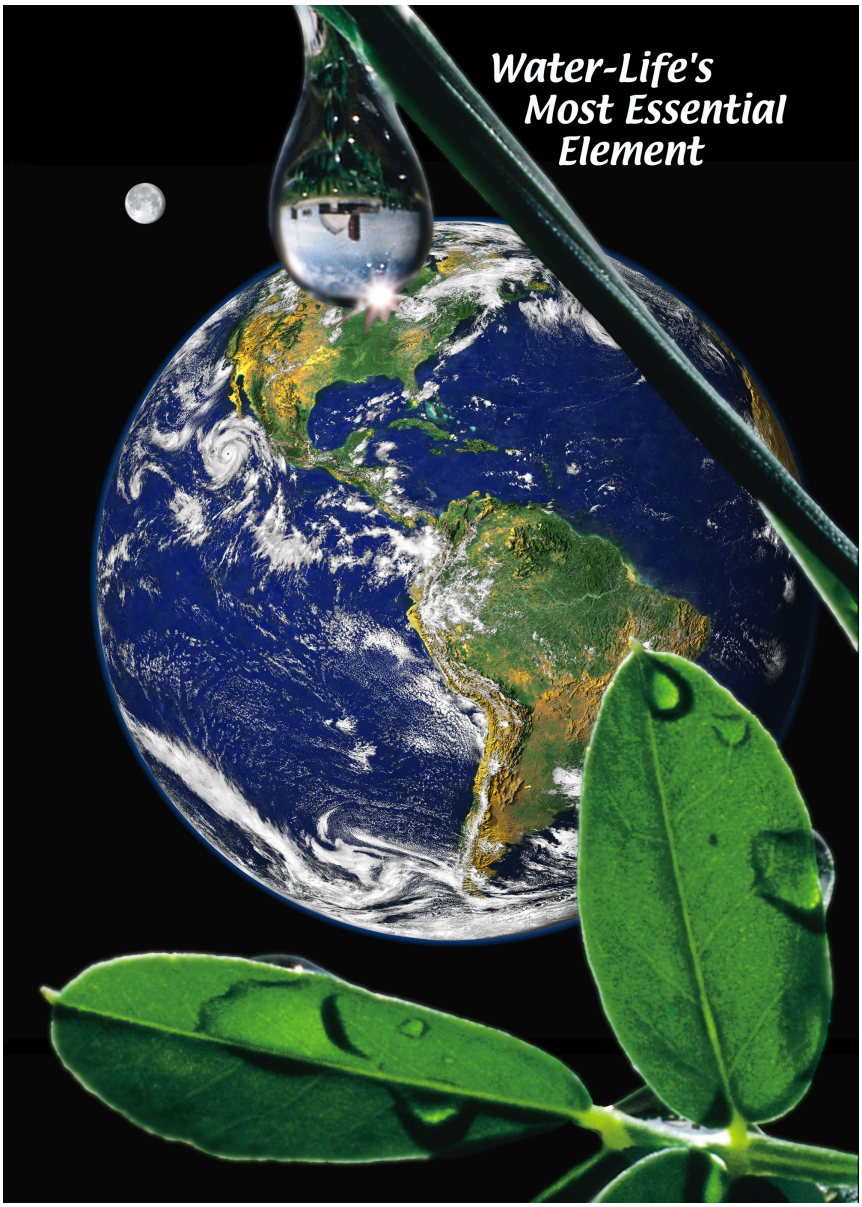
USDA United States Department of Agriculture

NRCS Natural Resources Conservation Service

September 2003

# National Water Quality Handbook



Issued September 2003

Trade names mentioned are for specific information and do not constitute a guarantee or warranty of the product by the Department of Agriculture or an endorsement by the Department over other products not mentioned.



(450–VI–NWQH, September 2003)

# National Water Quality Handbook

Contents:

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook



# Part 600 Introduction

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**



# Part 600 Introduction to National Water Quality Handbook

# Part 600 Introduction to National Water Quality Handbook

Contents:

# Part 600 Introduction to National Water Quality Handbook

## 600.0000 Purpose of handbook

Water quality is an important natural resource concern for the Nation. Being a lead natural resource technical agency, the Natural Resources Conservation Service (NRCS) has developed this handbook as a principal reference pertaining to water quality as it relates to all agricultural land uses. The handbook is the principal NRCS reference document for technical information and guidance in carrying out water quality responsibilities. This document consolidates pertinent procedures, guidelines, and other materials to facilitate finding relevant and reliable information. It provides clear guidelines for filing and cross-referencing applicable local, state, and national water quality related reference materials.

## 600.0001 Scope of handbook

The National Water Quality Handbook (NWQH) is designed to provide guidance in all aspects of water quality to NRCS personnel, Agency technical partners, and those who provide technical services to clients for NRCS. Guidance is provided to address water quality issues within the NRCS planning and implementation process. Agricultural related pollutants are addressed within this document or through references to other water quality technical materials.

Specific technical or procedural details for planning, such as conservation practice design criteria, are beyond the scope of this handbook. Detailed design information is retained in other NRCS handbooks and manuals and is referred to in appropriate sections of the NWQH. Also, water quality issues related to industrial and municipal waste pollutants are not within the scope of this document.

## 600.0002 Intended audience

The focus of this handbook is the NRCS field office, NRCS technical partners, and those providing technical services for NRCS. The NWQH includes technical and procedural guidance that is applicable at any organizational or technical level in support of NRCS water quality activities. This handbook is appropriate for basic orientation of NRCS water quality activities as well as advanced procedures for technical specialists.

## 600.0003 Structure

The National Water Quality Handbook consists of core water quality information as well as extensive cross-referencing to NRCS documents and publications and selected non-NRCS materials. Referenced documents that support and contribute to the handbook are referred to as *Key References*. The handbook leads the user through a logical sequence beginning with basic information and introductory material and progressing through planning and implementation procedures for more complex subjects. Key references are presented to allow the user to pursue more in-depth information than given in the handbook. A substantial part of the handbook is available electronically on the NRCS national Web page, (*http://www.nrcs.usda.gov*) which reflects updates, revisions, and the status of the document.

## 600.0004 Key handbook support references

This document and its specified support references are listed in section I of the Field Office Technical Guide (FOTG). Hardcopy materials of the NRCS National Water Quality Handbook and key references reside in each NRCS field office.

**Key references**

Agricultural Waste Management Field Handbook, Rev. 1

Field Office Technical Guide (FOTG) Sections I–V

Ground Water and Surface Water, A Single Resource USGS Circular 1139

National Agronomy Manual

National Engineering Handbook, Part 652, National Irrigation Guide

National Planning Procedures Handbook

Stream Corridor Restoration—Principles, Processes, and Practices (NEH Part 653)

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 601 Resource Management Framework

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 602 Introduction to Water Quality

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 603 Water Quality Assessment

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 604 Nutrients

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook



# Part 605 Pathogens

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 606 Pesticides

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 607 Temperature

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 608 Sediment

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 609 Dissolved Oxygen

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook



# Part 610 Salts

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 611 Heavy Metals

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 612 Other Surface and Ground Water Contaminants

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 613

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 614 Water Quality Monitoring System Design

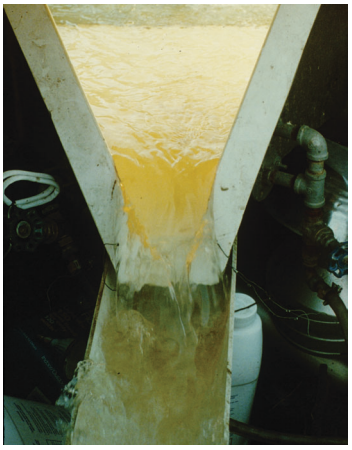
United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

National Water Quality Handbook

# Part 614

# Design of Water Quality Monitoring Systems



# Preface

## Purpose

The purpose of part 614 of the National Water Quality Handbook (NWQH) is to describe methods for monitoring the water quality response to land use and land management activities and conservation practices. These methods include how to design a monitoring study, how to set up a monitoring station, and how to analyze the water quality data. The information presented assumes that the reader has a basic understanding of water quality. A basic knowledge of statistical analysis also is useful, although part 615 of this handbook provides guidance in statistical analysis of water quality data.

Part 614 of the NWQH is needed at this time because:

- The effectiveness of programmatic activities needs to be determined. Water quality managers are constantly asking for evidence of the results of a program.
- Comprehensive guidance is needed. Many water quality managers are placed in the role of overseeing or designing monitoring projects, but a comprehensive guidance is lacking.
- Several water quality monitoring projects currently underway may require modification to show the results anticipated.

It is intended to assist those with direct or supervisory responsibilities in planning, implementing, and evaluating water quality monitoring projects.

## Structure of part 614

This part of the NWQH is formatted to directly assist in designing a water quality monitoring project. A 2-page worksheet using the steps in planning a monitoring study is at the end of chapter 1. This worksheet was organized to facilitate rapid and complete monitoring study design. Each step in the worksheet corresponds to a separate chapter in part 614. Each chapter includes examples to guide practice in applying the major concepts being described.

Part 615 of the handbook is concerned with the statistical analysis of monitoring results. It may be useful to review the introductory chapter in part 615 to perform some of the statistical operations described in part 614.

## Acknowledgments

Part 614 of the National Water Quality Handbook was written by **John C. Clausen**, Ph.D., College of Agriculture and Natural Resources, University of Connecticut. The concept for this project was developed by **James N. Krider**, former national environmental engineer, Natural Resources Conservation Service (NRCS), Washington, DC. Technical leadership was provided by **Bruce J. Newton**, limnologist, NRCS National Water and Climate Center, Portland, Oregon, and by **Frank Geter**, agricultural engineer, formerly with the NRCS Information Technology Center, Fort Collins, Colorado, and **Douglas Holy**, limnologist, Ecological Sciences Division, Washington, DC.

Many people contributed to this document. The following NRCS staff served on teams that reviewed drafts of parts 614 and 615:

**Stephanie Aschmann**, agroecologist, Watershed Science Institute, Lincoln, Nebraska
**William H. Boyd**, environmental engineer, National Water Management Center, Little Rock, Arkansas
**Richard Croft**, water quality specialist, Watershed Science Institute, Burlington, Vermont
**Carl DuPoldt**, agricultural engineer, Somerset, New Jersey
**David C. Moffitt**, envirnomental engineer, National Water Management Center, Fort Worth, Texas
**Gerald Montgomery**, biologist, Northern Plains Regional Office, Lincoln, Nebraska
**James D. Rickman**, (retired) agricultural engineer, Fort Worth, Texas
**Lynn Sampson**, biologist, East Lansing, Michigan
**Ron Schierer**, Northern Plains Regional Technical Team, Lakewood, Colorado
**Donald Stettler**, (retired) agricultural engineer, National Water and Climate Center, Portland, Oregon
**Howard Thomas**, (retired) economist, Portland, Oregon
**Barbara M. Vining**, pesticide and nutrient specialist, formally of NRCS, Indianapolis, Indiana
**James Wood**, water quality specialist, Burlington, Vermont

In addition, many attendees of a water quality monitoring workshop held in 1993 provided helpful input. They included **Jerry M. Bernard**, **Denise A. Tennessee**, **Gene Lindeman**, and **John Sutton**. Other comments were provided by **Dana Chapman**, **Lyle J. Steffen,** and **Ivan Wilkinson**.

**Joseph A. Neafsey**, water quality specialist, Storrs, Connecticut, was the contract project officer. **Mary R. Mattinson**, editor, and **Wendy R. Pierce**, illustrator, National Cartography and Geospatial Center, Fort Worth, Texas, provided editorial and graphic support.

The National Water Quality Handbook is the result of a collaborative effort of the NRCS Science and Technology Consortium. Special thanks to the following individuals for their leadership and support in the development and publishing of this essential technical reference:

**Wil Fontenot**, natural resource specialist, Lafayette, Louisianna
**Jon Werner**, national hydrologist, Conservation Engineering Division, Washington, D.C.
**Roy Mann**, (retired) resource conservationist, Portland, Oregon
**Ken Pfeiffer**, pest management specialist, National Water and Climate Center, Portland, Oregon
**Bruce Newton**, acting director, National Water and Climate Center, Portland, Oregon
**Lynn Betts**, communications director, Wildlife Habitat Management Institute, Des Moines, IA

# Part 614 National Water Quality Handbook

Contents:

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 1 Introduction

# Chapter 1 Introduction

Contents:



Tables

**Figures**



**Example**

# Chapter 1 Introduction

## 614.0100 General

Recognition of agriculture's contribution to nonpoint source (NPS) pollutant loadings to streams, lakes, estuaries, and ground water has led to increased emphasis on water quality monitoring in rural watersheds. Conservation Districts and the Natural Resources Conservation Service (NRCS) are often sponsors and cooperators, respectively, of studies and projects to reduce agricultural NPS loadings. The primary purpose of this handbook is to provide these entities and their partners with guidance for gathering and using water quality information to support planning and implementation activities.

Although opinions vary about the value of water quality monitoring, there is consensus that monitoring is relatively expensive. Therefore, it is imperative that monitoring be well designed. As stated by Ward, et al. (1986), appropriate designs of monitoring systems are needed to prevent a "data rich, but information poor" monitoring system. Part 614 of this handbook primarily addresses the design of intensive monitoring programs. Part 615 addresses the analysis of monitoring data to enable us to refine our understanding of water quality.

For most projects that involve water quality concerns, the NRCS planning process requires information obtained by monitoring to perform the planning steps. Current and historical data are needed to perform Phase I, which includes identifying problem areas, determining objectives and setting goals, inventorying resources, and analyzing resource data. The results of Phase I work are used in Phase II to formulate and evaluate alternatives and decide on a plan. Phase III, implementation and evaluation, requires water quality information collected through time to evaluate the effectiveness of the implemented alternative.

The collection of water quality information is extremely important as we learn how to address water quality resource concerns. Adaptive management requires that we observe the effects of natural resources management decisions so we can maximize learning and increase the knowledge base for future natural resources management decisions. Even during studies, data could be used to calibrate and refine planning tools, such as computer models. The success of such efforts should eventually reduce the need for costly water quality monitoring in the future.

State water quality agencies are generally most active in assisting local water quality monitoring. At the Federal level, the Office of Management and Budget has directed agencies to coordinate their data acquisition efforts with the U.S. Geological Survey (USGS)(OMB Circular M-92-01). The local USGS office should be involved in the design of project water quality monitoring.

# 614.0101 Definitions

The term water quality is used throughout this guide, so a definition is appropriate. Although many definitions for this term exist (APHA, et al. 1969; Rechard and McQuisten 1968; Veatch and Humphreys 1966), water quality can be broadly defined as the physical, chemical, and biological composition of water as related to its intended use for such purposes as drinking, recreation, irrigation, and fisheries.

The term water quality has different meanings to different users of the water, which can result in confusion among water quality managers. The term may be applied to a single characteristic of the water or to a group of characteristics combined into a water quality index.

A few other terms related to water quality are important to define.

***Water quality management*** can be defined as the management of the physical, chemical, and biological characteristics of water (Sanders, et al. 1983).

***Water quality monitoring,*** one function of water quality management, is the collection of information on the physical, chemical, and biological characteristics of water (Sanders, et al. 1983).

***Pollution*** refers to a condition of water within a water body caused by the presence of undesirable materials (APHA, et al. 1969).

***Contamination*** is the introduction of substances into water at a sufficient concentration to make the water unfit for its intended use (APHA, et al. 1969).

***Pollution control*** generally is associated with the regulation of pollutants.

# 614.0102 Monitoring purposes

Monitoring of water quality can serve many purposes. Each purpose is described using relevant examples.

## (a) Analyze trends

Monitoring on a regular basis has been used to determine how water quality is changing over time. A widely publicized example of trend analysis was that published by Smith and Alexander (1983) on stream chemistry trends at the USGS benchmark stations. Trend analysis was also used in several of the Rural Clean Water Program (RCWP) projects in the United States, including those in Vermont, Idaho, and Florida.

Monitoring of so called "baseline" conditions also has been used and is often recommended. Baseline generally is thought of as a pre-condition; that is, what the water quality conditions are that currently exist. Caution is recommended in using baseline monitoring. Unless such data are used for reconnaissance purposes or actually are the beginning of trend analysis, then baseline monitoring is not recommended except where the effects caused by climate are controlled in the design of the project. If, for example, the baseline data were collected during an abnormal year, the data could be biased.

## (b) Determine fate and transport of pollutants

Monitoring also is conducted to determine whether a pollutant may move and where it may go. For such projects, monitoring over a long period may not be needed. For example, if the objective is to determine whether a pesticide is leaving the root zone, a short-term (<5 years) study of intensive sampling would be sufficient.

Fate and transport studies typically require frequent sampling of all possible transport pathways in a relatively small area. These studies also are subject to climate influences and may require sophisticated sampling equipment.

## (c) Define critical areas

Water quality monitoring has been used to locate areas within watersheds exhibiting greater pollution potential than other areas. The results of such monitoring can then be used to target Resource Management Systems (RMSs). This type of monitoring has often been termed *reconnaissance* monitoring.

Targeting critical areas also could occur following interpretation of water quality data collected early in a project. For example, monitoring in a particular watershed could indicate that one of the subwatersheds may have the highest phosphorus concentrations and export as compared to the other monitored sub-watersheds. Supplemental investigation may reveal the source of the phosphorus, either natural or related to management. Based on these early findings from monitoring data, priority could be given to that subwatershed for implementation of RMSs.

Reconnaissance monitoring however, is generally conducted over a short time frame, and caution should be exercised to assure that decisions regarding targeting are not biased by unusual climate conditions during the period of monitoring.

## (d) Assess compliance

Water quality monitoring frequently has been used to determine compliance with water quality plans and standards. For example, bacteria monitoring has been used to determine the percentage of the time bacteria levels exceed a standard, such as 200 organisms per 100 milliliter. Compliance monitoring should consider climate conditions as well as the ability to link instream levels with actual sources before taking action.

## (e) Measure effectiveness of conservation practices

Monitoring to determine the effectiveness of individual conservation practices is typically conducted on a plot or field scale, or as close as possible to the practice. Water quality studies of individual practices can be conducted in a relatively short time frame (<5 years). However, some practices may take many years to show results.

An example of monitoring to assess the effectiveness of a conservation practice would be sampling above and below a filter strip being used to treat feedlot runoff. Another example of a practice suitable for monitoring would be field nutrient management, in which case, sampling of both the field soils and the field runoff would be conducted.

## (f) Evaluate program effectiveness

Water quality monitoring used to evaluate the effectiveness of a program in a watershed (e.g., Hydrologic Unit Areas, HUAs) is generally conducted on a watershed scale. Several land uses would probably be within the watershed. RMSs, implemented as a result of a water quality program, would most likely be staggered over time and managed with varying vigor. Monitoring for program effectiveness would be conducted over the long-term (>5 years).

Monitoring the effectiveness of a program is difficult because of the lack of control over exactly what happens and when it happens. Also, the staggering of events will most likely compensate each other. Finally, water quality responses to changes in practices may be gradual and take many years because of the buildup of the pollutant of concern in the watershed.

## (g) Make wasteload allocations

Monitoring of receiving water bodies would be needed to perform wasteload allocations. Though typically thought of for point sources, wasteload allocations are used in some parts of the United States for both point and nonpoint sources (e.g., Oregon). Monitoring could be used to determine how much additional (or less) agriculture or what conservation practice could be allowed in a watershed without exceeding a certain level or tropic state in a water body.

Monitoring to allocate loads from different sources requires a good knowledge of the actual contributions from the sources. For nonpoint sources, extensive monitoring may be needed to determine the actual source.

### (h) Model validation and calibration

Water quality monitoring may be needed to validate or calibrate models to local conditions. Also, it is used to verify a model's adequacy. In such tests, the values predicted by the model are compared to values observed by monitoring.

A major difficulty in model validation is that many models are developed to simulate long-term average conditions; whereas, most monitoring data are collected on a relatively short-term basis. In addition, many of the input variables used in a model, such as the hydraulic conductivity or wind speed, typically are not monitored.

### (i) Conduct research

Water quality monitoring is necessary for addressing specific research questions. An example would be a comparison of nitrate concentrations obtained from samples using various types of lysimeters including suction plate, porous cup, and zero-tension types. Such monitoring would normally be conducted by a research agency or university. The difference between research monitoring and other purposes of monitoring often is not great. However, research monitoring is not the purpose of this handbook.

### (j) Define water quality problem

Although discussed elsewhere in this guide, water quality monitoring may be required to give adequate definition to the water quality problem. For example, if a fishery is impaired in a water body, water quality monitoring will be needed to determine the cause of the impairment. Possible causes might include sediment, toxins, reduced dissolved oxygen, or temperature problems, to name a few.

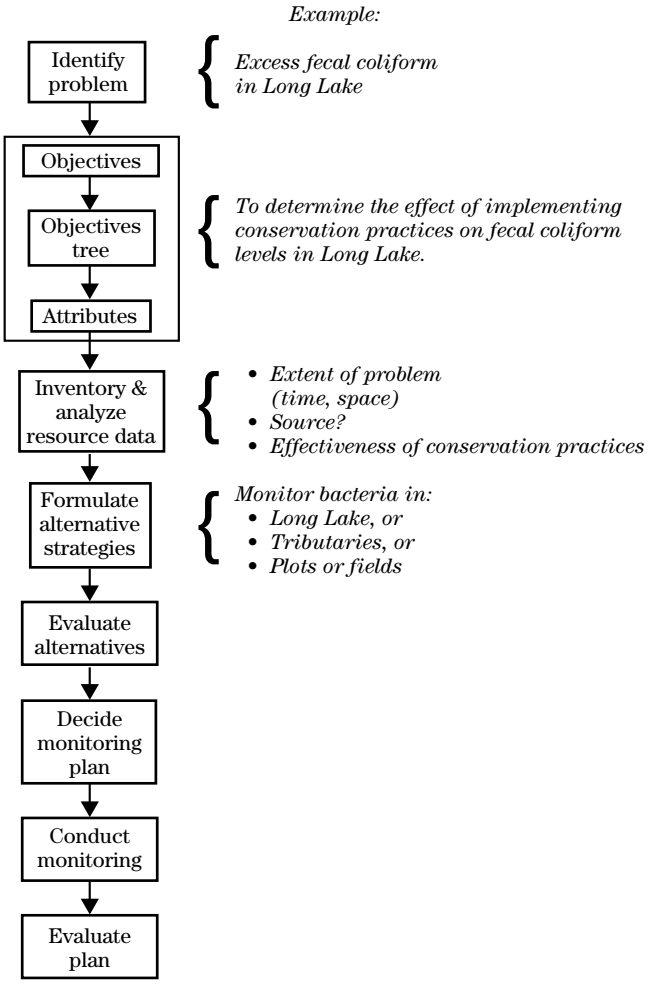
If monitoring to better define the water quality problem, the appropriate water quality characteristics must be monitored.

## 614.0103 Monitoring study design

Many outlines for developing a monitoring study have been made (Canter 1985; Ponce 1980; Sanders, et al. 1983; Solomon and Avers 12987; Tinlin and Everett 1978; Ward, et al. 1990; Whitfield 1988).

Water quality monitoring, like other tasks, can be viewed in a decisionmaking or planning context that begins with a definition of the problem and ends with an evaluation of the effectiveness of the plan (fig. 1–1).

**Figure 1–1** Steps in decisionmaking for a water quality monitoring system



This framework is similar to the 9-Step Planning Process (USDA-SCS 1993), although that process is primarily aimed at developing and implementing conservation practices. In some cases it may be desirable to develop the water quality monitoring plan within the context of the 9-Step Planning Process. The steps are:

Step 1 Identify problems
Step 2 Determine objectives
Step 3 Inventory resources
Step 4 Analyze resource data
Step 5 Formulate alternatives
Step 6 Evaluate alternatives
Step 7 Make decisions
Step 8 Implement plan
Step 9 Evaluate plan

This handbook uses 12 steps for developing a monitoring study (fig. 1–2). Chapters 2 through 13 describe these steps in detail. The complexity of each step varies with the type of system being designed; however, each step should be addressed for all monitoring projects.

The first step, defining the water quality problem, is necessary to assure that monitoring actually matches the problem. Setting objectives for monitoring clarifies the purposes of the project and keeps it on track. Knowledge of the overall project objectives assures that monitoring is consistent with the implementation goals. The statistical design is needed as an overall framework to ensure that the samples are being collected from the appropriate locations. The monitoring design must also include the scale of the project (plot, field, or watershed); the type of sample; the variables and locations to sample; and the frequency and duration of sampling. The type of monitoring station and its construction should be defined. The methods for collecting land use and management data need to be described, including how the water quality data and land use data will be linked. Finally, a system for managing the data should be described.

The 12 steps for developing a water quality monitoring design are similar in some ways to the 9-Step Planning Process. Water quality monitoring can be used to identify resource problems (step 1), formulate alternatives (step 5), and evaluate the effectiveness of the plan (step 9). In a side-by-side comparison, the first two steps of each method are analogous. Step 1 identifies problems, and step 2 determines objectives. The remaining steps in water quality monitoring design are included in step 3 of the 9-step process, which is to inventory resources. In actual practice, both frameworks would most likely be considered by the water quality specialist.

Example 1–1 is a case study for developing a water quality monitoring plan using the 12 water quality monitoring design steps. This case study is of the St. Albans Bay Rural Clean Water Program project in Northwestern Vermont (fig. 1–3). This project was one of 21 in the nation and one of 5 comprehensive monitoring and evaluation projects active from 1980 to 1990 (Cassell, et al. 1983). It contains physical, chemical, and biological monitoring.

**Figure 1–2** Steps in water quality monitoring system design

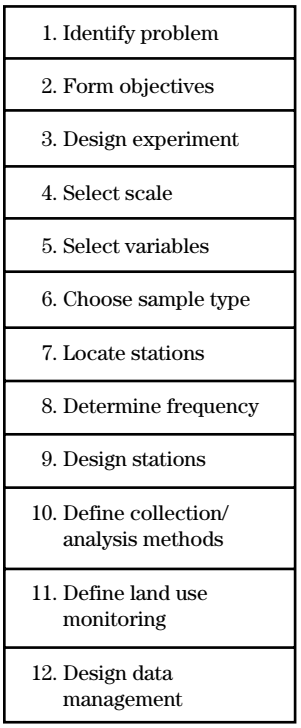
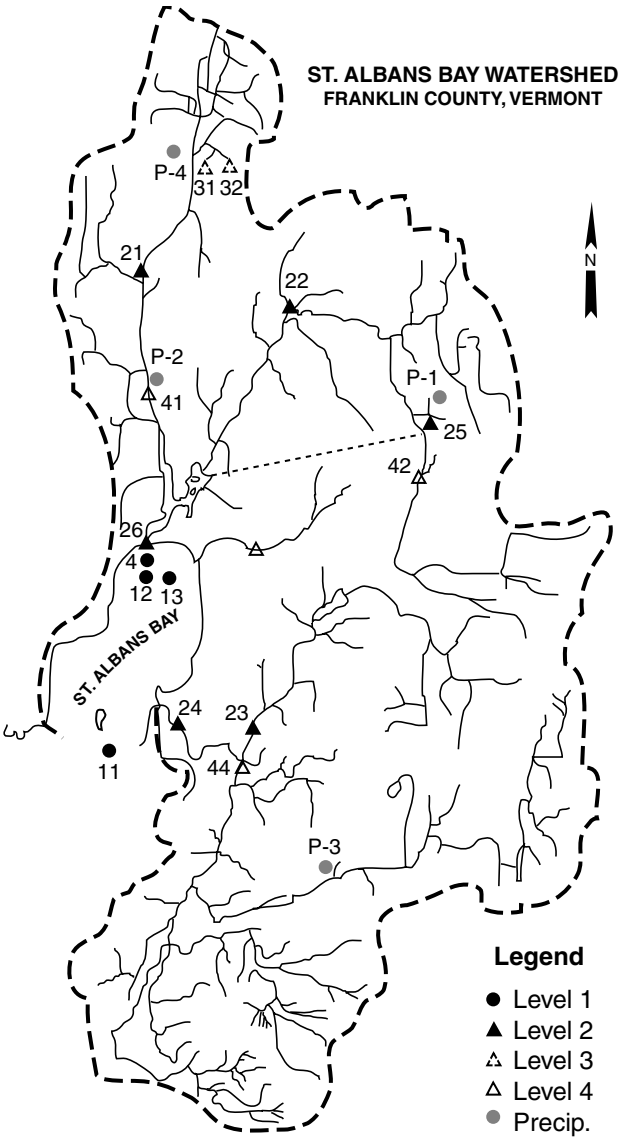


**Figure 1–3** St. Albans Bay watershed



**Example 1–1** Case study—St. Albans Bay RCWP

**Step 1 Water quality problem**

Recreation within St. Albans Bay was impaired because of excessive eutrophication. Also, a state park had closed because of reduced attendance associated with frequent beach closings resulting from coliform bacteria standard violations. A 1-year reconnaissance monitoring project by the state natural resource agency determined that both bacteria and phosphorus were coming from both point (wastewater treatment plant) and nonpoint (agricultural) sources.

**Step 2 Objectives**

Several monitoring objectives were defined:

- To document changes in the water quality of specific tributaries within the watershed resulting from implementation of manure management practices.
- To measure the changes in the amount of suspended sediment and nutrients entering St. Albans Bay resulting from implementation of water quality management programs within the watershed.
- To evaluate trends in the water quality of St. Albans Bay and the surface water within the St. Albans Bay watershed during the period of the RCWP Watershed Project.

Additional objectives were developed to address special projects in the study area. They included:

- To determine the role of an existing wetland, located between the point and nonpoint sources and the Bay, on the quality of water entering St. Albans Bay.
- To determine the role of Bay and wetland sediment on the quality of St. Albans Bay.
- To determine the effect of Bay circulation on the quality of St. Albans Bay.
- To determine the effect of individual BMPs, especially manure management, on exports to the Bay.
- To determine the effect of implementation of BMPs on aquatic organisms in the Bay and tributaries.

**Step 3 Statistical design**

Many statistical designs were used to meet the objectives. These designs were associated with four levels of study:

Level 1: Bay monitoring
Level 2: Tributary monitoring
Level 3: BMP monitoring
Level 4: Supplemental tributary monitoring

The primary statistical approach for the level 1 and 2 monitoring was trend analysis of data collected at each Bay (4) and tributary (4) station. In addition, since BMPs were not implemented at the same rate or intensity throughout the project area, paired regressions between tributary and bay stations were also used. An above-and-below paired

**Example 1–1** Case study—St. Albans Bay RCWP—Continued

watershed study was used for the level 3 monitoring. These types of statistical approaches are described in chapter 3 of this handbook. The level 4 monitoring had no statistical basis and was later dropped. There was no control watershed in the study area to serve as a hydrologic comparison for the treated watersheds. This lack of a control was found to be an important deficiency.

**Step 4 Scale of study**

The scale varied with the level of monitoring. Level 1 Bay stations were points along a nutrient gradient in the Bay. Level 2 and 4 tributary stations were of watershed scale ranging from 3,900 to 8,800 acres in area. The level 3 BMP monitoring used a field scale. The wetland study used point scale for samples within the wetland and a watershed scale for the wetland outlet. Sediment and circulation monitoring used point scales.

**Step 5 Variables selection**

The variable selected for study also varied with the level of study (table 1–1).

**Table 1–1** Variables monitored for the St. Albans Bay project

| Variable | Levels |
|---|---|
| Turbidity | 1, 2, 4 |
| Total suspended solids | 1 – 4 |
| Volatile suspended solids | 1 – 4 |
| Total phosphorus | 1 – 4 |
| Ortho-phosphorus | 1 – 4 |
| Ammonia-nitrogen | 1 – 4 |
| Total Kjeldahl nitrogen | 1 – 4 |
| Nitrate-nitrogen | 1 – 4 |
| Chlorophyll a | 1 |
| Fecal coliform | 1, 2, 4 |
| Fecal streptococcus | 1, 2, 4 |
| Temperature | 1, 2, 4 |
| Dissolved oxygen | 1, 2, 4 |
| pH | 1, 2, 4 |
| Conductivity | 1, 2, 4 |
| Secchi disc | 1 |
| Flow | 2, 3, 4 |
| Chloride | Wetland |
| Fish populations | 2 |
| Invertebrates | 2 |
| Periphyton | 2 |
| Precipitation | |

**Example 1–1** Case study—St. Albans Bay RCWP—Continued

**Step 6 Sample type**

The type of sample varied with the level of monitoring (table 1–2)

**Table 1–2** Sample types for the St. Albans Bay Project

| Level | Sample type |
|---|---|
| 1 | Grab - 2 depths<br>plankton - depth integrated |
| 2, 3 | time composite at point<br>grab - bacteria |
| 4 | grab |
| Wetland | grab<br>time composite at outlet |

**Step 7 Sampling location**

Sampling locations for all levels are shown in figure 1–3. Originally, three stations were located in St. Albans Bay. One station was associated with the closed beach; the other two represented an inner and outer bay component. A fourth station was added in the fourth year of the project to better characterize the nutrient gradient in the bay following the procedures described by Potash and Henson (1978). At each bay station, samples were taken at two points: one at the surface and one near the bottom. In addition, the extent and type of macrophyte growth were determined annually using aerial photography and a field survey.

Level 2 tributary stations were located along the four major tributaries to the bay at the lowest possible accessible site that passed a site selection criteria test. Samples were automatically collected in a tube at a single point at each cross section. Level 2 biological monitoring was conducted at the level 2 stations.

Two level 3 BMP stations were located with a ditch that drained two adjacent fields (fig. 1–4). The stations were located one up stream of the other, with the upper station serving as the control. At each station, samples were automatically collected in a tube at a point in the cross-section.

Level 4 stations were located at four tributaries as close to the bay as possible, and 15 wetland samples were located along stream channels at equal spacing. Additional wetland samples were located in the bay to better define a gradient (fig. 1–5).

**Example 1–1** Case study—St. Albans Bay RCWP—Continued

**Figure 1–4** Level 3 paired watershed

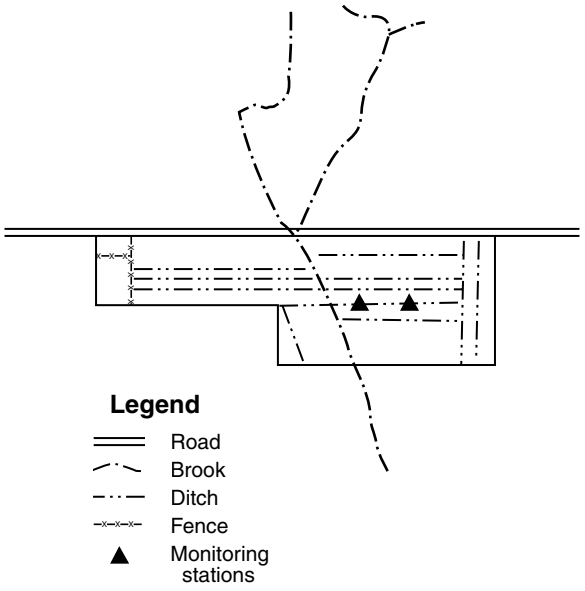
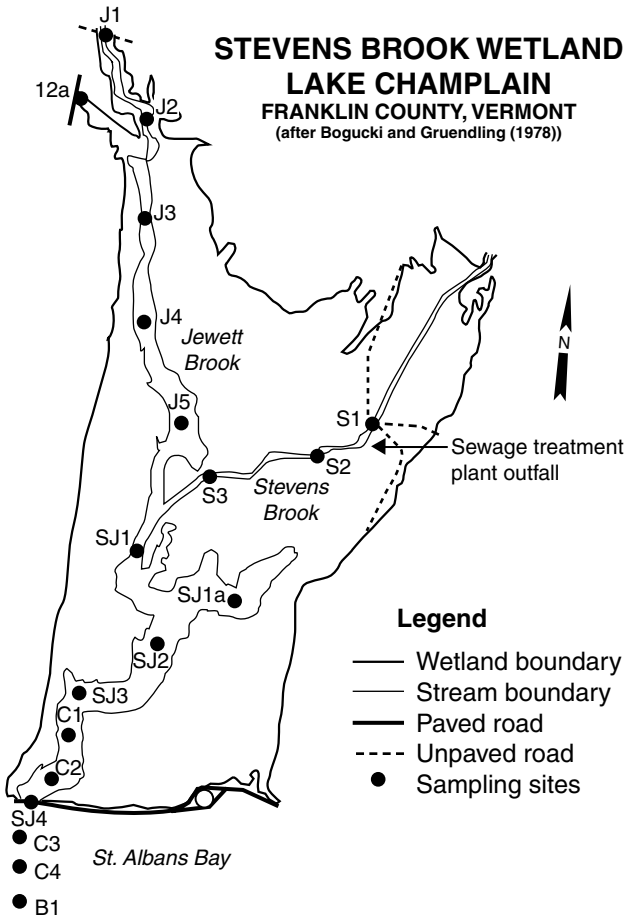



**Figure 1–5** Wetland sampling locations



**Example 1–1** Case study—St. Albans Bay RCWP—Continued

**Step 8 Sampling frequency and duration**

The number of samples collected also varied with the level of monitoring (table 1–3). The project was designed for a 10-year time frame.

**Table 1–3** St. Albans Bay monitoring frequency

| Level | Frequency |
|---|---|
| 1 | monthly (Oct – Apr)<br>biweekly (May – Jul)<br>weekly (Aug – Sep) |
| 2 | Two 48-hour and one 72-hour composite/week from 8-hour samples |
| bacteria | weekly |
| 3 | 4 hr composites |
| 4 | every 20 days |
| biological | every 5 years |
| periphyton | 3 times per week |
| benthos | 2 times per year |
| fish | 2 times per year |

**Step 9 Station type**

The type of station used varied with the level of sampling. Level 1 sampling was conducted at reference points in the Bay. A Kemmerer sampler was used to collect water samples. A Wisconsin sampling net was used to obtain plankton samples.

The level 2 stations were permanent structures located adjacent to the streams. Each station was heated, had 110 VAC power, but ran on batteries. Bubbler-type stage-height recorders and automatic samplers were used. Stilling wells were added to most stations.

The level 3 stations were temporary installations in field ditches that included a sharp-crested 120 degree v-notch weir, bubbler gage, and automatic sampler. The stations were heated with propane gas.

The level 4 sampling stations were grab sites as were the biological monitoring sites. Periphyton was collected on plastic slides. A Surber sampler was used to collect benthos in riffles. Hester-Dendy samplers were also used. Block nets and a back-pack electrofisher were used to collect fish samples.

**Example 1–1** Case study—St. Albans Bay RCWP—Continued

| | | |
|---|---|---|
| **Step 10** | **Sample collection and analysis** | Sample collection, preservation, and analysis followed EPA guidelines (USEPA 1983). Automatic samples were collected in tubing with a peristaltic pump and stored in acid-washed, distilled water rinsed bottles in refrigerated samplers. Bacteria samples were collected in sterilized bottles. Samples were preserved with acid and analyzed within EPA recommended holding times (USEPA 1983). A quality assurance and quality control plan was developed, and the success of quality control was reported quarterly. Field test kits were generally not used; however, in situ analysis was made of dissolved oxygen and conductivity. Daily field sheets were used, and each technician used individual field books. |
| **Step 11** | **Land use and management monitoring** | An elaborate program of land use and management monitoring was used in this study. A daily field log developed for each farm was left with the landowners. Twice each year the farm was visited, the logs were picked up, and any missing data were reconstructed. Data were collected on a field-by-field basis and included the date, amount, and type of applications of manure, fertilizer, and pesticide. In addition, baseline information was collected on soils, topography, stream courses, roads, and farm and field boundaries. Livestock numbers were also tracked for each farm. Annually, 35mm slides obtained from the Agricultural Stabilization and Conservation Service (ASCS) were consulted for land use changes in areas where land use data were missing. These flyovers include only cropland as part of program compliance by ASCS.<br><br>The entire system was managed in a Geographic Information System (GIS). Maps and tables were used to track land use and management activities, such as where manure was applied and whether it was incorporated. |
| **Step 12** | **Data management** | A computer-based data management system, Bayqual, was developed specifically for the project. Water quality and precipitation data were manually entered into the computer. Stage charts were digitized. All data were stored on a VAX computer with backup on a mainframe computer. Currently data are archived in both paper and computer disk format. Statistical analysis was conducted first on mainframe and then on PC computers. The PC revolution occurred in the middle of the project, and a general transfer of many data management activities to PC's occurred.<br><br>Data entry included a validation process that involved double-entry with an error checking program. Tests of reason were also programmed, such as the impossibility of orthophosphorus exceeding total phosphorus. Summaries of the data were presented quarterly and annually at project meetings. Written reports were also provided. This frequent reporting was found to be highly useful. |

# 614.104 References

American Public Health Association, American Society of Civil Engineers, American Water Works Association, and Water Pollution Control Federation. 1969. Glossary of water and wastewater control engineering.

Canter, L.W. 1985. River water quality monitoring. Lewis Publ., Inc., Chelsea, MI.

Cassell, E.A., J.C. Clausen, R.C. Furbush, M. Long, and D.W. Meals, Jr. 1983. St Albans Bay Watershed comprehensive monitoring and evaluation and socioeconomic monitoring and evaluation. Report No. 1, Background information. Vermont Water Resources Research Center, Burlington, VT.

Office of Management and Budget. 1992. Coordination of water data information. OMB Cir. M-92-01, Washington, DC.

Ponce, S.L. 1980. Water quality monitoring programs. USDA Forest Serv., Watershed Sys. Dev. Group Tech. Pap WDG-TP-00002, Fort Collins, CO.

Potash, M., and E.B. Henson. 1978. Significance of geographic positioning for water quality monitoring in Lake Champlain. Proc. Establishing water quality monitoring programs. Amer. Water Resourc. Assoc., pp. 58-73.

Rechard, P.A., and R. McQuisten. 1968. Glossary of selected hydrologic terms. Water Resourc. Res. Inst., Water Resourc. Serv. No. 1, Univ. WY, Laramie.

Sanders, T.G., R.C. Ward, J.C. Loftis, T.D. Steele, D.D. Adrian, and V. Yevjevich. 1983. Design of networks for monitoring water quality. Water Resourc. Pub., Littleton, CO.

Smith R.A., and R.B. Alexander. 1983. Evidence for acid-precipitation-induced trends n stream chemistry at hydrologic benchmark station. U.S. Geol. Surv. Circ. No. 910, Alexandria, VA.

Solomon R.M., and P.E. Avers. 1987. A water quality monitoring framework to satisfy legal requirements. Proc. Symp. on monitoring, modeling, and mediating water quality, Amer. Water Resourc. Assoc., Bethesda, MD, pp. 231-242.

Tinlin, R.M., and L.G. Everett. 1978. Establishment of ground water quality monitoring programs. In Establishment of water quality monitoring programs. Proc. Amer. Water Resourc. Assoc., Minneapolis, MN.

United States Department of Agriculture, Soil Conservation Service. 1993. National Planning Procedures Handbook, part 600.2. Washington, DC.

United States Environment Protection Agency. 1983. Methods for chemical analysis of water and wastes. EPA 600/4-79-020, Off. Res. and Dev., Cincinnati, OH.

Veatch, J.O., and C.R. Humphrys. 1966. Water and water use terminology. Thomas Print & Publ. Co., Kau Kauna, Wi.

Ward, R.C., J.C. Loftis, and G.B. McBride. 1986. The "data-rich but information poor" syndrome in water quality monitoring. Environ. Manag. 10(3):291-297.

Ward, R.C., J.C. Loftis, and G.B. McBride. 1990. Design of water quality monitoring systems. Van Nostrand Reinhold, New York, NY.

Whitfield, P.H. 1988. Goals and data collection designs for water quality monitoring. Water Resourc. Bul. 24(4):775-780.

## Steps In Planning A Water Quality Monitoring System

### Project Title _____

**1. Water Quality Problem**

_____

_____

**2. Objectives**

Project: _____

_____

_____

Monitoring: _____

_____

_____

_____

_____

**3. Statistical Design**

Plot _____ Above and below _____ Paired _____

Multiple _____ Trend _____

**4. Study Scale**

Stream: Plot _____ Field _____ Watershed _____

Ground water: Plot _____ Field _____ Watershed _____

Lake: Limnocorral _____ Bay _____ Lake-wide _____ Outlet _____

**5. Variables**

_____ _____

_____ _____

_____ _____

**6. Sample Type**

Grab _____ Composite _____ Integrated _____

Continuous _____ Time _____ Flow _____

## Steps In Planning A Water Quality Monitoring System (continued)

**7. Sampling Location**

Water body: _____ Location: _____

Water body: _____ Location: _____

Water body: _____ Location: _____

**8. Sampling Frequency and Duration**

n = _____ per _____ Duration _____

**9. Station Type**

Discharge _____ Concentration _____

Precipitation _____ Other _____

**10. Sample Collection and Analysis**

Preservation _____ Lab methods _____

Field methods _____

**11. Land Use And Management**

Monitoring method _____

_____

Data management _____

_____

Relating land treatment to water quality _____

_____

**12. Data Management**

Storage system _____

Validation _____

_____

Reporting frequency _____

By: _____ Date: _____

United States
Department of
Agriculture

**Natural
Resource
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 2 Water Quality Problem

# Chapter 2 Water Quality Problem

Contents:



Tables

# Chapter 2 Water Quality Problem

## 614.0200 Introduction

The first step in developing a water quality monitoring study is to define the water quality problem. The definition of the water quality problem is normally conducted before the design of the monitoring project. However, a redefinition or clarification of the water quality problem may often result as a monitoring design is developed or during actual monitoring.

In some cases a definite water quality problem may not exist, but rather a trend toward an emerging water quality problem is being monitored. For example, in Nebraska, monitoring of ground water nitrate concentrations has been used to identify trends toward exceeding a standard (Ehrman, et al. 1990). Chapter 2 describes defining the water quality problem. The *Water Quality Indicators Guide* by Terrell and Perfetti (1989) may be useful in using biological and habitat approaches to identify surface water quality problems.

## 614.0201 Characteristics

In formulating a water quality problem statement, the difference between a problem and a symptom needs to be distinguished. A water quality *problem* is a water quality issue requiring a solution, often stated in the form of a question. A *symptom* is a characteristic or condition of a water body indicating a problem or cause of the problem. For example, a poor fishery might be symptomatic of a sediment or dissolved oxygen problem. Excessive algal blooms might be symptomatic of excessive nutrient loadings. Every water quality problem typically has several symptoms.

The problem statement should be written in terms of a *use impairment.* Uses may include contact recreation, aesthetics, irrigation, fishing, or drinking. Ecological integrity is increasingly thought of as a use by some.

An indication of the impaired water body also helps to clarify the water quality problem statement. The type of water body could be described generically (e.g., lake, estuary, stream, vadose zone, ground water) or more specifically by name (e.g., Lucky Lake). Finally, identification of the cause of the problem and the source of that cause lend further definition to the problem statement. Table 2–1 summarizes some typical symptoms and problems and lists typical use impairments. Example water bodies are also summarized.

**Table 2–1** Water quality symptoms and problems

| Symptom | Problem | Use impairment | Water body | Cause | Source |
|---|---|---|---|---|---|
| Color | Algae, sediment, organic acids | Drinking | Lake | Erosion | Fields |
| Excess algae | Nutrients | Aesthetics | Lake | P, N | Animal waste |
| Excess macrophytes | Nutrients, abundant light | Recreation | Lake | P | Fertilizers |
| Hypoxia | Nutrients | Fishing | Estuary | N | Wastewater |
| Low biotic diversity | Toxics, nutrients | Fishing | Bay | PCB, pesticides | Contaminated sediment |
| Taste | Salinity, algae, metals | Drinking | Ground water | Salts | Geologic formation |
| Turbidity | Algae, sediment | Irrigation | Stream | Erosion | Return flows |

## 614.0202 Syntax

Based upon the characteristics of a water quality problem, a syntax for developing a water quality problem statement can be given. Thus, the water quality problem statement should include information about the problem, the use impairment, the specific water body, the cause of the problem, and the source of the causal agents. A suggested syntax for writing a water quality problem statement is:

**problem +**
**impaired use +**
**water body +**
**cause +**
**source**

A good example of a definition of a water quality problem is:

*The lack of recreation in St. Albans Bay is because of eutrophication caused by excessive phosphorus loading from agricultural sources.*

The problem has been stated with sufficient clarification to set monitoring and project objectives. The water quality problem is identified as eutrophication. A symptom of that problem, although not stated, might be algal blooms. The water body is St. Albans Bay. The cause identifies the driving factor for eutrophication, which in this case is phosphorus. A more complete discussion of causality is in part 615 of this handbook. Finally, the source of the pollutant is identified as agricultural in this case.

In many cases the actual source of the pollutant or the actual cause of the problem may not be known when designing the monitoring study. This is often the case where water quality data are limited or do not exist. In such cases the statement of the water quality problem may need to include some uncertainty. For example:

*The lack of recreation in St. Albans Bay is because of excess nutrients (N or P) from unknown sources.*

Another limitation may be knowledge of causality for the problem. The problem may be so new that a causal relationship has not been developed yet. As described in the preface, the actual purpose of monitoring may be to determine the source of the problem.

On the other hand, an example of a poor definition of a water quality problem is:

*Bad fishing.*

For this example, the real problem is unknown. Is fishing poor because of toxics, dissolved oxygen, sediment, food, or some other causal factor? Also, what is the source of the problem contributing to the causal factor? Therefore, to adequately define the problem, some knowledge of the condition of the resource must be available. Some data are needed. The problem must also be of a scale that is addressable by the project. For example, a study on a small plot in the watershed of a large lake will not allow determining whether the water quality problem of the lake has been corrected, but may address a water quality problem in a tributary to the lake.

The absence of a proper statement of the water quality problem is a common impediment to proper design and execution of a water quality monitoring study.

# 614.0203 References

Ehrman, R.L., M.L. Link, J.J. Gottala. 1990. Special protection areas: A new nonpoint source management option in Nebraska. J. Soil and Water Conserv. 45(2):263-264.

Terrell C.R., and P.B. Perfetti. 1989. Water quality indicators guide: Surface waters. USDA Soil Conserv. Serv., SCS-TP-161, Washington, DC.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 3 Objectives

# Chapter 3 Objectives

Contents:

# Chapter 3 Objectives

## 614.0300 Introduction

The second step in developing a water quality monitoring study, after defining the water quality problem, is to define the monitoring objectives. The objectives of a monitoring study must address the water quality problem. A well thought out objective or set of objectives drives the rest of the monitoring study design and is critical to a successful monitoring project. This chapter presents methods for formulating objectives and gives several examples of objectives. In addition, a process for organizing a multitude of objectives is provided.

Unfortunately, two types of objectives emerge when planning a monitoring project: management objectives and monitoring objectives. *Management objectives* refer to the goals of the project that monitoring is intended to assess. *Monitoring objectives* refer to obtaining knowledge about the system. Often these two types of objectives become confused; yet, both are important to the success of the project. Therefore, both types of objectives are presented in this chapter.

Setting objectives can be viewed as a series of three steps:

- Identifying the objective
- Developing an objective hierarchy
- Specifying attributes to measure the level of achievement of these objectives

## 614.0301 Forming objectives

Much time has been devoted to debating the differences among objectives, goals, and purposes. Although the distinction between goals and objectives has been made, the differences are subtle to most but the academician (Dickerson and Robershaw 1975, Keeney 1988, Keeney and Raiffa 1976). Therefore, for the purposes of this handbook, all these terms are grouped under the term *objective*.

### (a) Monitoring objectives

In general, an objective describes the answer to the following question: "What must be done?" It also states what is desired to accomplish. By definition, an objective includes an object as part of the statement. A useful syntax for writing an objective is:

*infinitive verb + object word or phrase + constraints*

The first component is the infinitive verb. An infinitive is a verb form that is usually preceded by the word *to*. An infinitive typically is used as a noun in objective statements. These infinitives allow determining whether or not they are achieved and are not subjective. Some examples for monitoring objectives are:

*To determine…*
*To evaluate…*
*To assess…*

The second component of an objective statement is the object. The object receives the action of the verb and answers the question, "What?" An example of a monitoring objective statement with an infinitive and a noun is:

*To determine the effects of implementing conservation practices…*

The third component of an objective statement is the constraints to the objective. This component is not necessary to make an objective statement. Constraints limit the objective statement to specified areas. The objective becomes constrained from the whole world

of opportunities or alternatives. Appropriate constraints can include the water quality variables to be sampled or the location of the study. For example, the completed monitoring objective could be:

*To determine + the effect of implementing conservation practices + on fecal coliform levels in Long Lake.*

Some constraints may be unnecessary and may overly limit the study design. For example, to limit the water quality variables to test for when the cause of pollution is unknown. The constraint would then interfere with determining the cause of the problem.

Coffee and Smolen (1990) suggest that monitoring objectives should specify the water quality variables, location of monitoring, the degree of causality, and the anticipated result of the management action.

### (b) Management objectives

For management objectives, the infinitives show a direction of preference; however, achievement of these objectives may be more subjective, depending upon how they are stated. The infinitives for management objectives include:

*To reduce…*
*To increase…*
*To eliminate…*

An example of a management objective statement with an infinitive and a noun is:

*To reduce bacterial loading…*

The completed management objective somewhat related to the monitoring objectives described above is:

*To reduce fecal coliform loading to Long Lake.*

This management objective is subjective. An example of a nonsubjective management objective is:

*To implement fecal coliform controls on 75 percent of the farms in the Long Lake watershed.*

## 614.0302 Objectives tree

Most projects have several objectives. These objectives may be complementary or even sometimes competitive. To achieve some overall general objective, several subobjectives may be needed. Thus, the subobjectives might be viewed as hierarchical.

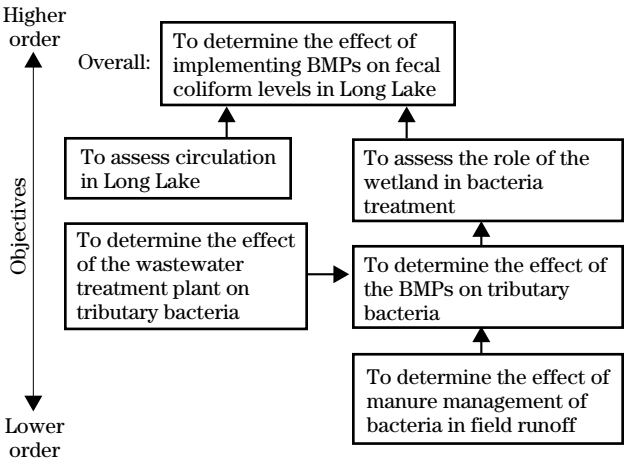
The relationships among objectives can be better understood by developing an objective tree. An objective tree displays all of the monitoring objectives in a hierarchical manner so that priorities can be established on which objective to tackle first. Two objectives in the tree are connected if *the achievement of one objective contributes directly to the achievement of the other objective*. Higher-order objectives are more general and stable than lower-order objectives. The lower-order objectives help to define the higher-level objectives more specifically and may change from time to time with expanding knowledge.

One way to develop the objective tree is to write each objective on a separate card and compare all possible combinations of card pairs using the statement: "Does the achievement of card A contribute directly to the achievement of card B?" If the answer is yes, the two objectives are connected in the direction indicated.

One of the advantages of developing the objective tree is that it shows the order in which objectives must be accomplished so that the overall objective can be attained.

An example of a monitoring objective tree is shown in figure 3–1. For this example, the system contains a wetland that receives tributary loadings before runoff outlets to the lake. The watershed has both point and nonpoint sources of bacteria. Also, the lake is not well-mixed and exhibits water quality gradients that appear to be influenced by wind-driven circulation patterns. In this case, before we could determine the effect of implementing BMPs in the watershed on the levels of bacteria in the lake, the circulation in the lake and the effect of the wetland would need to be assessed. Also, point and nonpoint sources of bacteria would need to be separated.

**Figure 3–1** Water quality monitoring objective tree




# 614.0303 Objective attributes

The final step in developing objective statements is to determine attributes for the objectives. Attributes define the level of achievement for each objective. Monitoring objectives are typically binary. They are either achieved or not achieved. For example, an assessment of the circulation patterns in Long Lake is either achieved or not. Another monitoring objective attribute could relate to time, such as:

*To determine circulation patterns in Long Lake in 1 year.*

One of the problems associated with binary attributes is that they have no intermediate steps upon which to evaluate progress.

Management or programmatic objectives may use other scaler quantities as attributes to measure their achievement. For instance, for the Long Lake example, an appropriate attribute for a management objective could be:

*…the percent of farms in the watershed receiving fecal coliform controls*

Another attribute could be:

*…the percentage change in bacteria loading to Long Lake.*

The attribute should be so stated that it helps answer the question, "…how do you know when you have monitored enough?"

In conclusion, monitoring objectives are often redefined after going through these three steps as well as after gaining experience in the monitoring project. Such changes are appropriate, expected, and should be encouraged.

## 614.0304 References

Coffee, S.W., and M.D. Smolen. 1990. The nonpoint source manager's guide to water quality monitoring. NC State Univ. Water Quality Group, Raleigh, NC.

Dickerson, S.L., and J.E. Robershaw. 1975. Planning and design: the systems approach. Chapter 2: Defining the problem. Lexington Books, Lexington, MA.

Keeney, R.L. 1988. Value-focused thinking and the study of values. *In* Decision making: descriptive, normative, and prescriptive interactions, chapter 21. D.E. Bell, J. Raiffa, and A. Tversky (Eds.) Cambridge Univ. Press, Cambridge, NY, pp. 465-494.

Keeney, R.L., and H. Raiffa. 1976. Decisions with multiple objectives: preferences and value tradeoffs. Chapter 2: The structuring of objectives. John Wiley & Sons, New York, NY.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 4 Statistical Designs

# Chapter 4 Statistical Designs

Contents:

# Chapter 4 Statistical Designs

## 614.0400 Introduction

Several experimental designs can be used to evaluate the effect of a conservation practice or a number of practices on water quality. The design selected depends primarily on the study objective. The study design must be determined before the project begins because the design of the project dictates most other aspects of the project including the study scale, the number of sampling locations, the sampling frequency, and the station type.

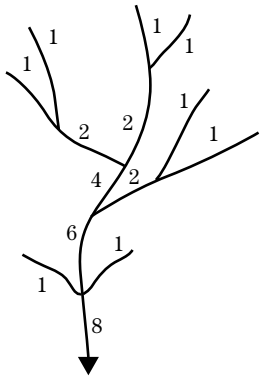
The study designs considered in this chapter include the reconnaissance, plot, single watershed, above-and-below, two watersheds, paired watershed, multiple watershed, and trend station. A more complete description of the statistical aspects of study designs is given in part 615 of this handbook.

## 614.0401 Reconnaissance

Reconnaissance or synoptic designs have been used to determine the magnitude and extent of the water quality problem or as a preliminary survey where no data exist. The term *synoptic* has been used to imply either obtaining a general view of water quality or obtaining samples at approximately the same time. Reconnaissance surveys differ greatly among the type of water body, whether stream, lake, or ground water. A properly stated objective also is critical for a reconnaissance survey. This type of monitoring is used to target critical areas as well.

Reconnaissance surveys are often grab sampling programs. For stream systems, one approach for determining sources of pollution was based on the number of contributing tributaries (Sanders, et al. 1983). In a downstream fashion, the number assigned to a stream segment is the sum of the numbers assigned to the upstream segments. The total number of segments at the most downstream station is used to select sampling locations. That number could be divided by two, four, and so forth, to obtain a desired number of sampling stations for the preliminary survey. The number obtained would describe which segment to sample. Example 4–1 illustrates this.

**Example 4–1** Sampling locations based on contributing tributaries




Determine the sampling locations for reconnaissance monitoring for the numbered stream segments. Assume two stations will be used.

$$\frac{8}{2} = 4$$

Sample segments 8 and 4

Other approaches might include designs based upon a percentage of the basin sampled, at known sources of pollution, and at shifts in land use or geology. One approach recommended by the World Health Organization (WHO) based the number of water quality stations on a percentage of the stream gaging stations, which in turn are based on a minimum density for different climate zones (WHO 1978). They also recommended "basic stations" to classify water quality and "auxiliary" stations to understand the assimilative capacity of streams. Basic stations were generally located at the mouth and major tributaries, at political boundaries, at water intakes, below outfalls, and below urban areas. In addition, when biological monitoring is being conducted, different stream habitats (riffle, pool) should be considered when selecting sampling stations.

Reconnaissance biological monitoring approaches, such as the Rapid Bioassessment Protocol I, must consider the major factors influencing aquatic organisms (Plafkin, et al. 1989). These factors include pollution sources, bottom types, stream habitats, flow characteristics, and other physical characteristics, such as shade (Klemm, et al. 1990). A biological reconnaissance is also important in determining ultimate sample sizes and taxa of importance. Reference stations are also recommended for reconnaissance biological monitoring.

The goal for stream reconnaissance surveys is often to locate the areas not meeting their intended uses and those that are the most polluted. Other design considerations in stream reconnaissance surveys are the frequency of sampling (chapter 9 and the number of locations needed per unit area.

Lake synoptic surveys typically involve collecting a large number of samples over a short time. Locations could be determined on an areal basis by overlaying a grid on the lake and sampling randomly located grid intersections. Other approaches include sampling bays or sampling longitudinally along lake gradients.

Design of ground water reconnaissance surveys depends on whether there is a local concern or more regional concern in knowledge about ground water quality. In local monitoring, monitoring wells are located above and below the potential pollution source. At a minimum the survey should have three wells located in a triangular array about the area of interest. This array allows the preliminary determination of flow direction. Additional wells could be added to further determine the extent of the contaminant plume. In regional reconnaissance surveys, wells could be located based on a grid bases as for lakes, or existing wells could be surveyed.

## (a) Advantages

Reconnaissance surveys are less expensive than fixed-station monitoring.

## (b) Disadvantages

Because of the frequent lack of statistical designs, reconnaissance surveys may miss important information. For example, stream grab sampling based on equal time intervals (e.g., weekly) often results in oversampling baseflow conditions and undersampling stormflow periods. As a result a smaller variability will be observed than actually exists. Also reconnaissance surveys have the potential to include judgment bias in the selection of sampling locations. Sampling just below outfalls, at tributaries, and at easily accessible locations, such as bridges, may give unrealistic representations of general water quality conditions.

# 614.0402 Plot

Plots have been used for conducting agricultural experiments in the United States since before 1900 (LeClerg, et al. 1962). They are generally small areas (fractions of an acre) that are replicated on the landscape or in the water. Plot size is a difficult decision. Generally, smaller plots that have many replicates are preferred to larger plots with fewer replicates (LeClerg, et al. 1962). For agronomy studies, three to six replicate plots have been recommended. A 0.01 acre plot might be 6 feet wide by 72.6 feet long (USDA 1979). On land, runoff plots might be used for studies of erosion, the surface transport of chemicals, or soil water nutrient status. In water, limnocorrals have been used in lakes to evaluate nutrient and acid additions. Plots are generally too small for ground water studies. The influence of the plot treatment on ground water below the plot may be insignificant in relation to other inputs to the ground water. However, field plots have been used to study water in the vadose zone.

For a plot design, all plots are treated alike except for the factor(s) under study. Plots are typically located across the slope in homogeneous areas, although such placement of plots can introduce a factor of bias (LeClerg, et al. 1962). Differences of an area can be accounted for by blocking. An example of blocking in a plot study is shown in figure 4–1. This example shows three replicates of four treatments. One treatment would be a control, the other three could be different rates of sludge applications, for example. Individual treatments would be randomly assigned to the plots. Blocking could be used to determine if there was an upslope-downslope effect.

## (a) Advantages

The greatest advantage to a plot study is that the treatments are replicated; most watershed studies have no true replicates. Also, plots generally allow control of several variables, such as soil type, including the treatment (Striffler 1965). Plots are generally small enough that precipitation should be uniform over the area. A major advantage of the plot design is that it has a control. A control is a plot that is monitored like all others, but does not receive the treatment.

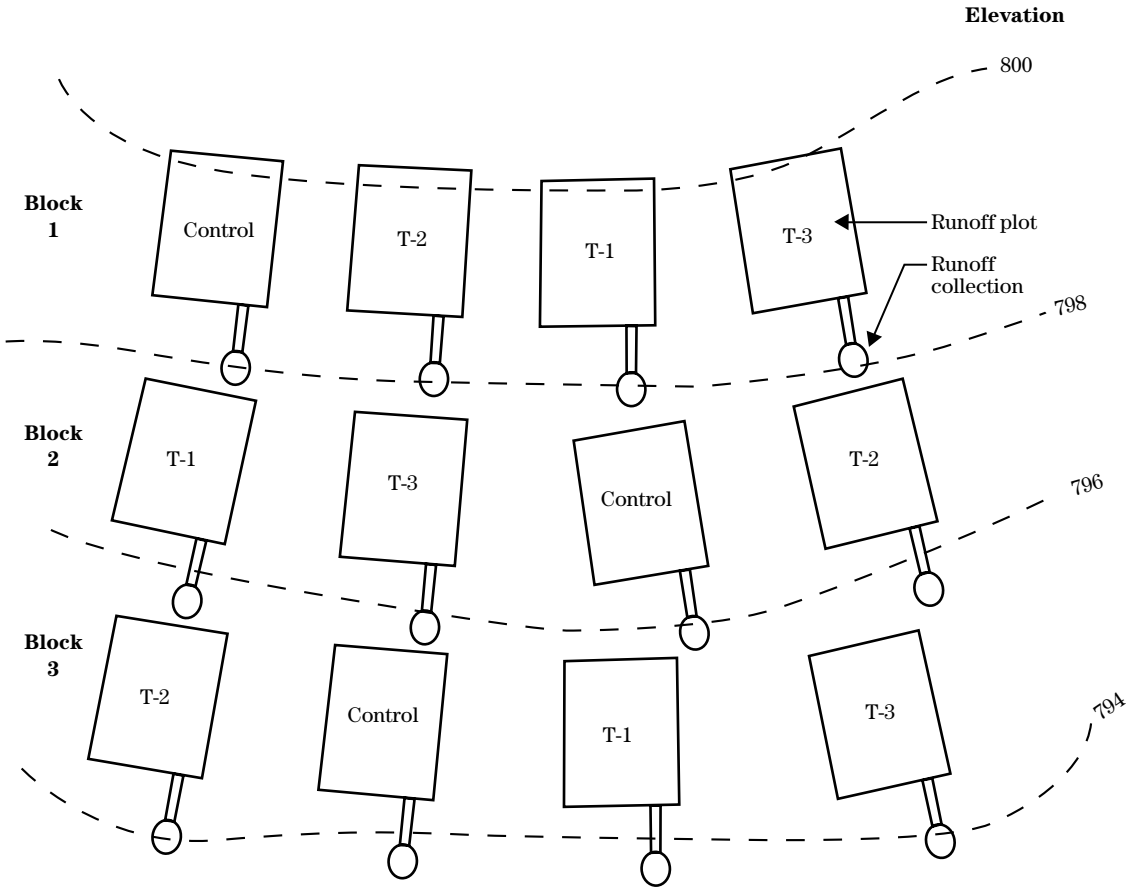
## (b) Disadvantages

The results from plot studies are not transferable to other watersheds, especially larger watersheds (Striffler 1965). Plots also may be too small a unit to adequately represent the hydro-ecosystem. Because of their small size, plots do not receive "real world" management. They must be separated from each other by some method to prevent cross-contamination of the treatment from one plot to another.

## (c) Statistical approach

The primary statistical approach is the analysis of variance of a randomized complete block design. The area where the plots are to be located is divided into blocks, with the number of blocks equal to the number of replicates chosen. Each block serves as a replication. Blocks are assumed to be homogeneous areas. For the example in figure 4–1, three blocks are shown at difference elevations. Each block contains all treatments. The treatments are assigned to plots within the blocks randomly. This design allows for the removal of the effect of the block that might be caused by differences in the field. Other more complicated designs are available including the Latin square and split plot designs, or a factorial arrangement of treatments (Snedecor & Cochran 1980). These designs are described in part 615 of this handbook.

**Figure 4–1** Layout example of a plot study with blocking

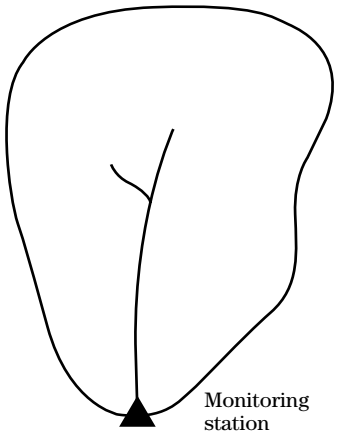


# 614.0403 Single watershed/ before-after

A single watershed has sometimes been used to evaluate the water quality effectiveness of a conservation practice (fig. 4–2). Water quality monitoring is conducted both before and after the practice is applied. The *before* period has sometimes been referred to as baseline data. Generally, this technique is not recommended and should be avoided (see 614.0403(b)).

However, a second manner in which a single watershed could be used was described by Striffler (1965). For this technique a water quality variable could be related to a climate variable(s), such as precipitation. The difference because of the conservation practice could be evaluated as a change in the relationship between the water quality characteristic and the climate variable. The interpretation of results would be somewhat constrained. For example, a result might be: "For an equal amount of monthly precipitation, the concentration declined." More specific results are generally needed, such as the percent reduction in a water quality variable resulting from the practice.

**Figure 4–2** Single watershed design




## (a) Advantages

The primary advantage of the single watershed design, with monitoring before and after a practice is implemented, is that it is the simplest of all designs. Only one monitoring station needs to be monitored. This design is applicable for most watersheds (Striffler 1965).

## (b) Disadvantages

This design should not be used because the effect by the practice cannot be separated from other confounding effects. As indicated in table 4–1 for the single watershed design, the effect because of the treatment (e.g., BMP) cannot be separated from year-to-year climate differences. If a dry year occurred when the practice was implemented, following a wet year when the watershed was in the pre-practice stage, stream concentration reduction would generally occur because of the climate differences. Also, an interaction would most likely occur between climate and the practice that could not be assessed by the study. For example, during a drought a field terrace might be expected to reduce sediment loading to a stream. However, during a wet year the terrace could be overtopped, resulting in increased suspended solids loading.

Using the alternative relationship approach described by Striffler (1965) on a single watershed is more complex, requires a longer calibration period, and is less precise than a paired watershed design. The single watershed design also has the disadvantage of not being able to transfer results to other areas.

**Table 4–1** Causal factors for alternative monitoring designs

| Design | Cause |
|---|---|
| Single watershed/ before-after | BMP or climate |
| Above-and-below watershed | BMP or watershed |
| Two watersheds | BMP or watershed |
| Paired watershed | BMP |

## (c) Statistical approach

The difference in water quality caused by the practice generally is expressed as the difference between the means for the two periods. A t-test is most often used for this type of comparison (Snedecor and Cochran 1980). An appropriate null hypothesis (Ho:) might be that the mean concentrations are equal between the two periods, for example:

$H_o$: *mean tss (period 1) = mean tss (period 2)*

As described further in part 615 of this handbook, rejection of the null hypothesis is desirable. Errors can be made in accepting the null hypothesis.

A paired t-test is not appropriate for this design because the samples collected are not paired in any meaningful way. For example, the water quality associated with months across years cannot be paired because of random components in water quality.

To perform a parametric t-test, the samples would need to be random, independent, normally distributed, and have equal variances. A nonparametric comparison of means could be used where data are not normally distributed.

The statistical approach for using the relationship between water quality and a climate variable would be similar to that described for the paired watershed below. The differences between the slopes and intercepts of the two regression relationships (one pre-practice, one post- practice) would be analyzed using analysis of covariance. Multivariate regressions that include flow or climate variables might improve these relationships.

Examples of the statistical approach to apply to a single watershed design are given in part 615 of this handbook.

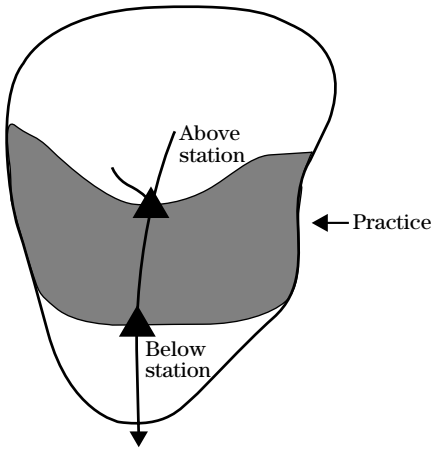
# 614.0404 Above-and-below watersheds

The above-and-below design is applied after the treatment is in place. This approach is sometimes viewed as a single watershed with monitoring above and below a practice (fig. 4–3), or in the case of ground water monitoring, upgradient and downgradient from the activity of interest. In actuality, two watersheds are being monitored, one nested within the other. In some cases the above station is erroneously thought of as "background water quality," and the below station is the one believed to be influenced by the practice.

This design is probably the most commonly used strategy in-ground water monitoring. Placement of the wells is important because ground water sites are three-dimensional. Gradients may occur in both vertical as well as horizontal directions.

If the above-and-below approach is applied both before and after the practice is installed, this approach can be analyzed as a paired watershed design as described below.

**Figure 4–3** Above-and-below watershed design



### (a) Advantages

The above-and-below approach is not as susceptible to year-to-year climatic differences as is the single watershed approach using before and after sampling. Also, it may be relatively easy to locate a watershed where a practice could be implemented between the above and below stations on a stream. This technique may be useful for isolating critical areas. The above-and-below design is well suited to biological as well as chemical/physical monitoring.

### (b) Disadvantages

Water quality measurements from nested watershed may not be independent. The water quality downstream is most likely a function of the upstream water quality. For example, a high concentration upstream would most likely result in a large concentration downstream.

A second major disadvantage of this design is that the differences between the above and below stations might be caused by inherent watershed differences (e.g., geology) or to some interaction between the practice and the watershed, and not only because of the practice itself (table 4–1). These various causal factors cannot be separated using this design; however, proper site selection may reduce this effect.

### (c) Statistical approach

The above-and-below design is analyzed as a t-test of the differences between paired observations at the above and below stations (see part 615). An appropriate null hypothesis might be:

$$H_o: \textit{difference} = 0$$

Parametric and nonparametric (distribution free) t-test approaches are available. A nonparametric analysis uses the rank of the data rather than the data itself (part 615).

Another approach would be to compare regressions between concentration and a climate variable, such as flow, for the above and below stations (Ponce 1980).

## 614.0405 Two watersheds

Two watersheds, one with the practice and one without, have been incorrectly used to evaluate the effects of a practice on water quality. This design should always be avoided. The two watershed design is not the same as the paired watershed design. There is no calibration period for the two watershed design when the two watersheds are in the identical treatment, but there is for the paired watershed approach.

### (a) Advantages

Two watersheds, each in a different land use, are relatively easy to locate.

### (b) Disadvantages

The differences in water quality between the two watersheds may be caused by the practice, inherent watersheds differences, or an interaction between these two factors, and there is no way to distinguish among these causal factors (table 4–1).

### (c) Statistical approach

Although a statistical examination of the water quality associated with two watersheds may not be appropriate, the water quality could be compared using the same approach as that for the nested watersheds. That is, a paired *t*-test or nonparametric *t*-test of treatment means could be used. In some cases regressions between water quality and a climate variable could be compared.

## 614.0406 Paired watersheds

Paired watersheds have been used for over 40 years to evaluate the effects of silvicultural practices on watershed quantity and quality (Wilm 1949). The basic approach requires a minimum of two watersheds and two periods of study. The two watersheds are called control and treatment; the two periods of study are referred to as calibration and treatment (fig. 4–4). The control watershed serves as a check over year-to-year or seasonal climate variations and receives no changes in management practices during the study.

During the calibration period, the two watersheds are treated identically and paired water quality data are collected. Such paired data could be annual means or totals, or for shortened studies, the observations could be seasonal, monthly, weekly, or event-based.

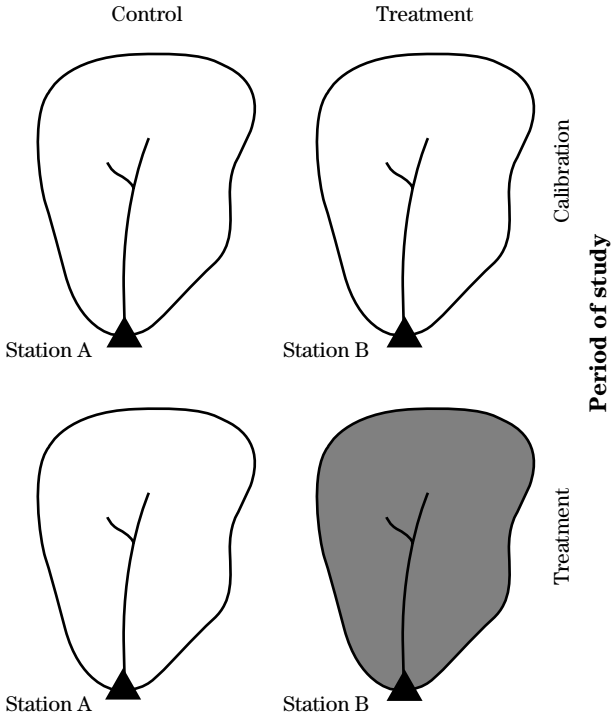
During the treatment period, one randomly selected watershed is treated with a practice while the control watershed remains in the original management. The reverse of this schedule is possible for certain practices. Both watersheds might already be treated with a conservation practice during the calibration period. During the treatment period, one of the watersheds could be treated with a traditional practice.

