A REVIEW OF THE WATER QUALITY MONITORING PROGRAMS IN THE NATIONAL PARKS IN CENTRAL COASTAL CALIFORNIA

Report to the National Park Service

2 April 2004

Report prepared by

Sarah Stafford & Professor Alex Horne
Ecological Engineering Group
Department of Civil & Environmental Engineering
University of California, Berkeley

1.0 SUMMARY

2.0 TABLE OF CONTENTS

1.0 summary *2*

2.0 table of contents *2*

2.1 Table of Figures **3**

2.2 Table of Tables **4**

3.0 Introduction *5*

3.1 Overview of parks monitoring systems in San Francisco Bay Area Network (SFBAN) **5**

3.2 Organization of the report **7**

3.3 Acknowledgements and authorship **7**

4.0 description of water quality monitoring parameters *7*

4.1 Physical and basic parameters **7**

4.1.1 Water discharges (flow) and storm flows 8
4.1.2 Temperature 13
4.1.3 Acidity, pH 17
4.1.4 Conductivity, Salinity and Total dissolved solids (TDS) 21
4.1.5 Total Suspended Solids (TSS) and Turbidity 27

4.2 Oxygen related parameters **33**

4.2.1 Dissolved oxygen (DO) 33
4.2.2 Biochemical oxygen demand (BOD), chemical oxygen demand (COD) and total organic carbon (TOC) 39

4.3 Nutrients **41**

4.3.1 Nitrogen, nitrate and ammonia 41
4.3.2 Phosphorous, phosphate, total phosphorous (TP) and orthophosphate (ORP) 47

4.4 Biological indicators **52**

4.4.1 Bacterial indicators, Fecal coliforms (FC), total coliforms (TC) and *E. coli* 52

5.0 Quality Assurance and Quality Control *56*

5.1 Field Parameters **56**

6.0 Discussion of water quality analysis *57*

6.1 The water quality information system (modified from ward et al., 1990) **57**

6.2 Establish a data analysis protocol (DAP) **59**

6.3 How to prepare the raw data record for analysis **60**

6.3.1 Censored and missing data 60
6.3.2 Replicates 60

6.4 Establishing Reference Conditions **60**

6.5 Statistical issues in the analysis of water quality data **61**

6.5.1 Distributions: Visualizing the data 61
6.5.2 Power, Efficiency, and sample size 62
6.5.3 Parametric vs. Nonparametric tests (Hirsch et al., 1991) 63

6.5.4 Hypothesis Testing (Hirsch et al., 1991) 63
6.5.5 Serial Correlation and sampling frequency (modified from (Kirchner, 2003) 65
6.5.6 Transformations and other data manipulations to reduce variance (Hirsch et al., 1991) 66
6.5.7 Concentration vs. Flux (Hirsch et al., 1991) 69
6.5.8 Trend Detection (modified from (Hirsch et al., 1991)) 69

7.0 Monitoring Critique 71

8.0 Summary of Recommendations 72

9.0 Bibliography ..... *Error! Bookmark not defined.*

Appendix A: Data distributions ..... *Error! Bookmark not defined.*

Appendix B: Time plots of Parks Data ..... *Error! Bookmark not defined.*

Appendix C: Sample site location tables ..... *Error! Bookmark not defined.*

Appendix d: Notes on data preparation and analysis ..... *Error! Bookmark not defined.*

2.1 TABLE OF FIGURES

Figure 1. Log flow (cfs) by month for GGNRA streams 10
Figure 2. Log flow (cfs) by month for PINN streams 11
Figure 3. Log flow (cfs) by month for PRNS streams 12
Figure 4. Distribution of Temperature (°C) in GGNRA Streams 16
Figure 5. Distribution of pH in GGNRA Streams 19
Figure 6. pH in GGNRA Streams by Month 20
Figure 7. Distribution of Alkalinity in GGNRA Streams 21
Figure 8. Distribution of Log Conductivity in GGNRA Streams 24
Figure 9. Log conductivity in GGNRA over time. 24
Figure 10. Specific Conductance in Stables-based Monitoring in GGNRA and PRNS 25
Figure 11. Distribution of log transformed salinity at GGNRA. 25
Figure 12. Salinity in GGNRA 26
Figure 13. Distribution of Log Transformed TSS in GGNRA Streams 29
Figure 14. Log TSS over time in GGNRA streams. 29
Figure 15. TSS in GGNRA streams as a function of month. 30
Figure 16. Log TSS as a function of Log Flow at LAG1 Nov, 1989- March, 1990. 31
Figure 17. Log TSS as a function of Log Flow at RDW4 32
Figure 18. Distribution of Log Turbidity (NTUs) in GGNRA Streams 33
Figure 19. Dissolved Oxygen as a function of Temperature 35
Figure 20. Distribution of DO in GGNRA Streams 36
Figure 21. Dissolved Oxygen by Month in GGNRA Streams 1978-1999. 37
Figure 22. DO over time in PRNS, GGNRA and PINN 38
Figure 23. BOD in Stables monitoring at PRNS and GGNRA 41
Figure 24. Distribution of log transformed ammonia in GGNRA streams 1993-2002. 45
Figure 25. Ammonia over time at GGNRA, PRNS and PINN 45
Figure 26. Distribution of Log Nitrate in GGNRA Streams 46
Figure 27. Nitrate at PNRS and GGNRA 1986 to 2002 46
Figure 28. Distribution of Log Transformed Total Phosphorous in GGNRA Streams. 49
Figure 29. Distribution of Log Transformed Orthophosphate in GGNRA. 49
Figure 30. Log ORP in GGNRA Streams as a Function of Month. 50
Figure 31. Total Phosphorous over Time in GGNRA 51
Figure 32. Orthophosphorous over time in GGNRA and PNRS. 51
Figure 33. Comparison of ORP and TP from same samples in Stables 1999-2001 data set. 52
Figure 34. Distribution of Log Transformed Fecal Coliforms in GGNRA 55
Figure 35. Fecal Coliforms in PRNS, GGNRA and PINN 55
Figure 36. Water Quality Information Flow (Ward et al., 1990) 1

Figure 37. Distribution of dissolved oxygen in GGNRA streams *62*
Figure 38. Log Transformed and Untransformed Flow by Month at GGNRA *67*

2.2 TABLE OF TABLES

Table 1. Monotonic trends in uncensored data..... ***Error! Bookmark not defined.***
Table 2. Step trends in uncensored water quality data *70*
Table 3. Monotonic trends in censored water quality data *70*
Table 4. Step trends in censored data..... *71*

3.0 INTRODUCTION

3.1 OVERVIEW OF PARKS MONITORING SYSTEMS IN SAN FRANCISCO BAY AREA NETWORK (SFBAN)

Point Reyes National Seashore (PRNS)

Starting in 1999, PRNS began surface water quality monitoring at 23 stream locations and three recreational ponds, as well as aquatic bioassessment at 12 locations. This monitoring effort is focused on identifying sources of pollution. The sites range from wilderness areas to directly downstream of dairy impact areas. Monitored water quality parameters include nitrate, nitrite, ammonia, conductivity, TSS, dissolved oxygen and pH. The aquatic bioassessment includes macroinvertebrate samples collected from Olema Creek and three coastal watersheds draining to Drakes Estero for a long-term biological tend analysis. Data from this water quality monitoring effort are stored in a Microsoft Access database. One report covering May 1999 to May 2001, "Point Reyes National Seashore Water Quality Monitoring Report, May 1999 – May 2001" has been published (Ketcham, 2001). This report identifies fecal coliform and unionized ammonia as primary indicators of water quality degradation, and conductivity and Total Suspended Solids (TSS) as secondary indicators. Much of the water quality monitoring effort has been focused on the Olema Creek watershed, which supports four federally threatened aquatic species and drains to Tomales Bay, itself listed as impaired by the RWQCB for pathogens, sediments, nutrients and mercury. The Olema Creek watershed is the subject of a detailed analysis, including the watershed hydrologic response and fecal coliform time-series in one of the appendices of the 1999-2001 report.

Golden Gate National Recreation Area (GGNRA)

Starting in WY1998, GGNRA began a winter water quality monitoring program targeting stable operations within the park. Monitored water quality parameters include fecal coliforms, ammonia, total phosphorous, orthophosphate, nitrate, specific conductance, total suspended solids, pH, dissolved oxygen, temperature, biological oxygen demand, copper, MBAS and flow. The main management issues facing both GGNRA and PRNS have to do with balancing the historical and cultural traditions of ranching and dairy establishments with the very high water quality needed for endangered species such as Coho salmon, steelhead trout, California freshwater shrimp and California red-legged frogs.

Water quality scoping meetings have identified accelerated erosion due to roads, trails, and other uses and developments as threats to the sediment balance and ecological health of several watersheds. A sediment budget study was conducted in the Redwood Creek watershed as well as a report of all sediment sources in the watershed. Grazing is no longer allowed on GOGA managed lands (GGNRA, 1999) but some of the impacts remain. Bacteria and nutrient inputs from equestrian operations, pet waste, agricultural operations and, potentially, sewer and septic systems can impact wildlife and public health as well as the overall ecological balance of water resources. Channel alteration (including dams and culverts) impacts the ecological health of park watersheds. These primary issues occur to varying extents within multiple park watersheds (Cooprider, 2004).

Pinnacles National Monument (PINN)

Monthly monitoring of creeks at approximately 40 different sites was initiated in 1997 at Pinnacles National Monument. Parameters measured include flow, dissolved oxygen, pH, salinity, conductivity, temperature, TDS (total dissolved solids), TSS (total suspended solids), nitrates, ammonia, TKN (total Kjeldahl nitrogen), organic nitrogen, total phosphorous, orthophosphorous, TOC (total organic carbon), general minerals, and a few metals (gold, copper, lead, zinc, antimony, nickel and barium). The samples have typically been collected during first flush (Dec/Jan.), peak flow (Feb/March), and baseflow (late May). Monitoring for *Escherichia coli* and fecal coliforms was initiated in 2001.

Monitoring is focused on the Chalone Creek Watershed that includes 95% of the parklands. The headwaters for Chalone Creek are, however, located outside of parklands. Monitoring has additionally been focused to some extent on the areas around and downstream of a landfill that is located within the Chalone Creek floodplain (Cooprider, 2004). Benthic macroinvertebrates have been sampled in the South Wilderness area downstream of this landfill site. Nitrates and phosphate analyses as well as GMA (general mineral analyses) were also conducted in the creek downstream of the landfill site. Peizometers were installed at the landfill site; only water level has been monitored to date.

The PINN water quality scoping meeting identified elevated levels of sediment, bacteria, and nutrients in surface waters as current issues. Due primarily to past land uses, threats of heavy metal contamination are also a concern. Some of these concerns are not well documented; therefore, one goal of a long-term monitoring plan is to clearly identify threats to water quality in order to better understand the extent of contamination and address it (Cooprider, 2004).

Elevated levels of nitrates and phosphates have been detected, possibly due to grazing which occurs in the riparian zone on land upstream of parklands. Annual rainfall for Pinnacles averages 16 inches per year, most of which falls in a few storms. Therefore, many streams are ephemeral and flashy showing just a trickle of flow in most years except for during 3 or 4 flood events. Some years some streams do not flow at all. Stream invertebrates in ephemeral streams are depauperate compared with permanent ones. The stream hydrology thus creates a very different set of management and monitoring challenges than are present at the GGNRA and PRNS. Therefore, even though all three parks are grouped together in the same network, it will be important to acknowledge that different monitoring strategies may be required. Data has been stored in the form of field reports and was only put into digital format in 2002. No database has yet been assembled, and no reports have been written.

Historical Data Sets

Sites within the GGNRA and PRNS have been monitored for various parameters since the 1950s, resulting in a huge data set of varying quality. Many of these sampling events targeted specific water quality areas, specifically pastoral and horse stable operations, and were not intended to document basic water quality conditions. Most available data up until 1999 has been entered into an Access database. However, with a

few exceptions, information regarding QA/QC, and intent of the data collection is not easily accessible. Several reports have been written including:

i. "1996 Fall Fish Kill Evaluation for Rodeo Lagoon, Golden Gate National Recreation Area, Marin Co."
ii. "Winter 1997-1998 Water Quality Monitoring at Golden Gate Dairy Tributary"
iii. "Golden Gate National Recreation Area Storm Water Monitoring Program 19997/1998"
iv. "Winter 1999-2000 Water Quality Monitoring at Golden Gate National Recreation Area Stables"
v. "Winter 2000-2001 Water Quality Monitoring at Golden Gate National Recreation Area Stables"
vi. "Point Reyes National Seashore Water Quality Monitoring Report, May 1999 – May 2001."

Information from the above reports is included in the following discussion of monitoring in the National Parks in the SFBAN.

3.2 ORGANIZATION OF THE REPORT

The report is divided into four main sections; water quality description, water quality analysis, a monitoring critique and recommendations. They can be summarized thus:

- *Water quality description.* Basic water quality monitoring parameters, including: what each one is, where it comes from and where it goes, why it is important, how it is measured, how it relates to other parameters and water quality as a whole, what typical levels and criteria limits are, if it has been measured at the parks and if so what it tells us.
- *Water quality analysis*. Data analysis, trend detection and hypothesis testing in the context of the historical data set and what can be improved on for future monitoring.
- *Monitoring Critique*. Evaluation of what is working and what isn't working in the current water quality monitoring network
- *Recommendations*. Key recommendations for future integrated monitoring

3.3 ACKNOWLEDGEMENTS AND AUTHORSHIP

This report was prepared as part of a grant from the GGNRA. It was researched and written primarily by doctoral student Sarah Stafford under the guidance of Professor Alex Horne. We wish to thank Mary Cooprider, Brannon Ketcham and Darren Fong for their assistance with this work.

4.0 DESCRIPTION OF WATER QUALITY MONITORING PARAMETERS

4.1 PHYSICAL AND BASIC PARAMETERS

4.1.1 Water discharges (flow) and storm flows

- *What is flow?* Flow or discharge is the volume of water per unit time flowing through a stream or river.
- *Why is flow important?* Contaminant concentrations are often correlated with flow, and in order to detect long-term trends it is useful to correct for the flow. The power and efficiency of statistical analysis will be enhanced if the variance of the data can be decreased. The best procedure is to analyze flow-adjusted concentrations (FAC). In addition, flow is needed to calculate the mass loadings of any contaminants to downstream waters and is needed in the Total Maximum Daily Load (TMDL) process (Rosenlieb et al., 2002).

The Freshwater Workgroup Subcommittee recommends that some measure of discharge for flowing water (streams and rivers), and waterbody level or stage for still waterbodies (lakes, reservoirs, etc.) be included as a fifth core water quality monitoring parameter in the NPS Vital Signs Program (Rosenlieb et al., 2002).

- *What influences changes in flow?* Changes in precipitation, groundwater levels, water use by riparian vegetation, and reservoir management all contribute to variability in stream flow.
- *How is flow related to other parameters and overall water quality?* In many cases, a great deal of the variability in water quality is a function of river discharge. The variability is due to two main factors, dilution and wash-off ((Hirsch et al., 1991). In the case of dilution, a solute such as ammonia or nitrate may be delivered to the stream at a relatively constant rate, but as flow increases, the concentration goes down, or as the flow decreases, the concentration goes up. The flux, or mass loading, however remains the same. This is typically seen in most of the major dissolved ions. In the case of wash-off, a constituent (for example a pesticide or some heavy metals) attached to a particle can be delivered to a stream primarily from overland flow in small temporary rivulets. In this case, both the concentration and the flux tend to rise as the flow increases. Similar to wash-off is the case where some areas of the watershed are not connected to the stream during the dry season, and when enough rain falls to restore connectivity to the system, the concentration of dissolved ions from nonpoint sources increases along with the increase in flow. Some constituents, such as phosphorous, can exhibit combinations of both types of behavior, since they may be coming both from fixed point sources as well as from overland flow. For a more comprehensive discussion of the relationship between discharge and concentration of dissolved solids, see section 5 of the *Study and Interpretation ofChemicalCharacteristics of Natural Water* (Hem, 1985).

In terms of how the GGNRA staff would make use of flow, flow-weighted and total load values, the two concepts of pollution control are important. These two are (1) toxicity-assessment related to a certain exposure time to an average and a maximum concentration and (2) the mass loading to a lentic or still water body such as a lake, estuary or the coastal ocean. In the case of such still waters the total load to the water body, usually expressed in grams of pollutant per square meter of lake or estuarine surface is often used to estimate eutrophication and associated nuisance algae and fish kill potential. Nitrate, phosphate and ammonia are the three main nutrients supporting algae growth and ammonia is both a stimulant and toxicant depending on load, concentration and some associated variables such as pH.

The third important reason to measure flow is that it controls the sediment movement in streams. In our attempts to clean or restore streams, it is often forgotten that, unlike some other aquatic habitats such as lakes, sediment transport is vital for stream health. Too much flow and erosion will destroy the stream and its riparian zone. Too little flow and silt will accumulate upstream while the downstream areas will be starved of sediment needed for insect habitat and fish spawning grounds. It is quite startling to see how much sediment including small silts and sand but also pebbles and quite large rocks are moved during the few hours of high flow each year. Thus, the suitable gravels for aquatic life at any point in a stream must be continually renewed from upstream. The only source of this life-giving sediment and gravel is from erosion of the upstream stream banks. Large infrequent storms such as the 20-year flood erode the most sediment but smaller annual storms are more frequent. Professor Geofory Mount at the University of California, Davis considers that in the semi-arid west both kinds of storms are equally important in the natural and continual evolution of our California streams (Mount, 1995).

- *How is flow (discharge) measured?*

Flow, strictly discharge (volume per time) is measured by taking the velocity of the current and multiplying it by the depth and width of the stream. Current velocity can be measured in various ways including using a bucket and stopwatch for small flows that pass through culverts. However, a submerged flowmeter with rotating cups or an electronic signal are most often used in small streams and rivers. To get an accurate measurement, the site should have parallel flow lines, a relatively uniform streambed, and uniform flow. The USGS midsection method (Buchanan and Somers 1969) provides a consistent method. Discharge is usually reported in cubic feet per second (cfs), but should also be reported in metric units, usually cubic meters per second (m^3/s).

- *What are typical flow (discharge) levels and criteria?* Typical flow levels vary widely depending on season and the specific waterbody. Water currents in small streams can exceed 200 cm/sec over waterfalls but values of 10-50 cm/s are more typical in flatter regions. Flow in pools can be very low and even circular so most flow measurements are made in non-pool areas. Since the absolute amount of flow (discharge) varies with the individual stream, some authors have related the water level to the stream size. For example, in some kinds of streams with regular annual flow patterns, bankful flow is a standard way to represent stream or river discharge. If enough data is available, an estimate of the bankful flow should be included in the table of site location descriptions. In flashy western streams where discharge varies considerably from year to year, bankful values are of less value. In some cases, flow ceases entirely during the dry season, resulting in an intermittent or ephemeral stream. Ephemeral streams are usually indicated on maps as a dotted blue line and, if accurate for the stream in current times, should also be noted in the site locations table.

- *What is the history and current state of flow measurement in the SFBAN?* Flow data has been collected for streams in the GGNRA since at least 1990 and at PRNS since at least 1997. Flow data has been collected during most sampling events, however the flow data has not always been placed in the same spreadsheet with the other data. At PINN, flow data is available from 1997 to present. In the stables-based monitoring for PRNS and GGNRA, flow was measured at the downstream-most site using a Swoffer 2100 flow meter. As can be seen in figures 1-3, flow varies widely with seasonal rainfall at all three

parks. It does not make sense to compare flow in one stream to flow in another stream or look at the average flow across the parks, since flow by itself is not reflective of impairment. It is more important to compare the flows of each stream over time (taking into account the season and rainfall) or to its bankful flow.

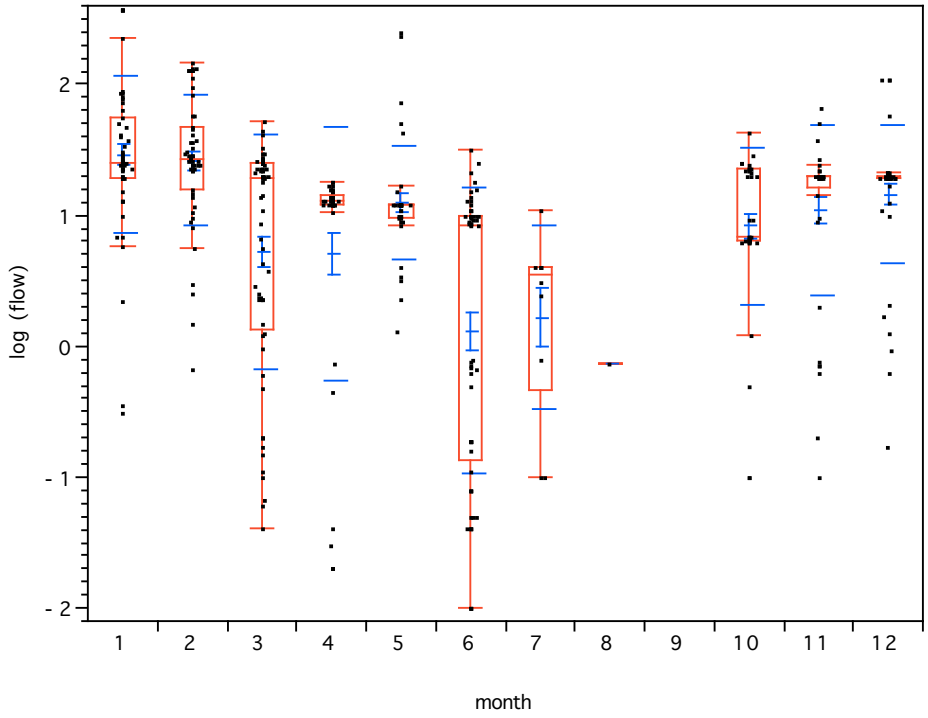


Figure 1. Log flow (cfs) by month for GGNRA streams

This is a log transform of flow by month for all the GGNRA data in the Access database. The medians and quantiles for each month are indicated by the red boxes and whiskers. The line inside of each box represents the median. The ends of the boxes are the 25th and 75th quantiles, so the box itself encompasses the interquartile range. The whiskers extending beyond the box represent the upper and lower quartiles + 1.5*(interquartile range).

The monthly means and standard deviations are indicated by the blue lines. The center blue line is the mean. The two inner blue lines are the confidence interval for the mean and the outer two blue lines are one standard deviation.

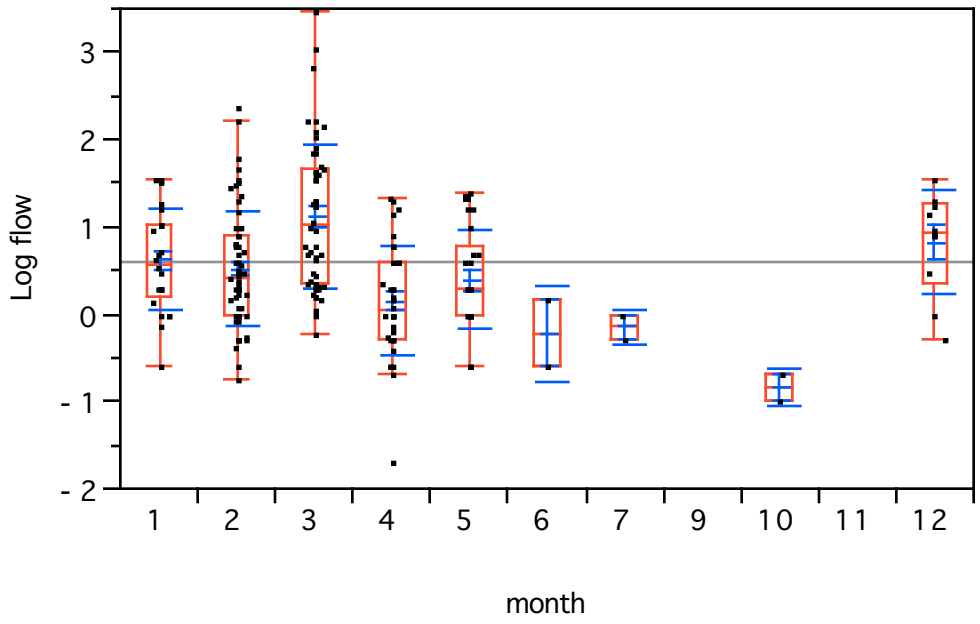


Figure 2. Log flow (cfs) by month for PINN streams.

The medians and quantiles for each month are indicated by the red boxes and whiskers. The red line inside of each box represents the median. The ends of the boxes are the 25th and 75th quantiles, so the box itself encompasses the interquartile range. The whiskers extending beyond the box represent the upper and lower quartiles + 1.5*(interquartile range). The monthly means and standard deviations are indicated by the blue lines. The center blue line is the mean. The two inner blue lines are the confidence interval for the mean and the outer two blue lines are one standard deviation.

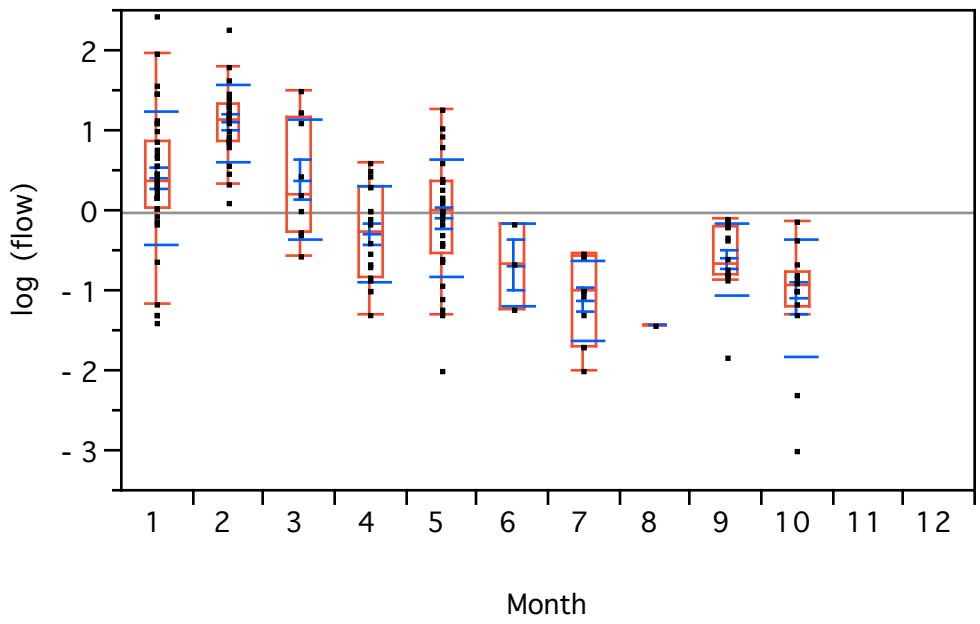


Figure 3. Log flow (cfs) by month for PRNS streams

The medians and quantiles for each month are indicated by the red boxes and whiskers. The red line inside of each box represents the median. The ends of the boxes are the 25th and 75th quantiles, so the box itself encompasses the interquartile range. The whiskers extending beyond the box represent the upper and lower quartiles + 1.5*(interquartile range). The monthly means and standard deviations are indicated by the blue lines. The center blue line is the mean. The two inner blue lines are the confidence interval for the mean and the outer two blue lines are one standard deviation.

A "Detailed Analysis of Olema Creek Watershed Response" is included as an appendix in the "Point Reyes National Seashore Water Quality Monitoring Report, 2000-2001". Olema Creek has the most complete set of discharge and rainfall information within the PRNS. The report concludes that the hydrologic response in this watershed is rapid and short-lived. The flow of Olema Creek ranges from 0.5 cfs in the summer to more than 2500 cfs in the winter. The bankful discharge is calculated at approximately 500 cfs. The rainfall-runoff response time in most of the tributaries is 15 to 45 minutes, depending on watershed size. The report further includes information on historical land-use in the watershed and channel conditions, all of which provide important context for interpreting water quality monitoring data.

- *What are recommendations regarding flow monitoring?* Discharge or flow of water should be measured at every sampling event. However, accurate discharge measurements are time-consuming. The usual way to overcome the problem is to construct a series of "stage curves." Discharge is measured at first in the normal way using a flow meter and the calculated flow related to a "staff gauge" a fixed stream depth rod located in a convenient place for observation on the stream bank. The rod usually is inscribed with large easy to read divisions in cm or inches. Most people are used to flood warnings based on inches or feet of river height rather than discharge. The exercise is repeated for several different water levels and a relationship between staff gauge height

and stream discharge is established. So long as major changes do not occur in the shape of the streambed, water discharge can easily be determined from a simple recording of the water elevation on the staff gauge. It is reasonably simple to automate the stream height level using a stilling well which is a metal cylinder that protects the automatic reading device from large stream debris but is still hydraulically connected to the creek. Stilling wells and staff gauges are best located near a hard bottom that will not change much over time. Upstream of bridges are favored spots.

Discharge and the station that it was recorded at should be included in the database entry for each observation. For ease of comparison, it would also be useful to include a column for the bankful flow and whether the stream is intermittent or not in the site locations table. It is also important to record if the measurements are representative of base flow conditions, first flush, winter runoff or other significant hydrological conditions. A column for information on rainfall and first flush events should be included. Ideally, storm samples aimed at capturing the worst conditions should be collected on the rising limb of the hydrograph, as pollutants are flushed from the land surface (Ketcham, 2001). A separate column should also be included for storm events, to record if measurements were made during the rising limb, the peak or the falling limb of the flood.

