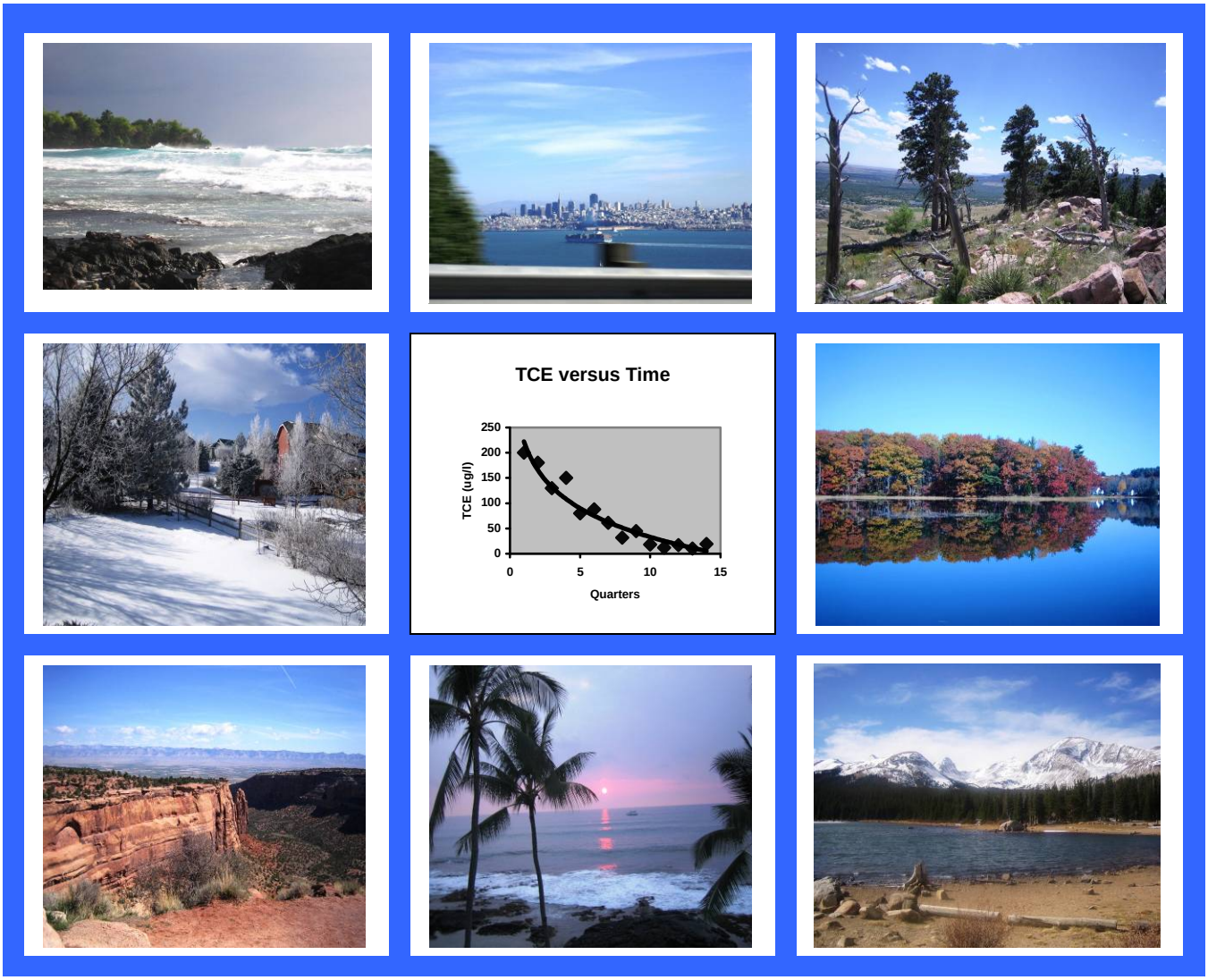
US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