For ground water monitoring, an above-and-below approach to the paired watershed design is recommended. During the calibration period, monitoring would take place upgradient and downgradient for both the control and treatment portions of the ground water formation being studied. During the treatment period, one of the areas bounded by wells would receive a practice, while the other control area would remain as before.

Guidelines for paired watershed studies include:

- Steady-state—The control watershed should be at or near a steady-state condition during the life of the study (Reinhart 1967). Steady-state is used here to mean that there are no gradual changes that would result in a trend in water quality. For example, a watershed that had a gradual shift in crop types would not make a good control.
- Size—The watersheds should be small enough to obtain a uniform treatment over the entire area (Reinhart 1967). The size will vary depending on climatic region. In humid areas the watersheds generally would be less than 5 square miles in area. In arid climates, they could be larger.
- Range—The calibration period should encompass the full range of observations expected (Reinhart 1967, Wilm 1949). Normally, this refers to wet and drought years. This allows reasonable comparison of treatment data to calibration data.
- Calibration length—The calibration period should be long enough to develop significant regression relationships between the two watersheds so that data for the treatment watershed can be predicted knowing data from the control watershed within certain error limits (Striffler 1965). Methods for determining the length of calibration are described in part 615.

**Figure 4–4** Paired watershed design



- Response—The designed treatment should be expected to have a large enough response to exceed prediction errors. At least a 10 percent change in the variable of interest is suggested (Hewlett & Pienaar 1973).
- Watershed similarity—The watersheds should be similar in size, slope, location, soils, and land cover (Hewlett 1971, Striffler 1965). They should also have been in the same land cover for a number of years before the study (Hewlett 1971). Chemical characteristics of the soils should be similar. However, no two watersheds are identical, nor can they be considered representative.
- Monitoring suitability—Each watershed should have a stable channel, a stable control section for monitoring, and should not leak around the gaging station at the watershed outlet (Reinhart 1967).

## (a) Advantages

The greatest advantage of the paired watershed approach is that variation not associated with the treatment, such as climate differences over years, are statistically controlled (Kovner & Evans 1954). Also, the control watershed eliminates the need to measure and understand all the mechanisms generating the response (Hewlett & Pienaar 1973). The water quality of runoff from the two watersheds need not be identical. Finally, the calibration phase can be done in reverse with the treatment period preceding the calibration period (Reinhart 1967).

## (b) Disadvantages

Several disadvantages to the paired watershed approach also apply to all the study designs.

- The variances in water quality data are not likely to be equal between time periods because the treatment on one of the watersheds is often quite drastic. It is also difficult to satisfy the assumptions of normality and independence of observations. Shortened calibrations may increase the likelihood of serially correlated data (Reinhart 1967).
- The treatment effect may be gradual and not constant with time (Reinhart 1967; Hewlett & Pienaar 1973). Thus overall comparisons may mask interesting results.
- The paired watershed experiment is costly and time consuming (Hewlett & Pienaar 1973).
- Long-term changes in the soils or vegetation may occur in the control watershed. Other catastrophes, such as fires, dust storms, hurricanes, and insect infestations, could occur, which could destroy the meaning of results. This disadvantage applies to all watershed designs.

## (c) Statistical approach

The basis of the paired watershed approach is that there is a quantifiable relationship between paired water quality data for the two watersheds and that this relationship will persist until a major change is made in one of the watersheds (Hewlett 1971). This does not require that the quality of runoff be the same for the two watersheds; but rather that the relationship between the water quality of the two sites, except for the influence of the treatment (practice), remains the same over time. In fact, most often the water quality is different between the two watersheds. This inherent difference between all watersheds further substantiates the need to use the paired watershed approach.

The primary statistical approach is to develop significant regression relationships between the control and treatment watersheds during both the calibration and treatment periods (see part 615). These two regression relationships are then compared for identical slopes and intercepts using analysis of covariance (Reinhart 1967). During the calibration period the significance of the regression is tested using analysis of variance for regression (Snedecor & Cochran 1980). Procedures for determining the length of the calibration period have been described by Wilm (1949), Kovner and Evans (1954), and Reinhart (1967) and are presented in part 615 of this handbook. An alternative analysis approach has been presented by Green (1979), Bernstein and Zalinski (1983), and Carpenter, et al. (1989).

# 614.0407 Multiple watersheds

The multiple watershed approach involves more than two watersheds (Clausen and Brooks 1983, Striffler 1965, Wicht 1967). Watersheds with the treatments already in place are selected from across the region of interest. The region could be as large as a state or as small as an individual field. Sampling of the runoff is conducted from these watersheds over a period of time.

As an example, multiple watersheds could be used as a method to assess the water quality effect of storing manure during the winter and not daily spreading as a conservation practice. About 15 watersheds in each treatment could be selected. That is, 15 fields or watersheds where daily spreading was occurring during the winter, and 15 fields where no spreading occurred. During runoff periods, these fields could be sampled for the concentrations of appropriate pollutants, such as nitrogen and phosphorus.

Another example could be a test of irrigation water management. Runoff from fields in flood irrigation could be compared to runoff from sprinkler irrigated fields.

## (a) Advantages

The greatest advantage of the multiple watershed approach is that the results are transferable to the region included in the monitoring. A second major advantage is that the true variability among watersheds is included in the variance for each treatment.

## (b) Disadvantages

The multiple watershed approach is difficult to conduct using intermittent streams or field runoff because sampling must be timed with stormflow periods. Also, mass calculations would only be point estimates, and annual mass calculations would be expensive to obtain using a large number of watersheds. However, the probability approach has been used to determine annual mass estimates, which could reduce the number of samples that need to be collected (Richards 1989).

## (c) Statistical approach

The basic statistical approach is the comparison of the means of two populations using the *t*-test. The testing would be for unpaired samples that may be of unequal sizes.

# 614.0408 Trend stations

Trend stations are single watersheds monitored over time. A trend is a persistent change in the water quality variable(s) of interest over time. In many cases the most appropriate design may be the use of long-term trend stations. Trend stations are single, independent watersheds where a group of conservation practices might be implemented gradually over time or where the response to a practice might take a long time.

It is important for trend analysis that there not be gaps in the data set, that methods of water quality analysis not change during the study, that hydrologic control at the monitoring station is stable, and that a causal link can be made between water quality and the watershed treatments. This implies that collection of hydrologic data and land use activities are crucial to trend analysis. In addition, for some trend analysis techniques, water quality data must be collected or aggregated to fixed time intervals (Valiela & Whitfield 1989; Montgomery & Reckhow 1984; Hirsch, et al. 1982).

The use of a control watershed for trend detection cannot be emphasized enough. The control should have a stable land use and no changes in practices during the life of the trend investigation.

Although models are sometimes used to simulate long-term trends, the purpose of this handbook is to discuss the applicability of monitoring and not modeling.

## (a) Advantages

A long-term trend station is relatively easy to establish for watersheds drained by permanent streams. For complex watersheds, conservation practices are typically installed at different times over several years. This prevents use of short-term designs. For example, it may take many years for water quality to respond to practices because of the residual storage of nutrients.

## (b) Disadvantages

A true commitment to long-term (>10 years) monitoring is difficult to achieve because of changing priorities and changing personnel within funding and monitoring agencies. A significant effort must be made for land use data tracking. Over the long term, the potential is greater for unwanted disturbances, such as a new road or urban development, to affect water quality.

## (c) Statistical approach

A large number of parametric and nonparametric techniques are available for detecting trends in water quality data. Several techniques should be used before reaching a conclusion (WHO 1978). These techniques are described below and discussed in detail with examples in part 615 of this handbook.

**Time plot**—A graph of the water quality versus time is useful in detecting obvious trends (WHO 1978).

**Least square fit regression**—A linear or nonlinear regression could be fit through the data, which would allow quantification of the slope or trend rate (WHO 1978).

**Comparison of annual means**—A t-test could be used to compare averages for shorter, equal time periods within the trend total period (WHO 1978). For example, annual means could be compared. An analysis of variance, followed by a multiple comparison test, would be a more appropriate method because the overall variance would be pooled (Snedecor and Cochran 1980).

**Cumulative distribution curves**—Two cumulative distribution curves (which portray the percent cumulative distribution as a function of concentration) for two different time periods could be compared for shifts to determine trends (WHO 1978).

**Q-Q plo**t—A Q-Q plot is a comparison of the quartiles of one data set plotted against those of another data set for the variable of concern. By comparing the data from different time periods, a shift in the data as compared to a y=x line can be determined (WHO 1978).

**Double mass analysis**—Typically used for precipitation records, double mass analysis is a comparison of the accumulated data from one station plotted against the accumulated averages of data from several stations. A break in the slope would indicate a change in that one station as compared to the others, which could be interpreted as a trend (Dunne & Leopold 1978).

**Paired regressions**—A "before" period can be compared to an “after” period by the comparison of the regression equations between data from a control trend station and a treatment trend station. This analysis is identical to the paired watershed analysis described above.

**Time series analysis**—Because water quality data collected at the same station may be autocorrelated, time series analysis could be used to detect trends (McLeod, et al. 1983). However, the forecasting features of time series analysis are not likely to be relevant (Vandaele 1983).

**Seasonal Kendall test**—This nonparametric approach is especially useful where seasonality exists in the data set. A seasonal Kendall slope estimator is used to determine the magnitude of the trend (Hirsch, et al. 1982).

Generally, when applying several approaches to trend detection, the results rarely vary in direction, although the statistical significance of these techniques will vary. All of the methods, except paired regressions, only provide information on whether a trend exists and not why it exists. Only the paired regression approach allows linking the trend to causes other than hydrologic because a control is used. An alternative approach would be to adjust the trend data set for hydrologic influences. This is described in part 615 of this handbook.

## 614.0409 References

Bernstein, B.B., and J. Zalinski. 1983. An optimum sampling design and power tests for environmental biologists. J. Environ. Mgmt. 16(1):35-43.

Carpenter, S.R., T.M. Frost, D. Heisey, and T.K. Kratz. 1989. Randomized intervention analysis and the interpretation of whole-ecosystem experiments. Ecology 70(4):1142-1152.

Clausen, J.C., and K.N. Brooks. 1983. Quality of runoff from Minnesota peatlands: II. A method for assessing mining impacts. Water Resour. Bull. 19(5):769-772.

Dunne, T., and L.B. Leopold. 1978. Water in environmental planning. W.H. Freeman and Co., San Francisco, CA.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York, NY.

Hewlett, J.D. 1971. Comments of the catchment experiment to determine vegetal effects on water yield. Water Resour. Bull. 7(2):376-381.

Hewlett, J.D., and L. Pienaar. 1973. Design and analysis of the catchment experiment. *In* Proc. Symp., Use of small watersheds, E.H. White (ed.), Univ. KY.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. Techniques of trend analysis for monthly water quality data. Water Resour. Res. 18(1):107-121.

Klemm, D.J., P.A. Lewis, F. Faulk, and J. Lazorcnak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. U.S. Environmental Protection Agency. Office of Research and Development, Washington, DC. EPA/600/4-90/030.

Kovner, J.L., and T.C. Evans. 1954. A method for determining the minimum duration of watershed experiments. Trans. Amer. Geophys. Union. 35(4):608-612.

LeClerg, E.L., W.H. Leonard, and A.G. Clark. 1962. Field plot technique. 2nd ed., Burgess Publ. Co., Minneapolis, MN.

McLeod, A.I., K.W. Hipel, and F. Comancho. 1983. Trend assessment of water quality time series. Water Resour. Bull. 19(4):537-547.

Montgomery, R.H., and K.H. Reckhow. 1984. Techniques for detecting trends in lake water quality. Water Resour. Bull. 20(1):43052.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Office of Water, Washington, DC. EPA/444/4-89-001.

Ponce, S.L. 1980. Water quality monitoring programs. USDA Forest Serv., Watershed Sys. Dev. Group. WSDG Tech. Pap. WSDG-TP-00002, Fort Collins, CO.

Reinhart, K.G. 1967. Watershed calibration methods. *In* Proc. Intern. Symp. on forest hydrology. W.E. Sopper and H.W. Lull (eds.), Pergamon Press, Oxford, pp. 715-723.

Richards, R.P. 1989. Determination of sampling frequency for pollutant load estimation using flow information only. *In* Proc. Symp., Design of water quality information systems. R.C. Ward, J.C. Loftis, and G.B. McBride (Eds.) CO Water Resour. Res. Inst. Inf. Ser. No. 61, Fort Collins, CO.

Sanders, T.G., R.C. Ward, J.C. Loftis, T.D. Steele, D.D. Adrian, and V. Yevjevich. 1983. Design of networks for monitoring water quality. Water Resour. Pub., Littleton, CO.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods. 7th Ed. IA State Univ. Press, Ames, IA.

Striffler, W.D. 1965. The selection of experimental watersheds and methods in disturbed forest areas. Publ. No. 66. I.A.S.H. Symp. of Budapest. pp. 464-473.

United States Department of Agriculture. 1979. Field manual for research in agricultural hydrology. Agric. Handb. 224, Washington, DC.

Valiela, D., and P.H. Whitfield. 1989. Monitoring strategies to determine compliance with water quality objectives. Water Resour. Bull. 20(1):127-136.

Vandaele, W. 1983. Applied time series and Box-Jenkins models. Academic Press, Inc., New York, NY

Wicht, C.L. 1967. The validity of conclusions from South African multiple watershed experiments. *In* Proc. Int. symp. on forest hydrol. W.E. Sopper and H.W. Lull (eds.), pp. 749-760.

Wilm, H.G. 1949. How long should experimental watersheds be calibrated? Amer. Geophys. Union Trans., part II. pp.618-622.

World Health Organization. 1978. Water quality surveys—A guide for the collection and interpretation of water quality data. IHD-WHO Working Group on Quality of Water. UNESCO/WHO, Geneva, Switzerland.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 5 Scale of Study

(450–VI–NWQH, September 2003)

# Chapter 5 Scale of Study

Contents:



Tables



Examples

# Chapter 5 Scale of Study

## 614.0500 Introduction

The fourth step in developing a water quality monitoring study is to determine the size or scale of the area to monitor. The study scale depends in part on: 1) study objectives, 2) available resources, 3) study duration, 4) type of water resource, and 5) the complexity of the project to monitor. These individual factors are described later in this chapter.

Although considered as a separate step, study scale is actually coupled with the statistical design. However, scale is provided as a separate subpart to force consideration of this decision in the overall design of a water quality monitoring study.

This chapter recognizes four scale categories—point, plot, field, and watershed—although it is acknowledged that the latter three scale types are in reality all watersheds.

For lake systems, the terminology is different. Plots are limnocorrals, fields are bays or regions, and watersheds are lakes. In ecology, scales are referred to as either microcosm (e.g., point), mesocosm (e.g., plot, limnocorral), and macrocosm (e.g., field, watershed, lake) (Odum 1984).

One potential barrier to selecting the appropriate scale of the project is where the monitoring objectives are not clearly stated. Contemplating the scale of the project often results in a clarification of the objectives in a feedback sense.

## 614.0501 Point scale

Points are the smallest scale considered for water quality monitoring and are characterized by obtaining single observations. The term "point scale" means a point in space, but not a point in time. Examples of point-scale monitoring include precipitation gages, snow samples, soil samples, most vadose zone lysimeters, and many lake samples. Ground water wells and stream samples are considered watershed-scale samples and not point-scale samples even though they may be taken at one location.

Point sampling is appropriate for trend monitoring, for problem definition or compliance monitoring, for research and fate and transport monitoring, or for evaluating certain types of models (table 5–1). Point samples are used in both vadose zone and lake studies (table 5–2). Point sampling is considered cheaper than larger scales, but the frequency of visits and the duration of sampling will vary greatly depending on the study objectives.

**Table 5–1** Objective by study scale matrix

| Objective | Point | Plot | Field | Watershed |
|---|---|---|---|---|
| 1. Baseline | | | X | X |
| 2. Trends | X | | | X |
| 3. Fate and transport | X | X | X | X |
| 4. Problem definition | X | | X | X |
| 5. Critical areas | | | X | X |
| 6. Compliance | X | | X | X |
| 7. BMP effectiveness | | X | X | |
| 8. Program effectiveness | | | | X |
| 9. Wasteload allocations | | | X | X |
| 10. Model evaluation | X | X | X | X |
| 11. Research | X | X | X | |

# 614.0502 Plot scale

Plots are mesocosm sampling units (LeClerg, et al. 1962). They are appropriate monitoring units if the objective is to replicate several treatments as part of a fate and transport study or if the effectiveness of a conservation practice or a model is evaluated (table 5–1). When considering the type of water body being studied, a plot scale is appropriate when investigating soil solution water or overland flow, but not for ground water, streamflow or lake studies. This is because these systems are larger than plot boundaries (table 5–2). An exception to the use of plots for lakes and streams would be the use of limnocorrals or seepage meters and artificial stream channels, which are actually plots of a mesocosm scale. Limnocorrals are floating water column enclosures that do not allow mixing with lake water (Odum 1984). Seepage meters are barrels placed on the land bottom that allow sampling of the flux through lake sediment (Lee 1977). Artificial streams divert some stream water into a controllable, constructed channel.

Plot studies work well for short duration (<5 years) studies, but may require a greater investment of personnel time and funds than other study scales. This, of course, depends upon the complexity of the study (table 5–3). The number of plots needed for an experimental study is a function of the number of treatments applied. A single treatment requires twice the number of plots as the number of replications because an equal (recommended) number of control plots is needed. For example, if the number of replicates determined based on the variability in runoff data were 5 (see chapter 9), the total number of plots needed would be 10. For two treatments, an additional five plots would be needed (table 5–4). The plot design is appropriate for evaluating a large number of individual practices (table 5–5).

From a water quality perspective, a critical requirement for the design of plots is that the treatment on each plot is isolated from all the other plots, or through monitoring, the effects of one plot are separated from the other plots by subtraction. For example, plots should be separated far enough apart so that a spray treatment on one plot could not drift onto other plots. Plots also may need to be isolated from overland flow from upslope areas. If the plot is designed for soil solution monitoring (e.g., via lysimeters), the plots may need to be configured to allow measurements of the soil solution of subsurface water entering the plot from above as well as at the bottom of the plot.

Several studies have used single replication plots; for example, one plot in treatment A, one plot in treatment B, and one control, for a total of three plots. This design is insufficient to determine the effects of the treatment. One can determine that the plots are different, but cannot distinguish between the difference as a result of the treatment or the individual plot.

**Table 5–2** Type of water resource by study scale matrix

| Waterbody | Point | Plot | Field | Watershed |
|---|---|---|---|---|
| Overland flow | | X | X | |
| Vadose zone | X | X | X | |
| Ground water | | | X | X |
| Streamflow | | X | | X |
| Lakes | X | X | X | X |

**Table 5–3** Relative cost and time requirements of various study scales

| | Point | Plot | Field | Watershed |
|---|---|---|---|---|
| Cost | Low | High | Low | Moderate to high |
| Frequency of visits | varies | events | events-weekly | weekly + |
| Duration | varies | <5 years | <5 years | >5 years |

**Table 5–4** Number of plots required based on the number of treatments (assuming replicates=5)

| Treatments | Plots |
|---|---|
| 1 | 10 |
| 2 | 15 |
| 3 | 20 |

**Table 5–5** Practice by study scale matrix

| Practice | Plot | Field | Watershed |
|---|---|---|---|
| **Vegetative/tillage practices** | | | |
| Conservation cropping | X | X | |
| Conservation tillage | X | X | |
| Contour farming | | X | |
| Cover crop | X | X | |
| Crop residue | X | X | |
| Crop rotation | | X | |
| Filter strip | X | X | |
| Mulching | X | X | |
| Hayland planting | X | X | |
| Riparian buffer | X | X | |
| Stripcropping, contour | | X | X |
| **Structural practices** | | | |
| Grassed waterway | | X | |
| Streambank protection | | | X |
| Terrace | | X | |
| **Management practices** | | | |
| Animal waste mgmt | X | X | |
| Irrigation mgmt | X | X | |
| Pasture/hayland mgmt | X | X | |
| Pesticide management | X | X | |
| Plant nutrient mgmt | X | X | |
| Woodland mgmt | | X | X |

**Example 5–1** Plot scale

The University of Rhode Island established 18 plots to monitor the water quality associated with turfgrass management (Morton, et al. 1988). Plots were 7 by 50 feet, were sloped at 2 to 3 percent, and had a 5-foot sod alley between them. Soil solution water was collected from 18 plots using ceramic lysimeter plates. The plots received six treatments consisting of three rates of nitrogen application and two irrigation rates per each nitrogen treatment. Each treatment was replicated three times. Overland flow collection occurred on 12 plots using an orifice flow splitter (10% of flow) to collection barrels.

This plot study determined that overwatering concurrently with fertilization can result in significantly higher nitrogen losses than controls. However, with scheduled irrigation, nitrogen losses were not different from controls. The study took 2 years to complete.

# 614.0503 Field scale

Monitoring on a field scale implies a larger area than an individual plot, although the entire plot design taken together could cover an area larger than a single field. The area of a field is difficult to state because it varies greatly in different parts of the United States. A field in humid (precipitation > evapotranspiration) areas is an area smaller than that required to produce a first order stream. In subhumid and arid areas (precipitation < evapotranspiration), a field typically would be larger, and many fields may occupy the area required to produce a first order stream.

Identical to the plot scale, a field scale monitoring project is appropriate if the objective was to investigate the fate and transport of a substance or the effectiveness of an individual conservation practice or a model (table 5–1). Field scale studies also are appropriate for ground water, vadose zone, and overland flow studies (table 5–2). The cost of monitoring a field scale project generally is not as great as either plot studies or watershed scale projects. Field scale projects are usually of short duration (<5 years), but could be longer.

Field scale projects are most suitable for evaluating individual practices on a field. For example, the practices may include field nutrient management, erosion control, or conservation cropping (table 5–5). If a field scale project is selected, it is important that the appropriate design (chapter 4) be matched to this scale. Monitoring a single field before and after a practice is installed is not an acceptable design unless the effects of climate over time are accounted for.

The scale of filter strips and many other constructed conservation practices, such as wetlands, lies somewhere between plot and field scales. Monitoring is usually conducted above and below the practice and typically has not been replicated.

For lake systems, different regions of a lake are synonymous with different fields on the land. Lake regions may be represented by bays, areas near sources, such as beaches, or gradient zones.

**Example 5–2** Field scale

Two fields were used in Vermont to determine the effect of conversion from conventional tillage to conservation tillage on pesticides in runoff (Clausen, et al. 1990). The two fields were compared using the paired watershed technique (subpart 614.02). During the calibration period, the two fields were moldboard plowed. During the treatment period, one field was conventionally tilled, while the other field was disk harrowed and planted with a conservation tillage planter. The two fields were 1.6 and 2.1 acres in area and had slopes ranging from 3 to 7 percent.

Field runoff was continuously monitored with heated,1.5-foot H-flumes and water level recorders. Flow proportional samples (0.1% of total flow) were obtained by tubing connected to the throat of the flume and to a storage carboy.

Using the paired watershed technique, conservation tillage was found to reduce runoff from the field. Therefore sediment loss and the mass export of the pesticides atrazine and cyanazine also were decreased.

# 614.0504 Watershed scale

A project scale larger than either plots or fields is needed if the monitoring objectives are to determine long-term trends, identify critical areas, examine standard compliance, make wasteload allocations, or verify watershed scale models (table 5–1). In addition, where a number of BMP systems are being installed in a watershed with the intent of improving downstream water quality, watershed scale monitoring is a necessity.

Watershed scale monitoring also is desired if the water resource system of concern is either ground water, a stream, or a lake/estuary (table 5–2). Watershed scale monitoring costs range from moderate to high depending on the size of the system being monitored. Large streams or lakes are more costly to monitor than smaller water bodies. Watersheds are studied for longer durations than are either plots or fields. For most individual BMPs, watersheds are not an appropriate scale of study. However, exceptions might include riparian buffers and streambank protection, which could be evaluated on a watershed basis (table 5–5). The watershed scale would be more appropriate for biological and habitat monitoring than smaller scales.

The most difficult decision regarding watershed scale projects is the selection of watershed size. Several factors influence the selection of watershed size including: drainage pattern, stream order, stream permanence, climate region, the number of manageable landowners, the homogeneity of land uses, and watershed geology and geomorphology.

No real relationship exists between a watershed area and most stream characteristics, including stream order, stream length, and drainage density (stream length /watershed area) (Harlin 1984). For example, the relationship $L=1.4A^{0.6}$ (where L=stream length and A=watershed area) has been found to be regional. The primary reasons for the lack of relationships are the differences in climate regions and geology across the U.S. It is not surprising that watershed area and watershed discharge would vary from humid climate regions to arid or subhumid regions. The ratio of potential evapotranspiration to precipitation has been used to distinguish between climate regions, with a ratio of one separating humid from subhumid areas (Holdridge 1962). If precipitation equals or is less than evapotranspiration, very little runoff would be expected and a larger basin would be needed to generate a permanent stream. On the other hand, if precipitation exceeds evapotranspiration, runoff would most likely occur, and a smaller basin would be needed to generate streamflow.

Streams draining small watersheds in humid regions (precipitation > evapotranspiration) are usually first or second order, intermittent, and < 500 acres in area. Moderately sized watersheds are from 500 to 5,000 acres in area, are permanent, and have third or fourth order streams. Stream order, according to Strahler's method (Ruhe 1975), is determined by numbering the smallest streams highest in the watershed as first order streams. The joining of two first order streams results in a second order stream, and so on.

Humid watersheds larger than 5,000 acres and less than 50,000 acres are considered large. Watersheds larger than 50,000 acres are considered very large and may be inappropriate for monitoring because of their likely heterogeneity in land uses.

The size of the watershed selected influences the response to implementation of conservation practices. For example, the export of phosphorus from agricultural watersheds generally decrease per unit area as the watershed size increases (T.-Prairie & Kalff 1986). This effect was not observed for forested watersheds. Comparing different agricultural land uses, this decreasing phosphorus export with increasing watershed area occurred for row crop and pasture watersheds, but not for mixed agricultural or non-row crop basins. The authors attributed this difference to a combination of decreasing sediment delivery ratios, a reduction of drainage density, and decreasing slope with increasing watershed area. Because an average of 84 percent of the total phosphorus exported from agricultural watersheds was found in the particulate rather than dissolved form, the decreasing sediment delivery would result in decreasing phosphorus delivery. For forested watersheds, less than 50 percent was in the dissolved form (T.-Prairie & Kalff 1986). Phosphorus yield from watersheds less than 5,000 acres was particularly sensitive to watershed size.

The importance of these findings is twofold. First, using markedly different watershed sizes for control and treatment areas could introduce a bias in response. If the practice installed influenced sediment delivery, a smaller watershed will react differently from a larger one. Second, because sediment delivery per unit area is greater in smaller watersheds, there may be differences in flushing of sediment stored in channels of different sized watersheds.

A final consideration may be whether the stream is intermittent or permanent. Intermittent streams appear to exhibit a first flush phenomenon after extended dry periods where concentrations of nutrients are higher than anticipated based on discharge measurements. Also, the biotic community in an intermittent stream is controlled, in large part, by the periodic lack of flow. Some biotic community changes may be influenced more by flow than water quality changes. This is not to say that intermittent watersheds are inappropriate for study. Intermittent watersheds are smaller, and therefore greater control over watershed land activities can be exercised.

**Example 5–3** Watershed scale

One of the objectives of the St. Albans Bay watershed RCWP was to determine the effect of implementing BMPs on the water quality of the bay and its tributaries. Water quality monitoring, both chemical and biological, was conducted in the bay and four tributaries. At each stream monitoring location, flow was continuously recorded and samples were taken at 8-hour intervals and composited. Bacteria grab samples were taken weekly.

The watersheds were 3,400, 6,000, 3,800, and 14,400 acres in area. Trend analysis applied to the bacteria data revealed that bacteria abundance declined significantly in all tributary streams by 60 to 70 percent. The decline was attributed to bacterial dieoff during manure storage and greater incorporation of manure applied to fields, both of which were BMPs.

## 614.0505 References

Clausen, J.C., F.I. Potter, III, W.S. Douglas, W.E. Jokela, and J.W. Williams. 1990. Conservation tillage use of pesticides for corn: surface water contamination. Vermont Water Resour. Res. Cent., Burlington, VT.

Harlin, J.M. 1984. Watershed morphometry and time to hydrograph peak. J. of Hydrol. 67:141-154.

Holdridge, L.R. 1962. The determination of atmospheric water movements. Ecology. 43(1):1-9.

LeClerg, E.L., W.H. Leonard, and A.G. Clark. 1962. Field plot techniques. 2nd. Ed. Burgess Pub. Co. Minneapolis, MN.

Lee, D.R. 1977. A device for measuring seepage flux in lakes and estuaries. Limnol. Oceanog. 22(1):140-147.

Morton, T.G., A.J. Gold, and W.M. Sullivan. 1988. Influence of overwatering and fertilization on nitrogen losses from home lawns. J. Environ. Qual. 17(1):124-130.

Odum, E.P. 1984. The mesocosm. BioScience. 34(9):558-562.

Ruhe, R.V. 1975. Geomorphology: geomorphic processes and surficial geology. Houghton Mifflin Co., Boston, MA.

T.-Prairie, Y., and J. Kalff. 1986. Effect of catchment size on phosphorus export. Water Resour. Bull. 22(3):465-470.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 6 Variable Selection

(450–VI–NWQH, September 2003)

# Chapter 6 Variable Selection

Contents:



Tables

# Chapter 6 Variable Selection

## 614.0600 Introduction

The term *variable* is used in this handbook to denote water quality characteristics that exhibit variability (e.g., algae counts, dissolved oxygen, nutrient concentrations). Although the term *parameter* is often used interchangeably with the term variable, in this handbook parameter is meant to be quantities that characterize statistical samples (mean, variance).

The selection of water quality variables to include in a project requires consideration of several factors. The tendency is to sample for more variables than are generally needed. The major reason for not sampling "full suite" is that there are trade-offs in the study design. Water quality monitoring is expensive, and resources committed to unnecessary water quality characteristics may be at the expense of a successful experimental design. Where funding is limited, fewer stations, and the number of samples at each station, can be monitored when more water quality variables are added to a project. As a final test in considering which water quality variables to include in a project, a written justification statement is recommended for each variable. If the justification is weak, the variable may be of low priority and might not be essential.

This chapter discusses the various factors that affect the selection of water quality variables. Also several methods for prioritizing variables are presented including: variable matrices, variable cross-correlations, and the probability of exceeding standards.

Water quality variables receive various names and are classified differently in different references. For this chapter, the naming conventions that appear in American Public Health Association's standard methods (APHA 1989) were used. Two excellent references describe the meaning of various water quality variables. They are Hem (1970) and McKee and Wolf (1963). Additional descriptions are in IHD-WHO (1978), McNeely et al. (1979), and Stednick (1991). The importance of biological characteristics is described in Cairns et al. (1982), Plafkin et al. (1989), Terrell and Perfetti (1989), and Weber (1973).

## 614.0601 Factors affecting variables

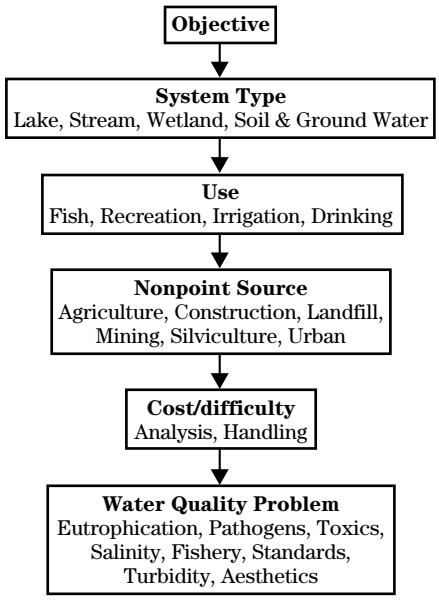
Considerations that influence the variables to sample include the study objectives, the type of water resource, the use or classification of the water body, the type of nonpoint source activity, the difficulty or cost in analysis of the variable, and the water quality problem. An overall schematic of these considerations is given in figure 6–1.

### (a) Objectives

A properly stated objective assists in defining the water quality variables to monitor. In fact, selecting the water quality variables may result in a redefinition or clarification of the objectives in a feedback manner. The *constraint* part of the objective may specifically mention the water quality variables (chapter 3). For example, the following objective statement from chapter 3 clearly indicates that the variable to measure is fecal coliform levels:

*To determine the effect of implementing conservation practices on fecal coliform levels in Long Lake.*

**Figure 6–1** Water quality variable selection



## (b) System type

The type of water resource being studied also influences the variables selected. Table 6–1 indicates that the appropriate variables of interest differ primarily between subsurface systems, such as soil water and ground water, and surface water systems, including lakes, streams, and wetlands. For example, chemical nutrients may be important to all systems, but particulate forms of nutrients are meaningful only for lake, stream, and wetland systems and not for soil water or ground water systems.

In addition, different variable selections may be made for intermittent or permanent stream systems (USDA 1976). Generally, more variables can be justified for a perennial water body than for an intermittent one. The biota in intermittent streams is limited by the flow regime, and therefore may not be good water quality indicators in that situation.

Tables 6–1 through 6–8 provide a list of potential water quality variables to consider when designing a monitoring program.

**Table 6–1** Water quality variable groups by water resource system type matrix (general guidelines; in some circumstances variables that are not marked should be considered)

| Variable | System type | | | | |
|---|---|---|---|---|---|
| | Lake | Stream | Wetland | Soil water | Ground water |
| **Physical** | | | | | |
| Dissolved oxygen | X | X | X | | |
| Discharge | X | X | X | X | X |
| Embeddedness | | X | | | |
| Habitat assessment | | X | | | |
| Riffle/pool ratio | | X | | | |
| Salinity | X | X | X | X | X |
| Secchi disk transparency | X | | | | |
| Specific conductance | X | X | X | X | X |
| Substrate characteristics | X | X | | | |
| Suspended solids | X | X | X | | |
| Temperature | X | X | X | | |
| Total dissolved solids | X | X | X | X | X |
| Turbidity | X | X | X | | |
| **Chemical** | | | | | |
| $BOD_5$ | X | X | X | | |
| Inorganic nonmetals: Cl, F | | X | X | X | X |
| Nutrients - N, P dissolved | X | X | X | X | X |
| total or particulate | X | X | X | | |
| Metals: As, Ca, Cd, Cr, Co, Cu, Fe, Hg, K, Pb, Mg, Mn, Na, Ni, Zn | X | X | X | X | X |
| pH | X | X | X | X | X |
| **Biological** | | | | | |
| Bacteria | X | X | X | X | X |
| Chlorophyll 'a' | X | X | | | |
| Indices (SCI, BI, IBI)* | X | X | | | |
| Invertebrates | X | X | X | | |
| Fish | X | X | | | |
| Macrophyton | X | X | X | | |
| Periphyton | X | X | | | |
| Plankton (algae) | X | X | | | |
| Protozoa | X | X | | | |

* SCI = Sequential Comparison Index
BI = Beck's Biotic Index
IBI = Index of Biotic Integrity

## (c) Designated use

Variable selection may be modified by the intended or designated use of a water body (US EPA, 1981b). A water body being used for recreation, including aesthetic uses, might emphasize variables associated with sediment, nutrients, toxic and biological characteristics because all these are visual or affect visual characteristics of water bodies. However, water having an irrigation use might not include biological variables (table 6–2). Water intended to be used for drinking, recreation, or fisheries might include analysis of biological and toxic substances.

## (d) Pollutant source

The nonpoint source of the water quality problem also influences variable selection, as will certain activities for those sources. The major nonpoint source categories include:

- agriculture
- construction
- landfill
- mining
- silviculture
- urban

**Table 6–2** Water quality variable groups by intended water resource use (general guidelines; in some circumstances variables that are not marked should be considered)

| Variable | Intended use | | | | |
|---|---|---|---|---|---|
| | Fish | Recreation contact | Aesthetics | Irrigation | Drinking |
| **Physical** | | | | | |
| Dissolved oxygen | X | | X | | X |
| Discharge | | | | | |
| Salinity | X | | | X | X |
| Secchi disk transparency | X | X | X | | |
| Specific conductance | | | | X | X |
| Suspended solids | X | X | X | X | X |
| Temperature | X | | | | |
| Total dissolved solids | X | | | X | X |
| Turbidity | X | X | X | X | X |
| **Chemical** | | | | | |
| $BOD_5$ | X | | X | | |
| Inorganic nonmetals: Cl, F | X | | | X | X |
| Nutrients - N, P dissolved | X | | X | X | |
| total or particulate | X | | X | | |
| Metals: As, Ca, Cd, Cr, Co, Cu, Fe, Hg, K, Pb, Mg, Mn, Na, Ni, Zn | X | X | | X | X |
| pH | X | | | X | X |
| **Biological** | | | | | |
| Bacteria | | X | | | X |
| Chlorophyll 'a' | X | | X | | X |
| Indices (SCI, BI, IBI) | X | | X | | |
| Invertebrates | X | | | | |
| Fish | X | | | | |
| Macrophyton | X | | X | | |
| Periphyton | | | | | |
| Plankton (algae) | X | | X | | X |
| Protozoa | | X | | | |

Within each of these categories are specific activities that influence certain water quality variables. Agricultural activities are shown in table 6–3. Almost all agricultural activities justify monitoring dissolved oxygen or BOD, flow, suspended solids, nutrients in all forms, and invertebrates. Most agricultural activities might also influence turbidity and bacteria. Pesticide monitoring requires fewer variables to analyze, although the metabolites should also be monitored. In addition, metals can be added with certain pesticides, such as copper sulfate or a zinc fungicide.

Pesticides in field runoff are carried in both dissolved and particulate forms. Generally, the concentration of the pesticide is greater in the particulate form; however, the annual mass export may be greater in the dissolved form.

Three forms of nutrients (total, dissolved, and particulate) are appropriate for most agricultural activities. However, all three forms may not need to be analyzed since they are highly related. Including the other forms in the monitoring study would require justification.

**Table 6–3** Water quality variable groups by nonpoint source activity (general guidelines; in some circumstances variables that are not marked should be considered)

| Variable | Activity | | | | | | |
|---|---|---|---|---|---|---|---|
| | Field runoff* | Pesticide | Fertilizer | Barnyard/feedlot | Stream access | Pasture | Animal waste |
| **Physical** | | | | | | | |
| Dissolved oxygen | X | | X | X | X | X | X |
| Discharge | X | X | X | X | X | X | X |
| Salinity | | | | | X | | X |
| Secchi disk transparency | X | | X | X | X | X | X |
| Specific conductance | | | | | | | |
| Suspended solids | X | | | X | X | X | X |
| Temperature | | | | | | | |
| Total dissolved solids | X | | | X | X | X | X |
| Turbidity | X | | | X | X | X | X |
| **Chemical** | | | | | | | |
| $BOD_5$ | X | | X | X | X | | X |
| Inorganic nonmetals: Cl, F | | | | | | | |
| Nutrients - N, P dissolved | X | | X | X | X | X | X |
| total or particulate | X | | | X | X | X | X |
| Metals: As, Ca, Cd, Cr, Co, Cu, Fe, Hg, K, Pb, Mg, Mn, Na, Ni, Zn | | | | | | | |
| pH | | | | | | | |
| **Biological** | | | | | | | |
| Bacteria | X | | | X | X | X | X |
| Chlorophyll 'a' | X | X | X | X | X | X | X |
| Indices (SCI, BI, IBI) | X | X | X | X | X | X | X |
| Invertebrates | X | X | X | X | X | X | X |
| Fish | | X | | | | | |
| Macrophyton | X | X | X | X | X | X | X |
| Periphyton | X | X | X | X | X | X | X |
| Plankton (algae) | X | X | X | X | X | X | X |
| Protozoa | | X | | | | | |

* Includes runoff from hayland, rangeland, and cropland.

An activity by variable matrix for additional nonpoint source categories is given in table 6–4. Most of the activities have the potential to directly influence discharge, sediment and nutrients. Therefore, additional indirect effects may occur to oxygen, transparency, and several biological characteristics. Landfill leachate may contain a wide range of water quality constituents; therefore, a large number of physical, chemical, and biological variables are usually monitored.

The water quality variables selected for mining operations would change with the type of mining. Acid mine drainage, associated with coal mining, might involve monitoring several physical variables, as well as metals and biological characteristics. Mining of taconite, sylvite, rock phosphate, and sand and gravel might imply other, more specific variables.

**Table 6–4** Water quality variable groups by construction, landfill, and mining activities (general guidelines; in some circumstances variables that are not marked should be considered)

| Variable | Activity | | |
|---|---|---|---|
| | Construction | Landfill | Mining |
| **Physical** | | | |
| Dissolved oxygen | | X | X |
| Discharge | X | X | X |
| Salinity | | X | |
| Secchi disk transparency | X | X | |
| Specific conductance | | X | X |
| Suspended solids | X | X | X |
| Temperature | X | | X |
| Total dissolved solids | X | X | X |
| Turbidity | X | | X |
| **Chemical** | | | |
| $BOD_5$ | | X | |
| Inorganic nonmetals: Cl, F | | X | X |
| Nutrients - N, P dissolved | X | X | |
| total or particulate | X | X | |
| Metals: As, Ca, Cd, Cr, Co, Cu, Fe, Hg, K, Pb, Mg, Mn, Na, Ni, Zn | | X | X |
| pH | | X | X |
| **Biological** | | | |
| Bacteria | | X | |
| Chlorophyll 'a' | X | X | X |
| Indices (SCI, BI, IBI) | X | X | X |
| Invertebrates | X | X | X |
| Fish | X | X | X |
| Macrophyton | | X | X |
| Periphyton | | X | X |
| Plankton (algae) | | X | X |
| Protozoa | | X | |

Several activities are associated with silvicultural operations (table 6–5). Of these activities, road construction, grazing, and site preparation have the greatest potential to influence the most water quality characteristics. Timber harvesting alone only influences the water quality variables affected by riparian vegetation removal. Transporting the timber out of the forest causes most of the potential water quality effects. However, water yield changes associated with timber harvesting can have additional water quality impacts.

Urban activities may influence several physical, chemical, and biological variables, as indicated in table 6–6. Impervious areas and combined sewer overflows (CSOs) influence the same variables directly and indirectly because their primary sources of pollutants are runoff from impervious surfaces.

**Table 6–5** Water quality variable groups by silvicultural activity (general guidelines; in some circumstances variables that are not marked should be considered)

| Variable | Activity | | | | |
|---|---|---|---|---|---|
| | Harvesting | Roads | Site preparation | Grazing | Pesticide |
| **Physical** | | | | | |
| Dissolved oxygen | X | X | X | X | |
| Discharge | X | X | | X | |
| Salinity | | | | | |
| Secchi disk transparency | | X | | X | |
| Specific conductance | | | | | |
| Suspended solids | | X | X | X | |
| Temperature | X | | | | |
| Total dissolved solids | | X | X | X | |
| Turbidity | | X | X | X | |
| **Chemical** | | | | | |
| $BOD_5$ | | | | | |
| Inorganic nonmetals: Cl, F | | | | | |
| Nutrients - N, P dissolved | | X | X | X | |
| total and particulate | | X | X | X | |
| Metals: As, Ca, Cd, Cr, Co, Cu, Fe, Hg, K, Pb, Mg, Mn, Na, Ni, Zn | | | | | |
| pH | | | | | |
| **Biological** | | | | | |
| Bacteria | | | X | | |
| Chlorophyll 'a' | | X | X | X | X |
| Indices (SCI, BI, IBI) | | X | X | X | X |
| Invertebrates | X | X | X | X | X |
| Fish | X | X | X | X | X |
| Macrophyton | X | X | X | X | X |
| Periphyton | X | X | X | X | X |
| Plankton (algae) | X | X | X | X | X |
| Protozoa | | | | | X |

**Table 6–6** Water quality variable groups by urban activity (general guidelines; in some circumstances variables that are not marked should be considered)

| Variable | Impervious areas | Lawns | Combined sewer overflows | Pets |
|---|---|---|---|---|
| **Physical** | | | | |
| Dissolved oxygen | X | X | X | X |
| Discharge | X | X | X | |
| Salinity | | | | |
| Secchi disk transparency | X | X | X | X |
| Specific conductance | X | | | |
| Suspended solids | X | | X | |
| Temperature | | | | |
| Total dissolved solids | | | | |
| Turbidity | X | | X | |
| **Chemical** | | | | |
| $BOD_5$ | X | | X | |
| Inorganic nonmetals: Cl, F | X | | X | |
| Nutrients - N, P dissolved | X | X | X | X |
| total or particulate | X | | X | X |
| Metals: As, Ca, Cd, Cr, Co, Cu, Fe, Hg, K, Pb, Mg, Mn, Na, Ni, Zn | X | | X | |
| pH | | | | |
| **Biological** | | | | |
| Bacteria | X | | X | X |
| Chlorophyll 'a' | X | X | X | X |
| Indices (SCI, BI, IBI) | X | X | X | X |
| Invertebrates | X | X | X | |
| Fish | X | X | X | |
| Macrophyton | X | X | X | |
| Periphyton | X | X | X | X |
| Plankton (algae) | X | X | X | X |
| Protozoa | X | | X | X |

## (e) Analysis difficulty

The difficulty or cost of analysis should be considered when selecting water quality variables. Table 6–7 presents some relative costs of analysis for specific water quality variables. These costs are relative to the cost of analyzing the sample for either pH or conductance. When water quality characteristics are highly related, but the analysis cost of one is much cheaper than the other, the less expensive variable could be selected. For example, analysis of turbidity is less costly than suspended solids, both of which are less expensive than total solids. Also, nitrate nitrogen is cheaper than ammonia nitrogen or total Kjeldahl nitrogen because digestion of the sample is not required.

The range and level of accuracy are also important. For example, Inductively Coupled Plasma (ICP) emission spectroscopy will determine elements cheaper, but not as accurately, as atomic absorption. Sample holding times also influence parameter selections. For example, nitrate and ortho-phosphate are recommended by the Environmental Protection Agency (USEPA 1983) to be analyzed within 48 hours of collection, whereas nitrate+nitrite and total phosphorus can be held for 28 days before analysis if preserved (see table 11–1 in chapter 11).

## (f) Water quality problem

Finally, the water quality problem itself influences the variables to sample. The major water quality problems are summarized in table 6–8 along with the appropriate water quality variables. Eutrophication problems require monitoring of several physical, chemical, and biologic characteristics. Excess algae might suggest sampling of dissolved oxygen and temperature, flow for mass balance purposes, turbidity or secchi disk transparency, nutrients, plankton abundance/type, and chlorophyll 'a' concentrations. Because many of these variables are related, not all would be needed to detect changes. Also, an index, such as Carlson's Trophic State Index (TSI) could be used (Carlson 1977). It combines some of these variables.

A problem associated with either a standard violation or a toxic substance might focus on monitoring that particular standard or toxicant.

**Table 6–7** Relative cost of analysis for water quality variables (based on Beetem et al. 1980)

| Variable | Cost ($/analysis) dissolved | total | particulate |
|---|---|---|---|
| **Ions** | | | |
| Ca, Mg | 4.70 | | 12.00 |
| Na, K, $SiO_2$ | 3.40 | | 10.00 |
| Cl | 5.35 | | |
| F | 5.25 | | |
| $SO_4$ | 5.80 | | |
| **Trace metals** | | | |
| As, Hg | 5.20 | | 22.70 |
| Cd, Co, Cu, Pb, Ni | 6.20 | | 10.00 |
| Cr | 10.50 | | 2.90 |
| Fe, Mn | 3.40 | | 10.00 |
| Zn | 4.20 | | 10.00 |
| **Physical** | | | |
| Alkalinity | | 3.55 | |
| pH | | 1.00 | |
| Specific conductance | | 1.00 | |
| Total solids | | 8.95 | |
| Turbidity | | 1.80 | |
| **Nutrients** | | | |
| $NH_3$, $NO_3$, $NO_2$ | | 3.40 | |
| TKN | | 8.90 | |
| Total P | | 9.55 | |
| $PO_4$ | | 3.40 | |

**Table 6–8** Water quality variable groups by water quality problem (general guidelines; in some circumstances variables that are not marked should be considered)

| Variable | Problem | | | | | | |
|---|---|---|---|---|---|---|---|
| | Aesthetics | Bacteria | Algae | Macrophytes | Salinity | Sediment | Toxics |
| **Physical** | | | | | | | |
| Dissolved oxygen | X | | X | | | | |
| Discharge | | | X | | | | |
| Salinity | | | | X | | | |
| Secchi disk transparency | X | | X | | | X | |
| Specific conductance | | | | | X | | |
| Suspended solids | X | | | | X | | |
| Temperature | | | | | | | |
| Total dissolved solids | | | | | X | | |
| Turbidity | X | | | | | X | |
| **Chemical** | | | | | | | |
| $BOD_5$ | | | | | | | |
| Inorganic nonmetals: Cl, F | | | | | X | | |
| Nutrients - N, P dissolved | X | | X | X | | | |
| total or particulate | X | | X | X | | | |
| Metals: As, Ca, Cd, Cr, Co, Cu, Fe, Hg, K, Pb, Mg, Mn, Na, Ni, Zn | | | | | | | |
| pH | | | | | | | |
| **Biological** | | | | | | | |
| Bacteria | | X | | | | | |
| Chlorophyll 'a' | X | | X | | | | X |
| Indices (SCI, BI, IBI) | X | | X | | X | | X |
| Invertebrates | | | | | X | | X |
| Fish | | | | | X | | X |
| Macrophyton | X | | | X | X | | |
| Periphyton | X | | X | | | | X |
| Plankton (algae) | X | | X | | | | X |
| Protozoa | | | | | | | X |

# 614.0602 Prioritizing variables

Because virtually hundreds of water quality variables exist and are therefore candidates for monitoring, a method for prioritizing their selection is important. The four basic approaches for prioritizing water quality variables are ranking, activity matrices, correlations, and probability of exceeding a standard.

## (a) Ranking

Sanders et al. (1983) suggest a hierarchical approach of:

- **Primary**—water quantity variables that serve as a carrier of water quality, e.g., discharge, volume, head
- **Secondary**—water quality variables that are the result of aggregated effects, e.g., temperature, pH, conduction, dissolved oxygen, turbidity, anions, cations
- **Tertiary**—water quality variables that produce aggregated effects, e.g., radioactivity, suspended matter

Variables higher in the hierarchy would be selected over lower-ranked variables. Greater priority should be placed on primary variables than on secondary variables when the number of variables to monitor need to be limited.

Another example of prioritizing suggests two levels of analysis (USEPA 1981 a, b). Level I is the minimum list of variables needed to evaluate program effectiveness associated with a particular water quality problem and use of the water resource. For example, chlorophyll ‘a’ would be the level I variable for a stream experiencing excessive algal growth and being used for drinking water. Level II includes more detailed, multiparameter variables. For the example above, nitrogen and phosphorus species would be added to the chlorophyll 'a' sampling.

## (b) Activity matrices

The water quality variable matrices given in tables 6–1 through 6–6 serve as a second method in selecting water quality variables. Ponce (1980) assigned values of 1, 2, or 3 to primary, secondary, and tertiary sampling priority codes in a forest management activity matrix with water quality variables. This method combines the ranking and activity matrices approaches. The activity matrix variable provides an initial list of variables to consider when planning the monitoring study.

## (c) Correlations

Correlations between variables can be used to reduce the variable list. A number of water quality variables are often correlated. Total phosphorus often is highly related to ortho-phosphorus. In lake systems, total phosphorus has been reported to be highly related to secchi disk transparency and chlorophyll 'a' (Reckhow & Chapra 1983). Other variables that might be expected to exhibit correlations are conductivity and dissolved solids and suspended solids and turbidity. Since these variables may be highly related, one variable could be dropped from the monitoring program or monitored less frequently.

Correlation coefficients are readily computed in most statistical packages. This topic is further discussed in part 615 of this handbook. The correlation coefficient ($r$) can be determined from:

$$r = \frac{\sum\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum\left(X_i - \overline{X}\right)^2 \sum\left(Y_i - \overline{Y}\right)^2}} \qquad [6–1]$$

where:

$\overline{X}$ and $\overline{Y}$ = the means of the variables X and Y, respectively

$X_i$ and $Y_i$ = individual values of variables X and Y, respectively

To use correlation coefficients, some monitoring data would have to be available either from a previous study or from preliminary monitoring in the watershed of interest.

Another consideration for correlated variables is the proximity of the range in values to the detection limit for that variable. Values below detection limits, termed censored values, require adjustments when calculating means and variances. Variables that do not include censored values are preferred.

Example 6–1 illustrates variable correlations.

**Example 6–1** Variable correlations

Muddy Bay is experiencing impairment caused by excessive sedimentation and eutrophication. Both nitrogen and phosphorus are believed to contribute to the problem. Appropriate variables include:

- Turbidity
- Total Suspended Solids (TSS)
- Volatile Suspended Solids (VSS)
- Total Phosphorus (TP)
- Ortho-Phosphate (OP)
- Total Kjeldahl Nitrogen (TKN)
- Ammonia Nitrogen ($NH_3$)
- Nitrate Nitrogen ($NO_3$)

Based on cost data, these analyses would cost a total of $40.45 per site visit (1980 dollars). You have $25 budgeted to monitor water quality per sampling period. Which parameters would you monitor?

Note that based on sampling in Muddy Bay during 1 year, the following correlation matrix was developed.

**Correlation matrix (*r*)**

| | Turbidity | TSS | TKN | $NO_3$ | TP |
|---|---|---|---|---|---|
| TSS | 0.577 | 1.000 | --- | --- | --- |
| VSS | 0.764 | 0.855 | --- | --- | --- |
| $NH_3$ | --- | --- | 0.836 | 0.281 | --- |
| $NO_3$ | --- | --- | -0.057 | 1.000 | --- |
| OP | --- | --- | --- | --- | 0.915 |

The correlations between TP and OP, TKN and $NH_3$, and TSS and VSS are significant and very high. Adequate monitoring could be achieved by choosing TSS, total P, and TKN for less than $25 to meet sedimentation and eutrophication objectives. In nitrogen-limited systems, measurement of $NO_3$ should be included.

## (d) Probability of exceeding standard

An alternative method for determining the priority of variables to monitor would be to select those with the highest probability of exceeding a particular standard (Moser & Huibregtse, 1976). To determine this probability requires knowledge of the mean ($\overline{X}$), standard deviation ($S$), and numerical standard value ($X_{std}$) not to be exceeded. The probability is determined from the Z-statistic as:

$$Z = \frac{X_{std} - \overline{X}}{S} \qquad [6–2]$$

Using a standard Z-table (appendix A), the probability would be obtained. Not all variables have adopted numerical values for standards. For example, nitrogen and phosphorus generally are not included in lists of numeric standards. In such cases a eutrophication value, such as 0.05 mg/L for total phosphorus could be used. Another alternative would be to set a concentration goal to achieve and substitute that for a standard value.

Example 6–2 illustrates this approach.

**Example 6–2** Probability of exceeding a standard

Using the St. Albans Bay data, the mean fecal coliform bacteria count for Jewett Brook in 1989 was 149 organisms/100 mL. The standard deviation was 493 organisms/100 mL. Using a water quality standard of 200 organisms/100 mL, what is the probability of exceeding the fecal coliform standard?

$$Z = \frac{200 - 149}{493} = 0.10$$

From a standard Z-table (appendix A), the probability would be 0.4602 or 46 percent. This probability may be higher than that for other water quality variables, and therefore would be given higher priority.

# 614.0603 References

American Public Health Association. 1989. Standard methods for the examination of water and wastewater. 17th Ed. American Public Health Assoc. Washington, DC.

Beetem, W.A., L.C. Friedman, G.R. Perryman, and C.A. Wottersan (eds). 1980. 1981 Water quality laboratory services catalog. U.S. Geol. Surv. Open-file Rep. 80-1279.

Cairns, J., Jr., A.L. Buikema, Jr., D.S. Cherry, E.E. Herricks, R.A. Matthews, B.R. Niederlehner, J.H. Rodgers, Jr., and W.H. van der Schalie. 1982. Biological monitoring in water pollution. Pergamon Press, Elmsford, NY.

Carlson, R.E. 1977. A trophic state index for lakes. Limnol. Ocean. 22:361- 369.

Hem, J.D. 1970. Study and interpretation of the chemical characteristics of natural water. U.S.. Geol. Surv. Water-Supply Pap. 1473. US GPO, Washington, DC.

IHD-WHO Working Group on the Quality of Water. 1978. Water quality surveys—a guide for the collection and interpretation of water quality data. Stud. and Rep. in Hydrol. No. 23. United Nations Educational, Scientific, and Cultural Organization. Paris, France, and World Health Organization, Geneva, Switzerland.

McKee, J.E., and H.W. Wolf. 1963. Water quality criteria. 2nd Ed. California State Water Quality Control Board. Publ. No. 3-A. Sacramento, CA.

McNeely, R.N., V.P. Neimanic, and L. Dwyer. 1979. Water quality sourcebook—a guide to water quality parameters. Inland Water Directorate, Water Quality Branch. Ottawa, Canada. Cat. No. EN 37- 541.

Moser, J.H., and K.R. Huibregtse. 1976. Handbook for sampling and sample preservation of water and wastewater. USEPA 600/4-76-049.

Plafkin, J.L., M.T. Barbour, K.D. Portor, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. USEPA 444/4-89-001. Office of Water. Washington, DC.

Ponce, S.L. 1980. Water quality monitoring programs. WSDG Tech Pap. No. WSDG-TP-00002. Watershed Systems Development Group, USDA Forest Service. Ft. Collins, CO.

Reckhow, K.N. and S.C. Chapra. 1983. Engineering approaches for lake management. Vol. 1: Data analysis and emphirical modeling. Butterworth Publ. Woburn, MA.

Sanders, T.G., R.C. Ward, J.C. Loftis, T.D. Steele, D.D. Adrian, and V. Yevjevich. 1983. Design of networks for monitoring water quality. Water Resourc. Pub. Littleton, CO.

Stednick, J.D. 1991. Wildland water quality sampling and analysis. Academic Press, Inc., San Diego, CA.

Terrell, C.R., and P.B. Perfetti. 1989. Water quality indicators guide: surface waters. USDA, Soil Conserv. Serv., SCS-TP-161. Washington, DC.

United States Department of Agriculture, Soil Conservation Service. 1976. A general guideline for the assessment of water quality. Tech. Rel. No. 58. Eng. Div., Washington, DC.

United States Environmental Protection Agency. 1981a. Conceptual framework for assessing agricultural nonpoint source projects. NC State Univ., Raleigh, NC.

United States Environmental Protection Agency. 1981b. Guidelines for evaluation of agricultural nonpoint source water quality projects. NC State Univ., Raleigh, NC.

United States Environmental Protection Agency. 1983. Methods for chemical analysis of water and wastes. US EPA 600/4-79-020. Environ. Monit. and Support Lab., Cincinnati, OH.

Weber, C.I. (Ed.) 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. US EPA 670/4-73-001. Off. Res. and Dev., Cincinnati, OH .

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 7 Sample Type

(450–VI–NWQH, September 2003)

# Chapter 7 Sample Type

Contents:

# Chapter 7 Sample Type

## 614.0700 Introduction

If water quality did not vary in space or in time, there would be little reason to collect more than one sample to describe the quality of a particular water body. However, water quality does vary spatially and temporally. Both random and deterministic components (fig. 7–1) are found in most water quality data. Variations in water quality data are caused by seasonal differences, trends, and the randomness associated with rainstorms. For example, suspended solids concentrations increase during stormflow, especially during the early part of the storm (Shelly & Kirkpatric 1975). Therefore, because of these temporal and spatial variations, samples must be taken from the entire population of water quality data possible.

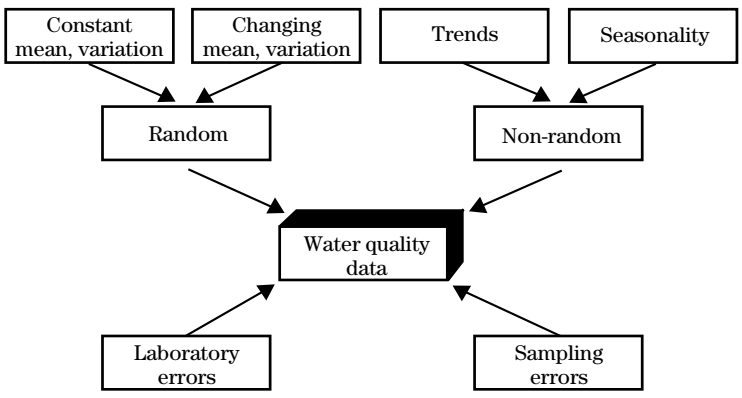
The four types of water quality samples that can be collected are grab, composite, integrated, or continuous. The sample type selected is governed by the study objectives, the variable to sample, and whether concentration or mass is the desired outcome. Composite samples are appropriate for most monitoring study objectives, whereas grab sampling is recommended for a few objectives directed toward reconnaissance sampling (table 7–1). Continuous samples are appropriate only for research and fate and transport studies.

The variable to sample influences the sample type as well. For example, bacteria samples must be taken as grab samples with sterilized bottles and cannot be stored in the field as a composite sample. The concentrations of other variables change dramatically during storage and therefore are inappropriate for compositing. These include all dissolved gases, chlorine, pH, temperature, and sulfide (APHA 1989). Water quality variables that correlate highly with stream velocity, especially those related to suspended sediment concentrations, may need to be sampled with depth integrated samplers. Grab samples may be insufficient to determine mass loading values unless the concentrations are correlated to discharge (Baun 1982).

**Table 7–1** Sample type as a function of monitoring study objective

| Objective | Grab | Integrated or composite | Continuous |
|---|---|---|---|
| 1. Baseline | X | X | |
| 2. Trend | X | X | |
| 3. Fate & transport | | X | X |
| 4. Problem definition | X | X | |
| 5. Critical areas | X | X | |
| 6. Compliance | X | X | |
| 7. Conservation practice effectiveness | | X | |
| 8. Program effectiveness | | X | |
| 9. Wasteload allocations | | X | |
| 10. Model evaluation | | X | |
| 11. Research | | X | X |

**Figure 7–1** Factors contributing to variability in water quality data



# 614.0701 Grab samples

A grab sample is a discrete sample that is taken at a specific point and time (APHA 1989; Ponce 1980). Grab samples may not be representative of the water quality of the body of water being sampled. For example, the water quality may vary with depth or distance from the streambank. Samples at a single location in a lake or a single well are really grab samples. For lakes and ground water, variable concentrations may vary with location and depth. For example, nitrate concentrations have been found to be stratified in some water table aquifers in the Midwest. Also, since water quality often varies with time, grab samples may not represent temporal variations.

Grab samples can be collected manually by hand or automatically with a sampler.

# 614.0702 Composite samples

A series of grab samples, usually collected at different times and lumped together, are considered composite samples. However, composite samples typically are taken only at one point. These samples can be either time-weighted or flow-weighted. The collection of composite samples generally is done with the aid of an automatic sampler, as described in chapter 9, although manual techniques could be used as well. A distinct advantage of the composite sample is that a savings in laboratory and field costs can be realized. Also, compositing will reduce sample-to-sample variability.

## (a) Time-weighted composite

Time-weighting is the most common type of water quality compositing. For this type of sample, a fixed volume of sample is collected at prescribed time intervals in either a large composite bottle or separate bottles for compositing later. With automatic samplers, the time interval can range from 1 minute to 100 hours, and the volume collected can range from 10 mL to 990 mL, although larger volumes are possible. Equation 9–1 in chapter 9 can be used to determine the number of samples (n) to take to make up a composite, where n is a function of the variability in the data and the desired precision. For water quality variables where the length of the composite time is greater than the prescribed holding times (USEPA 1983), the collection bottles may be pre-acidified for preservation.

## (b) Flow-weighted composite

Time-weighted compositing has been criticized as being inappropriate for mass loading calculations and inaccurate where the discharge and concentrations vary (Baun 1982; Shelly & Kirkpatric 1975). Also, the time interval may miss peak concentrations during peak discharges. Therefore, flow-weighted compositing is an alternative to time-compositing. Where flow-weighted compositing is used, a sample is taken after a specified volume ($l^3$) of flow has passed the monitoring station. This type of sampling requires automatic equipment that monitors stream stage and

calculates discharge. A number of automatic samplers offer this function, or a data logger can be used.

To sample in this manner, the stage-discharge relationship must be known for the monitoring location. Stage-discharge relationships require a great deal of effort to develop unless a calibrated flow devise, such as a weir or a flume, is used.

Flow-weighted compositing also can be achieved using certain types of passive samplers. A passive sampler is one that collects a water quality sample by action of the flow of water itself. A number of these types of devices are described further in chapter 9.

## 614.0703 Integrated samples

A specific type of grab sample is a depth-integrated sample (USGS 1977). Such a sample may account for velocity or stratification induced changes with depth, but temporal variations would not be integrated.

Multipoint sampling at a station may be necessary because of the horizontal and vertical variations in water quality. The U.S. Geological Survey recommends that streams should be sampled using a depth integrated sampler whenever practical (USGS 1977) except when the stream is too shallow to obtain that type of sample.

For variations across the stream, samples can be collected using either the Equal Width Increment (EWI) method or the Equal Discharge Increment (EDI) method. With the EWI method, depth integrated samples are collected at equally spaced intervals at the cross section. All subsamples are then composited. The EDI method requires knowledge of streamflow discharge by subsection in the cross section. The section is divided into equal discharge subsections, which are then sampled.

Depth-integrated samples may also be appropriate for both lake and ground water systems. In lakes, depth integration can be achieved by sampling each lake strata, by obtaining a sample of the entire water column with a hose, or by automatic devices or pulleys that collect at different depths over time.

Different ground water strata can be sampled with certain types of bailers or with multilevel wells and samplers.

## 614.0704 Continuous samples

Continuous sampling is rare in nonpoint source pollution studies and is typically used for research purposes (table 7–1). Continuous monitoring can be used for any water quality variable that is measured using electrometric methods (table 7–2). This would exclude analysis of metals and organics.

Several problems are encountered when using continuous sampling. Most electrodes are temperature dependent and have temperature limits beyond which they cease to function. Electrodes normally cannot be placed in areas of rapid water velocity, which influences readings by the probe. However, in-stream stilling wells can be used to reduce this effect.

Several manufacturers produce submersible, multiple recording probes for such variables as pH, dissolved oxygen, conductance, and depth. These probes have been widely used in lake systems.

**Table 7–2** The suitability of various water quality variables for continuous monitoring (based on APHA 1989)

| Suitable | Not suitable |
|---|---|
| Ammonia | Metals |
| Chloride | Organic compounds |
| Conductivity | Pesticides |
| Cyanide | |
| Dissolved oxygen | |
| Fluoride | |
| Inorganic nonmetals | |
| Nitrate | |
| pH | |
| Salinity | |
| Temperature | |

## 614.0705 References

American Public Health Administration. 1989. Standard methods for the examination of water and wastewater. 17th Ed. Amer. Pub. Health Assoc., Washington, DC.

Baun, K. 1982. Alternative methods of estimating pollutant loads in flowing water. Tech. Bull. No. 133. WI Dep. Nat. Res. Madison, WI.

Ponce, S.L. 1980. Water quality monitoring programs. WSDG Tech. Pap. WSDG-TP-0002. Watershed Systems Development Group, U.S. Dep. Agric., Forest Serv., Fort Collins, CO.

Shelley, P., and G. Kirkpatric. 1975. An assessment of automatic sewer flow samplers. U.S. Environ. Prot. Agency 600/2-75-065, pp. 327-330.

United States Environmental Protection Agency. 1983. Methods for chemical analysis of water and wastes. USEPA 600/4-79-020, Environmental Monitoring and Support Lab, Cincinnati, OH.

United States Geological Survey. 1977. National handbook of recommended methods for water-data acquisition. Chapter 5: Chemical and physical quality of water and sediment. Off. Water Data Coord., USGS, Reston, VA.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 8 Sampling Location

(450–VI–NWQH, September 2003)

# Chapter 8 Sampling Location

Contents:



Figures



Examples

# Chapter 8 Sampling Location

## 614.0800 Introduction

The question of where to sample is critical to a successful monitoring program. The factors that influence the location of sampling stations are:

- The study objectives and experimental design
- The type of water body (e.g., lake, stream, ground water)

Sampling locations may be viewed from two perspectives: macroscopic and microscopic. First, the overall watershed spatial locations must be determined. Second, the sampling locations within the system must be found. Because there are trade-offs between the number of sampling stations and the number of samples taken, some optimal sampling location strategies are based on travel distances and other such factors. Finally, when actually siting a station on the ground, some site selection criteria should be considered.

## 614.0801 Factors affecting locations

Definition of the study's objectives and the study design should aid in defining the general spatial sampling locations. As described in chapter 3, the monitoring study design indicates the basic sample locations. It is fairly obvious that needs differ in siting locations for plot studies versus a paired watershed design. Above-and-below or *nested* stations are particularly difficult to site. If these stations are too far apart, there may be no relationship between them. If they are located too close together, there may not be a detectable difference because of the treatment, especially in larger watersheds. Nested watersheds located too high in the watershed may exhibit poor relationships because the upper location may be intermittent. Above-and-below stations located lower in the watershed might be dominated by watershed processes not associated with the watershed treatment.

The most crucial element of sampling locations is siting the control station location. The control site must be stable and free from outside disturbances. For example, road ditch changes or repair must not be allowed to divert runoff into a control watershed. In biological monitoring this is termed the reference station.

The overall monitoring purpose, as described in the preface, influences sampling locations. For example, determining critical areas may require several watershed locations to isolate the major contributing sites. In contrast, long-term trend analysis or program evaluation may involve only one or two locations. Compliance monitoring would be located very close to the source. In contrast, fate and transport studies and wasteload allocations require downstream locations.

The type of waterbody also influences the sampling locations. To characterize a watershed outlet only requires one sampling station. To characterize ground water or the water quality of a lake would require several more sampling locations. Biological monitoring in any of these systems would require subsampling of different habitats or niches in the system.

Some specific recommendations have been made for locating sampling stations for biological monitoring (Klemm, et al. 1990)

- Select sampling locations with similar substrates, depth, physical characteristics, and velocity. If it is not possible to locate stations with similar habitats, artificial substrate samplers may be necessary.
- Include at least one reference station away from all possible discharge points.
- Include a station directly below the source of pollution. If the discharge is not mixed, include left-bank, midchannel, and right-bank substrations.
- Establish stations at various distances downstream from the source.
- Sampling locations for macroinvertebrates should be close to sites used for chemical and physical monitoring.
- Locations used for sampling should not be atypical, such as at bridges or dams. However, in urban areas such structures may be typical.
- Sampling nonpoint sources of pollution may require a number of stations along the water body impacted.

## 614.0802 Site selection criteria

The criteria used to determine sampling locations will be specific to the individual project, and will obviously change with the type of system (lake, ground water, stream) and the scale of system (plot, field, watershed) being monitored. However, the following generalized criteria can serve as a beginning point.

All sites

- Accessible all weather
- Power available
- Cooperative landowner
- Equipment protected from vandals
- Close to problem area

Streams

- Appropriate habitat
- Impermeable streambed
- Stable streambed
- Sufficient stream gradient
- Straight, uniform cross-section and approach
- Not at obstructions
- Not at meander
- Control at all stages
- Confined channel
- No road drainage influence
- Obtain stage-discharge at all stages
- Appropriate land use

Ground water

- Water table divide definable
- Barrier locations (stream, strata) known
- Direction of flow appropriate
- Water levels high or low as needed
- Stratified or mixed concentrations as needed
- Depth to confining layer known
- Away from large volume well drawdown

Lakes

- Stratification depths known
- Longitudinal gradient defined
- Bays and beaches considered
- Water circulation patterns known

Field/Plot

- Homogeneous land use
- Definable watershed
- Homogeneous soil

## 614.0803 Within system locations

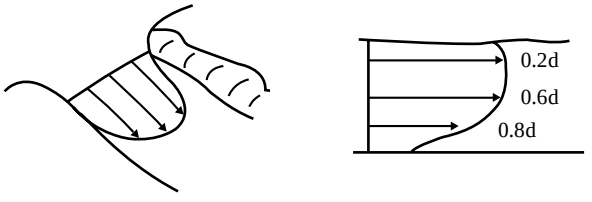
Once the overall sampling location is determined, a more specific location is needed to collect a representative sample (Canter 1985; Ponce 1980; Sanders, et al. 1983). These locations vary with system type.

### (a) Streams

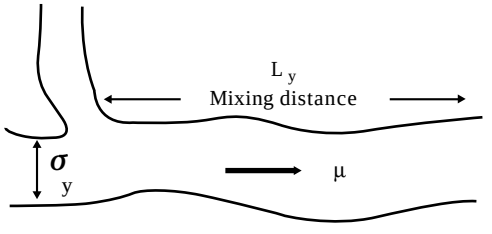
At a single stream cross section, water quality may vary vertically and horizontally for several reasons. Velocity profiles result in varying concentrations at a cross-section, especially for sediment and sediment-bound concentrations (fig. 8–1a). The stream velocity generally is greater in the center of the stream and just below the water surface. The mean velocity is considered to be at 0.6 times the depth from the water surface for water less than 1 foot deep and at the average of 0.2 and 0.8 times the depth for water more than 1 foot deep.

**Figure 8–1** Within stream sampling locations for physical/chemical monitoring

**a Velocity profiles**




**b Mixing zone**




Lateral mixing below tributary junctions may be incomplete, resulting in a plume following one streambank (fig. 8–1b). Meanders result in increased velocity near the outside bank and reduced velocity inside the meander near the point bar. Thus at a meander, lateral homogeneity would be small. The location of meanders also changes with flow stage.

Sampling locations must account for these vertical, horizontal, and longitudinal differences in water quality. Vertical and horizontal concentration differences are minimized where the stream is completely mixed; therefore, chemical sampling should be conducted at locations expected to be well mixed. Mixing is better in high velocity, turbulent stream sections and well below tributary inputs.

Mixing distances can be determined using equation 8–1 (Sanders, et al. 1983):

$$L_y = 2.17 \frac{\sigma_y^2}{d} \times \frac{\mu}{\mu^*} \qquad [8–1]$$

where:

$L_y$ = distance for complete lateral mixing
$\sigma_y$ = distance from farthest bank of stream to point of discharge
$d$ = depth of flow
$\mu$ = mean stream velocity
$\mu^*$ = shear velocity = $(gRS_e)^{0.5}$

where:

g = acceleration because of gravity
R = hydraulic radius = A/P
A = cross-section area
P = wetted perimeter
$S_e$ = slope of the energy gradient = approximately the streambed slope

The sampling station should be located downstream of a tributary, or other discharge to the stream, by a distance equal to or greater than the mixing distance.

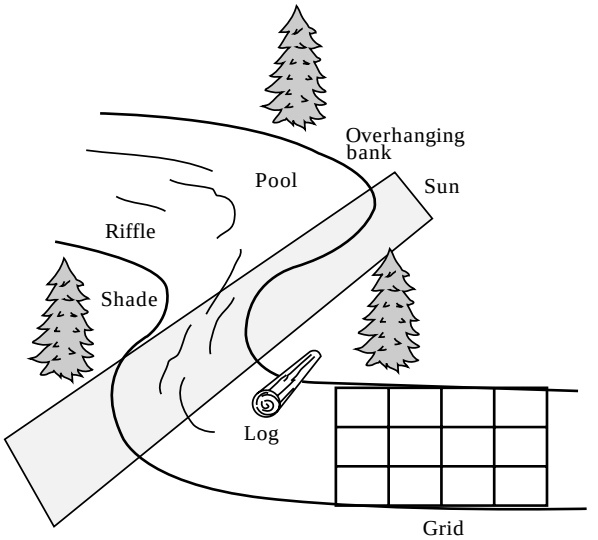
If differences in lateral concentrations still exist, compositing samples taken at locations across the stream can integrate these differences. Lateral locations can be width or flow integrated.

Differences in vertical gradients in streams also can be accounted for by the sampling technique. As described in chapter 10, a depth-integrating sampler, such as a

DH-48, can be used to obtain a grab sample. For automatic samplers, a floating sampling tube can be used.

Biological sampling within streams must consider the different stream habitats that occur as well as the mixing phenomena described. Stream systems contain pools, riffles, overhanging banks, logs, and debris that will all influence the biotic community (fig. 8–2). Within each of these habitats, stream velocity will further stratify biological communities. Shaded and sunny habitats will also differ. A good sampling program considers all of these habitats. For qualitative sampling, the biologist would make sure that each habitat was investigated. For quantitative sampling, a representative sample per unit area must be obtained from each habitat.

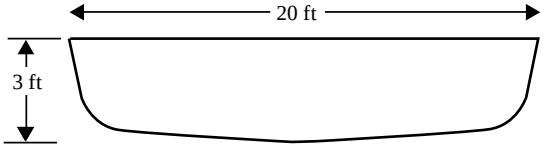
**Figure 8–2** Within stream sampling locations for biological monitoring




**Example 8–1** Mixing distances

A tributary to Mill River contains a large amount of sediment as compared to Mill River, which results in a sediment plume following one of the streambanks. How far downstream should a sampling station be located on Mill River to ensure complete mixing?

