# APPLICATION OF NATIONAL STREAM QUALITY ACCOUNTING NETWORK (NASQAN) STATION DATA FOR ASSESSING WATER QUALITY IN THE PEACE RIVER BASIN, FLORIDA

By Edward R. German and Donna M. Schiffer

---

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 87-4167

Tallahassee, Florida

1988

DEPARTMENT OF THE INTERIOR
DONALD PAUL HODEL, Secretary
U.S. GEOLOGICAL SURVEY
Dallas L. Peck, Director

---

For additional information
write to:

District Chief
U.S. Geological Survey
Suite 3015
227 N. Bronough Street
Tallahassee, Florida 32301

Copies of this report can
be purchased from:

U.S. Geological Survey
Books and Open-File Reports Section
Federal Center, Bldg. 810
Box 25425
Denver, Colorado 80225

## CONTENTS



## ILLUSTRATIONS

ILLUSTRATIONS--Continued

## ILLUSTRATIONS--Continued

## TABLES

# APPLICATION OF NATIONAL STREAM QUALITY ACCOUNTING NETWORK (NASQAN) STATION DATA FOR ASSESSING WATER QUALITY IN THE PEACE RIVER BASIN, FLORIDA

*By Edward R. German and Donna M. Schiffer*

## ABSTRACT

*Selected water-quality data for the Peace River basin at and upstream from Arcadia, Florida, were summarized to determine water-quality conditions, basin yields, and stream loadings. Data were examined for discernment of trends and to determine if the periodic sampling at the National Stream Quality Accounting Unit station at Arcadia was representative of the range of discharge conditions and of water quality throughout the basin.*

*Water quality varies within the basin with relatively high chemical yields from some subbasins. For example, mean total phosphorus yields along the Peace River for October 1974 through September 1982 increased from 0.9 ton per year per square mile at Bartow, to 1.3 tons per square mile at Zolfo Springs.*

*Data, as early as 1951, on four of eight stations show upward trends in specific conductance. However, since October 1974, data show that only two of the eight stations showed evidence of a significant trend.*

*Periodic sampling of National Stream Quality Accounting Unit station 02296750, Peace River at Arcadia, since October 1974 has been representative of the range of discharge. Increased frequency of sampling could probably reduce errors in computing annual mean values, with diminishing reduction in error at frequencies exceeding weekly to biweekly. Use of water-quality discharge relations could reduce the number of samples necessary for estimating mean annual values.*

*Intensive reconnaissance sampling in February and May 1983 provided more information on subbasin chemical yields and verified that National Stream Quality Accounting Unit data do not represent the entire basin. For example, dissolved-solids yield in the May 1983 sampling was 1.1 tons per day per square mile in a subbasin downstream from Bartow, and was only 0.09 ton per day per square mile at the outflow station for the entire basin. Some segments of the Peace River apparently served as sinks for nitrogen and phosphorus.*

# INTRODUCTION

The National Stream Quality Accounting Network (NASQAN) is a nationwide data-collection network that consists of 501 monitoring stations that measure streamflow and water-quality on a routine basis. This network was begun in 1974 with the specific purpose to obtain regional and nationwide overviews of the quality of the Nation's streams. Data from the network are used to: (1) account for the quantity and quality of water moving within the major rivers in the United States, (2) document on a large-scale or regional basis the variation in national stream quality, and (3) detect changes in stream quality with time (Ficke and Hawkinson, 1975, p. 1).

The spacing of NASQAN stations is based on a system of hydrologic subdivisions developed by the U.S. Water Resources Council (1970). In this system, drainage basins in the United States are divided into 21 regions, 222 subregions, and 352 accounting units, the latter two being progressively smaller parts of a region. The stations are located at points chosen to provide an accurate representation of the quality of water leaving an accounting unit.

The NASQAN data have been useful in national and regional generalizations regarding water quality of the Nation's rivers; however, little is known regarding the use of data from a single NASQAN station to assess water-quality conditions within a stream basin relative to such factors as land use. Furthermore, little is known about the adequacy of the established sampling frequency of the NASQAN program to define the full range of water-quality conditions on an annual or long-term basis. To address the above-noted unknowns, the U.S. Geological Survey has undertaken several investigations of application or use of NASQAN data for assessing water quality in river basins. These investigations began in 1981.

## Purpose and Scope

This report describes the results of a study to:

- Assess the use of water-quality data collected at NASQAN station 02296750 to represent, on both a spatial and temporal basis, the water quality for accounting unit 03100101 upstream of NASQAN station 02296750.
- Selectively describe, on both a spatial and temporal basis, the stream-water quality throughout the basin upstream of NASQAN station 02296750, Peace River at Arcadia.
- Determine sampling frequency needed to describe water-quality conditions on a continuing basis at any particular site upstream of NASQAN station 02296750.
- Describe methods of analysis of water-quality data used in the study to provide a potential guide to available exploratory data analysis techniques for future use.

The data used for this study include the NASQAN data and data from other stations in the basin available from other studies or networks. Most of the interpretation is based on data for October 1974 through September 1982 to coincide with the operation of the NASQAN station at Arcadia. Some pre-1974 data for other stations were used to look for the existence of long-term trends in water quality. The water-quality variables of primary interest for this assessment were specific conductance, dissolved oxygen (DO), selected nutrients and metals, dissolved solids, and major ions. Discharge records were used where available for computation of loads and yields, and flow-duration curves.

## Study Area

The study area consists of the drainage basin of the Peace River upstream from the U.S. Geological Survey gaging station at Arcadia, Fla., a total of 1,367 $mi^2$ (fig. 1). This area occupies 54 percent of the total drainage area of the Peace River (2,403 $mi^2$) and is referred to as the upper Peace River basin. The river has its headwaters among a group of lakes between Lakeland and Haines City. The headwaters area is drained by Saddle Creek and Peace Creek drainage canal. Peace River begins at the confluence of these two creeks, immediately northeast of Bartow.

The drainage basin includes parts of Polk, Hardee, and De Soto Counties (upper basin) and Charlotte County (lower basin). Peace River flows southward to Charlotte Harbor and thence, to the Gulf of Mexico. Major tributaries in the upper basin include Bowlegs, Payne, and Charlie Creeks.

Land use within the basin is predominantly forest and rangeland. Phosphate strip mines are in the upper basin. A small part of the basin is urbanized, and a small part is used in citrus cultivation.

## Data Base

For this study, two major computer data files, the U.S. Geological Survey National Water Data Storage and Retrieval System (WATSTORE) files and the U.S. Environmental Protection Agency STORET file (excluding Geological Survey data), were scanned for water-quality data. A summary of the data from both files indicated that, for stations with more than 10 samples, there were data for approximately 495 water-quality samples in STORET and 790 in WATSTORE. These tabulations do not include samples where only specific conductance, pH, dissolved oxygen, or water temperature were measured.

Most of the data used in this study were collected by the Geological Survey and are available in WATSTORE and in annual data reports of the Geological Survey. Data from STORET were collected mainly by the Florida Department of Environmental Regulation (FDER). These data were summarized for two sites, but were not used in trend analysis or for loading calculations because discharge for the sampling dates were not available. The additional data available from STORET not used in this study were for sites at which Geological Survey data were available.

Nine stations were used for much of the data analysis (table 1 and fig. 2). Available data at each station varied, but generally for each station specific conductance, some nutrients, major ions, and discharge data were available.

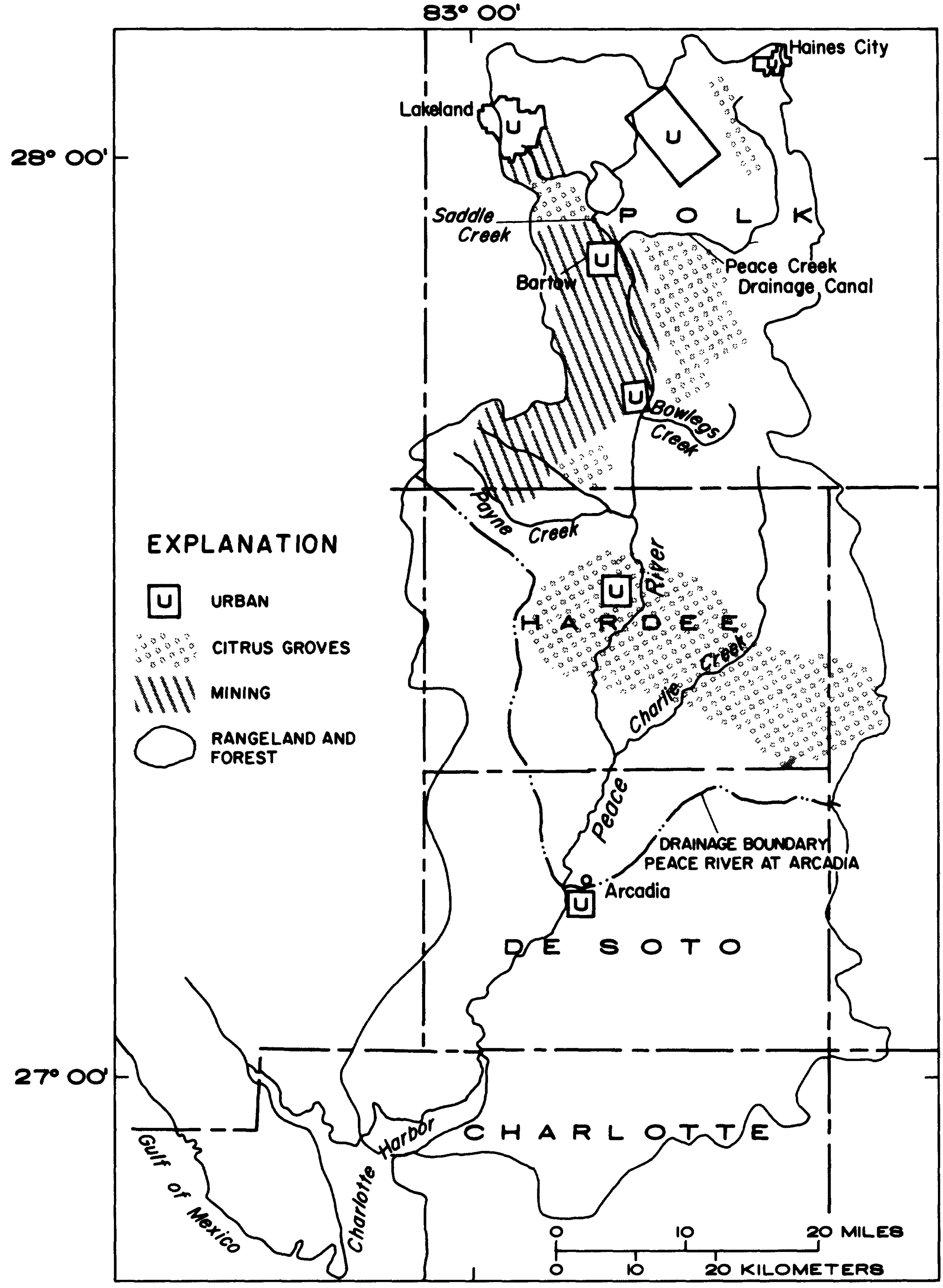


Figure 1.--Peace River basin and generalized land use.