Measurements of both contaminants and flow should ideally be made for several storms of different sizes, monitoring both flow and other parameters over the entire storm. For many pollutants, the main load comes in the first 25% of the discharge (usually the rising arm of the hydrograph) (Hem, 1985). The later water is relatively clean but this pattern is not always observed. In stream management, the first flush of pollutants can be contained in an economical manner, for example by diversion into constructed treatment/wildlife wetlands or excavatable detention basins and the cleaner water moved on if the storm surveys have been carried out.

The "Detailed Analysis of Olema Creek Watershed Response" provides an excellent example of the kind of hydrologic background information that could be prepared for each of the main watersheds in the parks.

4.1.2 Temperature

- *What is temperature?* Temperature is the specific degree of hot or cold as measured by a thermometer or other standard scale. It is a measure of the kinetic energy of the water.
- *Why is temperature important?* Temperature is important because living organism have a limited range of temperatures in which they can survive. In coastal California, common environmental changes such as the introduction of cattle causes loss of riparian zones that shade streams. The higher temperatures can kill fish. Temperature also affects water density and gas solubility, including how much oxygen can be dissolved in the water. At higher temperatures, water is able to hold less dissolved oxygen, increasing the stress on fish. In addition, pH, conductivity and the rates of biological and chemical reactions vary depending on the temperature (Hem, 1985; Shelton, 1994).

Temperature is one of the core required parameters recommended by the Freshwater Workgroup Subcommittee for the NPS Vital Signs Program (Rosenlieb et al., 2002).

- *What influences changes in temperature?* The temperature of the water is affected by several factors including the air temperature, humidity, percent shading, the turbidity or cloudiness of the water, as well as the temperature of groundwater and storm water inflows (Essig, 1998; Theurer et al., 1984). The flow is also important, since it takes longer for a larger volume of water to heat up, and water temperature is very sensitive to changes in air temperature when stream flow is low (Essig, 1998). In coastal California, the most important factor in small streams and rivers is the degree of shading provided by the trees and bushes of the riparian zone.
- *How is temperature related to other parameters and overall water quality?* High temperatures can increase the toxicity to fish of high ammonia levels and low DO levels (Horne and Goldman, 1994). High temperatures in streams increase the overall rate of metabolism for all the biota (Hem, 1985). Thus, oxygen is more rapidly consumed and can be reduced to critical levels at night, when plant and algal photosynthesis stops. Increased temperatures also favor kinds of algae, bacteria and invertebrates that are not good food for most fish, while decreasing those that are good fish food.
- *How is temperature measured and in what units is it reported?* Temperature can be measured by one of the handheld YSI meters (or similar from other manufactures) and should be reported in degrees Celsius. In smaller streams and rivers, temperature can vary spatially and temporally (Bilby, 1984; Essig, 1998). These changes can be best measured by inexpensive data loggers that can be set to record temperature every 15 minutes for 3 months without needing attention.

The USGS defines three temperatures of concern when monitoring continuous stream temperature: true stream temperature (TST), temperature near sensor (TNS), and temperature recorded (TRC) (Stevens et al., 1975). Ideally, all three of these temperatures would be the same, however this is not always the case. The true stream temperature is defined as an instantaneous measurement obtained in a shaded location in the main flow of the stream outside of the influence of tributaries or groundwater influx with a full immersion thermometer calibrated against an ASTM standard thermometer. It can also be calculated as the weighted average of a cross-section temperature profile. For reasons of safety and convenience, sensors for stream temperature recorders are often placed closer to shore than would represent true stream temperature (Essig, 1998). The actual temperature of the water surrounding the sensor reflects its position in the channel cross-section, and is known as the temperature near sensor (TNS). Whether it is the same as the TST may vary depending on flow and time of day (Essig, 1998). The temperature recorded (TRC) is the measurement read and recorded. If the thermometer or sensor has been calibrated, then TRC can be adjusted to TNS.

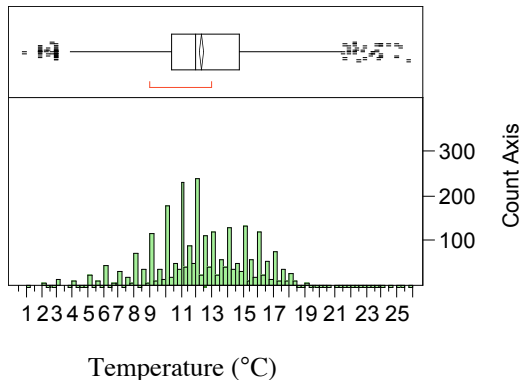
- *What are typical aquatic temperatures, and criteria levels?* The SF Bay Water Quality Control Plan states that "The natural receiving water temperature of inland surface waters shall not be altered unless it can be demonstrated to the satisfaction of the Regional Board that such alteration in temperature does not adversely affect benefical uses," and that " the temperature of any cold or warm freshwater habitat shall not be increased by more the 2.8C above natural receiving water temperatures," (Lee and Taylor, 1995). However, it does not specify the desired temperatures, since they vary depending on stream type and location. For streams in the parks with cold-water habitat, the temperature preferences of threatened species such as coho salmon and steelhead trout may provide some guidance.

The optimal thermal tolerance range for Coho salmon is from 11.4 to 16.6^{o}C (Coutant, 1977). Steelhead trout prefer slightly cooler temperatures, from 10 to13^{o}C (Bjornn and Reiser, 1991). However, salmonid and other fish can survive excursions above their tolerance range if the exposure is brief or if the increase is slow (days). Most salmonids can also survive much lower temperatures than their normal range but growth is slow and long-term population success is adversely affected. For example, for Coho the best growth is 11.8-14.6^{o}C, which is also a good temperature for migrations upstream (Brett, 1952; Reiser and Bjornn, 1979). In contrast, lower temperatures, from 4.4 to 9.4°C are best for spawning and the early life stages of the fry (Brett, 1952; Reiser and Bjornn, 1979).

In small streams, the temperature can vary as much as 10 ^{o}C over the diel (24 hour) cycle and a similar amount between shady and sunlit reaches (Bilby, 1984). A diel cycle (monitoring using a data logger) should be carried out in both sunny (exposed or riparian trees not present) and tree shaded reaches. Temperatures in smaller streams fluctuate more than temperatures in larger streams and lakes. In bigger streams cooler temperatures are often found where groundwater enters the stream and at the tail end of gravel banks. Fish use such areas as needed, but monitoring is not needed unless some special fishery need has been identified. In large pools and lakes, however, there can be considerable variation with depth, requiring several measurements to obtain a depth profile, especially in the summer when stratification is likely to occur. In larger rivers, coastal bays and estuaries, the temperature varies less, about 3°C per day, but is generally higher than in shaded streams since full sunlight reaches the water.

- *What is the history and current state of temperature monitoring in the SFBAN?*

Temperature at GGNRA and PRNS has been consistently measured in the field with hand-held YSI meters. As is shown in Figure 4, the median value for temperature in GGNRA streams is 12 °C with an interquartile range (IQR – see section 5.xxx for a definition and discussion) from 10.4 to 14.8 °C. In PRNS, the median value is 11.9 °C with an IQR from 10.3 to 14.2 °C. The National Park Service Coho Salmon and Steelhead Trout Restoration Project (CSRP) is conducting additional long-term temperature monitoring with data loggers to characterize the diurnal variations and thermal range of streams that are critical to the protection of coho salmon and steelhead trout (Ketcham, 2001). Temperature in PINN streams has been monitored starting in 1998 with a median value of 12 °C and an IQR from 10.9 to13.6 °C.



Quantiles

100.0%	maximum	25.800
99.5%		23.497
97.5%		18.327
90.0%		16.600
75.0%	quartile	14.800
50.0%	median	12.000
25.0%	quartile	10.400
10.0%		8.500
2.5%		5.020
0.5%		2.804
0.0%	minimum	1.000

Moments

Mean	12.407246
Std Dev	3.370943
Std Err Mean	0.0636707
upper 95% Mean	12.532092
lower 95% Mean	12.2824
N	2803

Figure 4. Distribution of Temperature (°C) in GGNRA Streams

The heights of the green bars in the histogram represent the number of times an observation was recorded, with the numbers on the left y-axis showing the counts. The outlier box-plot above the histogram shows the interquartile range within the box. The line across the middle of the box identifies the median sample value. The diamond represents the mean and 95% confidence interval. The lines extending from each end of the box, or the whiskers, encompass the quartiles +/- 1.5x(interquartile range). Points beyond the whiskers indicate extreme values that are possible outliers. The red bracket along the edge of the box identifies the shortest half, or the densest 50% of the observations. To the right of the histogram, the quantiles and moments are displayed. The total number of observations is listed as N.

- *What are recommendations involving temperature monitoring?* Water temperature is a basic parameter that needs to be measured at every sampling event. In addition, some one-time special temporal and spatial monitoring is needed. Extensive use of data loggers is recommended with each stream being characterized in detail over a year. It is important to note where in the stream the sensors are located and how representative of the overall temperature that cross-sectional location is.

Calibration and inter-calibration of the YSI meters should be carried out on a regular basis (Radtke et al., 1998a). Consult the USGS *National Field Manual* (Radtke et al., 1998a) for more complete QA/QC practices. It is good practice at the beginning of each year to place all the probes in the same bucket, leave them a minute to equilibrate, and then compare the readings. Sometimes they will not give identical readings. If so, note the differences and correct. Calibration to a standard value is important for temperature since fish health depends on it so a National Standard mercury-in-glass thermometer should be purchased to find how far from the real temperature the meters read. The Standard Thermometer should be stored in its metal container at room temperature and not exposed to extremes. Lacking the Standard Thermometer, alternative guaranteed thermometers may be used. The lower temperature end of the probes can also be checked by melting ice, the water from which will be at 0°C until all the ice is gone.

Thermistors included in other field-measurement instruments must be calibrated routinely (Radtke et al., 1998a). Accurate determination of other field measurements depends on accurate temperature measurements. This must be underscored for thermistors incorporated in specific electrical conductance, dissolved-oxygen, and pH instruments, such as the hand-held YSI, because these thermistors are used for automatic temperature compensation of the measurement being made.

4.1.3 Acidity, pH

- *What is acidity and pH?* The acidity of the water is measured by pH. It is the negative log of the effective concentration (activity) of the hydrogen ion (H+), and the scale runs from 0 to 14 (Hem, 1985; Radtke et al., 2003). A pH of 7 is considered neutral because it contains an equal concentration of H+ ions and hydroxide ions (OH-). A pH of less than 7 is considered acidic, and a pH of greater than 7 is basic. Since it is measured on a log scale, a change of one unit can make a big difference, for example going from a pH of 7 to 6 means a ten-fold increase in the concentration of H+.
- *Why is pH important?* At one time, limnologists considered pH to be the "master variable" since it affects the direction or thermodynamic equilibrium of many chemical reactions. The pH of water does affect many biological and chemical processes, most importantly the toxicity of copper and ammonia. However, it is now realized that although pH influences the direction of chemical reactions, their rates in the environment are often set by microbial action (Hem, 1985), which is more influenced by food and temperature.

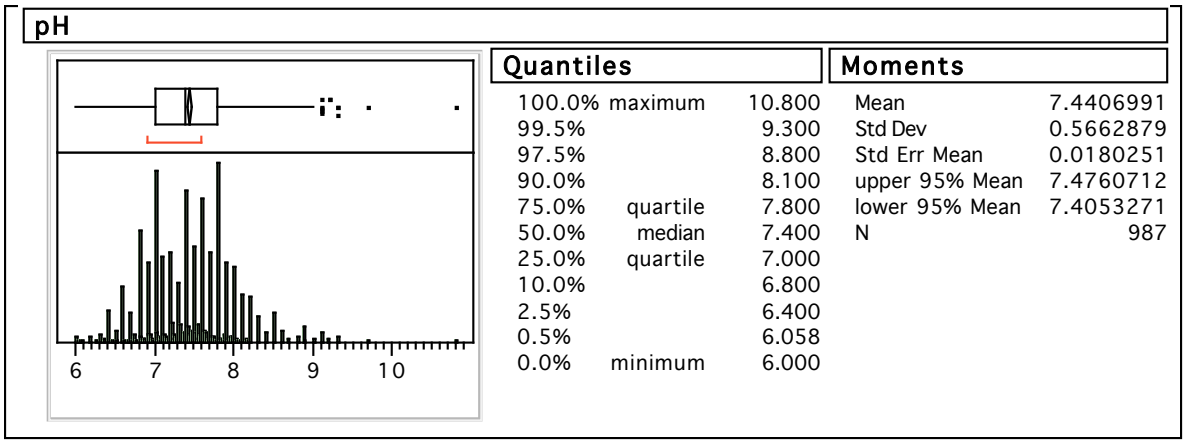
The Freshwater Workgroup Subcommittee for the NPS Vital Signs Program recommends that pH be included in the core set of required parameters for all parks (Rosenlieb et al., 2002).

- *What causes changes in pH?* In distilled water, small amounts of acid or alkali will cause very large changes in pH. If chemicals such as calcium are present in water, the acid or alkali will interact with the calcium, which acts as a buffer to reduce swings in pH (Hem, 1985). Soft water streams such as those in the calcium-poor granite streams of the Sierra Nevada are poorly buffered. In the California Coast Range streams however, the sedimentary rock provides ample calcium and alkalinity. In a well-buffered system with good alkalinity, such as is found in these parks, there shouldn't be much change in pH. In general, pH will go down as a result of acid-rain deposition or high levels of respiration in the water column. On rare occasions, pH may be artificially elevated due to construction and road run-off. The range of pH in Coastal Range streams is about 6 (slightly acid) in the upper reaches in winter to 9.5 (definitely alkaline) in nutrient rich pools when algae grow abundantly in summer. In the sea, pH is always slightly alkaline (~7.5) but is so well buffered that it changes little under normal conditions.
- *How is pH related to other parameters and overall water quality?* If algae are abundant, large diurnal fluctuations in pH and DO may occur due to algal photosynthesis and respiration. High levels of respiration will increase the concentration of dissolved carbon dioxide, which forms carbonic acid and decreases the acidity (Hem, 1985). Acidity of water is an important parameter but changes over time in Coastal California are usually small since they are well buffered by calcium in the soils and water. In

contrast, high levels of photosynthesis remove carbon dioxide and release alkaline ions raising the pH.

Temperature has a strong effect on hydrogen ion activity and must be taken into account when determining and recording pH measurements.

- *How is pH measured and what units is it reported in?* It is easily measured in the field with a hand-held YSI unit. It is reported in units of the base 10 log concentration of free protons in the water. Since it is a log scale, a change of one pH unit is actually a 10-fold increase in acidity. However, after much discussion limnologists accept that the average of two pH values, for example pH 7.0 and 7.6 is equal to 7.3 and not the average of the log of these two values.
- *What are typical pH values and criteria levels?* River water in areas not influenced by pollution generally has a pH in the range of 6.5 to 8.5 (Hem, 1985). Typical pH levels in the coastal central California region are in the range of 6 to 9 (Horne, Personal Communication), and the RWQCB water quality objective for pH is between 6.5 and 8.5 (Lee and Taylor, 1995). Typical pH levels observed in the PRNS area are from 7 to 8.5 (Ketcham, 2001).
- *What is the history and current state of pH monitoring in the SFBAN parks?* At GGNRA and PRNS pH has been measured in the field with Waterproof Oakton pH meters with a detection limit of 0.1 pH units. They are calibrated regularly (Ketcham, 2001) The median value in GGNRA is 7.4 with an IQR from 7.0 to 7.8, which is comfortably within the desired range. There have, however been values as extreme as 6 and 10.8. In the PNRS, the median value is 8.0 with an IQR from 7.7 to 8.3. However, the top 10% of values are above 9.0, which is excessively high. It has been suggested that some of these measurements may have been taken soon after completion of road or path work using concrete. However, unless such observations are documented at the time the measurement is made, we have no way to know if this is the case, or if something else was causing excessively high pH or if the probe was simply malfunctioning. Without a documented reason, it is not valid to exclude these values as outliers (Zar, 1999). At PINN, pH has been monitored since 1997 with a median value of 7.9 and IQR of 7.7 to 8.2. However, the upper 10% of values are above 8.5, which exceeds the RWQCB objectives. Values of pH greater than 9.5 are likely errors due to malfunctioning of the pH meter or probe or to calibration problems. Such errors are quite common with field pH meters, which are sensitive to various disturbances. The field measurements of pH at PINN are consistently higher than the lab measurements of pH. This may be due to an actual change in the pH from the field to the lab, or it may be that the pH meter needs to be calibrated.



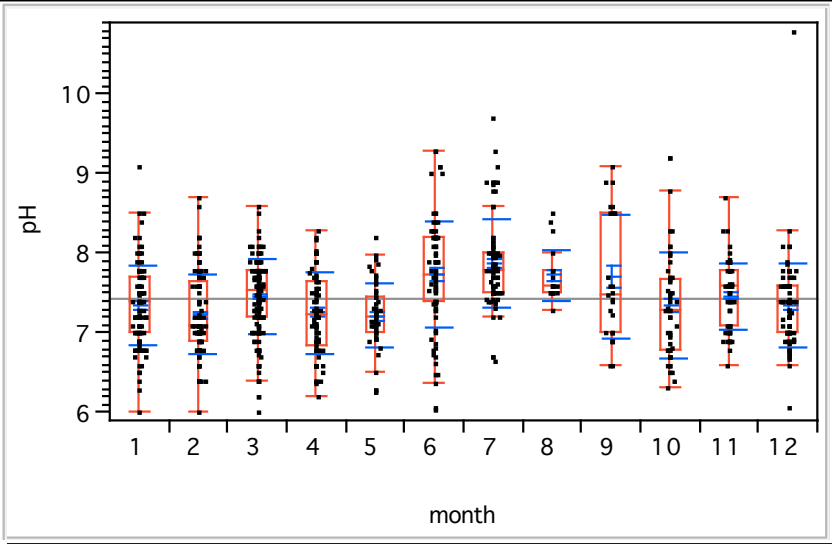
Quantiles		
100.0%	maximum	10.800
99.5%		9.300
97.5%		8.800
90.0%		8.100
75.0%	quartile	7.800
50.0%	median	7.400
25.0%	quartile	7.000
10.0%		6.800
2.5%		6.400
0.5%		6.058
0.0%	minimum	6.000

Moments	
Mean	7.4406991
Std Dev	0.5662879
Std Err Mean	0.0180251
upper 95% Mean	7.4760712
lower 95% Mean	7.4053271
N	987

Figure 5. Distribution of pH in GGNRA Streams

The heights of the green bars in the histogram represent the number of times an observation was recorded. The outlier box-plot above the histogram shows the interquartile range within the box. The line across the middle of the box identifies the median sample value. The diamond represents the mean and 95% confidence interval. The lines extending from each end of the box, or the whiskers, encompass the quartiles +/- 1.5x(interquartile range). Points beyond the whiskers indicate extreme values that are possible outliers. The red bracket along the edge of the box identifies the shortest half, or the densest 50% of the observations. To the right of the histogram, the quantiles and moments are displayed. The total number of observations is listed as N.

Oneway Analysis of pH By month



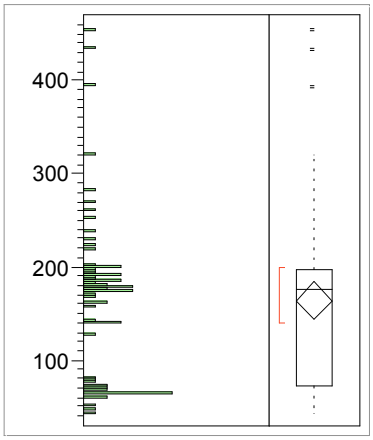
Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
1	6	6.8	7	7.4	7.7	8	9.1
2	6	6.74	6.9	7.1	7.65	7.9	8.7
3	6	6.8	7.2	7.53	7.8	8	8.6
4	6.2	6.6	6.8525	7.225	7.64	8	8.3
5	6.25	6.776	7	7.2	7.45	7.824	8.2
6	6.03	6.793	7.4	7.725	8.2	8.5	9.3
7	6.65	7.4	7.5	7.8	8.015	8.888	9.7
8	7.3	7.5	7.5	7.6	7.8	8.38	8.5
9	6.6	6.9	7	7.495	8.5	8.87	9.1
10	6.31	6.6	6.8	7.3	7.685	8.22	9.2
11	6.6	6.9	7.1	7.6	7.8	7.9	8.7
12	6.06	6.824	7	7.39	7.6	7.8	10.8

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1	117	7.34658	0.504041	0.04660	7.2543	7.4389
2	133	7.22684	0.494876	0.04291	7.1420	7.3117
3	165	7.45412	0.473718	0.03688	7.3813	7.5269
4	80	7.25337	0.514993	0.05758	7.1388	7.3680
5	53	7.21811	0.409939	0.05631	7.1051	7.3311
6	88	7.73273	0.677939	0.07227	7.5891	7.8764
7	85	7.87094	0.559012	0.06063	7.7504	7.9915
8	21	7.72381	0.320788	0.07000	7.5778	7.8698
9	30	7.70433	0.785821	0.14347	7.4109	7.9978
10	53	7.34094	0.665173	0.09137	7.1576	7.5243
11	69	7.46609	0.415367	0.05000	7.3663	7.5659
12	93	7.34860	0.531717	0.05514	7.2391	7.4581

Figure 6. pH in GGNRA Streams by Month



Quantiles

100.0%	maximum	454.00
99.5%		454.00
97.5%		438.00
90.0%		258.40
75.0%	quartile	196.00
50.0%	median	176.00
25.0%	quartile	72.00
10.0%		64.00
2.5%		45.20
0.5%		42.00
0.0%	minimum	42.00

Moments

Mean	163.79014
Std Dev	86.139362
Std Err Mean	10.222862
upper 95% Mean	184.179
lower 95% Mean	143.40129
N	71

Figure 7. Distribution of Alkalinity in GGNRA Streams

The heights of the green bars in the histogram represent the number of times an observation was recorded. The outlier box-plot above the histogram shows the interquartile range within the box. The line across the middle of the box identifies the median sample value. The diamond represents the mean and 95% confidence interval. The lines extending from each end of the box, or the whiskers, encompass the quartiles +/- 1.5x(interquartile range). Points beyond the whiskers indicate extreme values that are possible outliers. The red bracket along the edge of the box identifies the shortest half, or the densest 50% of the observations. To the right of the histogram, the quantiles and moments are displayed. The total number of observations is listed as N.

• *What are recommendations involving pH?* Monitoring of pH should be continued at all sampling events. The calibration of the pH meter in the laboratory before the day's work should involve two standard solutions, for example at pH 6 and 9 (other slightly different calibration values are OK). Further QA/QC methods can be found in the USGS *National Field Manual* (Radtke et al., 2003). When pH values over 8.5 or under 6.5 are measured in the field, it would be useful to look around for signs of recent construction or other activities that might affect pH and record these observations.

To determine an appropriate baseline specific to park streams, the data for pristine sites should be analyzed separately from the dairy sites.

4.1.4 Conductivity, Salinity and Total dissolved solids (TDS)

• *What are conductivity, salinity and TDS?* Total dissolved solids (TDS, mg/L) is often used interchangeably with salinity, but there is a subtle difference between the two. TDS refers to all of the dissolved solids in water including minerals, metals and organic matter. Salinity includes only salts with halides (fluorine, chlorine, bromine and iodine) such as potassium chloride, and bicarbonates such as sodium bicarbonate (Creek Connections, 2004c), thus excluding organics and many metals. TDS gives an

approximation of both the hardness (= acid buffering capacity) and the amount of dissolved salts in the water. Its complement is TSS (total suspended solids that measure insoluble salts). Together TDS and TSS make up total solids (TS). Measuring TDS directly by weighing the dry solids is tedious and so conductivity is usually used to approximate TDS and salinity. The relationship between specific conductance (SC) and TDS is very approximate (typically TDS ~ 0.6 x SC) depending on the specific make up of the dissolved solids in a water sample (Hem, 1985; Penoyer, 2003). As its name suggests, conductivity measures the ability of the water to carry an electrical current. Since the dominant ions dissolved in water are electrically charged, conductivity also monitors TDS and salinity. Conductivity depends on the temperature but is automatically compensated in most meters.

- *Why are specific conductance and TDS important?* They are useful as a general measure of water quality. Although they don't measure any one specific ion, such as ammonia or nitrate, they do detect large changes in the concentration of these ions and can be an alert to changes in the water quality. A constant level of TDS is essential for the maintenance of aquatic life because the density of total solids determines the osmotic strength of the water, and thus flow of water in and out of an organism's cells (Creek Connections, 2004c). Specific conductance is one of the core required parameters recommended by the Freshwater Workgroup Subcommittee for the NPS Vital Signs Program (Rosenlieb et al., 2002). The PRNS monitoring program uses conductivity as a secondary indicator of impaired conditions (Ketcham, 2001).
- *What influences changes in conductivity?* In rivers, the conductivity is mainly affected by the geology of the area through which the water flows (Behar et al., 1996; Creek Connections, 2004c; Hem, 1985). Close to the ocean, it is influenced by saltwater intrusion and tidal fluxes, making it a less useful parameter for detecting pollution. In general, groundwater usually has higher conductivity than pristine surface flows, since it spends more time in contact with soil particles (Creek Connections, 2004c). As a result, perennial streams with groundwater sources may have higher conductivity during low flow periods when most or all of the water comes from groundwater then during high flow periods when most of the water comes from surface runoff. However, if the surface runoff has high concentrations of TDS due to erosion or high nutrient loadings, and base flows come from a reservoir rather than groundwater, the opposite pattern could be seen, with elevated conductivity during high flow periods.
- *How is conductivity related to other parameters and overall water quality?* As a measure of TDS and salinity, dramatic changes in conductivity are indicative of changes in concentrations of TDS which can include nitrate, phosphate and metals as well as sodium. High conductivity warns of problems in freshwaters and gives an idea of the water quality in coastal bays and estuaries. Generally, a high or increased conductivity in a freshwater stream indicates pollution upstream. For example, any inflow from sewage (treated or raw) will contain chloride ions since almost all human food is salted with NaCl (table salt) (Behar et al., 1996). Runoff from highways is often slightly saline and soil disturbances from construction release other kinds of ions that increase conductivity. Since conductivity is a very sensitive measurement, it can show unexpected or hidden pollution in freshwater systems.

Conductivity is very highly correlated with discharge (Hem, 1985), so analysis of specific conductance must include consideration of discharge measurements. The

relationship between conductivity and discharge is not, however, always linear. In general, as discharge increases, solutes are diluted by the greater volume of water, and conductivity decreases. On a shorter temporal scale, the opposite pattern is observed in many streams, although closely studied in only a few places. There is a tendency for the water of a rising stage to have a considerably higher dissolved-solids concentration than the water passing the sampling point at an equal flow rate after the peak discharge has passed (Hem, 1985). The factors that control the concentration of water early in a flood event are different for different streams and sampling points. In general, however, when a sudden large inflow of water occurs upstream from a sampling point, the flood wave moving down the channel tends to push water already in the channel ahead of it. If the stream had a low or moderate flow rate before the rise began, the water in the channel would be relatively high in TDS, and as the wave moved downstream a large volume of this more highly mineralized water might accumulate in the wave front, resulting in higher conductivity than would be predicted by the flow rate (Hem, 1985).

- *How is conductivity measured and what units is it reported in?* It is measured with a meter and a probe containing two electrodes. A known voltage is applied to the electrodes and the voltage drop caused by the resistance of the solution is used to calculate its conductivity per centimeter. It is reported in micro-siemens per cm (uS/cm). Specific conductance is the conductance measured at 25°C. Since conductivity varies with temperature, it should always be reported as specific conductance. TDS can be measured directly by drying down a filtered water sample and weighing the residue, but more often it is calculated from the conductivity and temperature. It is reported in units of mg/L or ppm. Salinity is usually reported in parts per thousand (ppt) or Practical Salinity Units (PSU). Practical Salinity Units define salinity in terms of the conductivity of a standard solution containing 32 g potassium chloride at 15 C, which represents the salinity of typical seawater, 35 PSU or 35 ppt (UNESCO, 1985)
- *What are typical concentrations and criteria levels?* Distilled water has a conductivity in the range of 0.5 to 3 uS/cm and the conductivity of rivers generally ranges between 50 to 1500 uS/cm (Behar et al., 1996). Seawater has a specific conductivity on the order of 50,000 uS/cm or 35 PSU (Hem, 1985). The average concentration of TDS for the world's rivers is 100 mg/L, while North American rivers average 143 mg/L (Creek Connections, 2004c). Most inland streams supporting good mixed fisheries have a range of 150 to 500 uS/cm and industrial water can range as high as 10,000 uS/cm (Behar et al., 1996). However, ecosystems are adjusted to local conditions, and a large change in TDS concentrations may disrupt the system and increase its overall sensitivity (Creek Connections, 2004c). At about 1700 uS/cm, the salt levels become lethal to freshwater fish (Ketcham, 2001). Most streams measured within Point Reyes and Golden Gate are within the range of 20 to 400 uS/cm, with a few on the order of 600 to 8000 uS/cm where seawater tidal influence or high pollution exists. The PRNS 2001 report concludes that sample sites with simple mean conductivity above 850uS/cm are considered impacted by land use activity (Ketcham, 2001). In addition, sites with conductivity between 500 and 850uS/cm were concluded to be likely impacted by existing land use activity (Ketcham, 2001).
- *What is the history and current state of conductivity monitoring in the SFBAN parks?* Specific conductance has been measured at GGNRA, PRNS and PINN in the field with the hand-held YSI-30 and YSI-85 meters. The units are designed for field use and

calibrated on a regular basis. The detection limit for this meter is 0.1 uS/cm. In the GGNRA, the median specific conductance measured from 1997 to 2002 is 214 uS/cm with an IQR from 150 t0 600 uS/cm. In the PNRS, it is 314 uS/cm with an IQR from 203 to 663 uS/cm. At PINN, the median specific conductance measured from 1997 to 2001 is 285 uS/cm with an IQR from 192 to 491 uS/cm.