STATISTICAL ANALYSIS OF GROUNDWATER MONITORING DATA AT RCRA FACILITIES

UNIFIED GUIDANCE MARCH 2009

EPA 530/R-09-007



ENVIRONMENTAL PROTECTION AGENCY

OFFICE OF RESOURCE CONSERVATION AND RECOVERY

US EPA ARCHIVE DOCUMENT

This page left blank for printing

STATISTICAL ANALYSIS OF GROUNDWATER MONITORING DATA AT RCRA FACILITIES

UNIFIED GUIDANCE

OFFICE OF RESOURCE CONSERVATION AND RECOVERY

PROGRAM IMPLEMENTATION AND INFORMATION DIVISION

U.S. ENVIRONMENTAL PROTECTION AGENCY

MARCH 2009

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

This page intentionally left blank

DISCLAIMER

This Unified Guidance has been prepared to assist EPA's Regions, the States and the regulated community in testing and evaluating groundwater monitoring data under 40 CFR Parts 264 and 265 and 40 CFR Part 258. This guidance is not a rule, is not legally enforceable, and does not confer legal rights or impose legal obligations on any member of the public, EPA, the States or any other agency. While EPA has made every effort to ensure the accuracy of the discussion in this guidance, the obligations of the regulated community are determined by the relevant statutes, regulations, or other legally binding requirements. The use of the term "should" when used in this guidance does not connote a requirement. This guidance may not apply in a particular situation based on the circumstances. Regional and State personnel retain the discretion to adopt approaches on a case-by-case basis that differ from this guidance where appropriate.

It should be stressed that this guidance is a work in progress. Given the complicated nature of groundwater and geochemical behavior, statistical applications describing and evaluating data patterns have evolved over time. While many new approaches and a conceptual framework have been provided here based on our understanding at the time of publication, outstanding issues remain. The Unified Guidance sets out mostly classical statistical methods using reasonable interpretations of existing regulatory objectives and constraints. But even these highly developed mathematical models deal primarily with sorting out chance effects from potentially real differences or trends. They do not exhaust the possibilities of groundwater definition using other technical or scientific techniques (*e.g.*, contaminant modeling or geostatistical evaluations). While providing a workable decision framework, the models and approaches offered within the Unified Guidance are only approximations of a complex underlying reality.

While providing a basic understanding of underlying statistical principles, the guidance doesn't attempt to provide the reader with more thorough explanations and derivations found in standard texts and papers. It also doesn't comprehensively cover all potential statistical approaches, and confines itself to reasonable and current methods, which will work in the present RCRA groundwater context. While it is highly likely that methods promoted in this guidance will be applied using commercial or proprietary statistical software, a detailed discussion of software applications is beyond the scope of this document.

This document has been reviewed by the Office of Resource Conservation and Recovery (former Office of Solid Waste), U.S. Environmental Protection Agency, Washington, D.C., and approved for publication. Mention of trade names, commercial products, or publications does not constitute endorsement or recommendation for use.

"It is far better to have an approximate answer to the right question than a precise answer to the wrong question…" — John Hauser

US EPA ARCHIVE DOCUMENT

ACKNOWLEDGMENTS

EPA's Office of Solid Waste developed initial versions of this document under the direction of James R. Brown and Hugh Davis, Ph.D. of the Permits and State Programs Division. The final draft was developed and edited under the direction of Mike Gansecki, EPA Region 8. The guidance was prepared by Kirk M. Cameron, Ph.D., Statistical Scientist and President of MacStat Consulting, Ltd. in Colorado Springs, Colorado. It also incorporates the substantial efforts on the 1989 Interim Final Guidance of Jerry Flora, Ph.D. and Ms. Karen Bauer, both — at the time — of Midwest Research Institute in Kansas City, Missouri. Science Applications International Corporation (SAIC) provided technical support in developing this document under EPA Contract No. EP-WO-5060.

EPA also wishes to acknowledge the time and effort spent in reviewing and improving the document by a workgroup composed of statisticians, Regional and State personnel, and industry representatives — Dr. Robert Gibbons, Dr. Charles Davis, Sarah Hession, Dale Bridgford, Mike Beal, Katie Crowell, Bob Stewart, Charlotte Campbell, Evan Englund, Jeff Johnson, Mary Beck, John Baker and Dave Burt. We also wish to acknowledge the excellent comments by a number of state, EPA Regional and industry parties on the September 2004 draft. Finally, we gratefully acknowledge the detailed reviews, critiques and comments of Dr. Dennis Helsel of the US Geologic Survey, Dr. James Loftis of Colorado State University, and Dr. William Huber of Quantitative Decisions Inc., who provided formal peer reviews of the September 2004 draft.