Table 1.--Major stations used in analysis

| Map and site No. | USGS station No. | Name | Latitude | Longitude | Drainage area, in square miles |
|---|---|---|---|---|---|
| 1 | 02293986 | Peace Creek drainage canal | 27°55'23" | 81°42'28" | 160 |
| 3 | 02294491 | Saddle Creek | 27°56'17" | 81°51'05" | 135 |
| 4 | 02294650 | Peace River at Bartow | 27°54'07" | 81°49'03" | 390 |
| 6 | 02294898 | Peace River at Fort Meade | 27°45'04" | 81°46'56" | 465 |
| 10 | 02295420 | Payne Creek | 27°37'13" | 81°49'33" | 121 |
| 15 | 02295637 | Peace River at Zolfo Springs | 27°30'15" | 81°48'04" | 826 |
| 21 | 02296223 | Little Charlie Bowlegs Creek | 27°28'40" | 81°33'25" | 41.9 |
| 22 | 02296500 | Charlie Creek | 27°22'29" | 81°47'48" | 330 |
| 24 | 02296750 | Peace River at Arcadia | 27°13'19" | 81°52'34" | 1,367 |

During the 1983 water year (October 1982 through September 1983), the Survey, in cooperation with the FDER as part of independent investigation, collected reconnaissance data at 23 sites within the study area during the months of August, November, February, and May. Each of these data sets were collected within a span of 2 or 3 days and provided a "snapshot" of water quality throughout the basin for different seasons and flow conditions.

The water-quality data available for this study fall into two categories according to the sampling scheme. One category includes data from the recurring collection of samples at nine stations and is useful for assessment of water quality for a variety of flow conditions, and for changes over time (trends). The other category includes the data from the intensive reconnaissance, in which many stations were sampled nearly simultaneously, providing for a more comprehensive point-to-point comparison of water-quality at the time of reconnaissance sampling.

## Methods of Analysis

The various techniques used to analyze water-quality data in an effort to meet the objectives of the study are as follows:

- Graphical methods of data analysis, including box plots (Tukey, 1977) and Piper diagrams (Hem, 1970).
- Calculation of percentage of samples not meeting applicable State water-quality standards for selected constituents at several stations in the basin.
- Statistical distribution testing (testing for normality).
- Parametric and nonparametric analysis of variance of water-quality measures (identifying areal variation).

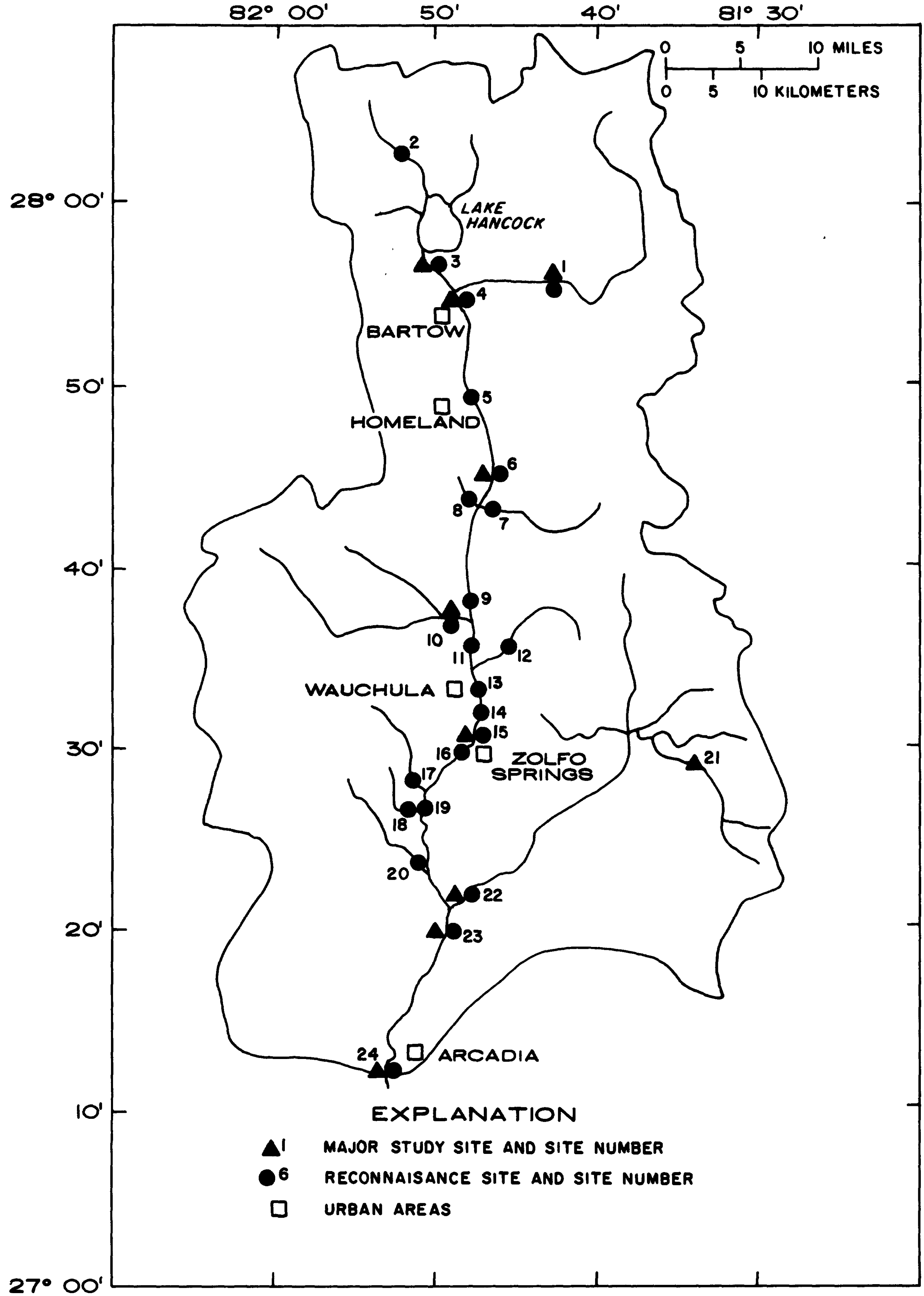


Figure 2.--Stations used in analysis.

- Regression analysis (relating water quality to stream discharge).

- Calculation of loads and yields of selected constituents, based on relations determined by regression analysis in the previous step.

- Statistical testing for short- and long-term trends in water-quality data (including testing flow adjusted data and seasonally grouped data).

- Analysis of sampling frequency required to adequately define water-quality conditions at a site (Monte Carlo simulation of sampling, and optimization of sampling based on the water-quality discharge relation).

- Analysis of data from intensive reconnaissance sampling and comparison of loads and yields computed from the data obtained from this indepth sampling (to determine the relation of a single-station sample, as in NASQAN, to the water quality throughout a drainage basin).

It was decided that some explanation of the methods used for this study might be beneficial to future investigators because of the various techniques used. Information about each data-analysis technique is discussed when the technique appears in the report so as to relate results of its use with actual data.

## ASSESSMENT OF WATER QUALITY IN THE PEACE RIVER BASIN, 1974-82

Objectives of this study included assessment of the adequacy of data collected at the NASQAN station at Arcadia to represent water quality throughout the basin, and the description of areal variability of water quality in the basin by use of data collected at the NASQAN station at Arcadia and water-quality data at other sites in the basin.

Several methods of data presentation and analysis have been used, as follows:

- Graphical representation of the major water-quality variables at each site (box plots and Piper diagrams).

- Comparison of conformance to State water-quality standards among sites within the Peace River basin.

- Analysis of statistical distributions for selected water-quality variables at individual sites.

- Statistical testing of data to determine if means, medians, or ranks of data are the same at all stations.

## Areal Variability of Water Quality

### Graphical Comparison of Water-Quality Data at Selected Sites

Figures 3 through 10 represent graphical summaries for comparison of selected water-quality data for sites in downstream order within the Peace River basin. The site numbers in the figures increase from left to right with the Peace River at Arcadia (site 24), the furthest downstream site, at the rightmost position. This type of data presentation, called a box plot or box-and-whisker graph, originated from Tukey (1977) for the purpose of graphically showing statistical distribution information about the data. The "box" is formed by the upper and lower quartiles (75 percent and 25 percent, respectively), also referred to as "hinges." An "0" indicates the location of one or more values that are "outside" (occurring only once in 20 samples from a normal distribution), and an "*" indicates the location of one or more "detached" values (occurring only once in 200 samples from a normal distribution). Median values are indicated by the dashed line within the "box," and mean values are indicated by a "+" (SAS Institute, Inc.,[1] 1980). It should be noted here that these definitions for the "outside" and "detached" values are not identical to Tukey's definitions. Box plots are useful for the determination of central tendency (shortness of the box), and detection of outliers ("0" and "*"), and can also be used to discover a lack of symmetry (the parts of the box are unequal) in the distribution of data. For any particular variable, they can illustrate similarities or differences (box sizes and shapes) among data sets at individual stations.

Because concentrations of dissolved oxygen (DO) fluctuate daily and seasonally, there is a greater probability for large variability in DO data than in data for other water-quality variables. Figure 3 shows that the Saddle Creek station (site 3) has the greatest range of DO concentrations, followed by Peace Creek drainage canal near Alturas (site 1). Together, the two streams gaged by these two stations form the headwaters of the Peace River and receive runoff from agricultural lands (Peace Creek) and outflow from Lake Hancock (Saddle Creek). Distributions of DO data collected downstream and in tributaries have smaller variances.

Distributions of specific-conductance values (fig. 4) in the main stem of the Peace River (sites 4, 6, 15, and 24) are similar, but waters that enter from tributaries have lower median values, particularly Little Charley Bowlegs Creek (site 21). Discharge contributions from tributaries to the total flow in the main river are small, therefore, the effect of low specific conductance water from tributaries on specific conductance of the main stem is minimal. Charlie Creek (site 22) is the last major tributary entering the river before Arcadia, but the generally lower values of specific conductance in Charlie Creek seem to have little effect on the distribution of specific conductance values at Arcadia, other than to lower the value found at the 25th and 50th percentile.

---

[1]Use of the trade name, Statistical Analysis System (SAS Institute, Inc.), in this report is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey.