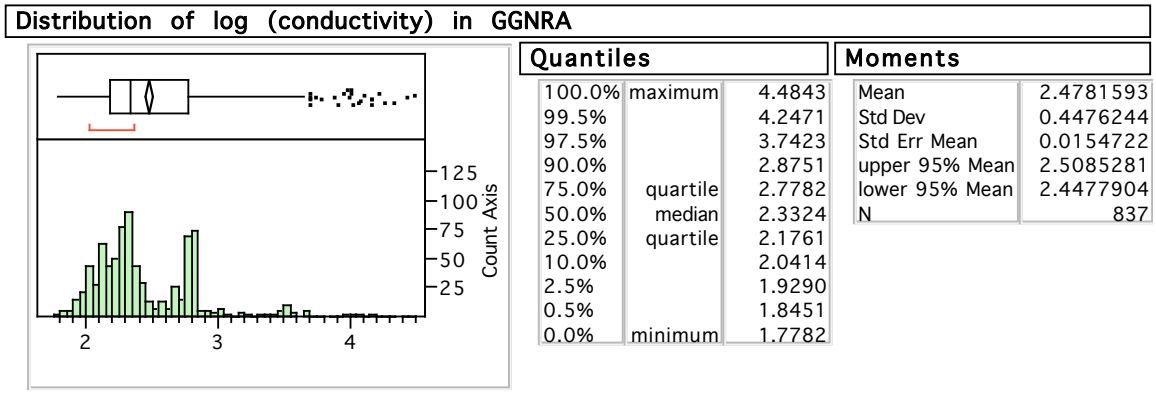


Figure 8. Distribution of Log Conductivity in GGNRA Streams

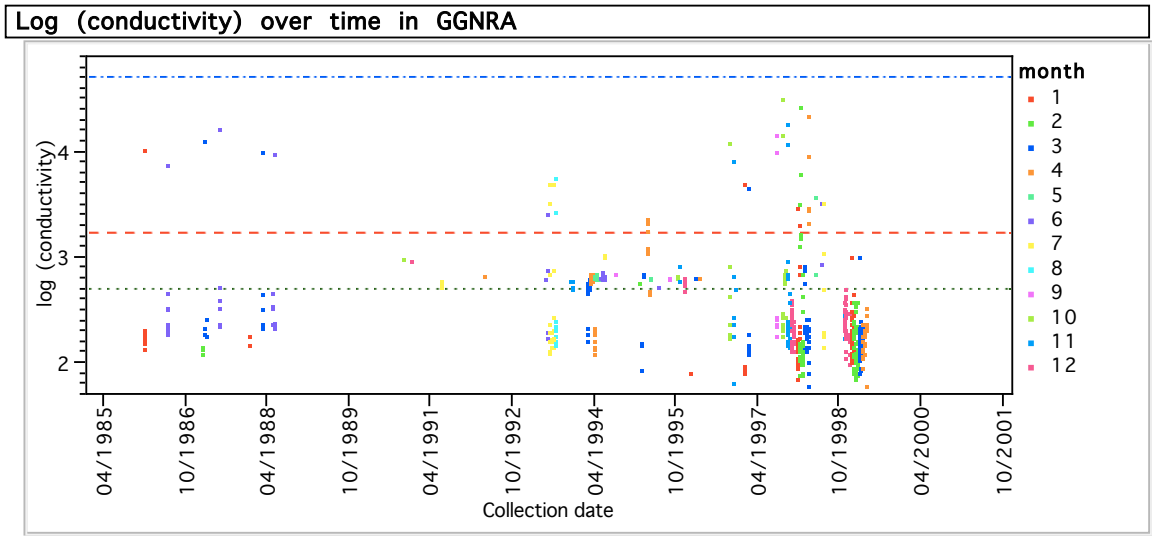


Figure 9. Log conductivity in GGNRA over time.

The blue line at the top represents the conductivity of seawater (~50,000 uS/cm). The red line represents the threshold where salinity becomes lethal to freshwater fish (1700 uS/cm). The green dotted line marks 500 uS/cm, measurements above which may indicate impacted sites.

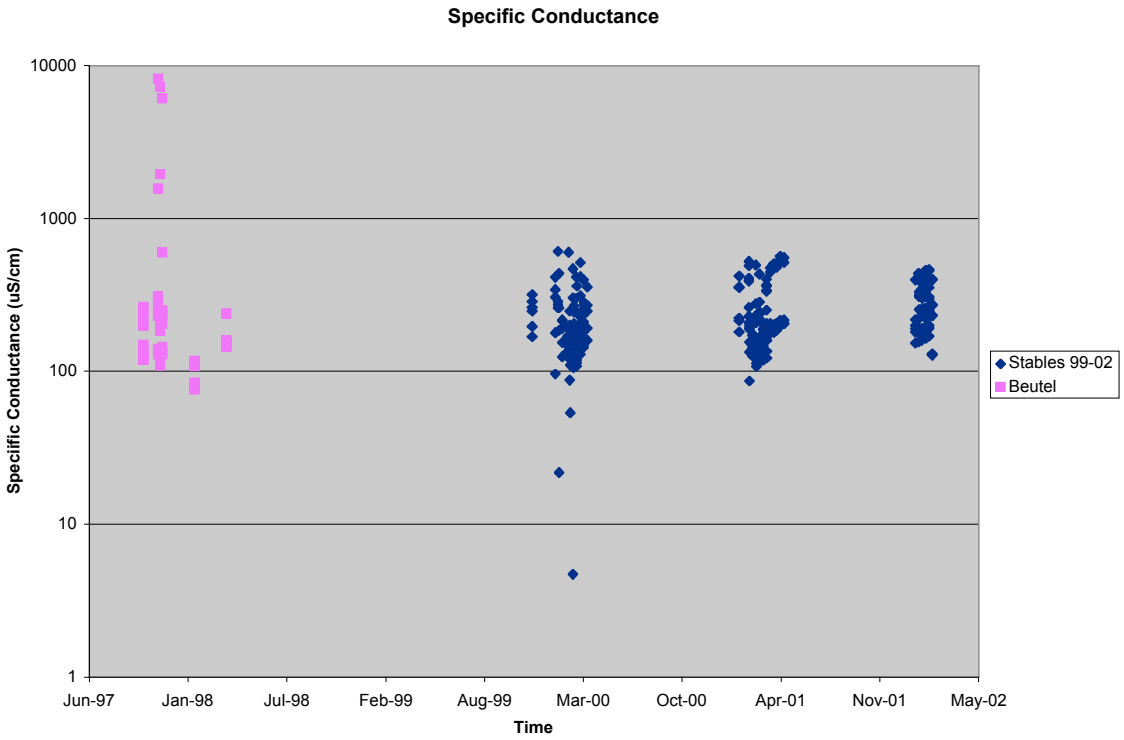
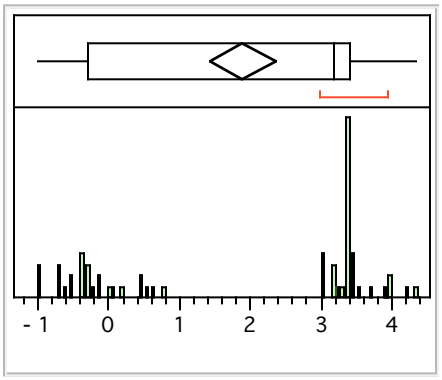


Figure 10. Specific Conductance in Stables-based Monitoring in GGNRA and PRNS

log (salinity) in GGNRA



Quantiles

100.0%	maximum	4.342
99.5%		4.342
97.5%		4.264
90.0%		3.796
75.0%	quartile	3.398
50.0%	median	3.176
25.0%	quartile	-0.281
10.0%		-0.651
2.5%		-1.000
0.5%		-1.000
0.0%	minimum	-1.000

Moments

Mean	1.896006
Std Dev	1.8869621
Std Err Mean	0.2358703
upper 95% Mean	2.3673551
lower 95% Mean	1.4246569
N	64

Figure 11. Distribution of log transformed salinity at GGNRA.

Note that this is a bimodal distribution. The high salinity values come from Rodeo Lagoon, and the low values come from streams.

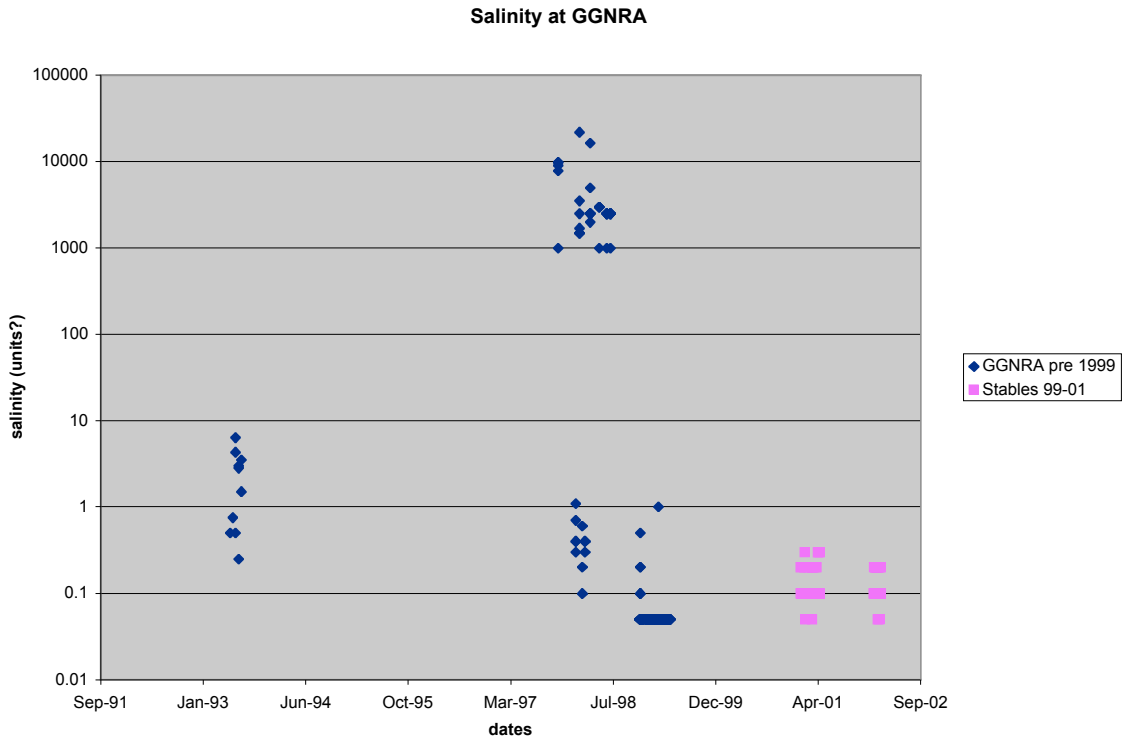


Figure 12. Salinity in GGNRA

- *What are recommendations involving conductivity monitoring?* Although it is not a particularly sensitive measure for some pollution events (Beutel 1998), it is easily measured at the same time as DO and temperature, and is useful in cases of more extreme disturbance to water quality. Specific conductance can serve as a surrogate for TDS, and is often best used as an early indicator parameter in baseline monitoring with more specific measurements of individual ions to determine cause and effect in follow-up sampling (MacDonald et al., 1991; Penoyer, 2003). Conductivity monitoring should be continued. Both the specific conductance and the raw conductivity values should be collected and reported. Specific conductance is more useful for comparison with other data, but both it and the TDS value reported by the multi-meters, are derived from the directly measured conductivity values. Not all manufacturer's instruments use the same built-in algorithms to obtain the calculated values, making it important to always collect and report the raw conductivity values along with the derived parameters (Penoyer, 2003).

Be sure to report units when recording conductivity, salinity and TDS.

As with other parameters, it would be useful to segregate the dairy sites from the pristine sites and determine the median values for each group of sites. It would then be possible to use the values from the pristine sites as a baseline to compare the dairy sites with. Specific to the case of conductivity, it would also be useful to segregate the sites according to tidal seawater influence, since elevated conductivity is normal and healthy

where seawater influence exists, but indicates pollution where seawater is not an influence.

4.1.5 Total Suspended Solids (TSS) and Turbidity

- *What are Total Suspended Solids (TSS) and turbidity?* TSS is made up of sediments and other materials suspended in flowing water. Turbidity is the capacity of suspended solids including clay, silt, finely divided organic and inorganic matter, soluble colored compounds, plankton, and microscopic organisms in water to scatter light (American Public Health Association et al., 1992). Turbidity can be thought of as the cloudiness of the water, caused by sediments absorbing and reflecting visible light (Creek Connections, 2004d).
- *Why are TSS and turbidity important?* Sediment may be the most important pollutant in streams of the Western United States (Cordone and Kelley, 1961; Larsen, 1999a). Tomales Bay, Lagunitas Creek, and Walker Creek are all listed on the 303(d) list as impaired by sediments. Turbidity is one of EPA's recommended parameters for nutrient assessments (Grubbs, 2001). High TSS (> 300 mg/L; ~ >40 NTU) inhibits fish feeding, can clog fish gills and can cover gravel spawning-beds (Creek Connections, 2004d; Horne, 2003). On the other hand, low turbidity (< 2 NTU) produces the clear sparkling water so aesthetically desired by the public (Creek Connections, 2004d; Horne, 2003). Low turbidity and TSS also allow the growth of algae on the streambed if riparian shading is not too great. There are both desirable and undesirable algae in streams but if the system is not polluted with nutrients, only the desirable species will grow. In streams, the most desirable species are the diatoms, the golden brown algae that are the "grass" of the aquatic world. The golden brown color of coastal California streams is due to a thin coating of diatoms on the rocks and cobble. These diatoms can move slowly and form a nutritious biofilm on the rocks that is a major food supply for valuable insect larvae such as mayflies, caddis flies and snails. In more open western streams, diatoms are as important as leaves from the riparian canopy in terms of providing food for insects and thus the fish that feed on them. Clear water with low TSS is vital since not only must sunlight reach the steam bed but sediment-laden water will scour the rocks killing the diatom biofilm.
- *Where do suspended solids come from and where do they go*? Most TSS is eroded from the stream banks or re-suspended from the streambed during high water. Some TSS is also washed into streams from the watershed during heavy runoff events. Thus, peak turbidity and TSS are common during floods. Some TSS can also come from algal and bacterial growth. As the water travels downstream, solids are deposited in the streambed as water slows down in low energy areas.
- *How do turbidity and TSS relate to other parameters and water quality as a whole?* The same general relationship that was discussed earlier between water quality and conductivity apply to water quality and the desired TSS and turbidity. Both TSS and turbidity are simple and common methods of direct monitoring of watershed performance since part of the function of a watershed is to hold on to most particulate matter. For example, the East Walker River below Bridgeport Reservoir has moderate water clarity (turbidity ~ 20 NTU, TSS ~ 10 mg/L) due to the presence of algae and zooplankton

washed out of the reservoir. These components of water clarity provide food for stream insects and small fish resulting in an abundance of "trophy quality" German brown trout below the dam. During a drawdown of the reservoir for maintenance, the water accidentally fell to a very low level and the reservoir sediments were eroded and suspended. The downstream turbidity suddenly rose to ~1,000 NTU (~3,000 mg/L TSS). As a result, many trout died and spawning beds were clogged with sediment for several miles downstream. Nowadays the turbidity and TSS downstream are more carefully monitored during reservoir water level changes and a hand-held turbidity meter has proven of great help in monitoring. Although the E. Walker River is an extreme case, it serves as a lesson for other streams in semi-arid climates.

Nonetheless, too low a level of TSS is not desirable since sediment transport is a major function of stream flow. Too little suspended solids produces a "hungry" flow that will erode more vigorously further downstream. In California coastal streams, muddy water with high TSS is common in winter during high water flow. This is a natural result of the dry climate and occasional storms. The dry climate leaves large areas of the watershed covered only by dry grass that provides little protection from rain. The speed of an unimpeded raindrop is sufficient to shatter the soil surface. When the rain drop hits wet soil it throws small soil particles up in the air and these particles are easily washed into small rills and thus to the stream. In contrast, if the raindrop hits a tree leaf, its velocity when it reaches the ground is too low to shatter the soil mantle. However, if the muddy floods are in winter, the insects and fish are adapted to cope with the problems. Many insect are in various refuges such as deeper in the gravel, present as eggs or in small side creeks that escape high and muddy flows. Therefore, the interpretation of high TSS or turbidity as pollution or as a natural event depends on the circumstances. In terms of conventional pollution, many contaminants, such as phosphorous, metals and pesticides are attached to particles and travel with them. Increased levels of TSS often also means increased levels of any particle-associated contaminants in depositional areas.

- *How are TSS and turbidity measured and in what units are they reported?* TSS is measured by filtering the water, then drying and weighing the filtered material. The GGNRA wet lab has the equipment necessary to do this. Results are reported in mg/L. Turbidity is an indirect measure of suspended solids and is measured with a portable turbidity meter. Turbidity is reported in Nephelomtric Turbidity Units (NTU). Consult the *National Field Manual* (Wilde and Gibs, 1998) for details on commonly used field instruments used in field monitoring of turbidity. Although TSS and turbidity rise when the stream carries more particles, there is no simple conversion from turbidity to TSS or vice versa (Wilde and Gibs, 1998). Since turbidity can be measured directly in the field or at the field laboratory, it is usually the measure of choice.
- *What are typical TSS levels and criteria limits?* According to the 1995 RWQCB Basin Plan "Waters shall be free of changes in turbidity that cause nuisance or adversely affect beneficial uses" (Lee and Taylor, 1995) Visible turbidity is found at greater than 5 NTU (Strausberg, 1983; Wilde and Gibs, 1998). The 2001 PRNS report recommends that over 50mg/L of TSS or 100 NTUs is cause for concern (Ketcham, 2001). A level of 40 mg/L TSS is about the level for concern unless the flow is storm water (Horne, 2003). For most fisheries, a level of > 300-400 mg/L TSS would actually clog fish gills and would also cover gravel beds where they spawn although some salmon swim through much higher TSS from glacial flour in Arctic rivers (Horne, 2003). The UC Cooperative

Extension Fact Sheet on Fishery Habitat (Larsen, 1999a; Lloyd, 1987) provides a summary of the effects of varying turbidity and TSS concentrations on salmonids.

- *Where has TSS been measured, for how long and what can we tell from this?* In the stables study, grab samples were collected at every sampling event in sealed Whirlpaks and stored on ice in the field. Samples were transferred to a refrigerator and analyzed with 7 days of collection by NPS personnel in accordance with Standard Methods (APHA 1985). At PINN, TSS has been measured since from 1997 to 2001, not including 1999. The median value is below the detection limit of 5mg/l. The upper IQR is 22mg/L and instances of excursions up to 425 mg/L have been recorded. At PRNS, TSS has been monitored since 1997, with a median value of 17 mg/L, and an IQR from 6 to 86 mg/L.

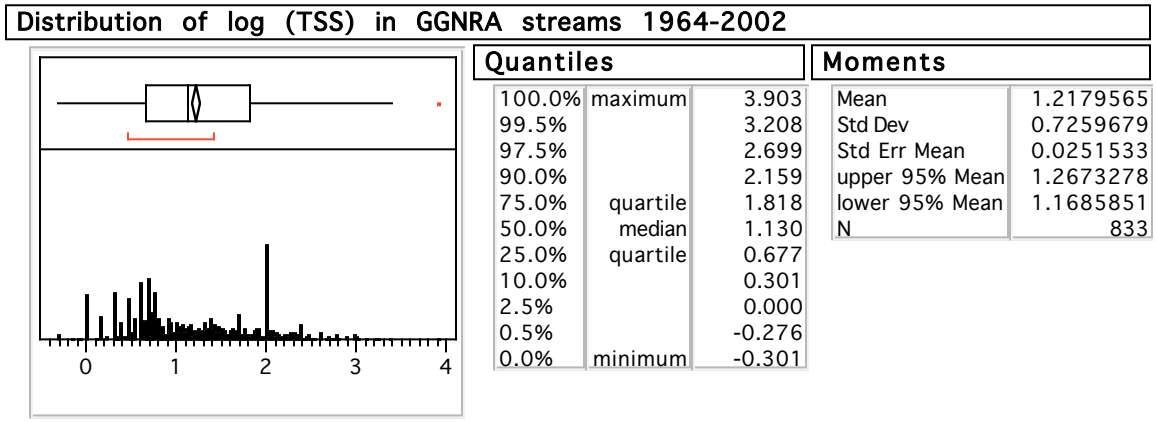


Figure 13. Distribution of Log Transformed TSS in GGNRA Streams

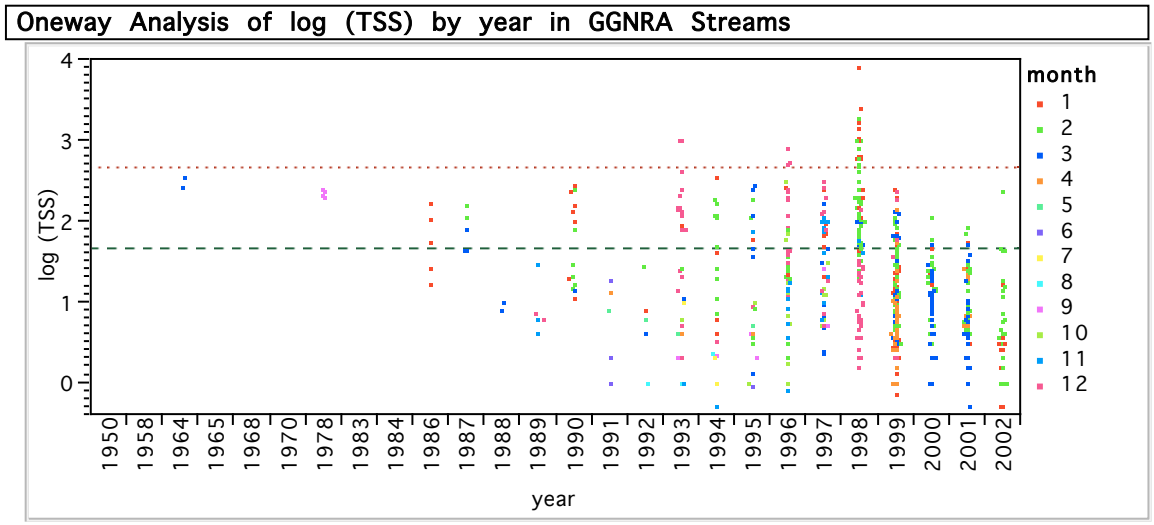
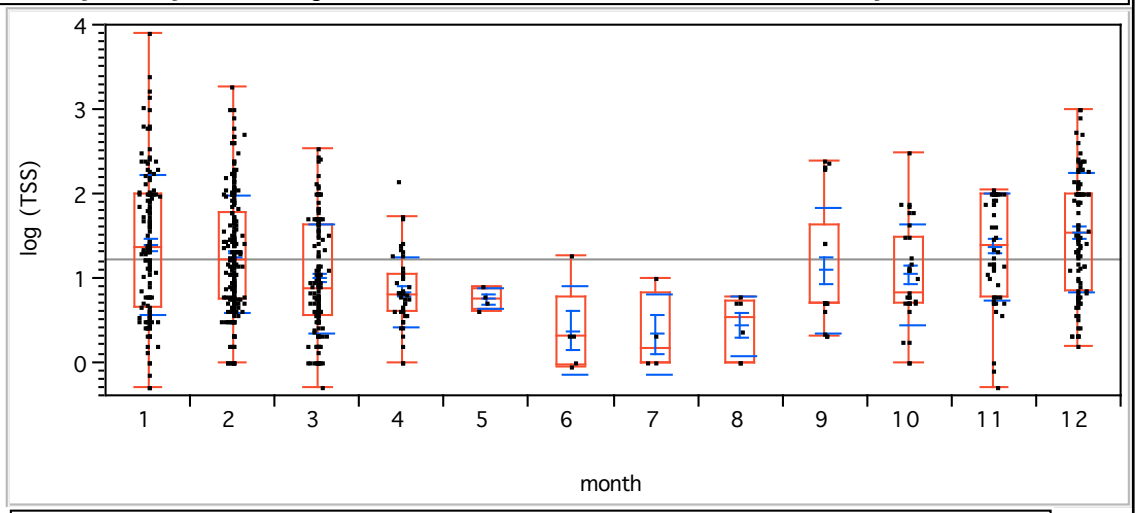


Figure 14. Log TSS over time in GGNRA streams.

The grey dashed line represents 40 mg/L of TSS, above which is possibly cause for concern. The red dotted line represents 400 mg/L of TSS, above which point sediments may clog fish gills. The color coded dots represent the month, with 1 representing January and 12, December.

Oneway Analysis of log (TSS)in GGNRA streams 1964-2002 By month



Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
1	-0.30103	0.39794	0.640424	1.358877	2	2.399032	3.90309
2	0	0.477121	0.740363	1.217484	1.778151	2.228384	3.256958
3	-0.30103	0.176091	0.544068	0.874094	1.623249	2	2.5302
4	0	0.388249	0.60206	0.789777	1.050821	1.383687	2.14145
5	0.60206	0.60206	0.626287	0.738561	0.871855	0.90309	0.90309
6	-0.04576	-0.04576	-0.02288	0.30103	0.778151	1.255273	1.255273
7	0	0	0	0.150515	0.825257	1	1
8	0	0	0	0.530349	0.716941	0.770852	0.770852
9	0.30103	0.313448	0.69897	0.69897	1.635232	2.344682	2.396896
10	0	0.230449	0.69897	0.819544	1.491362	1.869069	2.491362
11	-0.30103	0.650515	0.778151	1.396406	2	2	2.041393
12	0.176091	0.544068	0.845098	1.527976	2	2.37097	2.995635

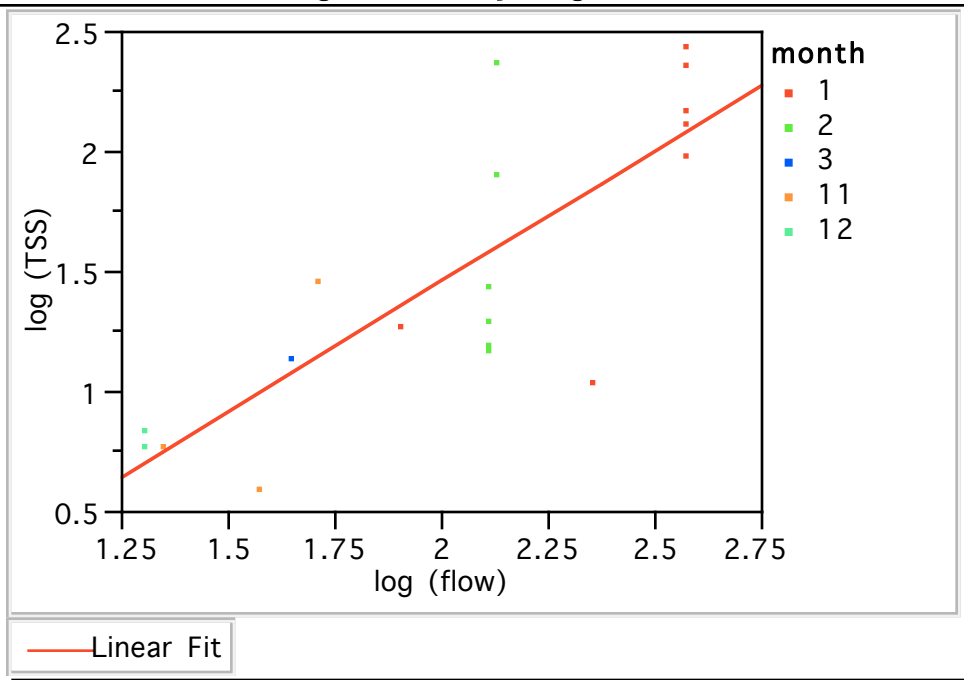
Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1	142	1.38273	0.834548	0.07003	1.244	1.5212
2	235	1.27769	0.695452	0.04537	1.188	1.3671
3	168	0.98978	0.650052	0.05015	0.891	1.0888
4	48	0.82452	0.424877	0.06133	0.701	0.9479
5	4	0.74557	0.127333	0.06367	0.543	0.9482
6	5	0.36232	0.525043	0.23481	-0.290	1.0142
7	4	0.32526	0.471681	0.23584	-0.425	1.0758
8	6	0.42175	0.356349	0.14548	0.048	0.7957
9	22	1.07923	0.746310	0.15911	0.748	1.4101
10	31	1.03503	0.598866	0.10756	0.815	1.2547
11	54	1.36250	0.630471	0.08580	1.190	1.5346
12	114	1.52688	0.705210	0.06605	1.396	1.6577

Figure 15. TSS in GGNRA streams as a function of month.

It appears that TSS varies with season and flow, however there are so few data points for the summer months, that it is not a fair comparison.