A special note of thanks is due to Dave Bartenfelder and Ken Lovelace of the EPA CERCLA program, without whose assistance this document would not have been completed.

US EPA ARCHIVE DOCUMENT

EXECUTIVE SUMMARY

The Unified Guidance provides a suggested framework and recommendations for the statistical analysis of groundwater monitoring data at RCRA facility units subject to 40 CFR Parts 264 and 265 and 40 CFR Part 258, to determine whether groundwater has been impacted by a hazardous constituent release. Specific statistical methods are identified in the RCRA regulations, but their application is not described in any detail. The Unified Guidance provides examples and background information that will aid in successfully conducting the required statistical analyses. The Unified Guidance draws upon the experience gained in the last decade in implementing the RCRA Subtitle C and D groundwater monitoring programs and new research that has emerged since earlier Agency guidance.

The guidance is primarily oriented towards the groundwater monitoring statistical analysis provisions of 40 CFR Parts 264.90 to .100. Similar requirements for groundwater monitoring at solid waste landfill facilities under 40 CFR Part 258 are also addressed. These regulations govern the detection, characterization and response to releases from regulated units into the uppermost aquifer. Some of the methods and strategies set out in this guidance may also be appropriate for analysis of groundwater monitoring data from solid waste management units subject to 40 CFR 264.101. Although the focus of this guidance is to address the RCRA regulations, it can be used by the CERCLA program and for improving remedial actions at other groundwater monitoring programs.

Part I of the Unified Guidance introduces the context for statistical testing at RCRA facilities. It provides an ***overview of the regulatory requirements***, summarizing the current RCRA Subtitle C and D regulations and outlining the statistical methods in the final rules, as well as key regulatory sections affecting statistical decisions. It explains the basic groundwater monitoring framework, philosophy and intent of each stage of monitoring — detection, compliance (or assessment), and corrective action — and certain features common to the groundwater monitoring environment. Underlying statistical ideas common to all statistical test procedures are identified, particularly issues involving false positives arising from multiple statistical comparisons and statistical power to detect contamination.

A new component of the Unified Guidance addresses issues of ***statistical design***: what factors are important in constructing a reasonable and effective statistical monitoring program. These include the establishment and updating of background data, designing an acceptable detection monitoring plan, and statistical strategies for compliance/assessment monitoring and corrective action. This part also includes a short summary of statistical methods recommended in the Unified Guidance, detailing conditions for their appropriate use.

Part II of the Unified Guidance covers diagnostic evaluations of historical facility data for the purpose of ***checking key assumptions*** implicit in the recommended statistical tests and ***for making appropriate adjustments to the data*** (*e.g.*, consideration of outliers, seasonal autocorrelation, or non-detects). Also included is a discussion of groundwater sampling and how hydrologic factors such as flow and gradient can impact the sampling program. Concepts of statistical and physical independence are compared, with caveats provided regarding the impact of dependent data on statistical test results. Statistical methods are suggested for identifying special kinds of dependence known as spatial and temporal variation, including reasonable approaches when these dependencies are observed. Tests for trends are also included in this part.

US EPA ARCHIVE DOCUMENT

Part III of the Unified Guidance presents a range of ***detection monitoring*** statistical procedures. First, there is a discussion of the Student's *t*-test and its non-parametric counterpart, the Wilcoxon rank-sum test, when comparing two groups of data (*e.g.*, background versus one downgradient well). This part defines both parametric and non-parametric prediction limits, and their application to groundwater analysis when multiple comparisons are involved. A variety of prediction limit possibilities are presented to cover likely interpretations of sampling and testing requirements under the RCRA regulations.

Substantial detailed guidance is offered for using prediction limits with retesting procedures, and how various retesting algorithms might be constructed. The final chapter of this Part considers another statistical method especially useful for ***intrawell*** comparisons, namely the Shewhart-CUSUM control chart. A brief discussion of analysis of variance [ANOVA] and tolerance limit tests identified in the RCRA regulations is also provided.