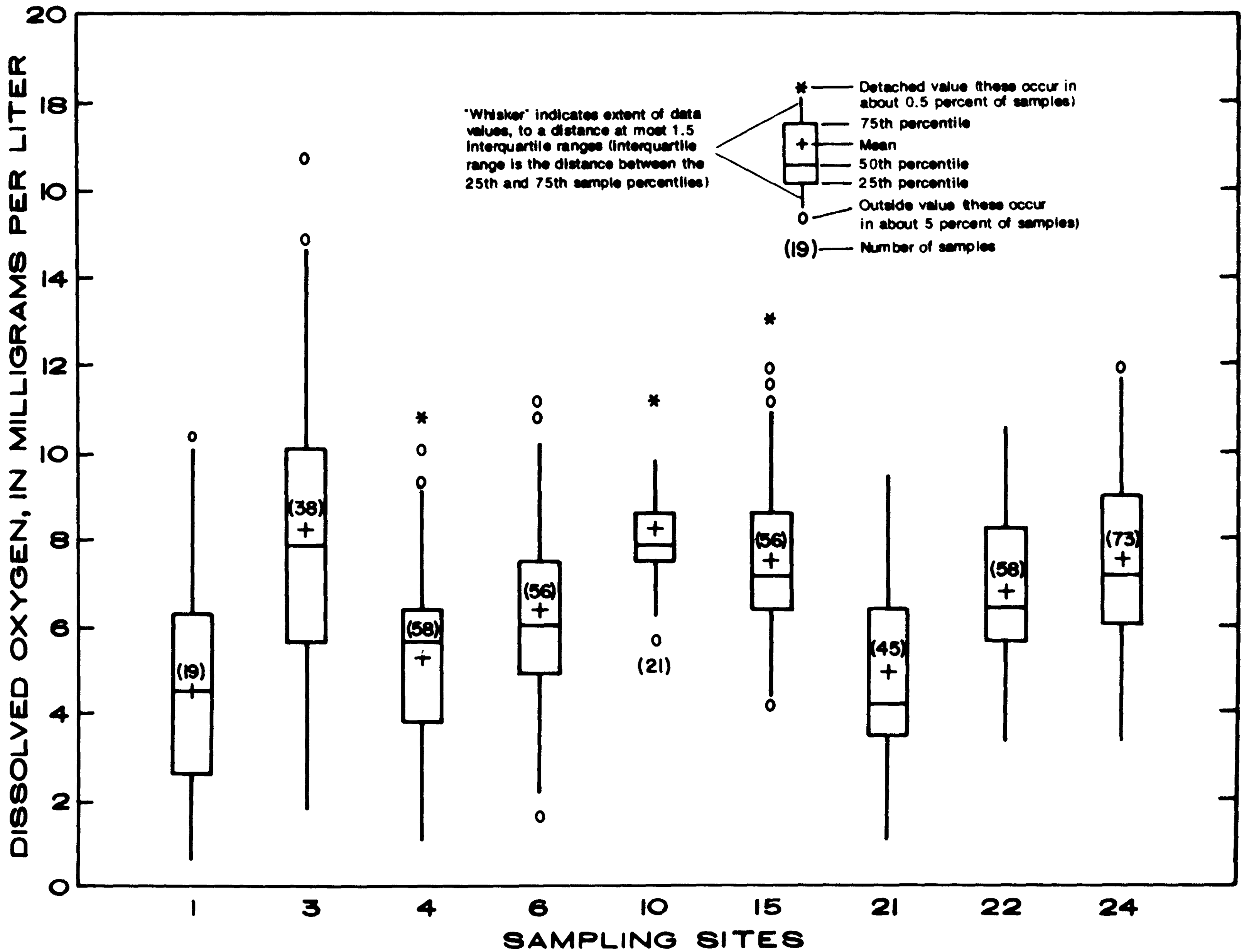


Figure 3.--Dissolved oxygen, October 1974 through September 1982.

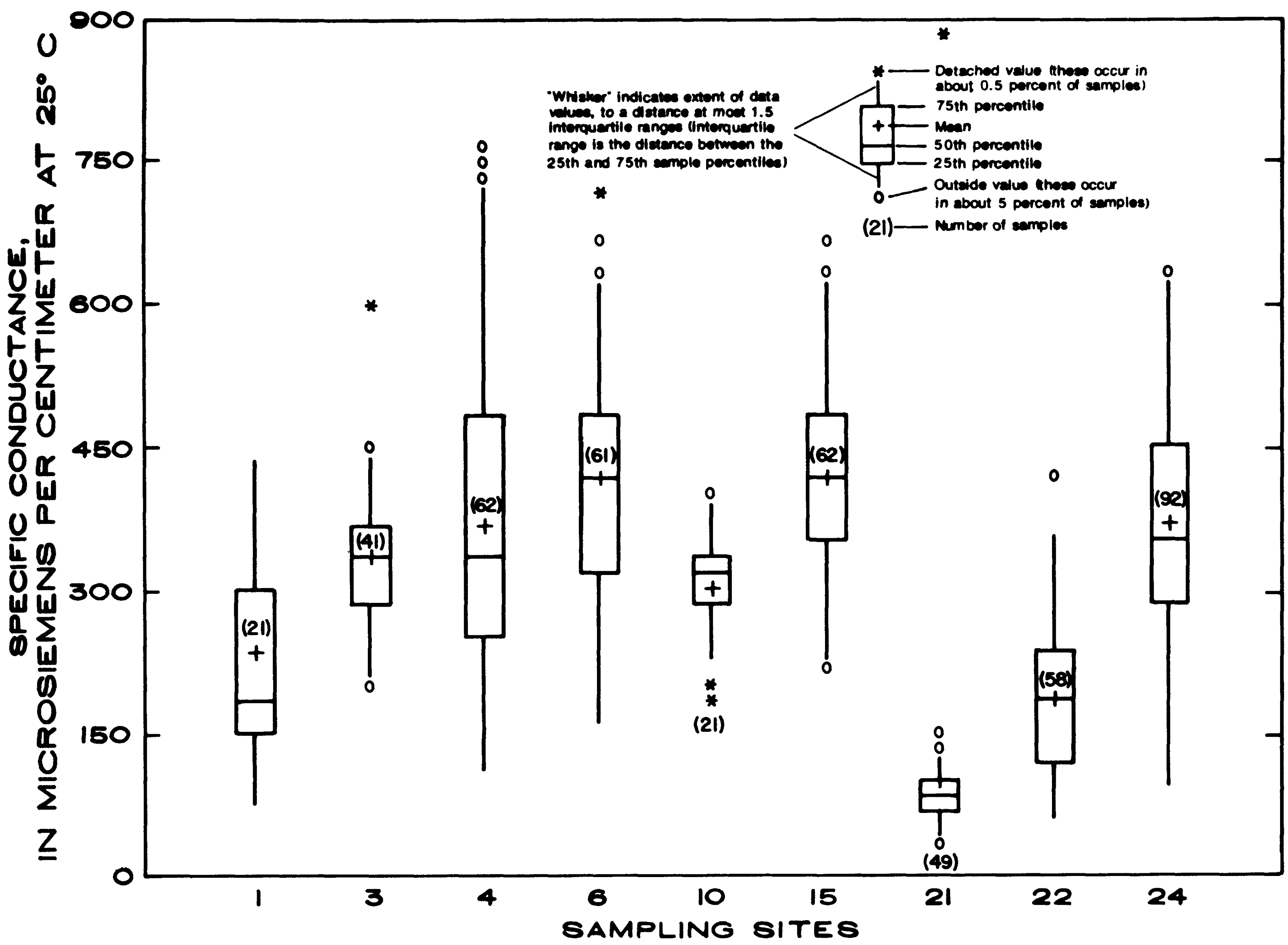


Figure 4.--Specific conductance, October 1974 through September 1982.

Box plots of pH for nine stations (fig. 5) in the Peace River basin indicate that distributions for these variables do not cover a broad range of values, with two exceptions. The first exception, Saddle Creek (site 3), has higher values and a larger range for pH than the other sites. Because of the location of the station (immediately south of Lake Hancock), water quality is subject to the quality of the lake water. DO and pH in the lake may fluctuate due to the presence of algae and vegetation, which can raise DO and pH during respiration. The second exception, Little Charlie Bowlegs Creek (site 21), has pH values that are lower than the pH values at the other sites. These low pH values may be due to organic acids leached from plant debris in swampy parts of the basin.

Box plots such as the one shown in figure 5 for site 1, and again in figure 6 for the same site, are the result of insufficient data to describe the distribution. A single line such as is shown in figure 6 for sites 1, 3, and 10 indicates that only one data point was available from which to make the plot. If no data are available, no plot is shown, as in figure 6 for site 21.

Total nitrogen distributions in figure 6 show a large range for sites 4 and 6 in comparison to the more narrowly defined range for the other stations. Organic nitrogen (fig. 7) has similar pattern to total nitrogen. Relatively low values are found in the plots for total nitrate nitrogen (fig. 8) indicating that, at most sites, much of the nitrogen is in the organic form.

Distributions for total phosphorus and dissolved chloride are illustrated in figures 9 and 10, respectively. The two streams that form the headwaters to the Peace River (sites 1 and 3) do not contribute enough of a phosphorus load to explain the distribution of total phosphorus found at Bartow (site 4). The higher values at the Bartow station are possibly due to strip mining in the basin. Runoff from these mining areas enters the river somewhere downstream of the two headwater stations and upstream of Bartow. The effect of high phosphorus concentrations at site 4, far upstream from the Arcadia station, is mitigated within the river by the inflow of waters with lower phosphorus concentrations. The higher chloride values for the Bartow station likely reflect agricultural or industrial land uses within the basin.

Relative concentrations of major ions in water are illustrated through the use of trilinear diagrams of anions, and cations, and a Piper diagram of both. These diagrams are a graphical method of typing water by ionic composition and can indicate whether or not a particular water is a mixture of other waters or similar to water at one or two other locations (Hem, 1970, p. 268).

Major cations found at four sites in the Peace River basin are shown by a Piper diagram in figure 11. These four sites were the only ones with a sufficient number of analyses to plot for comparison. From the diagram, it appears that the water at Arcadia (triangle symbol) generally plots between Peace River at Zolfo Springs (asterisk) and Charlie Creek (dot) waters, and is a mixture of these waters. Composition of cations was from 30 to 40 percent magnesium, 10 to 30 percent sodium and potassium, and 40 to 60 percent calcium. Water from the upper part of the basin, at Peace River at Bartow (plus sign), differed from water from the other sites because of a higher percentage of sodium and potassium, but a lower percentage of magnesium. At all four sites, calcium is the predominant cation.