Bivariate Fit of log (TSS) By log (flow) for LAG1



Linear Fit

log (TSS) = -0.720648 + 1.0910569 log (flow)

Summary of Fit

RSquare	0.653747
RSquare Adj	0.633379
Root Mean Square Error	0.365362
Mean of Response	1.496918
Observations (or Sum Wgts)	19

Analysis of Variance

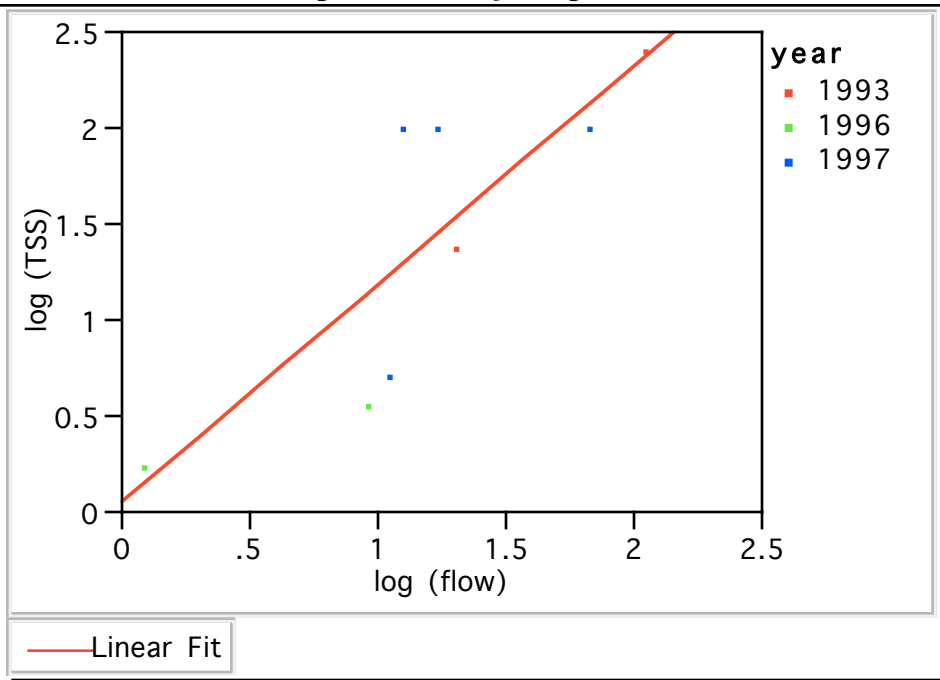
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	4.2846062	4.28461	32.0970
Error	17	2.2693149	0.13349	Prob > F
C. Total	18	6.5539211		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.720648	0.400295	-1.80	0.0896
log (flow)	1.0910569	0.192582	5.67	<.0001

Figure 16. Log TSS as a function of Log Flow at LAG1 Nov, 1989- March, 1990.

Bivariate Fit of log (TSS) By log (flow) at RDW4



Linear Fit

log (TSS) = 0.0496668 + 1.1379045 log (flow)

Summary of Fit

RSquare	0.686254
RSquare Adj	0.633963
Root Mean Square Error	0.492297
Mean of Response	1.409059
Observations (or Sum Wgts)	8

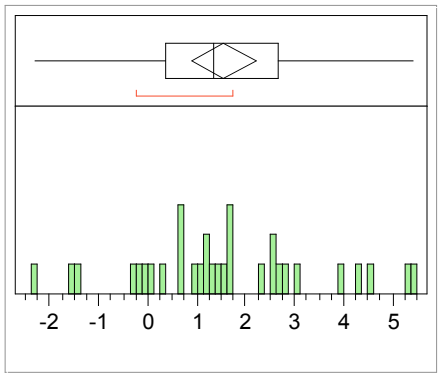
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	3.1806167	3.18062	13.1237
Error	6	1.4541369	0.24236	Prob > F
C. Total	7	4.6347536		0.0111

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.0496668	0.413647	0.12	0.9083
log (flow)	1.1379045	0.314107	3.62	0.0111

Figure 17. Log TSS as a function of Log Flow at RDW4



Quantiles

100.0%	maximum	5.407
99.5%		5.407
97.5%		5.407
90.0%		4.432
75.0%	quartile	2.672
50.0%	median	1.358
25.0%	quartile	0.370
10.0%		-1.077
2.5%		-2.303
0.5%		-2.303
0.0%	minimum	-2.303

Moments

Mean	1.5534239
Std Dev	1.8669291
Std Err Mean	0.3300296
upper 95% Mean	2.2265236
lower 95% Mean	0.8803242
N	32

Figure 18. Distribution of Log Turbidity (NTUs) in GGNRA Streams

The heights of the green bars in the histogram represent the number of times an observation was recorded. The outlier box-plot above the histogram shows the interquartile range within the box. The line across the middle of the box identifies the median sample value. The diamond represents the mean and 95% confidence interval. The lines extending from each end of the box, or the whiskers, encompass the quartiles +/- 1.5x(interquartile range). Points beyond the whiskers indicate extreme values that are possible outliers. The red bracket along the edge of the box identifies the shortest half, or the densest 50% of the observations. To the right of the histogram, the quantiles and moments are displayed. The total number of observations is listed as N.

- *What are recommendations regarding TSS monitoring?* It is recommended that a field turbidity meter measuring NTU be purchased (~ $800 to $1,200) and used as a routine monitoring tool. The field unit measurements should be compared with the laboratory measurements. In addition, the calibration of turbidity meters is dependent on the range and there is no simple range adjustor. So if very different turbidities are to be compared (e. g. 2 and 20 or 20 and 100 NTU) then care is needed to ensure that the values on the instrument really do show increase of 10 or more times as in the example shown).

The historical data needs to be examined, especially for flood flows to ensure that the typical annual history of water clarity (as TSS and/or turbidity) has been documented. If not, then a one-year project to establish a normal season should be attempted and the results of future less intensive routine monitoring could then be more easily interpreted.

4.2 OXYGEN RELATED PARAMETERS

4.2.1 Dissolved oxygen (DO)

- *What is dissolved oxygen (DO)?* Dissolved oxygen (DO) is the concentration of oxygen dissolved in the water column.
- *Why is DO important?* Dissolved oxygen is important to measure for several reasons. First, if there is not enough DO, fish and other respiring aquatic organisms can suffocate (Larsen, 1999b). Second, the DO level is indicative of the redox of the system (Hem, 1985) and tells you whether to expect to find nitrogen in the form of "harmless" nitrate or possibly toxic ammonia and nitrite. It is very important to realize that the oxygen "atmosphere" in water is very different from that in air. Terrestrial air breathing organisms from bacteria to deer have direct access to air with 21% oxygen by volume. Aquatic organisms have access to only about 0.5% oxygen by volume. Thus, oxygen is easily depleted in water but almost never in the air.

Dissolved oxygen is one of the core required parameters recommended by the Freshwater Workgroup Subcommittee for the NPS Vital Signs Program (Rosenlieb et al., 2002).

- *Where does DO come from and where does it go?* Sources of oxygen are from the atmosphere and algal photosynthesis (Hem, 1985; Larsen, 1999b). Oxygen in the water is used up by respiration by living organisms in the water, including fish, algae and bacteria. Oxygen is also used in the process of oxidizing many wastes, including ammonia and organic matter (Rittman and McCarty, 2001). Flux of oxygen from air to water is slow, so in stratified lakes a layer of oxygen-depleted water may form on the bottom while the surface water maintains higher DO levels (Horne and Goldman, 1994). An exception is well-mixed and turbulent streams with waterfalls and riffles where oxygen in rapidly introduced to the water. In such regions, oxygen depletion is rare. In contrast, in many normal streams the oxygen demand from falling leaves in autumn can depress DO even in the absence of human or agricultural pollution.
- *How does DO relate to other parameters and water quality as a whole?* DO is a concern in the summer and fall as temperature rises, water flow drops, and leaf fall adds oxygen demand. Oxygen demand is measured as BOD (Biological Oxygen Demand in mg/L of DO) and all organic matter from urea to leaves to dead fish has a measurable BOD. As temperature rises, the water is able to hold less DO (Weiss, 1970). Solubility of oxygen in water is also affected by the partial pressure of oxygen in the air and the dissolved –solids concentration in the water (Hem, 1985; Radtke et al., 1998b). Reduced flow generally means reduced turbulence and mixing, which limits flux of DO into the water. As leaves fall in late summer and early fall, they are coated by a microbial biofilm of bacteria and fungi that feed on the cellulose and lignins in the dead leaves. As the microbes feed, they use up large amounts of oxygen. Since the leaf biofilm is the major source of food for stream insects and thus fish, the leaf BOD and potential lower DO is a normal part of a stream's seasonal cycle. Additional BOD from stables or eutrophication, however, is not normal. In addition, microbial consumption of BOD can be more rapid at warmer temperatures.

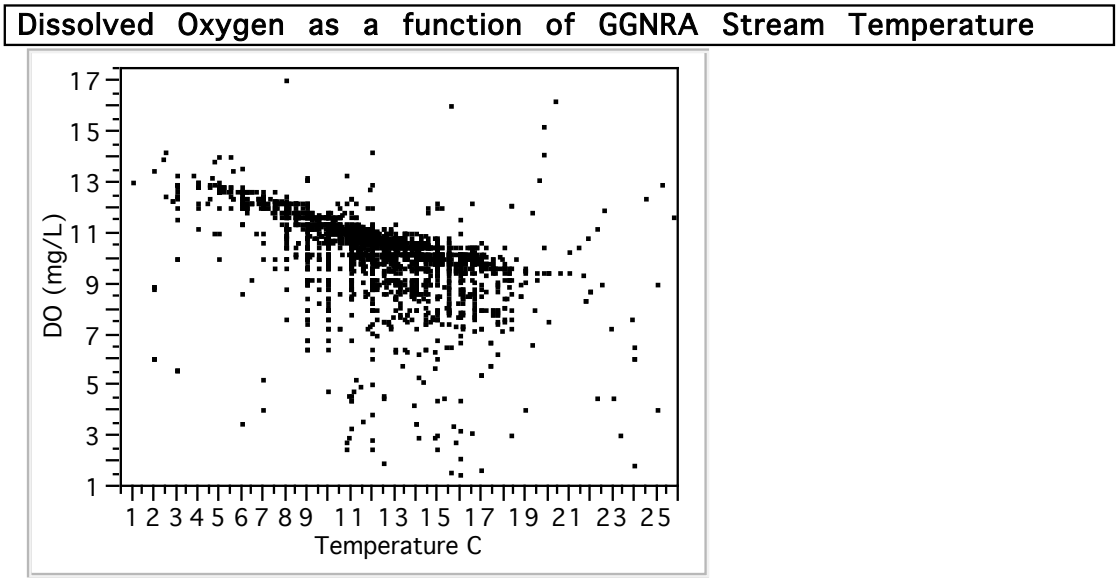


Figure 19. Dissolved Oxygen as a function of Temperature

DO is also a concern in eutrophic conditions. The water is supersaturated and contains an abundance of DO during the day while algae are photosynthesizing, but at night they continue respiring without producing oxygen and can create dramatic daily swings in the DO levels (Creek Connections, 2004a; Horne and Goldman, 1994). On top of that, when an algae bloom dies and sinks to the bottom, a large oxygen demand is created as it decays and can lead to anaerobic conditions. Such low DO is most common in semi-stagnant pools in summer and fall, especially if the stream is intermittent.

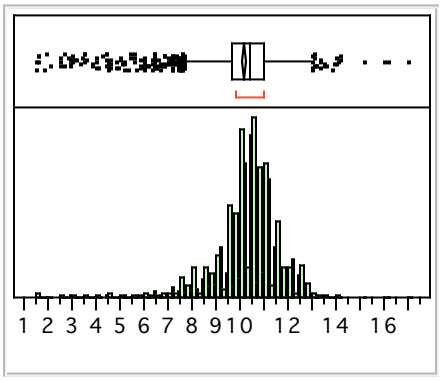
- *How is DO measured and what units is it reported in?* Since the equilibrium concentration of dissolved oxygen in the water is a function of the changing temperature of the water, DO is often reported as a percent of saturation in addition to the absolute concentration of mg/L (Radtke et al., 1998b). Water is considered 100% saturated with DO when the DO concentration is in equilibrium with the oxygen concentrations in the ambient air (Creek Connections, 2004a). It is not uncommon to have supersaturated conditions (110 to 200% DO), when photosynthetic algae and bacteria in the water produce oxygen, since it takes time for the oxygen to diffuse out of the water. DO is measured in the field with a YSI hand-held probe.
- *What are typical DO levels and criteria limits?* RWQCB criteria levels for DO in inland (fresh waters) are set 7.0 mg/L or above for cold water habitat and 5.0 mg/L or above for warm water habitat (Lee and Taylor, 1995). Estuaries can naturally have DO levels below 5 mg/L and no standard has been set by the USEPA or the RWQCB. However, such estuarine standards are under consideration at this time (2003) by the USEPA (EPA, 2001b). A level of > 3 mg/L seems to be a possible minimum for estuaries in California (Horne, 1998).

Generally, 0-2 mg/L or below 30% saturation is not enough DO to support aerobic life and is indicative of serious problems (Behar et al., 1996; Horne and Goldman, 1994). Almost all fish kills in natural waters (i.e. outside laboratory tests) are associated with DO < 2 mg/L (Horne, 2003). Such conditions can be expected in eutrophic pools where algae abound in warm still water or where excessive bacterial decomposition uses up the DO. With 2-4 mg/L of DO a few kinds of fish and insects can

survive, but the water is still only 20-60% saturated and this low level is usually associated with warm, still water and/or high nutrients and bacterial growth. Most fish are affected by a lack of oxygen when DO is less than 4 mg/L (Larsen, 1999b; Reiser and Bjornn, 1979). With 4-7 mg/L of DO, the water is good for many pond animals, acceptable for warm water fisheries, but still too low for cold water fisheries. Salmonids show signs of initial distress symptoms at 6 mg/L of DO (Larsen, 1999b; Reiser and Bjornn, 1979). Below 60% saturation is still considered poor water quality, from 60-79% is acceptable for most stream animals, and above 80% is very good for most fish (Behar et al., 1996). For good cold water habitat supporting freshwater salmonids, over 7 mg/L of DO is needed (Larsen, 1999b; Reiser and Bjornn, 1979).

It is important to note that with ample DO, fish can live through many other environmental threats, from ammonia to high TSS, which would otherwise kill them (Horne and Goldman, 1994). So maintaining a high DO can protect a fishery in many ways.

- *Where has DO been measured, for how long and what can we tell from this?* At PRNS, DO has been monitored in streams with the YSI-55 and YSI-85 since 1997, with a median value of 9.6 mg/L and an IQR from 8.1 to 10.5 mg/L. 10% of measurements were below 5 mg/L. The monitoring program at PRNS reports that low DO has been observed in association with isolated pools due to low flow, and that salmonids have been observed surviving in isolated pools with DO levels below 3 mg/L (Ketcham, 2001). In addition, ponds have been monitored since? (I have a table with locations and DO measurements but no dates). At GGNRA, DO was first monitored in 1978 and has been consistently monitored since 1983, with a median value of 10.4 and an IQR from 9.7 to 11.0 mg/L. At PINN, DO has been monitored since 1998 with a median of 9.8 mg/L and an IQR from 9.3 to 10.5 mg/L. 10% of measurements were below 7.5 mg/L, mostly in cases where flow had stopped.



DO

Quantiles		
100.0%	maximum	17.000
99.5%		13.553
97.5%		12.600
90.0%		11.800
75.0%	quartile	11.000
50.0%	median	10.400
25.0%	quartile	9.700
10.0%		8.400
2.5%		5.709
0.5%		2.700
0.0%	minimum	1.500

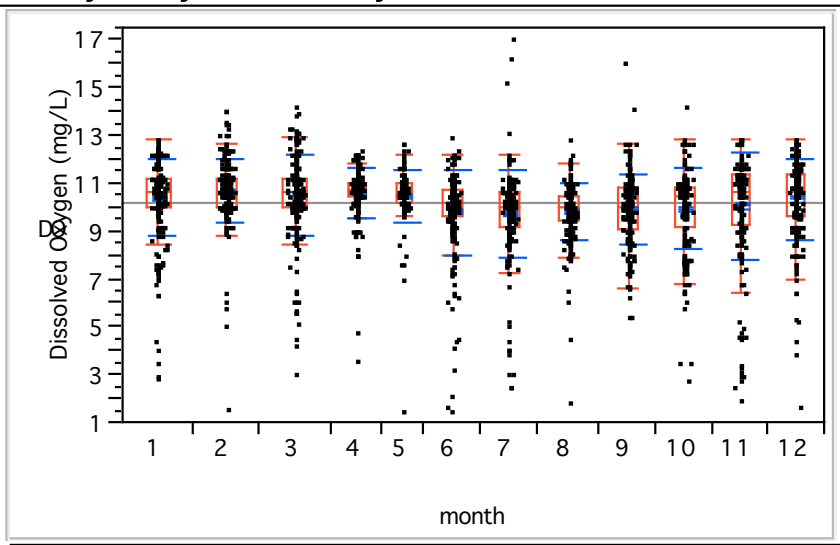
Moments	
Mean	10.171392
Std Dev	1.6223971
Std Err Mean	0.0324935
upper 95% Mean	10.235109
lower 95% Mean	10.107675
N	2493

Figure 20. Distribution of DO in GGNRA Streams

The heights of the green bars in the histogram represent the number of times an observation was recorded. The outlier box-plot above the histogram shows the interquartile range within the box. The line across the middle of the box identifies the median sample value. The diamond represents the mean and 95% confidence interval. The lines extending from each end of the box, or the whiskers, encompass the quartiles +/- 1.5x(interquartile range). Points beyond the whiskers indicate extreme values that are possible outliers. The red bracket along the edge of the

box identifies the shortest half, or the densest 50% of the observations. To the right of the histogram, the quantiles and moments are displayed. The total number of observations is listed as N.

Oneway Analysis of DO By month in GGNRA Streams



Quantiles

Month	Minimum	10%	25%	Median	75%	90%	Maximum
1	2.8	8.2	10	10.6	11.2	12.2	12.8
2	1.57	9.8	10	10.6	11.2	12.37	14
3	3	8.8	10	10.6	11.2	12.42	14.2
4	3.6	9.6	10.4	10.68	11	11.6	12.4
5	1.5	9.852	10.2	10.4	11	11.4	12.6
6	1.5	7.68	9.6	10.2	10.7	11.2	12.9
7	2.5	8.2	9.15	10	10.6	11	17
8	1.8	8.4	9.4	10	10.4	11	12.8
9	5.4	7.93	9.1	10	10.8	11.4	16
10	2.7	7.6	9.175	10.4	10.8	11.42	14.2
11	1.9	7.38	9.25	10.6	11.4	11.8	12.8
12	1.6	8.2	9.6	10.6	11.4	12.2	12.8

Means and Std Deviations

Month	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1	254	10.3627	1.60801	0.10090	10.164	10.561
2	232	10.6735	1.31074	0.08605	10.504	10.843
3	257	10.4842	1.65702	0.10336	10.281	10.688
4	181	10.5786	1.06410	0.07909	10.423	10.735
5	151	10.4605	1.08897	0.08862	10.285	10.636
6	195	9.7938	1.78841	0.12807	9.541	10.046
7	209	9.7122	1.82667	0.12635	9.463	9.961
8	218	9.7835	1.18201	0.08006	9.626	9.941
9	202	9.8799	1.45221	0.10218	9.678	10.081
10	198	9.9005	1.70400	0.12110	9.662	10.139
11	197	10.0513	2.22455	0.15849	9.739	10.364
12	199	10.3095	1.68417	0.11939	10.074	10.545

Figure 21. Dissolved Oxygen by Month in GGNRA Streams 1978-1999

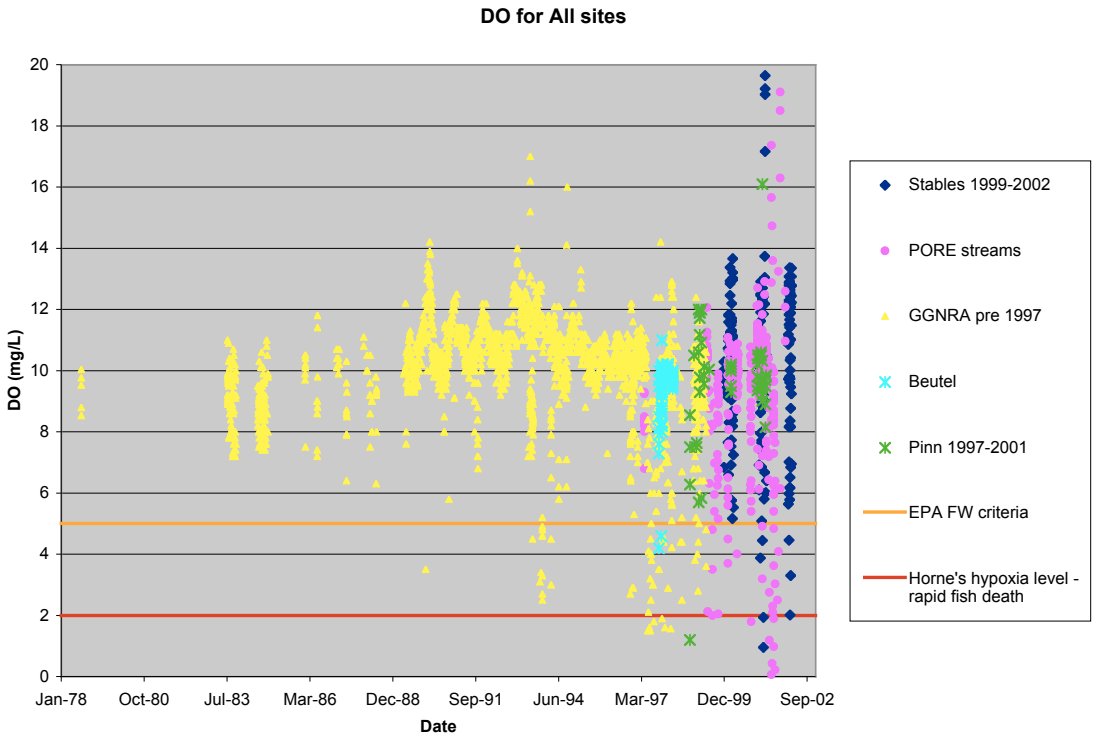


Figure 22. DO over time in PRNS, GGNRA and PINN

- *What are recommendations regarding DO monitoring?* Dissolved oxygen is an essential aspect of water quality, can be easily measured in the field with available equipment, and should continue to be monitored at all sampling events. Dissolved oxygen must be measured in situ, since degassing, mineral precipitation, and other chemical, physical, and biological reactions can cause the DO concentration of a water sample to change significantly within minutes after sample collection (Radtke et al., 1998b). DO can be measured using the standard YSI or other similar probes that measure DO, temperature and other variables depending on the manufacturer. However, special care should be taken in the calibration and operation of a DO probe. The manufacturers instructions on how to calibrate the probe should be carried out fully each morning. The membrane over the probe should be examined and if any bubbles are present or if the membrane has been used for several days, a new membrane should be put on. In the field, the probe should not be allowed to penetrate the sediments, especially if the DO is low (a REDOX probe is used for sediment oxygen levels). If the probe does accidentally hit anoxic sediments it should be allowed to equilibrate for at least a minute in higher DO water before the readings are recorded. A field kit comprising of a small burette and bottles containing the appropriate chemicals should be used to fully calibrate the DO probe by titration. The titration is simple. When I (Alex Horne) was working with the DWR we used it on small boats. Full calibration is especially important if low DO or fish kills are found or if the data is to be used for legal purposes. Further suggestions for QA/QC can be found in the USGS *National Field Manual* (Radtke et al., 1998b)

4.2.2 Biochemical oxygen demand (BOD), chemical oxygen demand (COD) and total organic carbon (TOC)

- *What are BOD and COD?* Biochemical oxygen demand (BOD) is a measure of the organic matter (TOC) in the water and how much oxygen would be required by bacteria to oxidize it to carbon dioxide (CO2) (Sawyer and McCarty, 1978). It represents the amount of oxygen consumed by bacteria and other organisms while they decompose organic matter under aerobic condtions (Delzer and McKenzie, 2003). Oxygen demand usually contains a rapidly consumed carbonaceous component (CBOD) and a more slowly consumed nitrogenous component (NBOD) (Delzer and McKenzie, 2003). Unless specified, references to BOD usually mean CBOD. BOD_5 is typically measured by letting a sealed sample sit at a specified temperature for 5 days and measuring the change in DO in the water. This does not measure the full amount of BOD in the water. By convention, twenty days is considered sufficient time to allow for a completer biochemical oxidation of organic matter, however a 20-day test is impractical when data are needed to address an immediate concern (Delzer and McKenzie, 2003). Chemical oxygen demand (COD) is a similar measure, but instead of allowing bacteria to oxidize the organic matter, a chemical oxidant is used, so the organic matter is oxidized more quickly and thoroughly (Hem, 1985; Rittman and McCarty, 2001). The measured COD is always higher than the BOD. Under natural conditions, the total COD would never be exerted since some components of COD (some lignins in wood, humic substances and high carbon ring compounds) would not decay for centuries. Thus, BOD is the preferred variable to measure in lakes and streams. Total organic carbon (TOC) more directly measure carbon and does not include nitrogenous BOD. TOC includes both dissolved organic carbon (DOC) and particulate organic carbon (POC). DOC is included in measurements of TDS and POC is included in measurements of TSS.
- *Why is BOD important?* BOD is important to measure, because it can be a predictor of low DO levels downstream. For example, the BOD in most unpolluted streams is about two mg/L and may peak at 5-10 mg/L in floods. In contrast, the BOD in the nearby soil water may be 20-50 mg/L as leaves are decomposed by microbes. Almost all soil water contains high BOD but it is used up by other microbes by the time it reaches the stream. As discussed earlier, even in normal streams the DO can easily drop in autumn (stream BOD 1-2 mg/L). Thus the effect of the addition of raw sewage (BOD 80-250 mg/L), animal waste (BOD ~ 3,500 mg/L) or winery waste (BOD up to 10,000 mg/L) is obviously serious. Even treated wastes (BOD 5 – 50 mg/L) can have a major effect if volumes are high and BOD towards the higher end. BOD is also often related to high nutrient levels via the process of eutrophication.
- *Where does BOD come from and where does it go?* Particulate and dissolved organic matter can enter streams through surface water runoff or through algal and bacterial growth within the stream or lake itself. Since any organic matter is by definition more reduced than carbon dioxide, oxygen is needed in decomposition. As the organic matter is oxidized, the BOD goes down.
- *How does BOD relate to other parameters and water quality as a whole?* BOD is a measure of the organic loading in a water body. Most bacteria need organic matter to grow, and higher levels of BOD support more bacterial growth. More importantly, BOD is a measure of how much oxygen will be used up, and so can be a warning of low DO

levels downstream or, in a lake at a later time. Higher BOD (>2 mg/L) in streams has a general relationship to poor water quality. High BOD itself may directly indicate sewage from homes or farms. High BOD may be indirectly caused by eutrophication that otherwise goes unnoticed and is due to leachate from septic tanks, farm runoff or recreational horse use.

- *How are BOD and TOC measured and in what units are theyt reported ?* Although BOD represents organic matter, it is measured as the change of DO in water as the organic matter is oxidized and so is reported as mg/L of DO consumed. The test is usually carried out by certified laboratories since it requires a 5-day incubation at a constant temperature, the addition of certain "starter" bacteria and the inhibition of ammonia oxidizing bacteria by chemical addition. However, the test can be carried out in field laboratories where approximate BOD is needed or a different kind of BOD is needed (for example actual site-specific BOD, 10 or 30 day BOD, or total BOD which includes the effects of ammonia oxidation).

Total organic carbon is measured by converting all carbon species to carbon dioxide and then measuring the total carbon dioxide (Hem, 1985). The total measured carbon dioxide minus the initial carbon dioxide is the organic carbon. DOC is measured by first filtering out the particulates.

- *What are typical BOD and DOC levels and criteria limits?* Typical raw sewage in urban areas usually has a BOD on the order of 250 mg/L. In rural areas, it is on the order of 80 mg/L. In a typical unpolluted stream, one wouldn't expect to find more than two mg/L of BOD. A stream with over 7 mg/L of BOD is likely to be negatively impacted (Horne, Personal Communication). The average DOC for rivers worldwide is 5.75 mg/L (Hem, 1985; Meybeck, 1982).
- *Where in the parks has BOD been measured, for how long and what can we tell from this?* BOD has only been monitored since 1999 at some of the sites in the stable studies. Most sites were below the detection limit of 2mg/L, however, a few sites had values up around 30 mg/L, which is definitely impaired. These same sites also had noticeable fecal coliforms (900 and 1600 MPN/100mL), but compared to other sites, these would not be considered impaired by fecal coliforms, but would be considered impaired by BOD. BOD provides additional information and should continue to be monitored. At PINN, TOC has been monitored since 1997 with a median value of 8.4 mg/L and an IQR from 6.5 to 12.3 mg/L.