Mill River has a mean velocity ($\mu$) of 1.5 feet per second. The average depth ($d$) of the stream is 3 feet, and the average width ($\sigma_y$) is 20 feet. The streambed slope ($S_e$) is 0.005 foot per foot, based on information from a topographic map.




$$R = \frac{A}{P} = \frac{3\text{ ft} \times 20\text{ ft}}{3\text{ ft} + 3\text{ ft} + 20\text{ ft}} = 2.31\text{ ft}$$

$$s^* = \sqrt{\left(32.2\text{ ft/s}^2\right)\left(2.31\text{ ft}\right)\left(0.005\right)}$$

$$\mu^* = 0.61\text{ ft/s}$$

$$L_y = \frac{\left(2.17\right)\left(20\text{ ft}\right)^2}{3\text{ ft}} \frac{1.5\text{ ft/s}}{0.61\text{ ft/s}}$$

$$L_y = 711\text{ ft} = 0.13\text{ mi}$$

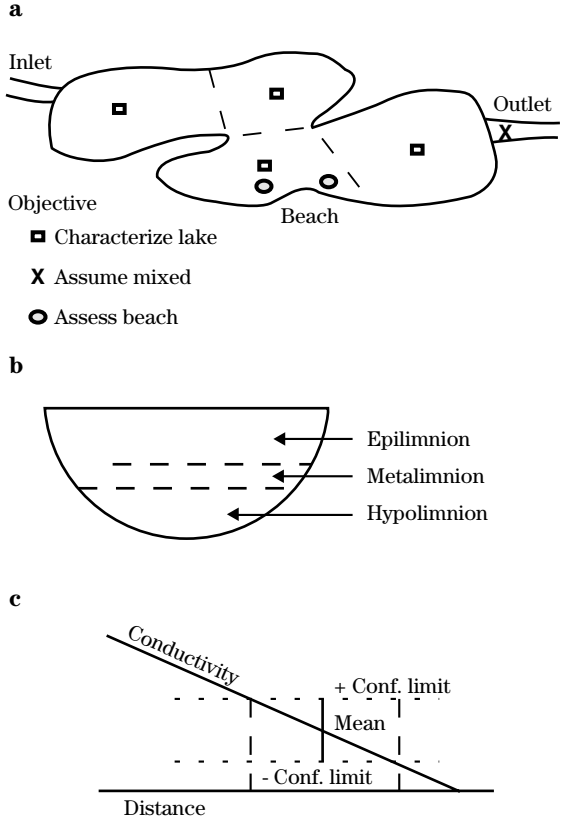
The monitoring station should be located at least 0.13 mile downstream from the tributary. This analysis assumes that the flow of the tributary is small in relation to the flow in Mill River.

## (b) Lakes

The water quality of lake systems also is heterogeneous because of vertical stratification, longitudinal gradients, and currents caused by winds and density differences. Furthermore, many lake basins are actually a combination of sub-basins or bays that have varying water quality. Near-shore water quality might be expected to be different from open water concentrations. Also, biotic populations in lakes are impacted by sediment types and some species are colonial.

Spatial variation within a lake is often greater when the lake has many bays or coves. In such cases samples may need to be located within each bay or section of the lake (fig. 8–3a). The objective of the study becomes very important in selecting lake sampling locations. Is it necessary to sample within the lake or is the outlet sufficient to fulfill the objectives?

**Figure 8–3** Lake sampling locations




Because of temperature, and therefore density differences, lakes may stratify into three layers: epilimnion, metalimnion, and hypolimnion (fig. 8–3b). Samples are needed from each stratified layer in the system to describe lake water quality at a particular point. Ideally, stratified random sampling should be used to determine the number of samples to collect in each layer (see ch. 9).

If information regarding individual layers is not needed, individual samples could be composited. An alternative approach is to collect a depth integrated sample using a hose or other similar device.

Longitudinal gradients may exist in some lakes, particularly riverine lakes or lakes that are long and narrow. If the objective includes defining the water quality gradient, the station location can be determined based on the variability at a station (Potash and Henson 1978). The procedure is to develop a linear regression with the variable being a function of the distance longitudinally through the lake (fig. 8–3c). Using the mean value and the 95 percent confidence limits, the distance either side of the station location is calculated from:

$$\pm \text{Distance} = \frac{\left[\left(\overline{X} \pm S_x t\right) - a\right]}{b} \qquad [8\text{–}2]$$

where:

$a$ and $b$ = the regression intercept and slope, respectively
$\overline{X}$ = the mean
$S_x$ = the standard deviation
$t$ = student's 't' at p = 0.05

Graphically, this represents the intercept of the upper and lower confidence limits with the regression line (fig. 8–3c). These intercepts could then be projected to the x-axis to determine the distances represented by the station. Stations with overlapping distances could be eliminated. Obviously, more stations will be needed in regions of greater concentration changes than in areas that have little gradient.

Biological monitoring in lakes must consider the spatial variability of biotic community of interest. Plankton will stratify within lakes. Blue-green algae may be more prevalent in surface water than in deeper water. Some zooplankton migrate diurnally from

shallow to deeper water. Fish seek layers of certain temperatures and dissolved oxygen concentrations. Horizontally, shallow, near-shore water contains different habitats than those of deeper water. Benthic organisms vary with lake sediment type. Certain species are colonial, growing in lake bottom villages.

Choice of biotic sampling locations must consider these variations. For plankton sampling, individual samples can be taken at different depths, or less accurately, a net can be towed vertically from a depth of no light to the surface. For quantitative benthic sampling, some estimate of spatial variability should be used to determine the number of samples needed. The same is true for macrophyte sampling.

**Example 8–2** Specific conductance of gradient

Conductivity data from Station #7 at Crown Point in Lake Champlain was used to determine the distance along Lake Champlain that the station represents (Potash & Henson 1978). The mean distance at the station was 112 miles. The value of $S_x t$ was 6.55.

The regression:

*Conductivity* = 110.3 – 0.13 (distance)

where distance is given in miles.

The confidence limits:

$$+\text{Distance} = \frac{(112 + 6.55) - 110.3}{-0.13}$$

$$+\text{Distance} = 63.5 \text{ mi}$$

$$-\text{Distance} = \frac{(112 - 6.55) - 110.3}{-0.13}$$

$$-\text{Distance} = 37.3 \text{ mi}$$

Station #7 would adequately describe the conductivity concentration gradient 63.5 miles in one direction and 37.3 miles in the other direction. Adjacent stations could be evaluated to determine if there is overlap with station #7. If there was, a station could be dropped while the gradient would still be adequately monitored.

## (c) Ground water

The location of sampling stations within ground water systems depends upon the objectives as well as the type of aquifer system being monitored. The objectives determine whether just the ground water concentrations or both concentration and flow for mass calculations are needed. For flow analysis, the well locations need to be expanded to determine the flow into and out of the area and the hydrogeologic properties of the aquifer. Several textbooks cover this subject (Davis & DeWiest 1970; Driscoll 1986; Domenico & Schwartz 1990; Freeze & Cherry 1979).

For concentration monitoring alone, the monitoring system is simplified as compared to flow monitoring. In siting ground water monitoring wells, the soils and geology, the direction of ground water flow, and the type of ground water system must be considered.

The two major types of aquifers are confined and unconfined (Davis & DeWiest 1970). Unconfined aquifers, also termed water table aquifers, are in direct contact with the atmosphere through the soil. Confined aquifers, also termed artesian aquifers, are separated from the atmosphere by an impermeable layer (fig. 8–4a).

Ground water monitoring also must consider vertical, horizontal, and longitudinal water quality differences. More commonly, ground water monitoring requires a two-staged approach. The first stage should be a hydrogeologic survey that determines the ground water surface elevations and flow directions. In some ground water investigations it may be important to locate the top of the ground watershed divide.

To investigate lateral ground water quality, sampling wells should be located upgradient and downgradient from the area of interest (fig. 8–4b). More than one well should be located above, within, and below the treatment area so that replications can be obtained. The actual number of wells needed to characterize the water quality of the aquifer can be determined from the formula in chapter 9. Before monitoring wells are sited, there must be knowledge of the general ground water flow direction. Preliminary estimates of flow direction can be obtained by triangulation using three driven well points.

The depth of the monitoring well also is important. If sampling nitrate in unconfined aquifers, it may be necessary to utilize multilevel wells because nitrate concentrations are often stratified with higher levels at the top of the aquifer (Eccles and Nicklen 1978). Such wells can be constructed in the same bore hole (fig. 8–4c) or in separate borings. Poor sealing between screens in the same borehold may make "nested" wells undesirable. For monitoring water table wells, the length of perforated screen should cover the full range of water levels anticipated.

It is important when locating the depth of all wells that the monitoring well be placed into the ground water of interest and not into a localized perched condition (fig. 8–4d).

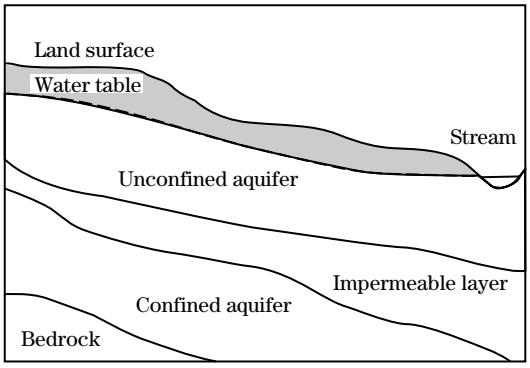
Using existing wells for monitoring presents several problems. Usually knowledge is lacking regarding well construction, screen length, and other such information. Also, the well could be contaminated. New monitoring wells, developed for the purpose of monitoring, are encouraged over existing wells.

Several geophysical techniques are available to characterize ground water conditions. Both surface and borehole techniques can be used. Surface techniques include (Driscoll 1986):

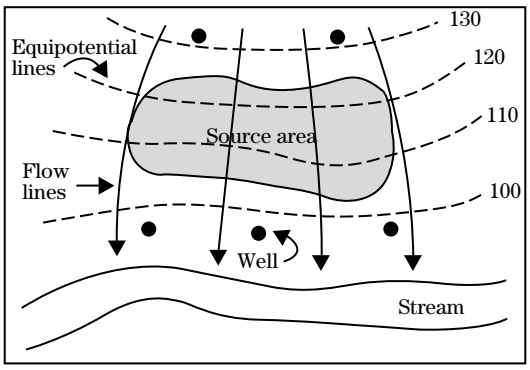
- seismic refraction/reflection
- gravimetric surveys
- electromagnetic surveys
- electrical resistivity

**Figure 8–4** Ground water sampling locations

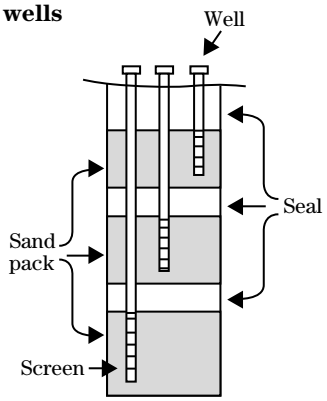
**a Ground water aquifers**



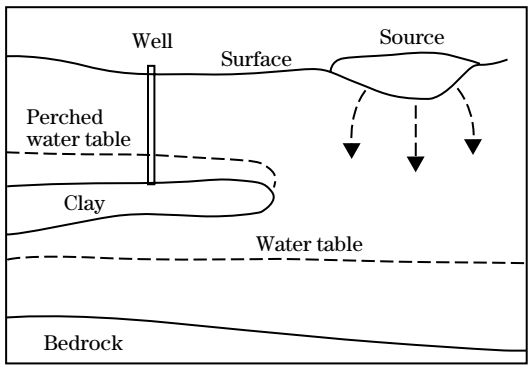

**b Monitoring source areas**




**c Multilevel wells**




**d Vertical locations**



All of these methods provide information on the geologic stratigraphy and presence of ground water. Seismic methods can be used to determine the depth to different geologic formations using a hammer and geophones. Gravity meters can be used to measure density differences in subsurface materials and are especially useful in locating bedrock.

Ground-penetrating radar is useful for shallow (<50 feet) investigations of subsurface materials. The device can be towed to obtain profiles of depths and distances. Resistivity is used to identify the depth to or thickness of subsurface strata. The depth to the water table can also be determined. Additional methods can be used in boreholes.

## 614.0804 Optimizing locations

Large monitoring programs generally include many sampling locations and many visits per location. The optimal number of stations and the number of visits per stations can be determined so that the variability about the mean is minimized. This has been described as a combination of a cost function and a statement of variability in the data (Hayne 1977; Mar, et al. 1986; Reckhow & Chapra 1983). A cost function could be:

$$C = C_o + SC_s + SpvC_v \qquad [8–3]$$

where:

$C$ = total cost of sampling
= total budget
$C_o$ = initial fixed cost
$C_s$ = cost of establishing site
$C_v$ = cost of visiting site
$S$ = number of sites
$pv$ = number of visits per site
= number of periods ($p$) times number of visits ($v$) per period

The number of visits ($v$) per site is a function of the variance caused by the number of sites, the number of visits, an interaction between site and visit, and an error term, such that:

$$v = \left( \frac{CK_v + C_s}{pC_v \left( pK_s + K_{S \cdot V} \right)} \right)^{\frac{1}{2}} \qquad [8–4]$$

where:

$$K_s = \frac{\sigma_s^2}{\sigma_e^2} \qquad [8–5]$$

$$K_v = \frac{\sigma_v^2}{\sigma_e^2} \qquad [8–6]$$

$$K_{S \cdot V} = \frac{\sigma_{S \cdot V}^2}{\sigma_e^2} \qquad [8–7]$$

where σ refers to the variance caused by the differences among sites (s), visits (v), a site by period interaction (s·v), and random error (e).

The number of sites can be determined based upon the optimum number of visits from:

$$S = \frac{C}{C_s + pvC_v} \qquad [8–8]$$

**Example 8–3** Optimizing sites and visits

A study was conducted by Hayne (1977) to determine the total number of small drainage basins that would describe the water quality in a river basin. Sampling sites were chosen randomly, and grab samples were collected and analyzed for total phosphorus.

A preliminary 1-year study using 13 4-week periods, 15 sites, and 2 randomly selected visits per period resulted in the following information:

Total cost = \$14,211.25
Per site cost = \$153.18
Per visit cost = \$79.47
Site variance = 0.01265
Visit variance = 0.06830
s·v variance = 0.04109
Error variance = 0.1153

Determine the optimum number of visits per site and the number of sites needed given the available budget. If the budget were doubled what would be the allocation between sites and visits?

$$K_s = \frac{0.01265}{0.1153} = 0.1097$$

$$K_v = \frac{0.06830}{0.1153} = 0.5924$$

$$K_{S\cdot V} = \frac{0.04109}{0.1153} = 0.3564$$

$$v = \left[\frac{14{,}211.25(0.5924) + 153.18}{13(79.47)\left[13(0.1097) + 0.3564\right]}\right]^{\frac{1}{2}}$$

$$v = 2.16 = 3$$

$$S = \frac{14{,}211.25}{153.18 + 13(3)(79.47)} = 4.4 = 5$$

For the budget of \$14,211.25, the optimal number of sites would be 5 and the number of visits per period would be 3 rather than the 2 used in the preliminary study.

If the budget were doubled, the number of sites could be increased to 9 and the number of visits per period would remain 3.

## 614.0805 References

Canter, L.W. 1985. River water quality monitoring. Lewis Publ., Inc., Chelsea, MI.

Davis, S.N., and R.J.M. DeWiest. 1970. Hydrogeology. John Wiley and Sons, Inc., NY.

Domenico, P.A., and F.W. Schwartz. 1990. Physical and chemical hydrogeology. John Wiley and Sons, Inc., NY.

Driscoll, F.G. 1986. Ground water and wells. 2nd Ed. Johnson Div., St. Paul, MN.

Eccles, L.A., and R.R. Nicklen, 1978. Factors influencing the design of a ground water quality monitoring network. Proc. Establishing water quality monitoring programs. Amer. Water Resour. Assoc.

Freeze, R.A., and J.A. Cherry. 1979. Ground water. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Hayne, D.W. 1977. Probability sampling of small streams: problems and results. *In* Watershed research in Eastern North America, A workshop to compare results, Vol. II, Rep. NSF/RA 770255. Smithsonian Inst., pp. 809-830.

Klemm, D.J., P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. Macroinvertabrate field and laboratory methods for evaluating the biological integrity of surface waters. Office of Research and Development, Washington, DC, EPA/600/4-90/030.

Mar, B.W., R.R. Horner, J.S. Richey, R.N. Palmer, and D.P. Lettenmaier. 1986. Data acquisition. Environ. Sci. Technol. 20(6):545-551.

Ponce, S.L. 1980. Water quality monitoring programs. WSDG Tech. Pap. No. WSDG-TP-00002. Watershed Sys. Dev. Group, USDA, Forest Serv., Ft. Collins, CO.

Potash M., and E.B. Henson. 1978. Significance of geographic positioning for water quality monitoring in Lake Champlain. Proc. Establishing water quality monitoring programs. Amer. Water Resour. Assoc., pp. 58-73.

Reckhow, K.N., and S.C. Chapra. 1983. Engineering approaches for lake management. Vol. 1: Data analysis and empirical modeling. Butterworth Publ., Woburn, MA.

Sanders, T.G., R.C. Ward, J.C. Loftis, T.D. Steele, D.D. Adrian, and V. Yevjevich. 1983. Design of networks for monitoring water quality. Water Resour. Pub., Littleton, CO.

United States Department of Agriculture, Soil Conservation Service. 1985. National Engineering Handbook, Section 4, Hydrology., chapter 14.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 9 Sampling Frequency and Duration

(450–VI–NWQH, September 2003)

# Chapter 9 Sampling Frequency and Duration

Contents:

# Chapter 9 Sampling Frequency and Duration

## 614.0900 Introduction

The most frequently asked questions when developing a water quality monitoring study are "How many samples and for how long?" Unfortunately, the correct response is: "It depends." Several factors affect the frequency of sampling. They include the objectives of the study, the type of waterbody being studied, the data variability, and the available resources. Table 9–1 summarizes general frequencies for various objectives for conducting a water quality study. Frequencies are given in relative terms to each other because a fixed time interval is inappropriate.

Long-term trend monitoring and programs evaluating program effectiveness on a watershed basis can use longer intervals between samples than other monitoring objectives. Frequent sampling or even a continuous recorder may be desirable for a study aimed at understanding a mechanism controlling certain water quality changes. The frequency of compliance monitoring should be approximately equal to the probability of exceeding a standard.

Sampling frequency is also affected by the aquatic system being studied. In general, the variance is greater; therefore, more samples are needed for studying streams than for lakes. Intermittent streams are often more variable than permanent streams. Ground water also is considered less variable than streams, but soil water samples can be highly variable (fig. 9–1).

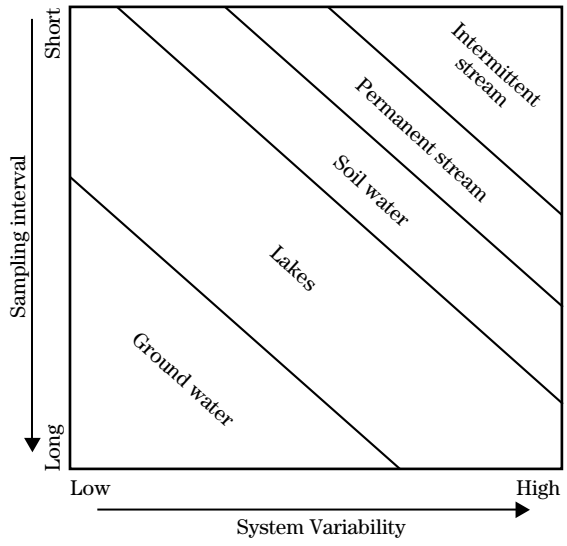
Financial resources typically limit the sampling frequency, although time, people, and laboratory capability can also limit sampling frequency. However, financial resources should not be allowed to dictate a sampling frequency. In cases where funds are limiting, a consideration should be given to eliminating extra parameters or stations. Compositing samples and passive sampling (chapter 10) can save substantial resources.

This chapter presents methods for calculating the sampling frequency. The primary sampling techniques described are simple random sampling and stratified random sampling.

**Table 9–1** Relative sampling frequency and objectives

| Objective samples | Relative interval between |
|---|---|
| 1. Baseline | Long |
| 2. Trends | Long |
| 3. Fate and transport | Short |
| 4. Problem definition | Short |
| 5. Critical areas | Short |
| 6. Compliance | Probability of exceeding standard |
| 7. BMP effectiveness | Short |
| 8. Program effectiveness | Long |
| 9. Wasteload allocations | Short |
| 10. Model evaluation | Short to long |
| 11. Research | Continuous to short |

**Figure 9–1** Sampling interval as a function of system type



# 614.0901 Simple random sampling

Sampling of water quality is needed to provide useful information about the entire population of water quality data that exists without measuring the entire population. Sampling saves time and money. Simple random sampling for water quality monitoring means that every water quality sample has an equal chance of being collected.

The calculation of sample size varies with the statistical objective of the monitoring study. Such objectives include an estimate of the mean, linear trend detection, and a step trend. The methods used to calculate sample sizes for each case are presented.

## (a) Estimate of the mean

One goal may be to be able to estimate the mean for a water quality variable with a certain amount of confidence in the estimate. The equation for calculating the sample size has been widely reported and is based on the variability and precision desired (Snedecor & Cochran 1980; Freese 1962; Moser & Huibregtse 1976; Ponce 1980; Rustagi 1983; Reckhow & Chapra 1983; Sanders, et al. 1983). The sample size can be calculated from the relationship:

$$n = \frac{t^2 S^2}{d^2} \qquad [9–1]$$

where:
$n$ = the calculated sample size
$t$ = Student's 't' (appendix B) at n-1 degrees of freedom and confidence level (p)
$S$ = the estimate of the population standard deviation
$d$ = the allowable difference from the mean

The standard deviation ($S$) is calculated as the square root of the variance ($S^2$) which is determined from (Snedecor & Cochran 1980):

$$S^2 = \frac{\Sigma X_i^2 - \frac{\left(\Sigma X_i\right)^2}{n}}{n-1} \qquad [9–2]$$

where:
$n$ = the sample size
$X_i$ = the value of the i$^{th}$ observation

If the coefficient of variation rather than the standard deviation is known, the following relationship may be used (Koch, et al. 1982; Moser & Huibregtse 1976):

$$n = \frac{t^2 CV^2}{\%\overline{X}^2} \qquad [9–3]$$

where:

$CV$ = the coefficient of variation = $\frac{S}{\overline{X}}$

$\%\overline{X}$ = the percent deviation allowed from the true mean

Ranges in coefficients of variation for select system type are given in table 9–2 for certain water quality variables. This formula should be used with a double iterative procedure as shown in the following examples.

If the variance ($S^2$) is not known, an approximation can be made based on the range in the data using equation 9–4 (Ponce 1980; Sanders, et al. 1983):

$$S^2 = \frac{\left(\text{Range}\right)^2}{4^2} \qquad [9–4]$$

where:
Range = the range from the smallest to the largest values expected to be encountered during the sampling period

**Example 9–1** Sample size using simple random sampling based on estimate of the mean

Based on historical monitoring in a stream, how many samples are needed to be within 10 and 20 percent of the true annual mean total phosphorus concentration? The following information was obtained from the existing monitoring program for 1 year:

| | |
|---|---|
| mean | = 0.886 mg/L |
| standard deviation | = 0.773 mg/L |
| variance | = 0.597 mg/L |
| maximum | = 4.1 mg/L |
| minimum | = 0.074 mg/L |
| n | = 165 |

The difference (d) for 10 percent and 20 percent would be:

d = 0.1 x 0.886 mg/L = 0.09 mg/L
d = 0.2 x 0.886 mg/L = 0.18 mg/L

The t-value would be 1.96 for >120 degrees of freedom and p=0.05 (appendix B). A two-tailed t-value can be obtained from most statistics books, such as table A–4 in Snedecor and Cochran (1980).

1st iteration—10%

$$n = \frac{(1.96)^2 (0.773)^2}{(0.09)^2} = 283$$

Because the t-value would not change for n=283 degrees of freedom, no additional iterations are necessary.

1st iteration—20%

$$n = \frac{(1.96)^2 (0.773)^2}{(0.18)^2} = 71$$

This result is a fourth of the 10% result. However, the t-value must be adjusted for the degrees of freedom.

2nd iteration—20%

$$n = \frac{(1.993)^2 (0.773)^2}{(0.18)^2} = 73$$

Therefore 73 samples should be taken to estimate the mean annual total phosphorus concentration within 20% of the true mean.

The variance could have been estimated based on the range as follows:

$$S^2 = \frac{\text{Range}^2}{16} = \frac{(4.1 - 0.074)^2}{16} = 1.013 \text{ mg/L}$$

This estimate of the variance is greater than the measured variance listed above, and would result in a larger sample size being taken.

## (b) Linear trend detection

Another goal may be to determine the number of samples needed to detect a linear trend in the water quality data (Ward, et al. 1990). The sample size may be calculated from:

$$n = \frac{12t^2S^2}{d^2} \qquad [9–5]$$

where:

$S$ = the standard deviation of the water quality data collected over time with any trend removed from the data

$d$ = the minimum magnitude of the trend

**Example 9–2** Sample size for trend detection

Using example 9–1, determine the number of samples needed to detect a trend of at least 0.5 mg/L per year.

1st iteration

$$n = \frac{12(1.96)^2(0.773)^2}{(0.5)^2} = 110$$

2nd iteration

$$n = \frac{12(1.981)^2(0.773)^2}{(0.5)^2} = 113$$

Therefore, 113 samples per year would be needed to detect a linear trend of 0.5 mg/L per year. The greater the trend, the fewer samples that would be needed.

**Table 9–2** Coefficients of variation[1] (dashes indicate data not available)

| Parameter | Agricultural streams | Lakes | Ground water | Treatment plant | Edge-of-field |
|---|---|---|---|---|---|
| Temperature | 0.7-1.2 | 0.4-0.7 | | 0.4-0.7 | |
| Dissolved oxygen | 0.2-0.6 | 0.2-0.4 | | 0.2-0.7 | |
| pH | 0.03-0.1 | 0.05-0.1 | | 0.03-0.1 | |
| Conductivity | 0.2-0.7 | 0.1-0.5 | | 0.2-1.3 | |
| Secchi disk | --- | 0.1-0.7 | | --- | |
| Fecal coliform | 0.9-27.1 | 1.6-9.5 | | 0.6-39.2 | |
| Fecal streptococci | 1.2-94.0 | 1.5-32.0 | | 0.9-11.2 | |
| Turbidity | 0.7-5.5 | 0.6-2.5 | | 0.4-3.8 | |
| Total suspended solids | 1.0-9.0 | 0.1-3.7 | | 0.3-3.4 | |
| Volatile suspended solids | 0.7-4.4 | 0.5-2.8 | | 0.3-2.2 | |
| Total phosphorus | 0.6-2.2 | 0.3-2.4 | | 0.3-0.9 | |
| Ortho phosphorus | 0.5-2.1 | 0.4-3.3 | | 0.5-1.4 | |
| Total Kjeldahl nitrogen | 0.4-1.8 | 0.1-1.4 | | 0.3-1.1 | |
| Ammonia nitrogen | 0.8-4.0 | 0.3-3.9 | | 0.4-2.2 | |
| Nitrate nitrogen | 0.1-4.8 | 0.7-2.0 | | 0.4-4.4 | |
| Chlorophyll 'a' | --- | 0.2-4.0 | | --- | |

[1] St. Albans Bay RCWP

### (c) Step trend

The goal may be to determine if there has been a change in the mean water quality between two time periods. This would be equivalent to a step trend (Sanders, et al. 1983). The number of samples needed to detect a stated change is determined from:

$$n = \frac{2t^2S^2}{d^2} \qquad [9–6]$$

where:
- $n$ = the size of each sampling period, which is assumed to be equal
- $S$ = the pooled standard deviation for both periods
- $d$ = the allowable difference (precision) from the mean

The total number of samples needed to detect the difference would be 2n.

**Example 9–3** Sample size for step trend

For example 9–2, determine the number of samples needed to detect a change in the mean total phosphorus concentrations between a pre-implementation period and a post-implementation period with 20 percent precision. No changes in the original sampling data were assumed.

$$\mathrm{d} = 0.2 \times 0.886\ \mathrm{mg/L} = 0.18\ \mathrm{mg/L}$$

$$\mathrm{n} = \frac{2(1.96)^2(0.773)^2}{(0.18)^2} = 141$$

$$2\mathrm{n} = 282$$

Therefore, 282 samples would need to be taken over the two time periods to detect a difference in the means between the two periods. Note that the level of precision would only be 20 percent; therefore, the difference would need to be greater than 20 percent to be detectable.

## 614.0902 Stratified random sampling

Instead of each water quality sample having the equal chance of being collected, there may be advantages to dividing the population of water quality samples into subgroups that are each more homogeneous than the whole data set. Samples could then be taken from each subgroup or strata. This type of sampling is termed *stratified random sampling* (Snedecor & Cochran 1980). More samples are allocated to subgroups that have greater variability. Two examples of appropriate applications of this technique would be:
- grouping by a flow period (snowmelt, summer low flow) or
- grouping by strata in a lake (epilimnion, hypolimnion).

The sample size for stratified random sampling can be calculated from the relationship (Reckhow & Chapra 1983):

$$n = \frac{t^2\left(\sum w_i S_i\right)^2}{d^2} \qquad [9–7]$$

where:
- $n$ = the total number of samples
- $t$ = Student's 't' at n–1 degrees of freedom
- $w_i$ = the proportional size of stratum $i$
- $S_i$ = the standard deviation of the water quality data for stratum $i$
- $d$ = the difference from the mean

The number of samples for each individual stratum is determined from:

$$n_i = \frac{nw_iS_i}{\sum(w_iS_i)} \qquad [9–8]$$

where:
- $n_i$ = the number of samples of stratum i

**Example 9–4** Stratified random sampling

Mudd Lake stratifies in the summer; therefore, it is desirable to subsample each layer to determine lake-wide phosphorus concentrations. Preliminary sampling resulted in the following information:

| | Thickness (ft) | Thickness (%) | Standard deviation (mg/L) |
|---|---|---|---|
| epilimnion | 14 | (35) | 0.012 |
| metalimnion | 6 | (15) | 0.005 |
| hypolimnion | 20 | (50) | 0.010 |

Determine the total number of samples and the number of samples within each stratum to be within 10 percent of the true mean at the 95 percent confidence level. The overall mean was 0.04 mg/L total phosphorus.

1st iteration

$$n = \frac{(1.96)^2\left[(0.35)(0.012)+(0.15)(0.005)+(0.50)(0.010)\right]^2}{\left[(0.10)(0.04)\right]^2}$$

$$n = 23.8 = 24$$

2nd iteration

$$n = \frac{(2.069)^2\,(0.00995)^2}{(0.004)^2}$$

$$n = 26.5 = 27$$

Allocate the 27 samples among the 3 strata by:

$$n_{epi} = \frac{27(0.35)(0.012)}{0.00995} = 11.4$$

$$n_{meta} = \frac{27(0.15)(0.005)}{0.00995} = 2.0$$

$$n_{hypo} = \frac{27(0.50)(0.010)}{0.00995} = 13.6$$

Therefore 11, 2, and 14 samples should be taken from the epilimnion, metalimnion, and hypolimnion, respectively.

# 614.0903 References

Freese, F. 1962. Elementary forest sampling. USDA Forest Serv., Agric. Handb. No. 232, Washington, DC.

Koch, R.W., T.G. Sanders, and H.J. Morel-Seytoux. 1982. Regional detection of change in water quality variables. Water Resour. Bull. 18(5):815-821.

Moser, J.H., and K.R. Huibregtse. 1976. Handbook for sampling and sample preservation of water and wastewater. U.S. Environ. Prot. Agency 600/4-76-049.

Ponce, S.L. 1980. Water quality monitoring programs. WSDG-TP-00002, Watershed Sys. Dev. Group, USDA Forest Serv., Ft. Collins, CO.

Reckhow, K.H., and S.C. Chapra. 1983. Engineering approaches for lake management. Volume 1: Data analysis and empirical modeling. Butterworth Publ., Woburn, MA.

Rustagi, K.P. 1983. Determination of sample size in simple random sampling. Forest Science. 29(1):190-192.

Sanders, T.G., R.C. Ward, J.C. Loftis, T.D. Steele, D.D. Adrian, and V. Yevjevich. 1983. Design of networks for monitoring water quality. Water Resour. Pub., Littleton, CO.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods. 7th ed. IA State Univ. Press, Ames, IA.

Ward, R.C., J.C. Loftis, and G.B. McBride. 1990. Design of Water Quality Monitoring Systems. Van Nostrand Reinhold, NY, pp. 1-13.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 10 Station Type

(450–VI–NWQH, September 2003)

# Chapter 10 Station Type

Contents:

**Figures**

**Example**

# Chapter 10 Station Type

## 614.1000 Introduction

The purpose of this section is to provide guidance on the design, operation, and maintenance of hydrologic and water quality monitoring stations. This chapter is divided into the types of monitoring to be conducted:

- discharge
- concentration
- precipitation
- soil water
- biota
- bottom sediment

Generally, several optional methods for conducting the monitoring are available for each type of monitoring station needed. Also, the costs of installation and operation of these stations differ.

When designing monitoring stations, three principles are recommended: redundancy, simplicity, and quality. Important hydrologic variables, such as stage, should be measured in more than one way. Power failures and the unexpected seem to influence any monitoring record. Whenever possible, the most simple alternative is often the best. Complicated monitoring station designs invite problems. Finally, whatever is done should be installed with high quality. A neat and sturdy monitoring setup will be a safe and reliable one.

Agricultural Handbook No. 224 (USDA 1979) is an important reference for designing monitoring stations. The U.S. Geological Survey has published a series of Techniques of Water Resource Investigations (TWI) reports that addresses many of the issues related to designing monitoring stations. A listing of TWI 1 through TWI 8 is given following the references. Other references are also listed at the end of this chapter.

The type of station desired will, of course, depend on the objectivities as well as other components of the study design. Not all study designs require a fixed station, especially biological monitoring. This chapter is intended to give guidance on possible approaches and the equipment currently available to achieve certain monitoring goals.

## 614.1001 Discharge stations

The type of discharge station to construct is a function of the scale of the project (plot, field, or watershed), the project duration, and the project budget.

### (a) Plot discharge

Two types of devices for measuring the amount of plot runoff are shown in figure 10–1. A simple, small plot design is shown in figure 10–la. Sheet metal (18 gauge) cutoff walls are driven into the soil. Overland flow from just within the plot flows into a rain gutter installed flush with the soil surface, and then into a collection jug. The lip on the rain gutter can be inserted into the soil to prevent underflow. The plot can be sized based on expected overland flow so that the volume of the jug will not be exceeded. For example, a 3 by 6 foot plot has been used in the northeast United States. This type of plot can be installed in about 20 minutes and removed during field cultivation. A tipping bucket device (Chow 1976; Johnson 1942) can be used at the bottom of the plot instead of a collection jug. In some cases a large barrel could be installed to capture all the flow. This sampler determines flow based on the volume of sample collected.

Runoff volumes from such small runoff plots are highly variable plot to plot; therefore, a large number of plots may be necessary to obtain a good estimate of runoff (see chapter 9).

An example of a runoff plot used for research purposes is shown in figure 10–lb. This type of plot used a multislot divisor. The total runoff volume is computed from the sample volume collected by the divisor (USDA 1979). Dressing, et al. (1987) describe an expensive sampler that determined flow based on the volume of sample collected.

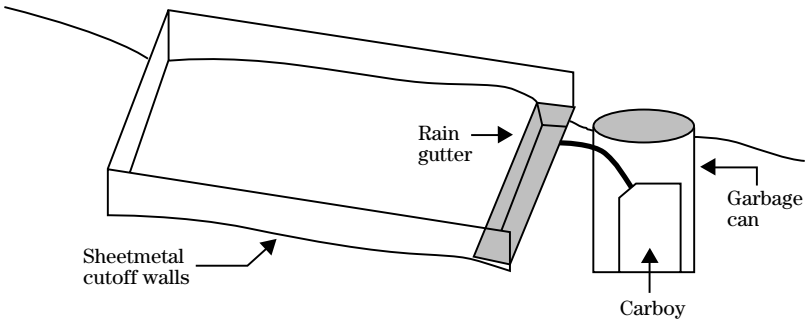
## (b) Edge-of-field discharge

Some of the devices described previously for plots can be enlarged for edge-of-field situations, especially that described by Dressing, et al. (1987). Because ponding of water on a field and high sediment and plant remains loads are undesirable, a flume, rather than a weir, is most often used for field discharge. The H-type flume is the most commonly used (fig. 10–2). This flume is so named because it was the eighth developed in a series starting with the A flume (Gwinn and Parsons 1976). The others include HS (small) and HL (large) flumes.

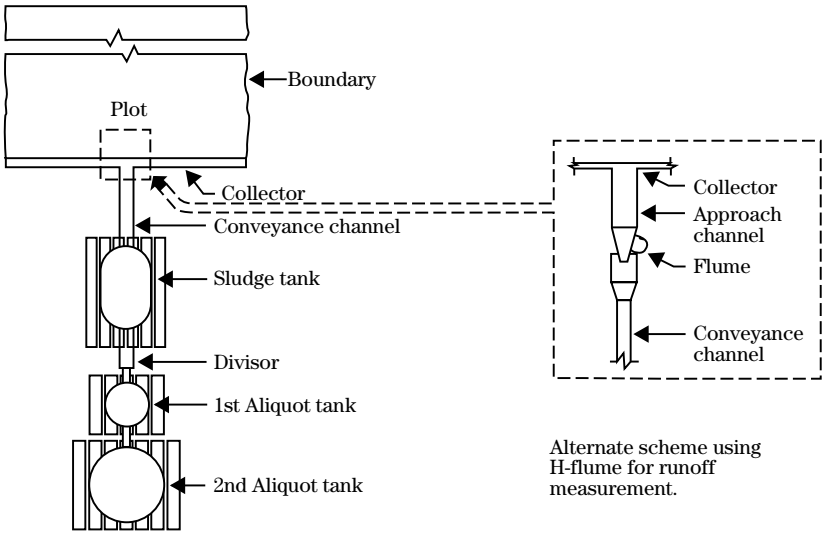
A complete description of the H-flume is given in Agricultural Handbook 224 (USDA 1979). The flume is often constructed of sheet metal; however, stainless steel flumes have been used for pesticide sampling (Smith, et al. 1985), and prefabricated fiberglass flumes are available as well. Rating tables and

**Figure 10–1** Runoff plots

**a Small-scale runoff plot**




**b Larger-scale runoff plot**

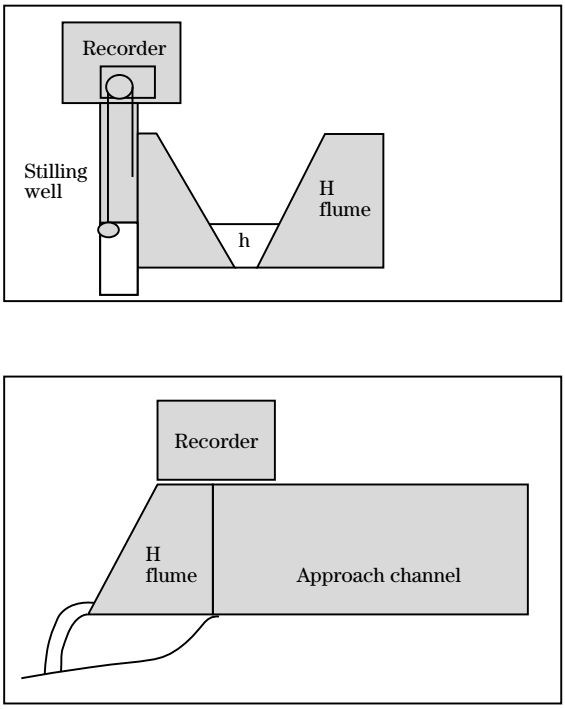



Alternate scheme using H-flume for runoff measurement.

equations are readily available (Gwinn and Parsons 1976; USDA 1979; Grant 1979). An approach channel to the flume is needed to reduce velocity and turbulence in the flow (fig. 10–2). A false side sloping floor (1:8) can be used when sedimentation in the flume is significant. The H-flume needs a method of recording stage, generally in a stilling well attached to the flume. Stage recording is described later in this chapter.

Other types of flumes have been used to measure edge-of-field runoff including Parshall and long-throated flumes (USDA 1979; Replogle & Clemmens 1981).

**Figure 10–2** Field runoff H-flumes




## (c) Stream discharge

Many options are available for determining discharge in streams. The selection of the type of station varies with individual site conditions, such as slope, sediment load, and stream size. The major options include flumes, wiers, and a natural channel. The use of existing structures, such as culverts, will also be discussed.

The practical limit to H-flumes is about a peak discharge of 100 cubic feet per second (5 ft head); however, larger flumes can be built onsite. Specialized flumes have been developed for use in the Western States where streams may be flashy and ephemeral (USDA 1979). Sufficient slope in the streambed is needed to prevent backwater into the flume and allow the freefall of water at the outlet opening.

Wiers are another common device used in streams for discharge measurement. Figure 10–3b shows several configurations for weir types. They include v-notched, rectangular, and Cipolletti wiers. Wiers can be constructed of wood, sheet metal, or concrete.

The practical size for a prefabricated weir made of plywood with a metal or plastic sharp crest is 5 cubic feet per second (1.3 ft head). Larger plywood weirs may fail. The weir must not leak. The weir plate should extend well into the streambed and be connected to a channel sill that extends upstream of the weir.

A natural channel is often necessary when flow is too large for an artificial structure. The basic features of recording discharge for a natural channel are shown in figure 10–4a. The cross-section is located at a control section; that is, a stable streambed and streambank location where the channel is straight. Also, stream gaging must be possible at or near the cross-section.

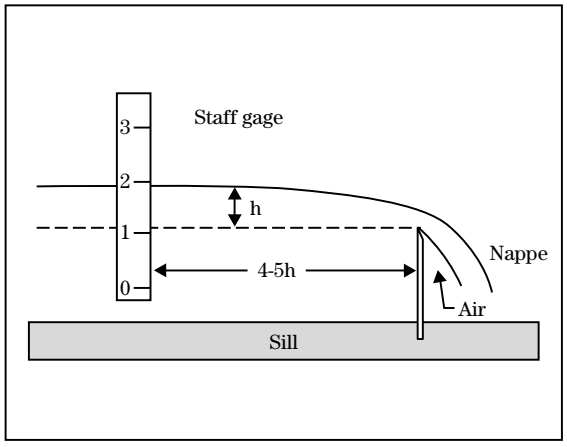
A basic setup for a natural channel includes a stilling well for stage measurement with intake pipes connected to the stream. The stilling well should not be placed in the stream because of velocity effects and icing problems, but rather should be installed in the streambank. The well diameter could range from a 12-inch PVC pipe to a 48-inch corrugated metal pipe (CMP). A gage house is either placed on top of the stilling well or, for large diameter culverts, is part of the well itself. The total cross-section area of the

intake pipes should be about 1 percent of the area of the stilling well. Venting the gage house helps to prevent moisture buildup.

For some study designs, using an open channel with point measurement of discharge may be sufficient to achieve the study objectives. However, such discharge monitoring does not give any information about the discharge between sampling dates. Existing structures, including culverts, dams, and spillways, are used for discharge measurements (USDA 1979). The author believes that culverts generally should be avoided for discharge measurements. At high flows, culverts can be submerged, a hydraulic jump may form at the culvert entrance, or the water level may drop because an entrance is constricted (fig. 10–4b). These conditions yield false stage values. Culverts also present problems by collecting debris and icing in winter.

**Figure 10–3** Weirs

**a Components of typical weir**




**b Weir types**

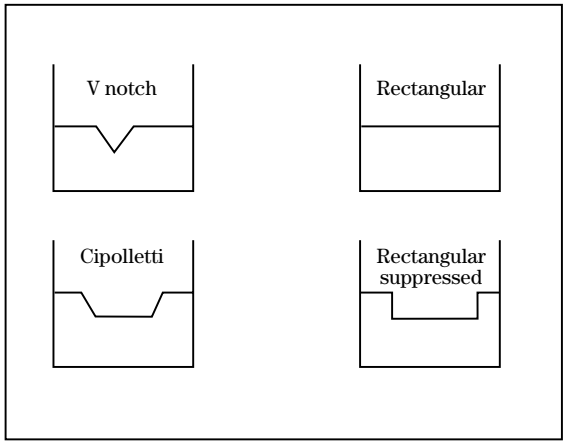
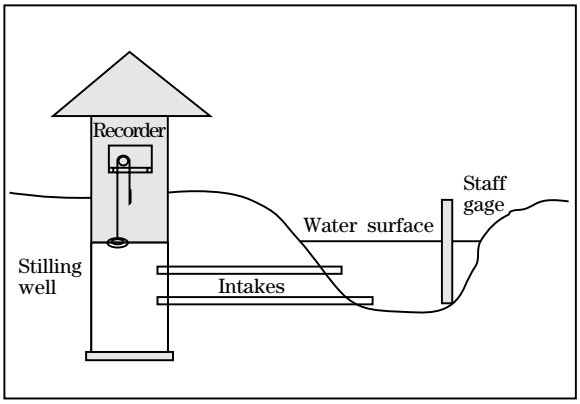
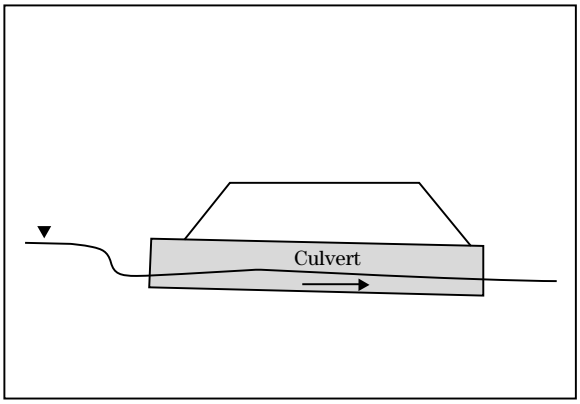



**Figure 10–4** Natural channel gaging station

**a Station cross-section**




**b Flow at culvert**



## (d) Staff gages

All discharge stations should include a staff gage. A staff gage is typically a vertical calibrated gage made of porcelain enameled steel (fig. 10–5). It should be so constructed or so placed as to not catch debris and to shift easily upward or downward. A point gage should be used in an instrument shelter, either with a separate float or using a graduated float tape. The outside staff gage reading is the true stage to which all recording gages should be set. The elevation of the staff gage should be checked periodically for shifts.

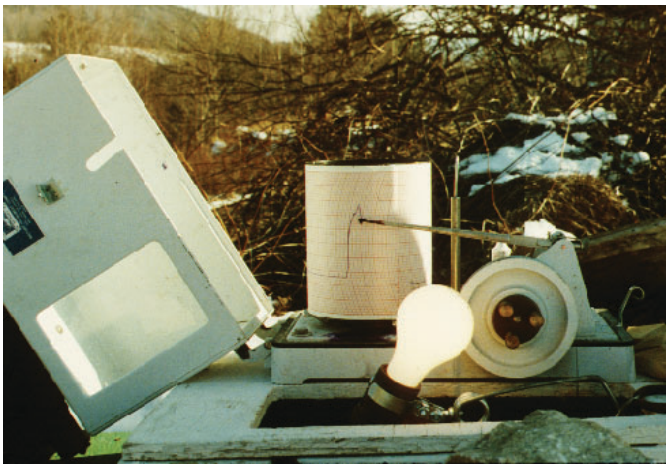
## (e) Stage recording

Stage is most often recorded in a stilling well, although bubbler gages have made this requirement unnecessary. The primary methods for recording stage are through the use of floats, bubblers, pressure transducers, and ultrasonic sensors (fig. 10–6). Several float-level recorders are highly reliable and remain the preferred method of stage recording for many hydrologists. Advantages of bubbler gages are that no stilling well is required and they can be easily combined with automatic water samplers. Almost all stage recorders available today allow for data logging. Those with programmable data loggers can control automatic water sampling. Pressure transducers and ultrasonic sensors are not widely used at this time; however, they are very useful for data logging.

**Figure 10–5** Porcelain staff gage

**Figure 10–6** Stage recorders (photos c, d, e, f courtesy Instrumentation Specialties Company)

**a Float-level**



**b Punch tape**


**c Bubbler**

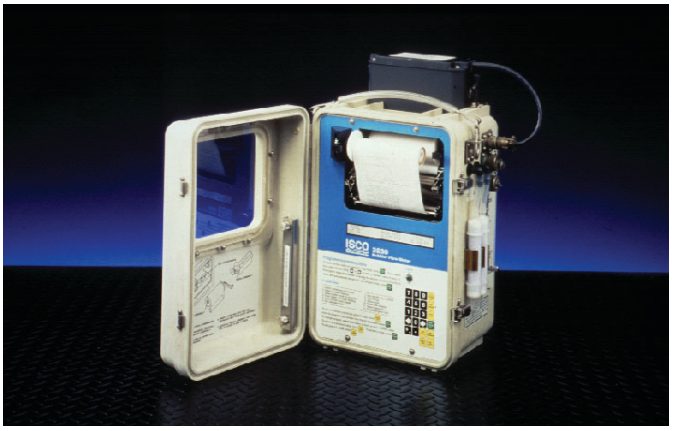
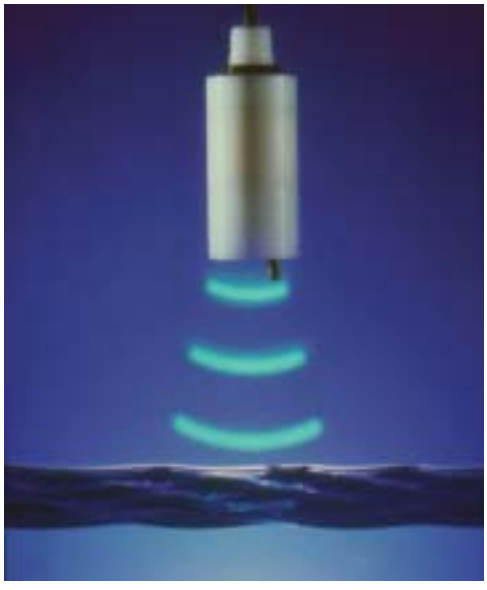
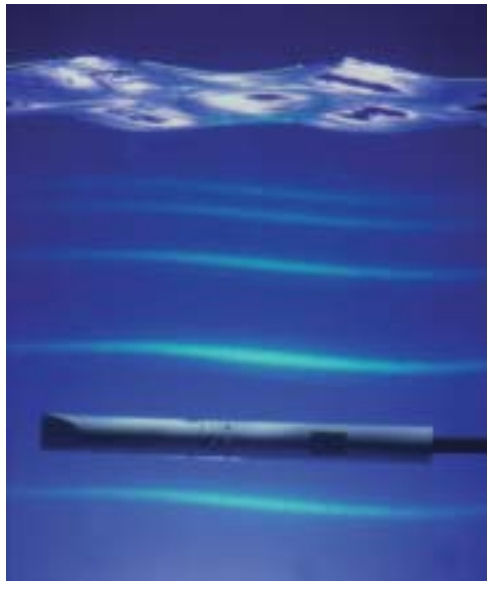


**Figure 10–6** Stage recorders—Continued

**d Ultrasonic**



**e Pressure transducer**



**f Bubbler**

## (f) Stage-discharge relationship

Because stage is only a measure of the height of the water, not the discharge in the stream, a stage-discharge relationship for the open channel station must be developed. Simultaneous measurement of stage and discharge is needed to develop the rating equation for an open channel stage recorder. Once the relationship is developed, stage measurements can be used to compute discharge. Discharge in open channels typically is determined using current meter measurements (fig. 10–7).

The primary method for determining discharge is the velocity-area method, although other techniques, such as the salt dilution method, exist as well (USDI, BOR 1977). The velocity-area method uses the equation:

$$Q = AV \qquad [10–1]$$

where:
$Q$ = discharge
$A$ = cross-sectional area of stream
$V$ = stream velocity

When conducting a discharge measurement using a velocity meter, the stream cross-section is divided into subsections and velocity measurements are taken at each subsection. For sections deeper than 2.5 feet, two velocity measurements are taken at 0.2 and 0.8 times the depth; otherwise a single velocity measurement is taken at 0.6 times the depth.

A good description of guidelines for making discharge measurements is given in Buchanan and Somers (1969) and the National Handbook for Recommended Methods for Water-Data Acquisition (USGS 1977). Example 10–1 shows the recommended steps for developing a rating equation.

The general form of the rating equation is:

$$Q = CH^b \qquad [10–2]$$

where:
$Q$ = the discharge ($ft^3/s$)
$C$ = the regression intercept, which is the discharge where $H = 1.0$
$H$ = the stage (ft)
$b$ = the slope of the regression

**Figure 10–7** Pygmy current meter

This equation should plot as a straight line on log-log paper. Figure 10–8 shows an example rating curve. Note: By convention, the discharge (Q) is plotted as the abscissa even though it is the dependent variable.

A minimum of 15 pairs of stage and discharge measurements should be used to develop the rating equation shown as points on figure 10–8. At times, two rating equations are developed; one for low flow and one for high flow. The ratings should be checked periodically because shifts in the equation may occur. Changes in the rating curve may be caused by scouring or filling the streambed, the growth of aquatic vegetation, or by icing. Figure 10-8 displays two of these cases. If scour occurs, the rating would be expected to move to the right and concave downward. That is, for an equal stage, the discharge would be greater after scouring. With filling, the rating would move left and concave upward (USGS 1977).

**Example 10–1** Developing a rating equation

Use the following steps to develop the rating equation:

1. Log transform paired values of Q and H
2. Perform a linear regression of Q vs. H with Q as the dependent variable.
3. Obtain intercept (C) and slope (b)
4. Add coefficients to the equation:

$$\log Q = \log C + b \log H$$

5. Transform equation to the form:

$$Q = CH^b$$

by taking the antilog of equation in step 4, so that:

$$Q = 10^c H^b$$

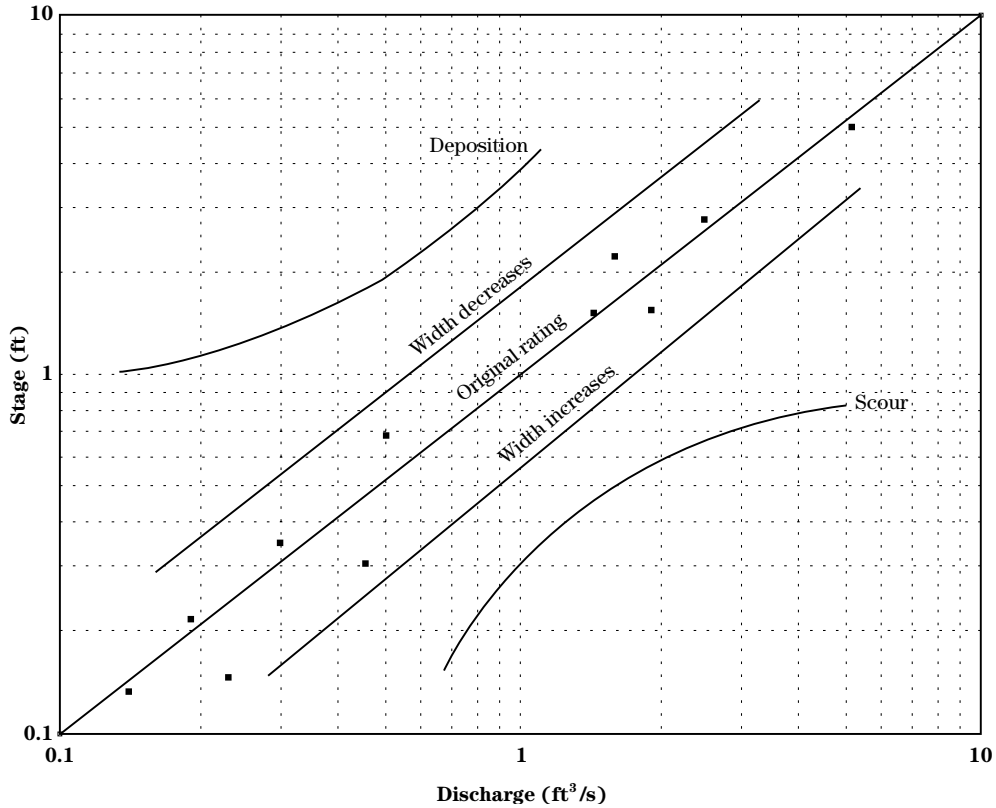
For example, if the intercept (C) was 0.05 and the slope (b) was 2.54, the equation would be:

$$Q = 10^{0.05} H^{2.54}$$

or

$$Q = 1.12 H^{2.54}$$

**Figure 10–8** Stage-discharge rating curve



If the width of the downstream control section increased, the intercept would be expected to increase. That is, for an equal discharge, the stage would be lower if the width decreased. The opposite would happen if the width of the control decreased.

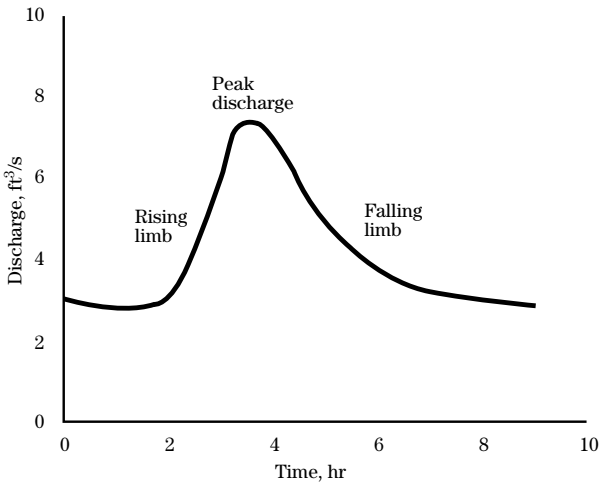
Several methods are available for extending the relationship for higher observed stages (Schulz 1976; USGS 1977). Also, additional adjustments can be made to the rating. For example, there is a histeresis effect of rising limb discharge exceeding falling limb discharge at the same stage (fig. 10–9). Assistance of agencies, such as the U.S. Geological Survey, may be necessary where large streams are involved.

## (g) Heating in cold climates

Year around monitoring is necessary in many cases. In cold climates, heating may be needed to guarantee sample collection. Heating design varies with the type of gaging station. Generally, heating requirements can be reduced by insulating. For many gaging stations, insulating means having the stilling well buried into the soil as far as possible. Weir plates can be kept open by covering with a wooden box during the winter. The box also can be heated for further protection.

Where electric power is present, heating is relatively easy. Heat lamps, light bulbs, space heaters, or stock tank heaters have all proven to prevent freeze-up.

**Figure 10–9** Stream hydrograph




Sample lines to automatic samplers can be prevented from freezing by wrapping with electrical heat tape. When electric power is not present, propane can provide heat. A regulator with a "fail safe" must be used with the pilot light to prevent gas leakage and possible explosions in the stilling well. A pilot light propane heater is shown in figure 10–10a. This type of system could heat a stilling well and instrument shelter on little gas. Catalytic propane heaters can be used to provide a more directed heat source, such as needed at the mouth of an H-type flume (fig. 10–10b). However, these heaters require much more gas than the smaller pilot light heater.

**Figure 10–10** Heating devices

**a Pilot light propane heater**

**b Catalytic propane heater**

# 614.1002 Concentration sampling

A variety of devices have been developed for taking samples for water quality analysis. Sampling may be either attended or unattended, and unattended sampling may be either passive or automated. The type of sampling device varies with the scale of the project, the objectives, and the project budget.

## (a) Grab samples

A grab sample is a discrete sample that is taken at a specific place and time. A series of grab samples lumped together are considered composite samples. Grab samples may not be representative of the water quality of the body of water being sampled for several reasons. Water quality may vary with depth or distance from the streambank.

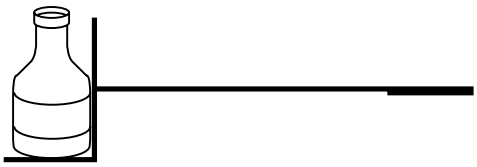
A grab sample typically is taken by hand with a sampling bottle. The bottle should be held just below the surface of the water to avoid contaminants in the surface film. The sample bottle can be connected to a holder on the end of a rod with plastic tubing to obtain a sample at some distance away (fig. 10–11a).

Sampling lake systems requires more specialized equipment. Frequently used samplers include Kemmerer, VanDoren, or Beta bottles. These samplers can obtain a sample from any depth in the water column. An inexpensive sampler consisting of a bottle with a pullable stopper (fig. 10–11b) has been described by Schwoerbel (1970) and WHO (1978). The same effect could be achieved by lowering a weighted, open bottle upside down, and inverting it with a second rope, allowing the air to escape and the bottle to fill with water.

Depth integrating samplers have been used especially for sediment sampling. For example, the DH-48 sampler (fig. 10–12) is designed to continuously obtain a sample as it is lowered to the streambed and then raised to the surface. In lakes, hoses have been used to obtain a sample of the total column of water. The hose is lowered into the water and allowed to fill. A rope attached to the bottom end is used to raise the lower end of the hose to the surface thereby collecting the entire sample of water in the hose. Pumps also have been used to sample lake water.

**Figure 10–11** Grab samplers

**a Rod sampler**



**b Lake sampler**

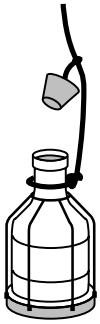


**Figure 10–12** DH-48 sampler

## (b) Passive samplers

A passive sampler collects a water quality sample by action of the flow of water itself. A tipping bucket discharge station is well suited to passive sampling (fig. 10–13). Slots or funnels under the tipping bucket have been used to collect water samples (Chow 1976; Johnston 1942; Russell 1945). H-flumes also have been widely used for passive sample collection. The Coshocton wheel (fig. 10–14) has been used to sample 1 percent of discharge for sediment sampling (USDA 1979). A splitter below a Coshocton wheel has been used to reduce the size of the sample to 0.1 percent of discharge (Coote & Zwerman 1972). Holes drilled in the mouth of an H-flume also have been used to collect stage-integrated samples through tubing.

Passive devices have been used for plot runoff. Most involve some sort of divisor and collection tank (Coote & Zwerman 1972; Dressing, et al. 1987; Geib 1933; Kohnke & Hickok 1943; USDA 1978) unless the plot is sized to collect the entire sample in a collection jug, as shown in figure 10–1.

The primary advantages of a passive sampler are that it can be unattended, requires little maintenance, and no power.

Stage samplers are another type of unattended passive sampler. Originally devised for suspended sediment sampling, a stage sampler consists of a series of bottles attached to a board arranged vertically at different stages (fig. 10–15). Each bottle has two tubes at different heights, which creates a siphon when filling. The disadvantages of this type sampler include collection of debris, some bias in size of sediment collected, sample taken near the water surface during the rising stage, and a filled bottle may have some mixture with later water (USDA 1979).

A single stage sampler was used by Schwer and Clausen (1989) to sample the outflow from dairy milkhouse waste pipes. Tubing was connected to the milkhouse drainage pipe with an extension collar. When the pipe flowed, part of the wastewater flowed through the tubing into a collection bottle. The bottle had a second tube as an air outlet.

**Figure 10–13** Tipping bucket with passive sampler

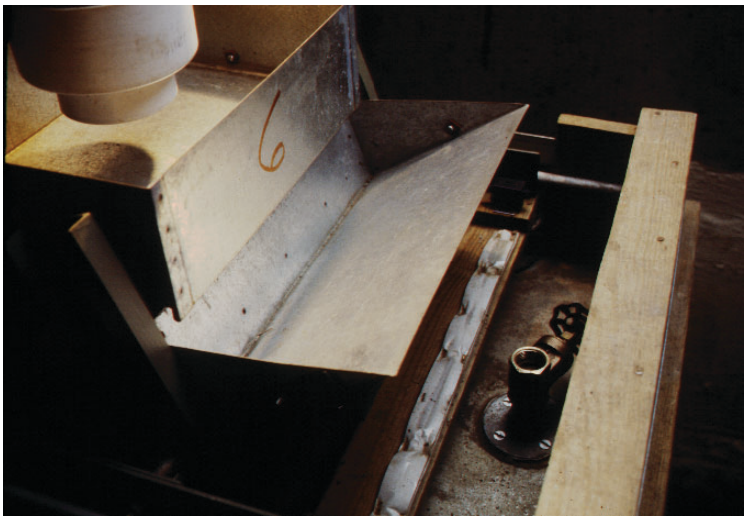
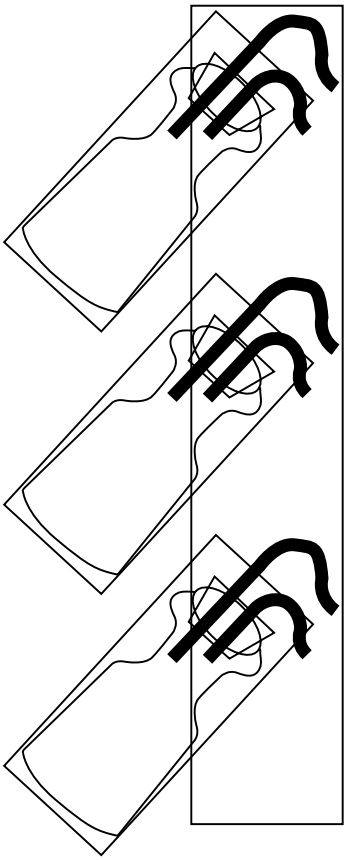


**Figure 10–14** Coshocton wheel

**Figure 10–15** Stage sampler



## (c) Automated samplers

Automated samplers are needed for larger streams and unattended sampling. These samplers typically allow programming of sample volume, time or flow interval between samples, and whether composite or discrete samples are taken. A summary of some of the older models available is in the National Park Service's publication "Automatic Water Samplers for Field Use" (NPS 1983). One of the common samplers in use is shown in figure 10–16. The ISCO sampler also can be connected to an ISCO flow meter to assist flow proportional sampling.

An inexpensive sampler developed in Canada is a submerged pipe section that has an opening operated by a solenoid. At timed intervals, a solenoid opens a port and allows a sample to enter the pipe. The volume of sample taken is proportional to the stage of the stream. The sample is removed by vacuum pump during a field visit.

The advantage of automated samplers is that they operate at all times, especially during runoff events, without attendance. However, these samplers are expensive and require maintenance.

One of the criticisms of pumping samplers is that a sample is taken from one point in the stream profile. Depth integrated intakes have been described by Eads and Thomas (1983) and McGuire, et al. (1980). These devices use a float to raise the intake with the stage and can collect a depth-integrated sample if the intake is perforated along its entire length (fig. 10–17).

**Figure 10–17** Depth-integrating intake

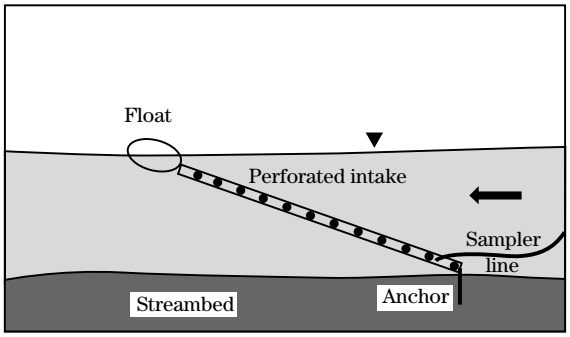
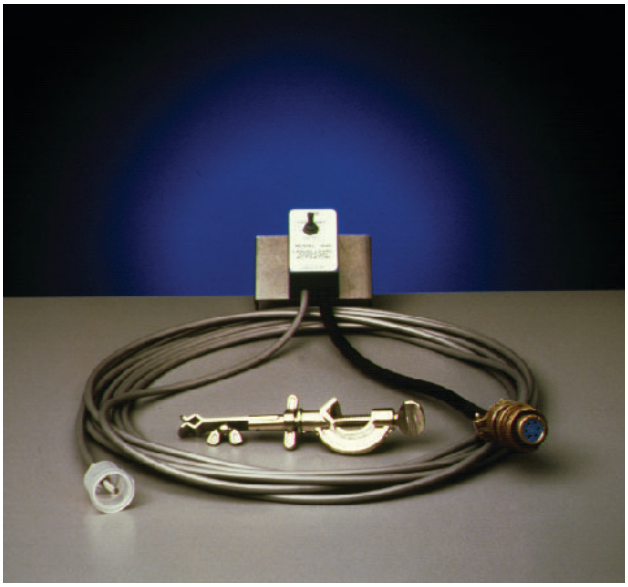



**Figure 10–16** ISCO automatic sampler (courtesy Instrumentation Specialties Company)

### (d) Actuated sampling

Actuated sampling is effective for sampling intermittent streams or for just sampling during storm events. Several options are available for initiating sampling during storms. Liquid level actuation has been used to initiate an ISCO sampling sequence (fig. 10–18). Precipitation sensors can also be used to initiate sampling. Programmable data loggers that also are monitoring stage could be used to initiate sampling. Various homemade float devices have been used to trip a switch and initiate samplers.

**Figure 10–18** ISCO liquid level actuator (courtesy Instrumentation Specialties Company)



## 614.1003 Precipitation monitoring

The extent of precipitation monitoring varies with the objectives of the study, but some precipitation monitoring is necessary in most monitoring projects. Precipitation data are useful for event sampling, for computing runoff coefficients for quality assurance programs, and for documenting rainfall conditions relative to a normal year. For most installations, both nonrecording and recording rain gages should be used. The nonrecording gage gives the total amount of precipitation; whereas the recording rain gage gives the time of precipitation. The total precipitation obtained by the recording rain gage should be adjusted to that measured in the nonrecording rain gage. A good background in precipitation monitoring is described in Agricultural Handbook 224 (USDA 1979), and guidance on maintenance is given in Weather Bureau Observing Handbook No. 2 (USWB 1970).

A variety of nonrecording and recording rain gages are commercially available. For the nonrecording gage, the National Weather Service standard 8-inch (20 cm) gage is most often used (fig. 10–19). For summer operation, a small amount of oil reduces evaporation. For winter operation, antifreeze can be added to the gage. The most common types of recording rain gages are either weighing bucket or tipping bucket (fig. 10–20). A weighing bucket gage can collect both rain and snow. For a tipping bucket gage to operate in the winter, it must be heated. However, the tipping bucket gage is easily adapted to data logging.

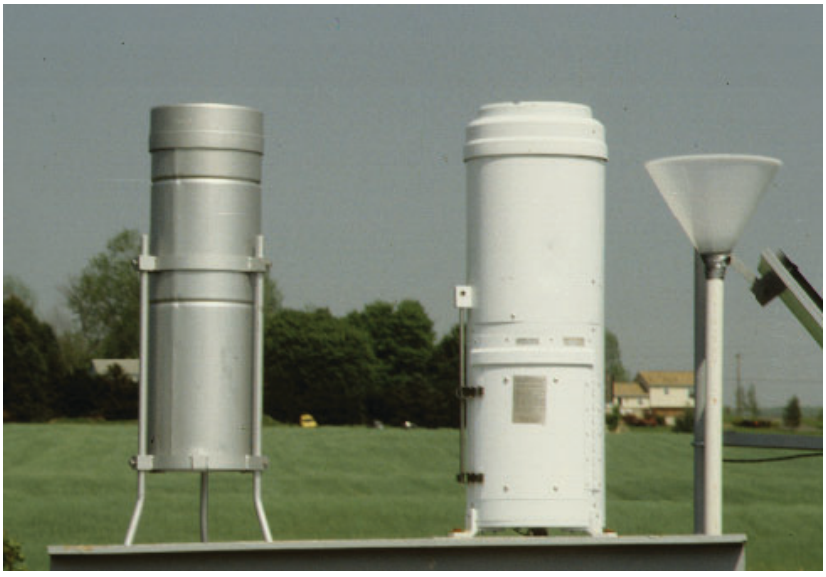
The location of the gage is important to precipitation monitoring. Recording and nonrecording gages should be placed at the same height and be leveled. The gages must be located in an opening where there is no obstruction within 45° of the lip of the gage. In areas of snowfall, the use of a windshield (fig. 10–21), such as an Alter shield, should be considered (USDA 1979). A windshield would be especially important in an open installation.

For some water quality studies, more than one gage may be necessary. The objective of precipitation monitoring must be considered when designing the

**Figure 10–19** Standard rain gage

**Figure 10–21** Standard rain gage altar shield

**Figure 10–20** Standard rain gage with tipping bucket and funnel gages



precipitation network. Other factors influencing the number and location of rain gages include topography, storm type, and the size of the area being studied. Monitoring in mountainous areas should definitely consider multiple gages.

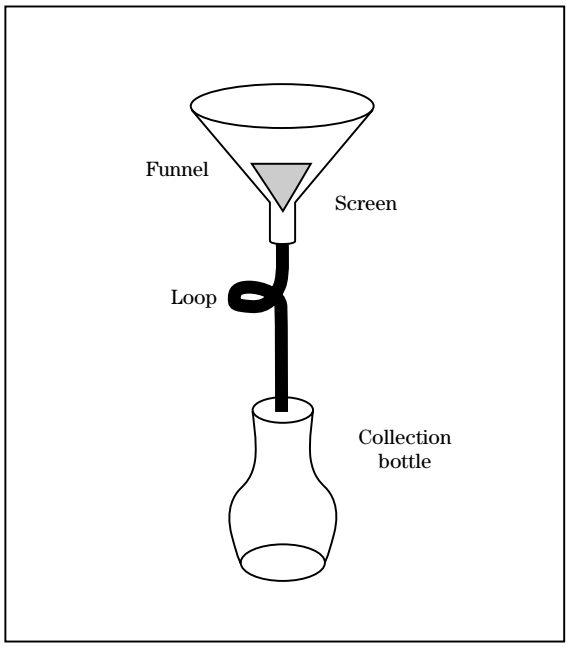
Knowledge of the quality of precipitation may be desired for some water quality studies. For example, studies examining the mass budget of nitrogen might consider N inputs in precipitation. Two common methods for sampling precipitation quality are wet-only collection and bulk precipitation.

Bulk precipitation can be easily sampled using a funnel gage (fig. 10–20) as described by Eaton et al (1973). A loop in the tubing leading from the funnel to the collection jug prevents evaporation (fig. 10–22a). A screen is recommended in the funnel opening to prevent large insects from entering the sample. Although this type of sampler is inexpensive and easy to construct, it collects any dry deposition that occurs on the funnel surface as well as rainfall. In addition, the funnel will not collect snow without bridging unless it is heated.

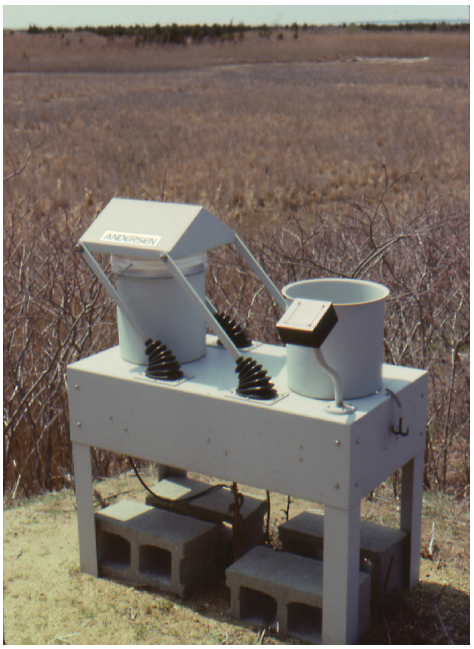
A wet-only sample can be obtained from a wet-dry deposition sampler as used by the NADP (Bigelow 1982). This sampler covers the precipitation bucket during dry periods thus preventing dry deposition from contaminating the sample (fig. 10–22b). A precipitation sensor opens the wet bucket during rainfall. The time of opening and closing the lid can be recorded on a rain gage that has a second pen attachment.

**Figure 10–22** Gages for precipitation chemistry

**a Funnel collector**




**b Wet-dry deposition sampler**



## 614.1004 Soil water sampling

Sampling the soil water may be useful for determining nutrient concentrations and possibly mass fluxes in the vadose zone of soils. A number of sampling techniques have been used to sample soil water. These samplers generally can be classified as tension and zero-tension. Tension lysimeters extract a sample of soil water at some suction and include porous ceramic cups, plate lysimeters (fig. 10–23a), and capillary-wick samplers. The zero-tension lysimeters collect gravitational water and have included funnels, pans, and troughs (fig. 10–23b).

Volumes of water collected in lysimeters are highly variable; therefore, a large number of lysimeters may be needed to adequately represent soil solution fluxes in an area. Water quality concentrations collected by tension and zero-tension lysimeters are different (Haines, et al. 1982).

## 614.1005 Biotic sampling

Biologic sampling includes collection and analysis of plankton, periphyton, macrophyton, macroinvertebrates, and fish. In addition, several techniques are available for determining primary production. Although not discussed in this guide, biotic sampling also may include bioassay.

### (a) Plankton

Plankton are organisms that move with the currents. Two major types of plankton are phytoplankton (plants) and zooplankton (animals). Knowledge of the phytoplankton is particularly useful in water quality monitoring studies because they are good indicators of nutrient enrichment.

Plankton are influenced by currents, temperature, light, turbidity, and various chemical variables, such as salinity, nutrients, and toxics (USEPA 1973). Most of these factors vary with depth, except in well-mixed systems.