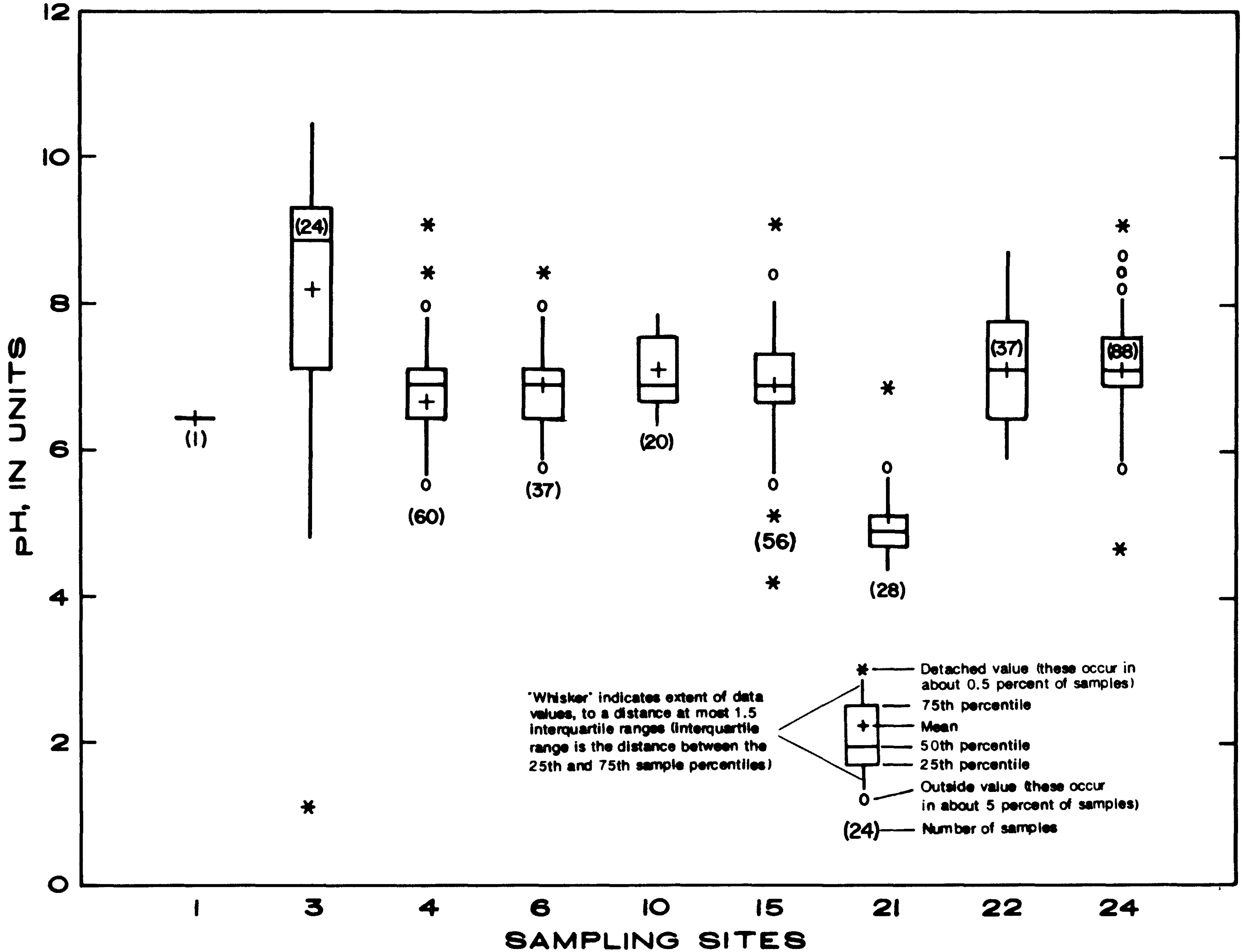


Figure 5.--pH, October 1974 through September 1982.

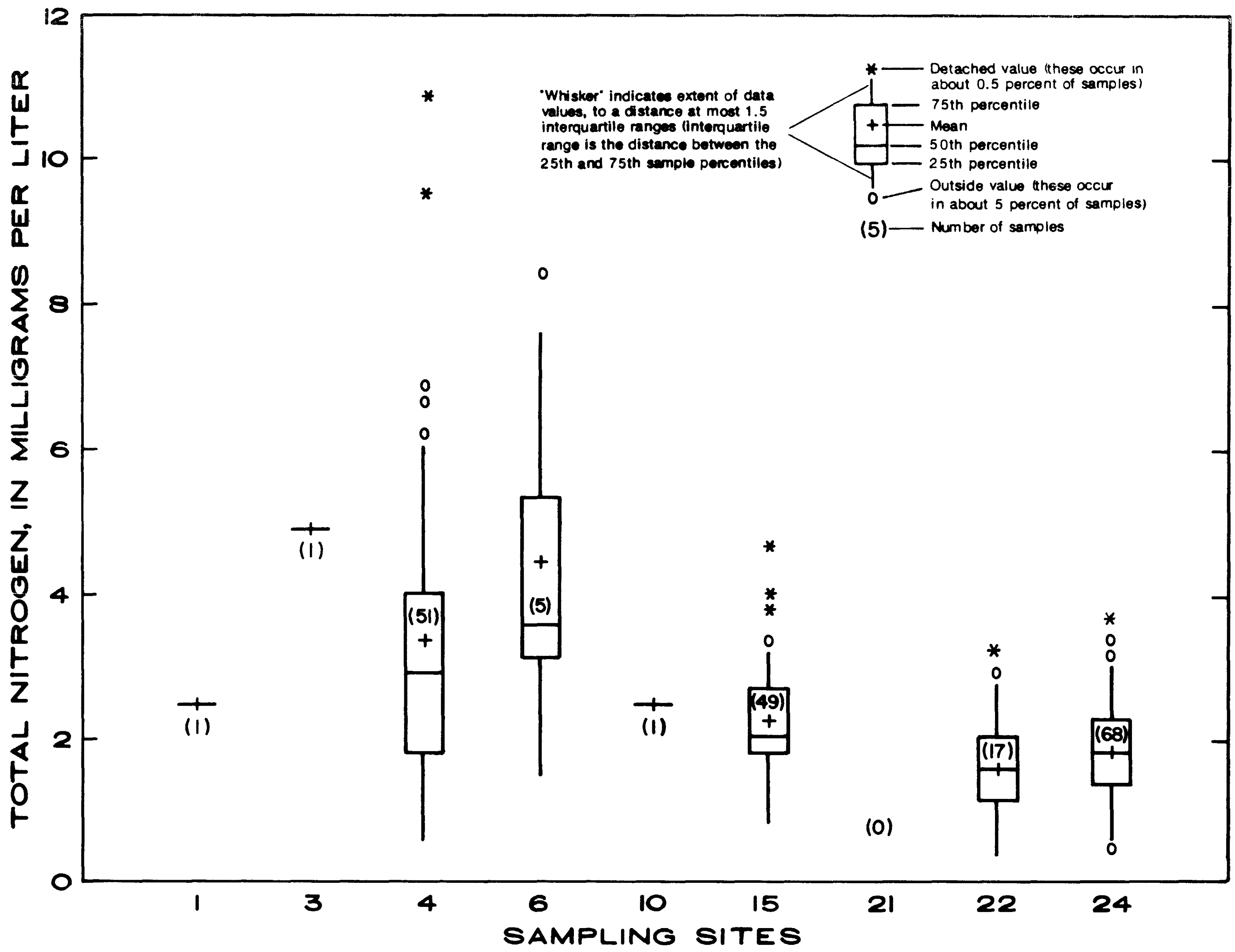


Figure 6.--Total nitrogen, October 1974 through September 1982.

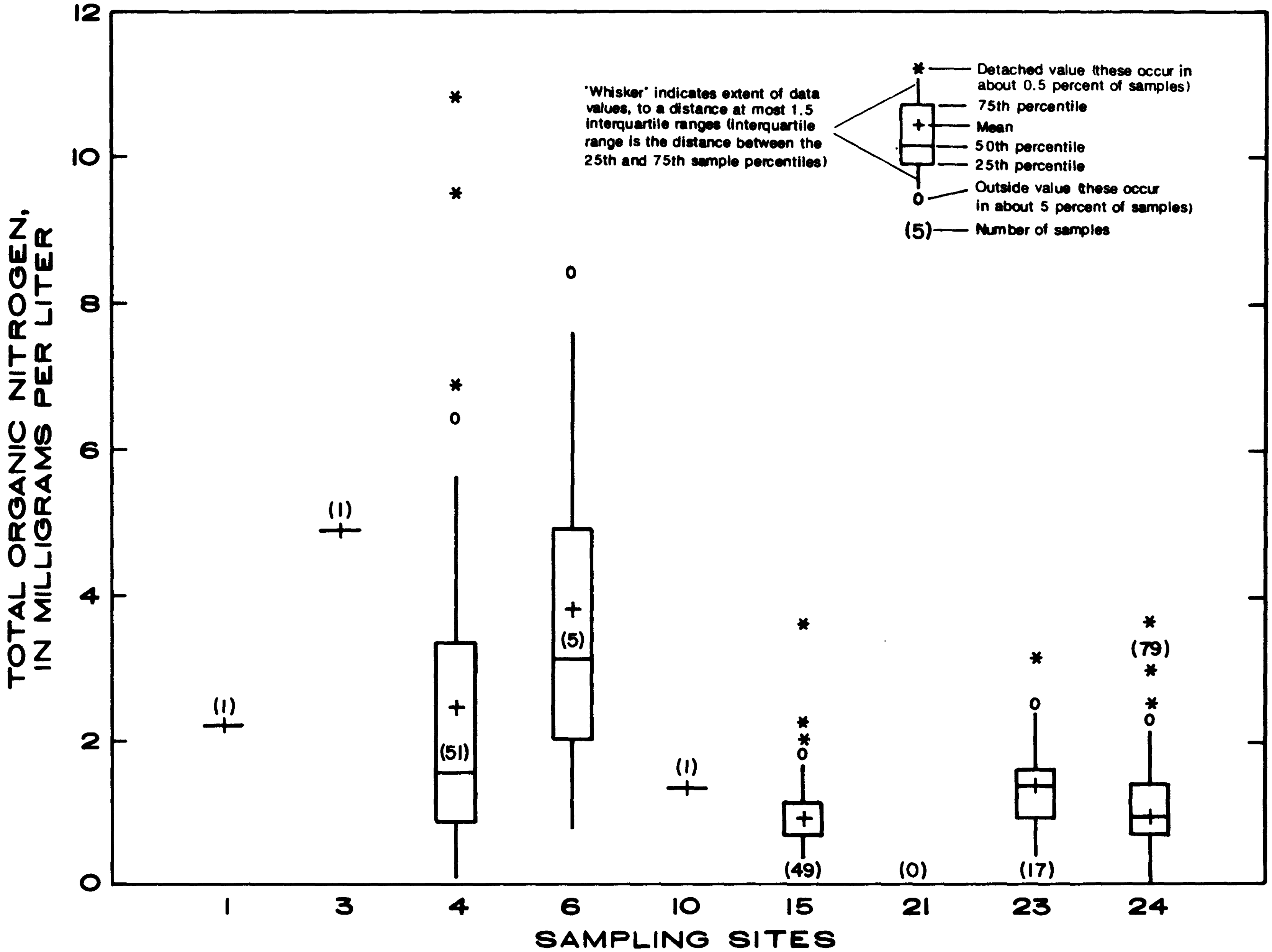


Figure 7.--Total organic nitrogen, October 1974 through September 1982.