Part IV of the Unified Guidance is devoted to statistical methods recommended for ***compliance*** or ***assessment monitoring*** and ***corrective action***. Compliance monitoring typically involves a comparison of downgradient well data to a groundwater protection standard [GWPS], which may be a limit derived from background or a fixed concentration limit (such as in 40 CFR 264.94 Table 1, an MCL, a risk-based limit, an alternate concentration limit, or a defined clean-up standard under corrective action). The key statistical procedure is the confidence interval, and several confidence interval tests (mean, median, or upper percentile) may be appropriate for compliance evaluation depending on the circumstances. The choice depends on the distribution of the data, frequency of non-detects, the type of standard being compared, and whether or not the data exhibit a significant trend. Discussions in this part consider fixed compliance standards used in a variety of EPA programs and what they might represent in statistical terms. Strategies for corrective action differ from those appropriate for compliance monitoring primarily because statistical hypotheses are changed, although the same basic statistical methods may be employed.

Since some programs will also utilize background as standards for compliance and corrective action monitoring, those tests and discussions under **Part III** detection monitoring (including statistical design in **Part I**) may pertain in identifying the appropriate standards and tests.

A ***glossary*** of important statistical terms, ***references*** and a subject ***index*** are provided at the end of the main text. The ***Appendices*** contain additional notes on a number of topics including previous guidance, a special study for the guidance, more detailed statistical power discussions, and an extensive set of ***statistical tables*** for implementing the methods outlined in the Unified Guidance. Some tables, especially those for prediction limit retesting procedures, have been extended within the Unified Guidance beyond published sources in order to cover a wider variety of plausible scenarios.

US EPA ARCHIVE DOCUMENT

TABLE OF CONTENTS

DISCLAIMER.....i
ACKNOWLEDGMENTS.....ii
EXECUTIVE SUMMARY.....iii
TABLE OF CONTENTS.....v

PART I. STATISTICAL DESIGN AND PHILOSOPHY

CHAPTER 1. OBJECTIVES AND POTENTIAL USE OF THIS GUIDANCE

1.1 OBJECTIVES..... 1-1
1.2 APPLICABILITY TO OTHER ENVIRONMENTAL PROGRAMS..... 1-3

CHAPTER 2. REGULATORY OVERVIEW

2.1 REGULATORY SUMMARY.....2-1
2.2 SPECIFIC REGULATORY FEATURES AND STATISTICAL ISSUES.....2-6
2.2.1 Statistical Methods Identified under §264.97(h) and §258.53(g).....2-6
2.2.2 Performance Standards under §264.97(i) and §258.53(h).....2-7
2.2.3 Hypothesis Tests in Detection, Compliance and Corrective Action Monitoring.....2-10
2.2.4 Sampling Frequency Requirements.....2-10
2.2.5 Groundwater Protection Standards.....2-12
2.3 UNIFIED GUIDANCE RECOMMENDATIONS..... 2-13
2.3.1 Interim Status Monitoring.....2-13
2.3.2 Parts 264 and 258 Detection Monitoring Methods.....2-14
2.3.3 Parts 264 and 258 Compliance/assessment Monitoring.....2-15

CHAPTER 3. KEY STATISTICAL CONCEPTS

3.1 INTRODUCTION TO GROUNDWATER STATISTICS.....3-2
3.2 COMMON STATISTICAL ASSUMPTIONS..... 3-4
3.2.1 Statistical Independence.....3-4
3.2.2 Stationarity.....3-5
3.2.3 Lack of Statistical Outliers.....3-7
3.2.4 Normality.....3-7
3.3 COMMON STATISTICAL MEASURES..... 3-9
3.4 HYPOTHESIS TESTING FRAMEWORK..... 3-12
3.5 ERRORS IN HYPOTHESIS TESTING..... 3-14
3.5.1 False Positives and Type I Errors.....3-15
3.5.2 Sampling Distributions, Central Limit Theorem.....3-16
3.5.3 False Negatives, Type II Errors, and Statistical Power.....3-18
3.5.4 Balancing Type I and Type II Errors.....3-22