**Figure 10–23** Soil water samplers

**a Porous cup and plate lysimeters**

**b Outlet to funnel lysimeters**

Plankton samples can be obtained by net, water bottle, or with a pump (Schwoerbel 1970). Various plankton nets are available for sampling, the most common of which is the Wisconsin plankton net (fig. 10–24a). Plankton nets collect what is termed *net plankton* because some plankton may pass through the net. These nets are generally used for qualitative analysis.

Plankton also can be collected with a water bottle, such as a Kemmerer (fig. 10–24b), VanDorn, or Beta bottle. A quantitative sample of plankton can be obtained because the volume of water collected is known. Water bottles obtain a sample of plankton from a particular location and layer; therefore, the number of samples needed is subject to the variability in sampling (ch. 9).

Plankton samples collected with a pump can be obtained from any depth and of any volume. However, the pump tubing should be cleaned between samples, and the pump may break apart some plankton.

Once collected, plankton should be preserved and enumerated using standard techniques (USEPA 1973). In some cases chlorophyll analysis should be performed on the plankton as an indicator of the biomass.

## (b) Periphyton

The periphyton are organisms that mostly are attached to underwater substrates, such as rocks or macrophytes. These organisms may be predominant in shallow and running bodies of water. They also indicate water quality conditions.

Artificial substrates are used to quantitatively collect periphyton samples. They include glass microscope slides or the Hester-Dendy sampler (fig. 10–24c). Samplers are left in the field for about 2 weeks and then removed. Zooplankton and macroinvertebrates may graze on the periphyton, which will result in an underestimate of periphyton growth. The resulting samples should be preserved and enumerated. Biomass analysis is often used to express the amount of periphyton present.

## (c) Macrophyton

Large aquatic plants are termed macrophyton. In many cases these plants are what many perceive to be the water quality problem. Macrophytes are influenced by light (turbidity), nutrients, and sediment. Qualitatively, macrophytes may be identified to species and classified as to the relative cover. Quantitative sampling might involve small plots with analysis of the number of stems or the biomass. Air photography often is used to delineate boundaries of plant communities.

## (d) Macroinvertebrates

Aquatic macroinvertebrates are animals that are large enough to be seen with the unaided eye and include insects, mollusks, worms, and crustaceans. Their presence is seasonally-dependant and influenced by type of substrate, light, oxygen content, water velocity, and various chemical constituents. They also are susceptible to various stressors. Because their locations vary, proper sampling is important. Quantitative sampling involves determining the numbers or biomass of macroinvertebrates per unit area. This type of information is often used to calculate an index, such as Beck's Biotic Index (Terrell & Perfetti 1989). Samples are collected using such devices as the Surber sampler (fig. 10–24d). These samplers are difficult to use in some habitats, such as rocky substrates.

Qualitative samples of macroinvertebrates also are taken. Such sampling allows determining what is present and the diversity of the community. Samples are collected using a wide variety of devices, including sediment samplers in deep water, such as the Ekman or Peterson dredge (fig. 10–25a & b). These types of samplers have several disadvantages (USEPA 1973).

Artificial substrates using baskets of rocks also have been used to collect macroinvertebrates. Drift nets are most commonly used to qualitatively assess the macroinvertebrate community. These nets come in various shapes (fig. 10–25c). Collected samples should be preserved before identification (Klemm et al. 1990).

The EPA's Rapid Bioassessment Protocols (RBP) are methods for assessing the biotic condition of streams in comparison to reference stations (Plafkin et al. 1989). Several indices are recommended using RBP level III.

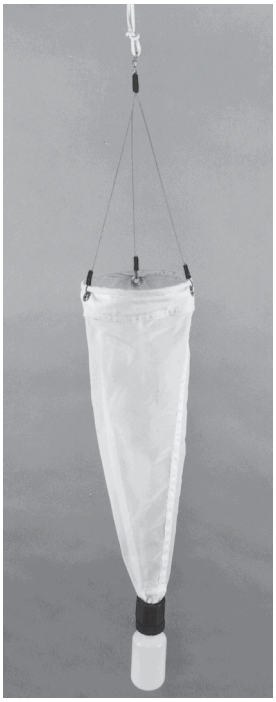
## (e) Fish

Water quality influences fish species, abundance, and health. Certain species of fish are sensitive to pollutants and serve as indicators of water quality. The species, abundance by species, size, growth rate, condition, reproductive success, and disease are of interest where fish are used in biomonitoring (USEPA 1973). Sampling of fish has been classified as either active or passive. Active sampling includes electrofishing and seines. Passive collection includes gill nets and trap nets. The various methods used to collect fish samples usually result in somewhat different species being collected. Fish are not located randomly throughout the water body; therefore, sampling must be adjusted.

The Rapid Bioassessment Protocol level V for fish describes methods for electrofishing and calculation of the Index of Biotic Integrity (IBI) and other metrics (Plafkin, et al. 1989).

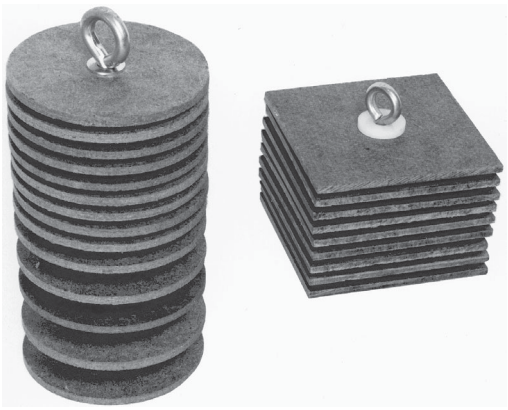
**Figure 10–24** Biotic samplers (courtesy Wildlife Supply Company)

**a Wisconsin plankton net**



**b Kemmerer water bottle**

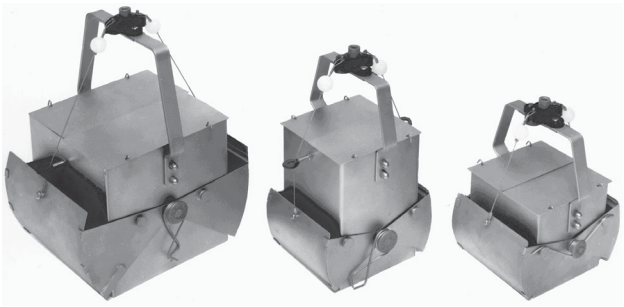
**c Hester-Dendy sampler**



**d Surber sampler**

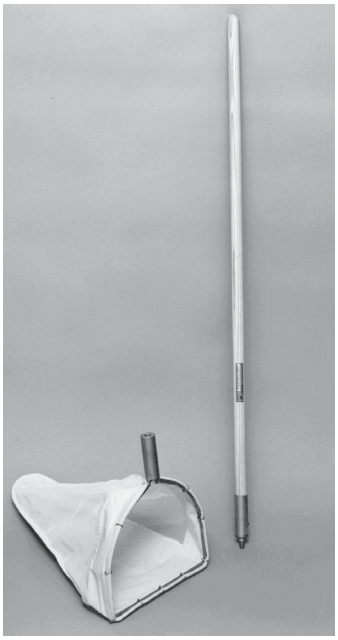
**Figure 10–25** Biotic and sediment samplers (courtesy Wildlife Supply Company)

**a Ekman dredge**



**b Peterson dredge**

**c D-type drift net**



## 614.1006 Sediment sampling

The sampling of sediment varies between running water and standing water. In running water, sediment has been divided into suspended sediment and bedload. Suspended sediment is carried by the water above the bed of the stream (USGS 1977). Bedload sediment is heavier than suspended sediment and moves along the bed of the stream.

Sampling of suspended sediment was previously described in this chapter. Suspended sediment-bedload sediment rating curves can be developed to estimate bedload transport. Bedload sampling is conducted by using bedload traps in the streambed or net samplers of a certain height, or it can be conducted by measuring changing cross sections in the stream.

In edge-of-field runoff, sediment is best sampled in some type of proportional sampler, such as the Coshocton wheel. Other bedload samplers have been developed for use with flumes. They consist of a slot across the flume that traps the bedload.

Sampling of sediment in standing water, such as lakes and ponds, generally is conducted with a type of coring device. The type of corer used varies with the depth of the water and the thickness and type of substrate. An example of a hand-held corer is shown in figure 10–26. Other types of corers include piston or drive samplers for deeper water.

In some cases lake sediment samples are obtained by diving, so that the sample remains undisturbed. The force of a sampler hitting the sediment may disturb the upper organic deposits, thereby biasing the sample. Sediment samples collected from standing water bodies are often analyzed for particle sizes, organic matter content, chemical content, dry weight, and volume.

**Figure 10–26** Hand-held sediment corer (courtesy Wildlife Supply Company)

## 614.1007 References

Bigelow, D.S. 1982. NADP instruction manual. Site operation. Natl. Atmos. Deposition Program, Nat. Resour. Ecol. Lab., CO State Univ., Ft. Collins.

Bodhaine, G.L. 1968. Measurement of peak discharge at culverts by indirect methods. United States Geological Survey, TWI 3-A3, 60 pp.

Buchanan, T.J., and W.P. Somers. 1969. Discharge measurement at gaging stations. Book 3, chap. A8 *in* Techniques of water-resources investigations of the United States Geological Survey.

Buchanan, T.J., et al. 1968, Stage measurements at gaging stations. United States Geological Survey, TWI 3-A7, 28 pp.

Carter, R.W., et al. 1968. General procedure for gaging streams. United States Geological Survey, TWI 3-A6, 13 pp.

Chow, T.L. 1976. A low-cost tipping bucket flowmeter for overland flow and subsurface stormflow studies. Can. J. Soil Sci. 56: 197-202.

Coote, D.R., and P.J. Zwerman. 1972. A conveniently constructed divisor for splitting low water flows. Soil Sci. Soc. Amer. Proc. 36:970-971.

Dressing, S.A., J. Spooner, J.M. Kreglow, E.O. Beasley, and P.W. Westerman. 1987. Water and sediment sampler for plot and field studies. J. Environ. Qual. 16(1):59.

Eads, R.E., and R.B. Thomas. 1983. Evaluation of a depth proportional intake device for automatic pumping samplers. Water Resour. Bull. 19(2):289-292.

Eaton, J.S., G.E. Likens, and F.H. Bormann. 1973. Throughfall and stemflow chemistry in a northern hardwood forest. J. Ecol. 61:495-498.

Geib, H.V. 1933. A new type of installation for measuring soil and water losses from control plots. J. Amer. Soc. Agron. 25(7):429-440

Grant, D.M. 1979. ISCO open channel flow measurement handbook. Instrumentation Spec. Co., Lincoln, NE.

Guy, H.P. 1969. Laboratory theory and methods for sediment analysis. United States Geological Survey, TWI 5-C1, 58 pp.

Guy, H.P., et al. 1970. Field methods of measurement of fluvial sediment. United States Geological Survey, TWI 3-C2, 59 pp.

Gwinn, W.R., and D.A. Parsons. 1976. Discharge equations for the HS, H, and HL flumes. J. Hydraul Div., Amer. Soc. Agric. Eng. 102:73-88.

Haines, B.L., J.B. Waide, and R.L. Todd. 1982. Soil solution nutrient concentrations sampled with tension and zero-tension lysimeters: Report of discrepancies. Soil Sci. Soc. Amer. J. 46:658-661.

Johnston, C .N. 1942. Tilt buckets for measuring runoff and erosion. Agric. Engr. 23:161-162.

Kennedy, E.J. 1985. Discharge ratings at gaging stations. United States Geological Survey, TWI 3-A10, 59 pp.

Kennedy, E.J. 1990. Levels at streamflow gaging stations. United States Geological Survey, TWI 3-A19, 27 pp.

Kilpatrick, F.A., et al. 1983. Use of flumes in measuring discharge. United States Geological Survey, TWI 3-A14, 46 pp.

Kilpatrick, F.A., et al. 1985. Measurement of discharge using tracers. United States Geological Survey, TWI 3-A16, 52 pp.

Klemm, D.J., P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. Macroinvertabrate field and laboratory methods for evaluating the biological integrity of surface waters. Office of Research and Development, Washington, DC, EPA/600/4-90/030.

Kohnke, H., and R.B. Hickok. 1943. An automatic aliquot runoff sampler. Soil Sci. Soc. Amer. Proc. 8:444-447.

Laenen, A. 1985. Acoustic velocity meter systems. United States Geological Survey, TWI 3-A17, 38 pp.

McGuire, P.E., T.C. Daniel, D. Stoffel, and B. Andraski. 1980. Sample intake position and loading rates from nonpoint source pollution. Envir. Mgmt. 4(1):73-77.

National Park Service, Water Resources Field Support Laboratory. 1983. Automatic water samplers for field use. U.S. Dept. of the Int. Natl. Park Serv., Water Resourc. Field Sup. Lab. Rep. No. 83-1, Fort Collins, CO.

Plafkin, J.L., M.T. Barbour, K.D. Porter, .S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macro-invertebrates and fish. U.S. EPA EPAl444/4-89001, Off. of Water. Washington, DC.

Replogle, J.A., and A.J. Clemmens. 1981. Measuring flumes of simplified construction. Trans. Amer. Soc. Agric. Eng. pp. 362-366.

Russell, M.B. 1945. A portable runoff measuring device. J. Amer. Soc. Agron. 37(8):589-594.

Schulz, E.F. 1976. Problems in applied hydrology. Water Resourc. Publ., Fort Collins, CO.

Schwer, C.B., and J.C. Clausen, 1989. Vegetative filter treatment of dairy milkhouse wastewater. J. Environ. Qual. 18(4):446-451.

Schwoerbel, J. 1970. Methods of hydrobiology. Pergamon Press, New York, NY.

Skougstad, M.W., et al. 1979. Methods for determination of inorganic substances in water and fluvial sediments. United States Geological Survey, TWI 5-A1, 626 pp.

Smith, C.N., D.S. Brown, J.D. Dean, R.S. Parrish, R.F. Carsel, and A.S. Donigian, Jr. 1985. Field agricultural runoff monitoring (FARM) manual. EPA/600/385/043, Environ. Res. Lab., Athens, GA.

Terrell, C.R., and P.B. Perfetti. 1989. Water quality indicators guide: Surface waters. USDA Soil Conserv. Serv. SCS-TP-161, Washington, DC.

United States Department of Agriculture. 1979. Field manual for research in agricultural hydrology. Agric. Handb. 224, Washington, DC.

United States Department of Interior, Bureau of Reclamation. 1977. Water measurements manual. 2nd edition.

United States Environmental Protection Agency. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. USEPA 670/4-73-001. Off. Res. & Dev., Cincinnati, OH.

United States Geological Survey. 1977. Methods for collection and analysis of aquatic biological and microbiological samples. Edited by P.E. Greeson, et al., TWI 5-A4, 332 pp.

United States Geological Survey. 1977. National handbook of recommended methods for water-data acquisition. Chapter 1: Surface water, Off. Water Data Coord., Reston, VA.

United States Geological Survey. 1987. Methods for the determination of organic substances in water and fluvial sediments, edited by R.L. Wershaw, et al., TWI 5-A3, 80 pp.

United States Weather Bureau. 1970. Weather Bureau observing handbook No. 2. Substation observations. U.S. Dep. Com., Environ. Sci. Serv. Admin., Silver Spring, MD.

World Health Organization. 1978. Water quality surveys—A guide for the collection and interpretation of water quality data. IHD-WHO Working group on water quality. UNESCO/WHO, Geneva, Switzerland.

United States
Department of
Agriculture

**Natural
Resource
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 11 Sample Collection and Analysis

(430–VI–NWQH, September 2003)

# Chapter 11 Sample Collection and Analysis

Contents:

**Tables**



**Figures**

# Chapter 11 Sample Collection and Analysis

## 614.1100 Introduction

Obtaining high quality data requires following appropriate techniques for obtaining water quality samples and analyzing them for their constituents. Equally important is the need to describe in detail how the work is being conducted so that others can duplicate the information. This chapter describes suggested techniques for collecting a water sample, and recommended quality assurance and quality control procedures for both the lab and the field. Two references may be helpful for volunteer monitoring (US EPA 1990, Simpson 1991)

## 614.1101 Sample collection

Different sample collection procedures should be followed depending upon the type of sample (grab, automatic) and whether the system is a lake, stream, or ground water. Generally, a bottle used for a grab sample should be rinsed with the sample water two or three times before filling unless the bottle contains a preservative, in which case there should be no rinsing (APHA 1989). If samples are collected from pipes under pressure, make sure that the system has been flushed for a sufficient period to guarantee that new water is being sampled. Bacteria samples are collected in sterilized bottles.

Collection of samples from wells can be complicated. Water within the well may be stagnant and not representative of surrounding ground water. The well should be purged for a sufficient amount (3 to 10 well-bore volumes) to ensure that the sample is representative of the ground water. More than 5 minutes may be required to remove over 80 percent of the well-bore volume when pumped at 1.3 gpm. Some recommend that well purging should be conducted at the rate of well replenishment. This would not be the case for well-mixed aquifers. Sampling for volatile organics may require special precautions and possibly no purging.

Sampling of volatile substances requires special sampling equipment in wells. The release of gases during pumping can change the pH of the water and. therefore, the solubility of metals. Oxidation of the sample during pumping can influence organics, sulfur, iron, ammonium, and manganese (Driscoll 1986).

Generally, all samples should be collected so that the bottle is completely full. This reduces volatilization losses. An exception to this would be if the sample was to be frozen, in which case room for expansion upon freezing should be left in the container. Sampling of toxic substances require extra precautions, including gloves, coveralls, aprons, eye protection, and in the case of toxic vapors, a respirator may be necessary.

The quantity of sample to collect is dependent upon the type of analyses to be conducted. Suggested volumes are given in table 11–1. The total volume should include a summation of the recommended volumes plus amounts for the quality assurance program. In addition, the analysis of a sample may need to be repeated. Therefore, it is generally recommended that the total recommended volume be doubled (Shelley 1977).

**Table 11–1** Recommended methods for sample collection and preservation (USEPA 1983)

| Measurement | Vol. req. (mL) | Container P=plastic; G=glass | Preservative | Maximum holding time |
|---|---|---|---|---|
| **Physical properties** | | | | |
| Color | 50 | P,G[1/] | Cool, 4 °C | 48 hrs |
| Conductance | 100 | P,G | Cool, 4 °C | 28 days |
| Hardness | 100 | P, G | $HNO_3$ to pH < 2 | 6 mos |
| Odor | 200 | G only | Cool, 4 °C | 24 hrs |
| pH | 25 | P,G | None req. | Analyze immediately |
| Residue | | | | |
| Filterable | 100 | P,G | Cool, 4 °C | 7 days |
| Nonfilterable | 100 | P,G | Cool, 4 °C | 7 days |
| Total | 100 | P,G | Cool, 4 °C | 7 days |
| Volatile | 100 | P,G | Cool, 4 °C | 7 days |
| Settleable matter | 1,000 | P,G | Cool, 4 °C | 18 hrs |
| Temperature | 1,000 | P,G | None req. | Analyze immediately |
| Turbidity | 100 | P,G | Cool, 4 °C | 48 hrs |
| **Metals** | | | | |
| Dissolved | 200 | P,G | Filter on site | 6 mos |
| | | | $HNO_3$ to pH <2 | |
| Suspended | 200 | | Filter on site | 6 mos |
| Total | 100 | P,G | $HNO_3$ to pH <2 | 6 mos |
| Chromium | 200 | P,C | Cool, 4 °C | 24 hrs |
| Mercury dissolved | 100 | P,G | Filter | 28 days |
| | | | $HNO_3$ to pH <2 | |
| Total | 100 | P G | $HNO_3$ to pH <2 | 28 days |
| **Inorganics, nonmetallics** | | | | |
| Acidity | 100 | P,G | Cool, 4 °C | 14 days |
| Alkalinity | 100 | P,G | Cool, 4 °C | 14 days |
| Bromide | 100 | P,G | None req | 28 days |
| Chloride | 50 | P,G | None req | 28 days |
| Chlorine | 200 | P,G | None req | Analyze immediately |
| Cyanides | 500 | P,G | Cool, 4 °C | 14 days |
| | | | NaOH to pH >12 | |
| | | | 0.6g ascorbic acid[6] | |
| Fluoride | 300 | P,G | None req | 28 days |
| Iodide | 100 | P,G | Cool, 4 °C | 24 hrs |

**Table 11–1** Recommended methods for sample collection and preservation (USEPA 1983)—Continued

| Measurement | Vol. req. (mL) | Container P=plastic; G=glass | Preservative | Maximum holding time |
|---|---|---|---|---|
| **Inorganics, nonmetallics** (continued) | | | | |
| Nitrogen | | | | |
| Ammonia | 400 | P,G | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| Kjeldahl, total | 500 | P,G | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| Nitrate plus Nitrite | 100 | PG | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| Nitrate | 100 | P,G | Cool, 4 °C | 48 hrs |
| Nitrite | 50 | P,G | Cool, 4 °C | 48 hrs |
| Dissolved oxygen | | | | |
| Probe | 300 | G bottle & top | None req. | Analyze immediately |
| Winkler | 300 | G bottle & top | Fix on site and store in dark | 8 hrs |
| Phosphorus | | | | |
| Ortho-phosphate Dissolved | 50 | P,G | Filter on site Cool, 4 °C | 48 hrs |
| Hydrolyzable | 50 | P,G | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| Total | 50 | P,G | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| Total dissolved | 50 | P,G | Filter on site Cool, 4 °C $H_2SO_4$ to pH <2 | 24 hrs |
| silica | 50 | P only | Cool, 4 °C | 28 days |
| sulfate | 50 | P,G | Cool, 4 °C | 28 days |
| sulfide | 500 | P,G | Cool, 4 °C add 2 mL zinc acetate plus NaOH to pH >9 | 7 days |
| sulfite | 50 | P,G | None req. | Analyze immediately |
| **Organics** | | | | |
| BOD | 1,000 | P,G | Cool, 4 °C | 18 hrs |
| COD | 50 | P,G | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| Oil & grease | 1,000 | G only | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| Organic carbon | 25 | P,G | Cool, 4 °C $H_2SO_4$ / HCl to pH < 2 | 28 days |
| Phenolics | 500 | G only | Cool, 4 °C $H_2SO_4$ to pH <2 | 28 days |
| MBAS | 250 | P,G | Cool, 4 °C | 48 hrs |
| NTA | 50 | P,G | Cool, 4 °C | 24 hrs |

## 614.1102 Sample preservation and transport

Once a sample is collected, it has the opportunity to change its composition through chemical, physical, and biological processes. Some changes may not be preventable, so rapid analysis is recommended in those situations (USEPA 1983).

Examples of physical changes include settling of solids, adsorption of certain cations on container walls, and loss of dissolved gases. Chemical changes could include precipitation, dissolution from sediments, complexation with other ions, and changes in valence state. Biological reactions could result in both the uptake and release of certain constituents. Microbial activity may change the species of nitrogen present (APHA 1989).

Preservation techniques are aimed at slowing biological activity, hydrolysis, volatility, and absorption. The primary preservation methods are acidification, refrigeration, filtration, and preventing light from reaching the sample (USEPA 1983; APHA 1989). Recommended preservation methods for most chemical properties of water are summarized in table 11–1. The appropriate sample volume, type of sampling container, and maximum holding time also are listed. A similar listing is given in the "Standard methods for the examination of water and wastewater" (APHA 1989).

Using a sample bottle that has the preservative already added may be useful for composite sampling. The sample becomes preserved immediately upon collection. Preservation of biological samples is also important (Klemm, et al. 1990). Without preservation predation within the sample may occur or the specimens may degrade. Generally, adding an equal volume of 95 percent ethanol to the sample results in an ethanol strength of 70 percent, which is adequate to preserve the sample (USEPA 1973). Plankton can be preserved with Lugol's solution (APHA 1989).

The sample container is also important. Glass containers may leach sodium and silica, and plastic containers may sorb organics (APHA 1989). Certain pesticides may adsorb to silicone rubber and tygon, but not high-density polyethylene or acrylic plastic (Topp and Smith 1992). Teflon and stainless steel are appropriate containers in certain cases.

Transportation to the laboratory should be direct. Transport should be done following some methods of preservation, such as cooling and keeping in the dark. Using dry ice for cooling is not recommended (APHA 1989).

## 614.1103 Methods of laboratory analysis

It is not within the scope of this handbook to describe methods of laboratory analysis for water quality variables. Two important references on this subject are Standard Methods for the Examination of Water and Wastewater (APHA 1989) and Methods for Chemical Analysis of Water and Wastes (USEPA 1983).

**Table 11–2** Water quality variables for which field test kits are available (Kunkle & Ricketts 1984)

| Water quality variables |
|---|
| Alkalinity, hardness |
| Ag, Al, Ba, Ca, Cd, Co, Cr, Cu, |
| Fe, Hg, K, Mg, Mn, Na, Ni, Pb, Zn |
| Ammonia, nitrate, nitrite |
| Total phosphorus, ortho-phosphate |
| Acidity, COD, color, pH, salinity |
| Dissolved oxygen, carbon dioxide |
| Turbidity, dissolved solids |
| Arsenic |
| Bromine |
| Chloride, chlorine |
| Cyanide chromate |
| DEEA |
| Detergents |
| EDTA/NTA |
| Fluoride |
| Formaldehyde |
| Gasoline |
| Hydrogen peroxide |
| Hydrogen sulfide |
| Iodine |
| Lignin |
| Molybdate |
| Ozone |
| pH |
| Phenol |
| Silica |
| Sulfate, sulfide |
| Tannin |
| Temperature |

## 614.1104 Field test kits

Many test kits are available for field analysis of a wide variety of water quality variables (tables 11–2 & 11–3). These kits range in level of sophistication and price. Field test kits are not considered as accurate as laboratory analyses, but may be useful in many situations (Kunkle & Ricketts 1984).

Kits function in one of three ways.

- *Color comparator* kits use the addition of a reagent to a sample, which results in a color development. The intensity of the color is compared to a color wheel or color tubes.
- *Colorimeter and spectrophotometer* kits use color development, which is read in battery powered colorimeters. Colorimeter kits are the most expensive kit.
- *Titration* kits use the addition of a reagent until a color change occurs.

Electric meters for field pH, conductivity, and dissolved oxygen are also available.

**Table 11–3** Partial list of manufacturers of field test kits

| Manufacturers | |
|---|---|
| Bausch and Lomb | (716) 338-8317 |
| CHEMetrics, Inc. | (703) 788-9026 |
| Ecologic Instrument | (516) 567-9000 |
| EM Science | (609) 423-6300 |
| Hach Company | (303) 669-3050 |
| Hellige, Inc. | (516) 222-0300 |
| In-Situ, Inc. | (307) 742-8213 |
| Kahl Scientific | (619) 444-2158 |
| LaMotte Chemical | (301) 778-3100 |
| Millipore Corp. | (617) 875-2050 |
| Soiltest, Inc. | (312) 869-5500 |
| Solomat | (203) 849-3111 |
| Spectrum Technologies, Inc. | (815) 436-4440 |
| Taylor Chemicals, Inc. | (301) 472-4776 |

## 614.1105 Quality assurance

Quality Assurance (QA) is the total integrated program for assuring the reliability of monitoring and measurement data (USEPA 1988). Quality assurance programs should allow determining statistical limits of confidence in the data (Taylor 1984). The program also should document the procedures that are followed (Dillaha, et al. 1988). Quality assurance is composed of quality control and quality assessment. Quality Control (QC) refers to activities conducted to provide high quality data (Lawrence and Chau 1987). Quality assessment refers to techniques used to evaluate the effectiveness of the program (Taylor 1984).

An overall outline for a quality assurance plan is given in figure 11–1.

**Figure 11–1** Outline of a quality assurance plan (USEPA 1988)

1. Cover page
2. Table of contents
3. Project description
   a. Objectives and scope
   b. Data usage
   c. Design and rationale
   d. Monitoring parameters and collection frequency
   e. Parameter table
4. Project organization and responsibility
5. Data quality requirements
   a. Precision
   b. Accuracy
   c. Representativeness
   d. Comparability
   e. Completeness
6. Sampling and laboratory procedures
7. Sample custody procedures
8. Calibration procedures and preventive maintenance
9. Documentation, data reduction and reporting
10. Data validation
11. Performance and system audits
12. Corrective action
13. Reports
14. Literature cited

## 614.1106 Quality control

Table 11–4 summarizes the major components of a quality control program. Good Laboratory Practices (GLPs) refer to general practices, such as glassware cleaning and preparation. Standard Operating Procedures (SOPs) are recipes for conducting analyses. These would include standard methods (APHA 1989) and approved methods (USEPA 1983). SOPs would also exist for sample handling (chain of custody records) and calibration and maintenance procedures.

Education and training refer to procedures used to support and verify the training of sampling and analysis personnel. This is especially important for safety training. Supervision includes the monitoring and review of techniques and data to allow for timely corrective actions.

**Table 11–4** Components of a quality control program (after Taylor 1984)

| |
|---|
| Good Laboratory Practices (GLPs) |
| Standard Operating Procedures (SOPs) |
| Education/training<br>sample custody procedures<br>calibration and maintenance |
| Supervision |

# 614.1107 Quality assessment

Quality assessment allows feedback on how well the quality control program is operating. Table 11–5 summarizes the components of a quality assessment program, and table 11–6 shows the indicators of quality data. Indicators of data quality include:

- precision
- accuracy
- representativeness
- comparability
- completeness

A description of each indicator follows.

**Table 11–5** Components of a quality assessment program

**Internal**

Duplicate samples
Standard additions (spikes)
Tests of sampling frequency
Tests of reason with comparable data
Missing analysis records
Standard curves
Internal audit

**External**

Exchange sample with other lab
External known materials
External audit

## (a) Precision

Precision is a measure of the closeness by which repeated measures of a given sample agree with each other. The Relative Standard Deviation (RSD) of duplicate samples provides the overall precision of the study, including random sampling errors and errors associated with sample preparation and analysis.

### (1) Frequency

Duplicate analysis should be performed for every 20th sample collected for which there is sufficient quantity for splitting or at least one per analytical run.

### (2) Calculation

The relative standard deviation, which also is the coefficient of variation, between the duplicates can be calculated as follows:

$$RSD = \frac{S}{\overline{X}} \times 100 \qquad [11–1]$$

where:

$S$ = standard deviation

$\overline{X}$ = the mean

### (3) Acceptance

An RSD of more than 10 percent could require notification of the onsite QA officer.

**Table 11–6** Quality control samples

| Indicator | Sample type | Frequency | Measure | Acceptance criteria (%) |
|---|---|---|---|---|
| Precision | Duplicate | 1/20 | RSD | 10 |
| Accuracy | Spike | 1/20 | % recovery | 90-110 |
| Representative | Multiple | Initial | n | ±20 |
| Completeness | All | Annual | % missing | <10 |
| Performance audit | EPA known | 4/yr | % recovery | 90-110 |

## (b) Accuracy

Accuracy (bias) is the degree of agreement between measured and true values. The percentage recovery of known standard additions to a sample provides the measure of accuracy for the study. The amount added should be sufficient to double the concentration.

### (1) Frequency

Every 20th sample collected in sufficient quantity for splitting should be spiked.

### (2) Calculation

Chemical recovery is calculated as follows:

$$\% \text{ Recovery} = \frac{A}{B+C} \times 100 \qquad [11\text{–}2]$$

where

$A$ = measured concentration of spiked sample
$B$ = measured concentration of unspiked sample
$C$ = concentration of known addition

### (3) Acceptance

A recovery of 90 to 110 percent is considered acceptable. Recovery less than this limit requires corrective action.

## (c) Representativeness

Representativeness refers to how well the results represent the sample and how well the samples represent the population. Representativeness can be assessed by examining the variability among samples. For example, to determine whether individual composite samples are sufficient to develop a weekly composite, the required number of samples could be calculated. Methods for calculating the number of samples are presented in chapter 9 and repeated here.

### (1) Calculation

Compute the required number of samples as follows:

$$n > \frac{t^2 S^2}{d^2} \qquad [11\text{–}3]$$

where:

$n$ = number of samples
$t$ = students 't' at a given confidence level
$S$ = sample standard deviation
$d$ = acceptable difference from the mean

## (d) Comparability

Certain data from the study can be compared to results obtained from other similar studies.

## (e) Completeness

Completeness can be measured as the percentage of total samples collected that were analyzed. Sufficient water volumes should be collected to allow re-analysis of a sample if beyond a standard curve or if lost in a laboratory accident. A measure of completeness is the percentage of missing data obtained in the study. The number of samples needed is governed by the study design.

## 614.1108 Sample custody procedures

Each sample should be dated and coded according to site, sample type, station number, and sample sequence. The actual sample containers should be labelled with a sample number for identification.

Transfer of sample custody takes place upon delivery of samples to the laboratory. At the time of delivery, the person delivering the samples signs over custody to a laboratory person receiving the samples. This transaction is recorded on forms for that purpose, and the records are maintained in the laboratory (fig. 11–2).

As part of the process of sample receipt, each sample is assigned a unique identification number that can include specific information on location, date, composite, and yearly sequence. For example, a sample numbered 10-011891-24-566 represents a sample taken at station 10, on January 18, 1991, a 24-hour composite, and is the 566th sample received by the laboratory in a calendar year. This final number, representing the sample received in a year, serves as the shorthand sample number and is used for overall tracking in the laboratory.

The sample number should be used in all laboratory books to identify the sample. Sample transfer forms may be needed for some studies where samples are sent to other labs. Some agencies employ the practice of prelabeling bottles before they go to the field.

## 614.1109 Calibration procedures and preventative maintenance

The primary pieces of laboratory equipment should be described in a quality assurance plan together with the calibration and maintenance procedures and schedules. Standard curves, using from 8 to 10 standards including blanks, should be developed the same day of analysis for most analyses. Each analytical run should include a set of standards.

The maintenance schedule should be included in a quality assurance plan. The options available if equipment breakdown occurs should be described.

**Figure 11–2** Laboratory chain of custody sheet

Custody sheet for samples collected on (date)_____
Relinquished by_____
Received by_____

Samples held until:
(+ 28 days)

| Lab No. | Sample Description | Procedure completed (indicate with date) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Acid | Filter | Digest | TKN | NH3 | NO3 | TP | CL | TSS |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |

Remarks:

# 614.1110 Performance and systems audits

The project should be subject to both performance audits and systems audits. The performance audit could consist of unknown samples submitted quarterly to the laboratory.

## (a) Calculation

Reported results are compared to known values. The percentage recovery for the known is calculated as:

$$\% \text{ Recovery} = \frac{R - T}{T} \times 100 \qquad [11–4]$$

where
$R$ = reported value
$T$ = true value

Performance within ± 25 percent should be acceptable. Performance beyond ± 25 is considered an out-of-control situation calling for corrective action.

Project supervisors should make unscheduled performance audits of all laboratory personnel to detect any deviations from standard operating procedures. A checklist of the audit should remain on file in the supervisor's office.

A systems audit consists of an onsite review of the entire project.

# 614.1111 Corrective action

Data quality assurance procedures should be designed to ensure that project personnel are able to quickly identify and correct analytical problems. Data failing to meet quality control requirements should be subject to repeated analysis where sufficient volume exists to retest the sample.

# 614.1112 Field quality assurance

## (a) Field equipment

Calibration of field equipment is necessary. In situ analysis of temperature, pH, dissolved oxygen, conductivity, and other ions use field instruments requiring maintenance and calibration. Some instruments, such as pH and dissolved oxygen meters, require daily or more frequent calibration. A record should be maintained of all calibrations.

Stage recorders should be calibrated against a permanent outside staff gage at every visit. The staff gages should be surveyed to a benchmark at least annually. Precipitation gages should be calibrated annually, and checked weekly. Well pressure transducers should be calibrated when they do not equal staff gage readings. Well top elevations should be surveyed annually to a temporary benchmark. Stage-discharge relationships should be constructed during the first year of the project by at least 15 discharge measurements using the velocity-area method. Annually, the stage-discharge relationship should be checked with at least five ratings. Annual runoff coefficients should be calculated as the percentage of precipitation that left the watershed as discharge. These coefficients could be compared to runoff coefficients calculated from U.S. Geological Survey water resources data collected from other watersheds in the same general area of the state.

## (b) Field logs

Daily field logs should be kept for each field visit. These logs record operating status, calibration checks, manual readings, and the name of the field visitor. They are often 1-page sheets (fig. 11–3) and are tailored to the individual project. A personal notebook (survey book) maintained by each field worker may be useful. Each field visit is recorded and additional notes are made on work to be done.

**Figure 11–3** Example daily field log

Checked _____
Date _____

**Daily Field Log**

| | | Station 1 | Station 2 | Station 3 | Comments |
|---|---|---|---|---|---|
| **General** | Time of visit | | | | |
| | Weir clear/chop | | | | |
| | Solar panels Ok? | | | | |
| | Batterries Ok? | | | | |
| | | | | | |
| | | | | | |
| **Sampler** | Sample volume Ok? | | | | |
| | Intake line Ok? | | | | |
| | Dessicant Ok?/replace | | | | |
| | Line in bottle? | | | | |
| | Sampler on? | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| **Stage** | Recorder stage (ft) | | | | |
| | Staff stage (ft) | | | | |
| | Point gage (ft) | | | | |
| | Display | | | | |
| | Enough paper? | | | | |
| | | | | | |
| | | | | | |

## (c) Field quality control samples

The four types of samples needed to assess field quality control include (Burger 1987):

- **Field duplicate—**Samples collected simultaneously at a location used to determine the variability associated with sample collection.
- **Trip blank**—Sample container taken to field and filled with distilled or deionized water and returned. This sample assesses contamination during transport or storage.
- **Sampler blank—**Sample obtained by passing deionized water through a nondedicated sampler, such as a portable pump. This blank is used to test contamination by a sampler.
- **Filtration blank—**Sample collected by field filtering apparatus using deionized water. This blank tests contamination by a filter and apparatus.

## (d) Field chain of custody

The sample custody procedures actually begin in the field. Proper labeling of sample bottles is critical. Some laboratories use pre-numbered bottle labels (Burger 1987).

# 614.1113 References

American Public Health Association. 1989. Standard methods for the examination of water and wastewater. 17th Ed., American Public Health Assoc., Washington, DC.

Burger, Robert M. 1987. Design and implementation of a successful quality control program for ground water sampling events. Symp. on monitoring, modeling, and mediating water quality, pp. 49-58, Amer. Water Resour. Assoc.

Dillaha, T.A., S. Mostaghimi, C.D. Eddleton, and P.W. McClellan. 1988. Quality control/quality assurance for water quality monitoring. Amer. Soc. Agric. Eng. Pap. No. 88-2648, St. Joseph, MI.

Driscoll, F.G. 1986. Ground water and wells. 2nd. Ed. Johnson Div., St. Paul, MN.

Klemm, D.J., P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. Macroinvertabrate field and laboratory methods for evaluating the biological integrity of surface waters. Office of Research and Development, Washington, DC, U.S. Environmental Protection Agency/600/4-90/030.

Kunkle, S.H., and T.E. Ricketts. 1984. Portable kits for water chemistry reconnaissance in the field. WRFSL Rep No. 84-2, Water Resour. Field Sup. Lab., Natl. Park Serv., Fort Collins, CO.

Lawrence J., and A.S.Y. Chau. 1987. Quality assurance for environmental monitoring. Symp. on monitoring, modeling, and mediating water quality, pp. 165-175. Amer. Water Resour. Assoc.

Shelley, P.E. 1977. Sampling of water and wastewater. US EPA-600/4-77-039. Off. Res. and Dev., Cincinnati, OH.

Simpson, J.T. 1991. Volunteer lake monitoring: a methods manual. US EPA 440/4-91-002. Off. Water. Washington, DC.

Taylor. J.E. 1984. Essential features of a laboratory quality assurance program. *In* Statistics in the environmental sciences, S.M. Gertz and M.D. London (Eds.). Amer. Soc. for Test. and Mat. pp. 66-73.

Topp, E., and W. Smith. 1992. Sorption of the herbicide Atrazine and Metolachlor to selected plastics and silicone rubber. J. Environ. Quality 21:316-317.

United States Environmental Protection Agency. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. EPA 670/4-73-001, Off. Res. and Dev., Cincinnati, OH.

United States Environmental Protection Agency. 1983. Methods for chemical analysis of water and wastes. EPA600/4-79-020, Off. Res. and Dev., Cincinnati, OH.

United States Environmental Protection Agency. 1988. Guide for the preparation of quality assurance project plans for the national estuary program. EPA 556/2-88-001, Off. Marine and Estuarine Prot., Washington, DC.

United States Environmental Protection Agency. 1990. Volunteer water monitoring: a guide for state managers. EPA 440/4-90-010. Off. Water. Washington, DC.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 12 Land Use and Management Monitoring

(450-VI-NWQH, September 2003)

# Chapter 12 Land Use and Management Monitoring

Contents:

# Chapter 12 Land Use and Management Monitoring

## 614.1200 Introduction

An essential element of water quality monitoring is the tracking of land use and management activities in the watershed being monitored. Land use and management data are needed to explain any water quality changes that may occur. The water quality changes must be attributed to the management practice and not to other confounding influences, such as climate or a point source. For watershed scale monitoring, the proximity of the land practices to the monitoring location can directly influence the water quality observed. A poor practice near the watershed outlet or downgradient can mask the influence of good practices upstream or upgradient.

This chapter presents methods for monitoring and managing land use and management data and provides checklists of recommended activities to monitor for the major sources of the nonpoint pollutant.

## 614.1201 Methods of monitoring

The four basic approaches for monitoring land treatment data are personal observations, field logs, personal interviews, and remote sensing. Any one project may use some or all of these approaches to track activities on the land, depending on the scale and complexity of the project.

Land treatment data can be either static or dynamic, point or diffuse. Static land treatment data do not change with time. Examples of this type data include soil type and slope. Dynamic land treatment data can vary with time and include the number of animals, cover crop, nutrient applications, and irrigation schedules. Most land treatment activities are considered diffuse or nonpoint. However, some activities, such as feedlots, manure stacks, and silage bunkers, can be viewed as potential point sources from a watershed scale perspective.

### (a) Personal observations

For small scale projects, such as plots or individual fields, tracking may best be accomplished by project personnel using personal observations. Routine site visits can include an analysis of the site conditions at the time of the visit. The type of information that can be collected through personal observations includes counts, timing of certain activities, site characteristics, and tests. Some examples are:

**Counts**
- Number of animals
- Crop type

**Site characteristics**
- Slope
- Slope length
- Soil type

**Timing**
- Planting date
- Harvest date
- Tillage dates
- Fertilizer applications
- Pesticide applications
- Irrigation schedules

**Tests**
- Yield test
- Soil test
- Application rates

A form for recording personal observations is highly recommended. It should include required check-offs to assure certain questions are not overlooked.

The windshield survey is another type of personal observation. This survey is useful in identifying land uses for areas where ownership is unknown and information is difficult to collect from traditional methods.

**(1) Advantages**
A major advantage of the personal observation is that the quality of the data is controlled by the observer. This means that the timing of the visit can be scheduled as well. Personal observation-type data are relatively inexpensive to obtain.

**(2) Disadvantages**
Timing is critical to certain types of land use observations. For example, pesticide applications occur on a short time frame and will most likely be missed by less frequent than daily visits. Also, the amount of an application, such as nutrient loading, can only be determined by being present during the application.

The potential for "judgment bias" in personal observations is great. Different individuals will most likely make different observations. Bias also can be introduced by personal schedules. Quantitative and randomized observations may help to reduce bias. Generally, a reliance on personal observation alone results in an incomplete data set of land treatment activities.

## (b) Field logs

The term *field log* is meant to include the various forms that would be left with the landowner or manager. The manager ideally would keep a record of activities. A copy of a manure/fertilizer log used in the St. Albans Bay RCWP is shown in figure 12–1. This particular log was given to each cooperating and noncooperating farm producer in the watershed. The log was placed inside a checkbook cover with a farm map showing numbered fields. The field logs were recovered twice yearly.

Some states require that the producer maintain a field log as part of a permit condition.

**(1) Advantages**
The major advantage of the field log is that the person performing the activities is keeping the records. This person is often the only one who knows when certain activities occur and how much occurred. Picking up a field log allows for additional interaction with the producer.

**(2) Disadvantages.**
A 100 percent compliance in good record keeping in the watershed is unlikely. Some producers will not fill out the log. Others will not complete the log with the level of detail or precision needed. For example, instead of indicating the exact date of a manure application on field No. 10, a producer may indicate “early spring.”

## (c) Personal interviews

A personal interview or one-on-one contact is an effective way to obtain land treatment data. A direct visit is preferred over a telephone interview. A form is recommended as a guide for the interview. Based on experience obtained in the St. Albans Bay RCWP, two visits per year yields much more reliable data than an annual visit. Meetings with producers were timed with less busy periods on the farm (e.g., mid-summer and mid-winter).

**(1) Advantages**
The major advantage of the personal interview is that the data is obtained from the person responsible for the land activity. Also, the interview facilitates obtaining information on subtle land use changes, such as rental lands, field boundary changes, and shifts in animal numbers.

**(2) Disadvantages**
A major disadvantage of the personal interview is that the quality of data obtained varies with both the interviewer and the interviewee. Some people are adept at questioning farm producers, while others are not. Similarly, some farm producers are reluctant to share management information. Another disadvantage is that the personal interview relies on "reconstructed" data based on the memories of the person interviewed.

**Figure 12–1** Example of a field log

### A. Manure application

| Date | Field ID (see map) | Amount applied (full spreader load) | Date incorporated | Time (approx.) | Comments |
|---|---|---|---|---|---|
| Example 4/23/82 | 3b | 1 1/2 | 4/23 | 10:30 am | • Evenly spread except wet spot on NE corner<br>• Planted corn 4/28 |
| | | | | | |
| | | | | | |
| | | | | | |

### B. Commercial fertilizer application (including lime)

| Date | Field ID | Formulation | Amount applied/ac | How applied | Comments |
|---|---|---|---|---|---|
| Example 4/23/82 | 21 (or all corn fields) | 10-20-10 | 4 lb/ac | broadcast | disced on 4/23 |
| | | | | | |
| | | | | | |
| | | | | | |

## (d) Remote sensing

For certain types of land use and treatment data, remote sensing techniques may serve as a primary data source or verification of other data sources. For example, the 35mm slides of cropland areas taken annually by FSA can provide a source of land cover information on a field basis. Satellite data would generally not be sufficient for monitoring land treatment, although it has been used to assess critical areas (Sivertun, et al. 1988).

### (1) Advantages

Remotely sensed data can give a permanent visual and spatial record of certain types of land use data, including land cover. Certain types of critical sources of nonpoint pollution, such as erosion, may be observable using remote sensing. Data that can be obtained by remote sensing eliminate reliance on the memories of individuals.

### (2) Disadvantages

Remotely sensed data have limited applications. Low level air photos can be used to distinguish some crop covers, but it is difficult to distinguish between others, such as forest and residential. Remotely sensed data will not provide timing information, such as manure or fertilizer applications.

# 614.1202 Management of land treatment data

The method employed to keep track of land use data varies with the situation, but the method used must be defined at the beginning of the project. Without attention to management of land treatment data, records will most likely be insufficient and full of gaps. The three methods for management of land treatment and land use data are ad hoc files, spread sheets/data bases, and geographic information systems.

## (a) Ad hoc files

A good filing system can be effectively used to track land use and treatment data. It is important that the results of land treatment monitoring be reported routinely and often. Failure to do so will result in data gaps remaining hidden, possibly until the end of the project when it will be too late to recover the data. Spatial data from ad hoc files should be transferred to and displayed on maps as a quality control check on how much information is actually being obtained.

## (b) Spreadsheets/data bases

Various computer spreadsheet and data base programs can be used to track land treatment data. Such programs are particularly efficient in attaching attributes to field IDs. The EPA has developed a PC software program, the Nonpoint Source Management System (NPSMS), to track management activities and water quality and implementation data (US EPA 1991). NPSMS actually has several separate files for tracking information. The *management* file stores information about the water quality problem and project goals. The *monitoring plan* file holds descriptions of the monitoring design, including stations, variables, and frequencies. The *annual report* file includes the annual water quality and implementation data. The system also includes the water body system for identifying the individual body of water involved.

Data bases, in particular, allow relating data between different files, such as land treatment files and water quality files.

## (c) Geographic information system (GIS)

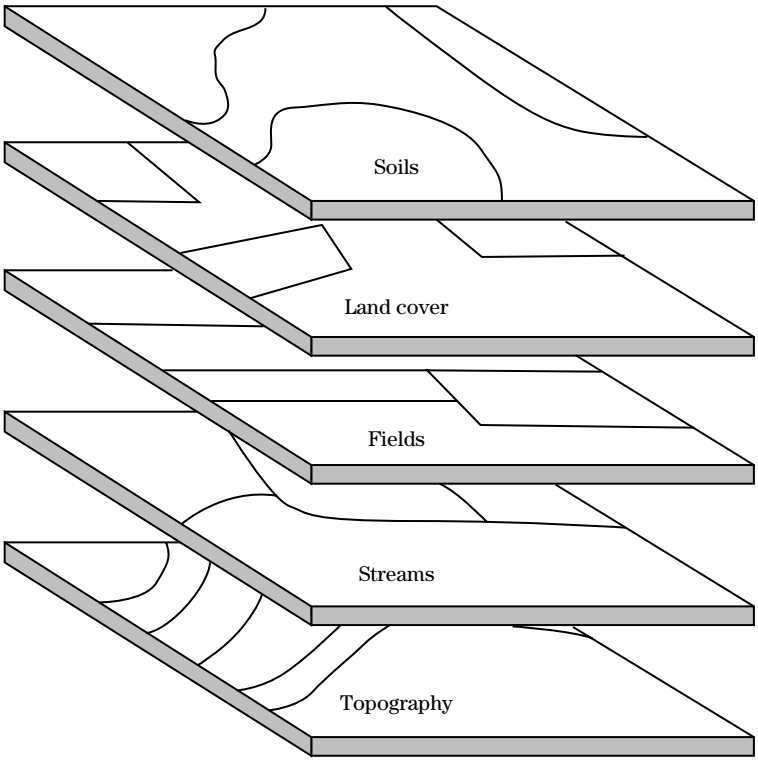
Geographic information systems are "…systems that integrate layers of spatially oriented information, whether manually or automatically…" (Walsh 1985). A GIS is ideally suited to track land use and treatment data. The primary advantage is that land treatment data can be displayed spatially and combined with other water quality related information.

GIS data can be stored as values for uniform grids (raster) or as strings of coordinates representing points, lines, and areas, including polygons (vector).

Land treatment data, such as land cover, can be overlaid on stream courses, soil types, and topography (fig. 12–2). A GIS also allows displaying and calculating new information from the combined data layers, such as where and how much animal waste was applied within 50 feet of a stream or where and how much animal waste was applied on soil hydrologic group D.

Because all the files in a GIS are relational, that is, two-dimensional tables can be related to each other based on a common characteristic, such as field ID, a GIS also serves as a data base for managing and reporting land treatment data.

**Figure 12–2** Geographic information systems data layers



### (1) Data entry

The most difficult aspect of using a GIS for managing land treatment data is the initial digitizing of the spatial data layers. Quality control is an important consideration in GIS data entry, just as it is for water quality analysis. Digitized information should go through an error checking system to make sure that the layer has been appropriately geo-referenced and lines and points are properly located. Just closing polygons is insufficient quality control. Other information added should also receive error checking (see chapter 13).

Once the data layers have been entered, attributes are easily added and data management is enhanced and powerful. Although the appropriate data layers would vary with each situation, several useful data layers are given in table 12–1 along with suggested priorities for most water quality monitoring situations.

Farm and field boundaries are almost essential as a data layer. Such data can be obtained from the farm plan photos with verification from the farm operator.

**Table 12–1** Frequently used data layers for a GIS

| Priority | Data layer |
|---|---|
| 1 | quadrangle basemap |
| 1 | farm and field boundaries |
| 1 | stream courses and other water bodies (or proximity class) |
| 1 | watershed boundary |
| 1 | soil series (or attribute of field) |
| 2 | topography or slope (or attribute of soil) |
| 2 | land cover/land use |
| 3 | transportation |
| 3 | geology |
| 3 | political boundaries |
| 4 | archeology |
| 4 | precipitation (where variable) |

Stream courses can be digitized as lines or bands, polygons, or grids, or a proximity zone to the watercourse can be used. For example, Sivertun, et al. (1988) used proximity bands of 0 to 150, 150 to 650, 650 to 3,300, and >3,300 feet to help identify critical areas in a watershed.

Soils data could be entered as the soil series or as some more general textural class either as a separate layer or as a field attribute. However, a separate soils layer is recommended. Topography could be entered as a data layer, either as points, polygons, or grid information, or the percent slope could be entered either as an attribute of the field boundary or the soil series. Topographic information is not necessary to track land use data, but is useful for displaying results in a 3-D format and identifying critical areas.

Land cover could be entered as a separate data layer; however, it is best entered as an attribute of a farm field because it is easily updated. Good land cover/land use maps are not readily available. Therefore, these maps are often developed from aerial photo interpretations, satellite imagery, or on-the-ground observations.

For the St. Albans Bay RCWP, a land use/land cover data layer was created from individual farm 9 by 9 1:660 scale farm plan photos, verifications from the farm operator, supplemental ASCS 35mm slides, and ground truthing of gaps in the data layer.

The use of satellite results is not accurate enough at this time to determine land use/land cover for water quality monitoring purposes. However, satellite data may be very useful when determining critical areas of high pollution potential (Guilliland and Baxter-Potter 1987).

Precipitation is an appropriate data layer when highly variable across the watershed in some cases. Irrigation networks may be useful in certain areas (Walsh 1985).

For ground water projects, information on ground water withdrawals and piezometric surfaces may be important management information.

**(2) Analysis**

After the data layers have become part of the GIS data base, attributes of dynamic data layers can be updated. For example, cover crop can be changed annually. The additions of nutrients, either as animal waste or fertilizers, can be updated on a weekly, monthly, or annual basis. From this data base, several types of land use and management information can be generated (table 12–2).

**Table 12–2** Land use and management data generated from a GIS

| | Units |
|---|---|
| **Land treatment data** | |
| Critical area under BMP | %, ac. |
| Animal units under BMP | %, No., No/ac |
| Fields under nutrient management | %, ac. |
| Fields under irrigation management | %, ac. |
| Area of land use (pasture, etc.) | %, ac. |
| Erosion control | %, ac. |
| **Animal waste data** | |
| Manure from storage | % |
| Manure incorporated | % |
| Barnyard management | No. |
| Milkhouse management | No. |

# 614.1203 Relationship between land use/treatment and water quality

The purpose in collecting land use and management information is to use that data to establish causal relationships with water quality. Causality involves several steps:

1. An association should exist between the water quality and land treatment data.
2. This association should be consistent across different data sets so that a general statement may be made about the relationship.
3. The association should be tested to make sure that one variable is responsive to the other variable. This responsiveness may require experimentation.
4. There must be a mechanism that logically explains the process that results in the relationship.

This section will focus on developing associations between land treatment and water quality data.

When developing a program for monitoring land treatment data for the purpose of relating that data to water quality, both temporal and spatial scales must be decided.

Water quality data are often collected at a much more frequent rate than land treatment data. For example, in the St. Albans Bay RCWP, water quality samples were collected every 8 hours, but land treatment information was collected twice a year. In one analysis associations were made of weekly phosphorus and manure application data (Hopkins & Clausen 1985). However, the danger in such associations is that they are confounded by the timing of agricultural practices. For example, animal waste is not applied to agricultural lands during wet seasons, but nutrient concentrations in streams are highest during the same wet periods. Thus a confounded association of manure applications and stream concentrations could exist. To resolve this problem, Meals (1992) used annual data for the associations.

The spatial scale of land treatment data also is important. Watershed-wide summaries were most useful in establishing land treatment-water quality relationships in Vermont (Meals 1992). However, an association of land use (corn, pasture, hay) and certain water quality variables for data summarized were within 150 feet of the streams for each watershed. Schlagel (1992) also pointed out that the spatial pattern within watersheds of changes in land treatment practices is important and could mask water quality changes.

The primary methods for establishing associations are described in part 615 of this handbook. Correlations serve as an initial tool.

When developing the monitoring plan, a list of land use and management data that will be used to relate to water quality data also should be developed. This list will obviously vary with the project.

## 614.1204 References

Gilliland, M.W., and W. Baxter-Potter. 1987. A geographic information system to predict non-point source pollution potential. Water Resourc. Bull. 23(2):281-291.

Hopkins, R.B., Jr., and J.C. Clausen. 1985. Land use monitoring and assessment for nonpoint source pollution control *In* Perspectives on nonpoint source pollution, EPA 440/5-85-001, Washington, DC, pp. 25-29.

Meals, D.W. 1992. Relating land use and water quality in the St. Albans Bay watershed, Vermont. Proc. Natl. RCWP Symp., EPA pp. 131-143.

Schlagel, J.D. 1992. Spatial and temporal change in animal waste application in the Jewett Brook watershed, Vermont: 1983-1990. Proc. Natl. RCWP Symp., EPA pp. 145-150.

Silvertun, A., L.E. Reinelt, and R. Castensson. 1988. A GIS method to aid in non-point source critical area analysis. J. Geogr. Inf. Sys. 2(4):365-378.

United States Environmental Protection Agency. 1991. Watershed monitoring and reporting for section 319 national monitoring program projects. Office of Water Washington, DC.

Walsh, S.J. 1985. Geographic information systems for natural resource management. J. Soil and Water Conserv. pp. 202-205.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 614
National Water Quality Handbook

# Chapter 13 Data Management

(450–VI–NWQH, September 2003)

# Chapter 13 Data Management

Contents:



Figures

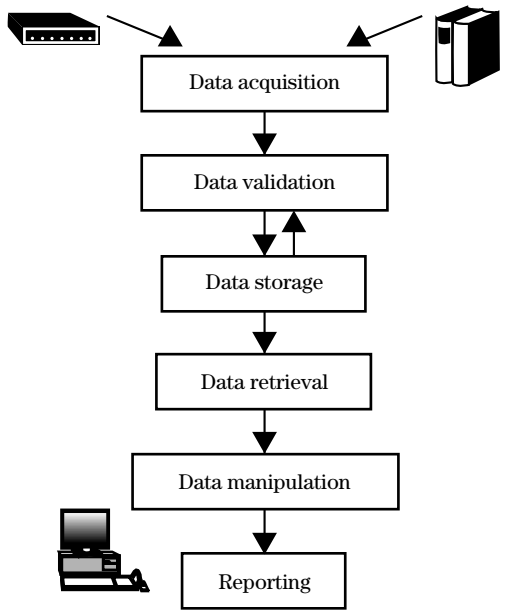
# Chapter 13 Data Management

## 614.1300 Introduction

Data management in water quality monitoring projects refers to a series of steps for handling data (fig. 13–1). The management of data has become increasingly important because efficient means are needed to deal with a large amount of numbers and the integrity of those numbers must be guaranteed. The processes in a data management system include acquisition, storage, validation, retrieval, manipulation, and reporting of data (Canter 1985; Sanders, et al. 1983; Ward, et al. 1990). The interpretation of data will be further described in part 615 of this handbook.

Advances in computers and software have made the process of data management much easier. Therefore computer applications will be described in this chapter.

**Figure 13–1** Data management system




## 614.1301 Data acquisition

The acquiring of data is meant to include its collection and entry into the data management system. Entry may begin indirectly from data entry sheets (fig. 13–2), which could be completed by either field or laboratory personnel. More direct entry of data has been made possible via the use of data loggers. This latter process bypasses the steps of manually entering data and therefore avoids transcription errors. Data from a data logger can be input using storage modules, cassette tapes, or telecommunication devices that have an interface with a computer system.

**Figure 13–2** Example data entry sheets

Data entry sheet
riparian zone restoration project

**Streams**

| STA | Date MM/DD/YY | Hours | Lab No. | Concentration (mg/l) | | | |
|---|---|---|---|---|---|---|---|
| | | | | TKN | $NH_3$ | $NO_2/NO$ | TP |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |

## 614.1302 Data storage

The storage of data should be viewed as a multilevel effort using manual and computerized technologies. Manual efforts should include safe storage of original laboratory notebooks, field notebooks, daily field logs, and any paper tapes and strip charts. A manual copy of all computerized data files should be printed on high quality paper and placed in safe storage. Smoke destroys a floppy disk, but not paper.

Laboratory notebooks should be considered a permanent record of data. The notebooks should be bound with numbered pages so that pages cannot be substituted or deleted. Pages should be dated and signed by operators. Entries should be made in ink. Errors should be crossed out so that they are legible, but not erased. The correction should be initialed and dated. Large blank spaces in the notebooks should have lines drawn through them. Standard curves should be drafted within the lab notebook.

Computerized data storage also is highly recommended. In the past, computerized data management systems were developed specifically for individual projects onsite, and could not be transferred to other locations. The availability of general spreadsheet software, such as Lotus 1-2-3, Quattro Pro, or Excel, has greatly changed the need to develop individual data management systems. In addition, data base management software is available. The following are recommendations for computerized spreadsheets and data base management systems use:

- Store data in ASCII format, preferably formatted in columns.
- Store data on floppy disks, not hard drives.
- Backup disks are essential; maintain one set onsite and one set offsite (at home).
- Store data in files of "convenient" blocks of data, such as annually. One disk could represent 1 year of data.
- Plan file naming conventions. A file name could include such information as project or study area, data type, data manipulations, and project year. For example, the file "SAQ23.S85" refers to the St. Albans Bay RCWP project (SA), flow data (Q), for the Level 2 tributary stations (2), for the third quarter of the year (3), sorted by station number and date (S), and for project year 1985. For this study separate formatted ASCII files were created for flow (Q) files, concentration data (C), mass data (M), stage data (S), and precipitation data (P) using the same file naming convention. Because knowing that the data files have been error checked is important, checking was done quarterly. However, many spreadsheets use their own filename extensions, such as XXXXXXXX.WQ1 for Quattro.
- Decide how to record missing data in the computer files. A -9.0 could be a code for missing data in cases where negative data does not exist (e.g., concentration, flow). The statistical package SAS uses a single period, '.' as an indicator of missing data.

Geo-referencing the location of water quality sampling stations by latitude and longitude (degrees, minutes, seconds) is further recommended. Such referencing is required by some data storage systems, such as STORET.

Data that are below detection limits are termed *censored* data. Data should be entered in the data management system that codes the data as below the detection limit. For example, a -8.0 could be used where negative data is not possible. The elimination of data below detection limits or the entry of the below detection limit data as either a 0, half the limit, or the limit itself is not recommended (Newman, et al. 1989).

## 614.1303 Data validation

All data reported should receive a 100 percent error check. Transcription errors can be checked by entering the data twice, preferably by two individuals. A computer program can compare the two data files and flag any inconsistencies for correction.

Also, the **COMP** command in DOS allows the comparison of the contents of two files in either the same or different directories. If the **COMP** command finds any mismatches, an error statement will be displayed.

Laboratory notebook calculations should be checked by a supervisor, who initials the notebooks as verified. Sample custody sheets should be reviewed to ensure that holding times, preservation, sample integrity, and equipment calibration requirements have been met.

Additional tests of reason can be applied to concentration values. For example, ammonia concentrations cannot exceed total Kjeldahl nitrogen values, and ortho-phosphorus cannot exceed total phosphorus values. Also, limits can be used as flags in the data set. For example, appropriate limits for pH are 0 to 14. A maximum limit for total phosphorus might be 5 mg/L for a lake. Standard laboratory curves should be analyzed for warning and control limits as described in Standard Methods (APHA 1989).

Data not meeting the requirements described above could be rejected and noted in the data files as missing data.

## 614.1304 Data retrieval

The retrieval of data from the data management system must consider the form of retrieval (paper report, data file, graph) as well as the intended use (statistical, quality control, share with others). Good records must be maintained on format for data storage so that others can review the data files. Readme.txt files stored on disks containing the data files are highly recommended.

## 614.1305 Data manipulation

Data generally require some form of manipulation before being reported. Common manipulations include:

- calculations of average values or mass exports
- sorting
- graphical presentations
- statistical analysis/ transformations

Common spreadsheet and data base programs facilitate the calculation of averages and mass exports. For example, Quattro Pro and Lotus allow entering a formula, i.e., equation, to apply to stored data or the use of functions (internal formulas) to apply to the data. These functions include mathematical, statistical, and logical operations.

The sorting of data is a common manipulation in a data management system. Frequently, data must be arranged by date or station number to report the results, input to a graph, or perform statistical analysis. Most spreadsheets have sorting commands. It may be desirable to search through the data system as well as sort the data.

Graphical presentations also are facilitated by spreadsheets, or a number of graphics packages are available.

Statistical manipulation of data will be very specific to the study design. However, most data receive routine univariate analysis, including the number of samples, mean, maximum, minimum, and standard deviation. These simple statistics can be determined in most spreadsheets. More sophisticated statistical analysis may require the use of other statistical packages.

If censored (below detection limits) data are in the data set, the mean and standard deviation for the data are strongly influenced by the manner in which the censored data is handled and the percentage of data that is censored. This is described further in part 615.

## 614.1306 Data reporting

Reporting data at the end of a monitoring study may seem obvious, but reporting during the progress of the study is very important for several reasons. Interim reporting encourages (requires) identifying data errors and data gaps. Frequent reporting aids in solving problems. Although it seems like it takes too much time, reporting should be at a minimum of quarterly either formally or informally. Progress reports should include data that have been screened, analyzed statistically, summarized and plotted. A few copies of the raw data should be made available to project sponsors and cooperators. The data could be shared as ASCII files on diskettes.

Guidelines for preparing reports are beyond the scope of this handbook. However, following the guidelines of an appropriate professional journal, especially regarding tables and figures, is recommended.

## 614.1307 References

American Public Health Association. 1989. Standard methods for the examination of water and wastewater. 17th ed., American Public Health Assoc., Washington, DC.

Canter, L.W. 1985. River water quality monitoring. Lewis Pub., Inc. Chelsea, MI.

Newman, M.C., P.M. Dixon, B.B. Looney, and J.E. Pinder, III. 1989. Estimating mean and variance for environmental samples with below detection limit observations. Water Resourc. Bull. 25(4):905-916.

Sanders, T.G., R.C. Ward, J.C. Loftis, T.D. Steele, D.D. Adrian, and V. Yevjevich. 1983. Design of networks for monitoring water quality. Water Resourc. Pub., Littleton, CO.

Ward, R.C., J.C. Loftis, and G.B. McBride. 1990. Design of water quality monitoring systems. Van Nostrand Reinhold. New York, NY.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

# Part 614
# National Water Quality Handbook

# Appendixes

(450–VI–NWQH, September 2003)

# Appendixes

Contents:

## Appendix A Distribution of Z [1]

Probability of a random value of $Z = (X - \mu)/s$ being greater than the values tabulated in the margins

| Z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| .0 | .5000 | .4960 | .4920 | .4880 | .4840 | .4801 | .4761 | .4721 | .4681 | .4641 |
| .1 | .4602 | .4562 | .4522 | .4483 | .4443 | .4404 | .4364 | .4325 | .4286 | .4247 |
| .2 | .4207 | .4168 | .4129 | .4090 | .4052 | .4013 | .3974 | .3936 | .3897 | .3859 |
| .3 | .3821 | .3783 | .3745 | .3707 | .3669 | .3632 | .3594 | .3557 | .3520 | .3483 |
| .4 | .3446 | .3409 | .3372 | .3336 | .3300 | .3264 | .3228 | .3192 | .3156 | .3121 |
| .5 | .3085 | .3050 | .3015 | .2981 | .2946 | .2912 | .2877 | .2843 | .2810 | .2776 |
| .6 | .2743 | .2709 | .2676 | .2643 | .2611 | .2578 | .2546 | .2514 | .2483 | .2451 |
| .7 | .2420 | .2389 | .2358 | .2327 | .2296 | .2266 | .2236 | .2206 | .2177 | .2148 |
| .8 | .2119 | .2090 | .2061 | .2033 | .2005 | .1977 | .1949 | .1922 | .1894 | .1867 |
| .9 | .1841 | .1814 | .1788 | .1762 | .1736 | .1711 | .1685 | .1660 | .1635 | .1611 |
| 1.0 | .1587 | .1562 | .1539 | .1515 | .1492 | .1469 | .1446 | .1423 | .1401 | .1379 |
| 1.1 | .1357 | .1335 | .1314 | .1292 | .1271 | .1251 | .1230 | .1210 | .1190 | .1170 |
| 1.2 | .1151 | .1131 | .1112 | .1093 | .1075 | .1056 | .1038 | .1020 | .1003 | .0985 |
| 1.3 | .0968 | .0951 | .0934 | .0918 | .0901 | .0885 | .0869 | .0853 | .0838 | .0823 |
| 1.4 | .0808 | .0793 | .0778 | .0764 | .0749 | .0735 | .0721 | .0708 | .0694 | .0681 |
| 1.5 | .0668 | .0655 | .0643 | .0630 | .0618 | .0606 | .0594 | .0582 | .0571 | .0559 |
| 1.6 | .0548 | .0537 | .0526 | .0516 | .0505 | .0495 | .0485 | .0475 | .0465 | .0455 |
| 1.7 | .0446 | .0436 | .0427 | .0418 | .0409 | .0~01 | .0392 | .0384 | .0375 | .0367 |
| 1.8 | .0359 | .0351 | .0344 | .0336 | .0329 | .0322 | .0314 | .0307 | .0301 | .0294 |
| 1.9 | .0287 | .0281 | .0274 | .0268 | .0262 | .0256 | .025n | .0244 | .0239 | .0233 |
| 2.0 | .0228 | .0222 | .0217 | .0212 | .0207 | .0202 | .0197 | .0192 | .0188 | .0183 |
| 2.1 | .0179 | .0174 | .0170 | .0166 | .0162 | .0158 | .0154 | .0150 | .0146 | .0143 |
| 2.2 | .0139 | .0136 | .0132 | .0129 | .0125 | .0122 | .0119 | .0116 | .0113 | .0110 |
| 2.3 | .0107 | .0104 | .0102 | .0099 | .0096 | .0094 | .0091 | .0089 | .0087 | .0084 |
| 2.4 | .0082 | .0080 | .0078 | .0075 | .0073 | .0071 | .0069 | .0068 | .0066 | .0064 |
| 2.5 | .0062 | .0060 | .0059 | .0057 | .0055 | .0054 | .0052 | .0051 | .0049 | .0048 |
| 2.6 | .0047 | .0045 | .0044 | .0043 | .0041 | .0040 | .0039 | .0038 | .0037 | .0036 |
| 2.7 | .0035 | .0034 | .0033 | .0032 | .0031 | .0030 | .0029 | .0028 | .0027 | .0026 |
| 2.8 | .0026 | .0025 | .0024 | .0023 | .0023 | .0022 | .0021 | .0021 | .0020 | .0019 |
| 2.9 | .0019 | .0018 | .0018 | .0017 | .0016 | .0016 | .0015 | .0015 | .0014 | .0014 |
| 3.0 | .0013 | .0013 | .0013 | .0012 | .0012 | .0011 | .0011 | .0011 | .0010 | .0010 |
| 3.1 | .0010 | .0009 | .0009 | .0009 | .0008 | .0008 | .0008 | .0008 | .0007 | .0007 |
| 3.2 | .0007 | .0007 | .0006 | .0006 | .0006 | .0006 | .0006 | .0005 | .0005 | .0005 |
| 3.3 | .0005 | .0005 | .0005 | .0004 | .0004 | .0004 | .0004 | .0004 | .0004 | .0003 |
| 3.4 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0003 | .0002 |
| 3.6 | .0002 | .0002 | .0001 | .0001 | .0001 | .0001 | .0001 | .0001 | .0001 | .0001 |
| 3.9 | .0000 | | | | | | | | | |

1/ Steel, R.G.D., and J.H. Torrie. 1960. Prnciples and procedures of statistics. McGraw-Hill, Inc., New York, NY. (Reproduced with permission of the McCraw-Hill Companies.)

**Appendix B** Distribution of *t* (two-tailed) [1]

| Degrees of Freedom | Probability of a Larger Value, Sign Ignored | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.500 | 0.400 | 0.20 | 0.10 | 0.050 | 0.025 | 0.010 | 0.005 | 0.001 |
| 1 | 1.000 | 1.376 | 3.078 | 6.314 | 12.706 | 25.452 | 63.657 | | |
| 2 | 0.816 | 1.061 | 1.886 | 2.920 | 4.303 | 6.205 | 9.925 | 14.089 | 31.598 |
| 3 | .765 | 0.978 | 1.638 | 2.353 | 3.182 | 4.176 | 5.841 | 7.453 | 12.941 |
| 4 | .741 | .941 | 1.533 | 2.132 | 2.776 | 3.495 | 4.604 | 5.598 | 8.610 |
| 5 | .727 | .920 | 1.476 | 2.015 | 2.571 | 3.163 | 4.032 | 4.773 | 6.859 |
| 6 | .718 | .906 | 1.440 | 1.943 | 2.447 | 2.969 | 3.707 | 4.317 | 5.959 |
| 7 | .711 | .896 | 1.415 | 1.895 | 2.365 | 2.841 | 3.499 | 4.029 | 5.405 |
| 8 | .706 | .889 | 1.397 | 1.860 | 2.306 | 2.752 | 3.355 | 3.832 | 5.041 |
| 9 | .703 | .883 | 1.383 | 1.833 | 2.262 | 2.685 | 3.250 | 3.690 | 4.781 |
| 10 | .700 | .879 | 1.372 | 1.812 | 2.228 | 2.634 | 3.169 | 3.581 | 4.587 |
| 11 | .697 | .876 | 1.363 | 1.796 | 2.201 | 2.593 | 3.106 | 3.497 | 4.437 |
| 12 | .695 | .873 | 1.356 | 1.782 | 2.179 | 2.560 | 3.055 | 3.428 | 4.318 |
| 13 | .694 | .870 | 1.350 | 1.771 | 2.160 | 2.533 | 3.012 | 3.372 | 4.221 |
| 14 | .692 | .868 | 1.345 | 1.761 | 2.145 | 2.510 | 2.977 | 3.326 | 4.140 |
| 15 | .691 | .866 | 1.341 | 1.753 | 2.131 | 2.490 | 2.947 | 3.286 | 4.073 |
| 16 | .690 | .865 | 1.337 | 1.746 | 2.120 | 2.473 | 2.921 | 3.252 | 4.015 |
| 17 | .689 | .863 | 1.333 | 1.740 | 2.110 | 2.458 | 2.898 | 3.222 | 3.965 |
| 18 | .688 | .862 | 1.330 | 1.734 | 2.101 | 2.445 | 2.878 | 3.197 | 3.922 |
| 19 | .688 | .861 | 1.328 | 1.729 | 2.093 | 2.433 | 2.861 | 3.174 | 3.883 |
| 20 | .687 | .860 | 1.325 | 1.725 | 2.086 | 2.423 | 2.845 | 3.153 | 3.850 |
| 21 | .686 | .859 | 1.323 | 1.721 | 2.080 | 2.414 | 2.831 | 3.135 | 3.819 |
| 22 | .686 | .858 | 1.321 | 1.717 | 2.074 | 2.406 | 2.819 | 3.119 | 3.792 |
| 23 | .685 | .858 | 1.319 | 1.714 | 2.069 | 2.398 | 2.807 | 3.104 | 3.767 |
| 24 | .685 | .857 | 1.318 | 1.711 | 2.064 | 2.391 | 2.797 | 3.090 | 3.745 |
| 25 | .684 | .856 | 1.316 | 1.708 | 2.060 | 2.385 | 2.787 | 3.078 | 3.725 |
| 26 | .684 | .856 | 1.315 | 1.706 | 2.056 | 2.379 | 2.779 | 3.067 | 3.707 |
| 27 | .684 | .855 | 1.314 | 1.703 | 2.052 | 2.373 | 2.771 | 3.056 | 3.690 |
| 28 | .683 | .855 | 1.313 | 1.701 | 2.048 | 2.368 | 2.763 | 3.047 | 3.674 |
| 29 | .683 | .854 | 1.311 | 1.699 | 2.045 | 2.364 | 2.756 | 3,.038 | 3.659 |
| 30 | .683 | .854 | 1.310 | 1.697 | 2.042 | 2.360 | 2.750 | 3.030 | 3.646 |
| 35 | .682 | .852 | 1.306 | 1.690 | 2.030 | 2.342 | 2.724 | 2.996 | 3.591 |
| 40 | .681 | .851 | 1.303 | 1.684 | 2.021 | 2.329 | 2.704 | 2.971 | 3.551 |
| 45 | .680 | .850 | 1.301 | 1.680 | 2.014 | 2.319 | 2.690 | 2.952 | 3.520 |
| 50 | .680 | .849 | 1.299 | 1.676 | 2.008 | 2.310 | 2.678 | 2.937 | 3.496 |
| 55 | .679 | .849 | 1.297 | 1.673 | 2.004 | 2.304 | 2.669 | 2.925 | 3.476 |
| 60 | .679 | .848 | 1.296 | 1.671 | 2.000 | 2.299 | 2.660 | 2.915 | 3.460 |
| 70 | .678 | .847 | 1.294 | 1.667 | 1.994 | 2.290 | 2.648 | 2.899 | 3.435 |
| 80 | .678 | .847 | 1.293 | 1.665 | 1.989 | 2.284 | 2.638 | 2.887 | 3.416 |
| 90 | .678 | .846 | 1.291 | 1.662 | 1.986 | 2.279 | 2.631 | 2.878 | 3.402 |
| 100 | .677 | .846 | 1.290 | 1.661 | 1.982 | 2.276 | 2.625 | 2.871 | 3.390 |
| 120 | .677 | .845 | 1.289 | 1.658 | 1.980 | 2.270 | 2.617 | 2.860 | 3.373 |
| × | .6745 | .8416 | 1.2816 | 1.6448 | 1.9600 | 2.2414 | 2.5758 | 2.8070 | 3.2905 |

1/ Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods, 7th ed. Iowa State Univ. Press, Ames. (No part of this appendix may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the prior written permission of the publisher.)

