16
	5	2.82	3.14
	10	2.71	3.00
30	1	2.91	3.24
	2	2.89	3.22
	3	2.88	3.20
	4	2.86	3.18
	5	2.84	3.16
	10	2.73	3.03
31	1	2.92	3.25
	2	2.91	3.24
	3	2.89	3.22
	4	2.88	3.20
	5	2.86	3.18
	10	2.76	3.06
46	1	3.09	3.45
	2	3.09	3.44
	3	3.08	3.43
	4	3.07	3.41
	5	3.06	3.40
	10	3.00	3.34

n	r	α	
		0.05	**0.01**
32	1	2.94	3.27
	2	2.92	3.25
	3	2.91	3.24
	4	2.89	3.22
	5	2.88	3.20
	10	2.78	3.09
33	1	2.95	3.29
	2	2.94	3.27
	3	2.92	3.25
	4	2.91	3.24
	5	2.89	3.22
	10	2.80	3.11
34	1	2.97	3.30
	2	2.95	3.29
	3	2.94	3.27
	4	2.92	3.25
	5	2.91	3.24
	10	2.82	3.14
35	1	2.98	3.32
	2	2.97	3.30
	3	2.95	3.29
	4	2.94	3.27
	5	2.92	3.25
	10	2.84	3.16
36	1	2.99	3.33
	2	2.98	3.32
	3	2.97	3.30
	4	2.95	3.29
	5	2.94	3.27
	10	2.86	3.18
37	1	3.00	3.34
	2	2.99	3.33
	3	2.98	3.32
	4	2.97	3.30
	5	2.95	3.29
	10	2.88	3.20
38	1	3.01	3.36
	2	3.00	3.34
	3	2.99	3.33
	4	2.98	3.32
	5	2.97	3.30
	10	2.91	3.22
70	1	3.26	3.62
	2	3.25	3.62
	3	3.25	3.61
	4	3.24	3.60
	5	3.24	3.60
	10	3.21	3.57

n	r	α	
		0.05	**0.01**
39	1	3.03	3.37
	2	3.01	3.36
	3	3.00	3.34
	4	2.99	3.33
	5	2.98	3.32
	10	2.91	3.24
40	1	3.04	3.38
	2	3.03	3.37
	3	3.01	3.36
	4	3.00	3.34
	5	2.99	3.33
	10	2.92	3.25
41	1	3.05	3.39
	2	3.04	3.38
	3	3.03	3.37
	4	3.01	3.36
	5	3.00	3.34
	10	2.94	3.27
42	1	3.06	3.40
	2	3.05	3.39
	3	3.04	3.38
	4	3.03	3.37
	5	3.01	3.36
	10	2.95	3.29
43	1	3.07	3.41
	2	3.06	3.40
	3	3.05	3.39
	4	3.04	3.38
	5	3.03	3.37
	10	2.97	3.30
44	1	3.08	3.43
	2	3.07	3.41
	3	3.06	3.40
	4	3.05	3.39
	5	3.04	3.38
	10	2.98	3.32
45	1	3.09	3.44
	2	3.08	3.43
	3	3.07	3.41
	4	3.06	3.40
	5	3.05	3.39
	10	2.99	3.33
250	1	3.67	4.04
	5	3.67	4.04
	10	3.66	4.03

TABLE B-16. (continued)

n	r	α	
		0.05	**0.01**
47	1	3.10	3.46
	2	3.09	3.45
	3	3.09	3.44
	4	3.08	3.43
	5	3.07	3.41
	10	3.01	3.36
48	1	3.11	3.46
	2	3.10	3.46
	3	3.09	3.45
	4	3.09	3.44
	5	3.08	3.43
	10	3.03	3.37
49	1	3.12	3.47
	2	3.11	3.46
	3	3.10	3.46
	4	3.09	3.45
	5	3.09	3.44
	10	3.04	3.38
50	1	3.13	3.48
	2	3.12	3.47
	3	3.11	3.46
	4	3.10	3.46
	5	3.09	3.45
	10	3.05	3.39
60	1	3.20	3.56
	2	3.19	3.55
	3	3.19	3.55
	4	3.18	3.54
	5	3.17	3.53
	10	3.14	3.49

n	r	α	
		0.05	**0.01**
80	1	3.31	3.67
	2	3.30	3.67
	3	3.30	3.66
	4	3.29	3.66
	5	3.29	3.65
	10	3.26	3.63
90	1	3.35	3.72
	2	3.34	3.71
	3	3.34	3.71
	4	3.34	3.70
	5	3.33	3.70
	10	3.31	3.68
100	1	3.38	3.75
	2	3.38	3.75
	3	3.38	3.75
	4	3.37	3.74
	5	3.37	3.74
	10	3.35	3.72
150	1	3.52	3.89
	2	3.51	3.89
	3	3.51	3.89
	4	3.51	3.88
	5	3.51	3.88
	10	3.50	3.87
200	1	3.61	3.98
	2	3.60	3.98
	3	3.60	3.97
	4	3.60	3.97
	5	3.60	3.97
	10	3.59	3.96

n	r	α	
		0.05	**0.01**
300	1	3.72	4.09
	5	3.72	4.09
	10	3.71	4.09
350	1	3.77	4.14
	5	3.76	4.13
	10	3.76	4.13
400	1	3.80	4.17
	5	3.80	4.17
	10	3.80	4.16
450	1	3.84	4.20
	5	3.83	4.20
	10	3.83	4.20
500	1	3.86	4.23
	5	3.86	4.23
	10	3.86	4.22