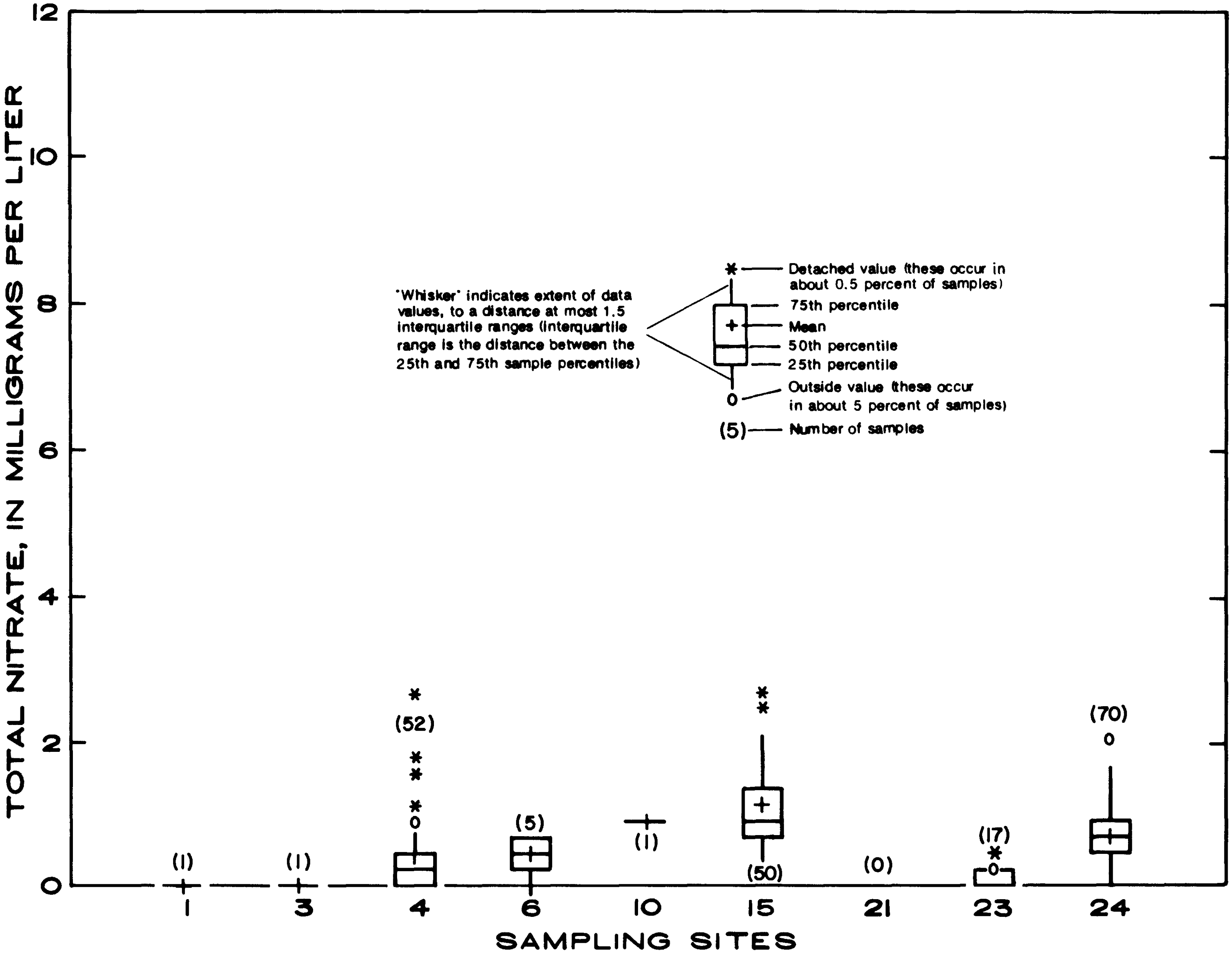


Figure 8.--Total nitrate nitrogen, October 1974 through September 1982.

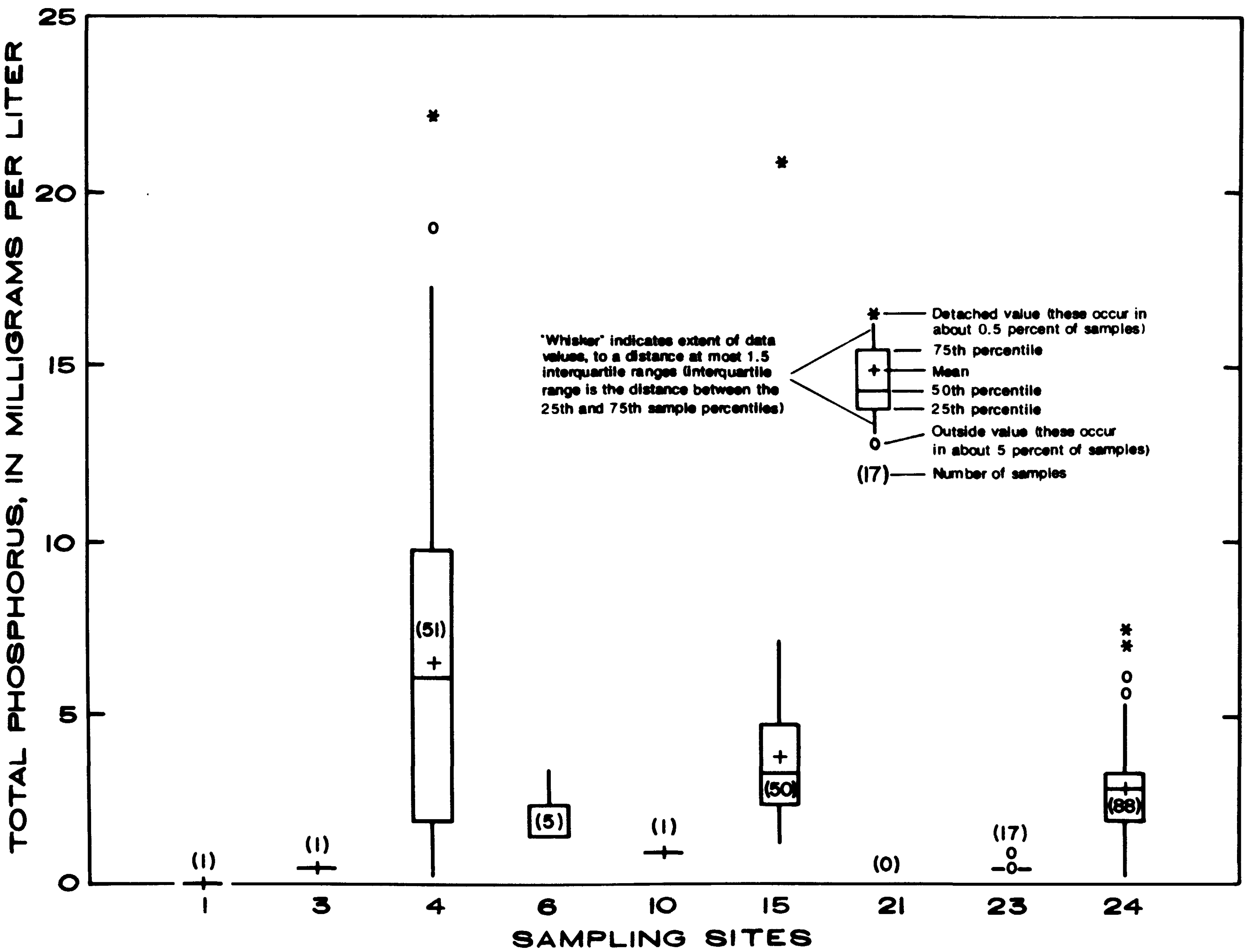


Figure 9.--Total phosphorus, October 1974 through September 1982.

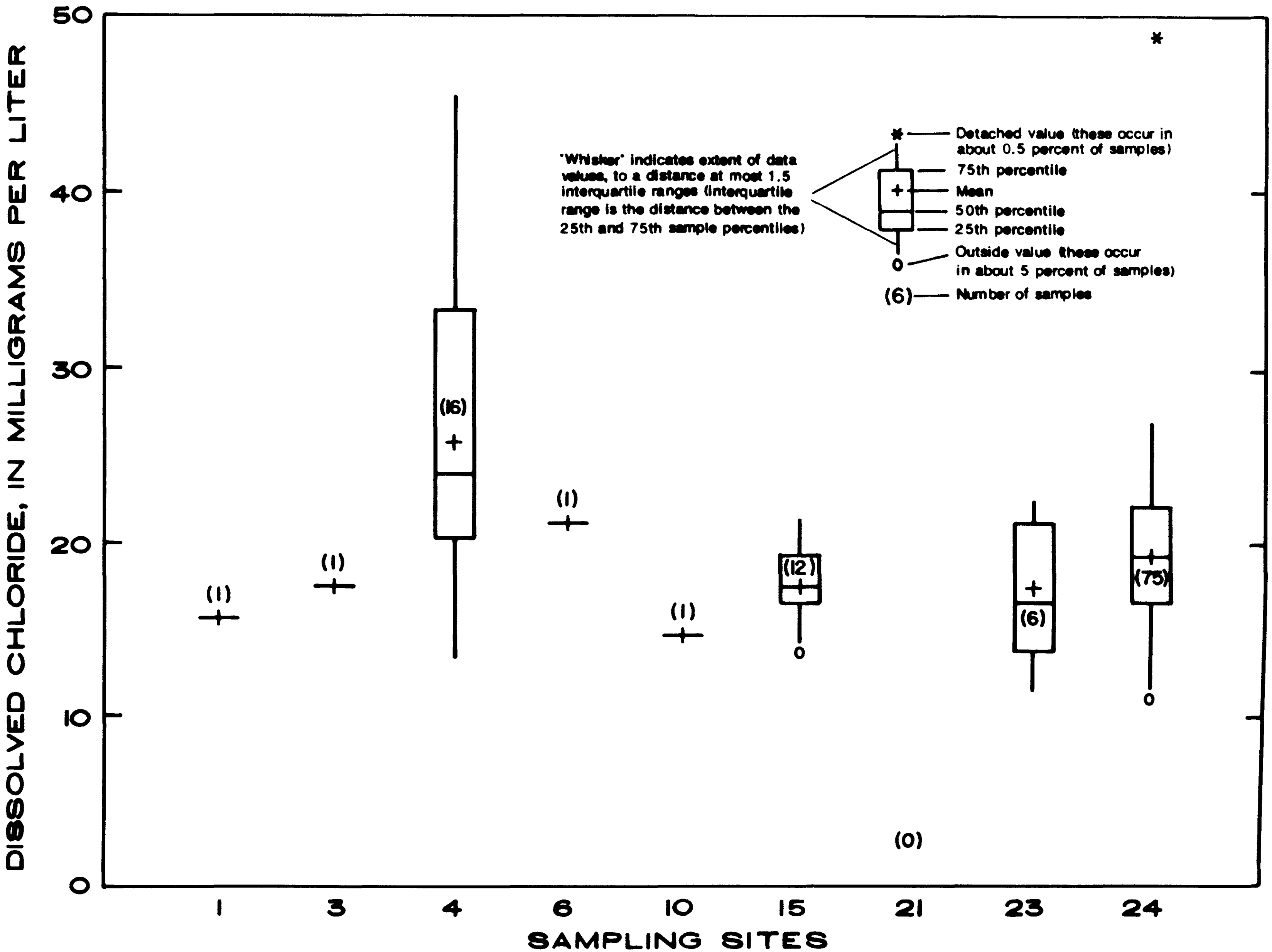


Figure 10.--Dissolved chloride, October 1974 through September 1982.