**Appendix C** Significance of r [1]

| df | 10% | 5% | 2% | 1% |
|---|---|---|---|---|
| 3 | 0.805 | 0.878 | 0.934 | 0.959 |
| 4 | .729 | .811 | .882 | .917 |
| 5 | .669 | .754 | .833 | .874 |
| 6 | .622 | .707 | .789 | .834 |
| 7 | .582 | .666 | .750 | .798 |
| 8 | .549 | .632 | .716 | .765 |
| 9 | .521 | .602 | .685 | .735 |
| 10 | .497 | .576 | .658 | .708 |
| 11 | .476 | .553 | .634 | .684 |
| 12 | .458 | .532 | .612 | .661 |
| 13 | .441 | .514 | .592 | .641 |
| 14 | .426 | .497 | .574 | .623 |
| 15 | .412 | .482 | .558 | .606 |
| 16 | .400 | .468 | .542 | .590 |
| 17 | .389 | .456 | .528 | .575 |
| 18 | .378 | .444 | .516 | .561 |
| 19 | .369 | .433 | .503 | .549 |
| 20 | .360 | .423 | .492 | .537 |
| 25 | .323 | .381 | .445 | .487 |
| 30 | .295 | .349 | .409 | .449 |
| 35 | .275 | .325 | .381 | .418 |
| 40 | .257 | .304 | .358 | .393 |
| 45 | .243 | .288 | .338 | .372 |
| 50 | .231 | .273 | .322 | .354 |
| 60 | .211 | .250 | .295 | .325 |
| 70 | .195 | .232 | .274 | .302 |
| 80 | .183 | .217 | .256 | .283 |
| 90 | .173 | .205 | .242 | .267 |
| 100 | .164 | .195 | .230 | .254 |
| 150 | .134 | .160 | .189 | .208 |
| 200 | .116 | .138 | .164 | .181 |
| 300 | .095 | .113 | .134 | .148 |
| 400 | .082 | .098 | .116 | .128 |
| 500 | 0.073 | 0.088 | 0.104 | 0.115 |

1/ Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods, 7th ed. Iowa State Univ. Press, Ames. (No part of this appendix may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the prior written permission of the publisher.)

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

# Part 614
# National Water Quality Handbook

# Index

(450–VI–NWQH, September 2003)

# Index

## O



## P



## Q



## R



## S



## T



## V

## W

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook



# Part 615 Water Quality Monitoring Data Analysis

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

National Water Quality Handbook

# Part 615 Analysis of Water Quality Monitoring Data

(450–VI–NWQH, September 2003)

# Preface

## Purpose

The purpose of part 615 of the National Water Quality Handbook (NWQH) is to provide guidance in the statistical analysis of water quality data that have been collected according to the designs described in part 614. Part 615 is concerned with the statistical analysis of monitoring results.

## Acknowledgments

Part 615 of the National Water Quality Handbook was written by **John C. Clausen**, Ph.D., College of Agriculture and Natural Resources, University of Connecticut. The concept for this project was developed by **James N. Krider**, former national environmental engineer, Natural Resources Conservation Service (NRCS), Washington, DC. Technical leadership was provided by **Bruce J. Newton**, limnologist, NRCS National Water and Climate Center, Portland, Oregon.

The following people contributed to the review of this document:

**Richard H. McCuen,** University of Maryland, College Park, Maryland
**Carl DuPoldt,** agricultural engineer, NRCS, Somerset, New Jersey
**Barbara M. Vining,** pesticide and nutrient specialist, formally with NRCS, Indianapolis, Indiana
**Howard Thomas,** (retired) economist, Portland, Oregon
**Robert C. Feurer,** Philadelphia Suburban Water Company, Bryn Bawr, Pennsylvania
**Lynn Sampson,** biologist, NRCS, East Lansing, Michigan
**William H. Boyd,** environmental engineer, NRCS National Water Management Center, Little Rock, Arkansas
**James C. Wood,** water quality specialist, NRCS, Burlington, Vermont
**Richard Croft,** water quality specialist, NRCS Watershed Science Institute, Burlington, Vermont
**Jorge A. Delgado,** USDA-Agricultural Research Service
**Ron Schierer,** resource conservationist, Greely, Colorado

**Joseph A. Neafsey**, water quality specialist, Storrs, Connecticut, was the contract project officer. **Mary R. Mattinson**, editor, and **Wendy R. Pierce,** illustrator, NRCS National Cartographic and Geospatial Center, Fort Worth, Texas, provided editorial and graphic support.

The National Water Quality Handbook is the result of a collaborative effort of the NRCS Science and Technology Consortium. Special thanks to the following individuals for their leadership and support in the development and publishing of this essential technical reference:

**Wil Fontenot**, natural resource specialist, Lafayette, Louisianna
**Jon Werner**, national hydrologist, Conservation Engineering Division, Washington, D.C.
**Roy Mann**, (retired) resource conservationist, Portland, Oregon
**Ken Pfeiffer**, pest management specialist, National Water and Climate Center, Portland, Oregon
**Bruce Newton**, acting director, National Water and Climate Center, Portland, Oregon
**Lynn Betts**, communications director, Wildlife Habitat Management Institute, Des Moines, IA

# Part 615 Analysis of Water Quality Monitoring Data

Contents:

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 1 Introduction

# Chapter 1 Introduction

Contents:

# Chapter 1 Introduction

## 615.0100 General

In National Water Quality Handbook (NWQH), part 614, the 12 steps for designing a water quality monitoring study were described. The overall purpose of part 615 is to provide assistance in how to analyze water quality data that have been collected according to the designs described in part 614. It is not the intention that part 615 replace a basic course or textbook on statistics; actually the reader would be much better prepared for this part of the handbook having had such a course.

Chapters 2 to 5 provide background information on statistical analysis; chapters 6 to 12 provide guidance on how to analyze data obtained from particular monitoring designs; and chapter 13 describes information on several available computer packages for statistics. The chapters include several examples that use both hand calculations and computer-generated output. Many computerized statistical packages are available today, and to save time and effort, the user is encouraged to invest in a package. Chapter 13 provides guidance on how to select statistical analysis software.

The Statistical Analysis System (SAS) software for a PC is used for illustration purposes throughout part 615 of the NWQH.

Table 1–1 summarizes the statistical procedures used in part 615 and indicates the chapter where that procedure is best described. Table 1–2 summarizes the purpose of the various statistics and statistical tests used in part 615 of the handbook.

## 615.0101 Steps in statistical analysis

As in part 614, there are several steps in conducting the statistical analysis of water quality data (fig. 1–1). The analysis of data begins with Exploratory data analysis (EDA), which is intended for the analyst to become familiar with the data (Tukey 1977). The next step is to test the appropriate assumptions for the statistical tests to be performed. The assumptions may include randomness, the type of distribution, the homogeneity of variances, and independence. The next step is to determine the appropriate hypotheses to test. This step may have already been completed as part of designing the study. The next step would be to conduct the actual statistical tests. Finally, the conclusions regarding the data are constructed. The following chapters are intended to assist the analyst through these steps of data analysis.

**Figure 1–1** Steps in data analysis for a water quality monitoring study

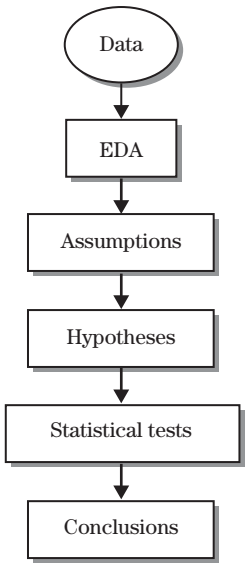


**Table 1–1** Summary of statistical procedures used in Part 615, by chapter

| Procedure | Chapter | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| **Basic statistics:** | | | | | | | | | | | |
| Mean | X | | X | | | | | | | | |
| Median | X | | X | | | | | | | | |
| Mode | X | | | | | | | | | | |
| Variance | X | | | | | | | | | | |
| Standard deviation | X | | | | | | | | | | |
| Standard error | X | | | | | X | | | | | |
| Coefficient of variation | X | | | | | | | | | | |
| Coefficient of skewness | X | | X | | | | | | | | |
| Kurtosis | | | X | | | | | | | | |
| Shapiro-Wilk W-statistic | | | X | | | | | | | | |
| Autocorrelation coefficient | | | X | | | | | | | | |
| **Statistical tests:** | | | | | | | | | | | |
| t-test | | | | | | | X | X | | | |
| Mann-Whitney U (nonparametric) | | | | | | | X | | | | |
| Wilcoxon paired sample (nonpar) | | | | | | | | X | | | |
| F ratio | | | X | | | | | | | X | |
| Analysis of variance | | | | | | X | | | | X | X |
| one-way | | | | | | X | | | | | |
| Kruskal-Wallis one-way (nonpar) | | | | | | | | | | | |
| two-way | | | | | | X | | | | | |
| Tukey's multiple comparisons | | | | | | X | | | | | X |
| Regression | | | | | | | | | X | | X |
| Coefficient of determination | | | | | | | | | X | | |
| Confidence intervals | | | | | | | | | X | | |
| Analysis of covariance | | | | | | | | | X | | |
| Kendall tau | | | | | | | | | | | X |

**Table 1–2** Summary of purpose of statistical procedures used in Part 615

| Procedure | Purpose |
|---|---|
| **Basic statistics:** | |
| Mean | measure of central tendency |
| Median | measure of central tendency |
| Mode | measure of central tendency |
| Variance | measure of dispersion of a random variable |
| Standard deviation | measure of dispersion |
| Standard error | measure of dispersion of a statistic |
| Coefficient of variation | standardized measure of dispersion |
| Coefficient of skewness | measure of symmetry |
| Kurtosis | measure of long tailedness (peakedness) of dispersion |
| Shapiro-Wilk W-statistic | test for normality |
| Autocorrelation coefficient | measure of independence of observations on a single random variable |
| **Statistical tests:** | |
| one-sample t-test | comparison of a single mean to a standard |
| two-sample t-test | comparison of two sample means |
| Mann-Whitney U (nonparametric) | nonparametric comparison of unpaired two-sample ranks |
| Wilcoxon paired sample (nonpar) | nonparametric comparison of paired ranks of differences |
| F ratio | test of homogeneity of variances |
| Analysis of variance | comparison of several means |
| one-way | comparison of several means for one factor |
| Kruskal-Wallis one-way (nonpar) | comparison of several means for one factor, nonparametric |
| two-way | comparison of several means for two factors |
| Tukey's multiple comparisons | determine which means are different for a rejected ANOVA test |
| Regression | relationship between two variables |
| Coefficient of determination | fraction of variation explained by relationship |
| Confidence intervals | measure of accuracy of a statistic |
| Analysis of covariance | comparison of regression slopes and intercepts among groups |
| Kendall tau | nonparametric measure of correlation for trend detection |

# 615.0102 References

Tukey, J.W. 1977. Exploratory data analysis. Addison-Wesley Publ. Co., Reading, MA.

United States Department of Agriculture, Natural Resources Conservation Service. 1996. Design of water quality monitoring systems**.** National Water Quality Handbook, Part 514.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 2 Basic Statistics

(450–VI–NWQH, September 2003)

# Chapter 2 Basic Statistics

Contents:



Tables

**Figures**



**Examples**

# Chapter 2 Basic Statistics

## 615.0200 Introduction

The understanding of basic statistics is important to the analysis of water quality data. For many, chapter 2 is a review of some of the foundations of statistics. Included in this chapter is the purpose of statistics, some statistical terms, definitions of data types, frequencies, measures of central tendency, and measures of dispersion.

## 615.0201 Purpose of statistics

In water quality monitoring, the use of statistics is important. For example, if our measurement of the quality of water averaged three this year and six next year, has the water quality really doubled in a year? In other words, is the number three different from the six and how confident can I be that they are or are not different?

Almost all water quality data collected are a sample. That is, we sample a certain portion of the entire population of water quality data available. For example, if we sample a well weekly from 2003 to 2008 for nitrate-N, that also means that we are not sampling the well during all other times. Assuming it takes at most 30 minutes to sample a well, we are sampling only 0.3 percent of the time during the week. We also are sampling between 2003 and 2008. We are not sampling before 2003 nor after 2008, which are times that also may be part of the entire population of water quality data. Therefore, the real purpose of statistics is to be able to make conclusions from a sampling of data for the entire population. Because we usually cannot measure the entire population, a sample is necessary. Statistics provide a systematic framework for analysis and summarization of the sample data.

## 615.0202 Statistical terms

A number of statistical terms used throughout this chapter are defined in this section.

***Observation***—A record representing a characteristic of a real-world object (EPA 1973). The record is generally a single number; for example, a chemical concentration or the number of macroinvertebrates found in a sample. The observation is the data you collect.

***Population***—The population is all possible values of a variable and is synonymous with universe (Steel and Torrie 1960).

***Sample***—A part of the population that should be representative of the population (Steele and Torrie 1960). A sample is a set of observations from the population.

***Random sample***—A sample that has an equal chance of being selected (Snedecor and Cochran 1980). Usually such a sample is collected to eliminate bias in the data.

## 615.0203 Data types

The two types of random variables that can be collected in water quality monitoring projects are continuous and discrete. The type of data selected influences the statistics applied and depends on the type of information being collected.

*Continuous data* means that all values within some range are possible (Steel and Torrie 1960). An example of continuous data would be concentrations. A nearly infinite number of values are possible within some range. More values become possible as detection equipment becomes more precise.

*Discrete data* means that the possible values can be only a certain set of numbers (Snedecor and Cochran, 1980). Examples include counts, categories, and binary data. The number of fish collected would be discrete data.

In addition to the continuous and discrete data, several scales can be used to measure water quality data. They include nominal, ordinal, interval, and ratio scales.

*Nominal data* include categories without ranking among the categories. The term nominal means that the category is called a name. Often, nominal data are binary, such as presence or absence. An example of nominal data would be taxa of macroinvertebrates present in a stream.

*Ordinal data* imply ordering (Ward et al. 1990). Ordinal variables measure the degree of something (Horowitz 1981). Trophic status—oligotrophic, mesotrophic, and eutrophic—is an example of an ordinal scale. However, the differences among the categories do not have to be equal.

*Interval data* also use ordering, but intervals between the categories are equal. Intervals or categories are used to describe the data. Interval data are used for data sets that do not have a true zero. For example, the intervals for temperature could be <25, 25–50, 50–75, and >75 degrees Fahrenheit. Intervals also are used to describe size classes of fish, such as <10, 10–20, 20–30, and >30 centimeters.

*Ratio data* are similar to interval data except that a true zero exists. Therefore, 500 is 5 times greater than 100. Concentration and flow data are ratio data.

## 615.0204 Frequencies

Water quality data can be presented in many ways. They include tables of raw data or frequencies, seasonal tables, and graphical pie charts or frequency diagrams. A raw data table is given in table 2–1 for algal counts in St. Albans Bay, Lake Champlain, Vermont.

This raw data can be summarized in a frequency table by establishing intervals in the data. For example, the raw algal data in table 2–1 were grouped into intervals of 2,500 organisms per milliliter and are summarized in table 2–2. The *frequency* is the number of observations for that class interval.

The frequency table can also be displayed as a frequency histogram. A histogram graphs frequency as a function of class intervals as rectangles on a graph (fig. 2–1).

Such data may also be presented as a cumulative frequency histogram. The *cumulative frequency* is the summation of all the frequencies up to and including the class interval plotted (fig. 2–2). The points are joined with a line forming a cumulative frequency polygon (Zar 1996).

The frequency histogram and the cumulative frequency polygon can be converted to relative frequency. This is done by changing the Y-axis to either a decimal or percentage scale by dividing the frequencies by the total sample size.

Frequency plots have several values, including:

- help assess the distribution type
- detect characteristics of the data (e.g., central tendency, dispersion)
- identify potential outliers
- assess the range of data

Although these forms of data presentation are useful, there are other ways to describe the data. They include describing a measure of central tendency and a measure of dispersion of the data.

**Table 2–1** Raw algal counts (organisms/mL) from St. Albans Bay, Vermont, 1985

| Date | Count | Date | Count |
|---|---|---|---|
| 1/23 | 25 | 8/6 | 1,564 |
| 3/19 | 125 | 8/13 | 6,384 |
| 4/23 | 410 | 8/20 | 10,062 |
| 5/14 | 1,883 | 8/27 | 6,305 |
| 5/30 | 770 | 9/4 | 39,861 |
| 6/11 | 2,229 | 9/10 | 6,755 |
| 6/18 | 519 | 9/17 | 15,074 |
| 6/25 | 899 | 9/25 | 36,823 |
| 7/2 | 882 | 10/1 | 29,448 |
| 7/9 | 565 | 10/8 | 45,283 |
| 7/16 | 826 | 10/15 | 1,336 |
| 7/23 | 299 | 11/5 | 1,000 |
| 7/30 | 547 | 12/4 | 56 |

**Table 2–2** Frequency table of algal counts in St. Albans Bay, Vermont

| Interval | Frequency | Interval | Frequency |
|---|---|---|---|
| 0 – 2,500 | 17 | 25,000 – 27,500 | 0 |
| 2,500 – 5,000 | 0 | 27,500 – 30,000 | 1 |
| 5,000 – 7,500 | 3 | 30,000 – 32,500 | 0 |
| 7,500 – 10,000 | 0 | 32,500 – 35,000 | 0 |
| 10,000 – 12,500 | 1 | 35,000 – 37,500 | 1 |
| 12,500 – 15,000 | 0 | 37,500 – 40,000 | 1 |
| 15,000 – 17,500 | 1 | 40,000 – 42,500 | 0 |
| 17,500 – 20,000 | 0 | 42,500 – 45,000 | 0 |
| 20,000 – 22,500 | 0 | 45,000 – 47,500 | 1 |
| 22,500 – 25,000 | 0 | | |

**Figure 2–1** Frequency histogram of algal counts in St. Albans Bay, Vermont

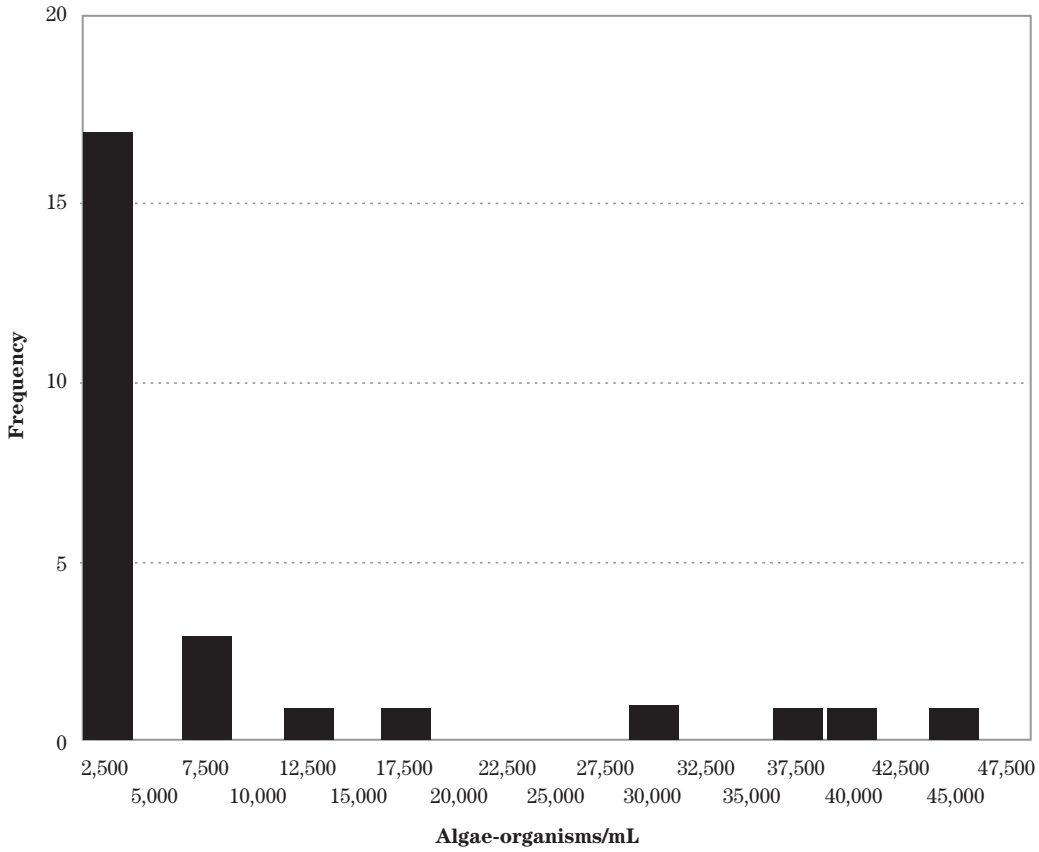
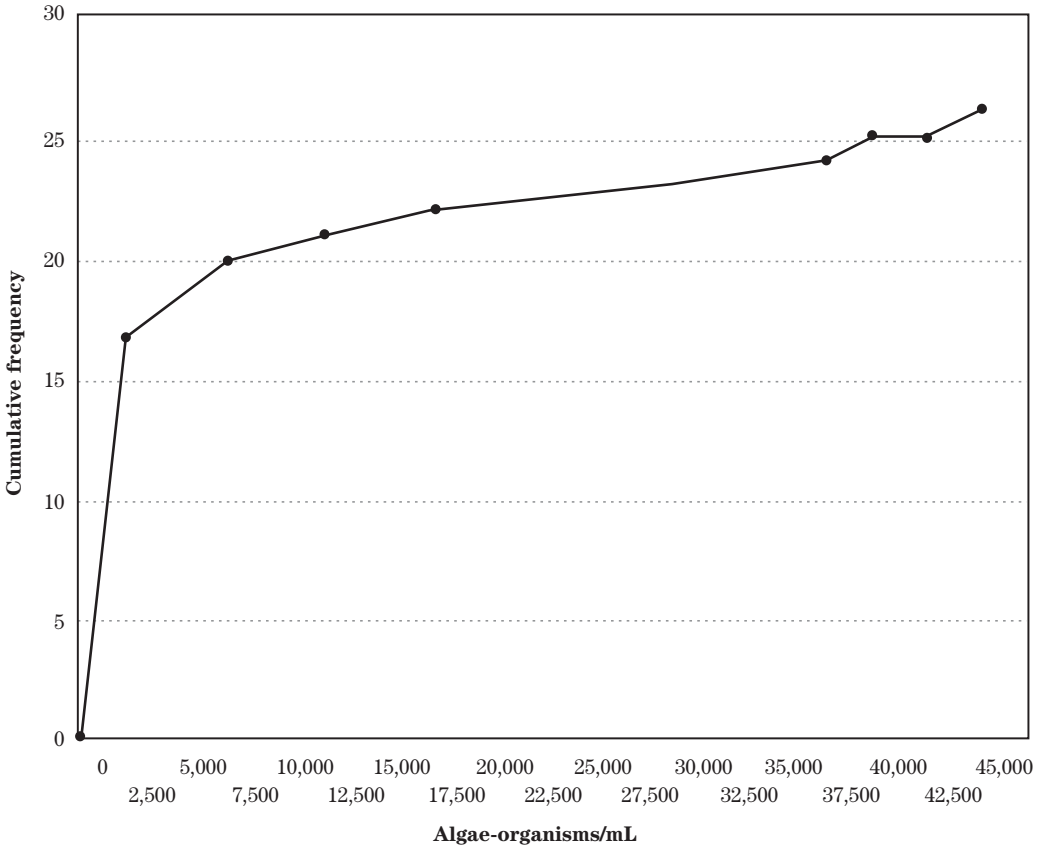


**Figure 2–2** Cumulative frequency of algal counts in St. Albans Bay, Vermont



# 615.0105 Measures of central tendency

Several measures of central tendency for a data set are available. The appropriate measure varies with the type of data (table 2–3). Example 2–1 illustrates the different measures of central tendency.

**Table 2–3** Measures of central tendency for data types

| Scale | Measure | Example |
|---|---|---|
| nominal | mode | taxa |
| ordinal | median | trophic state |
| interval | mean | fish age class |
| ratio | mean | concentrations |

## (a) Mean

The most commonly used measure is the arithmetic mean or average. The mean ( $\overline{X}$ ) is the sum of the observations ( $\sum \overline{X}_i$ ) divided by the number of observations (n):

$$\overline{X} = \frac{\sum X_i}{n} \qquad [2–1]$$

The mean is appropriate for interval and ratio data, but not nominal or ordinal types of information. Arithmetic means may not be the best measure of central tendency when distributions are skewed (long tail) left or right. If the data are censored, that is, there are observations below detection limits, the calculation of the mean is more rigorous. The mean for a censored distribution can be calculated from (Newman et al. 1989):

$$\overline{X} = \overline{X} - \sigma \frac{k}{n-k} \frac{f(\varepsilon)}{F(\varepsilon)} \qquad [2–2]$$

**Example 2–1** Measures of central tendency

**Given:** The algal count data from St. Albans Bay in table 2–1.

**Determine:** The mean, geometric mean, median, and mode.

**Solution:**

Mean: $$\overline{X} = \frac{25 + 125 + 410 + \text{K} + 56}{26} = 8{,}074 \text{ organisms/mL}$$

Geometric mean: $$\overline{X}_g = \text{anti}\log \frac{\sum_{i=1}^{n} X_i}{n} = \text{anti}\log 3.2343 = 1{,}715 \text{ organisms/mL}$$

Median—Because the data contain an even number of data values (n=26), the median is the mean of the two middle values.

$$\text{Median} = \frac{1{,}336 + 1{,}000}{2} = 1{,}168 \text{ organisms/mL}$$

Mode—No value occurred more than once in table 2–1; therefore, the mode does not exist for this data set.

where:
- n = total number of observations
- k = number of observations below the detection limit
- $\overline{X}$ = mean of all the values above the detection limit
- σ = standard deviation
- f(ε) = distribution function for the normal distribution
- F(ε) = cumulative distribution function for the normal distribution

ε is obtained from:

$$\varepsilon = \frac{DL - \hat{\mu}}{\sigma} \qquad [2–3]$$

where:
- DL = detection limit
- $\hat{\mu}$ = mean

In water quality data a geometric mean is often calculated. The geometric mean $\overline{X}_G$ is the $n^{th}$ root of the product of n values (Landwehr 1978, Zar 1996):

$$\overline{X}_G = \sqrt{X_1 X_2 \ldots X_n} \qquad [2–4]$$

The geometric mean is also obtained as the antilog of the mean of the log of the values, which is the typical manner of calculating the geometric mean:

$$\overline{X}_G = \text{anti}\log \frac{\sum_{i=1}^{n} X_i}{n} \qquad [2–5]$$

The geometric mean is only used when all the values are positive and is typically used as the measure of central tendency for log transformed data.

## (b) Median

A second measure of central tendency is the median ($\overline{X}_m$). The median is the value for which 50 percent of observations are greater and 50 percent are lesser. It is the midpoint of a frequency distribution. The median is an appropriate measure of centrality for ordinal data and is often used when the data are highly skewed. If a distribution is symmetrical, then the mean and the median will be the same.

$$\overline{X}_m = \begin{cases} X_{\frac{(n+1)}{2}} & \text{if n is odd} \\ \left( \frac{X_{\frac{n}{2}} + X_{\frac{n+1}{2}}}{2} \right) & \text{if n is even} \end{cases} \qquad [2–6]$$

## (c) Mode

The mode is the final measure of central tendency. It is the value that occurs most frequently. The mode is the only appropriate measure of central tendency for nominal data and quickly describes the most commonly occurring value.

## 615.0106 Measures of dispersion

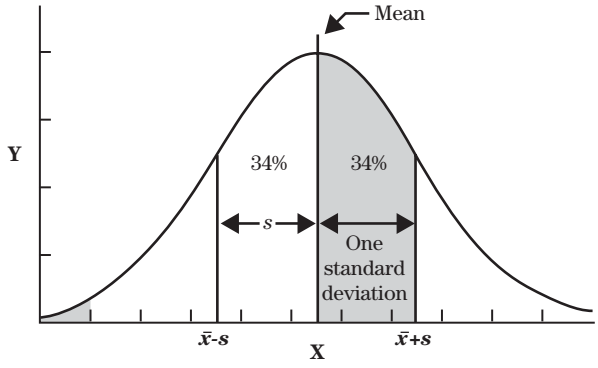
Measures of dispersion are useful to further understand a water quality data set. They indicate how spread out from the central tendency are the observations. The common measures of dispersion include the range, the variance (standard deviation is square root of variance), the standard error (standard deviation of a statistic, such as the mean), and the coefficient of variation.

A normal distribution has a preponderance of values around the mean and fewer observations at the extremes of the range of values. Such a distribution forms the typical bell-shaped curve (fig. 2–3).

The *range* is the distance from the smallest value to the largest value in the data set. It is the most simple of the measures of dispersion, but is subject to extreme values.

The *sample variance* is the sum of the squares of the deviations from the mean divided by the number of observations minus 1. Another term for variance is the mean square, which is the sums of squares divided by the degrees of freedom (n-1). The sample variance is represented by $s^2$, and the population variance is represented by $\sigma^2$.

**Figure 2–3** Normal distribution




The sample variance is calculated from:

$$s^2 = \frac{\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}}{n-1} \qquad [2–7]$$

where:
$X_i$ = value of the observation
n = number of observations

The population variance is the sum of the squares of the deviations from the mean divided by the number of observations, rather than the number of observations minus one. The population variance is rarely used in water quality studies because sampling is almost always being conducted. Some calculators compute the wrong variance.

The *standard deviation* (s) is the square root of the variance ( $\sqrt{s^2}$ ). The standard deviation carries the same units as the original data. The s is also called the root mean square. For normal distributions, one standard deviation on either side of the mean includes 68 percent of the observations and two standard deviations include 95 percent (fig. 2–3).

The *standard error of the mean* (SE), also termed the standard deviation of the mean, indicates the variability about the estimate of the mean:

$$SE = \frac{s}{\sqrt{n}} \qquad [2–8]$$

where:
s = standard deviation
n = number of observations

The standard error of the mean can be shown as an error bar in graphs summarizing mean values.

The *coefficient of variation* (CV) is a measure of the relative dispersion about the mean. It is defined as the standard deviation expressed as a percent of the mean:

$$CV = \frac{100 \times s}{\overline{X}} \qquad [2–9]$$

where:
s = standard deviation
$\overline{X}$ = mean

The advantage of the coefficient of variation is that it allows direct comparison of variations between variables or among studies.

The *coefficient of skewness* indicates how equally distributed or symmetrical the data are about the mean. It is defined as the cube of the deviations about the mean (SAS 1985):

$$g_1 = \frac{n\sum\left(X_i - \bar{X}\right)^3}{(n-1)(n-2)s^3} \qquad [2\text{–}10]$$

The coefficient of skewness is normally distributed with a mean of 0 and a standard deviation of:

$$\left(\frac{6n(n-1)}{(n-2)(n+1)(n+3)}\right)^{.05} \qquad [2\text{–}11]$$

If $g_1$ is greater than four times the standard deviation of the skewness coefficient, then the data are skewed. Snedecor and Cochran (1980) provide a table for determining the significance of the skewness coefficient (appendix B). The sign of the skewness coefficient indicates whether the data are positively skewed (upper tail extended) or negatively skewed (lower tail extended) (fig. 2–4).

*Kurtosis* is a measure of the long tailedness of the distribution. It is defined as the average of the deviations from the mean raised to the 4th power divided by the standard deviation to the 4th power (SAS 1985):

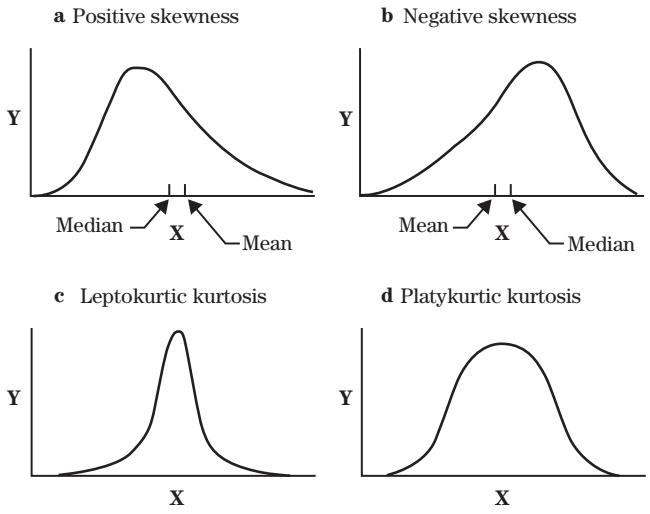
$$g_2 = \frac{n(n+1)\sum\left(X_i - \bar{X}\right)^4}{(n-1)(n-2)(n-3)s^4} - 3\frac{(n-1)^2}{(n-2)(n-3)} \qquad [2\text{–}12]$$

The kurtosis is normally distributed with a mean of –3 and a standard deviation of:

$$\sqrt{\frac{24}{n}}$$

If the ratio of $g_2$ to standard deviation is less than –2, then the distribution has shorter tails than a normal distribution. If the ratio is more than 2, then the distribution has longer tails than a normal distribution (fig. 2–4). Snedecor and Cochran (1980) provide a table for testing the kurtosis based on the sample size and level of confidence desired.

**Figure 2–4** Distributions showing skewness and kurtosis



**Example 2–2** Measures of dispersion

**Given:** Algal data in table 2–1

**Determine:** Range, variance, standard deviation, standard error, coefficient of variation, skewness, and kurtosis values.

**Solution:** Range: 45,283 – 25 = 45,258 organisms/mL

Variance: $$s^2 = \frac{6,334,495,000 - \frac{(209,930)^2}{26}}{26-1} = 185,590,000$$

Standard deviation: $$s = \sqrt{185,590,000} = 13,623 \text{ organisms / mL}$$

Standard error: $$SE = \frac{13,623}{\sqrt{26}} = 2,672 \text{ organisms / mL}$$

Coefficient of variation: $$CV = \frac{100(13,623)}{8,074} = 169\%$$

Skewness: $$g_1 = \frac{(26)(0.110873E+15)}{(25)(24)(23)(13,623)^3} = 1.90$$

Since the skewness coefficient is positive, the upper tail is extended. Based upon a table provided by Snedecor and Cochran (1980) (appendix B), this skewness is significant at a probability (p) = 0.01.

Kurtosis: $$g_2 = \frac{(26)(27)(0.38816E+19)}{(25)(24)(23)(13,623)^4} - \frac{3(25)^2}{(24)(23)}$$
$$= 2.335$$

The standard deviation of the kurtosis is: $$\sqrt{\frac{24}{n}} = \sqrt{\frac{24}{26}} = 0.9607$$
$$\frac{g^2}{0.9607} = 2.4$$

Since the ratio of $g_2$ to the standard deviation is greater than 2, the algae data have longer tails than a normal distribution (fig. 2–1).

# 615.0107 References

Horowitz, G. 1981. Sadistic statistics. Avery Publishing Group, Inc., Wayne, NJ.

Landwehr, J.M. 1978. Some properties of the geometric mean and its use in water quality standards. Water Resources Research 14(3):467-473.

Newman, M.C., P.M. Dixon, B.B. Looney, and J.E. Pinder, III. 1989. Estimating mean and variance for environmental samples with below detection limit observations. Water Resources Bulletin 25(4):905-916.

SAS Institute, Inc. 1985. SAS procedures guide for personal computers. SAS Institute, Inc. Cary, NC 27511.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods. 7th Ed. The Iowa State Univ. Press, Ames.

Steele, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York.

United States Environmental Protection Agency. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. EPA-670/4-73-001. Office of Research and Development. Cincinnati, OH.

Ward, R.C., J.C. Loftis, and G.B. McBride. 1990. Design of water quality monitoring systems. Van Nostrand Reinhold. New York.

Zar, J.H. 1996. Biostatistical analysis. 3rd Ed. Prentice Hall, Upper Saddle River, NJ 07458.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 3 Exploratory Data Analysis

(450–VI–NWQH, September 2003)

# Chapter 3 Exploratory Data Analysis

**Contents:**



**Figures**

**Examples**

# Chapter 3 Exploratory data analysis

## 615.0300 Introduction

The first step in water quality data analysis is exploratory data analysis (EDA). For most data sets, EDA is a necessary step. The basic purpose for EDA is to better become familiar with the data. EDA is "detective work" that examines the data for how it appears (Tukey 1977). EDA, as proposed by Tukey, relies heavily on pictures. It is intended to provide indications rather than confirmations of a specific test. The actual procedure used varies with the type of data being explored, whether univariate, bivariate, or multivariate. Not all techniques are appropriate for all data; however, a number of steps are often examined for routine EDA. They include writing the numbers, stem-and-leaf diagrams, schematic summaries, transformations, comparisons, plots of relationships, and smoothing data.

Chapter 3 explains the various approaches to EDA. It presents examples of each of the routine methods used in EDA.

## 615.0301 Writing numbers

The process of writing numbers may be as simple as the listing of the raw data in a table. Tukey (1977) suggests using colors to highlight differences in the numbers making visual inspection easier.

Table 2–1 in chapter 2 is an example of writing numbers. In this case the numbers were written according to date. An alternative presentation would be to write the numbers from lowest to highest.

## 615.0302 Stem-and-leaf diagrams

Stem-and-leaf diagrams summarize the data visually in a sideways frequency diagram. Each line in the diagram is a stem, and each data point is a leaf on the stem. The stem represents the first digit of an observation in the data set. The leaves indicate the number of observations at that stem and the digits for those observations. Significant figures to the right of the leaves often are dropped. Stem-and-leaf diagrams are presented in many ways. Such a diagram, as presented in output from the Statistical Analysis System software (SAS®) for the algal data in table 01–1, is given in figure 3–1.

In this diagram, each stem is a multiple of 10,000; indicated by a 10**+4 by SAS®. The 4 represents 40,000, 3 represents 30,000 and so on. The leaf of 05 indicates that there are two numbers of 40,000 or greater, after rounding to the nearest 1,000. The data are skewed toward the low values (Stem = 0). There are more values for the leaf column at the stem of 0 than other stem values. SAS® output indicates the number of leaves in each stem by a # column. SAS® output also gives a multiplication factor for the Stem.Leaf data if needed.

**Figure 3–1** Stem-and-leaf diagram for algal data from SAS® output

| Stem | Leaf | # |
|---|---|---|
| 4 | 05 | 2 |
| 3 | 7 | 1 |
| 2 | 9 | 1 |
| 1 | 05 | 2 |
| 0 | 00000111111111222667 | 20 |
| | --- +--- +--- +--- + | |

Multiply Stem.Leaf by 10**+4

## 615.0303 Schematic summaries

The stem-and-leaf diagram can also be summarized using five numbers: the median, maximum, minimum, and upper and lower hinges. The rank of the median can be determined from:

$$\text{median rank} = \frac{1+\text{count}}{2} \qquad [3–1]$$

The hinges are half-way from the extremes to the median and are determined by:

$$\text{hinge} = \frac{1+\text{median rank}}{2} \qquad [3–2]$$

The hinges are so-named because they represent folds in the data between the median and the extremes (Tukey 1977). Another way to characterize the lower and upper hinges is as the 25th and 75th percentile values. The upper hinge is the value that is three/fourths of the way along the values when ranked from lowest to highest.

These five numbers can be provided in a box, as below:

| | |
|---|---|
| Med | |
| $H_{low}$ | $H_{high}$ |
| Min | Max |

For example, using the algae data from table 2–1, the five-number summary would be:

| | |
|---|---|
| 1,168 | |
| 547 | 6,755 |
| 25 | 45,283 |

### (a) Box-and-whisker plot

Another more common schematic summary is the box-and-whisker plot. This is really a five-number summary in graphical form. The box extends from lower hinge to upper hinge and is crossed with a bar at the median (fig. 3–2). The 75th percentile means that 75 percent of all values are below that value. The whiskers extend from each end of the box to the respective extreme.

In some cases it is desirable to show some data values as farther out than others. H-spread is a term given to the differences between the hinges. A step is 1.5 times the box length (H-spread). An inner fence can be placed at one step outside the hinges; an outer fence is located at two steps outside the hinges.

The values located inside the inner fence, but closest to the inner fence are termed *adjacent*. Values between the inner and outer fences are termed *outside*. And values beyond the outer fences are *far out*.

The box-and-whisker plot is useful in conveying a concept of how even is the data above and below the median. In some cases the whisker may end at the adjacent values.

Boxplots are included in SAS® output using PROC UNIVARIATE PLOT (SAS® 1985). The boxplot for the algal data is shown in figure 3–3. Another boxplot from the output of JMP (SAS® Institute, Inc.) is given in figure 3–4.

The boxplots show that the data are highly skewed to the low values. The bottom and top of the box represent the 25th and 75th percentiles (hinges). The center horizontal line is drawn at the median, and a + is given at the mean (SAS® output). In the example, all these lines are so close that they are printed on the same line (fig. 3–3). The whiskers in SAS® extend to 1.5 the inter-quartile range (H-spread). Values more extreme, but within three interquartile ranges, are indicated with a zero. Values outside are indicated with an asterisk. For the example in figure 3–3, the three asterisks indicate three extreme values outside three interquartiles. The JMP outlier boxplot uses dots for points beyond the whiskers. The diamond indicates the 95 percent confidence intervals about the mean.

**Figure 3–2** Box-and-whisker plot

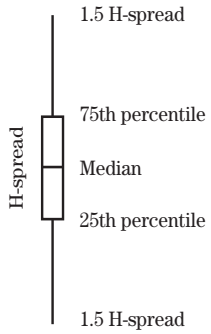



**Figure 3–3** Boxplot for the algal data from SAS® output

**Figure 3–4** Boxplot for the algal data from JMP output

# 615.0304 Transformations

Transformations of the data are sometimes needed to normalize the data or stabilize the variance. Transformations also change the appearance of the data into a form that may be more readily understandable (Tukey 1977). Some basic rules for different transformations have been described by Tukey (1977):

- Amounts and counts can never be less than zero, but can be large. A transformation may be useful if the ratio of the largest value to the smallest value is large (i.e., 100 or more). If the ratio is small (i.e., 1), the transformation will not modify the appearance of the data.
- Balances, values which can be both positive and negative, are usually not improved by transformations.
- Fractions and percentages may be better expressed with transformations.
- Grades, such as A, B, C, D, also may respond to complex transformations.

A common transformation for water quality data is the use of logarithms. The log distribution for concentration data makes sense because negative values do not exist, many values exist at lower concentrations, and a few values will exist at much higher concentrations (positively skewed). If plotted in a frequency diagram, the typical exponential decay curve results. Logs also are appropriate when the standard deviation in the data is likely to be proportional to the mean or for data that are proportional rather than additive on a linear scale (Snedecor and Cochran 1980, Sokal and Rohlf 1969). Logs tend to squeeze the data together and make it more symmetrical. A log transformation of zero does not exist; zeros can exist in a data set of mass export values. Also, when the data values are less than one, a log transformation gives negative numbers. In such cases the addition of a constant, such as log(X+1), is recommended (Steel and Torrie 1960, Zar 1984). However, the size of the constant added influences the estimate of the mean for the data set, as shown in example 3–1.

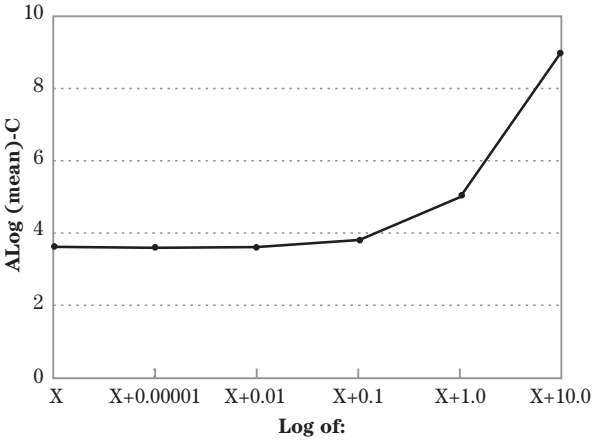
**Example 3–1** Log transformations with zero values

A $\log_{10}$ transformation was applied to the following values of X:

| | |
|---|---|
| 0.25 | 5.0 |
| 0.5 | 8.0 |
| 0.8 | 14.0 |
| 1.0 | 50.0 |
| 1.2 | 100.0 |

Additional transformations were made by adding the following constants: 10.0, 1.0, 0.1, 0.01, and 0.00001. The mean was obtained for each transformed data set as the antilog of the mean of the logged data minus the added constant. The results from these transformations are plotted in figure 3–5. These transformations indicate that adding smaller constants results in mean values that approach the true mean for the data set.

**Figure 3–5** The mean as a function of the size of the constant added in a log transformation



Counts, such as for bacteria data, can be re-expressed with logs and square roots, with root counts more often used (Tukey 1977). When the data numbers are small (<10) the square root transformation is recommended (Steel and Torrie 1960). If the counts are small, Snedecor and Cochran (1980) recommend the square root (X + 1) transformation.

Percentage data, based on counts, where the data range from 0 to 20 percent or 80 to 100 percent may be transformed with a square root (Steel and Torrie 1960). Percentage or decimal data based on binomial data can be re-expressed using an arc sine or inverse sine transformation.

For data that are skewed to the left, a value squared transformation has been recommended (Zar 1984).

Generally, when transformations are made, the mean is transformed back to the original scale, but variances or standard deviations should not be transformed back to the original scale (Steel and Torrie 1960).

Example 3–2 illustrates the transformation of the St. Albans Bay data from chapter 2.

**Example 3–2** Transformations

A $\log_{10}$ transformation was made of the St. Albans Bay algal data in table 2–1. The stem-and-leaf diagram for the transformed data indicates that the transformation removed much of the skewness in the algal count data, as compared to figure 3–1.

| Stem | Leaf | # |
|---|---|---|
| 4 | 5667 | 4 |
| 4 | 02 | 2 |
| 3 | 888 | 3 |
| 3 | 001233 | 6 |
| 2 | 56778999 | 8 |
| 2 | 1 | 1 |
| 1 | 7 | 1 |
| 1 | 4 | 1 |

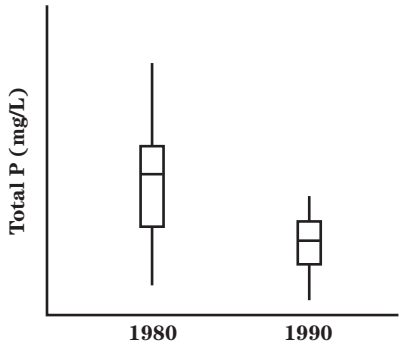
----+----

The box-and-whisker plot of the $\log_{10}$ transformed data also shows that the data are now more evenly distributed above and below the median as compared with figure 3–3. The absence of zeros and asterisks in the whiskers indicates that there are no values more extreme than three interquartile ranges. This shape is characteristic of a normal distribution.

## 615.0305 Comparisons

Different groups of data can be compared in several ways. They include side-by-side stem-and-leaf displays, tables of means or medians, and box-and-whisker plots. Transformations of scale often aid in the comparison among groups. For example, the box plots in figure 3–6 indicate that the phosphorus concentrations for 1990 were lower and less variable than for 1980. The width of the box can be used to reflect the sample size when comparing samples of different sizes (R.H. McCuen 1998, personal communication).

**Figure 3–6** Boxplots for two annual sets of phosphorus data



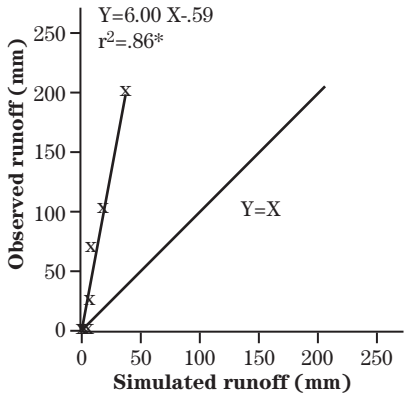

## 615.0306 Plots of relationship

Plots can be used to describe a relationship between a response variable (dependent) and a factor (independent) (Tukey 1977). The independent variable is usually shown as the abscissa (horizontal X-axis), and the dependent variable is shown as the ordinate (vertical Y-axis). Although default values in computer graphics programs make many decisions for us, there are some general rules that are useful in plotting relationships. These rules include guidance regarding the scale, shape, grid, and labeling of axis.

If comparing different plots with similar information, all plots should be at the same scale even though your graphics program may not default in this manner.

The shape of the plot is another important consideration. Plots can be taller than wide, wider than tall, and of equal dimensions. Taller than wide plots are useful for growth or decay phenomenon. Wider than tall plots facilitate reading from left to right and might be useful for scatter diagrams or time plots. Square plots may be useful in situations where the same units are plotted on each axis and the 45 degree line, representing Y = X, has some meaning. Figure 3–7 provides an example of such a graph—the comparison of observed data to data simulated by a model (Jamieson and Clausen 1988).

**Figure 3–7** Relationship of observed to predicted runoff (*=p=0.05)



The type of grid chosen for the graph influences the interpretation of the graph. Data that are extremely variable, such as suspended solids concentrations, might better be graphed using a log scale rather than a linear scale. Also, exponential relationships are straightened by plotting them on a log-log scale. If the data contain zero values, they cannot be plotted on a log scale unless a constant is added, as described in the previous section on transformations.

The labeling of axis, both in terms of the use of values and tick marks, influences interpretations from the graph. Generally, the number of tick marks and values shown on the graph are minimized because they can be distracting to the eye. An exception would be when the graph is used to pick off points. The origin of the graph, where Y = X, is generally zero to show the real magnitude of the values. This guidance is often abused in the media (e.g., stock market) to indicate larger variations than are really occurring.

One of the more common abuses of plots of relationship is termed *spurious correlation* (Kite 1989). This occurs when both axes have a variable in common. For example, a plot of mass export as a function of stream discharge is almost always guaranteed to show a positive relationship. This occurs because the values for the variable discharge are included in both axes. In regression, this would also violate the assumption of independence.

## 615.0307 Smoothing data

Smoothing data allow definition of general trends without looking at too much detail (Tukey 1977). Generally, the **Y** data are smoothed and the **X** data become intervals. Several techniques are used in smoothing. They include running medians or averages, eye smoothing, blurring, and splitting. An example of a water quality data set where smoothing might be useful is a time plot of concentration data (fig. 3–8).

The data as they appear are quite rough, and general trends are difficult to interpret. To use running medians or averages, take adjacent **Y** values and calculate a new smoothed point. Running implies that a central estimate is made for each point as opposed to creating intervals and deriving a central estimate for each interval. A running 3-day average of daily data is computed by determining the average of the days around Monday, then around Tuesday, and so on. For example, the median of three values running was used to develop the points for figure 3–9. The first three points were: 3, 7, and 10, which would result in a median of 7. New medians are calculated for the second set of three values by skipping the first number. In this case using the medians of a larger number of points may have provided a smoother picture of the data. The mean could be used rather than the median for smoothing. This method is also called the *moving average method.*

**Figure 3–8** Time plot of raw data

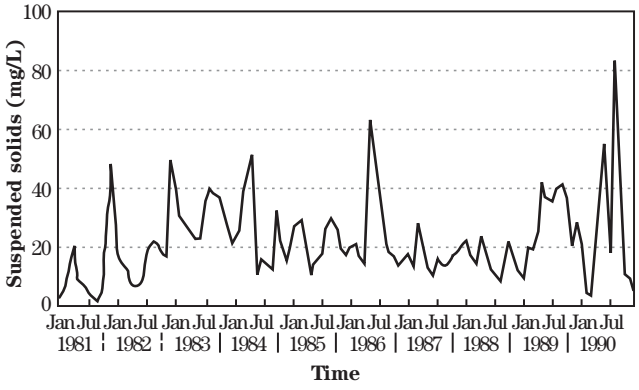
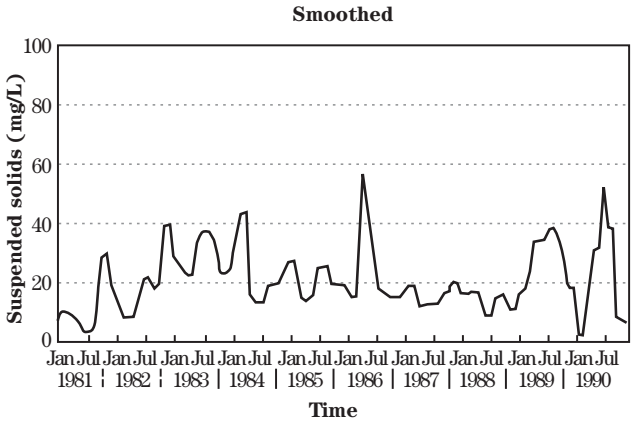



**Figure 3–9** Smoothed data using medians of three



*Eye smoothing* is drawing a smooth curve through the data. However, smoothing by eye allows bias to be used to meet the need or intent of the analyst. For example, figure 3–10 shows two curves fit to the smoothed data that are quite different from each other. Both lines are smoothing of the data. One line attempts to follow the peaks and valleys; the other suggests an even more general trend. This trend is contingent upon when monitoring began. Note that if the sampling began in 1983, a different trend might be suggested.

*Blurring* is a method of smoothing where the data points are replaced with vertical lines of some length showing their variability. In figure 3–11, the raw data have been blurred, which suggests a band of data rather than a line or a series of points.

**Figure 3–10** Smoothed data using the eye

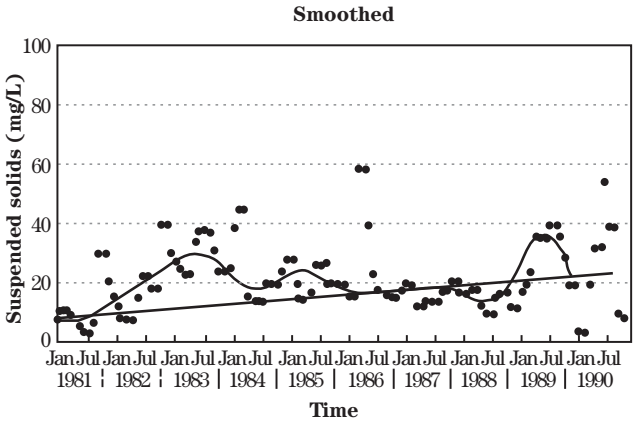
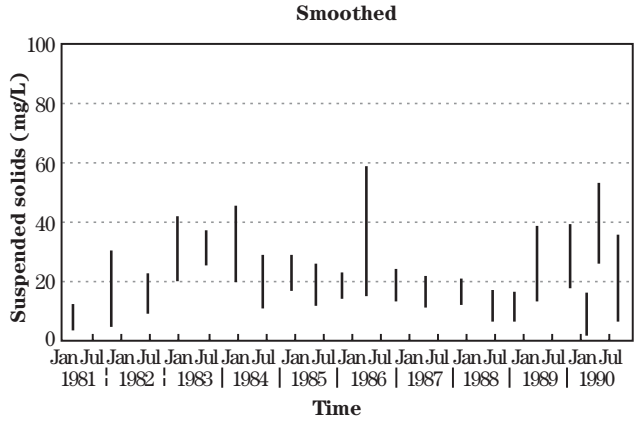



**Figure 3–11** Smoothed data using blurring




# 615.0308 References

Jamieson, C.A., and J.C. Clausen. 1988. Test of the CREAMS model on agricultural fields in Vermont. Water Resourc. Bul. 24(6):1219-1226.

Kite, G. 1989. Some statistical observations. Water Resourc. Bul. 25(3):483-490.

SAS Institute, Inc. 1985. SAS procedures guide for personal computers, Version 6 edition, SAS Institute, Inc., Cary, NC.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods. 7th Ed., The Iowa State Univ. Press, Ames, IA.

Sokal, R.R., and F.J. Rohlf. 1969. Introduction to biostatistics. W.H. Freeman and Co., San Francisco, CA.

Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York, NY.

Tukey, J.W. 1977. Exploratory data analysis. Addison-Wesley Publ. Co., Reading, MA.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 4 Statistical Assumptions

# Chapter 4 Statistical Assumptions

Contents:

# Chapter 4 Statistical Assumptions

## 615.0400 Introduction

When applying statistical analyses to water quality data, such as analysis of variance, we must be familiar with several underlying assumptions. It is important to know how to test if these assumptions have been violated and what to do if they are violated. All assumptions are difficult to meet exactly. It is more important to understand whether the violation of an assumption has a serious consequence on the probability statements made based on the assumption (Glass, et al. 1972). The main assumptions are: randomness, normality, homogeneity of variances, independence, and additivity.

Chapter 4 describes the various statistical assumptions made when performing statistical tests. The consequences of failing to exactly meet these assumptions are presented for each assumption. The usefulness of residual plots in evaluating assumptions is also detailed as is how to deal with missing data and extreme outliers.

## 615.0401 Assumptions

### (a) Randomness

The first assumption is that the water quality data are sampled randomly. *Randomness* means that the probability of obtaining a sample remains the same for all possible samples (Steel and Torrie 1960). The purpose of randomization is to design bias out of the study and increase the accuracy of the study (Hurlbert 1984). For example, if a stream was sampled only during stormflow periods, the study would be biased toward higher concentrations than if the stream were sampled mostly during low-flow periods. Water quality data have both random and deterministic components (Moser and Huibregtse 1976). Random components are introduced by precipitation events that are themselves random in most parts of the United States. Nonrandom components are related to trends or seasonality in the data (part 614, ch. 7, fig. 7–1).

Water quality samples may not be truly random for several reasons. Sampling is not randomized over all possible observations. For example, if sampling were done from 1980 to 1990, the sampling ignores what the water quality may have been for all time before 1980. By sampling within a shorter window than all time, there is a possibility that a nonrandom component is dominating water quality.

The lack of randomness may result in producing a lack of independence, heterogeneous variances, or nonnormal distributions (Sokal and Rohlf 1969). No specific test of randomness is available; however, proper design of the sampling program should ensure an appropriate level of randomness.

Sampling methods to maintain randomness are described in part 614 of this handbook in chapters 8 and 9.

## (b) Normal distribution

A second assumption is that the data come from a population with a particular frequency distribution of values, usually a normal distribution. Several methods are available for examining the normality of the data. They include graphical and statistical methods. The graphical approach is to plot the data in a cumulative frequency distribution. Normal data plot as a straight line on such a graph (fig. 4–1a). Data that are skewed (long tail) to the left have a cumulative frequency distribution that is concave upward (fig. 4–1b). Data that are skewed right have a cumulative frequency distribution that is concave downward (fig. 4–1c).

Within the Statistical Analysis System, a normal probability plot can be obtained from:

**PROC UNIVARIATE PLOT;**

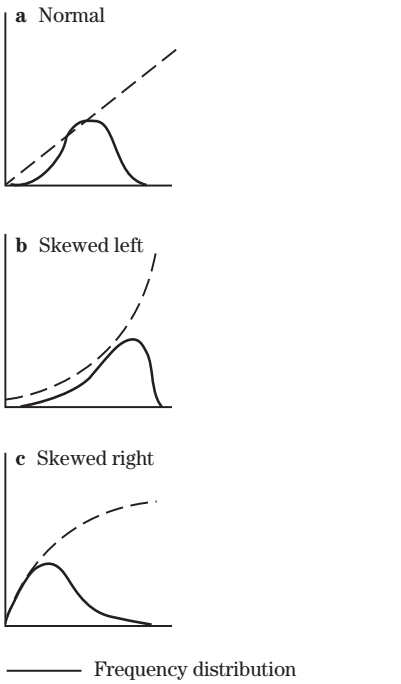
In addition to the normal probability plot, a stem-and-leaf plot and a boxplot are automatically produced.

Example 4–1 illustrates the cumulative frequency distributions for St. Albans Bay algal data in table 2–1 using the SAS® output. The log transformed data produces a straighter line on the normal probability plot than the untransformed data. This finding implies that the data follow a log normal distribution, and a log transformation should be used in subsequent statistical analysis.

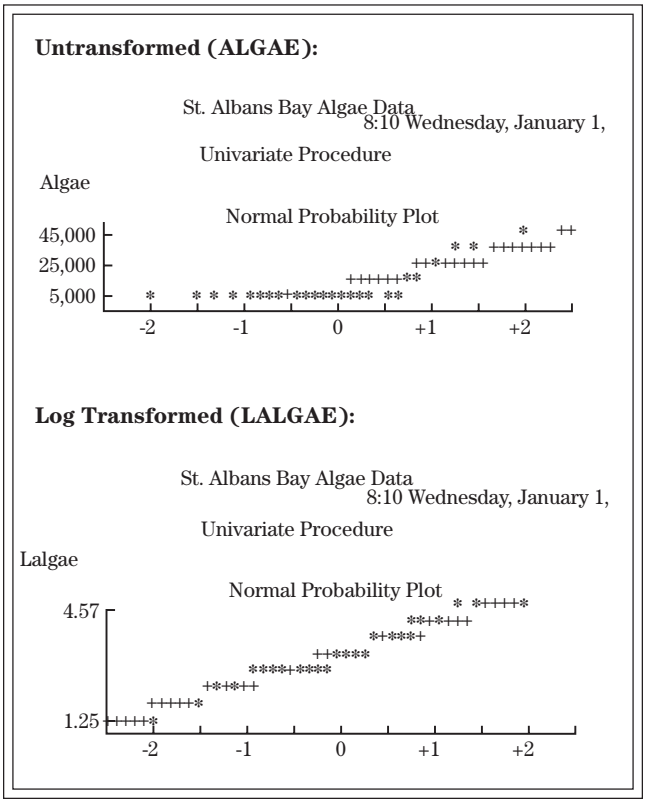
Among the statistical approaches for evaluating the normality of the data is the use of univariate statistics, such as the mean, median, skewness, and kurtosis. Generally, if the median and the mean are very different, the data may not be normally distributed. In addition, tests of either the skewness or the kurtosis will provide information regarding the normality of the distribution (see chapter 2).

Several statistical tests have been used for testing normality. One common test is the Chi-square goodness of fit (Snedecor and Cochran 1980; Sokal and

**Figure 4–1** Examples of frequency distributions




**Example 4–1** Cumulative frequency distributions for St. Albans Bay algal data from SAS® output



Rohlf 1969; Zar 1984). This tests the hypothesis that the sample came from a specific theoretical distribution.

The goodness of fit also may be tested using the Kolmogorov-Smirnov test (Zar 1984). Finally a test for normality can be accomplished by using the Shapiro-Wilk W-statistic. The W statistic has values ranging from 0 to 1; small values for W are significant and indicate nonnormality (Shapiro and Wilk 1965). The decision whether to use the Kolmogorov-Smirnov test is dependent on the sample size. For samples less than 2,000, the Shapiro-Wilk test should be used (SAS 1985). For larger samples, the Kolmogorov-Smirnov test should be used. Example 4–2 illustrates the test of normality using the Shapiro-Wilk W-statistic.

SAS® output provides the W statistic and its probability using the following command:

**PROC UNIVARIATE NORMAL;**

**Example 4–2** Test of normality for the St. Albans Bay algal data

| | Untransformed | Log transformed |
|---|---|---|
| Mean | 8075. | 3.234 |
| Median | 1168. | 3.063 |
| Skewness | 1.900 | –0.039 |
| Kurtosis | 2.334 | –0.389 |
| W:Normal | 0.626 | 0.959 |
| Prob<W | 0.0001 | 0.4018 |

For the St. Albans Bay algal data, the small W for the untransformed data indicates that the W is significant and nonnormal. The log-transformation of this data resulted in a large, nonsignificant W. The hypothesis that the data come from a normal distribution cannot be rejected. Therefore, the log transformed data are assumed to be normally distributed. Note also that the mean and median are closer and the skewness and kurtosis are smaller for the log transformed, as compared to the untransformed data.

Failure to exactly meet the assumption of normality is generally not considered to be a major problem (Glass, et al. 1972, Sokal and Rohlf 1969). The significance levels for t-tests and F-tests do not appear to be affected by nonnormality. That is to say that the probability of the Type I error is not increased significantly by failure to meet the assumption of normality (chapter 6). This is especially true for large data sets and when equal numbers of values are being compared. Skewed populations can affect the level of significance for one-tailed tests (Glass, et al. 1972). It is not considered necessary to use nonparametric approaches simply because the assumption of normality has not been exactly met. However, an appropriate transformation to better approximate normality is recommended.

## (c) Homogeneity of variances

In uses involving more than one data set, the equality of variances is an important assumption for several statistical tests. If there are two sample data sets that are being compared, the test of the homogeneity of variances is made by computing an F as the ratio between the larger variance divided by the smaller variance (Snedecor and Cochran 1980, Sokal and Rohlf 1969). The computed F is compared to a critical value for F from an F table (appendix C).

If three or more sample data sets are compared, Bartlett's test may be used. The ratio of the test statistic, B, to a correction factor is compared to the chi-square statistic (Snedecor and Cochran 1980, Zar 1984). For nonnormal distributions, some prefer the Levene's test for homogeneity of variances (Snedecor and Cochran 1980). The statistical program BMDP (Biomedical Computer Programs P-Series) computes both the Bartlett's test (BMDP9D) and the Levene's test (BMDP7D) (Dixon and Brown 1979).

A quick test is the $F_{max}$ test for which an F ratio is computed from $S^2_{max}/S^2_{min}$ and compared to a critical value for F that is given in various tables (appendix C) (Sokal and Rohlf 1969, Peterson and Hartley 1954).

The consequence of failing to meet the assumption of equal variances can be serious, especially when the sample sizes from the two groups are of unequal size (Glass, et al. 1972). When the sample sizes of the groups are equal, there is little effect on the probability

level of committing a Type I error (chapter 6). When the sample sizes of the groups being compared are unequal and the variances are heterogeneous, the probability of committing a Type I error may be seriously affected. The probability level may be underestimated when a smaller number of samples come from the more variable population. It may be overestimated when a smaller number of samples come from the less variable population (Glass, et al. 1972).

Transformations often help remove heterogeneous variances. If a transformation does not eliminate the problem, perhaps the data could be aggregated so that the number of samples among groups could be equalized. If this is not possible, a nonparametric approach may be desirable.

## (d) Independence

Another assumption is that the experimental errors are independently distributed (Sokal and Rohlf 1969, Steel and Torrie 1960). That is, if the data are arranged in some logical sequence, such as in the order of collection, the errors should follow each other randomly.

Randomization in sampling helps reduce the correlation of observations and their errors over time. This is a special concern in water quality sampling where high values are more likely to follow high values and low values follow low values.

If the errors are not independent, the F-test in analysis of variance (ANOVA) and the *t*-test results can be questioned. With positive serial correlations, the probability level of the Type I error is increased progressively with the size of the correlation. With negative correlations, the probability level of the Type I error is much lower than it really should be (Glass, et al. 1972).

If the sampled data are serially correlated, there is little that can be done. Randomization in the design of the experiment was insufficient. One alternative may be to aggregate the serial data in some logical manner, such as computing means or totals. For example, serially correlated weekly data could be aggregated to monthly data that may not be correlated. Another option would be to use Time Series Analysis (Vandaele 1983). This analysis assumes that the errors are not independent and are, in fact, correlated according to some time step. Although time series analysis has certain applications in water quality monitoring, such as trend analysis, it is a sophisticated statistical technique requiring special training.

Serial or auto correlation of the residuals can be determined from:

$$r_k = \frac{\sum_{t=1}^{N-k}\left(y_t - \bar{y}\right)\left(y_{t-k} - \bar{y}\right)}{\sum_{t=1}^{N}\left(y_t - \bar{y}\right)^2} \qquad [4–1]$$

where:
$r_k$ = autocorrelation coefficient for any lag k
y = observation at any time step t
N = total number of observations

In SAS® the autocorrelation coefficient may be obtained by:

```
PROC REG;
    MODEL Y=X / DW;
```

The DW stands for the Durbin-Watson d statistic that is a test of the hypothesis that autocorrelation is zero (SAS 1985).

## (e) Additivity

The assumption of additivity (also termed linearity) is normally applied to ANOVA and means that the effects of the treatment are additive, not multiplicative (Sokal and Rohlf 1969, Steel and Torrie 1970, Zar 1984). One way of viewing additivity is by writing the model for an ANOVA. A typical one-way ANOVA model would take the form:

$$\chi_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

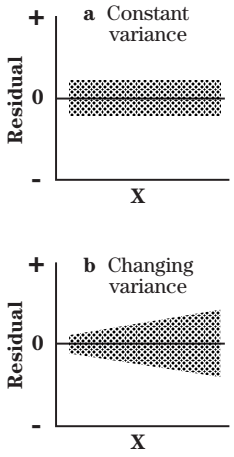
This equation states that an observed value ($\chi_{ij}$) equals the sum of an overall mean ($\mu$), a treatment deviation ($\alpha_I$), and a random error term ($\varepsilon_{ij}$) (Snedecor and Cochran 1980). The three factors in the equation are additive rather than multiplicative. Thus there would be no interaction in this particular model.

A test for nonadditivity has been suggested by Tukey (Snedecor and Cochran 1980). Log transformations of multiplicative effects promote additivity in the data.

## 615.0402 Residual plots

When using linear regression, an examination of a plot of the residuals, as a function of the independent variable, helps in the assessment of several of the assumptions including equal variances, independence, as well as the adequacy of the linear regression model (Afifi and Azen 1979, Draper and Smith 1981, Ponce 1980, Zar 1984). A residual is the deviation of a datum point from the regression line. For example, if the residuals are independent and of constant variance, then they should be scattered evenly about the horizontal line where the residual is zero (fig. 4–2a). If, on the other hand, the residuals appear to increase or decrease as X increases (fig. 4–2b), the variance may not be constant. A nonconstant variance implies that the regression model is inadequate.

**Figure 4–2** Residual plots for linear regression




## 615.0403 Missing data

Missing data are common in water quality sampling. Sometimes samples are missing because they were not collected. Possible reasons for not collecting samples include equipment failure, frozen conditions, or missing an event. Water samples that must be analyzed in a laboratory are subject to accidents or a quality assurance program that may render the sample as in error.

Missing data are important for some water quality monitoring designs, but not for all designs. Missing values may not be important for paired and unbalanced unpaired tests where the number of samples is adequate. The missing value merely eliminates a pair from the analysis and reduces the sample size. However, missing data may have important consequences on trend analysis.

As a cautionary note, the analyst must be aware of how missing data are coded when using computer statistical packages. Some packages read a blank as a zero. If a special value is used, such as –9, the computer may include that in calculations unless specifically informed otherwise. Each statistical package may have different requirements. SAS® for example recognizes a '.' as missing. One should also be aware that for some packages a missing value within a line (or case) may result in the elimination of the entire case.

Several techniques are used to estimate missing water quality data. They include linear interpolation, regression with another station or flow, and the use of several stations. In addition, more sophisticated measures are needed for missing blocks in randomized block designs (Snedecor and Cochran 1980, Zar 1984).

Linear interpolation uses the existing values adjacent to the missing value(s) and assumes that the missing value(s) is proportional to the difference between the known values.

For water quality data that are highly correlated to either other water quality data or flow data, missing values could be predicted using a regression equation. For missing flow data, a relationship with precipitation or with flow at a nearby station may provide an adequate predictor of the missing information.

Another approach is that several stations could be used to predict a single missing value if such data are available. For example, the concentration at a fourth station could be determined from the concentrations observed at three other stations and the means at all stations using the equation:

$$C_4 = \frac{1}{3}\left(\frac{\overline{C}_4}{\overline{C}_1}\times C_1 + \frac{\overline{C}_4}{\overline{C}_2}\times C_2 + \frac{\overline{C}_4}{\overline{C}_3}\times C_3\right) \qquad [4–2]$$

where:

$C$ = concentration at stations 1, 2, 3, and 4

$\overline{C}$ = mean for the respective station

## 615.0404 Extreme outliers

Water quality data sets generally contain values that appear to be extreme outliers. The initial response should be to verify that no mistake has been made in recording the observation. Upon occasion, an error has been made, but the true value cannot be determined. In this case the data could be declared missing.

Several methods are available for determining whether certain observations are outliers (e.g., Dunn and Clark 1987). For example, the maximum normed residual (MNR) can be calculated from:

$$MNR = \frac{Max\left|x - \overline{x}\right|}{\sqrt{\sum\left(x_1 - \overline{x}\right)^2}} \qquad [4–3]$$

where:

$x$ = outlier to be tested (Snedecor and Cochran 1980).

The calculated MNR is compared to a tabular MNR, which varies with the sample size and probability level. If the calculated MNR is less than the tabular MNR, the value is expected to occur more often than the probability level, and thus is not considered an extreme outlier.

## 615.0405 Summary

Table 4–1 provides a summary of the standard assumptions for parametric statistical tests and the appropriate methods for testing the assumption.

**Table 4–1** Statistical assumptions and tests

| Assumption | Test |
|---|---|
| Randomness | Sampling design |
| Normality | Graphical<br>Shapiro-Wilk<br>Kolmogorov-Smirnov |
| Equal variances | F ratio<br>Bartlett's<br>Levene's |
| Independence | Residual plot<br>Autocorrelation |
| Additivity | Tukey's |

## 615.0406 References

Afifi, A.A., and S.P. Azen. 1979. Statistical analysis: a computer oriented approach. (2nd ed). Academic Press. New York, NY.

Dixon, W.J., and M.B. Brown (Eds.). 1979. BMDP biomedical computer programs P-series. Univ. CA Press, Los Angeles, CA.

Draper, N.R., and H. Smith. 1981. Applied regression analysis. (2nd ed.) John Wiley and Sons. New York, NY.

Dunn, O.J., and V.A. Clark. 1987. Applied statistics: Analysis of variance and regression (2nd ed.) John Wiley and Sons, New York, NY.

Glass, G.V., P.D. Peckham, and J.R. Sanders. 1972. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. Rev. Educ. Research 42(3):237-288.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54(2):187-211.

Moser, J.H., and K.R. Huibregtse. 1976. Handbook for sampling and sample preservation of water and wastewater. U.S. EPA 600/4-76-049.

Peterson, E.S., and H.O. Hartley. 1954. Biometrika tables for statisticians. (Vol I.) Cambridge Univ. Press.

Ponce, S.L. 1980. Statistical methods used in water quality data analysis. WSDG Tech. Pap. WSDG-TP-00001. Watershed Systems Development Group, USDA Forest Service, Fort Collins, CO.

SAS Institute, Inc. 1985. SAS procedures guide for personal computers. (Ver. 6 ed.) Cary, NC.

Shapiro, S.S., and M.B. Wilk. 1965. An analysis of variance test for normality (complete samples). Biometrika 52:591-611.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods. (7th ed.) The IA State Univ. Press, Ames, IA.

Sokal, R.R., and F.J. Rohlf. 1969. Introduction to biostatistics. W.H. Freeman and Co., San Francisco, CA.

Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc. New York, NY.

Vandaele, W. 1983. Applied time series and Box-Jenkins models. Academic Press, Inc., New York, NY.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 5 Causality

# Chapter 5 Causality

Contents:



Table



Figures

# Chapter 5 Causality

## 615.0500 Introduction

Although the reasons for conducting water quality monitoring are varied (see part 614, chapter 1), many involve attempting to develop a cause-and-effect relationship between something that is done on the landscape (cause) and a response in water quality (effect). In statistical terms an experimental design is developed to determine the conclusion desired. An experimental design is a plan of the experimental units, treatments including a control, and the replications to achieve some objective. Four concepts provide a useful framework for designing water quality monitoring studies with causation in mind. These concepts are association, consistency, responsiveness, and mechanism (Mosteller and Tukey 1977).

This chapter describes these four concepts of causality. Examples are used to illustrate each of these requirements. Other features of designing experiments are also described.

## 615.0501 Association

An association between variables, such as water quality and land treatment, implies that these variables are paired in a related way across the population (Mosteller and Tukey 1977). An association is necessary, but not sufficient to show causality.

An association may be expressed in several ways including correlation and regression (Draper and Smith 1981) or a significance analysis of variance (ANOVA) model. Regression is appropriate when one variable is dependent on the other (Zar 1984). When two variables are associated, but one is not dependent upon the other, correlation analysis is used. For example, the association between runoff and rainfall is best analyzed by regression because runoff is dependent upon rainfall. However, the association between stream order and discharge is best explained by correlation. Discharge would be expected to be greater for higher order streams although there is no mathematical dependence of discharge on stream order.

Examples 5–1 and 5–2 help to illustrate the meaning of association.

**Example 5–1** Correlation

Water quality monitoring in the Jewett Brook watershed in Vermont revealed an association between stream discharge and various water quality variables (Hopkins and Clausen 1985). This association is represented by correlation coefficients of log-transformed data (table 5–1).

The correlations in table 5–1 do not necessarily imply dependence. Increased discharge may not cause increased concentrations in streamflow. Rather, other processes, for example snowmelt, can cause increases in both discharge and concentrations. Surely, increased stream concentrations do not cause increased discharge.