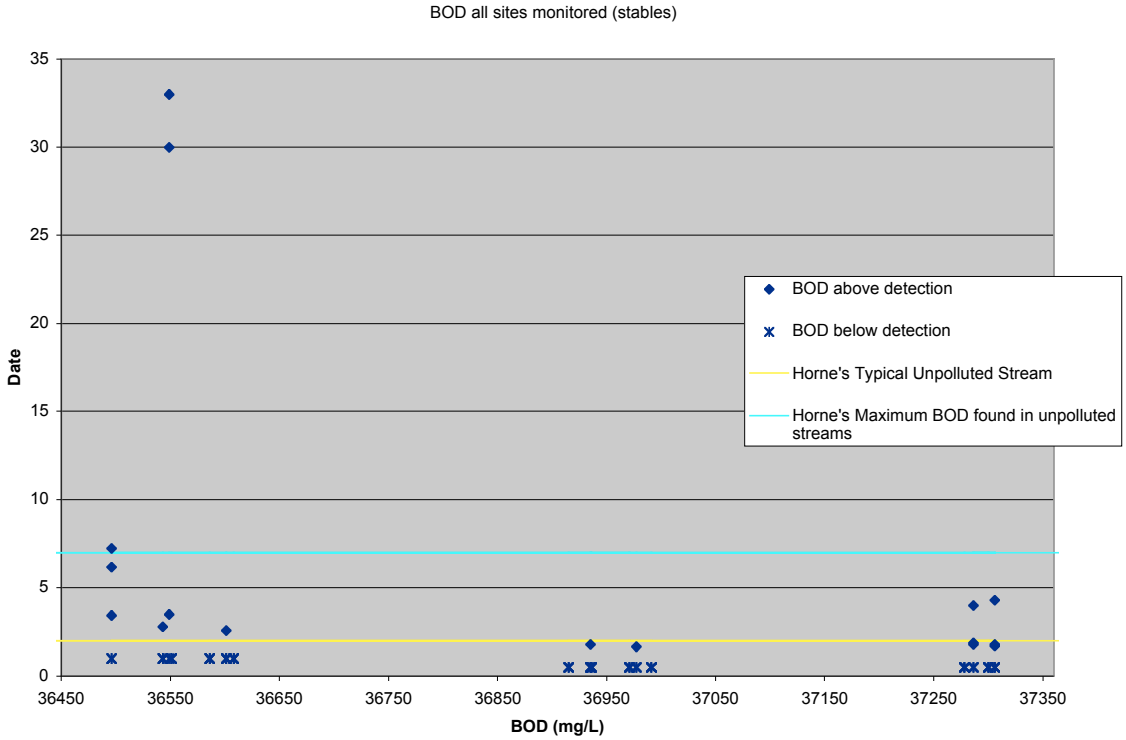


Figure 23. BOD in Stables monitoring at PRNS and GGNRA

- *What are recommendations regarding BOD monitoring?* BOD levels above 2 mg/L have been reported in a few sites and were due to pollution. Until a few years have gone by without such high BOD events, monitoring should be continued. A consideration to actual BOD should be given if models of the streams are to be used with regard to overall stream health (see earlier discussion on kinds of BOD test).

4.3 NUTRIENTS

EPA's recommended parameters for nutrient assessment are total phosphorous, total nitrogen, chlorophyll-a, and some measure of water clarity (e.g. Secchi depth for lakes and reservoirs and turbidity for rivers and streams) (Grubbs, 2001). Nitrogen and phosphorous are the main causal agents of enrichment, while the two response variables, chlorphyll-a and water clarity are early indicators of system over-enrichment for most waters.

4.3.1 Nitrogen, nitrate and ammonia

- *What is nitrogen?* Nitrogen is a nutrient. Since it is required to make proteins, enzymes and nucleic acids, it is the element needed by most organisms in the highest quantity after carbon and oxygen (Rittman and McCarty, 2001). In surface waters, it is predominately found in three main forms, nitrate, ammonia and organic nitrogen. Nitrate is the inorganic oxidized form (HNO3) and is often a limiting nutrient in eutrophic waters. Ammonia is the inorganic reduced form (NH3) and is toxic to fish at high

concentrations. Organic nitrogen is bound up with organic matter and not immediately bioavailable, but becomes important in systems with long residence times and/or nutrient cycling Some organic nitrogen is present as humic substances that color the water a yellow-brown and also act to chelate or bind potentially toxic heavy metals. When oxygen is low, a minor fraction of the nitrogen is sometimes found in the form of nitrite (HNO2), which is slightly less oxidized than nitrate but much more toxic to aquatic life. The majority of nitrogen, however, is found in the atmosphere in the form of inert nitrogen gas (N2) (Creek Connections, 2003). Plants, animals and bacteria cycle nitrogen between these forms and many others, depending on the available levels of light, oxygen and carbon (Horne and Goldman, 1994; Maier et al., 2000). The basic cycle starts with atmospheric nitrogen, which is fixed, or made bioavailable, through lightening, fertilizer manufacturing, or bacterial fixation. It is applied to fields and plants take it up and convert it to organic nitrogen, or if it was fixed by bacteria, it is taken up by plants or algae and converted into organic nitrogen. The plants and algae are eaten by animals or else die. Either process results in the release of organic nitrogen as well as the inorganic reduced form, ammonia. Some ammonia volatilizes, some is taken up again by plants and bacteria, and then converted back to organic nitrogen or nitrate, and under anaerobic conditions, some is denitrified and released back to the atmosphere as nitrogen gas.

- *Why is nitrogen important?* Nitrogen is important to monitor because as nitrate it contributes to eutrophication and low DO which are harmful to aquatic life. EPA recommends monitoring TN as an indicator of eutrophication (Grubbs, 2001). As un-ionized ammonia, nitrogen is directly toxic to fish. The monitoring program at PRNS has found toxic ammonia useful as a primary indicator of water quality degradation (Ketcham, 2001). Organic nitrogen is not directly harmful, and is more of a concern in receiving waters with long residence times such as lakes and estuaries where it will have time to enter the food chain and contribute to eutrophication than in streams where it is likely to pass through.
- *Where does nitrogen come from and where does it go?* Natural sources of nitrogen include the soil, animal wastes, and decomposing plants. The main human-influenced sources of nitrogen are sewage, fertilizers, and barnyard wastes. Sewage and barnyard wastes tend to have nitrogen primarily in the form of ammonia, while fertilizer runoff tends to have nitrogen primarily in the form of nitrate. If there is enough oxygen in the water, however, the ammonia will be converted to nitrates in the water, leading to a temporary drop in DO levels. Nitrate is very soluble and is flushed out of soils relatively quickly, while ammonia and organic nitrogen tend to be more associated with particles and surface runoff.
- *How does nitrogen relate to other parameters and water quality as a whole?* In regards to biotic production, surface waters are considered either nitrogen limited or phosphorous limited. The ratio of nitrogen to phosphorous (N:P ratio) used in algal and bacterial growth is approximately 10:1 (Horne and Goldman, 1994; Rittman and McCarty, 2001). Thus if the measured N:P ratio is greater than 10, the water is usually phosphorous limited, and a reduction in phosphorous levels will be the most effective method of reducing eutrophication. If the measured N:P ratio is less than 10, the water is usually nitrogen limited, and it will be more effective to try to reduce nitrogen levels. Eutrophication increases turbidity and leads to drastic variations in the DO levels. Low

DO levels act synergistically with ammonia to increase toxicity to fish. Photosynthesis during algae blooms also raises the pH, increasing the toxic fraction of ammonia.

- *How is nitrogen measured and in what units is it reported?* Nitrogen can be measured separately as nitrate, nitrite, ammonia and organic nitrogen or all together as total nitrogen (TN). Combined nitrate, nitrite and ammonia together are called total inorganic nitrogen (TIN), and is useful because it is the immediately bioavailable fraction of nitrogen. Combined ammonia and organic N are called Kjeldahl N. Ammonia comes in two forms, depending on the pH and temperature of the water. NH4+ is the dissociated ammonia ion (also called ionized ammonia) and is the primary form in most aquatic systems. Undissociated NH4OH (also called un-ionized ammonia) is the toxic (often written as NH3) form and is a greater fraction at higher pHs and higher temperatures. Total ammonia includes both forms, while un-ionized or undissociated ammonia includes only the toxic fraction, which is most important for fish health. For the health of the streams, TIN and ammonia are the most important forms of nitrogen to be monitoring. For lagoons and ponds, TKN or TN in the influent water is important as a source, but TIN is important as an indicator of eutrophication.
- *What are typical nitrogen levels and criteria limits?*

In the past, little guidance has been available regarding acceptable levels of total nitrogen in streams, since it is a problem of accumulation in the ecosystem rather than direct toxicity to particular organisms. Ammonia is usually considered separately since it is directly toxic. The drinking water standard is set at 10mg-N/L for nitrate, but that is much too high to be protective of many ecosystems. Recently the EPA has come out with new guidance documents for nutrient criteria development (EPA, 1999b; EPA, 1999c), including total nitrogen (TN) and total phosphorous (TP). PRNS, GGNRA and Pinnacles are all included in Aggregate Ecoregion III, subecoregion 6, for which the TN reference condition is 0.5 mg-N/L based on the 25th percentile of data available for 10 streams over a 10 year period. The TN reference condition in streams and rivers for the overall Aggregate Ecoregion III is 0.38 mg-N/L, also based on the 25th percentile over a 10 year period. EPA recommends that the criteria be based on the overall Aggregate Ecoregion reference condition, which in this case is 0.38 mg-N/L. For lakes and reservoirs, the reference condition is 0.4 mg-N/L of TN.

According to the 1995 San Francisco Bay RWQCB Basin Plan, receiving waters should not exceed an annual median of .025 mg-N/L of un-ionized ammonia or a maximum of 0.16 mg-N/L (Lee and Taylor, 1995). This is to protect against chronic toxic effects and to preclude a build up in receiving waters. EPA has also published a 1999 Update of Ambient Water Quality Criteria for Ammonia (EPA, 1999a), containing an Appendix C - Calculation of Freshwater Ammonia Criterion with a one hour CMC (criteria maximum concentration)for salmonids and not salmonids as well as a CCC (criterion continuous concentration) calculated from a thirty-day mean. These are both values for total ammonia and the calculation includes a pH as well as temperature dependence.

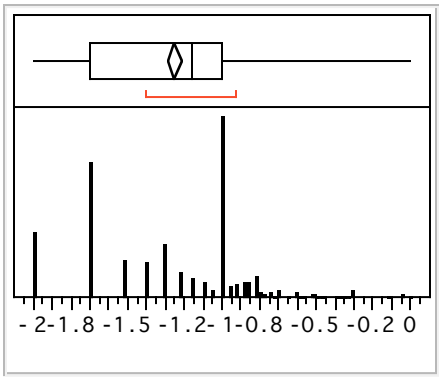
- *Where in the parks has nitrogen been measured, for how long and what can we tell from this?* In the stables study for both ammonia and nitrate, grab samples were collected at every sampling event. Ammonia samples were preserved with sulfate and both were stored on ice in the field. Samples were transferred to a refrigerator. Nitrate was analyzed within 48 hours of collection by NPS personnel with a Hach 2100

spectrophotometer using the cadmium reduction method. Ammonia was analyzed within 14 days of collection by NPS personnel with a Hach 2100 spectrophotometer using the salicylate method, measuring total ammonia. Unionized ammonia was calculated using the method defined in the Journal of Fisheries Research Board of Canada article titled *Aqueous Ammonia Equilibrium Calculations: Effect of pH and temperature* (Emerson, et al.), using pH and temperature field measurements taken at time of collection.

At PINN, nitrogen has been monitored starting in 1997, in five different parameters, field N (not sure if this is total N, NO3 or what?), Kjeldahl N, ammonia N, nitrate N and nitrite N. All of the nitrite measurements have been below the detection level of 0.5 mg/L. Most of the field N measurements were also below the detection limit of 2.5 mg/L, with a couple of samples right at the detection limit and one sample at 10mg/L. However, the field N measurements may not be reliable since, in some cases, both the field N and lab nitrate were measured for the same sample, and even though the field N was a non-detect, the lab nitrate measure was up around 5 or 6 mg/L. For the Kjeldahl N, there are only 16 samples. The median value is 0.76 mg/L, with an IQR from 0.57 mg/L to 1.55 mg/L. For ammonia, over 90% of the samples were below the detection limit of 0.1mg/L, with four samples going from 0.1 to 0.19 mg/L in 2001. Unfortunately, for two of the four samples with detectable ammonia concentrations, no pH data was collected, so it is not possible to calculate the toxic fraction. For nitrate, both the median and the lower IQR reflect the detection limits of 1.0 and 0.5 mg/L, respectively. The upper IQR is 3.2 mg/L, with values as high as 6.7 mg/L in some places.

At PRNS, nitrogen has been monitored since 1997 in the form of ammonia, nitrite, and nitrate. Over 75% of the ammonia samples were below the detection limits of 0.4 and 0.2 mg/L. The samples within the detection range went as high as 13 mg/L. Of 163 samples from 23 monitoring stations, 5 samples were associated with toxic conditions at 3 of the sampling sites. Toxic ammonia is considered a primary indicator of water quality degradation for PRNS because of its direct impact on aquatic species (Ketcham, 2001). Nitrite was mostly below the detection limits. Nitrate had a median value of 0.33 mg/L, with an IQR from 0.2 to 1.0 mg/L. Extreme values went as high as 11 mg/L. PRNS has found that stream nutrient concentrations are significantly diluted by winter flows. In several instances, tributaries showed elevated nutrient concentrations, but concentrations in the mainstem were below detection. This indicates that pollutant sources are limited within the monitored watersheds (Ketcham, 2001).

At GGNRA, nitrogen has been monitored in the forms of ammonia, nitrate, nitrite and "nitrogen" (listed in the Access database, but I'm not sure of units or test). Ammonia has been monitored since 1993. Over 60% of samples tested were below the detection limits of 0.01, 0.02 and 0.1 mg-N/L. The median value was 0.07 mg-N/L with an IQR from 0.02 to 0.1 mg-N/L. Nitrate has been monitored since 1986, with a median value of 0.8 mg-N/L and an IQR from 0.32 to 1.9 mg-N/L.



Distribution of log (NH3) in GGNRA Streams 1993-2002

Quantiles		
100.0%	maximum	0.000
99.5%		-0.041
97.5%		-0.370
90.0%		-0.824
75.0%	quartile	-1.000
50.0%	median	-1.155
25.0%	quartile	-1.699
10.0%		-1.699
2.5%		-2.000
0.5%		-2.000
0.0%	minimum	-2.000

Moments	
Mean	-1.250558
Std Dev	0.4239428
Std Err Mean	0.0199406
upper 95% Mean	-1.21137
lower 95% Mean	-1.289746
N	452

Figure 24. Distribution of log transformed ammonia in GGNRA streams 1993-2002.

The large spikes at –2, -1.7 and –1 are artifacts of the detection limits of 0.01, 0.02 and 0.1 mg-N/L.

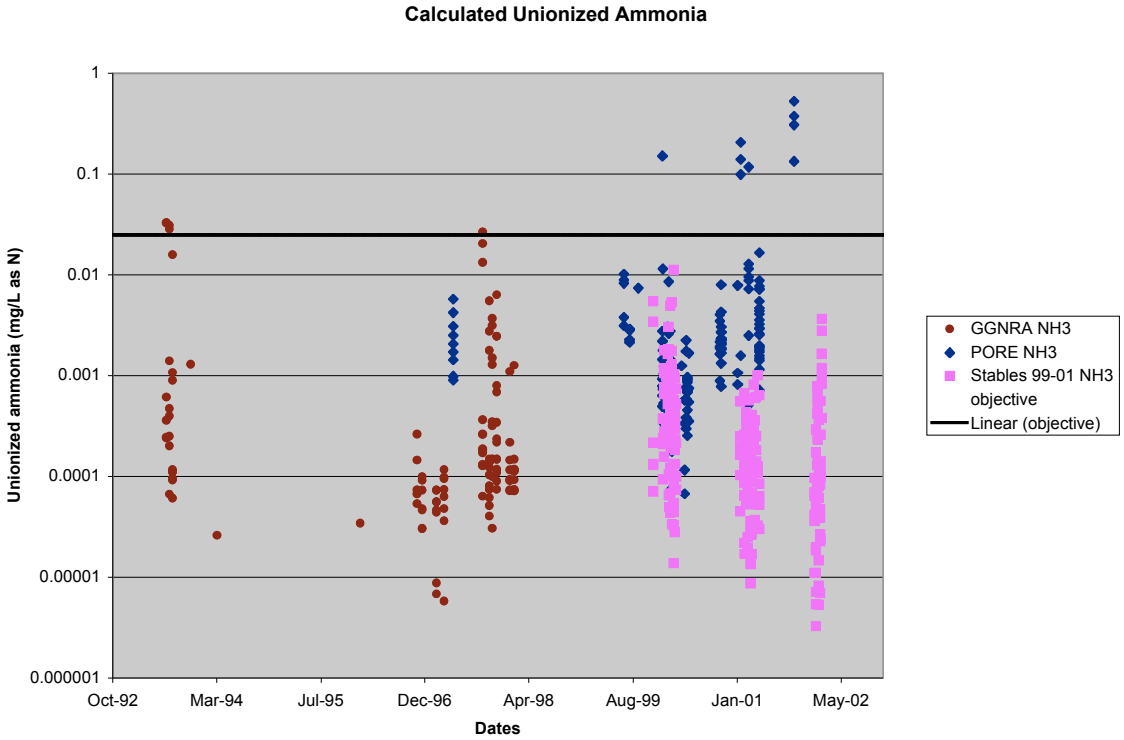
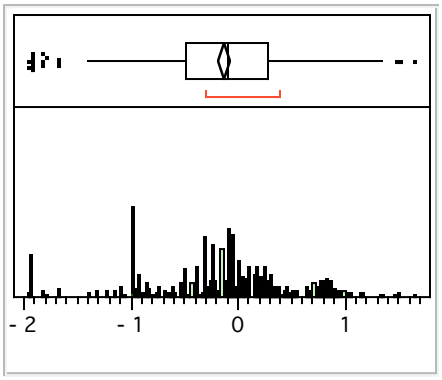


Figure 25. Ammonia over time at GGNRA, PRNS and PINN



Distribution of log (NO3) in GGNRA streams

Quantiles		
100.0%	maximum	1.653
99.5%		1.519
97.5%		0.999
90.0%		0.790
75.0%	quartile	0.279
50.0%	median	-0.097
25.0%	quartile	-0.495
10.0%		-1.000
2.5%		-1.921
0.5%		-1.923
0.0%	minimum	-1.959

Moments	
Mean	-0.13365
Std Dev	0.6947391
Std Err Mean	0.0286994
upper 95% Mean	-0.077284
lower 95% Mean	-0.190017
N	586

Figure 26. Distribution of Log Nitrate in GGNRA Streams

Note that the spikes at –1 and –1.92 are artifacts due to samples below the detection limit.

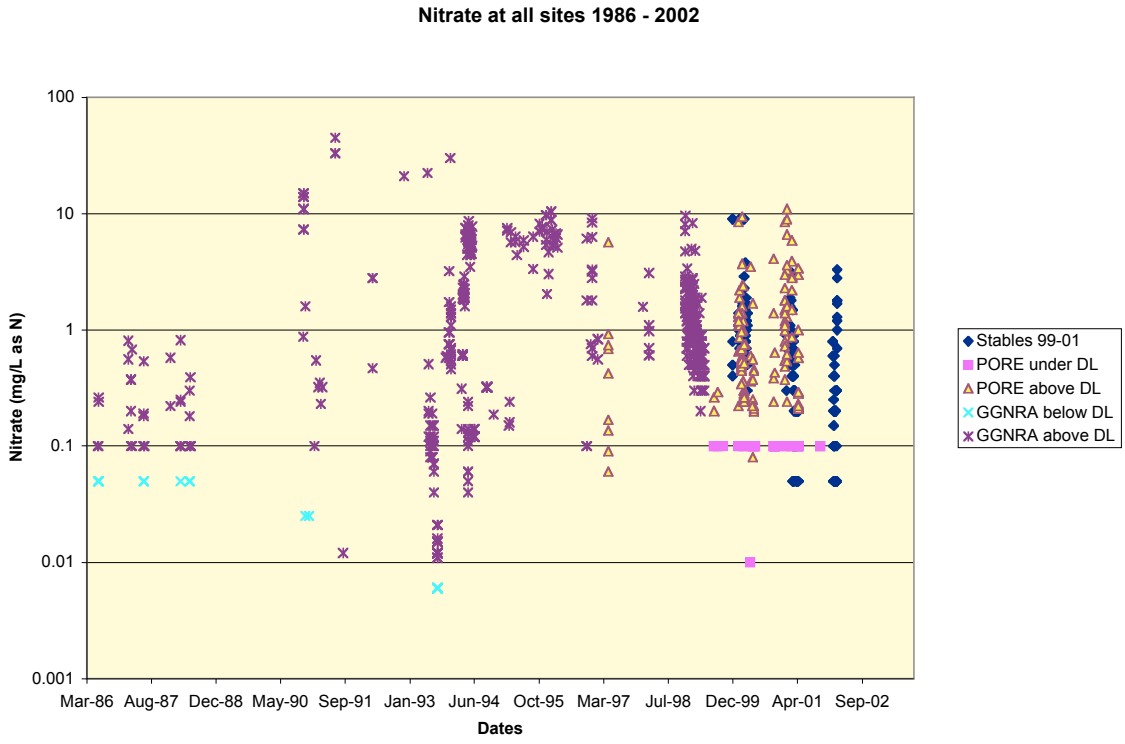


Figure 27. Nitrate at PNRS and GGNRA 1986 to 2002

- *What are recommendations regarding nitrogen monitoring?* Discontinue nitrite monitoring unless there is a reason to suspect its presence, such as: DO less than 1 mg/L, ammonia greater than 0.1 mg-N/L, or hydrogen sulfide odor.

At PINN, try filtering water with ~0.45 um HA Millipore filters before taking field N measurements. If the water is colored like tea, then direct nitrate colorimetric measurement may not work. If field N measurements continue to be unreliable, they should be discontinued.

Continue nitrate and ammonia measurements in streams at all parks. In receiving waters, such as lakes and lagoons, TN should be monitored to establish long-term trends. Ammonia should particularly be monitored during the summer when flows and DO levels are low, while temperatures are high.

Regardless of the form of the nitrogen measured, it should be reported in units of milligrams of nitrogen per liter (mg-N/L), so that there is the same amount of nitrogen in 2mg-N/L of nitrate as in 2mg-N/L of ammonia. If it were reported in mg/L of ammonia and mg/L of nitrate, it would be difficult to compare them, since one molecule of nitrate is much heavier than one molecule of ammonia. To convert from mg/L of ammonia to mg-N/L of ammonia, use the ratio of the molecular weight of nitrogen to the molecular weight of ammonia (14:18). To convert from mg/L of nitrate to mg-N/L of nitrate, use the ratio of the molecular weight of nitrogen to the molecular weight of nitrate (14:62).

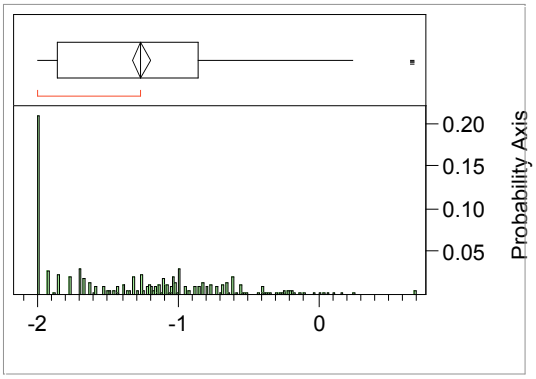
4.3.2 Phosphorous, phosphate, total phosphorous (TP) and orthophosphate (ORP)

- *What is phosphorous?* Phosphorous (P), like nitrogen, is a nutrient that is sometimes limiting in aquatic ecosystems (Creek Connections, 2004b; Horne and Goldman, 1994). Phosphorous is usually found as phosphate (H3PO4), which comes in organic and inorganic forms. Inorganic phosphate is the form that is readily bioavailable to primary producers and is also called reactive phosphate or orthophosphate.
- *Why is phosphorous important?* Phosphates do not pose any direct health risks to humans or aquatic life, but the indirect effects via increased eutrophication are significant. Total phosphorous is recommended along with total nitrogen by EPA as an indicator of eutrophication (Grubbs, 2001). Phosphorous, again like nitrogen, is a pollutant by accumulation, rather than performance, and so is more likely to be a problem in receiving waters, such as lakes, lagoons and estuaries, than in streams.
- *Where does phosphorous come from and where does it go?* Plants, algae and bacteria take up inorganic phosphate and convert it into organic phosphate as part of their tissues (Creek Connections, 2004b; Rittman and McCarty, 2001). As they die or are eaten by animals and excreted, the organic phosphate sinks to the bottom where some binds to sediments and is buried and some is converted by bacteria back to reactive phosphate to start the cycle all over again. Because there is no gaseous form of phosphorous, once it is in an aquatic system without a large outflow, it tends to cycle back and forth between the water column and the sediments without leaving the system (Horne and Goldman, 1994). Phosphates initially enter the system from soil sediments, fertilizer runoff, animal wastes and detergents (Creek Connections, 2004b).
- *How does phosphorous relate to other parameters and water quality as a whole?* As discussed in the nitrogen section, surface waters are considered either nitrogen limited or phosphorous limited with regards to eutrophication. The ratio of nitrogen to phosphorous (N:P ratio) used in algal and bacterial growth is approximately 10:1. Thus if the measured N:P ratio is greater than 10, the water is usually phosphorous limited, and a reduction in phosphorous levels will be the most effective method of reducing

eutrophication. If the measured N:P ratio is less than 10, the water is usually nitrogen limited, and it will be more effective to try to reduce nitrogen levels.

- *How is phosphorous measured and in what units is it reported?* Phosphates are measured either as total phosphates (TP), or as reactive phosphates (ORP) (Creek Connections, 2004b). ORP is easily measured with a colorimetric test, but total phosphates requires a digestive heating step to release all the bound organic phosphates and so is more time-consuming. Both tests are usually reported in mg-P/L, although sometimes they are reported in mg/L of phosphates. Phosphate is approximately 3 times heavier than phosphorous, so to convert between the two, either divide phosphate by 3 to get phosphorous or multiply phosphorous by 3 to get phosphate.
- *What are typical phosphorous levels and criteria limits?* Small oligotrophic streams may respond to P concentrations of 0.01 mg/L or less. In general any concentrations over 0.05 mg-P/L will have an impact, unless nitrogen is the limiting nutrient (Behar et al., 1996). Alex Horne (personal communication) sets the guideline for eutrophication in California surface waters at 0.1 mg-P/L. The recent EPA recommended total P criteria for Aggregate Ecoregion III streams and rivers is 22 ug-P/L or 0.02 mg-P/L, with a range of reference conditions from 0.01-.05 mg-P/L (EPA, 2000a). For Aggregate Ecoregion III lakes and reservoirs, the recommended criteria is set at .017 mg-P/L, with a range of reference conditions from 0.003 – 0.172 mg-P/L(EPA, 2001a). The value of 0.172 found in one of the reference sites is considered inordinately high, and EPA is planning to investigate further whether this is reflective of unique conditions, cultural enrichment or a statistical anomaly.
- *Where in the parks has phosphorous been measured, for how long and what can we tell from this?* In the Stables study, total phosphorous grab samples were collected at every sampling event in lab-supplied bottles. Samples were stored on ice and delivered to the lab within 6 hours of collection. Samples were preserved and analyzed by Sequoia within 28 days of collection using EPA method 365.3 ORP grab samples were also collected at every sampling event and stored on ice in the field. Samples were transferred to a refrigerator and analyzed by NPS personnel within 7 days of collection using the Hach 210 spectrophotometer and ascorbic acid method. PINN data?

Monitoring of orthophosphorous at PRNS found only low concentrations which were not useful in determining pollution sources. Has P been measured in lakes and lagoons?



Quantiles

100.0%	maximum	0.681
99.5%		0.681
97.5%		0.004
90.0%		-0.406
75.0%	quartile	-0.854
50.0%	median	-1.268
25.0%	quartile	-1.854
10.0%		-2.000
2.5%		-2.000
0.5%		-2.000
0.0%	minimum	-2.000

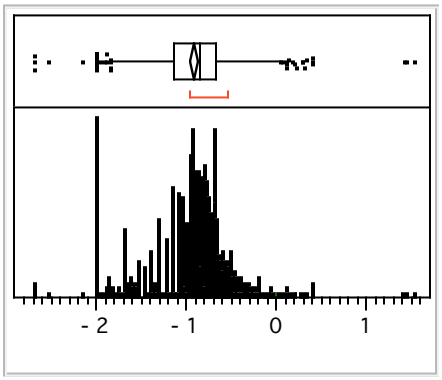
Moments

Mean	-1.26118
Std Dev	0.6030259
Std Err Mean	0.0339229
upper 95% Mean	-1.194436
lower 95% Mean	-1.327924
N	316

Figure 28. Distribution of Log Transformed Total Phosphorous in GGNRA Streams.

Note that while more of the distribution is visible, the spike at –2 is an artifact from samples below the detection limit.