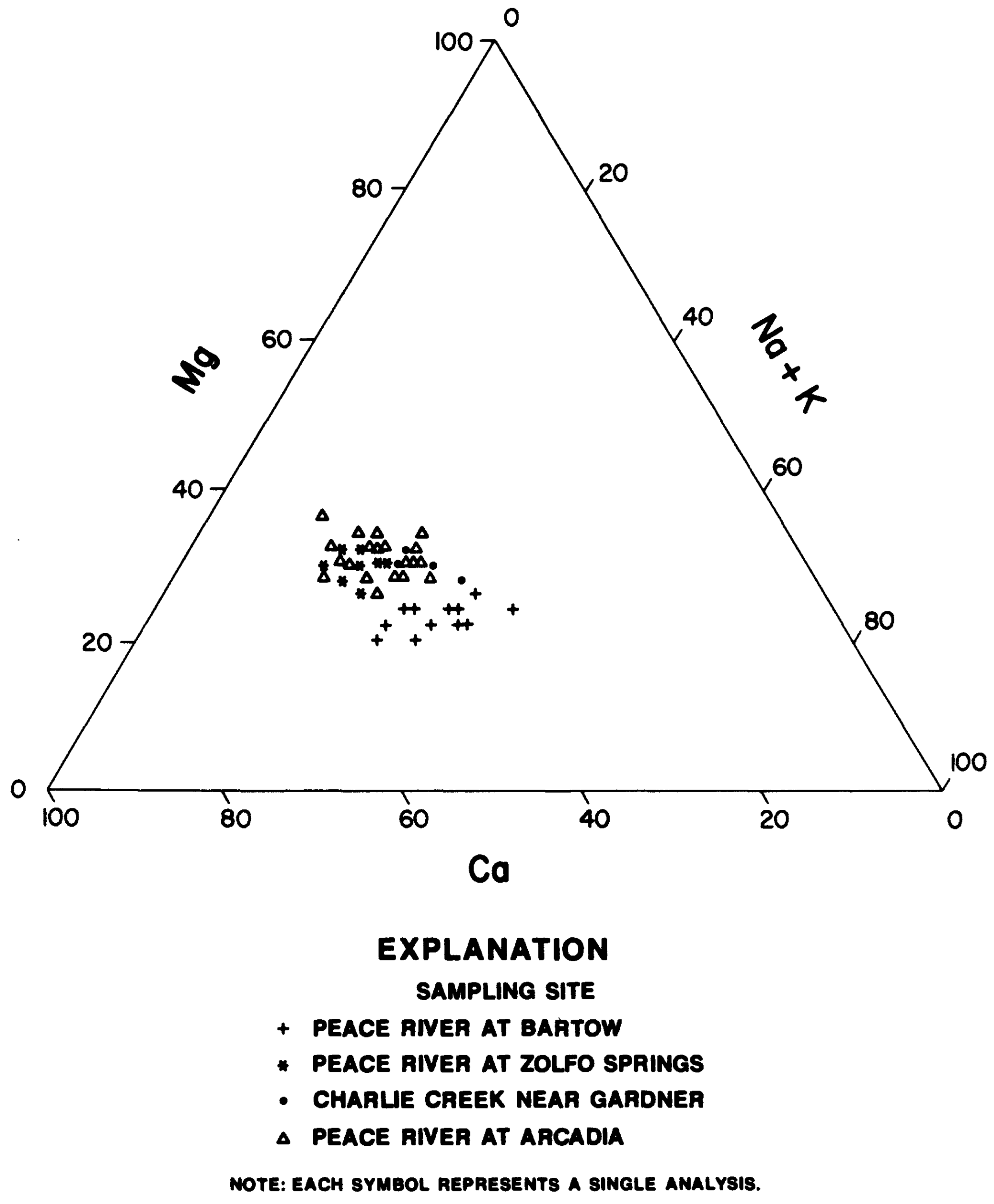


Figure 11.--Major cations at four sites, October 1974 through September 1982.

Major anion concentrations are plotted in figure 12. The composition of anions generally do not vary much between stations. The predominant anion is sulfate, ranging from 40 to 80 percent, followed by the carbonate bicarbonate group, which ranges from 20 to 60 percent. Chlorides make up only 20 to 40 percent of the anions. Limestone in the basin, particularly in mining areas and in the streambed, contribute to the higher percentage of carbonates and bicarbonates. The four analyses for anions shown for Charlie Creek (dot symbol) were too scattered to group together, but the two analyses in the lower third of the graph at about 20-percent sulfate would indicate a significant difference in the water from this tributary, specifically in the percentage of sulfate. The carbonate and bicarbonate percentage at the Peace River at Bartow (plus sign) tends to be lower than that of the other sites. The data for all sites are nearly centered in the diagram, but the lower chloride concentrations cause a shift to left of center.

Major anions and cations are combined for plotting in figure 13. The carbonate bicarbonate ion in each of the samples varies from 20 to 50 percent of the total anions, whereas the sulfate and chloride ions vary from 40 to 80 percent. The cations are predominantly calcium and magnesium, varying from about 60 to 80 percent of the total cations in each sample. The analyses for the Peace River at Bartow are dissimilar to the other stations shown (shifted to right on plot) because of the higher sodium and potassium concentration. This may be due to more agricultural runoff at the site, or to contribution of calcium downstream because of dissolution of limestone in the streambed, or to ground-water inflow.

## State Water-Quality Standards--Conformance of Surface Water to Standards Within the Peace River Basin

Water-quality data for eight sites were compared with standards set by the FDER (Florida Department of Environmental Regulation) for Class III waters (recreation, propagation, and maintenance of a healthy well-balanced population of fish and wildlife). The comparison was done to determine whether exceedance of standards in the upper parts of the basin would be reflected in exceedance at the NASQAN station (site 24). Conversely, if values for selected water-quality parameters exceed standards at the NASQAN station would this be an indication that data at a station upstream have exceeded standards?

No numeric standards exist for nutrients, only the general statement that "In no case shall nutrient concentrations of a body of water be altered so as to cause an imbalance in natural populations of aquatic flora or fauna" (Florida Department of Environmental Regulation, 1983). Total phosphorus concentrations within the basin vary widely as was seen in figure 9. Mackenthun (1973) has suggested a desired goal of 0.10 mg/L total phosphorus for "the prevention of plant nuisances in streams or other flowing waters not discharging directly to lakes or impoundments." Nearly all the values for the eight sites studied exceeded this limit (table 2).

A suggested limit of 1.0 mg/L was used for analysis of nitrogen values for this study (Dean Jackson, Florida Department of Environmental Regulation, oral commun., 1982). The percentage of samples exceeding this limit varied among the sites, from zero (all total nitrogen concentrations less than 1.0 mg/L) to 83 percent (33 of 40 samples had a total nitrogen concentration greater than 1.0 mg/L) at Peace River at Zolfo Springs (site 15). The

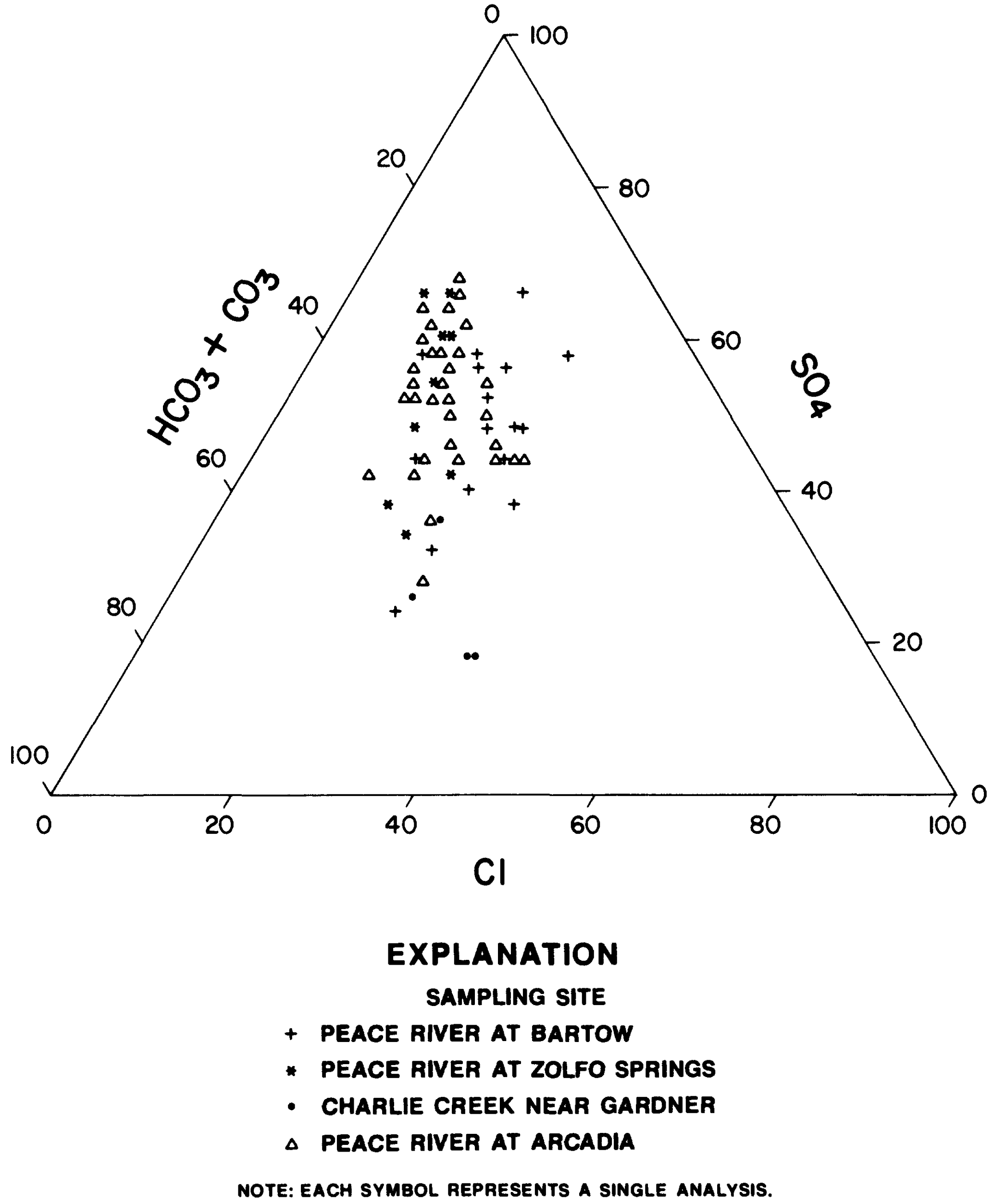


Figure 12.--Major anions at four sites, October 1974 through September 1982.