**Table 5–1** Correlations (r) between mean weekly discharge concentrations (mg/L) and discharge ($m^3/s$) n=52

| Variable | Correlation coefficient (r) |
|---|---|
| Total phosphorus | 0.37** |
| Total kjeldahl nitrogen | 0.44** |
| Total suspended solids | 0.61** |

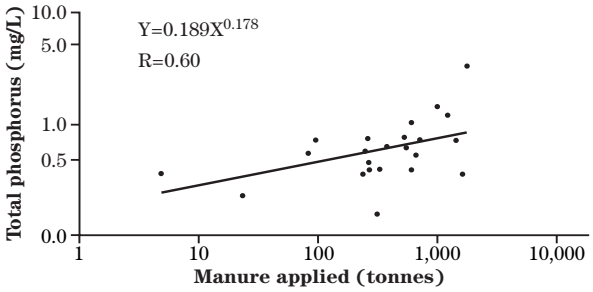
** Indicates p=0.01 (see chapter 6).

**Example 5–2** Regression

For the watershed described in example 5–1, land treatment data were also collected. These data included the amount of dairy cow manure applied in the watershed between each runoff event. A linear regression was developed between the concentration of total phosphorus in streamflow and the amount of manure applied in the watershed (fig. 5–1). This regression was significant based on analysis of variance for regression.

This association indicates that total phosphorus concentrations in the stream increase with increasing manure applications.

**Figure 5–1** Jewett Brook phosphorus concentration and manure applied in the watershed




# 615.0502 Consistency

Another requirement of causation is that the association between the variables is consistent from population to population in both direction and magnitude (Mosteller and Tukey 1977). To assess consistency, different data sets are needed of the same association. Consistency is shown in example 5–3.

**Example 5–3** Consistency

Figure 5–2 shows a relationship between either fecal coliform or fecal streptococcus abundance in Jewett Brook as a function of the percentage of the animal units in the watershed that are being managed with best management practices (BMPs). The major BMP used in this case was manure storage during the winter with spring manure spreading followed by rapid incorporation.

The association for fecal streptococcus was statistically significant, but the association for fecal coliform was not. Fecal coliform abundance appeared to be more variable than fecal streptococcus. To show consistency, compare this association to that derived from other data sets. Figures 5–3 through 5–5 show the association between bacteria abundance and the percent of animal units for three other watersheds in the same vicinity.

In all cases illustrated in figures 5–2 through 5–5, the bacteria abundance in the stream declined as the percentage of animal units being managed with BMPs increased. The same general relationship was observed in the LaPlatte River watershed about 50 miles away (Meals 1990). Ideally, this relationship should be tested across the United States to show consistency.

**Figure 5–2** Mean annual bacteria abundance and the percent of BMP animal units for the Jewett Brook watershed (n=6)

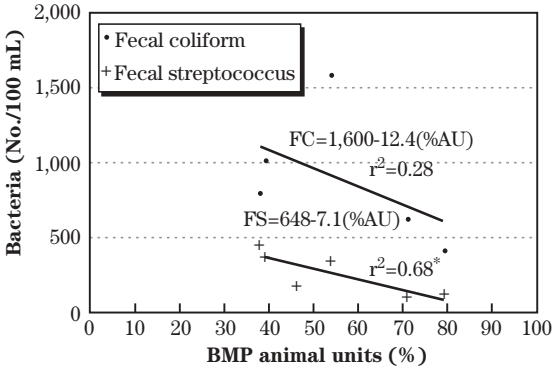
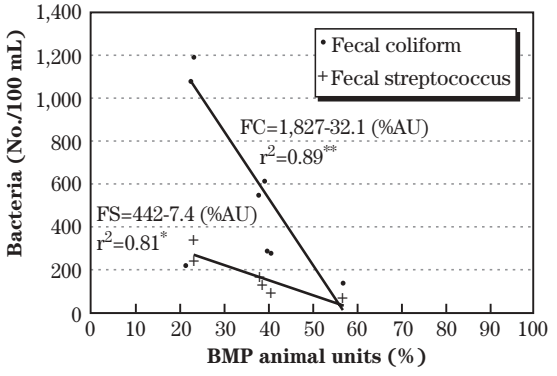



**Figure 5–3** Mean annual bacteria abundance and the percent of BMP animal units for the Stevens Brook watershed (n=6)




* Indicates p=0.05
** Indicates p=0.01
*** Indicates p=0.001

**Figure 5–4** Mean annual bacteria abundance and the percent of BMP animal units for the Rugg Brook watershed (n=6)

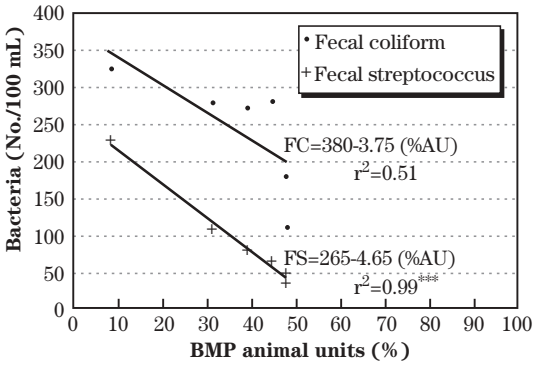
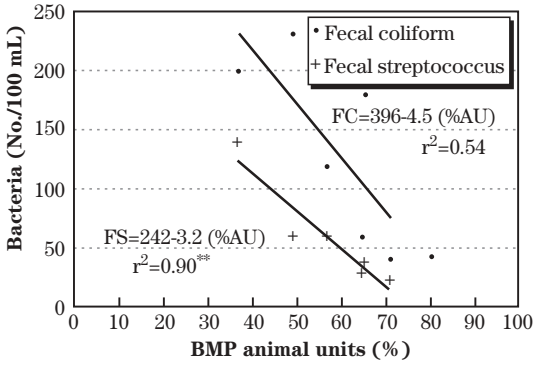



**Figure 5–5** Mean annual bacteria abundance and the percent of BMP animal units for the Mill River watershed (n=6)



## 615.0503 Responsiveness

Causality is also supported by the concept of responsiveness. By performing an experiment, the dependent variable should respond to manipulation of the independent variables (Mosteller and Tukey 1977). This concept requires that an experiment is performed where we intervene and change the x's and note whether the y's change in a corresponding manner. Example 5–4 illustrates this concept.

**Example 5–4** Responsiveness

For the bacteria example, we learned that the bacteria abundance in the streams draining agricultural watersheds was associated to the percent of the BMP animal units. The percent of BMP animal units is actually a surrogate variable for changes that occur in the management of bacteria from animal wastes. Included in these changes are longer storage of manure and incorporation of the manure soon after field application.

At a farm in the St. Albans Bay watershed, a paired watershed study was conducted at a field scale to determine the effect of best manure management on bacteria in runoff. During the calibration period both fields were spread with manure on top of ice and snow during the winter. During the treatment period, the upper field received manure in the winter again, but the lower field was spread with manure in the spring, which was immediately incorporated into the soil. This experiment could determine the change in bacteria abundance in runoff that resulted from the BMP of storage and incorporation. Bacteria abundance in runoff should respond to the application of manure on frozen ground.

## 615.0504 Mechanism

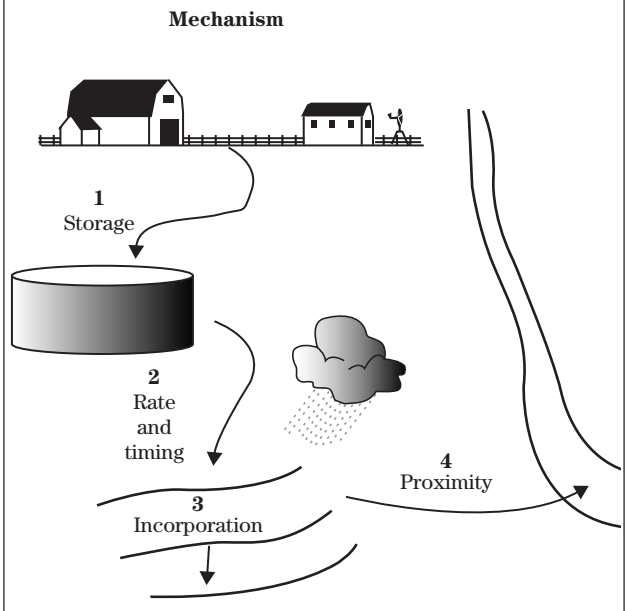
The final requirement for causality is adequate description of a mechanism that provides a step-by-step pathway from the cause to the effect, making the appropriate linkages along the way (Mosteller and Tukey 1977). Example 5–5 illustrates this point.

**Example 5–5** Mechanism

Figure 5–6 shows a logical mechanism that explains why bacteria abundance in the example stream may decline after the animal units begin to be managed.

Bacteria would have a tendency to die off, or otherwise decline in abundance, at several points along the pathway. First, bacteria would die off in storage in the manure pit or tank faster than in piled manure (Moore, et al. 1988). Second, the amount and timing of manure applied would be managed based on soil and crop needs. Third, much less manure would be available for runoff if it were incorporated into the soil. Fourth, manure would be applied at a safe distance from the stream off runoff-producing zones. All of these factors should contribute to lower bacteria abundance in streams draining agricultural watersheds that have animal waste BMPs.

**Figure 5–6** Mechanism for bacteria decreases




# 615.0505 Experimental design

Other considerations in analyzing cause and effect depend, in large part, on how the monitoring study is conducted. These factors include the time scale, system level, and reasonableness of treatment.

## (a) Time scale

The time scale is important for causality because we all investigate windows within the continuum of time. Numerous temporal cycles, such as diurnal, lunar, seasonal, annual, and astronomical, operate in the natural environment. All these cycles have the potential of influencing our perception of causality. These time scales also influence interpretation of trend data. The timing of flow occurrences during a study can influence our perception of water quality trends. For example, if a wet year occurred early in the study, flow, concentrations, and mass exports would be high during that year. If that year were followed by several years of lower flows, a decreasing trend in flow, concentrations, and mass would be likely.

To avoid or account for problems associated with time scales, the true natural variability must be determined before treatments are imposed. The response observed may be an increase in the variability rather than a change in the mean. Reference watersheds (controls) help account for time scale problems. The experimental design must consider time scale cycles.

## (b) System level

Biological systems can be studied at the ecosystem, community, population, individual, cell, and molecular level. Similarly, watersheds (catchments, drainage basins) can be investigated at the watershed, field, and plot level. Because the lower levels of systems are inherently easier to investigate, the tendency is to investigate at a lower level than is needed to answer the question. For example, interest is high in knowing the effect of implementation of BMPs in a watershed on water quality. However, the common approach to investigating cause-and-effect is to look at the effectiveness of an individual BMP on a field or plot basis.

This approach ignores processes that operate on a watershed basis, such as stream transport phenomenon. The project scale should be matched with the objective to avoid misconceptions about the system level being studied.

### (c) Reasonableness of treatment

When studying causality, the type of treatment applied should be reasonable and consistent with real world situations. Some treatments may be strong interventions, such as a catastrophe. An example of such treatment is the clearcut and herbicide treatment at Hubbard Brook Experimental Forest (Likens, et al. 1970). Following harvesting, the timber was left on the site and regrowth was prevented with herbicide applications. Stream concentrations increased dramatically in nitrate and cations. By comparison other treatments can be more gradual, such as a change in nutrient management on an agricultural field.

The interaction of the treatment with the environment may be more important than the main effect of the treatment. For example, certain erosion control practices may show no effect during small storms, but may be very effective during the larger, rarer storm events.

A final consideration in causality is understanding the number of variables contributing to a dependent variable. Most water quality issues are multivariate and not univariate. For example, stream phosphorus concentrations may be influenced by precipitation, antecedent moisture, previous stream loading of phosphorus, biological activity, temperature, geologic formation, land activities, and the time available for mineralization. Thus the cause of the level of phosphorus in a stream is potentially the effect of numerous factors that could be considered in the design of the study.

Some causal variables could be unexpected interferences. For example, the midnight dumping of septage, an accidental spill, or routine washing practices at a small point source can create havoc with an experimental design.

## 615.0506 References

Draper, N.R., and H. Smith. 1981. Applied regression analysis. (2nd. ed.) John Wiley and Sons, New York, NY.

Hopkins, R.B., and J.C. Clausen. 1985. Land use monitoring and assessment for nonpoint source pollution control. *In* Perspectives on Nonpoint Source Pollution, U.S. EPA 440/5-85-001, p. 25-29.

Likens, G.E., F.H. Bormann, N.M. Johnson, D.W. Fisher, R.S. Pierce. 1970. Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed-ecosystem. Ecological Monographs 40(1):23-47.

Meals, D.W. 1990. LaPlatte River watershed water quality monitoring and analysis program comprehensive final report. Prog. Rep. No. 12, VT Water Resourc. Res. Ctr., Univ. VT, Burlington.

Moore, J.A., J. Smyth, S. Baker, and J.R. Miner. 1988. Evaluating coliform concentrations in runoff from various animal waste management systems. Agric. Exp. Sta. Spec. Rep. 817, OR State Univ., Corvallis.

Mosteller, F., and J.W. Tukey. 1977. Data analysis and regression: a second course in statistics. Addison-Wesley Publ. Co., Reading, MA.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 6 Hypothesis Testing

(450–VI–NWQH, September 2003)

# Chapter 6 Hypothesis Testing

Contents:

**Examples**

# Chapter 6 Hypothesis Testing

## 615.0600 Introduction

Developing a hypothesis and testing that hypothesis are fundamental steps in data analysis for water quality monitoring studies. A *hypothesis* is a scientific statement about an assumption regarding the results expected from a study. A *statistical hypothesis* is a statement about a variable describing the distribution of the data, such as the mean (Snedecor and Cochran 1980, Steel and Torrie 1960, Zar 1984). Hypotheses are statements regarding population parameters, not sample statistics. We use hypotheses to draw inferences regarding the assumed population based on sample information. A test of a hypothesis, also termed a test of significance, is a procedure for determining whether a hypothesis should be rejected or accepted (Afifi and Azen 1979).

A *null hypothesis* is the primary hypothesis to be tested and is so termed because it is the hypothesis of no change. The null hypothesis is noted by $H_o$. Generally, rejecting the null hypothesis is desirable. An example of a null hypothesis is:

$H_o$: mean (year 1) = mean (year 2)

This seemingly reverse logic exists because data can be collected that can contradict the null hypothesis, but data cannot be obtained to directly accept the hypothesis.

An *alternative hypothesis*, denoted by $H_a$, is often the hypothesis of interest and is the statement that we may want to assume is true. An example of an alternative hypothesis is:

$H_a$: mean (year 1) • mean (year 2)

or possibly:

$H_a$: mean (year 1) < mean (year 2)

The various types of hypotheses used in water quality studies are described in this chapter. In addition, the consequences of making incorrect hypothesis decisions (error types) and the meaning of statistical significance are described.

## 615.0601 Error types

When performing a statistical test of a hypothesis, the decision can be wrong because probability, or chance, is involved. Two types of errors can occur. A Type I error can occur when the $H_o$ is rejected even though it is true (table 6–1). The probability of a Type I error is indicated by $\alpha$, which is usually a small value that should be decided before the study begins (Steel and Torrie 1960, Zar 1984). This is also termed the statistical significance of the study. Conversely, accepting the null hypothesis when it is true (a correct decision) has the probability of $1–\alpha$, which should be a high value.

A Type II error can occur when the $H_o$ is not rejected when it should be (table 6–1). The probability of a Type II error is indicated by $\beta$. Conversely, the probability of rejecting the null hypothesis when it is false has the probability of $1–\beta$, which is also called the power of the test (Steel and Torrie 1960, Zar 1984).

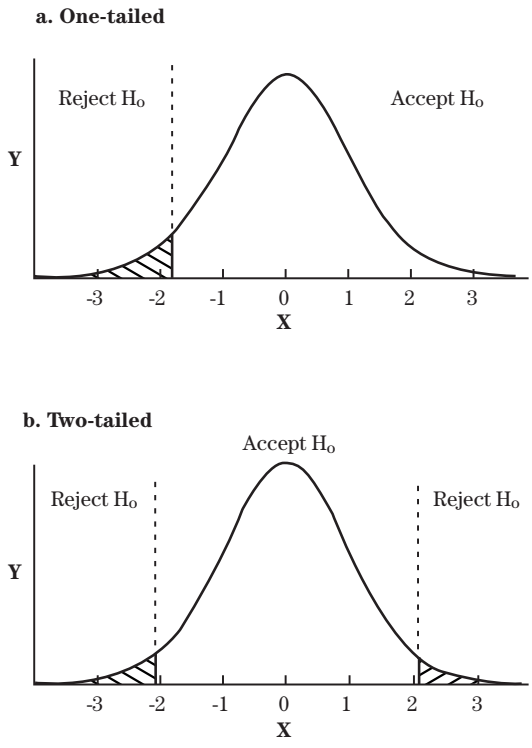
For a given number of samples, $\alpha$ is inversely related to $\beta$. This means that if we reduce the probability of rejecting the null hypothesis when it is true ($\alpha$), we increase the probability of accepting the null hypothesis when it is false ($\beta$). Both types of errors can be reduced by larger sample sizes.

Hypotheses will be used throughout the various chapters contained herein. However, some common hypotheses used and their appropriate applications are described further. Hypotheses may be categorized by the number of groups being compared. They are often distinguished as one-sample, two-sample, paired-sample, and multisample (Zar 1984).

**Table 6–1** Error types in statistical decisions

| Decision | Reality | |
|---|---|---|
| | $H_o$ is true | $H_o$ is false |
| Reject $H_o$ | Type I error<br>Prob = α termed<br>*significance level* | Correct decision<br>Prob = 1– β termed<br>*power* |
| Accept $H_o$ | Correct decision<br>Prob = 1–α termed<br>*confidence level* | Type II error<br>Prob = β |

**Figure 6–1** Distribution of *t* showing critical regions




# 615.0602 One-sample hypotheses

A test involving one sample is used when a population parameter (e.g., the mean) is compared to a fixed value that may either be known or hypothesized. Tests can be either one-tailed or two-tailed, depending upon the nature of the problem. These tests are termed one- or two-tailed because they refer to a comparison of a calculated *t* to a critical region of the *t*-distribution at a certain probability. A one-tailed test is used when the mean is to be compared to a fixed value, such as a water quality standard. A two-tailed test is used when the mean could lie on either side of a fixed value.

In figure 6–1a the *t*-distribution is shown for a one-tailed test. If the calculated *t* is greater than the critical *t* (see chapter 8 for a definition of *t*), the null hypothesis can be rejected at the probability used. This means that the mean is so different from the fixed value that it lies in the shaded area and has a very small probability of occurring if it were part of the fixed value's population.

The *t* distribution is used rather than the *z* distribution because the population standard deviation (σ) is unknown.

## (a) One-tailed

A one-tailed test is appropriate when the mean or some other population parameter is to be compared to some fixed value in a specific direction, such as a water quality standard (Snedecor and Cochran 1980, Zar 1984). We may test that the value is either significantly larger or significantly smaller than the fixed value, but we can only test one direction at a time. See example 6–1 for more information.

## (b) Two-tailed

A two-tailed test is appropriate when there is no reason to see whether a value is greater than or less than a fixed value. Therefore, an appropriate null hypothesis would be that the means are identical, and the alternative hypothesis would be that the means are

not equal (Steel and Torrie 1960, Zar 1984). In figure 6–1b, the distribution for $t$ is shown for a two-tailed test. In this case the calculated $t$ can be either positive or negative.

In some cases the appropriate value to compare to the mean might be a zero. This may happen when examining the change in something, such as the change in concentrations before and after some time period. See example 6–2 for more information.

**Example 6–1** One-sample hypothesis testing—one-tailed

Implementation of a nutrient management program on cropped fields might be expected to result in reduced ground water $NO_3$-N concentrations below the standard of 10 mg/L. An appropriate null hypothesis would be:

$H_o$: mean $NO_3$–N ≥ 10 mg/L

The alternative hypothesis might be:

$H_a$: mean $NO_3$–N < 10 mg/L

In this case it is desirable to reject the null hypothesis in favor of the alternative hypothesis.

**Example 6–2** One-sample hypothesis testing—two-tailed

When sampling the ground water in a field, we may be uncertain as to whether the $NO_3$–N in the ground water is improving or getting worse over time. An appropriate null hypothesis may be:

$H_o$: mean (year 1) = mean (year 2)

The alternative hypothesis would be:

$H_a$: mean (year 1) ≠ mean (year 2)

A two-tailed $t$-test would be appropriate to test these hypotheses. If the calculated $t$-value was greater than the critical value from a table, then the null hypothesis would be rejected. This $t$-value could be either positive or negative.

## 615.0603 Two-sample hypotheses

A two-sample hypothesis is used when testing for the differences between two populations sampled. Often we are testing for the difference between two means; although the two variances could be tested as well. Both one-tailed (example 6–3) and two-tailed (example 6–4) tests are appropriate for two-sample hypothesis; however, the two-tailed test is more commonly used.

**Example 6–3** Two-sample hypothesis testing—one-tailed

The nutrient management program described in Example 6–1 could result in a reduced mean concentration of nitrogen in a stream draining the treated watershed. Thus we are interested in detecting a difference in one direction only. The null hypothesis might be:

$H_o$: mean (year 2) > mean (year 1)

The alternative hypothesis might be:

$H_a$: mean (year 2) ≤ mean (year 1)

If we were less certain about the years, this could be a two-tailed test.

**Example 6–4** Two-sample hypotheses testing—two tailed

For long-term trend analysis we may not be certain as to whether the change from year 1 to year 2 might be an increase or a decrease. An appropriate null hypothesis might be:

$H_o$: mean (year 1) = mean (year 2)

The alternative hypothesis might be:

$H_a$: mean (year 1) ≠ mean (year 2)

These hypotheses could also be stated in terms of their differences. The null hypothesis would be:

$H_o$: mean (year 1) – mean (year 2) = 0

and the alternative hypothesis would be:

$H_a$: mean (year 1) – mean (year 2) ≠ 0

A *t*-test would be used to test the null hypothesis (chapter 8).

## 615.0604 Paired sample hypotheses

A paired sample hypothesis is appropriate when two samples are associated in some meaningful way. The two-sample hypotheses, described in the previous section, assume that the samples are independent and not associated in some way. For example, comparing the means of monthly observations from one year to the next would be a two-sample test. Months are not paired well from year to year because of climate differences. However, comparing the means of monthly observations from adjacent watersheds for the same year would be a paired sample test. The two adjacent watersheds would be similarly affected by climate from month to month during the year. The paired *t*-test is used to test the null hypothesis (chapter 9). Both the one-tailed and two-tailed hypotheses are used with paired comparisons. These tests are illustrated in examples 6–5 and 6–6.

The hypotheses for paired samples are expressed in several ways. One method is to assume that the difference between the means is zero. This is equivalent to stating that the means are equal.

**Example 6–5** Paired sample hypotheses testing—one tailed

An erosion control irrigation study was established to determine whether the newer sprinkler irrigation technique results in more than a 1 ton per acre reduction in erosion compared to the older flooded irrigation. To answer the question, paired plots were established with one plot from each pair being irrigated with a sprinkler and the other flooded. An appropriate null hypothesis is:

$H_o$: mean (sprink.)–mean (flood)=1 ton/acre reduction

The alternative hypothesis is:

$H_a$: mean (sprink)–mean (flood)>1 ton/acre reduction

We only wanted to know whether the change in irrigation practice was going to result in less erosion, so a one-tailed test was used.

**Example 6–6** Paired sample hypotheses testing—two tailed

For the above-and-below watershed design, samples collected at the above and below stations are associated because of the sampling time; therefore, they should be paired. An appropriate null hypothesis is:

$H_o$: mean (Lower) – mean (Upper) = 0

The alternative hypothesis would be:

$H_a$: mean (Lower) – mean (Upper) ≠ 0

The paired *t*-test would be used to test the null hypothesis (chapter 8).

## 615.0605 Multisample hypotheses

A multisample hypothesis is used when sampling is from three or more groups. The number of samples taken from each group is not required to be of equal size (unbalanced design). However, equal numbers of samples per group (balanced design) enhance the chance of rejecting the null hypothesis statistically. Example 6–7 illustrates a multisample hypothesis. If two samples are taken, either the *t*-test or ANOVA can be used because they yield identical results.

**Example 6–7** Multisample hypotheses testing

For a trend study being conducted over several years, we may be interested in comparing annual means. An appropriate null hypothesis might be:

$H_o$: mean (year 1)=mean (year 2)=…mean (year k)

The alternative hypothesis might be:

$H_a$: mean (year 1)≠mean (year 2)≠…mean (year k)

Analysis of variance (ANOVA) is used to test the null hypothesis (chapter 11) using the F-statistic. The test indicates whether all of the population means are different, but not which of those means are different. To answer this question, a multiple comparison test is needed (chapter 11).

## 615.0606 Nonparametric hypotheses

Nonparametric or distribution-free tests have the advantage that they do not assume that the populations are normal or have equal variances (Zar 1984). Nonparametric tests could be used in most cases where a parametric test may be used. A parametric test is better to use than a nonparametric test because it has greater power; that is, the probability of rejecting the null hypothesis is higher when it is false. A greater probability of a Type II error occurs when using nonparametric approaches. Nonparametric tests are described in detail in subsequent chapters.

Most nonparametric approaches require that the data be ranked from either lowest to highest or highest to lowest, and values are assigned the rank of 1, 2, and so forth. The rank, rather than the actual value, becomes the basis of comparison. Ranking eliminates the impact of outlyers in the tail regions of distributions.

The actual hypotheses stated will be the same as previously described; however, a nonparametric statistic is used to test the null hypothesis.

## 615.0607 Statistical significance

The significance level is the probability of committing a Type I error and is denoted as $\alpha$. By convention, an $\alpha$ of 0.05 is used because it is considered to be a small chance of committing a Type I error. However, in some cases an $\alpha$ of 0.01 is used. The selection of the significance level is somewhat arbitrary. Reporting the level of significance helps the reader in making their own conclusions regarding significance (Zar 1984). The significance level should be decided when the null hypothesis is constructed. Because the significance level is affected by the sample size, a smaller $\alpha$ might be used for a smaller experiment (Steel and Torrie 1960).

The concept of biological significance has two meanings. The first meaning is that a much higher $\alpha$ is acceptable in biological systems because we simply cannot get any better. An $\alpha$ of 0.2 is sometimes acceptable. The second meaning of biological significance is related to the interpretation of results. For example, just because the negative correlation is significant between elevation and abundance of macroinvertebrates, does it mean that high elevation causes lower abundance? This relationship may not have biological significance even though it may have statistical significance.

# 615.0608 Summary

Table 6–2 provides a summary of the appropriate null hypotheses and statistical test for various data types. In most cases we are interested in comparing means. However, in some cases a comparison of variances may be of greater interest. For example, we may want to know if a particular water quality constituent has become less variable over time.

**Table 6–2** Summary of hypotheses by type of data and appropriate test

| Data type | Rejection region | Null hypothesis | Test |
|---|---|---|---|
| One-sample | one-tailed | $\bar{x} > x_o$ | *t* |
| | two-tailed | $\bar{x} > x_o$ | *t* |
| Two-sample | one-tailed | $\bar{x}_1 > \bar{x}_2$ | *t* |
| | two-tailed | $\bar{x}_1 = \bar{x}_2$ | *t* |
| | | $\sigma_1 - \sigma_2 = 0$ | F ratio |
| Paired-sample | one-tailed | $\bar{x}_1 - \bar{x}_2 \leq x_o$ | *t* |
| | two-tailed | $\bar{x}_1 - \bar{x}_2 = 0$ | *t* |
| | | $\sigma_1 - \sigma_2 = 0$ | F ratio |
| Multisample | | $\bar{x}_1 = \bar{x}_2 = \bar{x}_k$ | F |
| | | $\sigma_1 = \sigma_2 = \sigma_k$ | Bartlett's |

# 615.0609 References

Afifi, A.A., and S.P. Azen. 1979. Statistical analysis: a computer oriented approach (2nd ed). Academic Press. New York, NY.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods (7th ed). The Iowa State Univ. Press. Ames, IA.

Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York. NY.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 7 Plot Designs

(450–VI–NWQH, September 2003)

# Chapter 7 Plot Designs

**Contents:**



**Tables**

# Chapter 7 Plot designs

## 615.0700 Introduction

Plots are generally small areas that are replicated on the land or water. In a plot design, all plots are treated alike except for the factors under study. Data from a plot design are usually organized into multiple data sets corresponding to control plots and treatment plots. A further description of the plot design is in part 614, chapter 4, of the National Water Quality Handbook (NWQH).

The principal tool for the analysis of plot data is the analysis of variance (ANOVA) (Snedecor and Cochran 1980, Sokal and Rohlf 1969, Steel and Torrie 1960, Zar 1984). Normally, more than two plots are used for plot studies because the treatment applied is replicated. The ANOVA procedure is needed to test multisample hypotheses, such as whether the means of several treatments are different.

When designing a plot study, two of the important decisions are selecting the treatment(s) to be tested and the number of replications for each treatment. Also, the number of observations per plot and whether blocking will be used need to be determined.

This chapter describes the methods used to analyze plot data. Hand calculations and SAS® programs are used to illustrate the statistical methods. Examples of parametric and nonparametric statistics are provided. All possible plot designs are not covered in this chapter. A statistical textbook should be consulted for more complicated designs. These other designs are mentioned in this chapter.

## 615.0701 Replications

Replications in plot studies can be of two kinds:

- number of replications (plots) per treatment
- number of observations (samples or subsamples) per plot

### (a) Replications per treatment

One of the most important initial decisions in a plot study is to determine the number of replications of each treatment to use. Often this decision is based upon economic considerations, such as not enough funding to have more than two replications per treatment. However, such judgments often result in studies with insignificant findings. It is far better to simplify the number of treatments tested rather than sacrifice the number of replications. The number of replications per treatment that would be desired is a function of the variability in the data, the precision desired, and the type of sampling used, as further described in NWQH, part 614, chapter 9. Example 7–1 illustrates the selection of the number of replications per treatment.

## (b) Observations per plot

A second decision is to determine the number of observations per plot. This decision is partly controlled by the objective of the study. For example, only one annual export value can be obtained from a plot per year, but sampling the soil generally requires that several soil samples be obtained per plot. Having more than one replicate per plot modifies the ANOVA used. In a randomized block design (for an example see fig. 4–1, NWQH, part 614, chapter 4) an interaction term between treatment and block is added, which represents the experimental error. A within-plot sampling error is also determined. The ANOVA for a randomized block design is discussed in several introductory statistics textbooks. Blocking is used when the blocks are believed to be significant. For example, soil type changes across the experimental area could be blocked if soil types contribute to the variability observed in the data being measured. All treatments would be assigned to each block. Blocking is further described in section 615.0704, Two-way ANOVA.

**Example 7–1** Replications per treatment

A plot study is being planned to assess the effect of different N fertilizer treatments on the export of $NO_3$–N in water. For this example the export in surface water and ground water are combined into one number. The methods described in NWQH, part 614, chapter 9, are used for this calculation, especially equation 9–1, which is repeated here:

$$n = \frac{t^2 s^2}{d^2} \qquad [9–1]$$

A published study, similar to the one planned resulted in the following:

| | | |
|---|---|---|
| mean $NO_3$–N export | = | 59 kg/ha |
| standard deviation | = | 7.05 kg/ha |
| n | = | 5 |

The difference (d) for 10 percent from the mean would be:

d = 0.1 x 59 kg/ha = 5.9 kg/ha

To determine the number of samples needed to estimate the mean value within 10 percent of the true mean, two iterations of equation 9–1 are needed. The *t*-value would be 2.776 for n–1 degrees of freedom, where n=5 from the published study (appendix A). Using equation 9–1, the following number of replications needed is calculated:

First iteration

$$n = \frac{(2.776)^2 (7.05)^2}{(5.9)^2} = 11$$

For the second iteration the *t*-value is 2.228 (appendix A) at 11–1 = 10 degrees of freedom.

Second iteration

$$n = \frac{(2.228)^2 (7.05)^2}{(5.9)^2} = 8$$

Based on this previous study, eight replications of each treatment would be recommended to estimate a mean value within 10 percent of its true value. If only a 20 percent difference were used, n would equal two replications per treatment.

## 615.0702 Assumptions

Because ANOVA is being used to analyze the plot data, the assumptions associated with ANOVA must be considered. First, it is assumed that the treatments have been assigned randomly to the plots. ANOVA also assumes that the errors are normally distributed, are independent, and have a common variance. Tests to determine if the plot data meet these assumptions are described in detail in chapter 4. In cases where the data do not meet these assumptions, you should first try a transformation of the data (chapters 3 and 4). For example, a log transformation may convert a non-normal distribution to an approximate normal distribution. If the transformation still does not result in meeting the assumptions of the test, then you should consider the use of nonparametric statistics.

## 615.0703 One-way ANOVA

In a one-way classification we are interested in only the effect of one factor on the water quality variable. To design this type of study, each plot is assigned one of the treatments at random with approximately the same number of plots receiving each treatment. This type of design is also termed a *completely randomized design.* Example 7–2 provides the calculations used to perform a one-way ANOVA of data from this design. This example has one observation per plot.

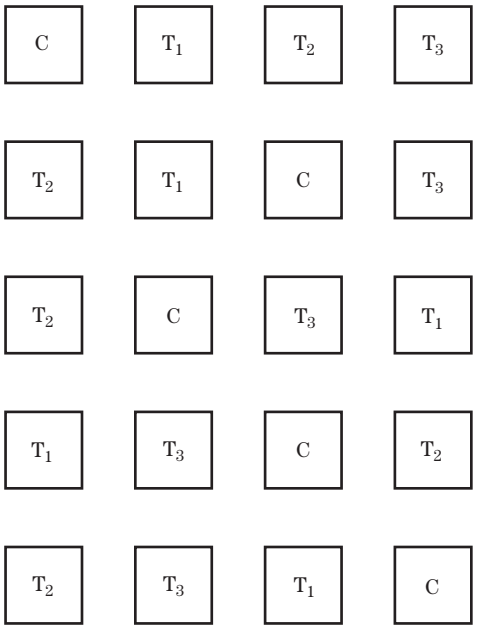
**Example 7–2** One-way ANOVA

A plot study was conducted to assess the effect of different N fertilizer treatments on the overall mass export of $NO_3$–N in water. The treatments included spring, split, and spring slow-release applications and a control plot with no fertilizer. There were five replications of each treatment. Figure 7–1 displays the plot layout and assignment of treatments. The data are summarized in table 7–1.

**Table 7–1** Annual $NO_3$–N export (kg/ha) from plots receiving different methods of N fertilizer applications

| | Control | Spring | Split | Slow release | $\Sigma X_j$= | $\Sigma X_j^2$= |
|---|---|---|---|---|---|---|
| Block 1 | 55 | 64 | 78 | 62 | 259 | 17,049 |
| Block 2 | 62 | 72 | 91 | 70 | 295 | 22,209 |
| Block 3 | 49 | 68 | 97 | 67 | 281 | 20,923 |
| Block 4 | 64 | 77 | 82 | 76 | 299 | 22,525 |
| Block 5 | 66 | 56 | 85 | 55 | 262 | 17,742 |
| $\Sigma X_i$ = | 296 | 337 | 433 | 330 | 1,396 | 100,448 |
| $\overline{X}$ = | 59 | 67 | 87 | 66 | | |
| $\Sigma X_i^2$ = | 17,722 | 22,969 | 37,723 | 22,034 | 100,448 | |

**Figure 7–1** Layout of plot design for fertilizer study



**Example 7–2** One-way ANOVA—Continued

The calculations for a one-way ANOVA are shown in table 7–2. The null hypothesis for this experiment would be:

$$H_o : \overline{X}_1 = \overline{X}_2 = \overline{X}_3 = \overline{X}_4$$

The alternative hypothesis is:

$$H_o : \overline{X}_1 \neq \overline{X}_2 \neq \overline{X}_3 \neq \overline{X}_4$$

For the calculations in table 7–2:
X = observation from table 7–1
i = ith treatment
j = jth replication
t = number of treatments
r = number of replicates per treatment

Hand calculations are shown in most beginning statistical books (e.g., Snedecor and Cochran 1980, Sokal and Rohlf 1969, Steel and Torrie 1960, Zar 1984). To perform these calculations by hand, initially determine $\Sigma X_i$, $\Sigma X_i^2$, and $\overline{X}$ for each treatment and overall (table 7–1). The additional calculations follow:

**Sums of squares**

Between treatments:

$$SS_{Bet} = \frac{\sum X_{ij}^2}{r} - \frac{\left(\sum X_{ij}\right)^2}{rt}$$
$$= \frac{296^2 + 337^2 + 433^2 + 330^2}{5} - \frac{(1{,}396)^2}{(5)(4)}$$
$$= 99{,}514.8 - 97{,}440.8 = 2.074$$

Total

$$SS_{Total} = \sum X_{ij}^2 - 97{,}440.8$$
$$= 100{,}448 - 97{,}440.8 = 3{,}007.2$$

Within treatment

$$SS_{Within} = SS_{Total} - SS_{Bet}$$
$$= 3{,}007.2 - 2{,}074 = 933.2$$

**Mean squares**

$$MS_{Bet} = \frac{SS_{Bet}}{df} = \frac{2{,}074}{4-1} = 691.333$$
$$MS_{Within} = \frac{SS_{Within}}{df} = \frac{933.2}{4(5-1)} = 58.325$$

**Table 7–2** One-way ANOVA

| Source of variation | Degrees of freedom | Sum of squares (SS) | Mean squares (MS) | F |
|---|---|---|---|---|
| Between treatments | t–1 | $\frac{\sum X_{ij}^2}{r} - \frac{\left(\sum X_{ij}\right)^2}{rt}$ | SS/df | $\frac{MS_{between}}{MS_{within}}$ |
| Within treatments | t(r–1) | by subtraction | SS/df | |
| Total | rt–1 | $\sum X_{ij}^2 - \frac{\left(\sum X_{ij}\right)^2}{rt}$ | | |

**Example 7–2** One-way ANOVA—Continued

### F-ratio

$$F = \frac{MS_{Bet}}{MS_{Within}} = \frac{691.333}{58.325} = 11.853$$

These calculations are summarized in table 7–3.

To determine whether to reject the null hypothesis of no difference between treatment means, the calculated F ratio is compared to the table F ratio for 3 and 16 degrees of freedom (appendix C). The table F is 3.24 and 5.29 for the 0.05 and 0.01 probability levels, respectively. Because the calculated F exceeds the table F, we can reject the null hypothesis with a 99 percent level of confidence. Therefore, the different fertilizer application methods most likely resulted in a difference in nitrate export. However, which treatments were different are not yet known. To determine which treatment means are different, the methods described in section 615.0707, Multiple mean comparisons, should be consulted.

Using SAS®, the appropriate program would be:

**SAS PC Program**

```
data nitrate;
        title 'ANOVA of Plot Data';
        infile 'a:nitrate.dat';
        input treat nitrate;
Proc ANOVA;
        class treat;
        model nitrate=treat;
run;
```

**Table 7–3** One-way ANOVA of fertilizer data

| Source of variation | Degrees of freedom | Sum of squares | Mean squares | F |
|---|---|---|---|---|
| Between | 3 | 2074.0 | 691.333 | 11.853 |
| Within | 16 | 933.2 | 58.325 | |
| **Total** | **19** | **3,007.2** | | |

# 615.0704 Two-way ANOVA

A two-way classification is useful when we are interested in the effect of two factors on the water quality variable. In plot studies, for example, plots that are adjacent to one another will have a tendency to give more similar results than plots located further away from each other. This may be because of some physical factor, such as soil heterogeneity. Another example would be if up slope plots have the potential to impact downslope plots. To account for this variability, the land can be subdivided into blocks of similar conditions. Blocks are sometimes referred to as replications.

When assigning treatments to the plots, they are assigned randomly within each block with a new randomization for each block. This type of design is termed a *randomized complete-block design*. The primary advantage of this design is that the variability contributed by field differences can be accounted for and eliminated from the treatment effect. Example 7–3 illustrates a two-way ANOVA.

**Example 7–3** Two-way ANOVA

For the N fertilizer experiment described in example 7–2, the plots were laid out in the field by placing them across four elevation transects (fig. 7–1). Treatments were randomly assigned to plots across each of the transects. In table 7–1, blocks are represented by rows. The calculations for a two-way ANOVA are shown in table 7–4.

**Sums of squares**

$$SS_{Blocks} = \frac{\sum X_{ij}^2}{t} - \frac{\left(\sum X_{ij}\right)^2}{rt}$$

$$= \frac{259^2 + 295^2 + 299^2 + 262^2}{4} - \frac{(1{,}396)^2}{(5)(4)}$$

$$= 97{,}778 - 97{,}440.8 = 337.2$$

**Table 7–4** Two-way ANOVA

| Source of variation | Degrees of freedom | Sum of squares (SS) | Mean squares (MS) | F |
|---|---|---|---|---|
| Blocks | r-1 | $\frac{\sum X_{ij}^2}{t} - \frac{\left(\sum X_{ij}\right)^2}{rt}$ | SS/df | $\frac{MS_{block}}{MS_{error}}$ |
| Treatments | t-1 | $\frac{\sum_j X_{ij}^2}{t} - \frac{\left(\sum X_{ij}\right)^2}{rt}$ | SS/df | $\frac{MS_{treatment}}{MS_{error}}$ |
| Error | (r-1)(t-1) | by subtraction | | |
| Total | rt-1 | $\sum X_{ij}^2 - \frac{\left(\sum X_{ij}\right)^2}{rt}$ | | |

Example 7–3 Two-way ANOVA—Continued

Treatments

$$SS_{Bet} = \frac{\sum X_{ij}^2}{r} - \frac{\left(\sum X_{ij}\right)^2}{rt}$$
$$= \frac{296^2 + 337^2 + 433^2 + 330^2}{5} - \frac{(1,396)^2}{(5)(4)}$$
$$= 99,514.8 - 97,440.8 = 2,074$$

Total

$$SS_{total} = \sum X_{ij}^2 - 97,440.8$$
$$= 100,448 - 97,440.8 = 3,007.2$$

Error

$$SS_{error} = SS_{total} - SS_{block} - SS_{Bet}$$
$$= 3,007.2 - 337.2 - 2,074 = 596$$

**Mean squares**

$$MS_{block} = \frac{SS_{block}}{df} = \frac{337.2}{5-1} = 84.3$$
$$MS_{Bet} = \frac{SS_{Bet}}{df} = \frac{2,074}{4-1} = 691.333$$
$$MS_{error} = \frac{SS_{error}}{df} = \frac{596}{(5-1)(4-1)} = 49.667$$

**F-ratio**

$$MS_{block} = \frac{SS_{block}}{df} = \frac{337.2}{5-1} = 84.3$$
$$F_{Block} = \frac{MS_{Block}}{MS_{error}} = \frac{84.3}{49.667} = 1.697$$
$$F_{treatment} = \frac{MS_{bet}}{MS_{within}} = \frac{691.333}{49.667} = 13.919$$

These calculations are summarized in table 7–5.

**Table 7–5** Two-way ANOVA of fertilizer data

| Source of variation | Degrees of freedom | Sum of squares | Mean squares | F |
|---|---|---|---|---|
| Blocks | 4 | 337.2 | 84.3 | 1.697 |
| Treatment | 3 | 2074.0 | 691.333 | 13.919 |
| Error | 12 | 596.0 | 49.667 | |
| **Total** | **19** | **3007.2** | | |

Based upon the ANOVA of the N fertilizer data, the block effect is not significant while the treatment effect was significant as before. A significant block effect would indicate that the design has been made more precise by blocking (Steel and Torrie 1960). Note that in the two-way ANOVA the error mean square has been reduced by apportioning some of the sums of squares to the block effect. This results in an overall higher treatment effect. If blocks are not different, they can be pooled into the error term, which results in an increase in the error degrees of freedom. However, in this example a higher significance was obtained with blocking than without it. Introductory statistical textbooks describe the calculation of the efficiency added by blocking.

Using SAS®, the appropriate program would be:

**SAS PC Program**

```
data nitrate;
        title 'ANOVA of Plot Data with Blocking';
        infile 'a:nitrate.dat';
        input block treat nitrate;
Proc ANOVA;
        class block treat;
        model nitrate = block treat;
run;
```

## 615.0705 Factorial

More complicated factorial design, split plot designs, and Latin squares are rare in water quality studies, but common in agronomic and soil investigations. An introductory statistics text should be consulted before planning one of these designs.

## 615.0706 Nonparametric ANOVA

If data are found to violate the assumptions of normality and especially homogeneous variances, a nonparametric approach may be used (Zar 1984). The Kruskal-Wallis test can be used for a one-way ANOVA. Other similar nonparametric tests, such as Friedman's, exist for a two-way ANOVA and more complicated designs. Because this test is based on rank rather than variance, the test statistic is determined from:

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1) \qquad [7–1]$$

where:

$n_i$ = number of observations in treatment i
$N$ = total number of observations
$R_i$ = sum of the ranks for each observation in treatment i

Observations are ranked from low (1) to high (N). The Kruskal-Wallis nonparametric ANOVA for the N fertilizer data is demonstrated in example 7–4.

**Example 7–4** Nonparametric ANOVA

For the N fertilizer data described in previous examples, determine the effects of the different fertilizer treatments on $NO_3$–N export using a nonparametric approach (see table 7–6).

If the calculated H is greater than the table H (appendix D, or $\chi^2$ for more than 5 groups) then the null hypothesis is rejected. In this case the table H is 5.78 at p = 0.05, and the null hypothesis that the nitrate exports are the same for each treatment is rejected.

$$H = \frac{12}{20(20+1)}\left[\frac{24^2}{5} + \frac{51^2}{5} + \frac{90^2}{5} + \frac{45^2}{5}\right] - 3(20+1)$$

$$H = 13.011$$

**Table 7–6** Annual $NO_3$–N export (kg/ha) from plots receiving different methods of N fertilizer applications (ranks are in parentheses)

| Control | Spring | Split | Slow release |
|---|---|---|---|
| 55 (2) | 64 (8) | 78 (16) | 62 (6) |
| 62 (5) | 72 (13) | 91 (19) | 70 (12) |
| 49 (1) | 68 (11) | 97 (20) | 67 (10) |
| 64 (7) | 77 (15) | 82 (17) | 76 (14) |
| 66 (9) | 56 (4) | 85 (18) | 55 (3) |
| R (24) | (51) | (90) | (45) |

## 615.0707 Multiple mean comparisons

From ANOVA we may have determined that the means are different; however, we do not know which of the means are statistically different from one another. Multiple comparison tests may be used to determine which of the means are different (Zar 1984). Although many such tests exist (e.g.; Duncan, LSD), the Tukey test is recommended for most cases and will be described further in example 7–5. The multiple comparison using the rank sums from the Kruskal-Wallis nonparametric ANOVA is described further in example 7–6.

## 615.0708 References

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods (7th ed.). The IA State Univ. Press, Ames.

Sokal, R.R., and F.J. Rohlf. 1969. Introduction to biostatistics. W.H. Freeman and Co., San Francisco, CA.

Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

**Example 7–5** Tukey multiple comparison test

For the N fertilizer data, it was determined that the mean $NO_3$–N exports were not equal. Using the Tukey multiple comparison test, determine for which groups the means are different.

First, the standard error is calculated from:

$$SE = \sqrt{\frac{S^2}{n}}$$

where:
SE = standard error
$S^2$ = variance (mean square error from the ANOVA)
n = number of observations per group

For the example without blocking, the standard error would be:

$$SE = \sqrt{\frac{58.325}{5}} = 3.415$$

Second, the means from table 05–1 should be arranged in increasing order and coded with a name or number, such as:

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| 59 | 66 | 67 | 87 |

The statistic q for each possible pair combination is calculated from:

$$q_\alpha = \frac{\overline{X_\beta} - \overline{X_A}}{SE}$$

If the calculated q is greater than the tabular q, the null hypothesis that the means are equal is rejected. The order of comparisons affects the conclusions. Therefore, the largest should be compared with the smallest first, then the second smallest and so on. The calculations are summarized in table 7–7.

The tabular q at rt–1 = 16 and k = 4 means degrees of freedom and p = 0.05 is 4.05 (appendix E). Therefore, group 4 is different from groups 1, 2, and 3, but no other groups are different. These results can be displayed by drawing a line under the groups that are not different, as shown above. More often the

**Example 7–5** Tukey multiple comparison test—Continued

means are listed in a table with letters following them and a notation that the means followed by the same letter are not different at p = 0.05, as follows:

| Treatment | $NO_3$–N Export (kg/ha) |
|---|---|
| Control | 59 a |
| Spring | 67 a |
| Split | 87 b |
| Slow release | 66 a |

The conclusion for this study would be that the split application resulted in significantly higher $NO_3$–N export from the plots than all other treatments, including the control.

Using SAS®, the following statement could be added below the Proc ANOVA statement. The mean values will also be printed.

**means tukey;**

**Table 7–7** Tukey's multiple comparison test of the N fertilizer data

| Comparison | Difference | q |
|---|---|---|
| 4 vs 1 | 87 – 59 = 28 | 8.20 |
| 4 vs 2 | 87 – 66 = 21 | 6.15 |
| 4 vs 3 | 87 – 67 = 20 | 5.86 |
| 3 vs 1 | 67 – 59 = 8 | 2.34 |
| 3 vs 2 | 67 – 66 = 1 | 0.29 |
| 2 vs 1 | 66 – 59 = 7 | 2.05 |

**Example 7–6** Multiple comparison using Kruskal-Wallas nonparametric ANOVA

A multiple comparison can also be made for a nonparametric ANOVA. The method is similar to that described for Tukey's in example 7–5, but uses the rank sums from the Kruskal-Wallis nonparametric ANOVA (Zar 1984). The standard error is determined from:

$$SE = \sqrt{\frac{n(nk)(nk+1)}{12}} = 13.23$$

where:
n = number of observations per k groups (Zar 1984)

The rank sums from the table 7–6, rather than the means, are used for arranging the data:

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| 24 | 45 | 51 | 90 |

The q statistic is determined as before. Table 7–8 displays nonparametric multiple comparison test of the N fertilizer data. In this case only the split treatment was higher than the control. There was no difference among all other treatments.

**Table 7–8** Nonparametric multiple comparison test of the N fertilizer data

| Comparison | Difference | q |
|---|---|---|
| 4 vs 1 | 90 – 24 = 66 | 4.99 |
| 4 vs 2 | 90 – 45 = 45 | 3.40 |
| 4 vs 3 | 90 – 51 = 39 | 2.95 |
| 3 vs 1 | 51 – 24 = 27 | 2.04 |
| 3 vs 2 | 51 – 45 = 6 | 0.45 |
| 2 vs 1 | 45 – 24 = 21 | 1.59 |

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 8 Single Watershed

(450–VI–NWQH, September 2003)

# Chapter 8 Single Watershed

Contents:

# Chapter 8 Single watershed

## 615.0800 Introduction

The single watershed design is used when a single station is monitored both before and after a watershed treatment occurs. As indicated in the National Water Quality Handbook (NWQH), part 614, chapter 4, the single watershed design is not recommended because any difference observed is difficult to attribute to the treatment rather than other influences that change over time, such as climate. However, the appropriate statistical approach when such a comparison is made is the unpaired *t*-test of pre and post data. This test actually determines the difference between the effects (Snedecor and Cochran 1980).

Comparisons between groups can be either paired or unpaired (independent). Paired comparisons occur when two samples can be paired in some meaningful way. For example, one pair could constitute an individual watershed measured before and after a treatment. However, in this case there is only one comparison and to make the test meaningful and valid, many watersheds (degrees of freedom) must be compared. It is generally not appropriate to pair observations, such as weekly or monthly data, from a single watershed across years. Because of climate variability, there is no reason to believe that the water quality of the 13th week or for July should be a valid pair across years. Therefore, the unpaired comparison is more common and is presented here.

## 615.0801 Unpaired comparison of means

The appropriate null hypothesis for the comparison of means is:

$$H_o : \overline{X}_1 - \overline{X}_2 = 0 \text{ or } \overline{X}_1 = \overline{X}_2$$

The appropriate alternative hypothesis would be:

$$H_a : \overline{X}_1 - \overline{X}_2 \neq 0 \text{ or } \overline{X}_1 \neq \overline{X}_2$$

The test of the significance of the difference between the means is based on the *t* distribution where *t* is defined as:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_d} \qquad [8–1]$$

where:

$\overline{X}$ = the mean for either group 1 or 2
$S_d$ = standard deviation of the difference between the means, which is determined from:

$$S_d = \sqrt{S_p^2 \frac{(n_1 + n_2)}{n_1 n_2}} \qquad [8–2]$$

for the case where $n_1 \neq n_2$ and is determined from:

$$S_d = \sqrt{2\frac{S_p^2}{n}} \qquad [8–3]$$

for the case $n_1 = n_2$.

$S_p^2$ is the pooled sample variance determined from:

$$S_p^2 = \frac{\left[\sum X_1^2 - \frac{\left(\sum X_1\right)^2}{n_1}\right] + \left[\sum X_2^2 - \frac{\left(\sum X_2\right)^2}{n_2}\right]}{(n_1 - 1) + (n_2 - 1)} \qquad [8–4]$$

where:

$S_p$ = pooled standard deviation

$S_p$ is calculated by pooling the individual standard deviations as calculated from equation 2–6 (Steel and Torrie 1960, Zar 1984). The *t*-test is appropriate when

the distributions are normally distributed and have equal population variances. Example 8–1 illustrates the analysis of a single watershed.

**Example 8–1** Single watershed analysis

Table 8–1 presents a summary of total phosphorus concentrations in watershed runoff for a before and after study of manure applications. The before period ($X_1$) occurred during the period when manure was applied to the watershed during the winter on ice and snow. The after period ($X_2$) represents samples that were taken during the period when manure was applied during the spring and incorporated into the soil. Each value listed in the table is the daily mean of eight 4-hour composite samples.

To determine whether the difference in phosphorus concentrations is significant between the two periods, the appropriate null hypothesis is:

$$H_o : \overline{X}_1 - \overline{X}_2 = 0 \text{ or } \overline{X}_1 = \overline{X}_2$$

The appropriate alternative hypothesis would be:

$$H_a : \overline{X}_1 > \overline{X}_2$$

The *t*-test assumes that the data are normally distributed and the groups have equal variances, so the data should first be tested for these assumptions.

**Table 8–1** Mean daily total phosphorus concentrations (mg/L) in watershed runoff from a period before and after implementation of best manure management

| - - - - Total phosphorus - - - - | |
|---|---|
| Before ($X_1$) | After ($X_2$) |
| (mg/L) | (mg/L) |
| 6.330 | 0.185 |
| 2.166 | 0.049 |
| 0.642 | 0.040 |
| 0.754 | 0.087 |
| 0.728 | 0.142 |
| 0.478 | 0.060 |
| 0.464 | 0.187 |
| 0.444 | 0.068 |
| 0.375 | 0.043 |
| 0.120 | 0.039 |
| 0.086 | 0.404 |
| 0.064 | 0.110 |
| 0.099 | 0.085 |
| 0.054 | 0.082 |
| 0.063 | 0.138 |
| 0.197 | 1.617 |
| 0.088 | 0.798 |
| 0.089 | 0.104 |
| 0.110 | 0.341 |
| 0.105 | 0.055 |
| 0.081 | 0.295 |
| | 0.090 |
| | 0.211 |
| | 0.151 |
| | 0.158 |
| | 0.047 |
| | 0.029 |
| | 0.027 |
| | 0.065 |
| | 0.152 |
| | 0.087 |
| | 0.041 |
| | 0.544 |
| | 0.296 |

**Example 8–1** Single watershed analysis—Continued

Using a statistical package, such as SAS®, a test for normality is made as described in chapter 4. Table 8–2 shows the test results for the total phosphorus data.

**Table 8–2** Test of normality for the total phosphorus data

| | Untransformed | | $Log_{10}$ transformed | |
|---|---|---|---|---|
| | $X_1$ | $X_2$ | $X_1$ | $X_2$ |
| Mean | 0.645 | 0.201 | –0.631 | –0.934 |
| Median | 0.120 | 0.097 | –0.921 | 1.014 |
| Skewness | 3.840 | 3.690 | 0.998 | 0.764 |
| Kurtosis | 15.70 | 15.78 | 0.531 | 0.404 |
| W:Normal | 0.445 | 0.559 | 0.884 | 0.954 |
| Prob < W | <0.001 | <0.001 | 0.015 | 0.204 |

Based on the nonsignificant Shapiro-Wilk W statistic, the data appear to be log-normally distributed. Therefore, the log transformation is used prior to the *t*-test. The next step is to calculate $S_d$, the standard deviation of the difference between means. Since $n_1$ does not equal $n_2$, equation 8–2 is used to calculate $S_d$. Table 8–3 provides a summary of calculations needed to determine $S_d$.

**Table 8–3** Summary of calculations for $log_{10}$ transformed phosphorus data

| | $X_1$ | $X_2$ |
|---|---|---|
| n | 21 | 34 |
| $\sum X$ | –13.242 | –31.764 |
| $\sum X^2$ | 14.595 | 35.465 |
| Log $\overline{X}$ | –0.631 | –0.934 |
| $\overline{X}$ | 0.234 | 0.116 |

First $S_p$ is calculated from equation 8–4:

$$S_p^2 = \frac{\left[14.595 - \frac{(13.242)^2}{21}\right] + \left[35.465 - \frac{(-31.764)^3}{}\right]}{(21-1)+(34-1)}$$

$$= \frac{6.245 + 5.790}{660} = 0.018235$$

$S_d$ is calculated from equation 8–2:

$$S_d = \sqrt{0.018235\frac{(21+34)}{(21)(34)}} = 0.037479$$

Student's *t* is calculated from equation 8–1:

$$t = \frac{0.637 - (-0.934)}{0.037479} = 8.085$$

From appendix A the table *t*-value is 2.006 for df = $(n_1–1)+(n_2–1)$ = 53 and p = 0.05. Therefore, since the calculated *t* is greater than the table *t*, the $H_o$ is rejected. The mean is determined on the log-transformed values. Therefore, to transform the mean back to original units, the antilog of the log mean is taken by taking the value 10 and raising it to the power of the log mean.

Based upon the *t*-test, this before and after study determined that the mean phosphorus concentration was significantly reduced by 50 percent after the implementation of the practice as compared to before the practice. Confidence limits can be added to this estimate of differences between means from:

$$\overline{X}_1 - \overline{X}_2 \pm t_\alpha S_{\overline{X}_1 - \overline{X}_2} \qquad [8–5]$$

Where the standard error is calculated from:

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}} \qquad [8–6]$$

or for log normal distributions when n is not large, consult page 170 of Gilbert (1987).

**Example 8–1** Single watershed analysis—Continued

For this example

$$S_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{0.0182}{21} + \frac{0.0182}{34}} = 0.037$$

The confidence limit is:

$$0.234 - 0.116 \pm 2.004(0.037) = 0.118 \pm 0.074$$

However, because of the limitations of this experimental design, it is possible that the differences are actually the result of some climate difference from the first year to the second. The design does not provide a way to correct for any deterministic features in the data, such as cyclic patterns or rainfall. For example, the change in concentrations might also be caused by a dry year following a wet year.

The SAS® program for the *t*-test in example 8–1 would be:

**SAS PC Program**

```
Data phos;
     title 'TTest of Phos Data';
     infile 'a:phos.dat';
     input trt phos;
logphos = log10(phos);
Proc TTEST;
     class trt;
run;
```

## 615.0802 Nonparametric two-sample test

If data violate the assumptions of normal distributions or equal variances, nonparametric or distribution-free approaches may be used (Zar 1984). The Mann-Whitney test is the nonparametric equivalent to the *t*-test for two-samples. As previously described for other nonparametric approaches, the ranks of the values are used rather than the values themselves. Ranking is done from highest to lowest, with the largest value in both groups given a value of 1 and so on.

The Mann-Whitney U statistic is calculated from:

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \qquad [8\text{–}7]$$

and

$$U' = n_1 n_2 - U \qquad [8\text{–}8]$$

where:
n = number of samples in each group
R = sum of the ranks for that group (Zar 1984)

If either U or U´ is equal to or greater than the table U, the $H_o$ is rejected at the appropriate α.

The data in table 8–1 are used in example 8–2, which is a nonparametric approach for single watershed analysis.

**Example 8–2** Nonparametric single watershed analysis

Table 8–4 provides the ranks for the data in table 8–1.

**Table 8–4** Ranks of total phosphorus concentrations for the before ($X_1$) and after ($X_2$) study of manure management

| | $X_1$ | $X_2$ | $X_2$ |
|---|---|---|---|
| | 1 | 20 | 32 |
| | 2 | 48 | 17 |
| | 7 | 52 | 23 |
| | 5 | 36 | 21 |
| | 6 | 24 | 49 |
| | 9 | 45 | 54 |
| | 10 | 19 | 55 |
| | 11 | 41 | 42 |
| | 13 | 50 | 22 |
| | 26 | 53 | 35 |
| | 37 | 12 | 51 |
| | 43 | 27 | 8 |
| | 31 | 38 | 15 |
| | 47 | 39 | |
| | 44 | 25 | |
| | 18 | 3 | |
| | 34 | 4 | |
| | 33 | 30 | |
| | 28 | 14 | |
| | 29 | 46 | |
| | 40 | 16 | |
| n | 21 | 34 | |
| R | 474 | 1066 | |

$$U = (21)(34) + \frac{21(21+1)}{2} - 474 = 471$$

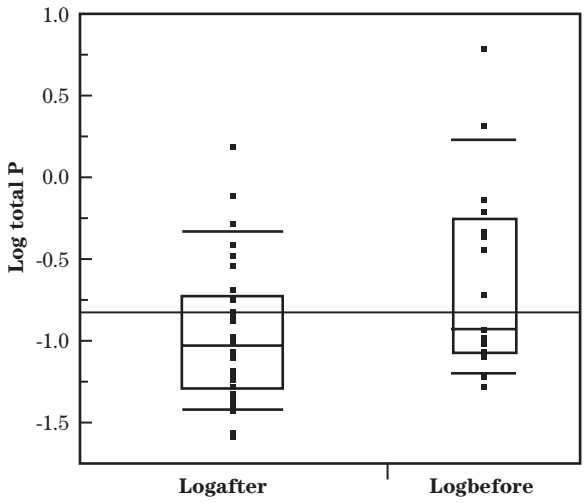
$$U' = (21)(34) - 471 = 243$$

The table value for U is 450 (α=0.05) (Zar, 1984). Since the calculated U is greater than the table U, the $H_o$ of equal concentrations is rejected. Using either parametric approaches with a transformation or nonparametric approaches, the conclusion was that there was a significant difference in the mean concentrations of total phosphorus in runoff.

## 615.0803 Presentation of results

The presentation of results from a before and after study is generally a presentation of means. Box plots (fig. 8–1) are also an appropriate presentation of the data. The bottom and top of the box represent the 25th and 75th percentiles, the center horizontal line is the median, and the outer lines are the 10th and 90th percentiles. In some cases time plots of the data can be used; however, since the data are not paired in a meaningful manner, the time plot could result in a misleading interpretation.

**Figure 8–1** Boxplots of phosphorus data




## 615.0804 References

Gilbert, R.O. 1987. Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold, New York, NY.

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods (7th ed.). The IA State Univ. Press, Ames, Iowa.

Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York, NY.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 9 Above and Below Watersheds

(450–VI–NWQH, September 2003)

# Chapter 9 Above and Below Watersheds

Contents:

# Chapter 9 Above and Below Watersheds

## 615.0900 Introduction

The above-and-below design is often thought of as a way to isolate the effect of a treatment. Theoretically, if we sample the water before it flows into an area and then again after it leaves an area, the difference in water quality will be a result of the treatment in the area. In some cases this may be true; however, the difference may be caused by watershed differences as well. An alternative is to conduct an above-and-below study before and after the treatment. Such a study becomes a paired watershed study as described in chapter 10.

The above-and-below design is actually one watershed physically nested within another. This design is applicable to streams as well as ground water systems. The appropriate statistical approach is the paired *t*-test of above-and-below data.

This chapter describes the assumptions used for the above-and-below design, provides examples of how to analyze the data using both parametric and nonparametric approaches, and gives examples of how to present the results from the study.

## 615.0901 Assumptions

The *t*-test assumes that the data are normally distributed and the two groups being compared are of equal variances (Snedecor and Cochran 1980, Steel and Torrie 1960, Zar 1984). If the data fail these assumptions, a transformation or nonparametric approach should be used. One of the conditions of the paired *t*-test is that pairs actually exist. Thus if data are collected at one station, but not the other, no pair exists. Flow occurring at one station, but not at the other still constitutes a pair since one of the values is a zero and the other is above zero. However, when there is no water to measure, a concentration value does not exist and, therefore, a concentrated pair does not exist.

## 615.0902 Paired comparison of means

The paired comparison of means assumes that the paired values are correlated in some way (Steel and Torrie 1960). Therefore, when one value of the pair was large, we would expect the other value to also be large. The variance is then computed on the difference between paired values rather than on the individual observations as for the unpaired example.

The appropriate null hypothesis of the paired comparison of means is the same as for the unpaired comparison described in chapter 8:

$$H_o : \overline{X}_1 - \overline{X}_2 = 0$$

The appropriate alternative hypothesis would be:

$$H_a : \overline{X}_1 - \overline{X}_2 \neq 0$$

The test of the significance of the difference between the means is based on the *t* distribution (Steel and Torrie 1960, Zar 1984) where *t* is defined as:

$$t = \frac{\overline{d}}{s_d} \qquad [9–1]$$

where:

$\overline{d}$ = the mean of the differences between the paired observations

$S_d$ = standard deviation of the difference between the means, which is determined from:

$$S_d^2 = \frac{\sum d_i^2 - \frac{\left(\sum d_i\right)^2}{n}}{n(n-1)} \qquad [9–2]$$

where:

$d_i$ = difference between the paired observation

$n$ = number of observation pairs

Example 9–1 illustrates the statistical analysis using the above-and-below process.

**Example 9–1** Above-and-below watershed analysis

Table 9–1 presents a summary of total phosphorus concentrations in watershed runoff above and below an area that received winter manure applications on ice and snow. Each value listed in the table is the daily means of eight 4-hour samples. The below data are the same as those listed as before data in table 8–1 in chapter 8. This example allows a direct comparison of the single watershed analysis to the above-and-below analysis since the data are real observations from a watershed in Vermont.

Determine whether a significant difference in phosphorus concentrations occurs between the above and below stations. The appropriate null hypothesis is:

$$H_o : \overline{X}_1 - \overline{X}_2 = 0$$

The appropriate alternative hypothesis would be:

$$H_a : \overline{X}_1 - \overline{X}_2 \neq 0$$

Because the *t*-test assumes that the data are normally distributed and the groups have equal variances, the data should first be tested for these assumptions.

Using a statistical package, such as SAS®, the data should be examined for normality. As shown in table 9–2, the data appear to be log-normally distributed. Therefore, the log transformation is used before the *t*-test. To calculate $S_d$ and t, the values in table 9–3 are calculated.

From appendix A, the table *t*-value is 2.101 for df = n–1 = 18 and p = 0.05. Therefore, since the calculated *t* is greater than the table *t*, the $H_o$ is rejected. The mean is determined on the log transformed values. To transform the mean back to original units, the antilog of the log mean is obtained by taking the value 10 and raising it to the power of the log mean. If a negative value had been obtained for the difference, a constant would need to be added

**Example 9–1** Above-and-below watershed analysis—Continued

to all difference values before a log transformation could be used because the log of a negative number does not exist.

Based upon the *t*-test, this above-and-below study determined that the phosphorus concentration was significantly increased in runoff by 0.173 mg/L as a result of the winter application of manure. However, because of the limitations of this experimental design, it may be possible that the differences may actually be the result of an inherent watershed difference between the upstream and downstream stations.

**Table 9–1** Mean daily total phosphorus concentrations (mg/L) in watershed runoff above and below an area receiving manure applications in the winter

| Total phosphorus (mg/L) | | | |
|---|---|---|---|
| Above | Below | Difference | Rank |
| 0.060 | 6.330 | 6.270 | 19 |
| 0.095 | 2.166 | 2.071 | 18 |
| 0.117 | 0.642 | 0.525 | 15 |
| 0.073 | 0.754 | 0.681 | 17 |
| 0.050 | 0.728 | 0.678 | 16 |
| 0.034 | 0.478 | 0.444 | 14 |
| 0.250 | 0.464 | 0.214 | 11 |
| 0.211 | 0.444 | 0.233 | 12 |
| 0.090 | 0.375 | 0.285 | 13 |
| 0.032 | 0.120 | 0.088 | 10 |
| 0.027 | 0.086 | 0.059 | 7 |
| 0.076 | 0.064 | –0.012 | 1 |
| 0.058 | 0.099 | 0.041 | 2 |
| 0.012 | 0.054 | 0.042 | 3 |
| 0.011 | 0.063 | 0.052 | 5 |
| 0.056 | 0.088 | 0.032 | 1 |
| 0.029 | 0.089 | 0.060 | 8 |
| 0.040 | 0.110 | 0.070 | 9 |
| 0.049 | 0.105 | 0.056 | 6 |
| 0.036 | 0.081 | 0.045 | 4 |

**Table 9–2** Test of normality for the difference in total phosphorus data

| | Untransformed | Log transformed |
|---|---|---|
| | d | d |
| Mean | 0.629 | –0.7634 |
| Median | 0.088 | –1.0555 |
| Skewness | 3.696 | 0.922 |
| Kurtosis | 14.442 | 0.158 |
| W:Normal | 0.449 | 0.888 |
| Prob < W | <0.001 | 0.029 |

**Table 9–3** Summary calculation for determining the value of *t*

| | log (d) |
|---|---|
| n | 19 |
| $\sum d_i$ | –14.5048 |
| $\sum d_i^2$ | 18.6759 |
| $\log \overline{X}$ | –0.763 |
| $\overline{X}$ | 0.173 (mean difference in mg/L) |

$$S_d^2 = \frac{18.6759 - \frac{(-14.5048)^2}{19}}{19(19-1)} = 0.0222$$

$$t = \frac{-0.763}{0.0222} = -34.369$$

The SAS® program for the *t*-test in example 9–1 would be:

**SAS PC Program**

```
Data phos;
     title 'TTest of Phos Data';
     infile 'a:phos.dat';
     input phos1 phos2;
diff=phos2-phos1;
logdiff = log10(diff);
Proc MEANS Mean Stderr T PRT;
     Var diff;
run;
```

## 615.0903 Nonparametric paired-sample test

If the data violate the assumptions of normal distributions or equal variances, nonparametric or distribution-free approaches may be used (Zar 1984) as was used for the unpaired comparison of means in chapter 8. The Wilcoxon paired sample test is the nonparametric equivalent to the *t*-test for paired samples. As previously described for other nonparametric approaches, the ranks *of the differences* between the values are used rather than the differences themselves. Ranking is done from lowest to highest with the smallest difference given a value of 1 and so on. The sign of the difference is also carried with the rank. Ranks are summed for both positive (T+) and negative (T–) ranks. The T values are compared to a tabular T value; if either value is less than or equal to the table T value, the $H_o$ of equal values is rejected.

The data in table 9–1 are used in example 9–2, which illustrates the nonparametric approach to analysis of the above-and-below design data.

**Example 9–2** Nonparametric above-and-below watershed analysis

T+ = 19 + 18 + ... + 1 = 209

T– = 1

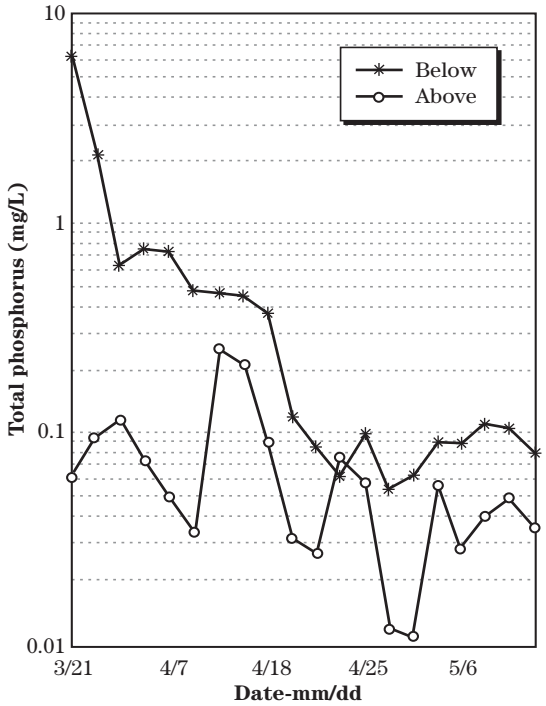
From appendix G, T at n = 20 df and p = 0.05 = 52. Since T– is less than the table T, the null hypothesis of equal concentrations above and below is rejected. Using either the $\log_{10}$ transformation or nonparametric approaches, the conclusion was that there was a significant difference in the mean total phosphorus concentrations in runoff at the below station as compared to those at the above station.

## 615.0904 Presentation of results

The presentation of results from an above-and-below study is usually a presentation of means. In this case the mean total phosphorus concentration was increased from 0.052 to 0.234 mg/L or 4.5 times. Box plots would be an informative graphic approach to presenting the comparison between above and below data.

In some cases time plots are useful in presenting the results. For example 9–1, the time plot in figure 9–1 reveals that the below station was consistently higher in phosphorus concentration than the above station. The plot also reveals that the difference was greater during the early part of the snowmelt season and became progressively less as time went on.

**Figure 9–1** Time plot of above-and-below phosphorus concentration data




## 615.0905 References

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods. 7th Ed. The IA State Univ. Press. Ames, IA.

Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York, NY.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 10 Paired Watersheds

(450–VI–NWQH, September 2003)

# Chapter 10 Paired Watersheds

Contents:



Tables

**Figures**



**Example**

# Chapter 10 Paired watersheds

## 615.1000 Introduction

The purpose of this chapter is to describe data analysis for the paired watershed design for conducting nonpoint source (NPS) water quality studies. The monitoring system design requires a minimum of two watersheds—control and treatment—and two periods of study—calibration and treatment (Green 1979, Hewlett 1971, Hewlett and Pienaar 1973, Ponce 1980, Reinhart 1967).

The control watershed accounts for year-to-year or seasonal climate variations. The management practices within the control watershed remain the same during the study. The treatment watershed has a change in management at some point during the study. During the calibration period, the two watersheds are treated identically, and paired water quality data are collected (table 10–1). Such paired data could be annual means or totals, or for shorter studies (<5 yr), the observations could be seasonal, monthly, weekly, or event-based (Reinhart 1967). During the treatment period, one watershed is treated with a best management practice (BMP) while the control watershed remains in the original management (table 10–1). The treated watershed should be selected randomly by such means as a coin toss.

The reverse of this schedule is possible for certain BMPs; the treatment period could precede the calibration period (Reinhart 1967). For example, the study could begin with two watersheds in two different treatments, such as **BMP** and **no BMP.** Later both watersheds could be managed identically to calibrate them. Since no calibration exists before the treatment occurs, this reversed design is considered risky because you will not find out if the watersheds are properly calibrated until the end of the study.

The basis of the paired watershed approach is that

- The relationship between paired water quality data for the two watersheds is quantifiable.
- This relationship is valid until a major change is made in one of the watersheds (Hewlett 1971). At that time, a new relationship will exist.

This basis does not require that the quality of runoff be statistically the same for the two watersheds. It does require that the relationship between paired observations of water quality remains the same over time except for the influence of the BMP. Often, in fact, the analysis of paired observations indicates that the water quality is different between the paired watersheds. This difference further substantiates the need to use a paired watershed approach. This is because the technique does not assume that the two watersheds are the same; it does assume that the two watersheds respond in a predictable manner together. Example 10–1 illustrates a paired watershed analysis.

**Table 10–1** Schedule of BMP implementation

| Period | Watershed | |
|---|---|---|
| | control | treated |
| Calibration | no BMP | no BMP |
| Treatment | no BMP | BMP |

# 615.1001 Calibration

The relationship between watersheds during the calibration period is described by a simple linear regression equation (fig. 10–1) between the paired observations, taking the form:

$$\text{treated} = b_o + b_1(\text{control}_i) + e \qquad [10–1]$$

where:

treated and control = flow, water quality concentration, or mass values for the appropriate watershed

$b_o$ and $b_1$ = regression coefficients representing the regression intercept and slope, respectively

e = residual error

Three important questions must be answered before shifting from the calibration period to the treatment period:

- Is there a significant relationship between the paired watersheds for all parameters of interest?
- Has the calibration period continued for a sufficient length of time?
- Are the residual errors about the regression smaller than the expected BMP effect?

In addition, the observations should cover the full range of observations expected during treatment.

## (a) Regression significance

The significance of the relationship between paired observations is tested using analysis of variance (ANOVA). The test assumes that the regression residuals are normally distributed, have equal variances between treatments, and are independent.