Distribution of Log (ORP) in GGNRA Streams 1978-2001



Quantiles

100.0%	maximum	1.537
99.5%		1.030
97.5%		0.101
90.0%		-0.414
75.0%	quartile	-0.658
50.0%	median	-0.839
25.0%	quartile	-1.125
10.0%		-1.654
2.5%		-2.000
0.5%		-2.632
0.0%	minimum	-2.699

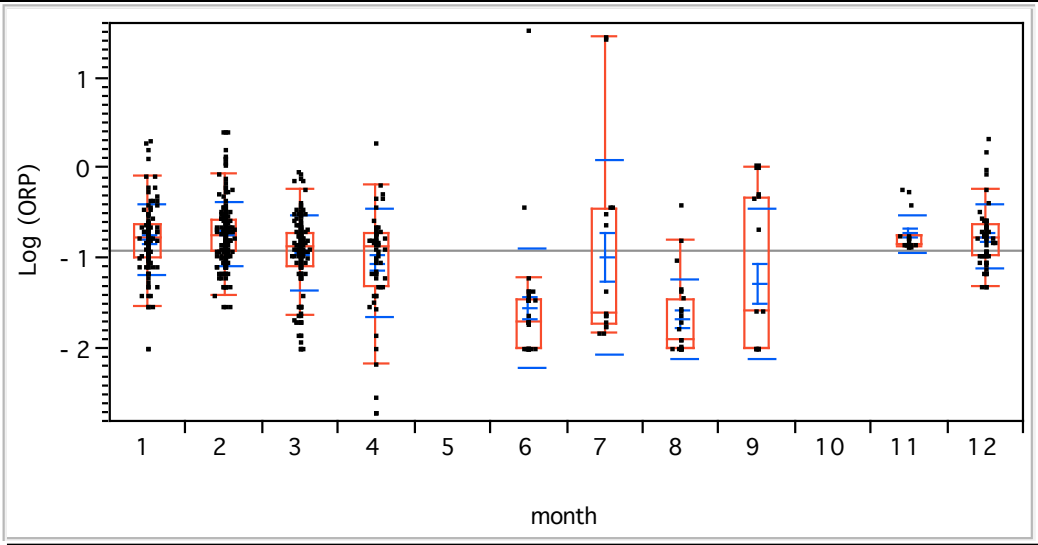
Moments

Mean	-0.911409
Std Dev	0.5126532
Std Err Mean	0.019732
upper 95% Mean	-0.872666
lower 95% Mean	-0.950153
N	675

Figure 29. Distribution of Log Transformed Orthophosphate in GGNRA.

Note the spike at –2 is due to samples below the detection limit.

Oneway Analysis of Log (ORP) By month



Quantiles

Level	Minimum	10%	25%	Median	75%	90%	Maximum
1	-2	-1.29611	-1	-0.78272	-0.63365	-0.35655	0.303196
2	-1.52288	-1.1433	-0.92082	-0.74473	-0.56864	-0.28079	0.39794
3	-2	-1.63827	-1.09691	-0.86996	-0.71568	-0.50864	-0.04576
4	-2.69897	-1.98539	-1.30103	-0.86996	-0.72125	-0.4367	0.278754
5	.	.	.	.	.	.	.
6	-2	-2	-2	-1.69897	-1.45593	-1.24684	1.537353
7	-1.82391	-1.82391	-1.73329	-1.61092	-0.45009	1.428209	1.454814
8	-2	-2	-2	-1.88606	-1.45337	-0.90734	-0.39362
9	-2	-2	-2	-1.56864	-0.31914	-0.10675	0.017033
10	.	.	.	.	.	.	.
11	-0.86646	-0.86646	-0.83863	-0.83863	-0.74473	-0.2582	-0.23657
12	-1.30103	-1.1549	-0.95861	-0.76955	-0.61979	-0.2518	0.330414

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1	110	-0.7953	0.39163	0.03734	-0.869	-0.721
2	191	-0.7329	0.34688	0.02510	-0.782	-0.683
3	162	-0.9404	0.40521	0.03184	-1.003	-0.878
4	60	-1.0534	0.60133	0.07763	-1.209	-0.898
5	0	.	.	.	.	.
6	31	-1.5522	0.67034	0.12040	-1.798	-1.306
7	16	-0.9866	1.08152	0.27038	-1.563	-0.410
8	24	-1.6776	0.43708	0.08922	-1.862	-1.493
9	13	-1.2868	0.82380	0.22848	-1.785	-0.789
10	0	.	.	.	.	.
11	17	-0.7281	0.20973	0.05087	-0.836	-0.620
12	51	-0.7627	0.35123	0.04918	-0.862	-0.664

Figure 30. Log ORP in GGNRA Streams as a Function of Month.

It appears that the bulk of ORP loading occurs during the rainy season. However, since the number of samples are not even month to month, it is not possible to make a statistically sound analysis.

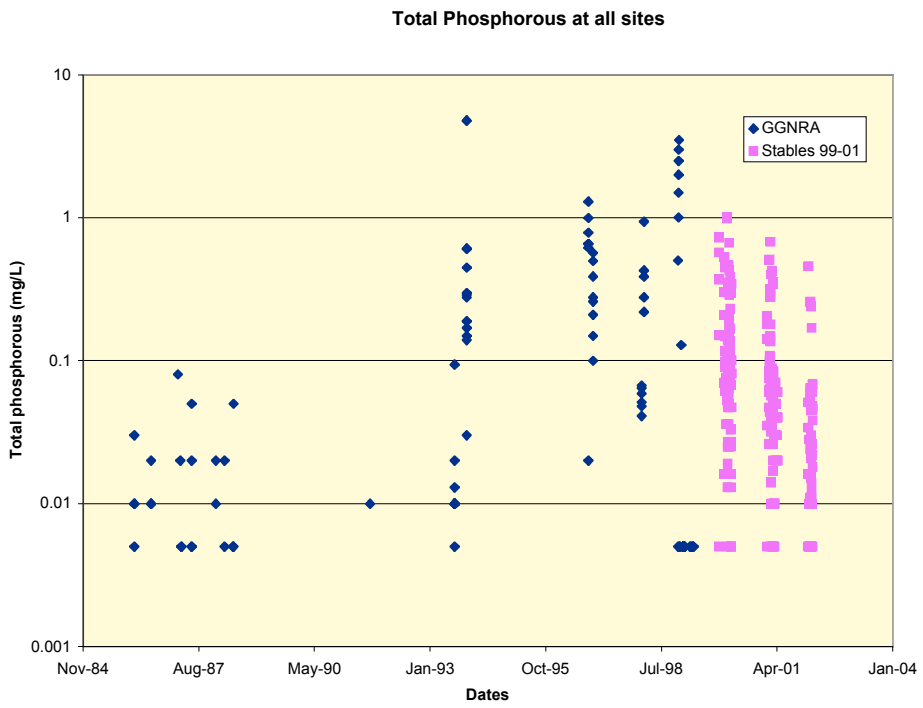


Figure 31. Total Phosphorous over Time in GGNRA

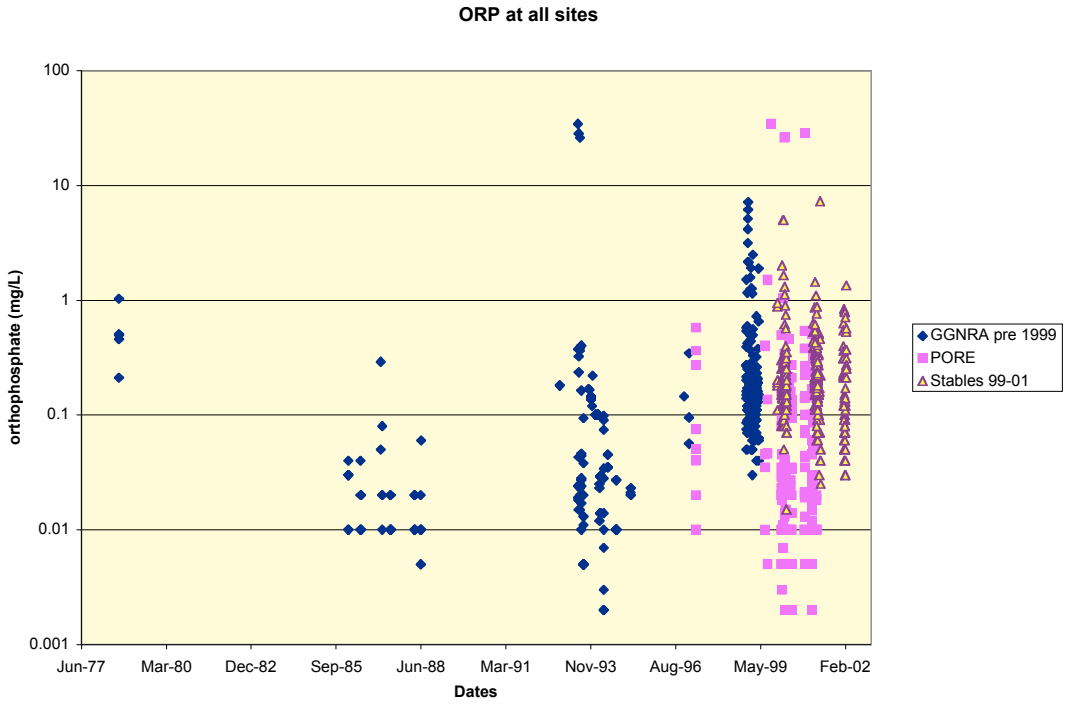


Figure 32. Orthophosphorous over time in GGNRA and PNRS

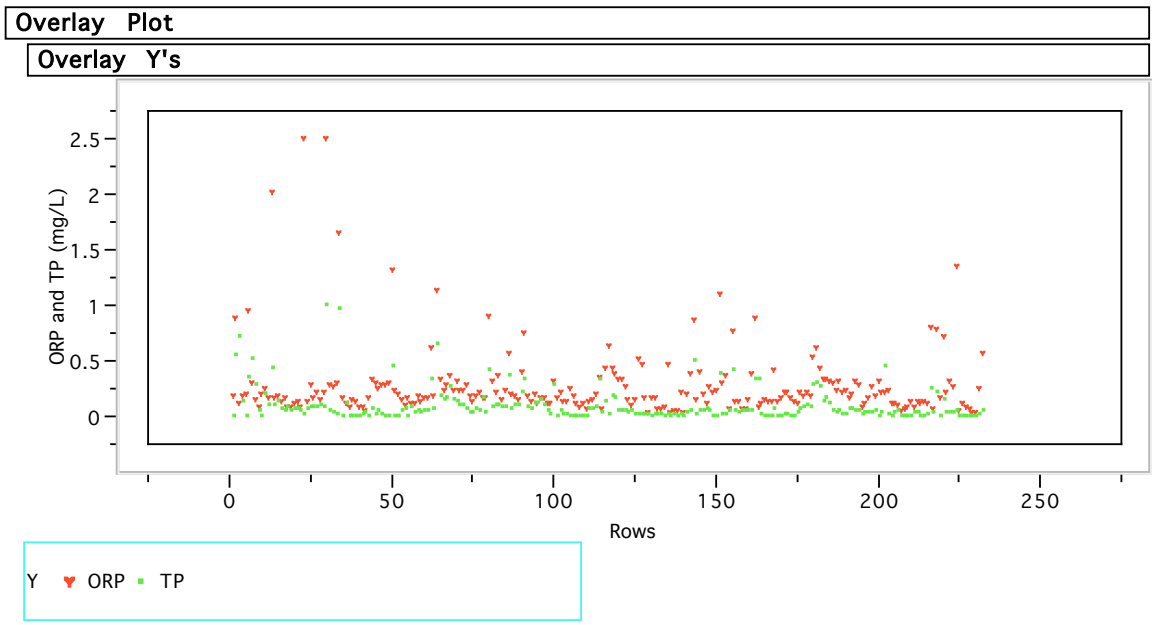


Figure 33. Comparison of ORP and TP from same samples in Stables 1999-2001 data set.

Note that in all instances ORP is greater than or equal to TP. It is not physically possible for ORP to be greater than TP, suggesting that either the data was entered incorrectly or there is a consistent bias in the analytical method for either ORP and TP.

What are recommendations regarding phosphorous monitoring? Evaluate apatite content of total phosphate measurements. If apatite is less than 50% of total P, most systems in the parks are already saturated with P. In which case, discontinue quarterly monitoring for management decisions, and replace with once every two years for long term trend detection. Also, investigate if TP and ORP data were switched in stables based monitoring for GGNRA and PRNS because, as shown in Figure 27, ORP values are consistently higher than TP values.

4.4 BIOLOGICAL INDICATORS

4.4.1 Bacterial indicators, Fecal coliforms (FC), total coliforms (TC) and *E. coli*

- *What are bacterial indicators?* There are four main types of fecal indicator bacteria frequently monitored in surface waters, including total coliforms (TC), fecal coliforms (FC), *Escherichia coli* (*E. coli*), and enterococci. The term *coliform* does not refer to a taxonomic classification but rather to a working definition used to describe a group of Gram-negative, facultative anaerobic rod-shaped bacteria that ferment lactose to produce acid and gas within 48 h at 35°C (Feng et al., 1998). Total coliforms refers to a group of bacteria that includes fecal coliforms as well as many nonfecal bacteria that are

widespread in the environment. All types may be found in human feces, as well as in animal manure, soil and rotting wood, which makes it not very useful for surface water monitoring. It is useful in drinking water monitoring because it is very protective. The water has to be essentially sterile to not contain total coliforms.

The term fecal coliforms refers to a subset of total coliforms more specific to feces, however it also contains one genus *Klebsiella*, with species that are not necessarily fecal in origin (Maier et al., 2000). Fecal coliforms is also a working definition based on the ability to ferment lactose at higher temperatures than total coliforms (Feng et al., 1998). *Klebsiella* are commonly associated with textile, pulp, and paper mill wastes in the absence of fecal contamination (Behar et al., 1996). In addition, some unidentified fecal coliforms genera have been found to grow in the sediments along riverbanks in tropical climates. Nevertheless, fecal coliforms was the primary bacterial indicator of fecal contamination for recreational waters until recently when EPA began promulgating the use of *E. coli* and enterococci as better indicators of health risk from water contact. Fecal coliforms continue to be widely used by many states as the primary fecal indicator bacteria.

E. coli is a species of fecal coliforms that is specific to the fecal material from warm-blooded animals including humans (Maier et al., 2000). It does not grow in the natural environment under ordinary circumstances, and studies have shown a close correlation between high *E. coli* counts and the incidence of digestive track illness at swimming areas (Cabelli, 1977; Dufour and Cabelli, 1984). EPA recommends *E. coli* as the best indicator of health risk from recreation in freshwaters (EPA, 1986; EPA, 2000b). However recent studies have shown that *E. coli* and enterococci are also able to multiply and grow in the algae *Cladophora* in streams and lakes (Byappanahalli et al., 2003), as well as in riparian sediments in tropical and subtropical climates (Desmarais et al., 2002).

Enterococci is another subgroup of fecal coliforms. They are useful as an indicator because they have the ability to survive in saltwater better than other fecal coliforms (Cabelli, 1977; Dufour and Cabelli, 1984). In this respect, they are more similar to many pathogens, and are more specific to human fecal contamination than other fecal coliforms. EPA recommends enterococci as the best indicator of health risk in salt water used for recreation (EPA, 1986; EPA, 2000b).

- *What influences changes in bacterial indicator concentrations?* In general, increased fecal loading will result in higher concentrations of indicator organisms. Fecal contamination may come from ineffective management of human wastes, such as leaking septic systems or untreated wastewater. Fecal contamination also comes from poor management of animal wastes, as well as from manure from dairies and ranches. Low levels of fecal contamination also come from wildlife.
- *Why are bacterial indicators important?* Bacterial indicators are important because they represent fecal pollution, which usually presents an increased health risk to people who come into contact with the water, as well as increased BOD, nitrogen and phosphorous loadings to the stream itself. Bacterial indicators are particularly important in surface waters where people are likely to come in contact with the water, as fecal pollution leads to an increased health risk from waterborne gastrointestinal pathogens. In addition, bacterial indicators are less sensitive to dilution than nitrogen and can be used to trace sources of fecal pollution further downstream from the source. The water quality

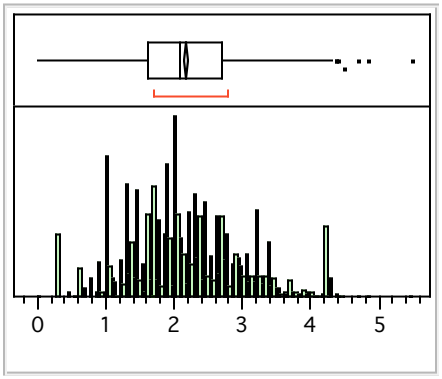
monitoring program at PRNS has identified fecal coliform as a primary indicator of water quality degradation (Ketcham, 2001)

- *How are bacterial indicators measured and in what units are they reported?* Fecal coliforms are distinguished from other bacteria in the lab by their ability to ferment lactose with the production of acid and gas at 44.5 C within 24 hours. However, there are some bacteria which are not fecal in origin which also have this ability and will be reported as fecal coliforms (Maier et al., 2000). Bacterial counts are usually reported as either colony forming units (cfu) per 100 ml or as the most probable number (MPN) per 100 ml. Both cfu and MPN are calculated by doing serial dilutions of the sample until some of the dilutions test negative. It is then assumed that the positive samples of that dilution each contained only one bacterial cell, and the MPN is calculated based on a random distribution (Blodgett, 2003). The term cfu is used because bacterial cells sometimes clump together, so it can't be assumed that each colony or positive result in a dilution comes from a single bacterial cell.
- *How do bacterial indicators relate to other parameters and water quality as a whole?* Bacterial indicators are usually indicative of fecal contamination, which can contribute to low DO, excess BOD, excess ammonia, total nitrogen and phosphorous, as well as pathogenic organisms.
- *What are typical bacterial indicator levels and criteria limits?* The 1995 RWQCB Basin Plan sets the freshwater enterococci geometric mean at 33 cfu/100ml with a single-sample maximum ranging from 61-151 cfu/100ml depending on frequency of usage (Lee and Taylor, 1995). It also sets a saltwater enterococci geometric mean at 35 cfu/100ml with a single-sample maximum ranging from 104 to500 depending on the frequency of usage. The *E. coli* geometric mean is set at 126 cfu/100ml with a single-sample maximum ranging from 235 to 576 cfu/100ml depending on usage. These are set to be protective of public health and not intended to be reflective of ecosystem health. USGS reports that the range of fecal coliforms found in uncontaminated surface waters is from < 1 to 5,000 cfu/100ml and that in fecal-contaminated surface waters, the range is from 200 to > 2,000,000 cfu/100ml. The range of *E. coli* in uncontaminated surface waters was from < 1 to 576 cfu/100ml, and in fecal-contaminated surface waters, 126 to > 2,000,000 cfu/100ml (Myers, 2003).
- *Where have bacterial indicators been measured, for how long and what can we tell from this?* At PRNS, fecal coliforms have been monitored and found useful in pollutant source tracking, since nutrients are so rapidly diluted in streams. In the stables study, grab samples were collected during every sampling event in lab-supplied bottles. Samples were stored on ice in the field and delivered to the lab within 6 hours of collection. Because of past coliform results in excess of upper detection limits, standard dilutions of 10:1 were requested.

Fecal coliforms have been monitored in GGNRA since 1986, with a median of 120 cfu/100ml, and an IQR from 40 to 500 cfu/100ml.

At PINN, coliforms (I'm not sure if this is total or fecal), have been monitored since 1999, with a median of 1600 MPN/100ml and an IQR of 850 to 2240 MPN/100ml. *E. coli* has also been monitored since 1999, with a median of 21 MPN/110ml and an IQR from 6 to 78 MPN/100ml.

log (FC) in GGNRA 1986-2002



Quantiles

100.0%	maximum	5.4771
99.5%		4.3010
97.5%		4.2041
90.0%		3.2553
75.0%	quartile	2.6990
50.0%	median	2.0792
25.0%	quartile	1.6021
10.0%		1.0792
2.5%		0.6021
0.5%		0.3010
0.0%	minimum	0.0000

Moments

Mean	2.1638191
Std Dev	0.8440577
Std Err Mean	0.0145962
upper 95% Mean	2.1924375
lower 95% Mean	2.1352008
N	3344

Figure 34. Distribution of Log Transformed Fecal Coliforms in GGNRA

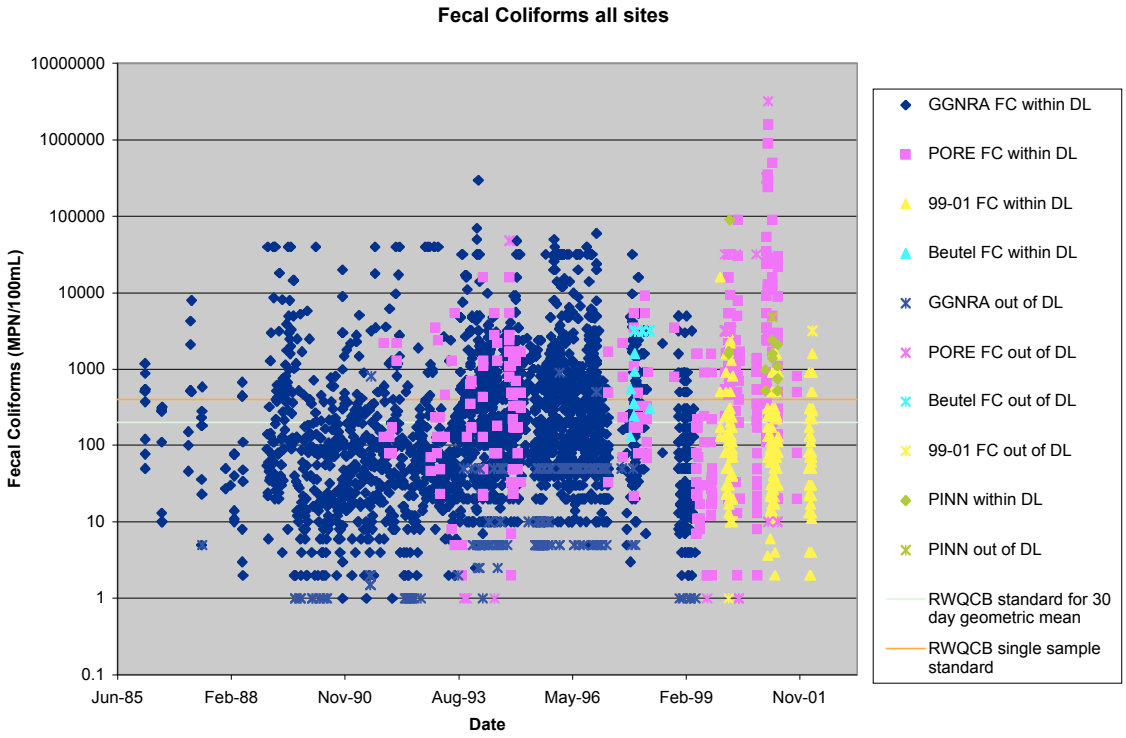


Figure 35. Fecal Coliforms in PRNS, GGNRA and PINN

- *What are recommendations regarding bacterial indicator monitoring?* For source tracking in ambient stream monitoring, continue measuring fecal coliforms. For recreational use monitoring programs, measurements of either enterococci or *E. coli* are recommended, since they are more closely linked to public health risks.

5.0 QUALITY ASSURANCE AND QUALITY CONTROL

Quality assurance (QA) and quality control (QC) are an integral component of any water quality monitoring effort. If the data are inaccurate or biased, then they are not useful in the production of information for management purposes, regardless of quantity. As a result, it makes sense to emphasize data quality over data quantity (Rosenlieb et al., 2002). This is best achieved by identifying the appropriate protocols, having staff well-equipped and well-trained on those protocols, and then following through with the appropriate number of quality control samples (Penoyer, 2003). In general QC samples should comprise 10% to 20% of the total (Rosenlieb et al., 2002).

5.1 FIELD PARAMETERS

Field measurements are intended to represent, as closely as possible, the natural condition of the surface-water system at the time of sampling. To ensure the quality of the data collected, the USGS *National Field Manual* (Wilde and Radtke, 2003) recommends the following:

- *Instrument Calibrations*
- Make field measurements only with calibrated instruments.
- Field instruments should be tagged or labeled with the date of last calibration.
- Each field instruments should have a permanent log book for recording calibrations and repairs. The log book should be reviewed before leaving for the field.
- Each instrument (meters and sensors) should be tested before leaving for the field. Measurement technique should be practiced if the instrument is new to the operator.
- Backup instruments should be readily available and in good working condition.
- *Quality Control Checks*
- Check precision in the field every tenth sample by repeating the measurement three times using separate aliquots from the same sample volume. Note that these should all be recorded as the same sample and only one value included in the data analysis.
- Document ability make accurate measurements by measuring known reference samples.
- *Allow time for sensors to equilibrate with sample before recording measurements.*
- Sensors have equilibrated adequately when instrument readings have stabilized, that is, when the variability among measurements does not exceed an established criterion. For surface waters, allow at least 60 seconds (or follow manufacurer's guidelines) for sensors to equilibrate, then take instrument readings until the stabilization criteria in table xxx are met. Record the median of the final three or more readings as the value to be reported for that measurement point.
- For sites at which variability exceeds the criteria of table xxx: Allow the instrument a longer equilibration time and record more measurements. To determine the value to be reported for that measurement point, use either the median of the final five or more measurements recorded, or apply knowledge of the site and professional judgment to select the most representative of the final readings. Be sure to note how the measurement was selected in the field notes.
- *Use a consistent method to locate the point(s) of measurement*

- Field measurements must accurately represent the body of water being studied. A measurement made from a still pool close to the stream bank may not accurately reflect the conditions in the main flow.
- Field measurements should usually be taken from the same location(s) at which samples are collected for lab measurements.
- Standard USGS procedures can be found in Chapter A4 of the *National Field Manual* (Wilde et al., 1999)
- *Use consistent methods to make the field measurements*
- In situ measurements, made by immersing a field-measurement sensor directly in the water body, are used to determine a profile of variability across a stream section. Measurements made directly in the water body are preferable in order to avoid changes that result from removing a water sample from its source. In situ measurement is necessary for determination of temperature, DO and any chemical properties of anoxic water.
- Depth-integrating and width-integrating sampling methods can be used to collect and composite samples that can then be subsampled for some field measurements. The same field measurements can also be performed on discrete samples collected with thief, bailer, or grab samplers. These samples can yield good data for conductivity, pH, turbidity, and alkalinity as longs correct procedures are followed and the water is not anoxic. Do not measure temperature or DO on subsamples.

6.0 DISCUSSION OF WATER QUALITY ANALYSIS

"*Contrary to the intuitive supposition that the longer the record, the better the understanding of the system, it often can be demonstrated that a briefer, more intensive investigation of hydrology and chemistry of a stream can form a more useful model than one derived from many years of routinely collected data.*" (Hem, 1985)

6.1 THE WATER QUALITY INFORMATION SYSTEM (MODIFIED FROM WARD ET AL., 1990)

A useful way to think about a water quality monitoring network is as a water quality information system (Ward et al., 1990). The purpose of monitoring is to gain information about the state of water quality, which can then be used for management purposes. This view helps to focus the purpose of monitoring and to reduce the common tendency for water quality monitoring programs to be, "data rich but information poor."

Data by itself is unfortunately not information. Samples can be collected, observations made, chemical analyses run, and data entered into tables, but until someone looks at it and interprets it, it is not information. In order for it to become useful, the data must be analyzed and the results made available to water quality managers or whoever needs the information. As sharing of water quality information becomes more common, it becomes important to have consistent methods for not only sample collection and laboratory analysis, but for the analysis of the resulting water quality data as well.

The following diagram of a water quality monitoring system (Ward et al., 1990) illustrates the information flow from the actual water quality in the environment through to a person or agency's accurate understanding of water quality conditions. The first three components, sample handling, laboratory analysis, and data handling, can be

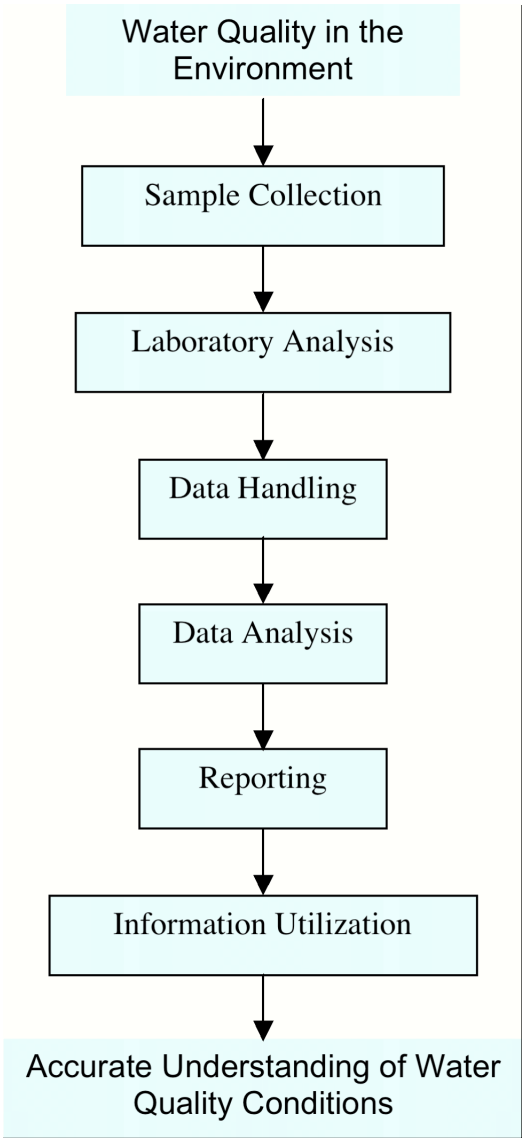


Figure 35. Water Quality Information Flow (Ward et al., 1990)

viewed as the *data generation* part of water quality monitoring. Often these three components are thought of as the entire water quality monitoring effort. However, the second three components, data analysis, reporting, and information utilization, make up the *information generation* part, which is equally important. Note that in this view, the water quality monitoring effort is not complete until the information is used. The focus of the rest of this section will be on information generation.