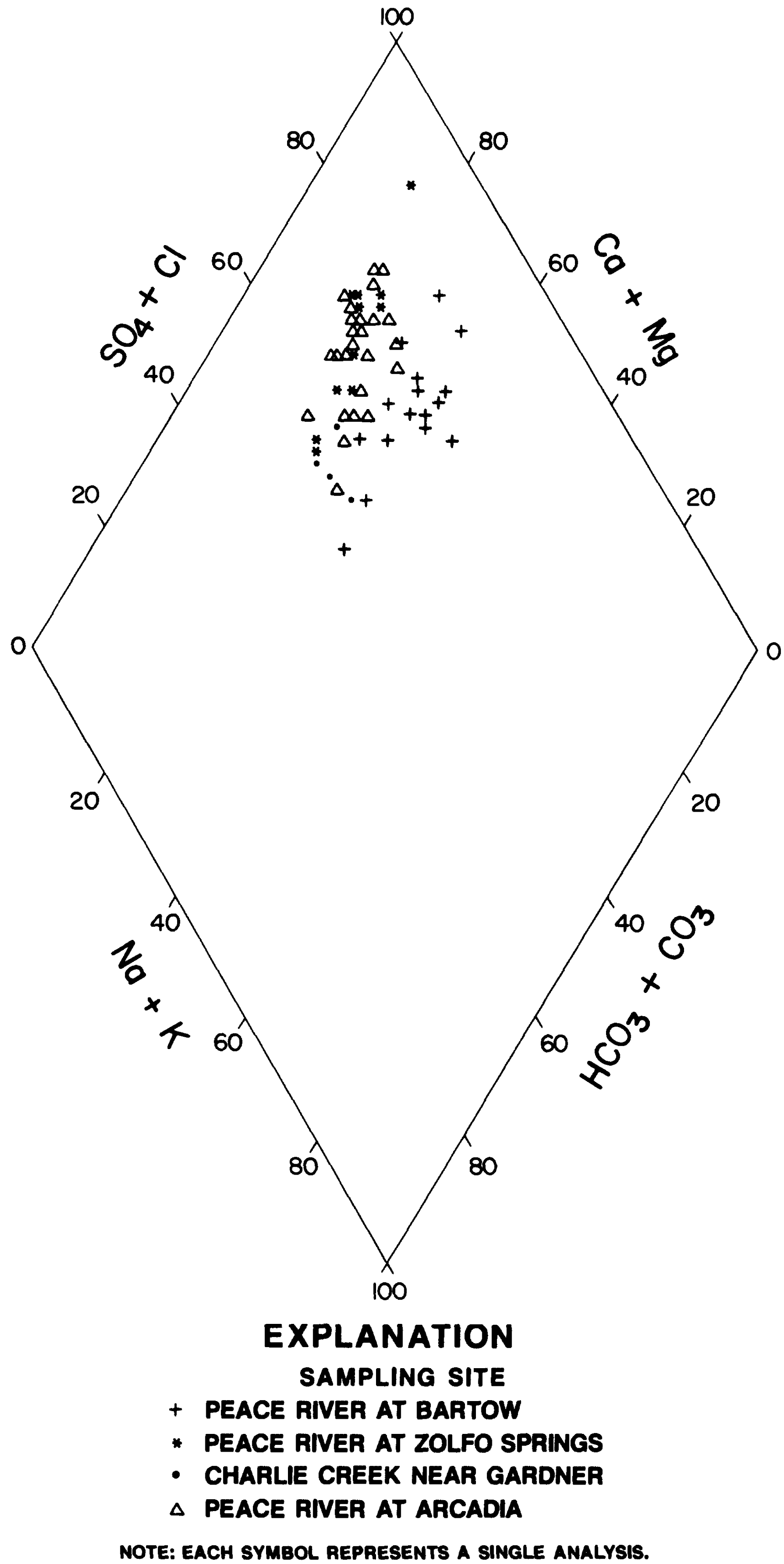


Figure 13.--Major cations and anions at four sites, October 1974 through September 1982.

Table 2.--Summary of number of samples that exceed Florida Department of Environmental Regulation standards and suggested standards for selected constituents, October 1974 through September 1982

[Number of samples exceeding criteria/number of samples. Criteria, indicated by parentheses, are in milligrams per liter and are included in regulations except for phosphorus and nitrogen. Criteria used here for phosphorus and nitrogen are not part of Florida Department of Environmental Regulation regulations, but have been suggested for eutrophication control (Dean Jackson, Florida Department of Environmental Regulation, oral commun., 1982)]

| Site No. | Phosphorous, total (0.1) | Nitrogen, total (1.0) | Lead, total recoverable (0.03) | Zinc, total recoverable (0.03) | Aluminum, total recoverable (1.5) |
|---|---|---|---|---|---|
| 1 | 1/1 | 0/2 | -- | -- | -- |
| 3 | 1/1 | 0/2 | -- | -- | -- |
| 4 | 53/55 | 9/40 | 1/17 | -- | 0/16 |
| 6 | 5/5 | 2/6 | -- | -- | -- |
| 10 | 1/1 | 2/2 | -- | -- | -- |
| 15 | 53/53 | 33/40 | 0/2 | -- | 0/2 |
| 22 | 17/17 | 1/20 | 0/4 | -- | 0/4 |
| 24 | 94/94 | 40/63 | 2/37 | 11/32 | 0/7 |

percentage of samples exceeding 1.0 mg/L total nitrogen at the NASQAN station (site 24) was 63 percent (40 of 63 samples). The lower total nitrogen concentrations at the NASQAN station may be because of the discharge of water low in nitrogen concentration from Charlie Creek (site 22). Total nitrogen at the Charlie Creek station exceeded 1.0 mg/L in only 5 percent (1 of 20) of the samples. These varying conformances to the suggested nitrogen standard are an indication that samples collected at the NASQAN station are not indicative of conformance to standards throughout the basin. Both higher and lower percentages of samples exceeded the limit elsewhere in the basin.

Few samples for metals were available for comparison to standards, except at the NASQAN station. At that station, lead exceeded the standard in 5 percent (2 of 37) of the samples, and zinc exceeded the standard in 34 percent (11 of 32) of the samples (table 2). These data indicate that water quality may sometimes not conform to standards at the NASQAN station. Whether or not this lack of conformance is representative of other sites in the basin cannot be determined because of the small number of samples for these metals at the other sites.

## Statistical Comparisons of Water-Quality Data at Selected Sites

The comparisons of water quality among selected sites involved two steps. First, the distribution of the water-quality data at each site was tested for normality, as a prerequisite to selecting the appropriate statistical procedure for further data analysis. Next, the data were analyzed to determine if water quality at all stations in the basin could be considered to be identical with a selected degree of probability.

Test for normality of statistical distributions.--An analysis was made of the distributions for each variable and station by using the UNIVARIATE procedure (SAS Institute, Inc., 1982a). The main objective of this analysis was to determine if the data (or natural-log transformed data) were normally distributed, and if the distributions were similar. Knowledge of data distribution is a necessary prerequisite to choosing statistical techniques for further analysis. Insufficient data prevented the use of the procedure for some variables, but most of the nine stations had data for at least one of the following variables: specific conductance, total nitrogen, nitrate nitrogen, total phosphorus, orthophosphorus, iron, lead, and dissolved solids. Two test procedures are used, depending on the number of analyses in the distribution. For a sample size of 50 observations or less, the Shapiro-Wilk test statistic, W, is used. For sample sizes larger than 50 observations, the Kolomogorov-Smirnov D-statistic is used (SAS Institute, Inc., 1982a). A probability level of 0.05 was used for the test of the hypothesis that the data values (by station and variable) are a random sample from a normal distribution.

Table 3 summarizes the results of the normality test on data for nine sites in the Peace River basin. All the selected sites in the table had data for specific conductance, but only four of the nine sites had data for all the other variables included in the table. The use of a natural log transformation of the data increased the number of stations having data that failed to reject the null hypothesis of the normality test, possibly due to the moderation of extremes by the transformation. Six out of nine sites had normally-distributed specific conductance values when the data were transformed, but the site with the most observations, Arcadia, did not have normally-distributed values, even after the transformation of the data.

The overall conclusion drawn from the normality testing is that the assumption that water-quality data are normally distributed cannot be made. Therefore, statistical tests comparing distributions must be restricted to those tests that do not depend on the underlying assumption that the sample data are normally distributed. For this reason, nonparametric tests were used to compare data among stations.

Tests for differences in water quality between sites.--Duncan's multiple-range test was used as the initial analysis of variance test (SAS Institute, Inc., 1982b). In this test, the means of the data for each site and variable are computed and tested for differences. Those means that are not (statistically) significantly different are grouped together. One site may fit in more than one group, or it may be in a group by itself. This test, however, is parametric, and has the underlying assumption that the data are normally distributed. On the basis of the previous analysis of the distributions, this test would not be valid for any of the water-quality variables except specific conductance. Six of the nine sites tested had passed the normality test, and these can be used in the Duncan test. The NASQAN station (site 24) was included in the test also, although the null hypothesis was rejected in the normality test, because the main objective of using the test was the comparison of the NASQAN station to the other stations within the basin.