APPENDIX C: FURTHER READINGS

C.1 Printed Matter

American Society for Testing and Materials (ASTM). 1994. *Standard Guide for the Contents of Geostatistical Site Investigation Report.* D5549-94e1. West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 1994. *Standard Guide for Selection of Simulation Approaches in Geostatistical Site Investigations.* D5924-96e1. West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 1996. *Standard Guide for Analysis of Spatial Variation in Geostatistical Site Investigations.* D5922-96e1. West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 1996. *Standard Guide for Selection of Kriging Methods in Geostatistical Site Investigations.* D5923-96e1. West Conshohocken, PA.

Connor, J.J., H.T. Shacklette, R.J. Ebens, J.A. Erdman, A.T. Miesch, R.R. Tidball, and H.A. Tourtelot. 1975. *Background Geochemistry of Some Rocks, Soils, Plants, and Vegetables in the Conterminous United States.* U.S. Geological Survey Professional Paper 574-F. Washington, DC.

California Environmental Protection Agency (Cal/EPA), Department of Toxic Substances Control (DTSC). 1992. *Supplemental Guidance for Human Health Multimedia Risk Assessments of Hazardous Waste Sites and Permitted Facilities.* State of California, Office of the Science Advisor. July.

California Environmental Protection Agency (Cal/EPA), Department of Toxic Substances Control (DTSC). 1994. *Preliminary Endangerment Assessment Guidance Manual.* State of California. January.

California Environmental Protection Agency (Cal/EPA), Department of Toxic Substances Control (DTSC). 1996. *Determining Ambient Concentrations of Metals for Risk Assessments at Hazardous Waste Sites and Permitted Facilities.* Interim final policy. State of California. June.

Davidson, J.R. Jr., N.L. Hassig, J.E. Wilson, and R.O. Gilbert. 2001. *Visual Sample Plan Version 1.0 User's Guide.* PNNL-13490. Richland, WA: Pacific Northwest National Laboratory. March.

Dixon, W.J. 1953. Processing Data for Outliers. *Biometrics* 9:74-89.

Fleishhauer, H.L., and N. Korte. 1990. Formulation of Cleanup Standards for Trace Elements with Probability Plots. *Environmental Management* 14(1): 95-105.

Gehan, E.A. 1965. A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples. *Biometrika* 52:203-223.

Gilbert, R.O., J.R. Davidson, Jr., J.E. Wilson, and B.A. Pulsipher. 2001. *Visual Sample Plan (VSP) Models and Code Verification.* PNNL-13450. Richland, WA: Pacific Northwest National Laboratory.

Gogolak, C.V., G.E. Powers, and A.M. Huffert. 1997. *A Nonparametric Statistical Methodology for the Design and Analysis of Final Status Decommissioning Surveys.* NUREG-1505. Washington, DC: U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Division of Regulatory Applications.

Hamblin, W.K. 1985. *The Earth's Dynamic Systems.* 4th ed. Minneapolis: Burgess.

Lepeltier, C. 1969. A Simplified Statistical Treatment of Geochemical Data by Graphical Representation. *Economic Geology* 64:538-550.

Moore, D.S., and G.P. McCabe. 1999. *Introduction to the Practice of Statistics.* 3rd ed. New York: W. H. Freeman.

Rosenbaum, S. 1954. Tables for a Nonparametric Test of Location. *Annals of Mathematical Statistics* 25:146-150.

Rosner, B. 1983. Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics* 25:165-172.

Shapiro, S. and M.B. Wilk. 1965. An Analysis of Variance Test for Normality. *Biometrika* 52:591-611.

Thompson, S.K., and G.A. Seber. 1996. *Adaptive Sampling*. New York: Wiley.

Sinclair, A.J. 1974. Selection of Threshold Values in Geochemical Data Using Probability Graphs. *J. Geochem. Explor.* 3:129-149.

United States Air Force (USAF). 1996. *Operable Unit 3 Remedial Investigation Report, Andersen Air Force Base, Guam.* Final. USAF-204-R. USAF Installation Restoration Program. December. Revised March 1997.

United States Environmental Protection Agency (U.S. EPA). 1985. *National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule.* 50 Federal Register 47912. November.

Walsh, J.E. 1958. Large Sample Nonparametric Rejection of Outlying Observations. *Annals of the Institute of Statistical Mathematics* 10:223-232.