The most basic information sought in most water quality monitoring programs is "How is the water quality at your park?" Clearly, this question can't be answered by one sample taken from one location at one moment on one day. Water quality is usually highly variable, both temporally and spatially. A sample taken at one place and time simply cannot tell the full story. However, it is also not possible to measure all possible water quality events or sites. In addition, there is no single test to run to tell if the water is "clean" or "dirty." In fact, what constitutes "good" water quality varies depending on they type of water body and the type of ecosystem it supports or the intended use of the water. To get from monitoring data to useful information, inferences must be made from particular samples to general conclusions. Statistical analysis provides an organized, replicable approach to quantify the unavoidable uncertainties in the inferences drawn from water quality data.

Many common statistical methods are based on the assumption that the data under analysis has come from a population[1] with a symmetric, bell-shaped Gaussian (also known as "normal") distribution. Water quality monitoring data, like most

[1] In this discussion, the term population is used to refer to a collection of measurements about which one wants to draw a conclusion (Zar, 1999). It could be used to refer to the concentration of nitrate in all the streams in a watershed, or any other variable with which we are concerned. It should not to be confused with a biological population of organisms.

environmental data, however, often comes from an asymmetric distribution skewed to the right, with the bulk of the data at very low concentrations (on the left) and a long tail stretching out to high concentrations (on the right) (Gilbert, 1987). As a result, the validity of classical statistical procedures applied to water quality data is questionable (Gilbert, 1987; Hirsch et al., 1991; Kirchner, 2003; Zar, 1999). Two other important assumptions in classical statistical methods that are frequently violated by water quality data include the assumption of uncorrelated data and the assumption of constant variance for populations under comparison (Gilbert, 1987). Problems that frequently plague water quality data, regardless of method of analysis, include large measurement errors, data near or below the measurement detection limits, missing and/or suspect data values, complex trends and patterns in concentration levels over time and space, and complicated cause-and-effect relationships (Gilbert, 1987). These problems further hinder straightforward statistical analysis. To overcome these difficulties, the savvy water quality data analyst must make frequent use of graphical methods to display the data, of nonparametric statistical methods to analyze the data, and of techniques to account for serial correlation in time series (Berthouex et al., 1981; Gilbert, 1987; Kirchner, 2003).

6.2 ESTABLISH A DATA ANALYSIS PROTOCOL (DAP)

To make the most effective use of limited time and resources, it is useful to first define, as specifically as possible, the questions to be answered, and then determine what data will be needed to answer the questions. In order to figure out what data needs to be collected, the questions should be translated into a quantitative framework, which includes the desired certainty in the results (statistical power) and the size of an effect worth documenting. Even if strict statistical significance is not of concern, it is still useful to go through the statistical planning process to get an idea of how many samples would be needed to see a given level of effect, and to know how much confidence to place in the results.

As part of the planning process, (Ward et al., 1990) recommend establishing a data analysis protocol (DAP). Having an explicit written document will help insure a basic minimum of consistent data analysis into the future, even if different people carry out the data analysis, and will make it possible for a third party to know what assumptions went in to the analysis and how data irregularities such as censored data, nondetects, missing values and outliers are dealt with. At a minimum, the DAP should include:

1. A statement of the exact information to be produced, for example:
    - Estimates of average conditions (median and IQRs or mean and s.d.)
    - Indications of changing conditions
    - Descriptions of extreme conditions
2. Procedures to be used to prepare a raw data record for graphical and statistical analysis, for example:
    - How to treat non-detects and missing values
    - How to treat values over the quantitation limit
    - How to incorporate quality control replicates
    - Transformations
        - Use log of concentrations?
        - Use calculated toxic fraction of ammonia instead of total

3. Means to visually summarize the behavior of water quality variables
4. Recommended statistical methods which yield the desired information.
5. Reporting formats for the resulting information.

(Ward et al., 1990) provides the following example:

Management goal: maintain or improve water quality
Monitoring goal: detect trends in water quality
Definition of water quality: variables listed in standards
Statistical methodology: kendall tau rank correlation test at 90% with Sen slope estimate
Statistical hypothesis: null – no trend exists, alternate – trend exists.
Data analysis result: conclusions regarding existence of trends in the data with estimates of trend magnitude
Information product: management goal is met when all slopes are zero or indicate improvement.

6.3 How to prepare the raw data record for analysis

6.3.1 Censored and missing data

Water quality records for areas with very low concentrations of measured parameters often have data values that are censored, or reported as less than or equal to the detection limit. In some cases, there are also instances of data values greater than or equal to the upper detection limit, which is the case for both ammonia and Fecal coliforms in some of the park data sets. One method for using this data is to treat it as categorical data rather than continuous data. Another method is to enter in one-half of the detection limit for values under and twice the detection limit for values over. In plotting the parks' data I used a different set of symbols to represent the censored data points, so that they can be seen, but recognized as outside the range of quantitation, even when different detection limits are used. Zeros should not be substituted for non-detects(Kirchner, 2003), since this would bias the results.

6.3.2 Replicates

Replicates from the raw data record should be averaged together and the single mean value used in their place for analysis, or else the median value should be used. The standard deviation or range of the replicates provides an estimate of the variability in the measurement technique.

6.4 Establishing Reference Conditions

EPA's *Technical Guidance Manual for Developing Nutrient Criteria for Lakes and Reservoirs* (EPA, 1999b) describes two ways of establishing a reference

condition. One method is to choose the upper 25th percentile (75th percentile) of a reference population of lakes (or streams). This is the preferred method. The 75th percentile is preferred because it is likely associated with minimally impacted conditions, will be protective of designated uses, and provides management flexibility. When reference lakes are not identified, the second method is to determine the lower 25th percentile of the all lakes within a region to attempt to approximate the upper 25th percentile of a reference population used in the preferred approach. The 25th percentile of the entire population was chosen by EPA to represent a surrogate for an actual reference population because data analyses to date indicate that the lower 25th percentile from an entire population roughly approximates the 75th percentile for a reference population (EPA, 2001a).

6.5 STATISTICAL ISSUES IN THE ANALYSIS OF WATER QUALITY DATA

In deciding the procedures for preparing the raw data record as well as the statistical methods to yield the desired information when looking at trends in water quality, there are four main decisions to make (entire discussion modified from Hirsch et al. 1991):

1. *Parametric vs. Nonparametric methods*
2. *Type of hypothesis to test*
3. *Concentration vs. Flux*
4. *Transformations*

6.5.1 Distributions: Visualizing the data

Regardless of what kind of hypothesis testing is used (if any), or whether parametric or nonparametric methods are used, the first step in understanding water quality data (or any other set of data) is to look at the distribution of the data (Kirchner, 2003; Zar, 1999). For example, to see what is happening with DO at PRNS, one could start by taking the DO data and creating a frequency distribution, or histogram, which simply plots the number of times a particular value of DO has been measured as seen in Figure X. Several questions immediately come up for the data analyst. Since several years of data are available, should they all be plotted together? In addition, data from many sites is available: should it all be plotted together or separately? The answer to these questions, depends on what question or questions the analyst is trying to answer. But even before comparing distributions from different times and places to see if they are the same or different, it is useful to simply look and see what kind of distribution is there. So let's start by assuming that all of the data belongs together. Now we have a distribution as is seen in Figure X. What can we say about it? How could we summarize it? Clearly, it has a particular shape. There is a value, which is found the most often, (the peak or mode), there is an average value (the mean), and there is a value in the middle (the median). These are all considered measures of central tendency, and in a normal or Gaussian distribution, they are all exactly the same value. However, many water quality variables are not normally distributed, and it is necessary to decide which measure of central tendency should be reported. Also important is how far the distribution spreads from that central value (the range). Does it spread equally on either side, as in a normal distribution or does it lean to one side or the other (skew)? Does it quickly trail off, or

does it have long tails (kurtosis)? The standard deviation (Std Dev) or variance is usually used to describe the spread or dispersion of a distribution around the mean. However, the interquartile range (IQR) can also be used to describe the spread of a distribution and is not as sensitive to outliers. The interquartile range encompasses the middle 50% of the distribution. The quartiles are the 25th and 75th quantiles. The 25th quantile gives the sample value that is the highest of the lowest 25% of the sample values. In addition to having an estimate of the central tendency and the dispersion of the distribution, it is also useful to have an estimate of the variability in the estimates of the mean, itself, or how well we know what the mean is. This is estimated by the standard error of the mean (Std Err Mean) and the 95% confidence intervals about the mean.

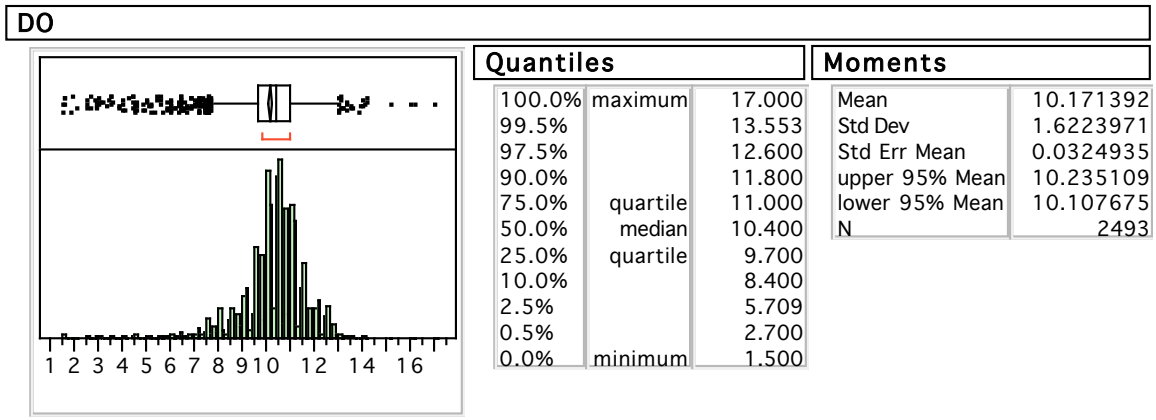


Figure 36. Distribution of dissolved oxygen in GGNRA streams

The heights of the green bars in the histogram represent the number of times an observation was recorded, with the numbers on the left y-axis showing the counts. The outlier box-plot above the histogram shows the interquartile range within the box. The line across the middle of the box identifies the median sample value. The diamond represents the mean and 95% confidence interval. The lines extending from each end of the box, or the whiskers, encompass the quartiles +/- 1.5x(interquartile range). Points beyond the whiskers indicate extreme values that are possible outliers. The red bracket along the edge of the box identifies the shortest half, or the densest 50% of the observations. To the right of the histogram, the quantiles and moments are displayed. The total number of observations is listed as N.

6.5.2 Power, Efficiency, and sample size

Power is the probability of rejecting the null hypothesis (of no trend) given a particular type and magnitude of actual trend [In other words, determine what would be an ecologically significant trend, and set your statistical power so that it will match]

Efficiency is a measure of estimation error. In particular, a procedure's relative efficiency can be measured by the ratio of the mean square errors of two competing procedures. For any given significance level, the most powerful test (for normally distributed residuals) is the parametric test. However, what is at issue in selecting a procedure is not the performance under some ideal set of conditions, (i.e. normality) but the range of performance abilities that occur for the range of distributions likely to exist

in the data to be analyzed. In the case of most water quality variables, they are not normal, and nonparametric methods are more effective.

6.5.3 Parametric vs. Nonparametric tests (Hirsch et al., 1991)

Parametric statistical tests are based on the assumption that the data under analysis comes from a normal distribution, and so can be accurately described by a mean value and the standard deviation around the mean. All other things being equal, parametric tests are more powerful than nonparametric tests when this assumption is met. However, water quality data are commonly skewed (Hirsch et al., 1982), which means that there are more data points on one side of the mean than the other, so it is not normally distributed. For example, in the case of a contaminant concentration, it is physically nonsensical to speak of negative concentrations. The mean may be very close to zero or the detection limit, in which case there will be many more instances of concentrations measured above the mean than below the mean. In cases of substantial departures from normality and large sample sizes, nonparametric procedures are more powerful, because they do not rely on assumptions that are not met by the data. However, it is less clear in cases with either small sample sizes or minor departures from normality. In the case of water quality data analysis it can be difficult to decide whether to use parametric or nonparametric methods because there are so many cases where the departure from normality is small enough to go unnoticed. According to a Monte Carlo analysis, nonparametric methods show a modest advantage in terms of efficiency and power for data sets that depart only slightly from normality. Hirsch et al. conclude that in analysis of multiple data sets, nonparametric methods are preferable to transformations for normality and parametric methods. "The credibility of results is enhanced if a single statistical method is used for all data sets in a study….Use of nonparametric methods avoids both the effort and the potential for real or perceived biases being imparted by the data analyst." However, in an analysis of a single record, parametric methods, including the use of transformations, can be very suitable and they are often more informative than nonparametric procedures in more complex applications. In conclusion, as a general rule of thumb, use nonparametric procedures when combining data from multiple stations, and use parametric procedure when analyzing a consistent data set from a single station.

Nonparametric terms can also be used to describe distributions. For example, instead of using the mean and standard deviation to describe the concentration of nitrate in a stream, the median and interquartile range (IQR) can be used. The median is simply the data point in the middle, with an equal number of data points above and below it. The median is less sensitive to outliers than the mean is (Zar, 1999). The IQR is defined as the difference between the 75th percentile and the 25th percentile, or the range including the middle 50% of data points.

6.5.4 Hypothesis Testing (Hirsch et al., 1991)

There are two options for hypothesis testing in trend detection. One is to look for a step-trend, and the other is to look for a monotonic trend. In most cases of general monitoring, the monotonic trend hypothesis is the default unless there is a reason to test for a step trend. A monotonic trend is simply a unidirectional change over time, which could be continuous or contain one or more discrete steps. It is definitely preferred if

records from multiple stations are to be compared. A step trend is testing the hypothesis that there are two distinct populations before and after, i.e., that there is a discontinuity in the record. It requires prior knowledge of the date of change and would be appropriate, for example, for a site before and after a change in management practices. It can also be used if for some reason the data record is naturally broken into multiple periods with a relatively long gap between them. (Hirsch et al., 1991) recommend that as a practical guideline, if the length of the gap is more than ~1/3rd of the entire data collection, a step trend test is preferable, even if the actual trend is linear. The decision to use a step trend test must be made before examining the actual data to avoid violating the assumptions going into the test.

When looking at water quality data, it is important to ask not just if there is a trend, but also how big the trend is. This is called the estimator of trend magnitude. For a monotonic trend, the estimator of trend magnitude is the regression slope. The parametric (see section 6.5.3 for discussion of parametric vs. nonparametric tests) test for a monotonic trend is regression analysis. The nonparametric test is the Mann-Kendall test for trend with the *Sen* slope estimator. For a step trend, the estimator of trend magnitude is the difference in the means before and after. Either the parametric two-sample t test or the nonparametric Mann-Whitney-Wilcoxon Rank Sum test with the associated Hodges-Lehmann estimator of trend magnitude can be used to test for a step trend.

A significant issue that comes up when analyzing historical data sets is deciding which time periods and records to include in the analysis. In multiple record studies, monotonic trend procedures are most appropriate – even if some of the records have extensive but not identical gaps, because comparable time periods can be more easily compared between all the records. The frequent problem of multiple starting dates, ending dates, and gaps in groups of records presents a significant problem in trend analysis studies. In order to correctly interpret the data, records examined in a multiple station study must be concurrent. The difficulty arises in selecting a period that is long enough to be meaningful but doesn't exclude too many shorter records. A further difficulty arises in deciding just how complete a record must be to include it in the analysis. One reasonable objective rule for deciding to include a record would be as follows: (1) divide the study period into three periods of equal length, (2) determine the coverage in each period (e.g., if the record is generally monthly, count the months for which there are data), (3) if any of the thirds has less than 20% of the total coverage then the record should not be included in the analysis.

Statistical significance and ecological importance (modified from (Griffith et al., 2001))

It is important to realize that when using common null hypothesis tests, obtaining a *statistically* significant result does not necessarily imply that one has found an *environmentally* significant result. A null hypothesis of no-difference between two populations does not specify how big the difference needs to be. However, in some cases, there may be a statistically significant difference, but the size or magnitude of the difference is so small that it is ecologically insignificant. On the other hand, there may be a large change with a significant impact on the ecosystem, but if the data are too noisy, then no statistically significant difference will be detected. In most cases the magnitude of the difference demonstrated usually has more to do with how environmentally significant or important the result is. Statistical significance, as measured by a p-value

simply says how likely it is that the demonstrated difference is not really a difference, but an artifact of sampling from a distribution with high variability. Interval testing, or the use of a null hypothesis that specifies the magnitude of difference that we care about, is a more useful alternative. "An advantage of this test is that the analyst must state the difference of practical significance to management, also the failure to reject the null no longer induces complacency." (Griffith et al., 2001)This is because the results now mean something, ecologically and environmentally.

Pitfalls of hypothesis testing (Griffith et al., 2001) (McBride et al., 1993) Null hypothesis testing has three problems that are applicable in environmental monitoring:

1. A conclusion that there is a significant result can often be reached merely by collecting enough samples (increasing sample size increases chance of rejecting the null);
2. A statistically significant result is not necessarily practically significant; and
3. Reports of the presence or absence of significant differences for multiple tests are not comparable unless identical sample sizes are used.

To avoid these problems, it is important to evaluate the power of the test, as well as the Least Significant Number, and the Minimum Detectable Difference (Zar, 1999). The Least Significant Number is the number of measurements that need to be included to detect a given difference at a given significance level (Kirchner, 2003). Power measures how reliably you can detect an effect or difference of a given size. The Minimum Detectable Difference is the magnitude of the effect or difference that would be consistently detected given a certain sample size and significance level (Kirchner, 2003). These can all be calculated in the planning stage to determine the frequency and number of measurements to be made.

6.5.5 Serial Correlation and sampling frequency (modified from (Kirchner, 2003)

Serial correlation, also called autocorrelation, occurs when the residuals, or "errors" from a time series are correlated with each other instead of being random noise. Temporal dependence, other than that caused by seasonality, is a kind of serial correlation, or redundancy in information from one observation to the next, and violates the assumption of independence inherent in most statistical tests. Environmental data can also be spatially autocorrelated, where measurements at nearby locations have correlated residuals. Serial correlation is very common in environmental time series and may be caused by any of the following (Kirchner, 2003): 1) the variable of interest has a long term trend which has been overlooked in the analysis, 2) the variable of interest varies seasonally, and this seasonal effect has been overlooked in the analysis, 3) the variable under study responds to a missing explanatory variable which is serially correlated, or 4) the variable of interest includes random noise that is serially correlated, or that has persistent effects, which is particularly common if the variable integrates or averages some other variable through time. For an example of this last cause, consider the case of chemical inputs to a lake. Changes in the chemical inputs to a lake will persist for timescales comparable to the lake's flushing time. Whenever the sampling period is shorter than the timescales on which the system responds to external forcing, many samples will effectively be redundant. As long as we recognize that these samples are

redundant, it is not a problem. However, nearly all statistical techniques assume that any random errors are independent. Instead, in the case of serial correlation, each error (or residual) depends on the previous residual. If this is not explicitly accounted for, then spurious trends are likely to be found statistically significant, and even if the correct trends are found, the uncertainties associated with the regression slope are severely underestimated. As a result, estimates of "goodness of fit" and statistical significance will be exaggerated. Unlike most statistical problems, the deceptive effects of serial correlation become stronger instead of weaker as more data is included.

(Ward et al., 1990) recommend reducing sampling frequency to once a quarter, unless looking for violations, to reduce serial correlation. The drawback to this approach is that while some explanatory variables may be serially correlated, there may be other variables of interest, which change on a shorter time scale, and for which it would be useful to have information. If using the same data for long-term evaluation and short-term management, it may be useful to average the measured values over each quarter, so that there is just one value per quarter. – this may also be useful in analyzing such large data sets with varying sampling frequencies.

Kirchner (2003) suggests several alternate methods for dealing with serial correlation in environmental data without reducing sampling frequency. In many cases, serial correlation arises because a time-dependent effect or an explanatory variable has been excluded from the analysis. Clearly then, the first step is to look and see if this is the case, and if possible include these variables in the analysis. If the serial correlation cannot be explained by other explanatory variables, then there are three basic ways of accounting for it in the statistical analysis: 1) subsample the data set, 2) average the data over longer periods, or 3) compute the regression taking into account autocorrelated residuals. The first two options are conceptually simple, but lose much of the information that has been so painstakingly gathered. The third option is conceptually more difficult, but retains as much information as possible and is not operationally difficult given modern computing capabilities. The Hildreth-Lu procedure (Kirchner, 2003; Neter et al., 1990) can be carried out with some training on a statistical software analysis program such as JMP.

To reduce the variability due to seasonality, analyze data from each season separately.

6.5.6 Transformations and other data manipulations to reduce variance (Hirsch et al., 1991)

- *Transformations*

Nonparametric statistical tests are invariant to monotonic transformations (such as the logarithm or square root). This means that in terms of significance levels the test results will be identical whether the test is applied to the raw data or the transformed data. Parametric statistics will, however, change when applied to transformed data, thus the decision to transform data, is important in terms of fitting various models that are useful in trend detection and analysis such as flow adjustment, or for computing significance levels of a parametric test. As an example, figure xxx.a shows flow in GGNRA plotted by month, and figure xxx.b shows the log transform of flow plotted by month.

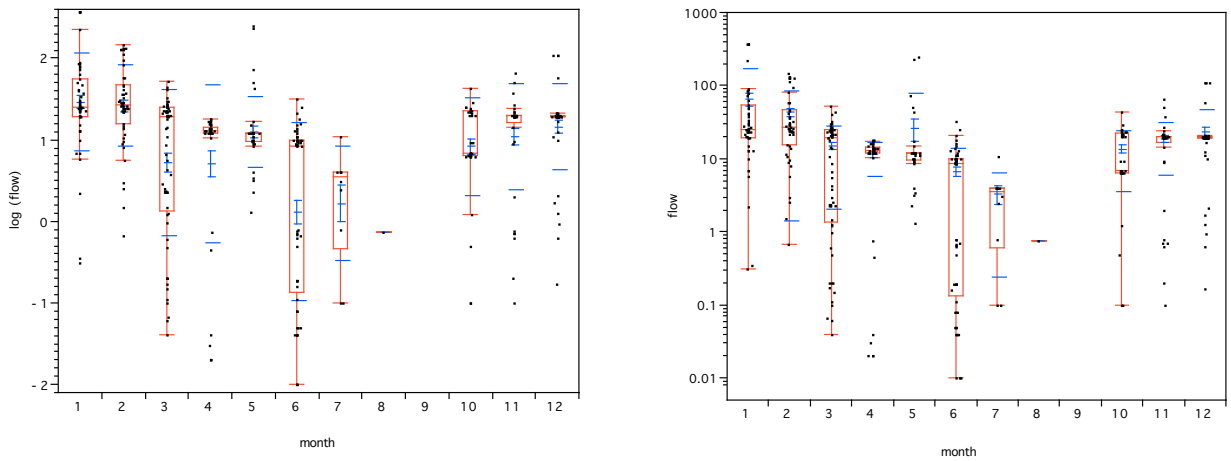


Figure 37. Log Transformed and Untransformed Flow by Month at GGNRA

The graph on the left is a log transform of flow by month for all the GGNRA data in the Access database. On the right is the same flow data plotted on a log axis, but not transformed. The medians and quantiles for each month are indicated by the red boxes and whiskers. The line inside of each box represents the median. The ends of the boxes are the 25th and 75th quantiles, so the box itself encompasses the interquartile range. The whiskers extending beyond the box represent the upper and lower quartiles + 1.5*(interquartile range). The monthly means and standard deviations are indicated by the blue lines. The center blue line is the mean. The two inner blue lines are the confidence interval for the mean and the outer two blue lines are one standard deviation. In comparing these two analyses, note that the axes are the same, so that medians and means can be visually compared. The median and quantiles (in red) are robust, and stay the same if log-transformed back to the original units, while the mean and standard deviation (in blue) are sensitive to outliers, and do change. Parametric tests rely on the use of the mean and standard deviation and so will give different results for transformed data than for untransformed data. Don't be distracted by the red whiskers; although they are based on the interquartile range, they do not stay the same, since they are multiplied by the factor 1.5.

It is often useful to express the history of a trend as a single linear equation. This is particularly true in the context of a multiple station trend analysis where the trend slopes are to be compared. If the actual trend is nonlinear (say exponential or quadratic), it is common for a linear trend fit to the data to predict negative values during some part of the period of record, which is clearly not a reasonable approximation of a long-term trend. This can be prevented by taking the log transformation of the data and using that for trend analysis. To use the log transformation does not have to mean that the trend is exponential, it simply provides an exponential trend approximation to the actual trend in the data. To make these trend slopes more interpretable, these log concentration slopes can be expressed in percent per year. If b is the estimated slope of a linear trend in natural log units, then the percent change from the beginning of the year to the of the year will be (e^b-1)*100. If slopes or step sizes in original concentrations are preferred, then

one can measure by some measure of central tendency in the data (mean or median) to express the step or slope in original units.

More robust and resistant results can be obtained if log transformations are used for data that typically have ranges of more than an order of magnitude at a given station. Transformations have been used with parametric and nonparametric tests when the range of variation is quite large. However, in multiple record analyses, the decision to transform should be based on the characteristics of the class of variables being studied and not on a case-by-case basis. Variables on which log transforms should typically be taken include: sediment concentrations, total concentration (suspended + dissolved) when suspended is substantial e.g., phosphorous and some metals; concentrations or counts of organisms; concentrations of substances that arise from biological processes (such as chlorophyll); and flux for virtually any constituent.

- *Removal of Variance due to discharge*

In many cases, a great deal of the variance in water quality variables, including both flux and concentration, is a function of river discharge. The power and efficiency of both parametric and nonparametric procedures will be enhanced if the variance of the data can be decreased. This is done by analyzing flow-adjusted concentrations (FAC), or concentrations with the flow effects removed. In the parametric case, this is done by using discharge as a covariate. In the nonparametric case, the variation due to flow is modeled by either regression or a robust curve-fitting procedure such as LOWESS (locally weighted scatter plot smoothing). Then the trend analysis is conducted on the residuals from this relationship, or the FAC.

- *Removal of Seasonal Variability*

The other major source of variability in water quality data is usually seasonal variation. Seasonal variation may be due to many different factors, including: yearly fertilizer and pesticide applications, differences in flow sources, such as rainwater with a high sediment load versus groundwater, connectivity of small tributaries in a watershed during the dry season, and seasonal temperature variations. Some of the variation that may initially be viewed as seasonal may in fact be statistically explained by variation in discharge. However in many cases, seasonality remains in the data set even after accounting for flow effects (Hirsch et al., 1982). Whether or not flow effects have been removed, it is a good idea to limit the seasonal variation in the data. This can be done by in parametric procedures by using trigonometric functions of times of the year as an explanatory variable. For nonparametric procedures, this can be done by performing the test on each of the different seasons, summing the test statistics and summing their expectations and variances. The overall test for trend is then carried out by using the summed test statistic with its expectation and variance. Both the Seasonal Kendall test (Hirsch et al.. 1982) and the Rank Sum test on grouped data (Bradley, 1968) make use of this approach. Software packages such as WQStat Plus are available that will perform both the seasonality adjustments and the flow adjustments easily, and would be a good investment to save time in data analysis for the parks.

6.5.7 Concentration vs. Flux (Hirsch et al., 1991)

Determination of which variable (concentration or flux) should be analyzed for trend depends on the question to be answered. For example, if the question is one of ambient quality in the stream, then concentration would be the appropriate variable to evaluate. The exposure of organisms that reside in the stream to potentially harmful chemicals is determined by concentrations and the time over which they persist. Flux is of no concern in this example. However, if storage of the water and its constituents is an important factor, then flux may be the more important variable to analyze. For example, the flux of relatively conservative components may be of interest in situations where the sampling site is upstream of a reservoir, lake or estuary where the water has a long residence time. In addition, studies focused on mass balances (changes in the sources and sinks of chemical species in a watershed) should also lead to analyses of flux. It may be appropriate to evaluate trends in both concentration and flux if there are multiple objectives for the study. Knowing the trends in one of these measures will not necessarily provide a clear indication of the trends one can expect in the other measure. For example, one may find a general upward trend in concentrations occurring at low discharges, but at high flows, the trends are either nonexistent or so small that they are obscured by the high variability of the high-flow conditions.

6.5.8 Trend Detection (modified from (Hirsch et al., 1991))

Tables 1 through 4 summarize recommendations for monotonic and step trend detection, depending on the type of data under analysis. Monotonic trends are to be used for gradual changes, and step trends are to be used before and after a change at a specific point in time. Censored data includes data that is either above or below the detection limit or for some other reason has to be changed from what is initially reported. Uncensored data includes all the data, without any gaps, omitted values, or below detection limits. The parameters classified as "mixed" in the first two tables have both parametric and nonparametric components that are typically executed in separate steps.

Regression on season uses a periodic function of time of year, as does Tobit regression on season. Tobit regression is a type of linear regression that considers both censored and non-censored values of the response variable, and uses maximum-likelihood estimation for determining slope and intercept of the modeled trend line (Hoppe, 2003).

Deseasonalizing is done by subtracting seasonal medians from each of the values to be regressed.

The Seasonal Kendall test is the Mann-Kendall test for trend done for each season, with the Seasonal Kendall test statistic being the sum of the several Mann-Kendall test statistics. The seasonal Kendall trend test accounts for seasonal variations in concentrations by comparing ranks of data from the same recurring time intervals; for example, in a four-season year, springtime values are compared only to other springtime values, summer values to summer values, and so forth.