CHAPTER 4. GROUNDWATER MONITORING PROGRAMS AND STATISTICAL ANALYSIS

4.1 THE GROUNDWATER MONITORING CONTEXT..... 4-1
4.2 RCRA GROUNDWATER MONITORING PROGRAMS.....4-3

US EPA ARCHIVE DOCUMENT

4.3 STATISTICAL SIGNIFICANCE IN GROUNDWATER TESTING 4-6
4.3.1 Statistical Factors 4-8
4.3.2 Well System Design and Sampling Factors 4-8
4.3.3 Hydrological Factors 4-9
4.3.4 Geochemical Factors 4-10
4.3.5 Analytical Factors 4-10
4.3.6 Data or Analytic Errors 4-11

CHAPTER 5. ESTABLISHING AND UPDATING BACKGROUND

5.1 IMPORTANCE OF BACKGROUND 5-1
5.1.1 Tracking Natural Groundwater Conditions 5-2
5.2 ESTABLISHING AND REVIEWING BACKGROUND 5-2
5.2.1 Selecting Monitoring Constituents and Adequate Sample Sizes 5-2
5.2.2 Basic Assumptions About Background 5-4
5.2.3 Outliers in Background 5-5
5.2.4 Impact of Spatial Variability 5-6
5.2.5 Trends in Background 5-7
5.2.6 Summary: Expanding Background Sample Sizes 5-8
5.2.7 Review of Background 5-10
5.3 UPDATING BACKGROUND 5-12
5.3.1 When to Update 5-12
5.3.2 How to Update 5-12
5.3.3 Impact of Retesting 5-14
5.3.4 Updating When Trends are Apparent 5-14

CHAPTER 6. DETECTION MONITORING PROGRAM DESIGN

6.1 INTRODUCTION 6-1
6.2 Elements of the Statistical Program Design 6-2
6.2.1 The Multiple Comparisons Problem 6-2
6.2.2 Site-Wide False Positive Rates [SWFPR] 6-7
6.2.3 Recommendations for Statistical Power 6-13
6.2.4 Effect Sizes and Data-Based Power Curves 6-18
6.2.5 Sites Using More Than One Statistical Method 6-21
6.3 HOW KEY ASSUMPTIONS IMPACT STATISTICAL DESIGN 6-25
6.3.1 Statistical Independence 6-25
6.3.2 Spatial Variation: Interwell vs. Intrawell Testing 6-29
6.3.3 Outliers 6-34
6.3.4 Non-Detects 6-36
6.4 DESIGNING DETECTION MONITORING TESTS 6-38
6.4.1 T-Tests 6-38
6.4.2 Analysis Of Variance [ANOVA] 6-38
6.4.3 Trend Tests 6-41
6.4.4 Statistical Intervals 6-42
6.4.5 Control Charts 6-46
6.5 SITE DESIGN EXAMPLES 6-46

US EPA ARCHIVE DOCUMENT

CHAPTER 7. STRATEGIES FOR COMPLIANCE/ASSESSMENT MONITORING AND CORRECTIVE ACTION

7.1 Introduction 7-1
7.2 Hypothesis Testing Structures 7-3
7.3 Groundwater Protection Standards 7-6
7.4 Designing A Statistical Program 7-9
7.4.1 False Positives and Statistical Power in Compliance/Assessment 7-9
7.4.2 False Positives and Statistical Power In Corrective Action 7-12
7.4.3 Recommended Strategies 7-13
7.4.4 Accounting for Shifts and Trends 7-14
7.4.5 Impact of Sample Variability, Non-Detects, And Non-Normal Data 7-17
7.5 Comparisons to Background Data 7-20

CHAPTER 8. SUMMARY OF RECOMMENDED METHODS

8.1 Selecting the right statistical method 8-1
8.2 Table 8.1 inventory of recommended Methods 8-4
8.3 Method Summaries 8-9