Hand calculations to test for the significance of the relationship are shown in Snedecor and Cochran (1980, p. 157) and in table 10–2. The values for the table are calculated from:

$$S_y^2 = \sum Y_1^2 - \frac{\left(\sum Y_i\right)^2}{n} \qquad [10–2]$$

$$S_x^2 = \sum X_1^2 - \frac{\left(\sum X_i\right)^2}{n} \qquad [10–3]$$

$$S_{xy} = \sum X_i Y_i - \frac{\left(\sum X_i\right)\left(\sum Y_i\right)}{n} \qquad [10–4]$$

$$S_{yx}^2 = \frac{S_y^2 - \frac{\left(S_{xy}\right)^2}{S_x^2}}{n-2} \qquad [10–5]$$

**Figure 10–1** Calibration period regression

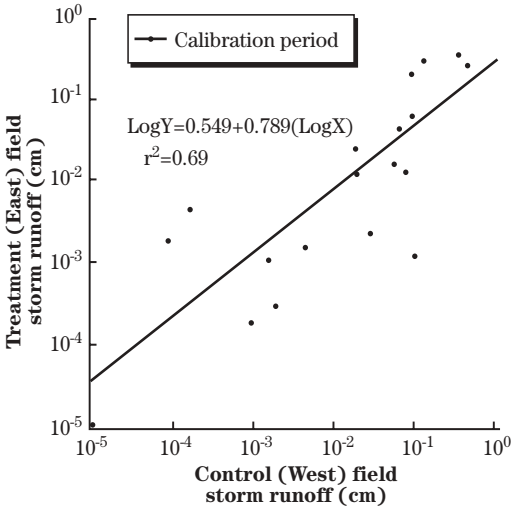



**Table 10–2** Analysis of variance for linear regression

| Source | Degrees of freedom | Sum of squares | Mean squares | F |
|---|---|---|---|---|
| regression | 1 | $\frac{\left(S_{xy}\right)^2}{S_x^2}$ | $\frac{\left(S_{xy}\right)^2}{S_x^2}$ | $\frac{\left[\frac{\left(S_{xy}\right)^2}{S_x^2}\right]}{S_x^2}$ |
| residual | n–2 | $S_y^2 \frac{\left(S_{xy}\right)^2}{S_x^{\,2}}$ | $S_{yx}^2$ | |
| total | n–1 | $S_y^2$ | | |

Also, the regression coefficients and coefficient of determination are determined from:

$$b_1 = \frac{S_{xy}}{S_x^2} \qquad [10–6]$$

$$b_o = \bar{Y} - b_1\bar{X} \qquad [10–7]$$

$$r^2 = \frac{\frac{(S_{xy})^2}{S_x^2}}{S_y^2} \qquad [10–8]$$

To perform the calculations by hand, initially calculate:

$$\Sigma X_i, \Sigma Y_i, \Sigma X_i Y_i, \Sigma X_i^2, \Sigma Y_i^2, \bar{X}, \bar{Y}$$

The mean squares (MS) are determined by dividing the sum of squares by the degrees of freedom (df).

Using SAS®, the appropriate program is shown as:

**SAS PC Program**

```
data flow;
     title 'Total Flow (cm)';
     infile 'fname.dat';
     input flow1 flow2;
logflow1=log10(flow1);
logflow2=log10(flow2);
Proc reg;
     Model logflow2=logflow1
               /P CLM;
run;
```

This program was used to generate table 10–4 in example 10–1.

## (b) Calibration duration

Methods for determining whether the length of the calibration period has been sufficient have been described by Wilm (1949), Kovner and Evans (1954), and Reinhart (1967). The ratio between the residual variance (mean squares, $S_{yx}^2$) for the regression and the smallest worthwhile difference (d) for the treatment watershed is used to determine if a sufficient sample has been taken to detect that difference, from (Kovner and Evans 1954):

$$\frac{S_{yx}^2}{d^2} = \left(\frac{n_1 n_2}{n_1 + n_2}\right)\left[\frac{1}{F\left(1 + F_{n_1+n_2} - 2\right)}\right] \qquad [10–9]$$

where:

$S_{yx}^2$ = estimated residual variance about the regression
$d^2$ = square of the smallest worthwhile difference
$n_1$ and $n_2$ = numbers of observations in the calibration and treatment periods ($n_1 = n_2$ for this calculation because $n_2$ is not known yet)
F = table value (p = 0.05) for the variance ratio at 1 and $n_1 + n_2 - 3$df (appendix C)

The difference (d) is selected based on experience and would vary with project expectations. If the left side of the equation is greater than the right side, then the number of samples taken was not sufficient to detect the difference.

## (c) Residual errors

The confidence bands for the regression equation allow determining the level of change needed to have a significant treatment effect. In other words, how far away from the calibration regression must the treatment data be to be significantly different? Confidence bands for the regression are determined from:

$$CI = \pm(t)(S_{yx})\sqrt{\frac{1}{n} + \frac{(X_i - \bar{X})^2}{S_x^2}} \qquad [10–10]$$

where:

CI = confidence interval
$S_{yx}$ = square root of $S_{yx}^2$
n and $S_x^2$ = factors have been previously defined
t = Student's t
$X_i$ = value at the point of comparison to compare to the mean on the regression line

Confidence limits can be generated in SAS® by adding / P CLM to the MODEL statement.

## 615.1002 Treatment

At the end of the treatment period the significance of the effect of the BMP is determined using analysis of covariance (ANCOVA). The analysis is actually a series of steps determining:

- significance of the treatment regression equation
- significance of the overall regression that combines the calibration and treatment period data
- difference between the slopes of the calibration and treatment regressions
- difference between the intercepts of the calibration and treatment regressions

The analysis can be computed by hand as shown in table 10–3 (Snedecor and Cochran 1980, p. 386). The summation's symbol ($\Sigma$) in table 10–3 is used to signify the addition of the column entries above it.

An example program using SAS® is shown below. This program contains both a test of the treatment regression in the PROC REG statement and a test comparing the regression lines in the PROC GLM statement.

**SAS PC Program**

```
Proc reg;
     model logflow2=logflow1;
run;
Proc glm;
     class period;
     model logflow2=logflow1 period
     logflow1 * period;
run;
```

**Table 10–3** Analysis of covariance for comparing regression lines

| Source | df | $S_x^2$ | $S_{xy}$ | $S_y^2$ | $b_1$ | df | SS | MS | F |
|---|---|---|---|---|---|---|---|---|---|
| Within calibration | $n_1$–1 | Eq 10–3 | Eq 10–4 | Eq 10–2 | Eq 10–6 | $n_1$–2 | $S_y^2 = \frac{(S_{xy})^2}{S_x^2}$ | Eq 10–5 | |
| Within treatment | $n_2$–1 | Eq 10–3 | Eq 10–4 | Eq 10–2 | Eq 10–6 | $n_2$–2 | $S_y^2 = \frac{(S_{xy})^2}{S_x^2}$ | Eq 10–5 | |
| | | | | Pooled | Error | $\Sigma$ | $\Sigma$ | SS/df | |
| | ——— | — | — | ——— | | | | | |
| Slopes | $n_1+n_2$–2 | $\Sigma$ | $\Sigma$ | $\Sigma$ | Eq 10–6 | $n_1+n_2$–3 | $S_y^2 = \frac{(S_{xy})^2}{S_x^2}$ | Eq 10–5 | |
| | | | Slope difference | | | 1 | Slope SS– Error SS | | MS/ Error MS |
| | | | | | | 1 | Combined SS- SlopeSS | | MS/ Slope MS |
| Intercepts | $n_1+n_2$–1 | combined data | | | | $n_1+n_2$-2 | $S_y^2 = \frac{(S_{xy})^2}{S_x^2}$ | | |

## 615.1003 Nonlinear/ multiple regression

At times the effect of the treatment may be nonlinear. Examples of nonlinear treatment effects include different responses to storm size or gradual vegetation changes. Swindel and Douglass (1984) described approaches for testing nonlinear treatment effects including quadratic approaches and fitting to a gamma distribution. Multiple regression may also be used for paired watershed studies (Hibbert 1969, Snyder 1962).

Regression through the origin can be used where zero flow is expected to occur from both watersheds at approximately the same time. This would occur for adjacent, equally sized watersheds, but not for watersheds of different sizes.

## 615.1004 Displaying results

The most common methods for displaying the results include a bivariate plot of paired observations together with the calibration and treatment regression equations (fig. 10–2). Another useful graph is a plot of deviations ($y_{observed} - y_{predicted}$) as a function of time during the treatment. The predicted values are obtained from the calibration regression equation.

Results should be provided of mean values for each period and each watershed. The overall results caused by the treatment can be expressed as the percent change based on the mean predicted and observed values.

## 615.1005 References

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York, NY.

Hewlett, J.D. 1971. Comments on the catchment experiment to determine vegetal effects on water yield. Water Res. Bul. 7(2):376-381.

Hewlett, J.D., and L. Pienaar. 1973. Design and analysis of the catchment experiment. *In* Proc. Symp., Use of Small Watersheds, E.H. White (ed.), Univ. KY.

Hibbert, A.R. 1969. Water yield changes after converting a forested catchment to grass. Water Resourc. Res. 5(3):634-640.

Kovner, J.L., and T.C. Evans. 1954. A method for determining the minimum duration of watershed experiments. Trans. AGU 35(4):608-612.

Ponce, S.L. 1980. Statistical methods commonly used in water quality data analysis. WSDG Tech. Pap. WSDG-TP-00001, USDA Forest Service, Fort Collins, CO.

Reinhart, K.G. 1967. Watershed calibration methods. *In* Proc. Intern. Symp. on Forest Hydrology, W.E. Sopper and H.W. Lull. (eds.), Pergamon Press. Oxford. pp. 715-723.

SAS Institute, Inc. 1986. SAS system for linear models. Cary, NC

Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods (7th ed.). IA State Univ. Press, Ames.

Snyder, W.M. 1962. Some possibilities for multivariate analysis in hydrologic studies. J. Geophysical Res. 67(2):721-729.

Swindel, B.F., and J.E. Douglass. 1984. Describing and testing nonlinear treatment effects in paired watershed experiments. Forest Science 30(2):305-313.

Wilm, H.G. 1949. How long should experimental watersheds be calibrated? Amer. Geophs. Union Trans, Part II, 618-622.

**Example 10–1** Paired watershed analysis

Data from a study in Vermont is used to illustrate the paired watershed approach. The purpose of the study was to compare changes in field runoff as a result of conversion of conventional tillage to conservation tillage. Information included:

- West watershed was the control and was 1.46 hectares (ha) in area.
- East watershed was the treatment field and was 1.10 ha.
- Conventional tillage was moldboard plow whereas conservation tillage was a single disk harrow.
- The calibration period was 1 year during which 49 paired observations of storm runoff were made.
- The treatment period was 3 years during which 114 paired observations of runoff were made.

The assumptions were tested for ANOVA. Data were log-transformed to approach normality based upon the Shapiro-Wilks (W) statistic. The equality of variances between periods was tested using the F-test. Residual plots were examined to check for independence of errors. The statistical package SAS® was used for all analyses (SAS 1986).

The regression coefficients of paired observations are calculated by hand as follows:

$$\sum X_i = -123.403$$
$$\sum Y_i = -180.704$$
$$\sum X_i Y_i = 533.553$$
$$\sum X_i^2 = 381.713$$
$$\sum y_i^2 = 814.847$$
$$\bar{X} = -2.518 \left(10^X = 0.003041\text{cm}\right)$$
$$\bar{Y} = -3.688 \left(10^Y = 0.000205\text{cm}\right)$$

Therefore,

$$S_y^2 = 148.441$$
$$S_{xy} = 78.463$$
$$S_x^2 = 70.933$$
$$S_{yx}^2 = 1.312$$

The resulting F statistic for this example would indicate that the regression adequately explains a significant amount ($p<0.001$) of the variation in paired data.

For the example, $S_{yx}^2$ was 1.312 (from table 10–4), $n_1 = n_2$ was 49, and F was 3.94. A 10 percent change from the mean was considered a worthwhile difference; therefore,

$$d = 0.10 \times \bar{X} = 0.10 \times \log 0.003041\text{cm}$$
$$\frac{S_{yx}^2}{d^2} = 20.7$$

The right side of equation 10–9 equals 6. Because 20.7 is greater than 6, the number of observations was not sufficient to detect a 10 percent change in discharge. Enough samples were taken to detect a 20 percent change in discharge:

$$\frac{S^2}{d^2} = 5.2$$

**Table 10–4** Analysis of variance for regression of treatment watershed runoff on control watershed runoff

| Source | df | MS | F | p |
|---|---|---|---|---|
| model | 1 | 86.79 | 66.17 | 0.0001 |
| error | 47 | 1.31 | | |
| total | 48 | | | |

**Example 10–1** Paired watershed analysis—Continued

To perform the calculations for determining analysis of covariance (ANCOVA) by hand, determine the following for the example treatment data:

$$\sum X_i = -358.14$$
$$\sum Y_i = -416.05$$
$$\sum X_i Y_i = 1{,}408.37$$
$$\sum X_i^2 = 1{,}352.54$$
$$\sum Y_i^2 = 1{,}623.43$$
$$\bar{X} = -3.1416$$
$$\bar{Y} = -3.650$$
$$n = 114$$

Therefore,

$$S_y^2 = 135.00$$
$$S_{xy} = 101.32$$
$$S_x^2 = 227.43$$

The treatment period regression was found to be significant based on the analysis of variance for regression (table 10–5).

The analysis of covariance obtained in SAS® output summarizes the significance of the overall model, compares the two regression equations, the regression intercepts, and the slopes (table 10–6). The ANCOVA indicates that the overall treatment and calibration regressions were significantly different and that the slopes and intercepts of the equations also were different. The difference in slopes is evident in figure 10–2. The slight differences in F values between the hand calculation method and the SAS® output are caused by rounding errors.

For the example, the plot of deviations indicates that for most paired observations, the observed value was less than that predicted by the calibration regression equation (fig. 10–3).

In the example, a 64 percent reduction in mean runoff was attributed to the treatment (table 10–7).

The ANCOVA is completed for the example in table 10–8.

Since the slopes were found to be different, the differences in intercepts do not have any real meaning and do not need to be calculated. That is, if slopes are different, intercepts generally are different. However, the calculation for the test of intercepts is presented to show the method. The combined data are determined by summing the $\Sigma X_i, \Sigma Y_i, \Sigma X_i Y_i, \Sigma X_i^2$, and $\Sigma Y_i^2$ values for both the calibration and treatment periods and calculating new values for $S_y^2$, $S_{xy}$, and $S_x^2$. The calculation of F for the intercept uses the slope MS in the denominator. The F for the slope test uses the error MS in the denominator. A significant difference in intercepts, but not slopes indicates an overall parallel shift in the regression equation.

**Table 10–5** ANOVA for regression of treatment watershed runoff on control watershed runoff for the treatment period

| Source | df | MS | F | p |
|---|---|---|---|---|
| model | 1 | 45.13 | 56.25 | 0.0001 |
| error | 112 | 0.80 | | |
| total | 113 | | | |

**Table 10–6** ANCOVA for comparing calibration and treatment regressions

| Source | df | MS | F | p |
|---|---|---|---|---|
| model | 3 | 43.99 | 46.17 | 0.001 |
| error | 159 | 0.95 | | |
| overall | 1 | 103.09 | 108.18 | 0.0001 |
| intercept | 1 | 5.47 | 5.74 | 0.0178 |
| slope | 1 | 23.42 | 24.58 | 0.0001 |

**Example 10–1** Paired watershed analysis—Continued

**Figure 10–2** Treatment and calibration period regressions

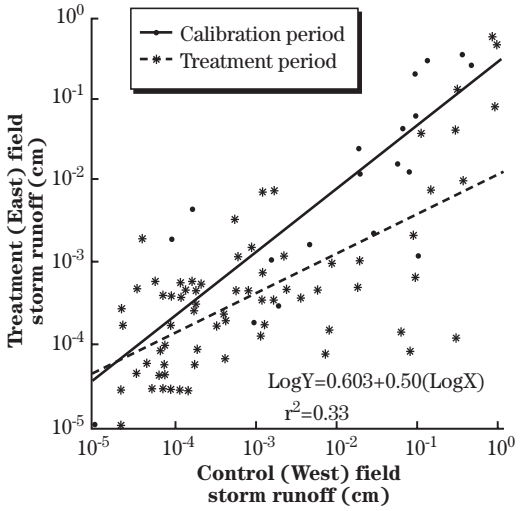



**Table 10–7** Mean values by period and watershed

| | Runoff (cm) x $10^{-2}$ | |
|---|---|---|
| Calibration | | |
| Control | 0.30 | |
| Treatment | 1.63 | |
| Treatment | | |
| Control | 0.08 | |
| Treatment | 0.04 | |
| Predicted | 0.11 | –64% |

**Figure 10–3** Observed deviations from predicted discharge

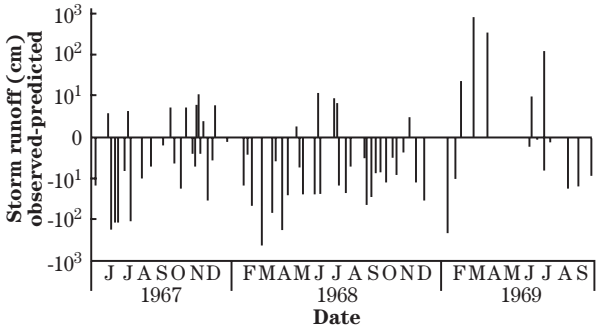



**Table 10–8** Example analysis of covariance for comparing regression lines

| Source | df | $S_x^2$ | $S_{xy}$ | $S_y^2$ | $b_1$ | df | SS | MS | F |
|---|---|---|---|---|---|---|---|---|---|
| Within calibration | 48 | 70.933 | 78.463 | 148.441 | 1.106 | 47 | 61.650 | 1.3117 | |
| Within treatment | 113 | 227.430 | 101.315 | 135.000 | 0.445 | 112 | 89.866 | 0.8024 | |
| | | | | Error | | 159 | 151.516 | 0.9529 | |
| Slopes | 161 | 298.363 | 179.778 | 283.441 | 0.603 | 160 | 175.116 | 1.0945 | |
| | | | | Slope difference | | 1 | 23.600 | 23.600 | 24.77*** |
| | | | | | | 1 | 5.8453 | 5.8453 | 5.34* |
| Intercepts | 162 | 311.671 | 178.762 | 283.492 | | 161 | 180.961 | | |

*** indicates significance at p=0.001
* indicates significance at p=0.05

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 11 Multiple Watersheds

(450–VI–NWQH, September 2003)

# Chapter 11 Multiple Watersheds

Contents:



Tables



Figures

**Examples**

# Chapter 11 Multiple Watersheds

## 615.1100 Introduction

The purpose of this chapter is to describe data analysis for the multiple watershed approach for conducting nonpoint source water quality studies. The multiple watershed approach is a study involving more than two watersheds in the design. Wicht (1967) described this approach that was intended to overcome some of the disadvantages of the paired watershed approach. These disadvantages included:

- Inability to always find a stable control watershed
- Uncertainty in predicting the length of the calibration period
- Risk that meteorological conditions may change at the same time as when treatment begins
- Progressive long-term response, such as during major land use changes

In addition, extrapolation of the results from paired watershed studies to broader areas or regions can be questioned, and there is no true replicate in paired watershed investigations.

For the multiple watershed approach, the treatments are intended to be applied to a series of watersheds that have comparable geology, topography, and initial vegetative cover, and are subject to the same or related uncontrolled climate influences (Wicht 1967).

Striffler (1965) also described a multiwatershed method that used multiple regression analysis to assess the relationship between a dependent variable, such as sediment yield, and several independent variables, such as watershed area, soil or vegetative types, and precipitation. Many watersheds selected represent different levels for the independent variables. A major advantage of such an approach is that a large range of watershed conditions is being sampled. Sampled watersheds also can vary in size and other characteristics, such as varying levels of a disturbance.

However, a different approach is more appropriate for nonpoint source pollution studies. Watersheds that have the treatment already in place could be selected across a region of interest. The size of the region would be dictated by the objectives of the study, but could be as large as a state or perhaps limited to an ecoregion or smaller unit. Once the watersheds were selected, sampling of the appropriate water quality variables would be conducted over a period of time. Clausen and Brooks (1983a, b) used such an approach when comparing the water quality associated with different types of wetlands and when comparing mined to unmined bogs.

This chapter describes the assumptions made in a multiple watershed experiment, presents examples of how to analyze the data from such designs using both parametric and nonparametric approaches, and gives examples of how to present the results from the study.

## 615.1101 Assumptions

The primary statistical approach for comparing groups of watersheds is the analysis of variance. Therefore, the assumptions made are the same as those previously described for analysis of variance (ANOVA). The major assumptions are:

- Water quality data are sampled randomly.
- Data come from a normal distribution.
- Variances are homogeneous across groups.
- Experimental errors are independently distributed.
- Treatment effects are additive.

The approaches used to test these assumptions are described in chapter 4. When using nonparametric approaches, the assumption of normality is no longer appropriate.

## 615.1102 Number of watersheds

One of the first decisions to make when designing a multiple watershed monitoring study is determining the number of watersheds in each group to monitor. Part 614, chapter 9, National Water Quality Handbook, describes procedures for estimating the number of sampling units for water quality monitoring. The basic requirements are knowledge of the variance among watersheds and a desired precision to achieve in the study. Clausen and Brooks (1983a) found that 15 watersheds of each type were sufficient to determine differences in the water quality of different peatland types.

# 615.1103 Comparison of groups

The original analytical approach suggested was a series of paired comparisons between different pairs of watersheds using covariance analysis as for the paired watershed technique (Wicht 1967). Both parametric and nonparametric approaches can be used to compare the results from several groups. The methods of analysis are quite similar to those used for plot studies (part 615, chapter 7). Example 11–1 illustrates parametric method of data analysis.

**Example 11–1** Parametric multiple watershed analysis

The multiple watershed approach was used to assess the water quality effects associated with paving dairy barnyards in Vermont. The objective of the study was to determine the effect of paving on runoff water quality within a 26,000-acre watershed. Five paved and five unpaved barnyards were sampled for runoff on an event basis nine times over 1 year. During each rainfall event, one or two grab samples were collected from each barnyard. Samples were analyzed for phosphorus (table 11–1), nitrogen, and suspended solids; however, only the total phosphorus concentration data are used in this example. Missing concentration data occurred during the study when either there was no runoff or the sample was destroyed during the analysis process.

Using PROC UNIVARIATE in SAS® (SAS 1995) the phosphorus concentration data were found to be log normally distributed (table 11–2). A P-value <0.05 for the unlogged data (i.e., the data prior to log transformation) indicated that the distribution may be significantly different from a normal distribution based on the Shapiro-Wilk W-statistic.

PROC ANOVA was used to test the null hypothesis that the mean phosphorus concentrations were the same in runoff from the paved and unpaved barnyards. The resulting ANOVA (table 11–3) indicated that there was a significant difference between barnyard types, and the null hypothesis is rejected.

The log mean and antilog mean phosphorus concentrations for the barnyard data are reported in table 11–4. The antilog was obtained by taking 10 to the power of the log value. These results indicate that the paved barnyards in this watershed had runoff phosphorus concentrations that were about two times greater than that in runoff from the unpaved barnyards.

**Example 11–1** Parametric multiple watershed analysis—Continued

**Table 11–1** Phosphorus concentration of runoff from paved and unpaved barnyards

| Date | Paved | Unpaved | Date | Paved | Unpaved |
|---|---|---|---|---|---|
| | ---- mg/L ---- | | | ---- mg/L ---- | |
| 6/12 | 20.20 | 1.90 | 3/20 | 12.40 | 90.00 |
| | 67.80 | 16.10 | | 50.00 | 22.30 |
| | 3.40 | 4.90 | | 13.30 | --- |
| | 38.20 | 5.10 | | 192.50 | 17.70 |
| | 25.70 | 23.30 | | 132.50 | 29.00 |
| 6/13 | 36.70 | 7.20 | 6/2 | 13.40 | 13.40 |
| | 132.7 | 14.70 | | 52.00 | 13.80 |
| | 12.20 | 25.50 | | 17.00 | 36.70 |
| | 80.70 | 7.20 | | 134.30 | 8.60 |
| | 32.70 | 20.30 | | 7.40 | 15.03 |
| 9/27 | 22.00 | 53.00 | 6/8 | 10.30 | 17.70 |
| | --- | 18.20 | | 105.50 | 27.60 |
| | 19.80 | 40.85 | | 47.30 | 18.80 |
| | 59.20 | 23.30 | | 68.80 | 6.40 |
| | --- | 35.90 | | 17.20 | 19.10 |
| 10/5 | 22.90 | 13.30 | 8/7 | --- | 19.00 |
| | 54.15 | 19.00 | | 63.35 | 26.00 |
| | 38.30 | 44.40 | | 93.02 | --- |
| | 73.70 | 14.60 | | 86.68 | 9.80 |
| | 96.60 | 43.10 | | 83.02 | 22.20 |
| 11/5 | 82.27 | 27.78 | | | |
| | 50.75 | 25.11 | | | |
| | 47.01 | 22.10 | | | |
| | 44.34 | 7.48 | | | |
| | 35.79 | 18.16 | | | |

**Table 11–2** Univariate statistics for barnyard phosphorus concentration data

| | Unlogged | Log10 |
|---|---|---|
| Skewness | 2.04 | –0.17 |
| Kurtosis | 4.68 | –0.02 |
| W-statistic | 0.777 | 0.986 |
| P-value | <0.001 | 0.877 |

**Table 11–3** ANOVA for barnyard phosphorus concentration data

| Source of variation | Degrees freedom | Sum of squares | Mean squares | F | P > F |
|---|---|---|---|---|---|
| Between | 1 | 2.627 | 2.627 | 21.53 | <0.0001 |
| Within | 83 | 10.127 | 0.122 | | |
| Total | 84 | 12.754 | | | |

**Table 11–4** Mean total phosphorus concentrations of runoff from the paved and unpaved barnyards

| | Log mean | Mean |
|---|---|---|
| | ------ mg/L ------ | |
| Paved | 1.4099 | 25.70 |
| Unpaved | 1.0927 | 12.38 |

## 615.1104 Nonparametric approaches

The nonparametric approaches described in chapters 7 to 9 are also appropriate for multiple watersheds data analysis. For the comparison of two groups, the Mann-Whitney or Wilcoxon rank-sum test may be used. For the comparison of more than two groups, the Kruskal-Wallis nonparametric analysis of variance may be appropriate.

Example 11–2 uses the Wilcoxon rank-sum test for the barnyard phosphorus data analyzed in example 11–1. From the previous example it was determined that the data were not normally distributed, which serves as justification for performing nonparametric analysis.

**Example 11–2** Nonparametric multiple watershed analysis using the phosphorus barnyard data

For the data in example 11–1, test the null hypothesis that the median phosphorus concentrations are the same for the paved and unpaved barnyards. The alternative hypothesis would be that the median concentrations are different.

Using JMP (SAS 1995), the box-and-whisker plots in figure 11–1 were obtained. This boxplot shows the median, the 25th and 75th quartiles framing the box, and two lines indicating the 10th and 90th percentiles.

Output for the Wilcoxon rank-sums test is given in table 11–5. This analysis indicates that the medians are significantly different and the null hypothesis is rejected. The median phosphorus concentration for the paved barnyard runoff of 47.2 mg/L was 2.5 times greater than the median of 19.0 mg/L for the unpaved barnyard. These results are similar to the parametric results presented in example 11–1.

**Figure 11–1** Boxplots of the paved and unpaved barnyard phosphorus concentration data

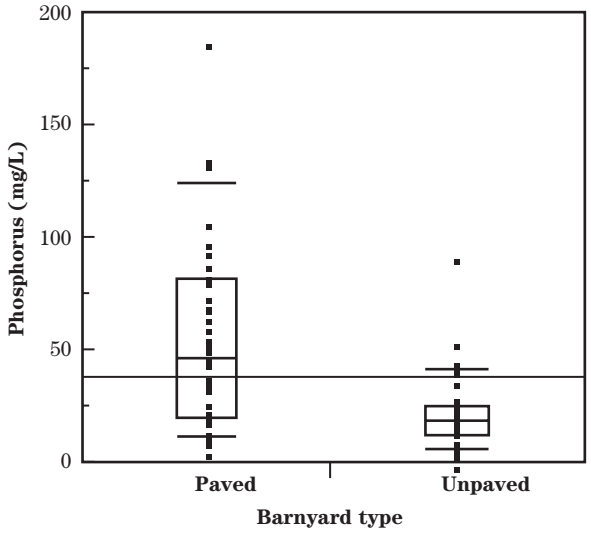



**Table 11–5** Wilcoxon rank-sum test for the barnyard phosphorus data using JMP

| Level | Count | Score sum | Score mean | Mean-mean0 Std0 |
|---|---|---|---|---|
| Paved | 42 | 2273.5 | 54.1310 | 4.105 |
| Unpaved | 43 | 1381.5 | 32.1279 | –4.105 |

2-sample test, normal approximation

| S | Z | Prb>\|Z\| |
|---|---|---|
| 2273.5 | 4.10501 | 0.0000 |

1-way test, chi-square approximation

| ChiSquare | DF | Prob>ChiSq |
|---|---|---|
| 16.8872 | 1 | 0.0000 |

## 615.1105 Presentation of results

The presentation of results depends in part on the number of groups being compared. However, side-by-side boxplots, as shown in figure 11–1, are a favored method of presenting results because they display graphically the data distributions. When viewing boxplots, if the boxes do not overlap each other, the groups are usually different.

## 615.1106 References

Clausen, J.C., and K.N. Brooks. 1983a. Quality of runoff from Minnesota peatlands: I. a characterization. Water Resourc. Bul. 19(5):763-767.

Clausen, J.C., and K.N. Brooks. 1983b. Quality of runoff from Minnesota peatlands: II. a method for assessing mining impacts. Water Resources Bul. 19(5):769-772.

SAS Institute, Inc. 1995. JMP statistics and graphics guide. Ver. 3.1, SAS Institute, Inc. Cary, NC.

Striffler, W.D. 1965. The selection of experimental watersheds and methods in disturbed forest areas. Publ. No. 66, I.A.S.H. Symp. of Budapest, pp. 464-473.

Wicht, C.L. 1967. The validity of conclusions from South African multiple watershed experiments. *In* Forest Hydrology, W.E. Sopper and H.W. Lull (eds.), Proc. Intern. Symp. on Forest Hydrol., Pergamon Press, Oxford.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 12 Trend Analysis

(450–VI–NWQH, September 2003)

# Chapter 12 Trend Analysis

Contents:



Tables

# Chapter 12 Trend Analysis

## 615.1200 Introduction

Several techniques have been applied to detect trends in water quality data. A trend as used in this chapter is intended to mean a persistent increase or decrease in a hydrologic or water quality variable over time (Erlebach 1978). Trend analysis methods range from the simple to the complex. Different techniques can be used to select different types of trends, such as monotonic and step trends. *Monotonic trends* are continuing increases or decreases over time (Helsel and Hirsch 1992). *Step trends* are comparisons of two nonoverlapping periods of data, perhaps caused by some intervention or time gap between the two periods. Trends may also be persistent or not persistent, and some trends may exhibit seasonality.

Some trend detection techniques require a continuous time-series of data. Thus, interruptions in the temporal data set must be eliminated for these detection techniques. Several methods are available for replacing missing data.

The true first step in trend analysis is actually exploratory data analysis (EDA) as described in chapter 3. Thus, the data should be examined using such techniques as stem-and-leaf diagrams and box-and-whisker plots. Transformations, such as the log transformation, of the data may be needed to bring out the trend as well as to meet certain statistical assumptions. Finally, some smoothing approaches may be useful in detecting trends.

In this chapter several techniques for trend detection are presented along with examples. Both parametric and nonparametric approaches are used. Generally, more than one trend method should be used when evaluating water quality data. The different techniques show trends in different ways. The existence of a trend does not mean causality. In fact, a major weakness of relying on trend analysis for an experimental design is that no causality can be inferred from a trend alone. The trend must be explained by other data in conjunction with the trend data. Hipel and McLeod (1994) present methods for testing causality between two time series.

## 615.1201 Missing data

Several techniques are used for dealing with missing data in a water quality data set. They include linear interpolation, regression analysis, and seasonal adjustment modeling. Linear interpolation may be appropriate if only one or two adjacent data points are missing. The missing data could be estimated by a linear interpolation between the known values before and after the gap. Regressions between water quality observations at different stations or between a water quality variable and flow may also be used to fill in missing data (Dunne and Leopold 1978). The gap in the missing data can be filled using the regression and the known independent values.

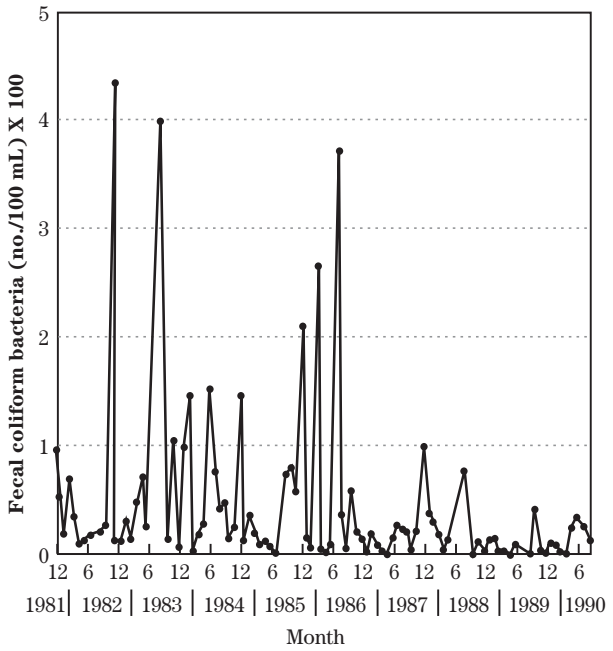
In seasonal adjustment modeling, the data are broken up into long-term, seasonal, and nonseasonal components (McLeod, et al. 1983). A missing data point is calculated from an equation representing the summation of influences related to long-term (median), stable seasonal (e.g., monthly), and irregular nonseasonal components.

## 615.1202 Time plots

Perhaps the first step in trend analysis is to plot the data versus time. Figure 12–1 is a time plot of fecal coliform bacteria counts obtained from Jewett Brook in the St. Albans Bay watershed from 1981 to 1990.

The fecal coliform data suggest that there is a monotonic decrease in fecal coliform abundance over time. The data also indicate that the variance of the data about this trend is also decreasing. A log transformation could be used to decrease the differences in the variance over time. The data may also show seasonality, but such variations are not obvious. Using just a time plot, the rate of the decrease cannot be obtained.

**Figure 12–1** Fecal coliform bacteria in Jewett Brook in the St. Albans Bay watershed, Vermont




## 615.1203 Least squares regression

A parametric regression analysis can be used to test the null hypothesis that the slope of the regression is 0 (i.e., no trend). This test requires the assumptions of normality, constant variance, and independence of errors. In example 12–1 the fecal coliform data previously described are tested using regression.

**Example 12–1** Determination of least squares regression of the fecal coliform data over time for Jewett Brook

Figure 12–2 contains frequency histograms and box plots for the fecal coliform and $\log_{10}$ fecal coliform data. The distribution and boxplots suggest that the untransformed data are not normally distributed and that the $\log_{10}$ transformed data are normally distributed. The univariate statistics are shown in table 12–1. Since the P-value for the untransformed data is less than 0.05, the data are not normally distributed. With a $\log_{10}$ transformation, the data appear to be normally distributed and the $\log_{10}$ transformed will be used for further analysis. Tests for normality are described in detail in chapter 4.

Figure 12–3 is a plot of the $\log_{10}$ transformed fecal coliform data as a function of month including a regression line.

The following linear regression equation was obtained using the statistical package JMP (SAS 1995):

Log fecal = 2.673 (month) – 0.0074

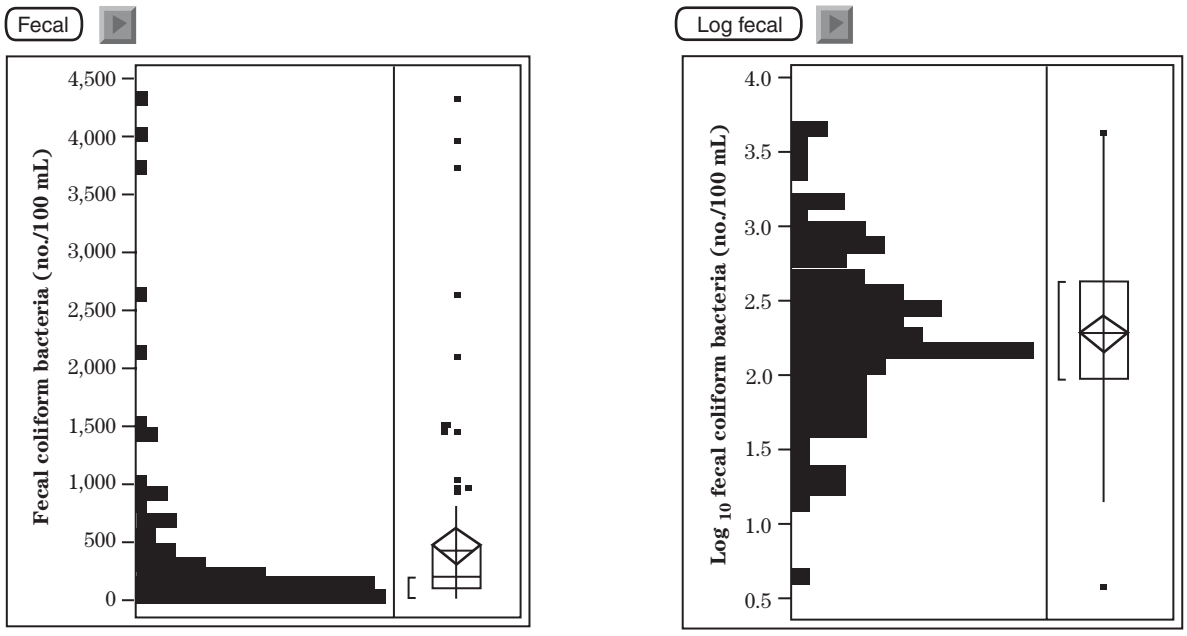
The analysis of variance for the regression is shown in table 12–2. The ANOVA indicates that the regression is significant. The $H_o$: slope = 0 is rejected. Also, based on the *t*-statistic, the slope of the regression is significantly different from zero. The results of the *t*-test are shown in table 12–3.

Using the slope of –0.0074, the fecal coliform bacteria are decreasing at a rate of 0.98 colonies per month (antilog of –0.0074).

**Table 12–1** Univariate statistics for fecal coliform data

| | Untransformed | $\text{Log}_{10}$ transformed |
|---|---|---|
| Mean (No./100 mL) | 458 | 2.285 |
| Median (No./100 mL) | 190 | 2.280 |
| Shapiro-Wilk W | 0.550 | 0.985 |
| P<W | 0.000 | 0.786 |

**Figure 12–2** Frequency histograms and box plots for fecal coliform data from JMP



**Example 12–1** Determination of least squares regression of the fecal coliform data over time for Jewett Brook—Continued

Annual means were used for the fecal coliform data. Boxplots for each year are shown in figure 12–4. Generally, years might be expected to be different when the boxes do not overlap each other, as for 1984 versus 1989. An analysis of variance indicates that the means are different at p=0.05 (table 12–4).

To determine which means were different, annual means were compared using the Tukey-Kramer honestly significant difference (HSD) test (SAS 1995). Only the means for 1984 and 1989 were different. In this example the comparison of annual means does not show a definite trend, but rather a high year early and a low year later. Additional methods of trend analysis are recommended to further analyze the data.

**Figure 12–3** Regression of log fecal coliform data over time

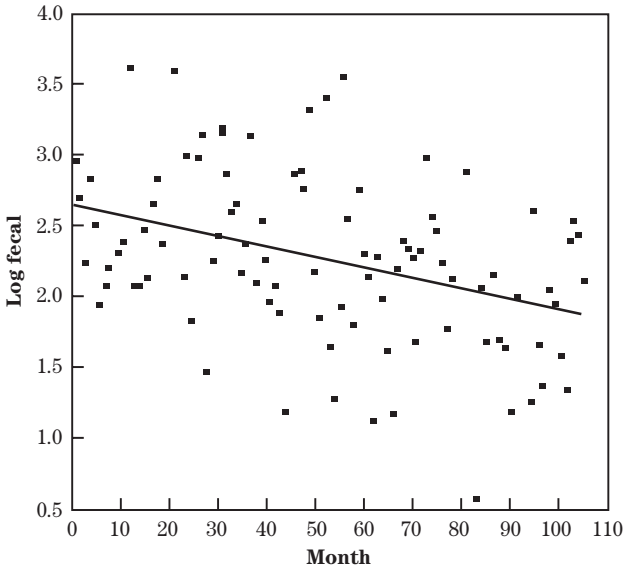



**Table 12–2** Analysis of variance for regression of log fecal coliform over time

| Source | DF | Sums of squares | Mean squares | F |
|---|---|---|---|---|
| Model | 1 | 4.750 | 4.750 | 16.378 |
| Error | 94 | 27.265 | 0.290 | |
| | | | P>F | 0.0001 |

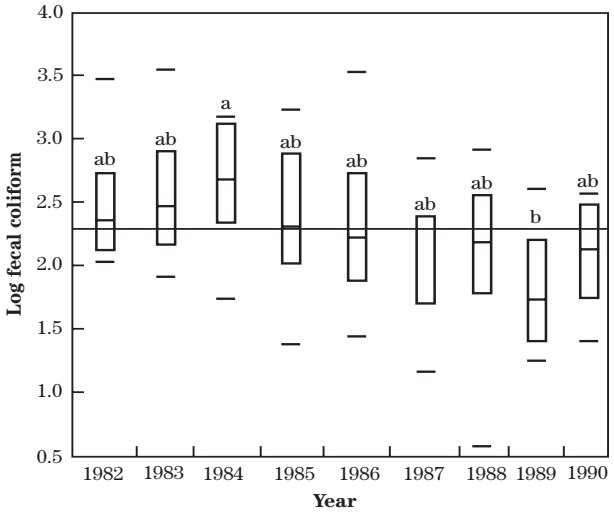
**Table 12–3** T-test of slope different from zero for fecal coliform trend data

| Term | Estimate | Std error | t ratio | Prob>t |
|---|---|---|---|---|
| Intercept | 2.673 | 0.111 | 24.18 | 0.0000 |
| Month | –0.0074 | 0.002 | –4.05 | 0.0001 |

**Table 12–4** Analysis of variance across years for fecal coliform data

| Source | DF | Sums of squares | Mean squares | F |
|---|---|---|---|---|
| Model | 9 | 6.626 | 0.762 | 2.494 |
| Error | 86 | 25.390 | 0.295 | |
| | | | P>F | 0.0139 |

**Figure 12–4** Annual boxplots for fecal coliform data [1/]




[1] Boxes with the same letter are not significantly different at p=0.05.

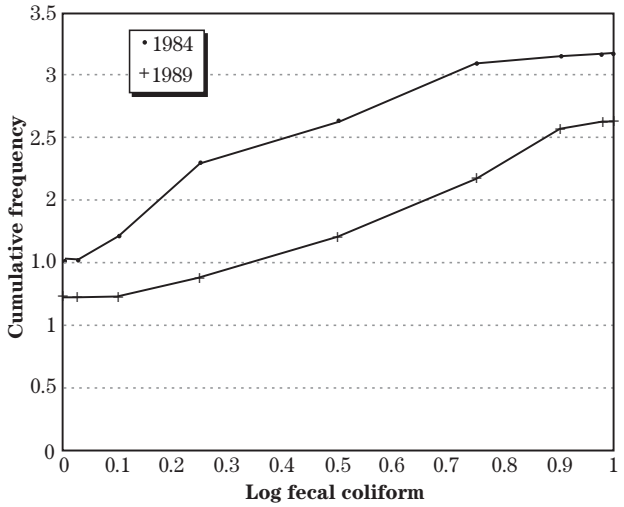
## 615.1204 Comparison of annual means

A comparison of means may be used to infer trends. Means across years, or some other unit of time such as every 2 or 3 years, may be compared for the analysis. The decision of what time unit to use is based partly on the degrees of freedom for the time unit as well as some scientific reasoning for dividing the time series. It is important that the units of time be equal for the analysis (UNESCO 1978).

## 615.1205 Cumulative distribution curves

The comparison of cumulative distribution curves may also be used to determine trends. Using the fecal coliform data, cumulative distribution curves were created for each year. By comparing the various curves, such as the 1984 curve to the 1989 curve, the decrease in fecal coliform bacteria is evident from 1984 to 1989 (fig. 12–5). These data could be tested using the Kolmogorov-Smirnov Goodness of fit (Zar 1996). The differences between these two curves is partly because of their individual means.

**Figure 12–5** Cumulative frequency curves for the fecal coliform data



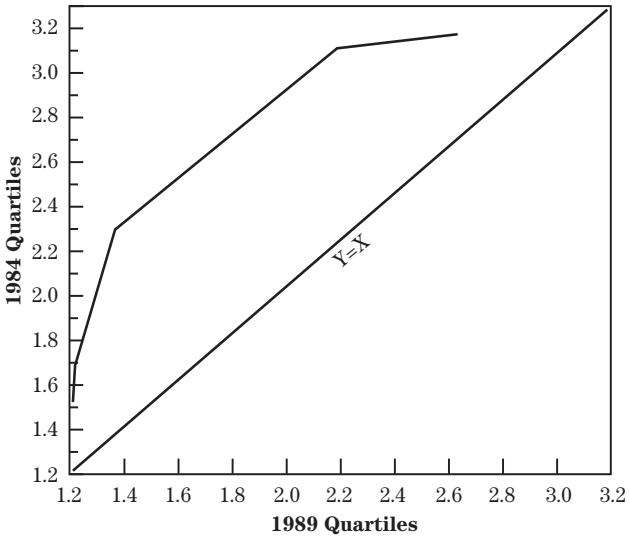
## 615.1206 Q-Q plots

For a Q-Q plot, the percentile (quartile) of one data set is plotted against another. For distributions that are similar, the points should follow along a line defined by Y = X (UNESCO 1978). The 1984 fecal coliform data in table 12–5 are plotted against the 1989 data in figure 12–6. The 1984 quartiles are clearly higher than the 1989 quartiles, indicating a trend toward decreasing fecal coliform in the stream.

**Table 12–5** Univariate statistics for fecal coliform data for 1984 and 1989

| Quartile | $Log_{10}$ fecal coliform (No./100mL) 1984 | 1989 |
|---|---|---|
| 0% (min) | 1.53 | 1.23 |
| 10% | 1.73 | 1.24 |
| 25% | 2.32 | 1.39 |
| 50% | 2.66 | 1.72 |
| 75% | 3.12 | 2.19 |
| 100% (max) | 3.18 | 2.63 |

**Figure 12–6** Q-Q plot of $log_{10}$ fecal coliform data for 1984 and 1989




## 615.1207 Double mass analysis

Double-mass curves are plots of accumulated values for a water quality station of interest as a function of an average from a number of stations or a control or reference station. This type of trend analysis requires data from several different locations, preferably in close proximity to each other.

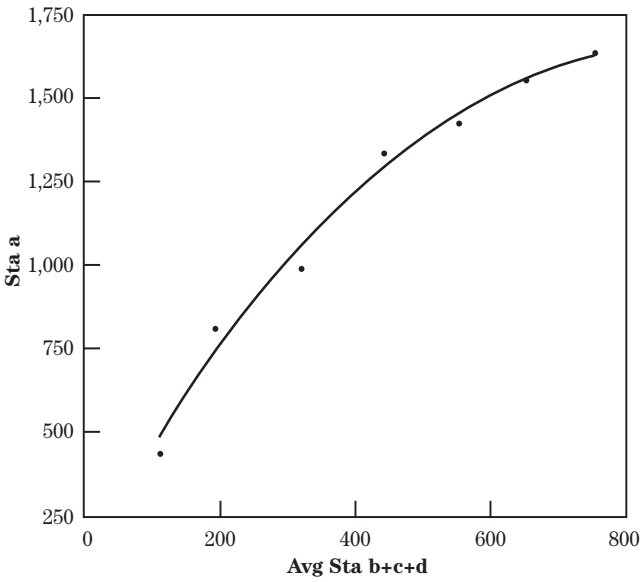
Double mass analysis is commonly used to assess changes in precipitation stations (Dunne and Leopold 1978). It is a visual tool that can be used to describe changes in one station in reference to a control station(s). A break in the slope of the line may indicate a trend or intervention. A comparison of slopes can be evaluated statistically (chapter 10) by analysis of covariance as pointed out by Dingman (1994).

A double mass curve of the fecal coliform data is shown in figure 12–7. In this case the cumulative coliform counts from a watershed receiving animal waste treatment (Sta a) are plotted as a function of the average among several stations (Sta b+c+d) that did not have watershed treatments. From this example the double mass analysis indicates that fecal coliform levels have fallen off gradually as compared to the average at the other three stations. A series of plots could be developed to check the other stations for trends (e.g., b vs. a+c+d, etc.)

## 615.1208 Paired regression analysis

Paired regressions can be used to infer trends if the data from two stations, one a control and one a treatment, are grouped into before and after time periods. Such data analysis was described in detail in chapter 10. A significant change in the paired regressions could signify a trend.

**Figure 12–7** Double-mass analysis of fecal coliform data




## 615.1209 Nonparametric approaches

Several nonparametric approaches are used in trend detection. The primary advantages of nonparametric approaches are that there are no assumptions regarding the distribution, censored data, outlyers, and missing data (Hirsch, et al. 1982). However, both parametric and nonparametric approaches assume that the data are not autocorrelated (i.e., that one observation is not related to the next observation).

### (a) Kendall's tau

Kendall's tau is a measure of correlation between a water quality variable and time for monotonic trends (Helsel and Hirsch 1992). Like most nonparametric approaches the procedure is based on rank, rather than the actual values. Although the calculation of tau is on many statistical packages (chapter 13), in example 12–2 a hand calculation is performed.

When seasonality or flow effects are removed from the trend, Spearman's rho test may be superior to the Kendall test (Hipel and McLeod 1994)

**Example 12–2** Kendall's tau for August fecal streptococcus data

The fecal streptococcus data from Jewett Brook used in the previous example was used for this example. To simplify the calculations, only the data for August is used (table 12–6).

The null hypothesis is that there is no correlation (trend) between bacteria level and time. The alternative hypothesis is that they are correlated. Kendall's S is calculated from:

$$S = P - M \qquad [12–1]$$

where:
P = number of pluses or the number of times the y's increase as the x's increase
M = number of minuses or the number of times the y's decrease as the x's increase (Helsel and Hirsch 1982)

**Table 12–6** Fecal streptococcus data for August from Jewett Brook, St. Albans Bay Watershed, VT

| Date | No./100 mL |
|---|---|
| 8/82 | 200 |
| 8/83 | 4,000 |
| 8/84 | 430 |
| 8/85 | 390 |
| 8/86 | 370 |
| 8/87 | 237 |
| 8/88 | 790 |
| 8/89 | 60 |
| 8/90 | 140 |

**Table 12–7** Summary of pluses and minuses for calculation of Kendall's tau for the August fecal streptococcus data for Jewett Brook, St. Albans Bay watershed, VT

| 200 | 4,000 | 430 | 390 | 370 | 237 | 790 | 60 | 140 |
|---|---|---|---|---|---|---|---|---|
| + | – | – | – | – | + | – | + | |
| + | – | – | – | + | – | – | | |
| + | – | – | + | – | – | | | |
| + | – | + | – | – | | | | |
| + | – | – | – | | | | | |
| + | – | – | | | | | | |
| – | – | | | | | | | |
| – | | | | | | | | |

To calculate the values, first compare 200/100 mL to all other values. For example, since 4,000 is greater than 200, a + is recorded, then 430 is greater than 200, a + is recorded, and so on. This can be summarized in a matrix format (table 12–7). Summing the pluses and minuses yields 11 P's and 25 M's.

$$S = 11 - 25 = -14.$$

Tau is calculated from:

$$\tau = \frac{S}{n\frac{(n-1)}{2}} \qquad [12–2]$$

$$\tau = \frac{-14}{9\frac{(9-1)}{2}} = -0.389$$

From appendix H, for S = (x)=–14 and n=9, p = 2 x 0.090 = 0.180. Because the calculated tau is greater than the table tau, the null hypothesis of no change is rejected because tau is significantly different from zero. The alternative hypothesis that there is a significant trend is accepted.

For a data set with seasonality (for example, months across years are different), the seasonal Kendall test may be used (Hirsch et al. 1982). For each season a separate S is calculated. They then are summed across seasons.

A seasonal slope estimator (B) can be calculated as the median of all the slopes between all possible data pairs within the same season (Helsel and Hirsch 1992). The individual slopes are calculated using equation 12–3 (Hirsch, et al. 1982):

$$d_{ijk} = \frac{\left(x_{ij} - x_{ik}\right)}{j-k} \qquad [12–3]$$

where:
i = 1, 2, …, 12 months
j = k+1, 2, …, n years
k = 1, 2, …, n–1 years

The slope estimator is determined in chapter 13 in the WQStat II package.

## 615.1210 Summary

Table 12–8 summarizes the trend methods described in this chapter and whether they are suitable for missing data, censored data, or seasonality.

**Table 12–8** Summary of trend detection techniques

| Trend method | Missing data | Censored data | Seasonality | Comments |
|---|---|---|---|---|
| Time plot | ok | ok | ok | |
| Least squares regression | ok | no | no | |
| Annual means | ok | no | no | |
| Cumulative distribution | ok | no | no | |
| Q-Q plots | ok | no | no | |
| Double mass analysis | ok | no | no | |
| Paired regressions | ok | no | ok | |
| Nonparametric seasonal Kendall | ok | ok | ok | Distribution free |

## 615.1211 References

Dingman, S.L. 1994. Physical hydrology. Prentice Hall. Englewood Cliffs, NJ.

Dunne, T., and L.B. Leopold. 1978. Water in environmental planning. W. H. Freeman and Co., San Francisco, CA.

Erlebach, W.E. 1978. A systematic approach to monitoring trends in the quality of surface waters. *In* Establishment of water quality monitoring programs, American Water Resources Association, pp. 7-19.

Helsel, D.R., and R.M. Hirsch. 1992. Chapter 12. Trend analysis. *In* Statistical methods in water resources, Studies in Environmental Science 49, Elsevier, New York, NY.

Hipel, K.W., and A.I. McLeod. 1994. Time series modeling of water resources and environmental systems. Elsevier Science BV. Amsterdam, The Neatherlands.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. Techniques of trend analysis for monthly water quality data. Water Resources Research 18(1):107-121.

McLeod, A.I., K.W. Hipel, and F. Comancho. 1983. Trend assessment of water quality time series. Water Resources Bulletin 19(4):537-547.

SAS Institute, Inc. 1995. Statistics and graphical guide. Ver. 3.1, SAS Institute, Inc., Cary, NC.

UNESCO/WHO. 1978. Water quality surveys - a guide for the collection and interpretation of water quality data. IHD-WHO Working Group on Quality of Water, UNESCO/WHO, Geneva, Switzerland.

Zar, J.H. 1996. Biostatistical analysis. Prentice Hall. Upper Saddle River, NJ.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

Part 615
National Water Quality Handbook

# Chapter 13 Statistical Packages

(450–VI–NWQH, September 2003)

# Chapter 13 Statistical Packages

Contents:



Tables

# Chapter 13 Statistical Packages

## 615.1300 Introduction

Several statistical software packages were developed specifically to aid in water quality data analysis. These packages include WQStat II (Loftis 1989), DETECT (Cluis 1989), SDS (Gaugush 1993), and ESTREND (Shertz, et al. 1991). In addition, numerous statistical software packages are available to assist in data analysis of most any type data. This chapter describes the packages available for water quality data analysis so that their usefulness for your particular situation can be determined. Statistical packages generally available for personal computers are described as well.

## 615.1301 Sample size and sampling frequency estimator

A sample size estimator has been developed by Region 6 of the United States Environmental Protection Agency (USEPA). This Windows program is downloadable free of charge from:

**www.epa.gov/Arkansas/6wq/ecopro/watershd/monitrng/sampling/sampling.htm**

This program estimates sample sizes for linear and step trends, estimation of means and differences between means. One major advantage of the software is that it performs iterative procedures.

# 615.1302 Water quality statistical software

## (a) WQStat Plus

WQStat II was developed at Colorado State University (Loftis 1989, Ward, et al. 1990). The most recent version is WQStat Plus. The package is IBM-PC compatible and includes both data management and data analysis capabilities. Although data for any frequency of time series can be used in this program, WQStat creates either a monthly or quarterly data file for analysis. Data can be either manually entered, or the program can read various files.

The following summary statistics are provided by WQStat Plus as part of an exploratory data analysis (EDA):

- mean
- median
- standard deviation
- number of data points
- skewness and significance
- kurtosis and significance
- frequency histogram
- correlogram (autocorrelation)

A time series plot also can be obtained as well as indicators of seasonality:

- seasonal box-and-whisker plot
- annual box-and-whisker plot
- Kruskal-Wallis test for seasonality

For trend detection the program determines:

- Kendall tau
- seasonal Kendall test
- seasonal Kendall slope estimator
- analysis of covariance

An analysis is also provided across groups using medians. This analysis allows comparison of sites or time periods within a single site. The following nonparametric approaches are used:

- Wilcoxon Signed Rank test
- Mann-Whitney test
- Kruskall-Wallis test

The package also provides an analysis of extreme values, such as the proportion of values exceeding a standard.

Example 26–1 gives an application of WQStat Plus using the fecal streptococcus time series data for Jewett Brook in the St. Albans Bay watershed in Vermont, used in chapter 12.

WQStat Plus is available from Intelligent Decision Technologies, 203 South Main Street, Longmont, Colorado 80501, **www.idt-ltd.com**.

**Example 13–1** WQStat using the fecal streptococcus data from the St. Albans Bay RCWP

Monthly mean fecal coliform values for Jewett Brook for the period December 1981 through August 1990 were entered into WQStat using a Lotus 1-2-3 file. A plot of the time series is shown in figure 12–1 in chapter 12.

Following the WQStat main menu, the summary statistics shown in table 13–1 were obtained.

The package produces a time plot and a seasonal Box-and-Whisker Plot. The Box-and-Whisker data indicate that seasons (months) are not greatly different (table 13–2).

An annual Box-and-Whisker Plot is provided. For the fecal streptococcus data, the annual Box-and-Whisker Plot indicates that the median and quartiles appear to decrease with time (table 13–3).

The program produces histograms and correlogram plots. The autocorrelations output is presented in table 13–4.

The autocorrelations shown in table 13–4 indicate that no significant serial correlation exists within the data. The highest autocorrelation was for the lag 9-month period, but it was not significant.

The Kruskal-Wallis test for seasonality using medians indicated that seasonality was not significant in the fecal streptococcus data (table 13–5).

The Seasonal Kendall test for trend was used since there were more than 5 years of data (table 13–6).

For this example, WQStat indicated that there was a declining trend in fecal streptococcus in Jewett Brook of 30 organisms per 100 mL per year, which is significant.

**Table 13–1** WQStat Mean / Skew values for the fecal streptococcus data

| Mean | Skew values (No / 100 mL) |
|---|---|
| Mean | 458.010 |
| Median | 158.000 |
| Standard deviation | 783.943 |
| Number of data points | 96 |

Skew test for normality
(skew value = 3.400)

| Confidence level | Test | Significance |
|---|---|---|
| 98% | 3.400>0.579 | significant |
| 90% | 3.400>0.397 | significant |
| 80% | 3.400>0.306 | significant |

Kurtosis test for normality
(Kurtosis value = 15.11)

| Confidence level | Test | Significance |
|---|---|---|
| 98% | 15.11>4.42 | significant |
| 90% | 15.11>3.79 | significant |
| 80% | 15.11>3.53 | significant |

**Example 13–1** WQStat using the fecal streptococcus data from the St. Albans Bay RCWP—Continued

**Table 13–2** WQStat seasonal box and whiskers for the fecal streptococcus data

| Season | Minimum | Interquartile | Median | Interquartile | Maximum |
|---|---|---|---|---|---|
| 1/1-2/1 | 1.4E+01 | 1.3E+02 | 1.5E+02 | 5.3E+02 | 9.8E+02 |
| 2/1-3/1 | 7.9E+01 | 1.6E+02 | 2.1E+02 | 3.6E+02 | 1.5E+02 |
| 3/1-4/1 | 3.4E+01 | 5.5E+01 | 1.5E+02 | 6.9E+02 | 2.7E+03 |
| 4/1-5/1 | 2.5E+01 | 4.8E+01 | 6.5E+01 | 3.4E+02 | 4.8E+02 |
| 5/1-6/1 | 1.6E+01 | 2.1E+01 | 1.3E+02 | 2.9E+02 | 7.1E+02 |
| 6/1-7/1 | 8.4E+01 | 9.9E+01 | 1.5E+02 | 2.1E+02 | 1.5E+03 |
| 7/1-8/1 | 1.7E+01 | 1.8E+02 | 2.8E+02 | 7.6E+02 | 3.7E+03 |
| 8/1-9/1 | 1.4E+02 | 2.4E+02 | 4.0E+02 | 7.9E+02 | 4.0E+03 |
| 9/1-10/1 | 2.0E+01 | 6.9E+01 | 2.2E+02 | 4.8E+02 | 7.6E+02 |
| 10/1-11/1 | 4.0E+00 | 1.0E+02 | 2.1E+02 | 3.5E+02 | 8.0E+02 |
| 11/1-12/1 | 5.1E+01 | 1.7E+02 | 2.4E+02 | 4.2E+02 | 4.3E+03 |
| 12/1-1/1 | 2.6E+01 | 7.6E+01 | 1.5E+02 | 1.5E+03 | 2.1E+03 |

**Table 13–3** WQStat annual box and whiskers for the fecal streptococcus data

| Season | Minimum | Interquartile | Median | Interquartile | Maximum |
|---|---|---|---|---|---|
| 1981 | 9.6E+02 | 9.6E+02 | 9.6E+02 | 9.6E+02 | 9.6E+02 |
| 1982 | 9.8E+01 | 1.3E+02 | 2.2E+02 | 6.1E+02 | 4.3E+03 |
| 1983 | 7.6E+01 | 1.5E+02 | 2.9E+02 | 7.1E+02 | 4.0E+03 |
| 1984 | 3.4E+01 | 1.8E+02 | 4.6E+02 | 1.2E+03 | 1.5E+03 |
| 1985 | 1.7E+01 | 9.2E+01 | 2.0E+02 | 7.8E+02 | 2.1E+03 |
| 1986 | 2.1E+01 | 5.8E+01 | 1.6E+02 | 4.9E+02 | 3.7E+03 |
| 1987 | 1.4E+01 | 3.1E+01 | 1.9E+02 | 2.3E+02 | 1.0E+03 |
| 1988 | 4.0E+00 | 6.5E+01 | 1.5E+02 | 3.9E+02 | 7.9E+02 |
| 1989 | 1.7E+01 | 2.6E+01 | 5.3E+01 | 1.5E+02 | 4.3E+02 |
| 1990 | 2.5E+01 | 3.4E+01 | 1.3E+02 | 2.8E+02 | 3.6E+02 |

**Example 13–1** WQStat using the fecal streptococcus data from the St. Albans Bay RCWP—Continued

**Table 13–4** WQStat autocorrelations for the fecal streptococcus data

| | |
|---|---|
| Rho 1 : | –0.0232 |
| Rho 2 : | –0.0996 |
| Rho 3 : | 0.1379 |
| Rho 4 : | 0.0825 |
| Rho 5 : | –0.0045 |
| Rho 6 : | 0.0403 |
| Rho 7 : | 0.0529 |
| Rho 8 : | –0.0347 |
| Rho 9 : | 0.1988 |
| Rho10 : | 0.0816 |
| Rho 11 : | 0.0048 |
| Rho 12 : | 0.0235 |
| Rho 13 : | –0.0651 |
| Rho 14 : | –0.0469 |
| Rho 15 : | 0.0434 |
| Rho 16 : | –0.0051 |
| Rho 17 : | –0.0099 |
| Rho 18 : | –0.0379 |
| Rho 19 : | 0.1106 |
| Rho 20 : | 0.0232 |
| Rho 21 : | 0.0004 |
| Rho 22 : | 0.0219 |
| Rho 23 : | –0.0296 |
| Rho 24 : | –0.0175 |

Boundary value = 0.2041

**Table 13–5** WQStat Kruskal-Wallis test for seasonality for the fecal streptococcus data (test statistic = 11.62)

| Confidence level | Test | Significance |
|---|---|---|
| 95% | 11.62<19.68 | not significant |
| 90% | 11.62<17.28 | not significant |
| 75% | 11.62<13.70 | not significant |

**Table 13–6** WQStat seasonal Kendall test for the fecal streptococcus data (test statistic = –3.987)

| Confidence level | Test | Significance |
|---|---|---|
| 95% | -3.987<-1.960 | significant |
| 90% | -3.987<-1.645 | significant |
| 80% | -3.987<-1.282 | significant |

Seasonal Kendall slope estimate:
Slope = –30.00000 units/year

## (b) DETECT

The program DETECT was developed in Quebec, Canada, to utilize nonparametric approaches to detect trends in water quality data (Cluis 1989). This package is IBM-PC compatible and is somewhat directed toward Canada's national water quality data collection program (NAQUADAT). A typical input file contains the date (YY MM DD), the concentration, and the discharge (optional). Mass loading information may be input as well. The input file must be in columns with a row in a strict FORTRAN format:

(12X, I2, 1X, I2, 1X, I2, 16X, F12.6, F12.6)

This format is designed to read as: 12 spaces, YY, one space, MM, one space, DD, 16 spaces, concentration in 12 spaces with 6 following decimal, and discharge in 12 spaces with 6 following decimal (optional). Concentration data should be in milligrams per liter and discharge in cubic meters per second.

Graphic analysis includes a time plot, double-mass curves, and the CUSUM function. Double-mass curves show the accumulated sum of the concentration or discharge as a function of accumulated time (days from first observation). The CUSUM function, or cumulative sum, is the summation of the deviations of the observations from the mean plotted as a function of time.

$$\mathrm{CUSUM}(X_t) = \sum_{j=1}^{t} X_j - j(\bar{X}) \qquad [13–1]$$

where:
t = time (Cluis 1989, Hipel and McLeod 1994)

DETECT allows elimination of high and low outliers. Among the tests in DETECT is one for seasonality based on ANOVA. Missing values may be replaced using three different options:

- Temporal interpolation
- Seasonal mean
- Concentration-discharge relationship

Persistence in the trend is examined using autocorrelation coefficients. The appropriate test for trend recommended in the user's manual is suggested based on:

- Type of trend–monotonic or stepwise
- Persistence–Markovian or none
- Seasonality

The following trend tests are available:
Lettenmaier/Spearman (Lettenmaier 1976)
Hirsch and Slack (Hirsch and Slack 1984)
Spearman/Kendal (Helsel and Hirsch 1992)
Kendall seasonality (Helsel and Hirsch 1992)
Lettenmaier/Mann-Whitney (Lettenmaier 1976)
Mann-Whitney (Lettenmaier 1976)

Example 13–2 shows an application of DETECT using the fecal streptococcus time series data for Jewett Brook in the St. Albans Bay watershed in Vermont, used in chapter 12.

**Example 13–2** DETECT using the fecal streptococcus data from the St. Albans Bay RCWP

Monthly mean fecal coliform values for Jewett Brook for the period December 1981 through August 1990 were prepared for entry into DETECT by editing a file in a DOS editor to put it in the proper format. A plot of the time series generated by DETECT is shown in figure 13–1.

The outliers were not eliminated from the data set for this example. As indicated in the manual, nonparametric tests yield stable results even with outliers present.

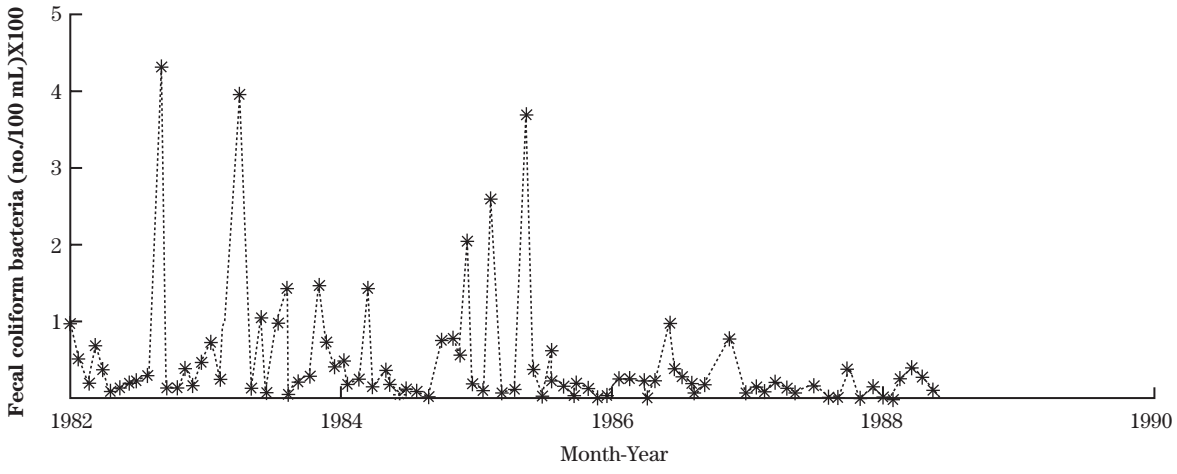
The double mass curve generated by DETECT is shown in figure 13–2. This plot shows the accumulated sum of fecal coliform abundance as a function of the accumulated time in months. The plot contains several lines. A general mean line is drawn from the origin to the upper right hand corner of the graph. The individual points are shown as X's. The general mean slope can be compared to groups of points in the double-mass curve. Slope of a group of points less than the general mean line indicates that the mean of these points would be less than the general mean. The lines above and below the mean line are termed *rails* and are two standard deviations from the mean line based on deviations from only its side of the line. Rails located far from the mean line indicate large variability in the data. If no trend is present, points on the double-mass curve are located on both sides of the mean line randomly. This in not the case in this example, indicating a trend is most likely present.

The CUSUM function is shown in figure 13–3. This plots the summation of the deviations from the general mean line in the previous figure.

Departures on one side of the line at Y=0 indicate a likely trend, as in this case. If the curve is parabolic, a monotonic linear trend is suggested. If the curve includes discontinuous lines, a stepwise trend is suggested. In this case a monotonic trend is suspected. The analysis of variance in table 13–7 tests whether monthly means are different as a test of seasonality.

The ANOVA in table 13–7 indicates no difference among months. Also, a Bartlett's test of the equality of variances is performed, which indicates in this case that the variances may not be equal across groups. Some data was missing in the fecal streptococcus data set, and the interpolation option was selected to fill missing data.

**Figure 13–1** Time series of fecal coliform data from DETECT



**Example 13–2** DETECT using the fecal streptococcus data from the St. Albans Bay RCWP—Continued

Autocorrelation correlation coefficients were used for an analysis of persistence (table 13–8). The autocorrelation coefficient is significant if the value is at least two times the standard deviation. In this case there was no significant persistence since no autocorrelations were significant. If the lag 1 r was significant and the lag 2 r was not, this would be termed Markovian persistence.

Using the decision tree in the program, the data displayed a monotonic trend without persistence or seasonality. Therefore, the Kendall test was used for analysis of the trend. Table 13–9 shows the results from Kendall's test as displayed by DETECT.

**Figure 13–2** Double-mass curve for the fecal streptococcus data

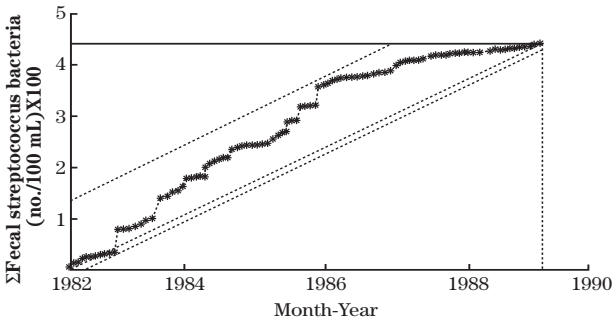



**Figure 13–3** CUSUM function for the fecal streptococcus data

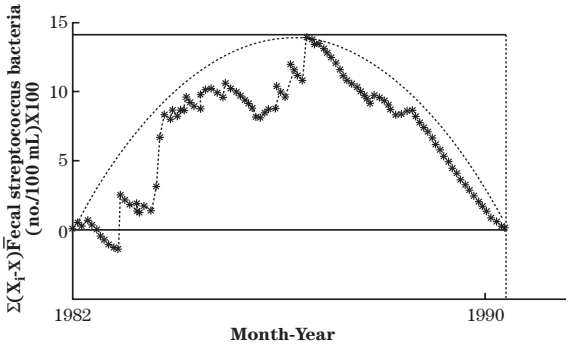



**Table 13–7** ANOVA table for equality of means for the fecal streptococcus data

| Source | df | MS | F |
|---|---|---|---|
| Month | 11 | 0.61241E-06 | 0.995 |
| Error | 84 | 0.61568E-06 | |
| Total | 95 | 0.61530E-06 | |

Equality of means accepted
No seasonality
Equality of variances is rejected!

**Table 13–8** Autocorrelation coefficients for the fecal streptococcus data

| | Lag | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| coeff. | 0.16 | –0.02 | 0.10 | 0.12 |
| std. dev | 0.10 | 0.10 | 0.10 | 0.10 |

**Table 13–9** DETECT Kendall's test for trend for the fecal streptococcus data

| | |
|---|---|
| statistic | –1375.63 |
| test value | –3.86 |
| signif. level | 0.00 |

Comment: Decreasing monotonic trend detected.

## (c) SDS

The Sampling Design Software (SDS) was developed by the U.S. Army Corps of Engineers. This software is used to determine sample sizes, variance components, optimization of stratified samples among strata, and clustering of groups to increase efficiency of sampling (Gaugush 1993).

The sample size determination can be based on multivariable sampling using either simple random or stratified sampling. The decision analysis is based on the mean, coefficient of variation, precision level, acceptable error, and the costs of sampling. The program displays the sample sizes and costs for each variable at different precision and error levels.

The variance component program determines the contribution to the variability in a water quality variable from different factors, such as station, date, and depth. The analysis attempts to determine which factors are most important in sampling and, therefore, which factors should dominate the design. For example, if most of the variance was explained by date, the station and depth subsampling could be reduced.

The number of samples applied to different strata can be optimized using error analysis in the program. The percent variance for each strata is compared to the percent of the number of samples and a percent optimum number of samples. Generally, more samples are allocated to strata with the higher variability.

Cluster analysis can be used to identify redundancy in the sampling program. For example, if a number of water quality stations are producing the same type of information, one or more could be dropped.

## (d) ESTREND

ESTREND (Shertz, et al. 1991) is used by the U.S. Geological Survey for nonparametric trend analysis at its various water quality stations. The program is written for UNIX and has been commonly used on Prime™ computers.

Table 13–10 summarizes the characteristics and capabilities of the various water quality statistical packages.

**Table 13–10** Summary of characteristics and capabilities of water quality statistical packages

| | WQStat II | DETECT | SDS |
|---|---|---|---|
| **Data manager** | | | |
| Data type | monthly, seasonal | monthly | summary |
| ASCII import | X | X | X |
| Lotus 1-2-3 import | X | | |
| Manual entry | X | | |
| Missing data | | X | |
| **Data analyses** | | | |
| EDA | X | | |
| Trends | X | X | |
| Group comparisons | X | | |
| Extreme values | X | | |
| Autocorrelation | | X | |
| Sample sizes | | | X |
| **Graphics** | | | |
| time plot | X | X | |
| double-mass | | X | |
| CUSUM | | X | |

# 615.1303 General statistics

Many statistical packages are commercially available that will perform the statistical analyses described in NWQH Part 615. Table 13–11 provides a summary of some of the capabilities and features of these packages, and table 13–12 summarizes the statistical methods included in each package.

**Table 13–11** Summary of cost (2001) and capabilities of general statistical software packages

| Statistical package | Cost ($) | Win/ Mac | Graphics | Documentation | Comments |
|---|---|---|---|---|---|
| Analyse-it | 125 | W | X | on-line | Plug-in for MS Excel, www.analyse-it.com (note British spelling) |
| DataDesk | £399 1/ | W/M | X | manual | www.longman.net/datadesk-activstats |
| Instat | 79 | W/M | X | | www.graphpad.com |
| JMP | 895 | W/M | X | manual | www.jmpdiscovery.com |
| Quick Statistica | 495 | W | | | www.statsoft.com |
| SAS | | W | X | manual | Primarily for mainframe computers, www.sas.com |
| SPSS | 858 2/ | W/M | | | www.spss.com |
| Statistica | 1095 | W | | | www.statsoft.com |
| Statistix | 495 | W | | | www.statistix.com |
| SYSTAT | 1299 | W | | | www.spss.com |
| WINKS Basic | 99 | W | X | manual | www.texasoft.com/homepage |

1/ Price in British Pounds.
2/ GSA Schedule.

**Table 13–12** Summary of statistical methods included in software packages (O = blank)

| Package | Descriptive/ univariate | Boxplot | Test of normality | Regression/ correlation | t-test | ANOVA | Multiple comparisons | ANCOVA | Nonparametric |
|---|---|---|---|---|---|---|---|---|---|
| Analyse-it | X | X | X | X | X | X | O | O | X |
| DataDesk | X | X | X | X | X | X | X | X | X |
| Instat | X | O | X | X | X | X | X | O | X |
| JMP | X | X | X | X | X | X | X | X | X |
| Quick Statistica | X | X | X | X | X | X | O | X | X |
| SAS | X | X | X | X | X | X | X | X | X |
| SPSS | X | X | X | X | X | X | X | X | X |
| Statistica | X | X | X | X | X | X | X | X | X |
| Statistix | X | X | X | X | X | X | O | X | O |
| SYSTAT | X | X | X | X | X | X | O | X | X |
| WINKS | X | X | X | X | X | X | X | X | X |

## 615.1304 References

Cluis, D. 1989. Detection of trend and norm violation. User's manual—DETECT and EXCEED, Ver. 2, INRS-Eau, Quebec, Canada.