LOWESS is locally weighted scatterplot smoothing. The LOWESS curve represents a nonlinear, smoothed relation between two variables (instantaneous discharge and each water quality parameter). The method uses a series of weighted least squares

regressions; observations are weighted by both distance from the fitted line and the magnitude of residuals from the previous regression. LOWESS is more desirable than simple regression because it makes no assumptions of data linearity or normality (Hoppe, 2003). Flow may be replaced by a transformation of flow in any of these analyses.

The Seasonal Rank Sum test is the Rank Sum test (also known as the Mann-Whitney "U" test (Kirchner, 2003), done for each season, with the Seasonal Rank Sum test statistic being the sum of the several test statistics.

Table 1. Options for testing monotonic trends in uncensored water quality data

Options for testing monotonic trends in uncensored water quality data		
	Not Flow Adjusted	Flow Adjusted
Fully parametric	Regression of concentration on time and season	Regression of concentration on time, season and flow
Mixed	regression of deseasonalized concentration on time	Seasonal Kendall on residuals from regression of concentration on flow
Nonparametric	Seasonal Kendall	Seasonal Kendall on residuals from LOWESS of concentration of concentration on flow.

Table 1. Step trends in uncensored water quality data

Options for testing step trends in uncensored water quality data		
	Not Flow Adjusted	Flow Adjusted
Fully parametric	Analysis of covariance of concentration on season and group (before and after)	Analysis of covariance concentration on season, flow and group
Mixed	Two-sample t test on deseasonalized concentration	Seasonal Rank Sum on residuals from regression of concentration on flow
Nonparametric	Seasonal Rank Sum	Seasonal Rank Sum on residuals from LOWESS of concentration on flow

Table 2. Monotonic trends in censored water quality data

Options for testing for monotonic trends in censored water quality data		
	Not Flow Adjusted	Flow Adjusted
Fully parametric	TOBIT regression of concentration on time and season	TOBIT regression of concentration on time, season and flow
Nonparametric	Seasonal Kendall	no test available

Table 3. Step trends in censored data

Options for testing for step trends in censored water quality data		
	Not Flow Adjusted	Flow Adjusted
Fully parametric	TOBIT analysis of covariance of concentration on season and group	TOBIT analysis of variance of concentration on season, flow and group
Nonparametric	Seasonal Rank Sum	no test available

7.0 MONITORING CRITIQUE

1. PRNS is setting an excellent example for others to follow
    a. Document WQ monitoring plan including parameters to be measured, sites to be measured, how often and why, as well as table of actions taken based on monitoring results. GGNRA has some of this, but could use improvements. Pinnacles hasn't started any of this yet.
    b. Maybe better to have more thorough every two year report than superficial annual report?
2. Have master file including other agencies monitoring in park watersheds, i.e., one place to find all info for a particular area.
3. Metadata needs, rainfall, hydrologic information, watershed uses. Sampling objectives.

4. Specific issues encountered in data analysis
1. Data needs to be all in one database with consistent site names and units
    a. Sites Table
        i. Column including upstream and surrounding land usage in site locations table
        ii. Column for site observations (with observations in it, even if just to say, looks normal)
        iii. Column indicating control or not
        iv. Column indicating intermittent or not
        v. Column indicating type of water body, e.g., stream mainstem, tributary, pond, lagoon, lake.
        vi. For streams column indicating bankful flow
    b. Data entry/ Monitoring
        i. include column for source of data and any reports utilizing it.
2. Data Analysis
        i. Include annual data analysis/report writing in WQ monitoring plan, data is useless without conversion into information.
            1. Maintain Excel Charts (or some other graphic/visual display) where new data can be included to quickly and easily see how it compares with past levels.

2. Purchase a statistical program such as WQStatPlus or JMP that will facilitate seasonally and flow-adjusted data analysis for the detection of long-term trends. These programs will also do a much better job than Excel of displaying data and linking graphics to the raw data.

- Consolidate site locations tables for GGNRA and PRNS
- Pinnacles
- Create table of site locations for Pinnacles with UTMN and UTME coordinates and site descriptions, rationales
- Create table or document describing analytical methods used for parameters monitored at Pinnacles

Data Reporting Recommendations:

PRNS queries are great, but FLOW should be included. It is included in Field Parameters report, but not in report by season, report by year, report by station, or report by watershed.

Several authors (Kirchner, 2003; McBean and Rovers, 1998; Zar, 1999) recommend using median and interquartile ranges to describe the central tendency and spread of the data, rather than (or in addition to) mean and standard deviations, since most water quality data are not normally distributed. Notice that most of the data are plotted with a base 10 log scale. The median and interquartile range remain the same even when the data is transformed, but the mean and standard deviation change. The difficulty with maximum and minimum values is that as the data set grows, the most extreme values just keep extending and don't necessarily reflect common conditions. They are useful for immediate management decisions, since they often indicate something is wrong, but are not as useful for an overall sense of the water quality conditions.

8.0 SUMMARY OF RECOMMENDATIONS

- All Parks
  3. Data all in one database with consistent site names and units
     a. Sites Table (See Appendix C for an example) (Put in Section 6 – Monitoring Critique)
        i. Column including upstream and surrounding land usage in site locations table
        ii. Column for site observations (with observations in it, even if just to say, looks normal)
        iii. Column indicating control or not
        iv. Column indicating intermittent or not
        v. Column indicating type of water body, e.g., stream mainstem, tributary, pond, lagoon, lake.
        vi. For streams column indicating bankful flow

b. Data entry/ Monitoring table (See Appendix C for an example) (Also section 6)
    i. include column for source of data and any reports utilizing it.
c. Specific Parameters (Section 4, Parameters, except MBAS?)
    i. Flow
        1. record at every sampling event.
        2. Where will not cause damage to stream bed, install staff gauges or stilling wells.
        3. Where possible, include bankful flow data in site locations chart.
        4. Conduct storm surveys - Analyze flow and contaminant data over storms of different sizes to determine when the biggest contaminant load is and if appropriate to divert some water for treatment before release.
        5. For most important watersheds in each park, collect background information such as is found in "Detailed Analysis of Olema Creek Watershed Response"
    ii. Temperature
        1. Record at every sampling event
        2. Calibrate YSI meters at least once a year
        3. Conduct one time special temporal and spatial monitoring over course of a year. Extensive use of data-loggers is recommended with each water body being characterized in detail over a year.
    iii. pH
        1. record at every sampling event
        2. calibrate pH meter in lab daily before collecting field samples.
        3. if unusual value is measured, look around site and note conditions
    iv. Conductivity, salinity and TDS
        1. Continue conductivity monitoring
        2. categorize sites according to seawater influences
    v. TSS
        1. Acquire a field turbidity meter measuring NTUs (~800-12000) and use as routine monitoring tool.
        2. Calibration of turbidity meters depends on range. Comparison of turbidities over more than a ten-fold range should only be done with great care.
        3. Field unit measurements should be compared with lab measurements.
        4. historical data should be examined for the typical annual history of water clarity. If not available from historical data, then a one-year project to establish a normal season should be carried out. This will provide a baseline with

which to compare the results of less intensive future monitoring.

vi. DO
    1. Continue collecting DO measurements at all sampling events.
    2. DO probe should be calibrated daily, following manufacturer's instructions.
    3. the probe membrane should be replaced every few days, or if any bubbles are present.
    4. the probe should never be allowed to penetrate the sediments, especially when DO is low. If the probe does accidentally hit anoxic sediments it should be allowed to reequilibrate at least a minute before readings are resumed.

vii. BOD
    1. BOD levels above 2 mg/L have been reported in a few sites and were due to pollution. Until a few years have gone by without such high BOD events, monitoring should be continued. A consideration to actual BOD should be given if models of the streams are to be used with regard to overall stream health (see earlier discussion on kinds of BOD test).

viii. Nitrogen -
    1. Nitrate and ammonia should continue to be measured on a regular basis
    2. nitrite monitoring should be discontinued
    3. Ammonia samples should be split in two, and one sample diluted before analysis, since many samples were over the detection limit.
    4. Also include TN every two years for long term trend detection
    5. At PINN, field N measurements should be done more carefully or else eliminated since they are not reliable. QA/QC depends on what method is being used.

ix. Phosphorous
    1. Evaluate apatite content of total phosphate measurements, if less than 50%, most systems already saturated with P, discontinue quarterly monitoring for management decisions, replace with once every two years for long term trend detection.
    2. Also, investigate if TP and ORP data were switched because ORP values are consistently higher than TP values.

x. Bacterial indicators
    1. Switch recreational monitoring to enterococci rather than fecal coliforms?
    2. Continue fecal coliforms in general monitoring

xi. Drop MBAS monitoring – too general, difficult to interpret false positives, unless suspected contamination.

xii. Continue/implement BOD monitoring

4. Data Analysis
   i. Include annual data analysis/report writing in WQ monitoring plan, data is useless without conversion into information.
      1. Maintain Excel Charts (or some other graphic/visual display) where new data can be included to quickly and easily see how it compares with past levels.
      2. Purchase a statistical program such as WQStatPlus or JMP that will facilitate seasonally and flow-adjusted data analysis for the detection of long-term trends.
   ii. Analyze all parameters according to site type: dairy, wilderness, etc.

- PRNS and GGNRA
- Consolidate site locations tables for GGNRA and PRNS
- Pinnacles
- Create table of site locations for Pinnacles with UTMN and UTME coordinates and site descriptions, rationales
- Create table or document describing analytical methods used for parameters monitored at Pinnacles

9.0 BIBLIOGRAPHY

American Public Health Association, American Water Works Association, and Water Environment Federation (1992): Standard methods for the examination of water and wastewater (18th ed.), pp. p. 2-8 to 2-11, American Public Health Association, Washington, D.C.

Behar, S., Dates, G., and Byrne, J. (1996): *Testing the waters: chemical and physical vital signs of a river*. River Watch Network. Montpelier, VT.

Berthouex, P. M., Hunter, W. G., and Pallesen, L. C. (1981): Wastewater treatment: A review of statistical applications. Environmetrics 81: Selected Papers. SIAM-SIMS Conference Series No. 8, pp. 77-99.

Bilby, R. E. (1984): Characteristics and frequency of cool-water areas in a western Washington stream. *Journal of Freshwater Ecology* **2**, 593-602.

Bjornn, T. C., and Reiser, D. W. (1991): Habitat requirements of salmonids in streams. *Influences of Forest and Rangeland Management on Salmonid Fishes and Their Habitats. American Fisheries Society Spec. Pub.* **19**, 83-138.

Blodgett, R. (2003): *Bacteriological Analytical Manual Online Appendix 2: Most Probable Number from Serial Dilutions*. U.S. Department of Health and Human Services, Center for Food Safety and Applied Nutrition. Accessed 3/10/2004 at

Bradley, J. V. (1968): *Distribution-Free Statistical Tests*. Prentice Hall. Englewood Cliffs, N.J.

Brett, J. R. (1952): Temperature tolerance in young Pacific salmon, Genus Oncorhynchus. *J. Fish. Res. Bc. Can.* **9**, 265-323.

Byappanahalli, M. N., Shively, D. A., Nevers, M. B., Sadowsky, M. J., and Whitman, R. L. (2003): Growth and survival of Escherichia coli and enterococci populations in the macro-alga Cladophora (Chlorophyta). *Fems Microbiology Ecology* **46**, 203-211.

Cabelli, V. J. (1977): Indicators of recreational water quality: Philadelphia, American Society for Testing and Materials, ASTM STP 635, 222-238 p.

Cooprider, M. (2004): San Francisco Area Network Water Quality Planning Meetings Summary (Draft): San Francisco, National Park Service, San Francisco Area Network, 67 p.

Cordone, A. J., and Kelley, D. W. (1961): The influences of inorganic sediment on the aquatic life of streams. *California Fish and Game* **47**, 189-228.

Coutant, C. C. (1977): Compilation of Temperature Preference Data. *Journal of the Fisheries Research Board of Canada* **34**, 739-745.

Creek Connections (2003): *Nitrogen Information Sheet*. Creek Connections, Allegheny College. Accessed March 25, 2004 at http://creekconnections.allegheny.edu/Chemistry/NitrateMain.html

Creek Connections (2004a): *Dissolve Oxygen Information Sheet*. Creek Connections, Allegheny College. Accessed March 25, 2004 at http://creekconnections.allegheny.edu/Chemistry/DissolvedOxygenMain.html

Creek Connections (2004b): *Phosphorous Information Sheet*. Creek Connections, Allegheny College. Accessed March 25, 2004 at http://creekconnections.allegheny.edu/Chemistry/PhosphorousMain.html

Creek Connections (2004c): *Total Dissolved Solids Information Sheet*. Creek Connections: Allegheny College. Accessed March 25, 2004 at http://creekconnections.allegheny.edu/Chemistry/TDSMain.html

Creek Connections (2004d): *Turbidity Information Sheet*. Creek Connections: Allegheny College. Accessed March 25, 2004 at http://creekconnections.allegheny.edu/Chemistry/TurbidityMain.html

Delzer, G. C., and McKenzie, S. W. (2003): *Five-day biochemical oxygen demand: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A7, section 7.0*. Accessed March 22, 2004 at http://pubs.water.usgs.gov/twri9A/

Desmarais, T. R., Solo-Gabriele, H. M., and Palmer, C. J. (2002): Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. *Applied and Environmental Microbiology* **68**, 1165-1172.

Dufour, A. P., and Cabelli, V. J. (1984): Health effects criteria for fresh recreational waters: Cincinnati, Ohio, U.S. Environmental Protection Agency, EPA 600/1-84-004, 33 p.

EPA, U. S. (1986): Ambient water quality criteria for bacteria - 1986: Washington, D.C., Office of Water Regulations and Standards Division, EPA 440/5-84-002, 15 p.

EPA, U. S. (1999a): 1999 Update of Ambient Water Quality Criteria for Ammonia: Washington, DC, United States Environmental Protection Agency, Office of Water, EPA-822-R-99-014, p.

EPA, U. S. (1999b): Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs (Draft), United States Environmental Protection Agency, Office of Water, EPA 822-D-99-001, p.

EPA, U. S. (1999c): Nutrient Criteria Technical Guidance Manual: Rivers and Streams (Draft), United States Environmental Protection Agency, Office of Water, EPA 822-D-99-003, p.

EPA, U. S. (2000a): Ambient Water Quality Criteria Recommendations: Information Supporting the Development of State and Tribal Nutrient Criteria: Rivers and Streams in Nutrient Ecoregion III, United States Environmental Protection Agency, Office of Water, EPA 822-B-00-016, p.

EPA, U. S. (2000b): Improved enumeration methods for the recreational water quality indicators: Enterococci and *Escherichia coli*: Washington, D.C., Office of Science and Technology, EPA/821/R-97/004, 49 p.

EPA, U. S. (2001a): Ambient Water Quality Criteria Recommendations: Information Supporting the Development of State and Tribal Nutrient Criteria: Lakes and Reservoirs in Nutrient Ecoregion III: Washington, DC, United States Environmental Protection Agency, Office of Water, EPA 822-B-01-008, p.

EPA, U. S. (2001b): Nutrient Criteria Technical Guidance Manual: Estuarine and Coastal Marine Waters: Washington, DC, United States Environmental Protection Agency, Office of Water, EPA-822-B-01-003, p.

Essig, D. A. (1998): The Dilemma of Applying Uniform Criteria in a Diverse Environment: An Issue Analysis: Boise, Idaho, Idaho Division of Environmental Quality, Water Quality Assessment and Standards Bureau, 34 p.

Feng, P., Weagant, S. D., and Grant, M. A. (1998): *Bacteriological Analytical Manual Online*. U.S. Food & Drug Administration Center for Food Safety and Applied Nutrition. Accessed 2/13/2004 at http://vm.cfsan.fda.gov/~ebam/bam-4.html

Gilbert, R. O. (1987): *Statistical methods for environmental pollution monitoring*. Van Nostrand Reinhold Company, Inc. New York.

Griffith, L. M., Ward, R. C., McBride, G. B., and Loftis, J. C. (2001): Data analysis considerations in producing 'comparable' information for water quality management purposes, National Water Quality Monitoring Council, 01-01, 45 p.

Grubbs, G. (2001): Memo: Development and Adoption of Nutrient Criteria into Water Quality Standards, United States Environmental Protection Agency, Office of Science and Technology, p.

Hem, J. D. (1985): Section 4: Significance of properties and constituents reported in water analyses: *Study and interpretation of the chemical characteristics of natural water (3rd ed.)*. U.S. Geological Survey. Accessed March 24, 2004 at http://water.usgs.gov/pubs/wsp/wsp2254/html/pdf.html

Hirsch, R. M., Alexander, R. B., and Smith, R. A. (1991): Selection of methods for the detection and estimation of trends in water quality. *Water Resources Research* **27**, 803-813.

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

Hoppe, M. C. (2003): 1-5. Ground and Surface Water Quality: *New York - New Jersey Highlands Regional Study Technical Report*. USDA Forest Service. Accessed April 2, 2004 at http://www.fs.fed.us/na/highlands/atechnical_report/index.html

Horne, A. J. (2003): Personal Communication, In S. Stafford (Ed.), Berkeley, CA.

Horne, A. J., and Goldman, C. R. (1994): *Limnology*. McGraw-Hill, Inc. New York.

Ketcham, B. J. (2001): Point Reyes National Seashore Water Quality Monitoring Report: May1999 - May 2001, National Park Service, p.

Kirchner, J. (2003): EPS 120 Reader: Analysis of Environmental Data, pp. 157. Berkeley.

Larsen, R., E. (1999a): *Fact Sheet No. 28: Fishery Habitat: 3) Sediment and Pollutants*. UC Davis, University of California Cooperative Extension. Accessed March 25, 2004 at http://danr.ucop.edu/uccelr/h28.htm

Larsen, R., E. (1999b): *Fact Sheet No. 29: Fishery Habitat: 4) Dissolved Oxygen*. UC Davis, University of California Cooperative Extension. Accessed March 25, 2004 at

Lee, R. Y., and Taylor, K. A. (1995): Water Quality Control Plan for the San Francisco Bay Basin (Basin Plan), California Regional Water Quality Control Board, San Francisco Bay Region, p.

Lloyd, D. S. (1987): Turbidity as a water quality standard for salmonid habitats in Alaska. *North American Journal of Fisheries Management* **7**, 34-35.

MacDonald, L. H., Smart, A. W., and Wissmar, R. C. (1991): Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska, U.S. Environmental Protection Agency, Region 10, EPA 9109-9-91-001, 155 p.

Maier, R. M., Pepper, I. L., and Gerba, C. P. (2000): *Environmental Microbiology*. Academic Press. New York.

McBean, E. A., and Rovers, F. A. (1998): *Statistical Procedures for Analysis of Environmental Monitoring Data and Risk Assessment.* Prentice Hall PTR. Upper Saddle River, New Jersey.

McBride, G. B., Loftis, J. C., and Adkins, N. C. (1993): What do significance tests really tell us about the environment? *Environmental Management* **17**, 423-432.

Meybeck, M. (1982): Carbon, nitrogen and phosphorous transport by world rivers. *American Journal of Science* **282**, 401-450.

Mount, J. F. (1995): *California Rivers and Streams*. University of California Press. Berkeley, CA.

Myers, D. N. (2003): *Fecal Indicator Bacteria: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A7. section 7.1*. U.S. Geological Survey. Accessed March 25, 2004 at http://pubs.water.usgs.gov/twri9A7/

Neter, J., Wasserman, W., and Kutner, M. H. (1990): Ch.13: *Applied Linear Statistical Models*, Richard D. Irwin, Inc.

Penoyer, P. (2003): *Vital Signs Long-Term Aquatic Monitoring Projects: Part C: Draft Guidance on WRD Required and Other Field Parameter Measurements, General Monitoring Methods and Some Design Considerations in Preparation of a Detailed Study Plan*. National Park Service, Water Resources Division. Accessed March 24, 2004 at http://science.nature.nps.gov/im/monitor/vsmTG.htm#TechGuide

Radtke, D. B., Busenberg, E., Wilde, F. D., and Kurklin, J. K. (2003): *pH: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A6, section 6.4*. U.S. Geological Survey, Office of Water Quality. Accessed March 23, 2004 at http://pubs.water.usgs.gov/twri9A6

Radtke, D. B., Kurklin, J. K., and Wilde, F. D. (1998a): *Temperature: U.S. Geologic Survey Techniques of Water-Resources Investigations, book 9, chap. A6, section 6.1*. U.S. Geologic Survey, Office of Water. Accessed March 23, 2004 at http://pubs.water.usgs.gov/twri9A6/

Radtke, D. B., White, A. F., Davis, J. V., and Wilde, F. D. (1998b): *Dissolved oxygen: U.S. Geologic Survey Techniques of Water-Resources Investigations, book 9, chap. A6., section 6.2*. U.S. Geologic Survey Office of Water Quality. Accessed March 22, 2004 at http://water.usgs.gov/owq/FieldManual/Chapter6/6.2_contents.html

Reiser, D. W., and Bjornn, T. C. (1979): Habitat requirements of anadromous salmonids: Portland, OR, USDA Forest Service, Pacific NW Forest and Range Exper. Station, PNW-96, 54 p.

Rittman, B. E., and McCarty, P. L. (2001): *Environmental Biotechnology: Principles and Applications*. McGraw-Hill.

Rosenlieb, G., Penoyer, P., Jackson, B., and Kimball, D. (2002): *Recommendations for Core Water Quality Monitoring Parameters and Other Key Elements of the NPS Vital Signs Program Water Quality Monitoring Component*. National Park Service, Freshwater Workgroup Subcommittee. Accessed March 24, 2004 at http://science.nature.nps.gov/im/monitor/vsmTG.htm#TechGuide

Sawyer, C. N., and McCarty, P. L. (1978): *Chemistry for environmental engineering*. McGraw-Hill. New York.

Shelton, L. R. (1994): *Field guide for collecting and processing stream-water samples for the National Water-Quality Assessment Program*. U.S. Geological Survey. Sacramento, Ca.

Stevens, H. H., Ficke, J. F., and Smoot, G. F. (1975): Chapter D1: Water Temperature-Influential Factors, Field Measurements, and Data Presentation ed.), pp. 65: *Techniques of Water-Resources Investigations of the United States Geological Survey*, USGS, Arlington, VA.

Strausberg, S. I. (1983): Turbidity interferes with accuracy in heavy metal concentrations. *Industrial Wastes* **29**, 16-21.

Theurer, F. D., Voos, K. A., and Miller, W. J. (1984): Instream water temperature model. Instream Flow Information Paper No. 16, U.S. Fish and Wildlife Service, FWS/OBS-84/15, p.

UNESCO (1985): The International System of Units (SI) in Oceanography. *Technical Papers in Marine Science* **45**, 124.

Ward, R. C., Loftis, J. C., and McBride, G. B. (1990): *Design of Water Quality Monitoring Systems*. Van Nostrand Reinhold. New York.

Weiss, R. F. (1970): Solubility of Nitrogen, Oxygen and Argon in Water and Seawater. *Deep-Sea Research* **17**, 721-&.

Wilde, F. D., and Gibs, J. (1998): *Turbidity: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A6, section 6.7*. U.S. Geological Survey, Office of Water Quality. Accessed March 25, 2004 at http://pubs.water.usgs.gov/twri9A6

Wilde, F. D., and Radtke, D. B. (2003): *General Information and Guidelines: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A6, section 6.0*. USGS Water Resources, Office of Water Quality. Accessed March 24, 2004 at http://pubs.water.usgs.gov/twri9A6/

Wilde, F. D., Radtke, D. B., Gibs, J., and Iwatsubo, R. T. (1999): *Collection of water samples: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A4*. USGS Water Resources, Office of Water Quality. Accessed March 24, 2004 at http://pubs.water.usgs.gov/twri9A4/

Zar, J. H. (1999): *Biostatistical Analysis*. Prentice-Hall. New Jersey.

APPENDIX A: DATA DISTRIBUTIONS

See word documents: 'Pinnacles Data Charts.doc', 'PORE distributions.doc', 'GGNRA jmp histograms.doc', 'Beutel97-98Distributions.doc', and 'Stables Distributions.doc'.

APPENDIX B: TIME PLOTS OF PARKS DATA

See Excel file 'charts9-28.xls' for revised time plots and 'Data Charts.xls' for unrevised time plots.

APPENDIX C: SAMPLE SITE LOCATION TABLES

APPENDIX D: NOTES ON DATA PREPARATION AND ANALYSIS

The data analysis here includes data from the GGNRA pre 1997 database as well as from the PORE Access database. It also includes the data sets described below: "cleaned up stables data" and "stationsNOTinDB." Not all parameters from these datasets are included, but all of the applicable data is included in the analysis of the parameters discussed. For all of the data sets except for "Cleaned up Stables data," there are some quality control duplicates and triplicates included, which should be consolidated into one value for more rigorous results. This was not done within the scope of the current project due to time constraints.

The file STREAM.xls contains data from Beutel's report from winter 1997-1998. Some of this data is already in the GGNRA pre 1997 Access database, namely the sheets for:

- Rodeo
- Ten. Valley
- Redwood

The other sheets are not included in either the PORE or GGNRA Access databases. These include:

- Pine Gulch
- Olema
- Lagunitas

The data that is not already in the databases has now been compiled in a sheet titled "stationsNOTinDB"

Although the GGNRA Access database is titled Pre-1997, it does include some data up to April of 1999.

The following data sets are also not included in either of the Access databases, but are included in the current analysis:

- 99stables_WQDATA (winter 1999-2000)
- 00_stablesWQDATA (winter 2000-2001)
- 01finaldata (winter 2001-2002)

Data from the above files have been compiled together in the new sheet "stables 99-01" in the file "cleaned up stables data". Quality control duplicates and triplicates have been deleted and one record left containing the average value.

The files rodwq98 and rodwq1 is not included in either of the Access databases and includes data from Rodeo Lagoon synoptic studies. It should be analyzed separately since it is a lagoon, not a stream.

GGNRA includes very impacted, urban areas as well as more pristine remote areas. For the purpose of establishing general background values, it is important to separate sites according to level of disturbance. There is currently no easy way to do this, the way the data has been stored. For future use, a column should be included in the sites table indicating watershed characteristics and level of disturbance. In addition, there are some discrepancies between site names that have been monitored by different persons within GGNRA and PRNS. In an attempt to straighten these out, I have created a new file (stationlistDUP) to be distributed to persons involved in monitoring to determine which sites are the same or close enough together to be analyzed together. I have also merged the station lists from GGNRA and PNRS into one table, including several new columns to facilitate future data analysis.

A note of warning: Excel records dates in one of two formats. Both formats convert the date into a single number, but one format starts with Jan.1, 1900 as day 1 and the other format starts with Jan.1, 1904…When moving data from one table to another, make sure that the same type of formatting is being used, or else all your dates will be three years off. You can verify the type of year format being used by looking under tools at the options window. If you click on the calculations tab, under workbook options, there is a box labeled 1904 date system. Checking it makes day 1 start in 1904. Leaving it blank makes day 1 start in 1900.

Notes on Pinnacles data analysis.

For data analysis, quality control samples (duplicates, replicates) need to be averaged, and just one value reported. Units should be included in column headings. A separate table or document should include specific methods and detection limits. Written words, asterisks, and non-numeric symbols should not be included in same cell as numerical data, for example sheet 1998 NO3 column. It's not clear to me what is meant by "* As a value of P, NO3, N, etc". For values that are below or above the detection limit, a column can be made right before with the greater than or less than symbol in it. Some kind of data entry check/validation should be performed. For example, there were a few pH values that read 78 or 82 instead of 7.8 and 8.2.

Specific questions

- What are units for salinity? Often reported as ppt (parts per thousand), but sometimes as percent (parts per hundred) and sometimes as mg/L or ppm
- Is "Field N" nitrate, total N? why is it often non-detect when nitrate is 5-6 mg/L?
-

Sheet 1998
-went through and checked date/location/time for duplicates. If more than one with same date/location/time, average values and leave only one.
Only one duplicate found, for Upper Chalone Creek on 9/8/1997
-removed words from data cells, (NO3)
-created new working columns with _ the value of below detection limits and average value for cells with multiple numbers i.e. 1.2/1.1, as well as given value when within detection limits

Sheet 1999
-went through and checked for duplicates as for Sheet 1998. Found none. Found several with multiple measurements on same day, but left in – highly variable flows.
-created new working columns with _ the value of below detection limits and average value for cells with multiple numbers i.e. 1.2/1.1, as well as given value when within detection limits

Sheet 2000
-checked for duplicates as above. Found none. Left in multiple measurements from different times on same day as above.
-created new working columns with _ the value of below detection limits and average value for cells with multiple numbers i.e. 1.2/1.1, as well as given value when within detection limits

Sheet 2001
-contains multiple measures of depth and velocity for average flow calculation. Created new sheet, 2001A containing only average values. Checked for duplicates as above. Found none. Left in multiple measurements from different times on same day.
-created new working columns in 2001A with _ the value of below detection limits, or twice the value of above detection limits, and average value for cells with multiple numbers i.e. 1.2/1.1, as well as given value when within detection limits.

Sheet 98to2001
-A new sheet with the modified data sheets from above ready to export to JMP for analysis.

Pinnacles98to2001.jmp is the above data now in JMP format.
-the following parameters were put into log scale for analysis, using the column labeled _ DL where applicable:

flow
coliforms
E. coli
Field N
Field P
Kjeldahl N
Ammonia N

Nitrate N
Nitrite N
Total P
Ortho P
TDS
TSS

Note that an alternative way to look at the data would be to use the original column instead of the _ DL column, but transform the data into categorical data instead of continuous, to handle the below detection values.

-pH values that were way out of range, e.g., 78 were modified to 7.8, assuming that decimal point was left out during data entry.