PART II. DIAGNOSTIC METHODS AND TESTING

CHAPTER 9. COMMON EXPLORATORY TOOLS

9.1 Times Series Plots 9-1
9.2 Box Plots 9-5
9.3 Histograms 9-8
9.4 Scatter Plots 9-13
9.5 Probability Plots 9-16

CHAPTER 10. FITTING DISTRIBUTIONS

10.1 Importance of Distributional Models 10-1
10.2 Transformations to normality 10-3
10.3 Using the Normal Distribution as a Default 10-5
10.4 Coefficient of Variation and Coefficient of Skewness 10-9
10.5 Shapiro-Wilk and Shapiro-Francía normality Tests 10-13
10.5.1 Shapiro-Wilk Test (n ≤ 50) 10-13
10.5.2 Shapiro-Francía Test (n > 50) 10-15
10.6 Probability Plot Correlation Coefficient 10-16
10.7 Shapiro-Wilk Multiple Group Test of Normality 10-19

CHAPTER 11. TESTING EQUALITY OF VARIANCE

11.1 Box Plots 11-2
11.2 Levene's Test 11-4
11.3 Mean-Standard Deviation Scatter Plot 11-8

US EPA ARCHIVE DOCUMENT

CHAPTER 12. IDENTIFYING OUTLIERS

12.1 SCREENING WITH PROBABILITY PLOTS 12-1
12.2 SCREENING WITH BOX PLOTS 12-5
12.3 DIXON'S TEST 12-8
12.4 ROSNER'S TEST 12-10

CHAPTER 13. SPATIAL VARIABILITY

13.1 INTRODUCTION TO SPATIAL VARIATION 13-1
13.2 IDENTIFYING SPATIAL VARIABILITY 13-2
13.2.1 Side-by-Side Box Plots 13-2
13.2.2 One-Way Analysis of Variance for Spatial Variability 13-5
13.3 USING ANOVA TO IMPROVE PARAMETRIC INTRAWELL TESTS 13-8

CHAPTER 14. TEMPORAL VARIABILITY

14.1 TEMPORAL DEPENDENCE 14-1
14.2 IDENTIFYING TEMPORAL EFFECTS AND CORRELATION 14-3
14.2.1 Parallel Time Series Plots 14-3
14.2.2 One-Way Analysis of Variance for Temporal Effects 14-6
14.2.3 Sample Autocorrelation Function 14-12
14.2.4 Rank von Neumann Ratio Test 14-16
14.3 CORRECTING FOR TEMPORAL EFFECTS AND CORRELATION 14-18
14.3.1 Adjusting the Sampling Frequency and/or Test Method 14-18
14.3.2 Choosing a Sampling Interval Via Darcy's Equation 14-19
14.3.3 Creating Adjusted, Stationary Measurements 14-28
14.3.4 Identifying Linear Trends Amidst Seasonality: Seasonal Mann-Kendall Test 14-37

CHAPTER 15. MANAGING NON-DETECT DATA

15.1 GENERAL CONSIDERATIONS FOR NON-DETECT DATA 15-1
15.2 IMPUTING NON-DETECT VALUES BY SIMPLE SUBSTITUTION 15-3
15.3 ESTIMATION BY KAPLAN-MEIER 15-7
15.4 ROBUST REGRESSION ON ORDER STATISTICS 15-13
15.5 OTHER METHODS FOR A SINGLE CENSORING LIMIT 15-21
15.5.1 Cohen's Method 15-21
15.5.2 Parametric Regression on Order Statistics 15-23
15.6 USE OF THE 15% AND 50% NON-DETECT RULE 15-24

PART III. DETECTION MONITORING TESTS

CHAPTER 16. TWO-SAMPLE TESTS

16.1 PARAMETRIC T-TESTS 16-1
16.1.1 Pooled Variance T-Test 16-4
16.1.2 Welch's T-Test 16-7
16.1.3 Welch's T-Test and Lognormal Data 16-10
16.2 WILCOXON RANK-SUM TEST 16-14
16.3 TARONE-WARE TWO-SAMPLE TEST FOR CENSORED DATA 16-20

US EPA ARCHIVE DOCUMENT

CHAPTER 17. ANOVA, TOLERANCE LIMITS, AND TREND TESTS

17.1 Analysis Of Variance [ANOVA] 17-1
17.1.1 One-Way Parametric F-Test *17-1*
17.1.2 Kruskal-Wallis Test *17-9*
17.2 Tolerance Limits 17-14
17.2.1 Parametric Tolerance Limits *17-15*
17.2.2 Non-Parametric Tolerance Intervals *17-18*
17.3 Trend Tests 17-21
17.3.1 Linear Regression *17-23*
17.3.2 Mann-Kendall Trend Test *17-30*
17.3.3 Theil-Sen Trend Line *17-34*