Gaugush, R.F. 1993. Sample design software user's manual. Instructions report W-93-1, Vicksburg, MS, U.S. Army Engineer Waterways Experiment Station.

Helsel, D.R., and R.M. Hirsch. 1992. Trend analysis, chapter 12. *In* Statistical Methods in Water Resources, Studies in Environmental Science 49, Elsevier, New York, NY.

Hipel, K.W., and A.I. McLeod. 1994. Time series modeling of water resources and environmental systems. Elsevier Science. The Neatherlands.

Hirsch, R.M., and J.R. Slack. 1984. A nonparametric trend test for seasonal data with serial dependence. Water Resources Research 20(6):727-732.

Lettenmaier, D.P. 1976. Detection of trend in water quality data from records with dependent observations. Water Resources Research 12(5):1037-1046.

Loftis, J.C. 1989. User's manual—WQStat II. A water quality statistics program, CO State Univ.

Shertz, T.L., R.B. Alexander, and D.J. Ohe. 1991. Computer program estimate trend (ESTREND). Water Resources Investigations WRI 91-4040.

Ward, R.C., J.C. Loftis, and G.B. McBride. 1990. Design of water quality monitoring systems. Van Nostrand Reinhold, New York.

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

# Part 615
# National Water Quality Handbook

# Appendixes

(450–VI–NWQH, September 2003)

# Appendixes

**Contents:**

## Appendix A Distribution of *t* (two-tailed) 1/

| Degrees of Freedom | Probability of a Larger Value, Sign Ignored | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.500 | 0.400 | 0.20 | 0.10 | 0.050 | 0.025 | 0.010 | 0.005 | 0.001 |
| 1 | 1.000 | 1.376 | 3.078 | 6.314 | 12.706 | 25.452 | 63.657 | | |
| 2 | 0.816 | 1.061 | 1.886 | 2.920 | 4.303 | 6.205 | 9.925 | 14.089 | 31.598 |
| 3 | .765 | 0.978 | 1.638 | 2.353 | 3.182 | 4.176 | 5.841 | 7.453 | 12.941 |
| 4 | .741 | .941 | 1.533 | 2.132 | 2.776 | 3.495 | 4.604 | 5.598 | 8.610 |
| 5 | .727 | .920 | 1.476 | 2.015 | 2.571 | 3.163 | 4.032 | 4.773 | 6.859 |
| 6 | .718 | .906 | 1.440 | 1.943 | 2.447 | 2.969 | 3.707 | 4.317 | 5.959 |
| 7 | .711 | .896 | 1.415 | 1.895 | 2.365 | 2.841 | 3.499 | 4.029 | 5.405 |
| 8 | .706 | .889 | 1.397 | 1.860 | 2.306 | 2.752 | 3.355 | 3.832 | 5.041 |
| 9 | .703 | .883 | 1.383 | 1.833 | 2.262 | 2.685 | 3.250 | 3.690 | 4.781 |
| 10 | .700 | .879 | 1.372 | 1.812 | 2.228 | 2.634 | 3.169 | 3.581 | 4.587 |
| 11 | .697 | .876 | 1.363 | 1.796 | 2.201 | 2.593 | 3.106 | 3.497 | 4.437 |
| 12 | .695 | .873 | 1.356 | 1.782 | 2.179 | 2.560 | 3.055 | 3.428 | 4.318 |
| 13 | .694 | .870 | 1.350 | 1.771 | 2.160 | 2.533 | 3.012 | 3.372 | 4.221 |
| 14 | .692 | .868 | 1.345 | 1.761 | 2.145 | 2.510 | 2.977 | 3.326 | 4.140 |
| 15 | .691 | .866 | 1.341 | 1.753 | 2.131 | 2.490 | 2.947 | 3.286 | 4.073 |
| 16 | .690 | .865 | 1.337 | 1.746 | 2.120 | 2.473 | 2.921 | 3.252 | 4.015 |
| 17 | .689 | .863 | 1.333 | 1.740 | 2.110 | 2.458 | 2.898 | 3.222 | 3.965 |
| 18 | .688 | .862 | 1.330 | 1.734 | 2.101 | 2.445 | 2.878 | 3.197 | 3.922 |
| 19 | .688 | .861 | 1.328 | 1.729 | 2.093 | 2.433 | 2.861 | 3.174 | 3.883 |
| 20 | .687 | .860 | 1.325 | 1.725 | 2.086 | 2.423 | 2.845 | 3.153 | 3.850 |
| 21 | .686 | .859 | 1.323 | 1.721 | 2.080 | 2.414 | 2.831 | 3.135 | 3.819 |
| 22 | .686 | .858 | 1.321 | 1.717 | 2.074 | 2.406 | 2.819 | 3.119 | 3.792 |
| 23 | .685 | .858 | 1.319 | 1.714 | 2.069 | 2.398 | 2.807 | 3.104 | 3.767 |
| 24 | .685 | .857 | 1.318 | 1.711 | 2.064 | 2.391 | 2.797 | 3.090 | 3.745 |
| 25 | .684 | .856 | 1.316 | 1.708 | 2.060 | 2.385 | 2.787 | 3.078 | 3.725 |
| 26 | .684 | .856 | 1.315 | 1.706 | 2.056 | 2.379 | 2.779 | 3.067 | 3.707 |
| 27 | .684 | .855 | 1.314 | 1.703 | 2.052 | 2.373 | 2.771 | 3.056 | 3.690 |
| 28 | .683 | .855 | 1.313 | 1.701 | 2.048 | 2.368 | 2.763 | 3.047 | 3.674 |
| 29 | .683 | .854 | 1.311 | 1.699 | 2.045 | 2.364 | 2.756 | 3,.038 | 3.659 |
| 30 | .683 | .854 | 1.310 | 1.697 | 2.042 | 2.360 | 2.750 | 3.030 | 3.646 |
| 35 | .682 | .852 | 1.306 | 1.690 | 2.030 | 2.342 | 2.724 | 2.996 | 3.591 |
| 40 | .681 | .851 | 1.303 | 1.684 | 2.021 | 2.329 | 2.704 | 2.971 | 3.551 |
| 45 | .680 | .850 | 1.301 | 1.680 | 2.014 | 2.319 | 2.690 | 2.952 | 3.520 |
| 50 | .680 | .849 | 1.299 | 1.676 | 2.008 | 2.310 | 2.678 | 2.937 | 3.496 |
| 55 | .679 | .849 | 1.297 | 1.673 | 2.004 | 2.304 | 2.669 | 2.925 | 3.476 |
| 60 | .679 | .848 | 1.296 | 1.671 | 2.000 | 2.299 | 2.660 | 2.915 | 3.460 |
| 70 | .678 | .847 | 1.294 | 1.667 | 1.994 | 2.290 | 2.648 | 2.899 | 3.435 |
| 80 | .678 | .847 | 1.293 | 1.665 | 1.989 | 2.284 | 2.638 | 2.887 | 3.416 |
| 90 | .678 | .846 | 1.291 | 1.662 | 1.986 | 2.279 | 2.631 | 2.878 | 3.402 |
| 100 | .677 | .846 | 1.290 | 1.661 | 1.982 | 2.276 | 2.625 | 2.871 | 3.390 |
| 120 | .677 | .845 | 1.289 | 1.658 | 1.980 | 2.270 | 2.617 | 2.860 | 3.373 |
| ∞ | .6745 | .8416 | 1.2816 | 1.6448 | 1.9600 | 2.2414 | 2.5758 | 2.8070 | 3.2905 |

1/ Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods, 7th ed. Iowa State Univ. Press, Ames. 

**Appendix B** Table for testing skewness (one-tailed) 1/

| Size of sample *n* | - - Percentage points - - 5% | 1% | Standard deviation |
|---|---|---|---|
| 25 | 0.711 | 1.061 | 0.4354 |
| 30 | .661 | .982 | .4052 |
| 35 | .621 | .921 | .3804 |
| 40 | .587 | .869 | .3596 |
| 45 | .558 | .825 | .3418 |
| 50 | .533 | .787 | .3264 |
| 60 | .492 | .723 | .3009 |
| 70 | .459 | .673 | .2806 |
| 80 | .432 | .631 | .2638 |
| 90 | .409 | .596 | .2498 |
| 100 | .389 | .567 | .2377 |
| 125 | .350 | .508 | .2139 |
| 150 | .321 | .464 | .1961 |
| 175 | .298 | .430 | .1820 |
| 200 | .280 | .403 | .1706 |
| 250 | .251 | .360 | .1531 |
| 300 | .230 | .329 | .1400 |
| 350 | .213 | .305 | .1298 |
| 400 | .200 | .285 | .1216 |
| 450 | .188 | .269 | .1147 |
| 500 | 0.179 | 0.255 | 0.1089 |

1/ Snedecor, G.W., and W.G. Cochran. 1980. Statistical methods, 7th ed. Iowa State Univ. Press, Ames. (No part of this appendix may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the prior written permission of the publisher.)

## Appendix C Values of F [1/]

| Denominator *df* | Probability of a larger *F* | Numerator *df* 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.050 | 161.40 | 199.50 | 215.70 | 224.60 | 230.20 | 234.00 | 236.80 | 238.90 | 240.50 |
| | 0.010 | 4052.00 | 4999.50 | 5403.00 | 5625.00 | 5764.00 | 5859.00 | 5928.00 | 5982.00 | 6022.00 |
| 2 | 0.050 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 9.35 | 19.37 | 19.38 |
| | 0.010 | 98.50 | 99.00 | 99.17 | 99.25 | 99.30 | 99.33 | 99.36 | 99.37 | 99.39 |
| 3 | 0.050 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 |
| | 0.010 | 34.12 | 30.82 | 29.46 | 28.71 | 28.24 | 27.91 | 27.67 | 27.49 | 27.35 |
| 4 | 0.050 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 |
| | 0.010 | 21.20 | 18.00 | 16.69 | 15.98 | 15.52 | 15.21 | 14.98 | 14.80 | 14.66 |
| 5 | 0.050 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 |
| | 0.010 | 16.26 | 13.27 | 12.06 | 11.39 | 10.97 | 10.67 | 10.46 | 10.29 | 10.16 |
| 6 | 0.050 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 |
| | 0.010 | 13.75 | 10.92 | 9.78 | 9.15 | 8.75 | 8.47 | 8.26 | 8.10 | 7.98 |
| 7 | 0.050 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 |
| | 0.010 | 12.25 | 9.55 | 8.45 | 7.85 | 7.46 | 7.19 | 6.99 | 6.84 | 6.72 |
| 8 | 0.050 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 |
| | 0.010 | 11.26 | 8.65 | 7.59 | 7.01 | 6.63 | 6.37 | 6.18 | 6.03 | 5.91 |
| 9 | 0.050 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 |
| | 0.010 | 10.56 | 8.02 | 6.99 | 6.42 | 6.06 | 5.80 | 5.61 | 5.47 | 5.35 |
| 10 | 0.050 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 |
| | 0.010 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 |
| 11 | 0.050 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 |
| | 0.010 | 9.65 | 7.21 | 6.22 | 5.67 | 5.32 | 5.07 | 4.89 | 4.74 | 4.63 |
| 12 | 0.050 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 |
| | 0.010 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.64 | 4.50 | 4.39 |
| 13 | 0.050 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 |
| | 0.010 | 9.07 | 6.70 | 5.74 | 5.21 | 4.86 | 4.62 | 4.44 | 4.30 | 4.19 |
| 14 | 0.050 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 |
| | 0.010 | 8.88 | 6.51 | 5.56 | 5.04 | 4.69 | 4.46 | 4.28 | 4.14 | 4.03 |

*See footnote at end of table.*

**Appendix C** Values of F [1/] —Continued

| Numerator *df* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ | P |
| 241.90 | 243.90 | 245.90 | 248.00 | 249.10 | 250.10 | 251.10 | 252.20 | 253.30 | 254.30 | 0.050 |
| 6056.00 | 6106.00 | 6157.00 | 6209.00 | 6235.00 | 6261.00 | 6287.00 | 6313.00 | 6339.00 | 6366.00 | 0.010 |
| 19.40 | 19.41 | 19.43 | 19.45 | 19.45 | 19.46 | 19.47 | 19.48 | 19.49 | 19.50 | 0.050 |
| 99.40 | 99.42 | 99.43 | 99.45 | 99.46 | 99.47 | 99.47 | 99.48 | 99.49 | 99.50 | 0.010 |
| 8.79 | 8.74 | 8.70 | 8.66 | 8.64 | 8.62 | 8.59 | 8.57 | 8.55 | 8.53 | 0.050 |
| 27.23 | 27.05 | 26.87 | 26.69 | 26.60 | 26.50 | 26.41 | 26.32 | 26.22 | 26.13 | 0.010 |
| 5.96 | 5.91 | 5.86 | 5.80 | 5.77 | 5.75 | 5.72 | 5.69 | 5.66 | 5.63 | 0.050 |
| 14.55 | 14.37 | 14.20 | 14.02 | 13.93 | 13.84 | 13.75 | 13.63 | 13.56 | 13.46 | 0.010 |
| 4.74 | 4.68 | 4.62 | 4.56 | 4.53 | 4.50 | 4.46 | 4.43 | 4.40 | 4.36 | 0.050 |
| 10.05 | 9.89 | 9.72 | 9.55 | 9.47 | 9.38 | 9.29 | 9.20 | 9.11 | 9.02 | 0.010 |
| 4.06 | 4.00 | 3.94 | 3.87 | 3.84 | 3.81 | 3.77 | 3.74 | 3.70 | 3.67 | 0.050 |
| 7.87 | 7.72 | 7.56 | 7.40 | 7.31 | 7.23 | 7.14 | 7.06 | 6.97 | 6.88 | 0.010 |
| 3.64 | 3.57 | 3.51 | 3.44 | 3.41 | 3.38 | 3.34 | 3.30 | 3.27 | 3.23 | 0.050 |
| 6.62 | 6.47 | 6.31 | 6.16 | 6.07 | 5.99 | 5.91 | 5.82 | 5.74 | 5.65 | 0.010 |
| 3.35 | 3.28 | 3.22 | 3.15 | 3.12 | 3.08 | 3.04 | 3.01 | 2.97 | 2.93 | 0.050 |
| 5.81 | 5.67 | 5.52 | 5.36 | 5.28 | 5.20 | 5.12 | 5.03 | 4.95 | 4.86 | 0.010 |
| 3.14 | 3.07 | 3.01 | 2.94 | 2.90 | 2.86 | 2.83 | 2.79 | 2.75 | 2.71 | 0.050 |
| 5.26 | 5.11 | 4.96 | 4.81 | 4.73 | 4.65 | 4.57 | 4.48 | 4.40 | 4.31 | 0.010 |
| 2.98 | 2.91 | 2.85 | 2.77 | 2.74 | 2.70 | 2.66 | 2.62 | 2.58 | 2.54 | 0.050 |
| 4.85 | 4.71 | 4.56 | 4.41 | 4.33 | 4.25 | 4.17 | 4.08 | 4.00 | 3.91 | 0.010 |
| 2.85 | 2.79 | 2.72 | 2.65 | 2.61 | 2.57 | 2.53 | 2.49 | 2.45 | 2.40 | 0.050 |
| 4.54 | 4.40 | 4.25 | 4.10 | 4.02 | 3.94 | 3.86 | 3.78 | 3.69 | 3.60 | 0.010 |
| 2.75 | 2.69 | 2.62 | 2.54 | 2.51 | 2.47 | 2.43 | 2.38 | 2.34 | 2.30 | 0.050 |
| 4.30 | 4.16 | 4.01 | 3.86 | 3.78 | 3.70 | 3.62 | 3.54 | 3.45 | 3.36 | 0.010 |
| 2.67 | 2.60 | 2.53 | 2.46 | 2.42 | 2.38 | 2.34 | 2.30 | 2.25 | 2.21 | 0.050 |
| 4.10 | 3.96 | 3.82 | 3.66 | 3.59 | 3.51 | 3.43 | 3.34 | 3.25 | 3.17 | 0.010 |
| 2.54 | 2.53 | 2.46 | 2.39 | 2.35 | 2.31 | 2.27 | 2.22 | 2.18 | 2.13 | 0.050 |
| 3.94 | 3.80 | 3.66 | 3.51 | 3.43 | 3.35 | 3.27 | 3.18 | 3.09 | 3.00 | 0.010 |

**Appendix C** Values of F [1/] —Continued

| Denominator *df* | Probability of a larger *F* | Numerator *df* 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 0.050 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 |
| | 0.010 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.14 | 4.00 | 3.89 |
| 16 | 0.050 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 |
| | 0.010 | 8.53 | 6.23 | 5.29 | 4.77 | 4.44 | 4.20 | 4.03 | 3.89 | 3.78 |
| 17 | 0.050 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 |
| | 0.010 | 8.40 | 6.11 | 5.18 | 4.67 | 4.34 | 4.10 | 3.93 | 3.79 | 3.68 |
| 18 | 0.050 | 4.41 | 3.35 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 |
| | 0.010 | 8.29 | 6.01 | 5.09 | 4.58 | 4.25 | 4.01 | 3.84 | 3.71 | 3.60 |
| 19 | 0.050 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 |
| | 0.010 | 8.18 | 5.93 | 5.01 | 4.50 | 4.17 | 3.94 | 3.77 | 3.63 | 3.52 |
| 20 | 0.050 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 |
| | 0.010 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.70 | 3.56 | 3.46 |
| 21 | 0.050 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.37 |
| | 0.010 | 8.02 | 5.78 | 4.87 | 4.37 | 4.04 | 3.81 | 3.64 | 3.51 | 3.40 |
| 22 | 0.050 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.34 |
| | 0.010 | 7.95 | 5.72 | 4.62 | 4.31 | 3.99 | 3.76 | 3.59 | 3.45 | 3.35 |
| 23 | 0.050 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.32 |
| | 0.010 | 7.88 | 5.66 | 4.76 | 4.26 | 3.94 | 3.71 | 3.54 | 3.41 | 3.30 |
| 24 | 0.050 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 |
| | 0.010 | 7.82 | 5.61 | 4.72 | 4.22 | 3.90 | 3.67 | 3.50 | 3.36 | 3.26 |
| 25 | 0.050 | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.28 |
| | 0.010 | 7.77 | 5.57 | 4.68 | 4.18 | 3.85 | 3.63 | 3.46 | 3.32 | 3.22 |
| 26 | 0.050 | 4.23 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.39 | 2.32 | 2.27 |
| | 0.010 | 7.72 | 5.53 | 4.64 | 4.14 | 3.82 | 3.59 | 3.42 | 3.29 | 3.18 |
| 27 | 0.050 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.37 | 2.31 | 2.25 |
| | 0.010 | 7.68 | 5.49 | 4.60 | 4.11 | 3.78 | 3.56 | 3.39 | 3.26 | 3.15 |
| 28 | 0.050 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.45 | 2.36 | 2.29 | 2.24 |
| | 0.010 | 7.64 | 5.45 | 4.57 | 4.07 | 3.75 | 3.53 | 3.36 | 3.23 | 3.12 |

*See footnote at end of table.*

**Appendix C** Values of F [1/] —Continued

| Numerator *df* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ | P |
| 2.54 | 2.48 | 2.40 | 2.33 | 2.29 | 2.25 | 2.20 | 2.16 | 2.11 | 2.07 | 0.050 |
| 3.80 | 3.67 | 3.52 | 3.37 | 3.29 | 3.21 | 3.13 | 3.05 | 2.96 | 2.87 | 0.010 |
| 2.49 | 2.42 | 2.35 | 2.28 | 2.24 | 2.19 | 2.15 | 2.11 | 2.06 | 2.01 | 0.050 |
| 3.69 | 3.55 | 3.41 | 3.26 | 3.18 | 3.10 | 3.02 | 2.93 | 2.84 | 2.75 | 0.010 |
| 2.45 | 2.38 | 2.31 | 2.23 | 2.19 | 2.15 | 2.10 | 2.06 | 2.01 | 1.96 | 0.050 |
| 3.59 | 3.46 | 3.31 | 3.16 | 3.08 | 3.00 | 2.92 | 2.83 | 2.75 | 2.65 | 0.010 |
| 2.41 | 2.34 | 2.27 | 2.19 | 2.15 | 2.11 | 2.06 | 2.02 | 1.97 | 1.92 | 0.050 |
| 3.51 | 3.37 | 3.23 | 3.08 | 3.00 | 2.92 | 2.84 | 2.75 | 2.66 | 2.57 | 0.010 |
| 2.38 | 2.31 | 2.23 | 2.16 | 2.11 | 2.07 | 2.03 | 1.98 | 1.93 | 1.88 | 0.050 |
| 3.43 | 3.30 | 3.15 | 3.00 | 2.92 | 2.84 | 2.76 | 2.67 | 2.58 | 2.49 | 0.010 |
| 2.35 | 2.28 | 2.20 | 2.12 | 2.08 | 2.04 | 1.99 | 1.95 | 1.90 | 1.84 | 0.050 |
| 3.37 | 3.23 | 3.09 | 2.94 | 2.86 | 2.78 | 2.69 | 2.61 | 2.52 | 2.42 | 0.010 |
| 2.32 | 2.25 | 2.18 | 2.10 | 2.05 | 2.01 | 1.96 | 1.92 | 1.87 | 1.81 | 0.030 |
| 3.31 | 3.17 | 3.03 | 2.88 | 2.80 | 2.72 | 2.64 | 2.55 | 2.46 | 2.36 | 0.010 |
| 2.30 | 2.23 | 2.15 | 2.07 | 2.03 | 1.98 | 1.94 | 1.89 | 1.84 | 1.78 | 0.050 |
| 3.26 | 3.12 | 2.98 | 2.83 | 2.75 | 2.67 | 2.58 | 2.50 | 2.40 | 2.31 | 0.010 |
| 2.27 | 2.20 | 2.13 | 2.05 | 2.01 | 1.96 | 1.91 | 1.86 | 1.81 | 1.76 | 0.050 |
| 3.21 | 3.07 | 2.93 | 2.78 | 2.70 | 2.62 | 2.54 | 2.45 | 2.35 | 2.26 | 0.010 |
| 2.25 | 2.18 | 2.11 | 2.03 | 1.98 | 1.94 | 1.89 | 1.84 | 1.79 | 1.73 | 0.050 |
| 3.17 | 3.03 | 2.89 | 2.74 | 2.66 | 2.58 | 2.49 | 2.40 | 2.31 | 2.21 | 0.010 |
| 2.24 | 2.16 | 2.09 | 2.01 | 1.96 | 1.92 | 1.87 | 1.82 | 1.77 | 1.71 | 0.050 |
| 3.13 | 2.99 | 2.85 | 2.70 | 2.62 | 2.54 | 2.45 | 2.36 | 2.27 | 2.17 | 0.010 |
| 2.22 | 2.15 | 2.07 | 1.99 | 1.95 | 1.90 | 1.85 | 1.80 | 1.75 | 1.69 | 0.050 |
| 3.09 | 2.96 | 2.81 | 2.66 | 2.58 | 2.50 | 2.42 | 2.33 | 2.23 | 2.13 | 0.010 |
| 2.20 | 2.13 | 2.06 | 1.97 | 1.93 | 1.88 | 1.84 | 1.79 | 1.73 | 1.67 | 0.050 |
| 3.06 | 2.93 | 2.78 | 2.63 | 2.55 | 2.47 | 2.38 | 2.29 | 2.20 | 2.10 | 0.010 |
| 2.19 | 2.12 | 2.04 | 1.96 | 1.91 | 1.87 | 1.82 | 1.77 | 1.71 | 1.65 | 0.050 |
| 3.03 | 2.90 | 2.75 | 2.60 | 2.52 | 2.44 | 2.35 | 2.26 | 2.17 | 2.06 | 0.010 |

**Appendix C** Values of F 1/ —Continued

| Denominator *df* | Probability of a larger *F* | Numerator *df* 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 29 | .050 | 4.18 | 3.33 | 2.93 | 2.70 | 2.55 | 2.43 | 2.35 | 2.28 | 2.22 |
| | .010 | 7.60 | 5.42 | 4.54 | 4.04 | 3.73 | 3.50 | 3.33 | 3.20 | 3.09 |
| 30 | .050 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 |
| | .010 | 7.56 | 5.39 | 4.51 | 4.02 | 3.70 | 3.47 | 3.30 | 3.17 | 3.07 |
| 40 | .050 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.25 | 2.18 | 2.12 |
| | .010 | 7.31 | 5.18 | 4.31 | 3.83 | 3.51 | 3.29 | 3.12 | 2.99 | 2.89 |
| 60 | .050 | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 2.04 |
| | .010 | 7.08 | 4.98 | 4.13 | 3.65 | 3.34 | 3.12 | 2.95 | 2.82 | 2.72 |
| 120 | .050 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.17 | 2.09 | 2.02 | 1.96 |
| | .010 | 6.85 | 4.79 | 3.95 | 3.48 | 3.17 | 2.96 | 2.79 | 2.66 | 2.56 |
| × | .050 | 3.84 | 3.00 | 2.60 | 2.37 | 2.21 | 2.10 | 2.01 | 1.94 | 1.88 |
| | .010 | 6.63 | 4.61 | 3.78 | 3.32 | 3.02 | 2.80 | 2.64 | 2.51 | 2.41 |

1/ Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill, Inc., New York, NY. (Reproduced with permission of the McCraw-Hill Companies.)

**Appendix C** Values of F [1/] —Continued

| Numerator *df* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ | P |
| 2.18 | 2.10 | 2.03 | 1.94 | 1.90 | 1.85 | 1.81 | 1.75 | 1.70 | 1.64 | .050 |
| 3.00 | 2.87 | 2.73 | 2.57 | 2.49 | 2.41 | 2.33 | 2.23 | 2.14 | 2.03 | .010 |
| 2.16 | 2.09 | 2.01 | 1.93 | 1.89 | 1.84 | 1.79 | 1.74 | 1.68 | 1.62 | .050 |
| 2.98 | 2.84 | 2.70 | 2.55 | 2.47 | 2.39 | 2.30 | 2.21 | 2.11 | 2.01 | .010 |
| 2.08 | 2.00 | 1.92 | 1.84 | 1.79 | 1.74 | 1.69 | 1.64 | 1.58 | 1.51 | .050 |
| 2.80 | 2.66 | 2.52 | 2.37 | 2.29 | 2.20 | 2.11 | 2.02 | 1.92 | 1.80 | .010 |
| 1.99 | 1.92 | 1.84 | 1.75 | 1.70 | 1.65 | 1.59 | 1.53 | 1.47 | 1.39 | .050 |
| 2.63 | 2.50 | 2.35 | 2.20 | 2.12 | 2.03 | 1.94 | 1.84 | 1.73 | 1.60 | .010 |
| 1.91 | 1.83 | 1.75 | 1.66 | 1.61 | 1.55 | 1.50 | 1.43 | 1.35 | 1.25 | .050 |
| 2.47 | 2.34 | 2.19 | 2.03 | 1.95 | 1.86 | 1.76 | 1.66 | 1.53 | 1.38 | .030 |
| 1.83 | 1.75 | 1.67 | 1.57 | 1.52 | 1.46 | 1.39 | 1.32 | 1.22 | 1.00 | .050 |
| 2.32 | 2.18 | 2.04 | 1.88 | 1.79 | 1.70 | 1.59 | 1.47 | 1.32 | 1.00 | .010 |

**Appendix D** Critical Values of the Kruskal-Wallis *H* Distribution [1/]

| $n_1$ | $n_2$ | $n_3$ | ∝ = 0.10 | 0.05 | 0.02 | 0.01 | 0.005 | 0.002 | 0.001 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 2 | 4.571 | | | | | | |
| 3 | 2 | 1 | 4.286 | | | | | | |
| 3 | 2 | 2 | 4.500 | 4.714 | | | | | |
| 3 | 3 | 1 | 4.571 | 5.143 | | | | | |
| 3 | 3 | 2 | 4.556 | 5.361 | 6.250 | | | | |
| 3 | 3 | 3 | 4.622 | 5.600 | 6.489 | (7.200) | 7.200 | | |
| 4 | 2 | 1 | 4.500 | | | | | | |
| 4 | 2 | 2 | 4.458 | 5.333 | 6.000 | | | | |
| 4 | 3 | 1 | 4.056 | 5.208 | | | | | |
| 4 | 3 | 2 | 4.511 | 5.444 | 6.144 | 6.444 | 7.000 | | |
| 4 | 3 | 3 | 4.709 | 5.791 | 6.564 | 6.745 | 7.318 | 8.018 | |
| 4 | 4 | 1 | 4.167 | 4.967 | (6.667) | 6.667 | | | |
| 4 | 4 | 2 | 4.555 | 5.455 | 6.600 | 7.036 | 7.282 | 7.855 | |
| 4 | 4 | 3 | 4.545 | 5.598 | 6.712 | 7.144 | 7.598 | 8.227 | 8.909 |
| 4 | 4 | 4 | 4.654 | 5.692 | 6.962 | 7.654 | 8.000 | 8.654 | 9.269 |
| 5 | 2 | 1 | 4.200 | 5.000 | | | | | |
| 5 | 2 | 2 | 4.373 | 5.160 | 6.000 | 6.533 | | | |
| 5 | 3 | 1 | 4.018 | 4.960 | 6.044 | | | | |
| 5 | 3 | 2 | 4.651 | 5.251 | 6.124 | 6.909 | 7.182 | | |
| 5 | 3 | 3 | 4.533 | 5.648 | 6.533 | 7.079 | 7.636 | 8.048 | 8.727 |
| 5 | 4 | 1 | 3.987 | 4.985 | 6.431 | 6.955 | 7.364 | | |
| 5 | 4 | 2 | 4.541 | 5.273 | 6.505 | 7.205 | 7.573 | 8.114 | 8.591 |
| 5 | 4 | 3 | 4.549 | 5.656 | 6.676 | 7.445 | 7.927 | 8.481 | 8.795 |
| 5 | 4 | 4 | 4.619 | 5.657 | 6.953 | 7.760 | 8.189 | 8.868 | 9.168 |
| 5 | 5 | 1 | 4.109 | 5.127 | 6.145 | 7.309 | 8.182 | | |
| 5 | 5 | 2 | 4.623 | 5.338 | 6.446 | 7.338 | 8.131 | 6.446 | 7.338 |
| 5 | 5 | 3 | 4.545 | 5.705 | 6.866 | 7.578 | 8.316 | 8.809 | 9.521 |
| 5 | 5 | 4 | 4.523 | 5.666 | 7.000 | 7.823 | 8.523 | 9.163 | 9.606 |
| 5 | 5 | 5 | 4.940 | 5.780 | 7.220 | 8.000 | 8.780 | 9.620 | 9.920 |
| 6 | 1 | 1 | ----- | | | | | | |
| 5 | 2 | 1 | 4.200 | 4.822 | | | | | |
| 6 | 2 | 2 | 4.545 | 5.345 | 6.182 | 6.982 | | | |
| 5 | 3 | 1 | 3.909 | 4.855 | 6.236 | | | | |
| 5 | 3 | 2 | 4.682 | 5.348 | 6.227 | 6.970 | 7.515 | 8.182 | |
| 6 | 3 | 3 | 4.538 | 5.615 | 6.590 | 7.410 | 7.872 | 8.628 | 9.346 |

*See footnote at end of table.*

**Appendix D** Critical Values of the Kruskal-Wallis *H* Distribution [1/] —Continued

| $n_1$ | $n_2$ | $n_3$ | | ∝ = | 0.10 | 0.05 | 0.02 | 0.01 | 0.005 | 0.002 | 0.001 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | 4 | 1 | | | 4.038 | 4.947 | 6.174 | 7.106 | 7.614 | | |
| 6 | 4 | 2 | | | 4.494 | 5.340 | 6.571 | 7.340 | 7.846 | 8.494 | 8.827 |
| 5 | 4 | 3 | | | 4.604 | 5.610 | 6.725 | 7.500 | 8.033 | 8.918 | 9.170 |
| 5 | 4 | 4 | | | 4.595 | 5.681 | 6.900 | 7.795 | 8.381 | 9.167 | 9.861 |
| 6 | 5 | 1 | | | 4.128 | 4.990 | 6.138 | 7.182 | 8.077 | 8.515 | |
| 6 | 5 | 2 | | | 4.596 | 5.338 | 6.585 | 7.376 | 8.196 | 8.967 | 9.189 |
| 5 | 5 | 3 | | | 4.535 | 5.602 | 6.829 | 7.590 | 8.314 | 9.150 | 9.669 |
| 5 | 5 | 4 | | | 4.522 | 5.661 | 7.018 | 7.936 | 8.643 | 9.458 | 9.960 |
| 6 | 5 | 5 | | | 4.547 | 5.729 | 7.110 | 8.028 | 8.859 | 9.771 | 10.27I |
| 5 | 6 | 1 | | | 4.000 | 4.945 | 6.286 | 7.121 | 8.165 | 9.077 | 9.692 |
| 6 | 6 | 2 | | | 4.438 | 5.410 | 6.667 | 7.467 | 8.210 | 9.219 | 9.752 |
| 6 | 6 | 3 | | | 4.558 | 5.625 | 6.900 | 7.725 | 8.458 | 9.458 | 10.150 |
| 5 | 6 | 4 | | | 4.548 | 5.724 | 7.107 | 8.000 | 8.754 | 9.662 | 10.342 |
| 5 | 5 | 5 | | | 4.542 | 5.765 | 7.152 | 8.124 | 8.967 | 9.948 | 10.524 |
| 6 | 6 | 6 | | | 4.643 | 5.801 | 7.240 | 8.222 | 9.170 | 10.187 | 10.889 |
| 7 | 7 | 7 | | | 4.594 | 5.819 | 7.332 | 8.378 | 9.373 | 10.516 | 11.310 |
| 8 | 8 | 8 | | | 4.595 | 5.805 | 7.355 | 8.465 | 9.495 | 10.805 | 11.705 |
| 2 | 2 | 1 | 1 | | ----- | | | | | | |
| 2 | 2 | 2 | 1 | | 5.357 | 5.679 | | | | | |
| 2 | 2 | 2 | 2 | | 5.667 | 6.167 | (6.667) | 6.667 | | | |
| 3 | 1 | 1 | 1 | | ----- | | | | | | |
| 3 | 2 | 1 | 1 | | 5.143 | | | | | | |
| 3 | 2 | 2 | 1 | | 5.556 | 5.833 | 6.500 | | | | |
| 3 | 2 | 2 | 2 | | 5.544 | 6.333 | 6.978 | 7.133 | 7.533 | | |
| 3 | 3 | 1 | 1 | | 5.333 | 6.333 | | | | | |
| 3 | 3 | 2 | 1 | | 5.689 | 6.244 | 6.689 | 7.200 | 7.400 | | |
| 3 | 3 | 2 | 2 | | 5.745 | 6.527 | 7.182 | 7.636 | 7.873 | 8.018 | 8.455 |
| 3 | 3 | 3 | 1 | | 5.655 | 6.600 | 7.109 | 7.400 | 8.055 | 8.345 | |
| 3 | 3 | 3 | 2 | | 5.879 | 6.727 | 7.636 | 8.105 | 8.379 | 8.803 | 9.030 |
| 3 | 3 | 3 | 3 | | 6.026 | 7.000 | 7.872 | 8.538 | 8.897 | 9.462 | 9.513 |
| 4 | 1 | 1 | 1 | | ----- | | | | | | |
| 4 | 2 | 1 | 1 | | 5.250 | 5.833 | | | | | |
| 4 | 2 | 2 | 1 | | 5.533 | 6.133 | 6.667 | 7.000 | | | |
| 4 | 2 | 2 | 2 | | 5.755 | 6.545 | 7.091 | 7.391 | 7.964 | 8.291 | |
| 4 | 3 | 1 | 1 | | 5.067 | 6.178 | 6.711 | 7.067 | | | |

*See footnote at end of table.*

**Appendix D** Critical Valuesof the Kruskal-Wallis *H* Distribution 1/ —Continued

| $n_1$ | $n_2$ | $n_3$ | | a = | 0.10 | 0.05 | 0.02 | 0.01 | 0.005 | 0.002 | 0.001 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 3 | 2 | 1 | | 5.591 | 6.309 | 7.018 | 7.455 | 7.773 | 8.182 | |
| 4 | 3 | 2 | 2 | | 5.750 | 6.621 | 7.530 | 7.871 | 8.273 | 8.689 | 8.909 |
| 4 | 3 | 3 | 1 | | 5.589 | 6.545 | 7.485 | 7.758 | 8.212 | 8.697 | 9.182 |
| 4 | 3 | 3 | 2 | | 5.872 | 6.795 | 7.763 | 8.333 | 8.718 | 9.167 | 8.455 |
| 4 | 3 | 3 | 3 | | 6.016 | 6.984 | 7.995 | 8.659 | 9.253 | 9.709 | 10.016 |
| 4 | 4 | 1 | 1 | | 5.182 | 5.945 | 7.091 | 7.909 | 7.909 | | |
| 4 | 4 | 2 | 1 | | 5.568 | 6.386 | 7.364 | 7.886 | 8.341 | 8.591 | 8.909 |
| 4 | 4 | 2 | 2 | | 5.808 | 6.731 | 7.750 | 8.346 | 8.692 | 9.269 | 9.462 |
| 4 | 4 | 3 | 1 | | 5.692 | 6.635 | 7.660 | 8.231 | 8.583 | 9.038 | 9.327 |
| 4 | 4 | 3 | 2 | | 5.901 | 6.874 | 7.951 | 8.621 | 9.165 | 9.615 | 9.945 |
| 4 | 4 | 3 | 3 | | 6.019 | 7.038 | 8.181 | 8.876 | 9.495 | 10.105 | 10.467 |
| 4 | 4 | 4 | 1 | | 5.564 | 6.725 | 7.879 | 8.588 | 9.000 | 9.478 | 9.758 |
| 4 | 4 | 4 | 2 | | 5.914 | 6.957 | 8.157 | 8.871 | 9.486 | 10.043 | 10.429 |
| 4 | 4 | 4 | 3 | | 6.042 | 7.142 | 8.350 | 9.075 | 9.742 | 10.542 | 10.929 |
| 4 | 4 | 4 | 4 | | 6.088 | 7.235 | 8.515 | 9.287 | 9.971 | 10.809 | 11.338 |
| 2 | 1 | 1 | 1 | 1 | - - - - - | | | | | | |
| 2 | 2 | 1 | 1 | 1 | 5.785 | | | | | | |
| 2 | 2 | 2 | 1 | 1 | 6.250 | 6.750 | | | | | |
| 2 | 2 | 2 | 2 | 1 | 6.600 | 7.133 | (7.533) | 7.533 | | | |
| 2 | 2 | 2 | 2 | 2 | 6.982 | 7.418 | 8.073 | 8.291 | (8.727) | 8.727 | |
| 3 | 1 | 1 | 1 | 1 | - - - - - | | | | | | |
| 3 | 2 | 1 | 1 | 1 | 6.139 | 6.583 | | | | | |
| 3 | 2 | 2 | 1 | 1 | 6.511 | 6.800 | 7.400 | 7.600 | | | |
| 3 | 2 | 2 | 2 | 1 | 6.709 | 7.309 | 7.836 | 8.127 | 8.327 | 8.618 | |
| 3 | 2 | 2 | 2 | 2 | 6.955 | 7.682 | 8.303 | 8.682 | 8.985 | 9.273 | 9.364 |
| 3 | 3 | 1 | 1 | 1 | 6.311 | 7.111 | 7.467 | | | | |
| 3 | 3 | 2 | 1 | 1 | 6.600 | 7.200 | 7.892 | 8.073 | 8.345 | | |
| 3 | 3 | 2 | 2 | 1 | 6.788 | 7.591 | 8.258 | 8.576 | 8.924 | 9.167 | 9.303 |
| 3 | 3 | 2 | 2 | 2 | 7.026 | 7.910 | 8.667 | 9.115 | 9.474 | 9.769 | 10.026 |
| 3 | 3 | 3 | 1 | 1 | 6.788 | 7.576 | 8.242 | 8.424 | 8.848 | (9.455) | 9.455 |
| 3 | 3 | 3 | 2 | 1 | 6.910 | 7.769 | 8.590 | 9.051 | 9.410 | 9.769 | 9.974 |
| 3 | 3 | 3 | 2 | 2 | 7.121 | 8.044 | 9.011 | 9.505 | 9.890 | 10.330 | 10.637 |
| 3 | 3 | 3 | 3 | 1 | 7.077 | 8.000 | 8.879 | 9.451 | 9.846 | 10.286 | 10.549 |
| 3 | 3 | 3 | 3 | 2 | 7.210 | 8.200 | 9.267 | 9.876 | 10.333 | 10.838 | 11.171 |
| 3 | 3 | 3 | 3 | 3 | 7.333 | 8.333 | 9.467 | 10.200 | 10.733 | 10.267 | 11.667 |

1/ Zar, J.H. 1996. Biostatistical analysis. 3rd ed., Prentice Hall, Upper Saddle River, NJ 07458.

**Appendix E** Upper Percentage Points of the Studentized Range, $q_a = \frac{\bar{x}_{max} - \bar{x}_{min}}{s_{\bar{x}}}$ 1/

| Error df | α | p = number of treatment means | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 5 | .05 | 3.64 | 4.60 | 5.22 | 5.67 | 6.03 | 6.33 | 6.58 | 6.80 | 6.99 | 7.17 | 7.32 | 7.47 | 7.60 | 7.72 | 7.83 | 7.93 | 8.03 | 8.12 | 8.21 |
| | .01 | 5.70 | 6.97 | 7.80 | 8.42 | 8.91 | 9.32 | 9.67 | 9.97 | 10.24 | 10.48 | 10.70 | 10.89 | 11.08 | 11.24 | 11.40 | 11.55 | 11.68 | 11.81 | 11.93 |
| 6 | .05 | 3.46 | 4.34 | 4.90 | 5.31 | 5.63 | 5.89 | 6.12 | 6.32 | 6.49 | 6.65 | 6.79 | 6.92 | 7.03 | 7.14 | 7.24 | 7.34 | 7.43 | 7.51 | 7.59 |
| | .01 | 5.24 | 6.33 | 7.03 | 7.56 | 7.97 | 8.32 | 8.61 | 8.87 | 9.10 | 9.30 | 9.49 | 9.65 | 9.81 | 9.95 | 10.08 | 10.21 | 10.32 | 10.43 | 10.54 |
| 7 | .05 | 3.34 | 4.16 | 4.68 | 5.06 | 5.36 | 5.61 | 5.82 | 6.00 | 6.16 | 6.30 | 6.43 | 6.55 | 6.66 | 6.76 | 6.85 | 6.94 | 7.02 | 7.09 | 7.17 |
| | .01 | 4.95 | 5.92 | 6.54 | 7.01 | 7.37 | 7.68 | 7.94 | 8.17 | 8.37 | 8.55 | 8.71 | 8.86 | 9.00 | 9.12 | 9.24 | 9.35 | 9.46 | 9.55 | 9.65 |
| 8 | .05 | 3.26 | 4.04 | 4.53 | 4.89 | 5.17 | 5.40 | 5.60 | 5.77 | 5.92 | 6.05 | 6.18 | 6.29 | 6.39 | 6.48 | 6.57 | 6.65 | 6.73 | 6.80 | 6.87 |
| | .01 | 4.74 | 5.63 | 6.20 | 6.63 | 6.96 | 7.24 | 7.47 | 7.68 | 7.87 | 8.03 | 8.18 | 8.31 | 8.44 | 8.55 | 8.66 | 8.76 | 8.85 | 8.94 | 9.03 |
| 9 | .05 | 3.20 | 3.95 | 4.42 | 4.76 | 5.02 | 5.24 | 5.43 | 5.60 | 5.74 | 5.87 | 5.98 | 6.09 | 6.19 | 6.28 | 6.36 | 6.44 | 6.51 | 6.58 | 6.64 |
| | .01 | 4.60 | 5.43 | 5.96 | 6.35 | 6.66 | 6.91 | 7.13 | 7.32 | 7.49 | 7.65 | 7.78 | 7.91 | 8.03 | 8.13 | 8.23 | 8.32 | 8.41 | 8.49 | 8.57 |
| 10 | .05 | 3.15 | 3.88 | 4.33 | 4.65 | 4.91 | 5.12 | 5.30 | 5.46 | 5.60 | 5.72 | 5.83 | 5.93 | 6.03 | 6.11 | 6.20 | 6.27 | 6.34 | 6.40 | 6.47 |
| | .01 | 4.48 | 5.27 | 5.77 | 6.14 | 6.43 | 6.67 | 6.87 | 7.05 | 7.21 | 7.36 | 7.48 | 7.60 | 7.71 | 7.81 | 7.91 | 7.99 | 8.07 | 8.15 | 8.22 |
| 11 | .05 | 3.11 | 3.82 | 4.26 | 4.57 | 4.82 | 5.03 | 5.20 | 5.35 | 5.49 | 5.61 | 5.71 | 5.81 | 5.90 | 5.99 | 6.06 | 6.14 | 6.20 | 6.26 | 6.33 |
| | .01 | 4.39 | 5.14 | 5.62 | 5.97 | 6.25 | 6.48 | 6.67 | 6.84 | 6.99 | 7.13 | 7.25 | 7.36 | 7.46 | 7.56 | 7.65 | 7.73 | 7.81 | 7.88 | 7.95 |
| 12 | .05 | 3.08 | 3.77 | 4.20 | 4.51 | 4.75 | 4.95 | 5.12 | 5.27 | 5.40 | 5.51 | 5.62 | 5.71 | 5.80 | 5.88 | 5.95 | 6.03 | 6.09 | 6.15 | 6.21 |
| | .01 | 4.32 | 5.04 | 5.50 | 5.84 | 6.10 | 6.32 | 6.51 | 6.67 | 6.81 | 6.94 | 7.06 | 7.17 | 7.26 | 7.36 | 7.44 | 7.52 | 7.59 | 7.66 | 7.73 |
| 13 | .05 | 3.06 | 3.73 | 4.15 | 4.45 | 4.69 | 4.88 | 5.05 | 5.19 | 5.32 | 5.43 | 5.53 | 5.63 | 5.71 | 5.79 | 5.86 | 5.93 | 6.00 | 6.05 | 6.11 |
| | .01 | 4.26 | 4.96 | 5.40 | 5.73 | 5.98 | 6.19 | 6.37 | 6.53 | 6.67 | 6.79 | 6.90 | 7.01 | 7.10 | 7.19 | 7.27 | 7.34 | 7.42 | 7.48 | 7.55 |
| 14 | .05 | 3.03 | 3.70 | 4.11 | 4.41 | 4.64 | 4.83 | 4.99 | 5.13 | 5.25 | 5.36 | 5.46 | 5.55 | 5.64 | 5.72 | 5.79 | 5.85 | 5.92 | 5.97 | 6.03 |
| | .01 | 4.21 | 4.89 | 5.32 | 5.63 | 5.88 | 6.08 | 6.26 | 6.41 | 6.54 | 6.66 | 6.77 | 6.87 | 6.96 | 7.05 | 7.12 | 7.20 | 7.27 | 7.33 | 7.39 |
| 15 | .05 | 3.01 | 3.67 | 4.08 | 4.37 | 4.60 | 4.78 | 4.94 | 5.08 | 5.20 | 5.31 | 5.40 | 5.49 | 5.58 | 5.65 | 5.72 | 5.79 | 5.85 | 5.90 | 5.96 |
| | .01 | 4.17 | 4.83 | 5.25 | 5.56 | 5.80 | 5.99 | 6.16 | 6.31 | 6.44 | 6.55 | 6.66 | 6.76 | 6.84 | 6.93 | 7.00 | 7.07 | 7.14 | 7.20 | 7.26 |
| 16 | .05 | 3.00 | 3.65 | 4.05 | 4.33 | 4.56 | 4.74 | 4.90 | 5.03 | 5.15 | 5.26 | 5.35 | 5.44 | 5.52 | 5.59 | 5.66 | 5.72 | 5.79 | 5.84 | 5.90 |
| | .01 | 4.13 | 4.78 | 5.19 | 5.49 | 5.72 | 5.92 | 6.08 | 6.22 | 6.35 | 6.46 | 6.56 | 6.66 | 6.74 | 6.82 | 6.90 | 6.97 | 7.03 | 7.09 | 7.15 |
| 17 | .05 | 2.98 | 3.63 | 4.02 | 4.30 | 4.52 | 4.71 | 4.86 | 4.99 | 5.11 | 5.21 | 5.31 | 5.39 | 5.47 | 5.55 | 5.61 | 5.68 | 5.74 | 5.79 | 5.84 |
| | .01 | 4.10 | 4.74 | 5.14 | 5.43 | 5.66 | 5.85 | 6.01 | 6.15 | 6.27 | 6.38 | 6.48 | 6.57 | 6.66 | 6.73 | 6.80 | 6.87 | 6.94 | 7.00 | 7.05 |
| 18 | .05 | 2.97 | 3.61 | 4.00 | 4.28 | 4.49 | 4.67 | 4.82 | 4.96 | 5.07 | 5.17 | 5.27 | 5.35 | 5.43 | 5.50 | 5.57 | 5.63 | 5.69 | 5.74 | 5.79 |
| | .01 | 4.07 | 4.70 | 5.09 | 5.38 | 5.60 | 5.79 | 5.94 | 6.08 | 6.20 | 6.31 | 6.41 | 6.50 | 6.58 | 6.65 | 6.72 | 6.79 | 6.85 | 6.91 | 6.96 |
| 19 | .05 | 2.96 | 3.59 | 3.98 | 4.25 | 4.47 | 4.65 | 4.79 | 4.92 | 5.04 | 5.14 | 5.23 | 5.32 | 5.39 | 5.46 | 5.53 | 5.59 | 5.65 | 5.70 | 5.75 |
| | .01 | 4.05 | 4.67 | 5.05 | 5.33 | 5.55 | 5.73 | 5.89 | 6.02 | 6.14 | 6.25 | 6.34 | 6.43 | 6.51 | 6.58 | 6.65 | 6.72 | 6.78 | 6.84 | 6.89 |
| 20 | .05 | 2.95 | 3.58 | 3.96 | 4.23 | 4.45 | 4.62 | 4.77 | 4.90 | 5.01 | 5.11 | 5.20 | 5.28 | 5.36 | 5.43 | 5.49 | 5.55 | 5.61 | 5.66 | 5.71 |
| | .01 | 4.02 | 4.64 | 5.02 | 5.29 | 5.51 | 5.69 | 5.84 | 5.97 | 6.09 | 6.19 | 6.29 | 6.37 | 6.45 | 6.52 | 6.59 | 6.65 | 6.71 | 6.76 | 6.82 |
| 24 | .05 | 2.92 | 3.53 | 3.90 | 4.17 | 4.37 | 4.54 | 4.68 | 4.81 | 4.92 | 5.01 | 5.10 | 5.18 | 5.25 | 5.32 | 5.38 | 5.44 | 5.50 | 5.54 | 5.59 |
| | .01 | 3.96 | 4.54 | 4.91 | 5.17 | 5.37 | 5.54 | 5.69 | 5.81 | 5.92 | 6.02 | 6.11 | 6.19 | 6.26 | 6.33 | 6.39 | 6.45 | 6.51 | 6.56 | 6.61 |

**Appendix E** Upper Percentage Points of the Studentized Range, $q_a = \frac{\bar{x}_{max} - \bar{x}_{min}}{s_{\bar{x}}}$ 1/ — Continued

| Error df | α | p = number of treatment means | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 30 | .05 | 2.89 | 3.49 | 3.84 | 4.10 | 4.30 | 4.46 | 4.60 | 4.72 | 4.83 | 4.92 | 5.00 | 5.08 | 5.15 | 5.21 | 5.27 | 5.33 | 5.38 | 5.43 | 5.48 |
| | .01 | 3.89 | 4.45 | 4.80 | 5.05 | 5.24 | 5.40 | 5.54 | 5.65 | 5.76 | 5.85 | 5.93 | 6.01 | 6.08 | 6.14 | 6.20 | 6.26 | 6.31 | 6.36 | 6.41 |
| 40 | .05 | 2.86 | 3.44 | 3.79 | 4.04 | 4.23 | 4.39 | 4.52 | 4.63 | 4.74 | 4.82 | 4.91 | 4.98 | 5.05 | 5.11 | 5.16 | 5.22 | 5.27 | 5.31 | 5.36 |
| | .01 | 3.82 | 4.37 | 4.70 | 4.93 | 5.11 | 5.27 | 5.39 | 5.50 | 5.60 | 5.69 | 5.77 | 5.84 | 5.90 | 5.96 | 6.02 | 6.07 | 6.12 | 6.17 | 6.21 |
| 60 | .05 | 2.83 | 3.40 | 3.74 | 3.98 | 4.16 | 4.31 | 4.44 | 4.55 | 4.65 | 4.73 | 4.81 | 4.88 | 4.94 | 5.00 | 5.06 | 5.11 | 5.16 | 5.20 | 5.24 |
| | .01 | 3.76 | 4.28 | 4.60 | 4.82 | 4.99 | 5.13 | 5.25 | 5.36 | 5.45 | 5.53 | 5.60 | 5.67 | 5.73 | 5.79 | 5.84 | 5.89 | 5.93 | 5.98 | 6.02 |
| 120 | .05 | 2.80 | 3.36 | 3.69 | 3.92 | 4.10 | 4.24 | 4.36 | 4.48 | 4.56 | 4.64 | 4.72 | 4.78 | 4.84 | 4.90 | 4.95 | 5.00 | 5.05 | 5.09 | 5.13 |
| | .01 | 3.70 | 4.20 | 4.50 | 4.71 | 4.87 | 5.01 | 5.12 | 5.21 | 5.30 | 5.38 | 5.44 | 5.51 | 5.56 | 5.61 | 5.66 | 5.71 | 5.75 | 5.79 | 5.83 |
| × | .05 | 2.77 | 3.31 | 3.63 | 3.86 | 4.03 | 4.17 | 4.29 | 4.39 | 4.47 | 4.55 | 4.62 | 4.68 | 4.74 | 4.80 | 4.85 | 4.89 | 4.93 | 4.97 | 5.01 |
| | .01 | 3.64 | 4.12 | 4.40 | 4.60 | 4.76 | 4.88 | 4.99 | 5.08 | 5.16 | 5.23 | 5.29 | 5.35 | 5.40 | 5.45 | 5.49 | 5.54 | 5.57 | 5.61 | 5.65 |

1/ Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill, Inc., New York, NY. (Reproduced with permission of the McCraw-Hill Companies.)

## Appendix F Wilcoxon two-sample rank test (Mann-Whitney test) 1/

| $n_2$ = larger $n$ | $P$ | $n_1$ = smaller $n$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 4 | .05 | | | 10 | | | | | | | | | | | |
| | .01 | | | --- | | | | | | | | | | | |
| 5 | .05 | | 6 | 11 | 17 | | | | | | | | | | |
| | .01 | | --- | --- | 15 | | | | | | | | | | |
| 6 | .05 | | 7 | 12 | 18 | 26 | | | | | | | | | |
| | .01 | | --- | 10 | 16 | 23 | | | | | | | | | |
| 7 | .05 | | 7 | 13 | 20 | 27 | 36 | | | | | | | | |
| | .01 | | --- | 10 | 17 | 24 | 32 | | | | | | | | |
| 8 | .05 | 3 | 8 | 14 | 21 | 29 | 38 | 49 | | | | | | | |
| | .01 | --- | --- | 11 | 17 | 25 | 34 | 43 | | | | | | | |
| 9 | .05 | 3 | 8 | 15 | 22 | 31 | 40 | 51 | 63 | | | | | | |
| | .01 | --- | 6 | 11 | 18 | 26 | 35 | 45 | 56 | | | | | | |
| 10 | .05 | 3 | 9 | 15 | 23 | 32 | 42 | 53 | 65 | 78 | | | | | |
| | .01 | --- | 6 | 12 | 19 | 27 | 37 | 47 | 58 | 71 | | | | | |
| 11 | .05 | 4 | 9 | 16 | 24 | 34 | 44 | 55 | 68 | 81 | 96 | | | | |
| | .01 | --- | 6 | 12 | 20 | 28 | 38 | 49 | 61 | 74 | 87 | | | | |
| 12 | .05 | 4 | 10 | 17 | 26 | 35 | 46 | 58 | 71 | 85 | 99 | 115 | | | |
| | .01 | --- | 7 | 13 | 21 | 30 | 40 | 51 | 63 | 76 | 90 | 106 | | | |
| 13 | .05 | 4 | 10 | 18 | 27 | 37 | 48 | 60 | 73 | 88 | 103 | 119 | 137 | | |
| | .01 | --- | 7 | 14 | 22 | 31 | 41 | 53 | 65 | 79 | 93 | 109 | 125 | | |
| 14 | .05 | 4 | 11 | 19 | 28 | 38 | 50 | 63 | 76 | 91 | 106 | 123 | 141 | 160 | |
| | .01 | --- | 7 | 14 | 22 | 32 | 43 | 54 | 67 | 81 | 96 | 112 | 129 | 147 | |
| 15 | .05 | 4 | 11 | 20 | 29 | 40 | 52 | 65 | 79 | 94 | 110 | 127 | 145 | 164 | 185 |
| | .01 | --- | 8 | 15 | 23 | 33 | 44 | 56 | 70 | 84 | 99 | 115 | 133 | 151 | 171 |
| 16 | .05 | 4 | 12 | 21 | 31 | 42 | 54 | 67 | 82 | 97 | 114 | 131 | 150 | 169 | |
| | .01 | --- | 8 | 15 | 24 | 34 | 46 | 58 | 72 | 86 | 102 | 119 | 137 | 155 | |
| 17 | .05 | 5 | 12 | 21 | 32 | 43 | 56 | 70 | 84 | 100 | 117 | 135 | 154 | | |
| | .01 | --- | 8 | 16 | 25 | 36 | 47 | 60 | 74 | 89 | 105 | 122 | 140 | | |
| 18 | .05 | 5 | 13 | 22 | 33 | 45 | 58 | 72 | 87 | 103 | 121 | 139 | | | |
| | .01 | --- | 8 | 16 | 26 | 37 | 49 | 62 | 76 | 92 | 108 | 125 | | | |
| 19 | .05 | 5 | 13 | 23 | 34 | 46 | 60 | 74 | 90 | 107 | 124 | | | | |
| | .01 | 3 | 9 | 17 | 27 | 38 | 50 | 64 | 78 | 94 | 111 | | | | |
| 20 | .05 | 5 | 14 | 24 | 35 | 48 | 62 | 77 | 93 | 110 | | | | | |
| | .01 | 3 | 9 | 18 | 28 | 39 | 52 | 66 | 81 | 97 | | | | | |
| 21 | .05 | 6 | 14 | 25 | 37 | 50 | 64 | 79 | 95 | | | | | | |
| | .01 | 3 | 9 | 18 | 29 | 40 | 53 | 68 | 83 | | | | | | |
| 22 | .05 | 6 | 15 | 26 | 38 | 51 | 66 | 82 | | | | | | | |
| | .01 | 3 | 10 | 19 | 29 | 42 | 55 | 70 | | | | | | | |
| 23 | .05 | 6 | 15 | 27 | 39 | 53 | 68 | | | | | | | | |
| | .01 | 3 | 10 | 19 | 30 | 43 | 57 | | | | | | | | |
| 24 | .05 | 6 | 16 | 28 | 40 | 55 | | | | | | | | | |
| | .01 | 3 | 10 | 20 | 31 | 44 | | | | | | | | | |
| 25 | .05 | 6 | 16 | 28 | 42 | | | | | | | | | | |
| | .01 | 3 | 11 | 20 | 32 | | | | | | | | | | |
| 26 | .05 | 7 | 17 | 29 | | | | | | | | | | | |
| | .01 | 3 | 11 | 21 | | | | | | | | | | | |
| 27 | .05 | 7 | 17 | | | | | | | | | | | | |
| | .01 | 4 | 11 | | | | | | | | | | | | |
| 28 | .05 | 7 | | | | | | | | | | | | | |
| | .01 | 4 | | | | | | | | | | | | | |

1/ Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill, Inc., New York, NY. (Reproduced with permission of the McCraw-Hill Companies.)

**Appendix G** Wilcoxon's signed rank test (tabulated values of T are such that smaller values, regardless of sign, occur by chance with stated probability) 1/

| Pairs n | - - Probability - - .05 | .02 | .01 |
|---|---|---|---|
| 6 | 0 | — | — |
| 7 | 2 | 0 | — |
| 8 | 4 | 2 | 0 |
| 9 | 6 | 3 | 2 |
| 10 | 8 | 5 | 3 |
| 11 | 11 | 7 | 5 |
| 12 | 14 | 10 | 7 |
| 13 | 17 | 13 | 10 |
| 14 | 21 | 16 | 13 |
| 15 | 25 | 20 | 16 |
| 16 | 30 | 24 | 20 |
| 17 | 35 | 28 | 23 |
| 18 | 40 | 33 | 28 |
| 19 | 46 | 38 | 32 |
| 20 | 52 | 43 | 38 |
| 21 | 59 | 49 | 43 |
| 22 | 66 | 56 | 49 |
| 23 | 73 | 62 | 55 |
| 24 | 81 | 69 | 61 |
| 25 | 89 | 77 | 68 |

1/ Steel, R.G.D., and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill, Inc., New York, NY. (Reproduced with permission of the McCraw-Hill Companies.)

**Appendix H** Quantiles (p-values) for Kendall's tau correlation coefficient (p = Prob[S ≥ x] = Prob [S≤ –x]) 1/

| x | Number of data pairs = n 4 | 5 | 8 | 9 | x | Number of data pairs = n 6 | 7 | 10 |
|---|---|---|---|---|---|---|---|---|
| 0 | 0.625 | 0.592 | 0.548 | 0.540 | 1 | 0.500 | 0.500 | 0.500 |
| 2 | 0.375 | 0.408 | 0.452 | 0.460 | 3 | 0.360 | 0.386 | 0.431 |
| 4 | 0.167 | 0.242 | 0.360 | 0.381 | 5 | 0.235 | 0.281 | 0.364 |
| 6 | 0.042 | 0.117 | 0.274 | 0.306 | 7 | 0.136 | 0.191 | 0.300 |
| 8 | | 0.042 | 0.199 | 0.238 | 9 | 0.068 | 0.119 | 0.242 |
| 10 | | 0.0083 | 0.138 | 0.179 | 11 | 0.028 | 0.068 | 0.190 |
| 12 | | | 0.089 | 0.130 | 13 | 0.0083 | 0.035 | 0.146 |
| 14 | | | 0.054 | 0.090 | 15 | 0.0014 | 0.015 | 0.108 |
| 16 | | | 0.031 | 0.060 | 17 | | 0.0054 | 0.078 |
| 18 | | | 0.016 | 0.038 | 19 | | 0.0014 | 0.054 |
| 20 | | | 0.0071 | 0.022 | 21 | | 0.0002 | 0.036 |
| 22 | | | 0.0028 | 0.012 | 23 | | | 0.023 |
| 24 | | | 0.0009 | 0.0063 | 25 | | | 0.014 |
| 26 | | | 0.0002 | 0.0029 | 27 | | | 0.0083 |
| 28 | | | <0.0001 | 0.0012 | 29 | | | 0.0046 |
| 30 | | | | 0.0004 | 31 | | | 0.0023 |
| 32 | | | | 0.0001 | 33 | | | 0.0011 |
| | | | | | 35 | | | 0.0005 |
| | | | | | 37 | | | 0.0002 |

1/ Helsel, D.R., and R.M. Hirsch. 1992. Chapter 12, Trend analysis. In Statistical methods in water resources, Studies in Environmental Science 49, Elsevier, New York, NY.

## Appendix I Conversion Factors

### Length

| From: | To: | Multiply by: |
|---|---|---|
| foot | inch | 12 |
| foot | meter | .3048 |
| inch | centimeter | 2.54 |
| kilometer | mile | 0.621 |
| meter | yard | 1.094 |
| mile | kilometer | 1.6093 |
| yard | inch | 36 |

### Area

| From: | To: | Multiply by: |
|---|---|---|
| acre | $ft^2$ | 43,560 |
| acre | hectare | 0.405 |
| $ft^2$ | $m^2$ | 0.0929 |
| hectare | acre | 2.471 |
| hectare | $m^2$ | $10^4$ |
| $mile^2$ | $kilometer^2$ | 2.59 |

### Volume

| From: | To: | Multiply by: |
|---|---|---|
| $ft^3$ | liter | 28.317 |
| $ft^3$ | gallon | 7.481 |
| gallon | liter | 3.785 |
| $m^3$ | $ft^3$ | 35.314 |
| $m^3$ | liter | 1,000 |

### Discharge

| From: | To: | Multiply by: |
|---|---|---|
| $ft^3/s$ | gpm | 448.83 |
| $ft^3/s$ | $m^3/s$ | .0283 |
| $m^3/s$ | liter/s | 1,000 |
| $m^3/s$ | gpm | 15,850 |

### Mass

| From: | To: | Multiply by: |
|---|---|---|
| pound | kilogram | 0.4536 |
| ton | pound | 2,000 |
| tonnes | pound | 2,205 |
| pound/ac | kg/ha | 1.1208 |
| $ft^3$ - water | pound | 62.4 |

Temperature

$$°F = \frac{9}{5}(°C) + 32$$

$$°C = \frac{5}{9}(°F) - 32$$

### Concentration

| From: | To: | Multiply by: |
|---|---|---|
| mg/L | ppm | 1.0 |
| ppm | ppb | 1,000 |
| mg/L | mg/kg | 1.0 |
| ug/L | $mg/m^3$ | 1.0 |
| $g/m^3$ | mg/L | 1.0 |
| lb/ac | kg/ha | 1.120851 |
| % solution | mg/L | 1 x 104 |

### Metric

To convert SI prefixes

| From: | To: | Multiply by: |
|---|---|---|
| Suffix | mega (M) | $1 \times 10^6$ |
| Suffix | kilo (k) | 1,000 |
| Suffix | hecto (c) | 100 |
| Suffix | deca | 10 |
| Suffix | Suffix | 1 |
| Suffix | deci | .1 |
| Suffix | centi | .01 |
| Suffix | milli | .001 |
| Suffix | micro | .000001 |

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

# Part 615
# National Water Quality Handbook

# Index

(450–VI–NWQH, September 2003)

# Index

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 616

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 617 Examples/Case Studies

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 618

United States
Department of
Agriculture

Natural
Resource
Conservation
Service

# National Water Quality Handbook

# Part 619 Glossary

United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

National Water Quality Handbook

# Part 619 Glossary

# Glossary

**Aerobic** The presence of oxygen.

**Alternate hypothesis** Any hypothesis alternative to the one under a test.

**Anaerobic** The absence of oxygen.

**Analysis of variance** An analysis of the total variation displayed by a set of observations, measured by the sums of squares of deviations from the mean. The variation is usually separated into components associated with sources of interest.

**Aquifer** A geologic formation containing water, usually able to yield appreciable water.

**Baseflow** A part of stream discharge not attributed to direct runoff from precipitation or snowmelt and usually contributed by subsurface flow.

**Baseline** Initial or background water quality conditions. Also a surveyed line.

**Bedload** Sediment, not in suspension, moving along the streambed by rolling or bouncing.

**Benthos** The assemblage of organisms living on or at the bottom of a body of water.

**Best Management Practice** A practice or combination of practices found to be the most effective, practicable (including economic and institutional considerations) means of preventing or reducing the amount of pollution generated by nonpoint sources to a level compatible with water quality goals.

**Blurring** An exploratory data analysis technique of smoothing by replacing data points with short vertical lines of appropriate length beginning with the median of the residuals.

**Calibration** The beginning period of time for a paired watershed design somewhat synonymous with a baseline period.

**Catchment** The area providing runoff to a lake, stream, or well (drainage area, drainage basin, watershed).

**Coefficient of determination** The square of the correlation coefficient. Decimal fraction of percent of variance explained.

**Coefficient of variation** The standard deviation of a distribution divided by the mean.

**Coliform bacteria** A group of bacteria predominantly found in the intestines of animals, but also occasionally found elsewhere.

**Composite sample** A combination of individual samples taken at selected intervals or volumes to minimize variability.

**Concentration** The amount of a substance dissolved or suspended in a unit volume of water.

**Conductance** The measure of the conducting ability of a solution that is equal to the reciprocal of the resistance.

**Confidence level** The measure of probability ($\alpha$) of the truth of a statement.

**Confidence limits** The values of an upper and lower *t* of a confidence interval. The interval has a probability ($\alpha$) that the value will lie between the upper and lower limits.

**Confined aquifer** An aquifer that is surrounded by formations of less permeable or impermeable material that is isolated from the atmosphere. (Artesian aquifer)

**Conservation practice** A specific treatment, such as a structural or vegetative measure, or management technique, commonly used to meet specific needs in planning and implementating conservation for which standards and specifications have been developed. Conservation practices are listed in the Field Office Technical Guide (FOTG), section IV, which is based on the National Handbook of Conservation Practices (NHDP).

**Contamination** An introduction of a substance into water in a sufficient concentration to make the water unfit for its intended use.

**Continuous data** Data for which all values in some range are possible, such as height and weight.

**Control** In a study, a standard for comparison against which other treatments are compared, but is either untreated or receives a standard treatment. Also, a stable cross section in a stream that controls flow upstream.

**Critical area** An area within a watershed determined to be an important source of a pollutant.

**Current meter** A devise for measuring the velocity of flowing water.

**Discharge** The rate or volume of water flowing at a specific cross section within a specified time.

**Discharge rating curve** A curve showing the relationship between the stage at a cross section and the discharge at that cross section.

**Discrete data** Data for which the possible values are fixed, such as counts.

**Dispersion** The mixing of the concentration of a substance in the water with another body of water due to the flow of water.

**Dissolved oxygen** The oxygen dissolved in water, expressed in milligrams per liter or percentage saturation.

**Drainage basin** See catchment.

**Drainage density** The density of natural drainage channels in a given area, expressed as length per unit area.

| | |
|---|---|
| **Effluent stream** | A stream that receives water from saturated ground water. |
| **Epilimnion** | The upper waters of a thermally stratified lake. |
| **Equipotential line** | A contour line that connects points of equal head for the water table or equipotential surface. |
| **Error** | The difference between an occurring value and its true or expected value. |
| **Eye smoothing** | Drawing s smooth curve through points of data on a graph. |
| **Field** | A small agricultural unit implying a management area. |
| **Filter strip** | A conservation practice that is a strip of vegetated land established downslope of a nonpoint source of pollution with the purpose of reducing the pollutant. |
| **Flow line** | A line indicating the direction of ground water flow toward the point of discharge. Flow lines are perpendicular to equipotential lines and together they form a flow net. |
| **Flume** | An open conduit for flow. |
| **Frequency distribution** | A listing of the way the frequencies of members of a population are distributed according to the values of the variable. The distribution is usually shown in a table. |
| **Gage** | A device for determining the water level. |
| **Grab sample** | A single sample taken at a certain time and place. |
| **Ground water** | Subsurface water in the saturated zone below the water table. |
| **Hydrograph** | A graph showing discharge as a function of time for a given location on a stream. |
| **Hypolimnion** | The bottom waters of a thermally stratified lake. |
| **Hypothesis** | A hypothesis concerning the parameters or form of the probability distribution for a designated population. |
| **Intermittent stream** | A stream or portion that flows only in direct response to precipitation. |
| **Interval scale** | A measurement with a constant interval size, but no true zero, such as temperature (arbitrary zero) and time. |
| **Kurtosis** | The extent to which a unimodal frequency curve is peaked. |
| **Least squares regression** | Estimation of regression parameters by minimizing a quadratic form. |

| | |
|---|---|
| **Limnocorral** | A device used in lakes that isolates the water column from surrounding water. |
| **Load** | The quantity of material entering a receiving body of water. |
| **Lysimeter** | A device used to measure the water quantity or quality draining through the soil. |
| **Macroinvertebrate** | A large animal without a backbone that can be observed without the aid of magnification. |
| **Macrophyton** | A large plant that can be observed without the aid of magnification. |
| **Mean** | The arithmetic average of the values for a variate. |
| **Median** | That value of the variate which divides the total frequency into two halves. |
| **Mesocosm** | A medium-sized experimental unit with boundaries. |
| **Metalimnion** | The middle layer of a thermally stratified lake. |
| **Mode** | The value of the variate that has the greatest number of members of the population. |
| **Model** | A description of a system; often mathematical. |
| **Nonparametric statistics** | Better termed distribution-free statistics. Testing a hypothesis that does not depend on the form of the underlying distribution. |
| **Nonpoint source** | A diffuse location with no particular point of origin. |
| **Null hypothesis** | A hypothesis under test that determines the probability of the Type I error. Also a hypothesis under a test of no difference. |
| **Objective** | A statement describing what is to be accomplished that contains an infinitive verb and an object. |
| **Observation** | Data that are collected or analyzed. |
| **Ordinal scale** | Data that consist of an ordering or ranking of measurements, such as A is bigger than B. |
| **Parametric statistics** | A statistical test that assumes the distribution type is known. |
| **Perennial stream** | A stream that flows continuously all seasons of a year and during both wet and dry years. |
| **Periphyton** | Small or microscopic aquatic plants attached to submerged objects. |
| **Phytoplankton** | Small or microscopic aquatic plants. |

| | |
|---|---|
| **Piezometer** | An instrument for measuring pressure head in the soil. |
| **Plankton** | Small or microscopic aquatic organisms that are floating, or weakly motile and generally considered to be at the mercy of the currents. |
| **Plot** | A small experimental unit with boundaries. |
| **Pollutant** | An undesirable substance in water, soil, or air at sufficient concentrations to impair the intended use of the resource. |
| **Pollution** | A condition caused by the presence of harmful or objectionable substances in water. |
| **Population** | A collection of individuals. |
| **Random sample** | A sample collected from a population where every sample has an equal probability of being selected. |
| **Rating** | A relation between stage and discharge of a stream. |
| **Ratio scale** | Measurements having a constant interval size and a true zero point, such as lengths, weights, volumes, and rates. |
| **Reconnaissance survey** | A survey to obtain a general view of water quality; may imply samples collected at approximately the same time (synoptic survey). |
| **Regression** | A statistical method to investigate relationships between two components. |
| **Replication** | The execution of an experiment more than once. |
| **Resource management system** | A combination of conservation practices and resource management, for the treatment of all identified resource concerns for soil, water, air, plants, and animals, that meets or exceeds the quality criteria in the Field Office Technical Guide (FOTG) for resource sustainability. |
| **Responsiveness** | In establishing cause-and-effect, the evidence that the dependent variable is related to the independent variable. |
| **Runoff** | That portion of precipitation or irrigation found in surface channels and streams. |
| **Runoff coefficient** | The ratio of the depth of runoff from a watershed to the depth of precipitation. |
| **Sample** | A part of all the possible measurements in some larger group, such as the population. |
| **Sampler** | A device used to obtain an aliquot of water. |
| **Significance** | The probability of committing a Type I error ($\alpha$). Biological significance refers to an underlying assumption about relationships. |

| | |
|---|---|
| **Skewness** | A measure of asymmetry in a frequency distribution. |
| **Smoothing** | The process of removing fluctuations in a series of data. |
| **Specific conductance** | The ability of water to conduct electricity across a specific length at a specified temperature. |
| **Stage** | The elevation of the water surface above some datum. |
| **Stage-discharge relation** | The relationship between stream stage and discharge at a gaging station. |
| **Standard deviation** | A measure of dispersion of a frequency distribution that is the square root of the variance. |
| **Statistic** | A summary value calculated from a sample of observations. |
| **Statistical error** | See Error. |
| **Statistics** | The science of collecting, analyzing, and interpreting data. |
| **Steady-state** | Conditions that are averaging constant over time. |
| **Stilling well** | A chamber with small inlets connected to a water body used for measuring the water level. |
| **Streamflow** | Water flowing in a stream channel. (Stream discharge) |
| **Surface runoff** | The portion of runoff that reaches a stream by traveling over the surface of the land. (Overland flow) |
| **Suspended solids** | Solids in suspension in water. |
| **Synoptic survey** | See reconnaissance survey. |
| **Tensiometer** | An instrument filled with water with a porous cup used for measuring the soil water potential. |
| **Turbidity** | A condition in water caused by suspended matter that causes the scattering and absorption of light. |
| **Unconfined aquifer** | An aquifer where the water table is exposed to the atmosphere. (Water table aquifer) |
| **Vadose zone** | Zone of soil between the surface and the water table that is not saturated. |
| **Variance** | The mean of the squares of the deviations from the mean. |
| **Velocity meter** | A meter used to measure stream velocity. |
| **Water quality** | The physical, chemical, and biological properties of water with respect to its suitability for an intended use. |

| | |
|---|---|
| **Water quality management** | The management of the physical, chemical, and biological characteristics of water. |
| **Water quality monitoring** | The collection of information on the characteristics of water. |
| **Water quality standards** | A rule established by an agency or units of government; often numerical. |
| **Water table** | The upper surface of the saturated zone in a soil that is at atmospheric pressure. |
| **Water-level recorder** | A device used for recording the water elevation over time. |
| **Watershed** | The area contributing water to a stream, lake, or well. |
| **Weir** | A device used in a stream with a damming crest and an opening of some known geometric shape, such as a V-notch. |
| **Zooplankton** | Small or microscopic aquatic animals. |