Zhou, Xiao-Hua, Sujuan Gao, and Siu L. Hui. 1997. Methods for Comparing the Means of Two Independent Log-Normal Samples. *Biometrics* 53:1129-1135.

C.2 Internet Resources

Department of Energy pages on the DQO process, case studies, software, and training: URL: http://dqo.pnl.gov/ (Pacific Northwest National Laboratory). URL: http://www.hanford.gov/dqo/ (Hanford site).

Microsoft TerraServer – USGS topographic maps (digital raster graphics) and orthorectified aerial photographs (downsampled digital orthophoto quadrangles) available free in tiled format: URL: http://www.terraserver.microsoft.com.

U.S. EPA Office of Environmental Information (formerly, in part, Center for Environmental and Information Statistics). Information on environmental quality, status, and trends: URL: http://www.epa.gov/oei/quality.htm.

U.S. EPA Quality System DQO guidance documents: URL: http://www.epa.gov/quality1/.

Water Science Glossary of Terms – U.S. Geological Survey: URL: http://ga.water.usgs.gov/edu/dictionary.html.

Water Quality Glossary – Water Quality Association: URL: http://www.wqa.org/glossary.cfm.

Water Words Dictionary – Nevada Division of Water Resources: URL: http://water.nv.gov/Water%20planning/dict-1/ww-index.htm.

Hanford's Environmental Restoration Project – Bechtel Hanford, Inc.: URL: http://www.bhi-erc.com.

HyperStat Online, an introductory-level hypertext statistics book: URL: http://davidmlane.com/hyperstat/index.html.

The Internet Glossary of Statistical Terms – The Animated Software Co.: URL: http://www.animatedsoftware.com/statglos/statglos.htm.

The "new statistics" of resampling (drawing repeated samples from the given data, or population suggested by the data) – Resampling Stats: URL: http://www.resample.com.

NFESC Environmental Services and BRAC page – numerous links to guidance/policy documents and technical information: URL: http://enviro.nfesc.navy.mil/.

Statistical courses, software, and additional information: URL: http://www.statistics.com/.

Statistics Glossary – Centre for Applied Statistics, Lancaster University, U.K.: URL: http://www.cas.lancs.ac.uk/glossary_v1.1/main.html

C.3 Computer Software

Data Quality Assessment Statistical Toolbox (DataQUEST) (G-9D). User's Guide and Software. United States Environmental Protection Agency. 1997. EPA/600/R-96/085. Washington, DC: Office of Research and Development. URL: http://www.epa.gov/quality/dqa.html.

Data Quality Objectives Decision Error Feasibility Trials (DQO/DEFT). United States Environmental Protection Agency. 1994. EPA QA/G-4D. EPA 600/R-96/056. Washington, DC: Office of Research and Development. URL: http://www.epa.gov/superfund/programs/dfa/decsupp.htm.

Data Quality Objectives Decision Error Feasibility Trials (DEFT) Software –User's Guide. United States Environmental Protection Agency (U.S. EPA). 2001. EPA QA/G-4D. EPA/240.B-01/007. URL: http://www.epa.gov/quality.

DUMPStat. Groundwater Detection Monitoring Statistics. Discerning Systems, Inc. URL: http://www.dsi-software.com/dumpstat.html.

ELIPGRID-PC (Hot Spot Probability Calculations): Upgraded Version. Davidson, J. R. 1995. ORNL/TM-13103. Oak Ridge, TN.: Oak Ridge National Laboratory. December. URL: http://dqo.pnl.gov/software/elipgrid.htm.

Environmental Stats for S-PLUS. Seattle: Insightful Corporation. (Millard, S.P., and N.K. Neerchal. 2001. *Environmental Statistics with S-PLUS*. Boca Raton, FL: CRC Press.) URL: www.insightful.com.

GEOPACK (geostatistical software). United States Environmental Protection Agency. 1990. EPA/600/8-90/004. Washington, DC: Environmental Sciences Division. URL: http://www.epa.gov/ada/csmos/models/geopack.html.

Geostatistical Environmental Assessment Software (Geo-EAS) 1.2.1. United States Environmental Protection Agency. 1988. EPA/600/4-88/033a. Washington, DC: Environmental Sciences Division. URL: http://www.epa.gov/superfund/programs/dfa/decsupp.htm.

MARSSIM (Multi-Agency Radiation Survey and Site Investigation Manual) software tools. United States Environmental Protection Agency. URL: http://www.epa.gov/radiation/marssim/tools.htm.

SCOUT: A Data Analysis Program (a univariate and multivariate data analysis tool). United States Environmental Protection Agency. 1993. Washington, DC: Office of Research and Development, Environmental Monitoring Systems Laboratory. URL: http://www.epa.gov/nerlesd1/databases/scout/abstract.htm.

SigmaPlot and *SPSS.* Chicago: SPSS Inc. URL: www.spss.com.

StatView 5.0. Cary, N.C.: SAS Institute, Inc. URL: www.statview.com.

Visual Sample Plan (sampling design software) from Department of Energy: URL: http://dqo.pnl.gov/VSP/Index.htm.