CHAPTER 18. PREDICTION LIMIT PRIMER

18.1 Introduction to Prediction Limits 18-1
18.1.1 Basic Requirements for Prediction Limits *18-4*
18.1.2 Prediction Limits With Censored Data *18-6*
18.2 Parametric Prediction Limits 18-7
18.2.1 Prediction Limit for m Future Values *18-7*
18.2.2 Prediction Limit for a Future Mean *18-11*
18.3 Non-Parametric Prediction Limits 18-16
18.3.1 Prediction Limit for m Future Values *18-17*
18.3.2 Prediction Limit for a Future Median *18-20*

CHAPTER 19. PREDICTION LIMIT STRATEGIES WITH RETESTING

19.1 Retesting Strategies 19-1
19.2 Computing Site-Wide False Positive Rates [SWFPR] 19-4
19.2.1 Basic Subdivision Principle *19-7*
19.3 Parametric Prediction Limits with Retesting 19-11
19.3.1 Testing Individual Future Values *19-15*
19.3.2 Testing Future Means *19-20*
19.4 Non-Parametric Prediction limits with retesting 19-26
19.4.1 Testing Individual Future Values *19-30*
19.4.2 Testing Future Medians *19-31*

CHAPTER 20. MULTIPLE COMPARISONS USING CONTROL CHARTS

20.1 Introduction to control charts 20-1
20.2 Basic Procedure 20-2
20.3 Control Chart Requirements and Assumptions 20-6
20.3.1 Statistical Independence and Stationarity *20-6*
20.3.2 Sample Size, Updating Background *20-8*
20.3.3 Normality and Non-Detect Data *20-9*
20.4 Control Chart Performance Criteria 20-11
20.4.1 Control Charts with Multiple Comparisons *20-12*
20.4.2 Retesting in Control Charts *20-14*

US EPA ARCHIVE DOCUMENT

PART IV. COMPLIANCE/ASSESSMENT AND CORRECTIVE ACTION TESTS

CHAPTER 21. CONFIDENCE INTERVALS

21.1 PARAMETRIC CONFIDENCE INTERVALS 21-1
21.1.1 Confidence Interval Around a Normal Mean 21-3
21.1.2 Confidence interval Around a Lognormal Geometric Mean 21-5
21.1.3 Confidence Interval Around a Lognormal Arithmetic Mean 21-8
21.1.4 Confidence Interval Around an Upper Percentile 21-11
21.2 NON-PARAMETRIC CONFIDENCE INTERVALS 21-14
21.3 CONFIDENCE INTERVALS AROUND TREND LINES 21-23
21.3.1 Parametric Confidence Band Around Linear Regression 21-23
21.3.2 Non-Parametric Confidence Band Around a Theil-Sen Line 21-30

CHAPTER 22. COMPLIANCE/ASSESSMENT AND CORRECTIVE ACTION TESTS

22.1 CONFIDENCE INTERVAL TESTS FOR MEANS 22-1
22.1.1 Pre-Specifying Power In Compliance/Assessment 22-2
22.1.2 Pre-Specifying False Positive Rates in Corrective Action 22-9
22.2 CONFIDENCE INTERVAL TESTS FOR UPPER PERCENTILES 22-18
22.2.1 Upper Percentile Tests in Compliance/Assessment 22-19
22.2.2 Upper Percentile Tests in Corrective Action 22-20

APPENDICES[1]

A.1 REFERENCES
A.2 GLOSSARY
A.3 INDEX
B HISTORICAL NOTES
C.1 SPECIAL STUDY: NORMAL VS. LOGNORMAL PREDICTION LIMITS
C.2 CALCULATING STATISTICAL POWER
C.3 R SCRIPTS
D STATISTICAL TABLES

[1] The full table of contents for Appendices A through D is found at the beginning of the Appendices

US EPA ARCHIVE DOCUMENT