Table 3.--Results of normality tests on raw data and logs of data for nine stations, October 1974 through September 1982

[Results are shown only for stations at which more than 10 values of the indicated parameter were available]

| Site No. | Specific conduct- ance | Lead, total | Iron, total | Dis- solved solids | Nitro- gen, total | Ni- trate, as N | Phos- phorus, total | Ortho- phos- phorus |
|---|---|---|---|---|---|---|---|---|
| 1 | 21 R F | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 3 | 43 R F | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 4 | 66 R F | 17 R R | 17 R R | 16 F F | 54 R F | 56 R R | 55 R R | 56 R R |
| 6 | 65 F F | 0 | 0 | 1 | 5 | 5 | 5 | 5 |
| 10 | 21 F R | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 15 | 66 F F | 2 | 2 | 10 | 51 R F | 53 R F | 54 R F | 53 R F |
| 21 | 53 R R | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 | 63 F F | 4 | 4 | 4 | 17 F F | 17 R R | 17 F F | 17 F F |
| 24 | 103 R R | 40 R R | 51 R R | 80 F R | 73 R F | 94 R R | 94 F R | 75 F R |

NOTE: Null hypothesis: Input data values are a random sample from a normal distribution.

F = Fail to reject null hypothesis.
R = Reject null hypothesis.

54 R F = number of analyses, result of test on raw data, and result of test on log transformation of data.

When the data do not meet the assumptions for parametric analysis of variance, nonparametric tests may be used. Procedures available in SAS were used to compare the data for the nine stations in table 1. These tests evaluate the probability that data for the entire group of stations are from the same population, but do not indicate the number of different populations that might be present, or provide a grouping of stations that have the same distribution of data. Identification of groups is possible using nonparametric procedures, but is not provided by the SAS procedures used. Were the study to be continued, nonparametric grouping of the data would be the next logical step in analyzing the data.

Results of the Duncan multiple-range testing, as used to determine grouping of specific conductance at seven sites, are shown in table 4. Five groupings were found for the seven sites, indicating the variability of the data within the basin. The mean value shown is the mean of the log values; the mean specific conductance ranged from 166 $\mu$S/cm for group E to 402 $\mu$S/cm for group A. The groups in which the station at Arcadia is included are groups B and C. In group C, the two sites in the upper reach of the basin, Saddle Creek (site 3) and Peace River at Bartow (site 4), were significantly different from the NASQAN station (site 24). In group B, Peace River at

Table 4.--Results of Duncan's multiple range test on specific conductance values, October 1974 through September 1982

[Means with the same Duncan grouping letter are not significantly different at a 5-percent significance level]

| Duncan grouping | Mean | Number of samples | Site |
|---|---|---|---|
| A | 6.0155 | 66 | 15 |
| B A | 5.9786 | 65 | 6 |
| C B | 5.8435 | 66 | 4 |
| C B | 5.8283 | 104 | 24 |
| C | 5.7699 | 43 | 3 |
| D | 5.3663 | 25 | 1 |
| E | 5.1141 | 63 | 22 |

Fort Meade (site 6) replaces Saddle Creek in the group. In either case, four of six sites possible were not included in a group with the NASQAN station, indicating that the means were significantly different. Therefore, the specific conductance at the Arcadia NASQAN station is not representative of specific conductance at all sites within the basin, and does not represent the entire basin.

Two nonparametric tests were used for further analysis of the data, including data found not to be normally distributed. These two tests, the Brown-Mood median test and the Kruskal-Wallis test, from the SAS procedure NPAR1WAY (SAS Institute, Inc., 1982b), perform an analysis of variance on ranks and rank scores (median test) and ranks (Kruskal-Wallis).

The median test (Brown-Mood) was designed to examine whether several samples come from populations having the same median (Conover, 1971). A "grand median" is computed based on all the samples (all sites), and each value (by water-quality variable) is then compared to this median. The number of values above and below this median are tallied for each station and a contingency table is constructed. A chi-square test statistic is then computed. The null hypothesis and its alternate for this test are as follows:

$H_0$: All populations have the same median.

$H_1$: At least two of the populations have different medians.

The major disadvantage of this test is that it is based only on whether the value is above or below the median, and does not consider how far above or below the median it is.

The null hypothesis and alternate hypothesis of the Kruskal-Wallis test are as follows:

$H_0$: All of the populations distribution functions are identical (shape and mean),

$H_1$: At least one of the populations tends to yield higher values than at least one of the other populations, (or)

The k populations do not have identical means.

The Kruskal-Wallis test uses more information obtained from the observations than does the median test. The test statistic is a function of the ranks of the observations in the combined sample (all sites).

The results of the application of the median test and the Kruskal-Wallis test are summarized in table 5. The two tests agree in rejecting or failing to reject their null hypotheses, except for dissolved solids. The median test indicated that the data for the five stations had the same median value, but the Kruskal-Wallis test indicated that the populations for each of the five sites were not identical. Each test provides a different type of information about the population, but the Kruskall-Wallis test tells more, and is a more powerful test because it is more discriminating in detecting differences in ranks of data (Daniel, 1978).

Table 5.--Results of nonparametric statistical tests for selected water-quality variables, October 1974 through September 1982

| | Number of stations | Number of observations | Median test results[1] | Kruskal-Wallis result[2] |
|---|---|---|---|---|
| Specific conductance | 9 | 480 | Reject | Reject |
| Nitrogen, total | 7 | 202 | do. | Do. |
| Nitrate, total | 7 | 227 | do. | Do. |
| Phosphorus, total | 7 | 226 | do. | Do. |
| Iron, total recoverable | 4 | 74 | Fail to reject | Fail to reject |
| Lead, total recoverable | 4 | 63 | do. | Do. |
| Orthophosphorus, total | 7 | 208 | Reject | Reject |
| Dissolved solids | 5 | 111 | Fail to reject | Do. |

[1]Reject indicates at least two stations have different medians. Fail to reject indicates there is insufficient evidence to show that the medians are significantly different.

[2]Reject indicates at least one station has higher values than at least one other station. Fail to reject indicates there is insufficient evidence to show that the means are significantly different.

For all other water-quality variables considered, the null hypothesis was rejected, with two exceptions--total lead and total iron. The smaller number of sites having data tested (four) and observations used in the tests (63 and 74, for lead and iron, respectively) probably increased the likelihood of failing to reject the null hypothesis. In general, larger samples increase the sensitivity of statistical tests. Both of the nonparametric tests agree with the parametric Duncan means test in indicating the specific conductance values at four different sites in the basin are not all from the same population.

The results of these statistical tests indicate that water-quality data for the NASQAN station at Arcadia do not represent the entire basin for most water-quality variables. The Duncan multiple range test, which identifies groupings of data, should not be used for most water-quality variables because of non-normal distribution of the data. The nonparametric tests used do not group data, though modification of the tests could be used to do so. These modifications were beyond the scope of this study, and further categorization of water quality in the basin with these data was not attempted.

The statistical procedures used here are not the only ones applicable to comparing data among sites. Other techniques, not used here but potentially useful, include data ranking followed by analysis of variance and the Duncan means test. The ranking of the data removes the restriction of normality.

## Chemical Constituent Loads and Basin Yields

The use of constituent loads and basin yields is another method of examining the uniformity of water quality in the basin. Comparison of loads and yields (rather than concentrations) has the advantage of considering the effect of stream discharge on water quality. Possible sources and sinks of constituents may be located using the estimated loads and yields at sampling points in the basin.

### Relation of water-quality properties and stream discharge

The relation of water physical properties and chemical constituents to stream discharge is of interest because it not only helps to describe variation in water quality but also is a tool for further analysis. Least-squares regression procedures were used to investigate quantitatively the relation between the water-quality variables, such as specific conductance, total nitrogen and total phosphorus, and stream discharge.

Regression analysis requires initial selection of a model, that is, a functional relation describing the dependent variable (water quality) in terms of the independent variable (stream discharge). The choice of a model is usually not obvious, but must be done on a "trial and error" basis for each station. Smith and others (1982) describe some commonly used models which they evaluated for describing total phosphorus-stream discharge relations. These functions, and some extensions of these functions, were evaluated for use in describing water quality-streamflow relations for stations with sufficient record in the Peace River basin.

The functions evaluated in this study (and listed below) are those suggested by Crawford and others (1983, p. 11-12), where Y equals concentrations (or specific conductance), Q equals discharge, and the coefficients a, b, and c are computed by the least-squares method. Logarithm (base e) transformation of the data is indicated by log(Y) or log(Q).

$$Y = a + b(Q): \text{ linear} \quad (1)$$

$$Y = a + b(\log(Q)): \text{ linear-log} \quad (2)$$

$$Y = a + b/(1 + c^{*}(Q)): \text{ hyperbolic} \quad (3)$$

$$Y = a + b/Q: \text{ inverse} \quad (4)$$

$$Y = a + b(Q) + c(Q^2): \text{ quadratic} \quad (5)$$

$$\log(Y) = a + b(\log(Q)): \text{ log-log} \quad (6)$$

$$\log(Y) = a + b(\log(Q)) + c(\log(Q))^2: \text{ log-log quadratic} \quad (7)$$

The coefficient $c^{*}$ in equation 3 was not computed by least squares, but was selected in eight increments from $c^{*} = 10^{-x}$ to $c^{*} = 10^{(-x + 3.5)}$, where x is equal to 2.5 + integer of $\log_{10}$ of mean discharge. This provided 14 possible models.

Selection process for the best of the 14 models at each station was to first compare the coefficients of determination ($R^2$ among the functions that do not use logarithms of Y (equations 1 through 5). The model with the highest $R^2$ is probably the best in explaining overall variations of Y as a function of discharge. Coefficients of determination for the log(Y) models (equations 6 and 7) cannot be compared with those for the non-log(Y) models because the log transformation of Y affects the distribution of Y, and thus the numerical value of $R^2$. The only way to compare log(Y) models with Y models is to examine the ratio of standard errors of estimate to the mean value of Y (called the coefficient of variation), and also to examine plots of Y and log(Y) against discharge to see if variation of Y and log(Y) are constant throughout the range of discharge. In these comparisons, the model with the lowest coefficient of variation and the most constant variation of Y or log(Y) would be preferred.

Table 6 summarizes the best models for relating specific conductance to discharge at the eight sites with sufficient periodic data in the Peace River basin. The best specific conductance model was either the quadratic or the hyperbolic form, and the best log specific conductance model was generally the quadratic. Comparing the coefficients of variation for log models and nonlog models shows that, except at site 22, there is little reason to select one form over the other.

A plot of specific conductance against discharge on log and linear scales is shown in figure 14 for site 6. This plot, typical of those for other sites, shows the lack of constant variation of specific conductance over the range of discharge (see plot on linear scale). Few data are available for discharges greater than 300 $ft^3/s$, but it seems likely that variation in specific conductance is much less at these high discharges than at lower discharges. Variation of logs of specific conductance (see plot on log scale), however, is much more constant over the range of discharge. For this reason, the log model is preferred over the linear model.

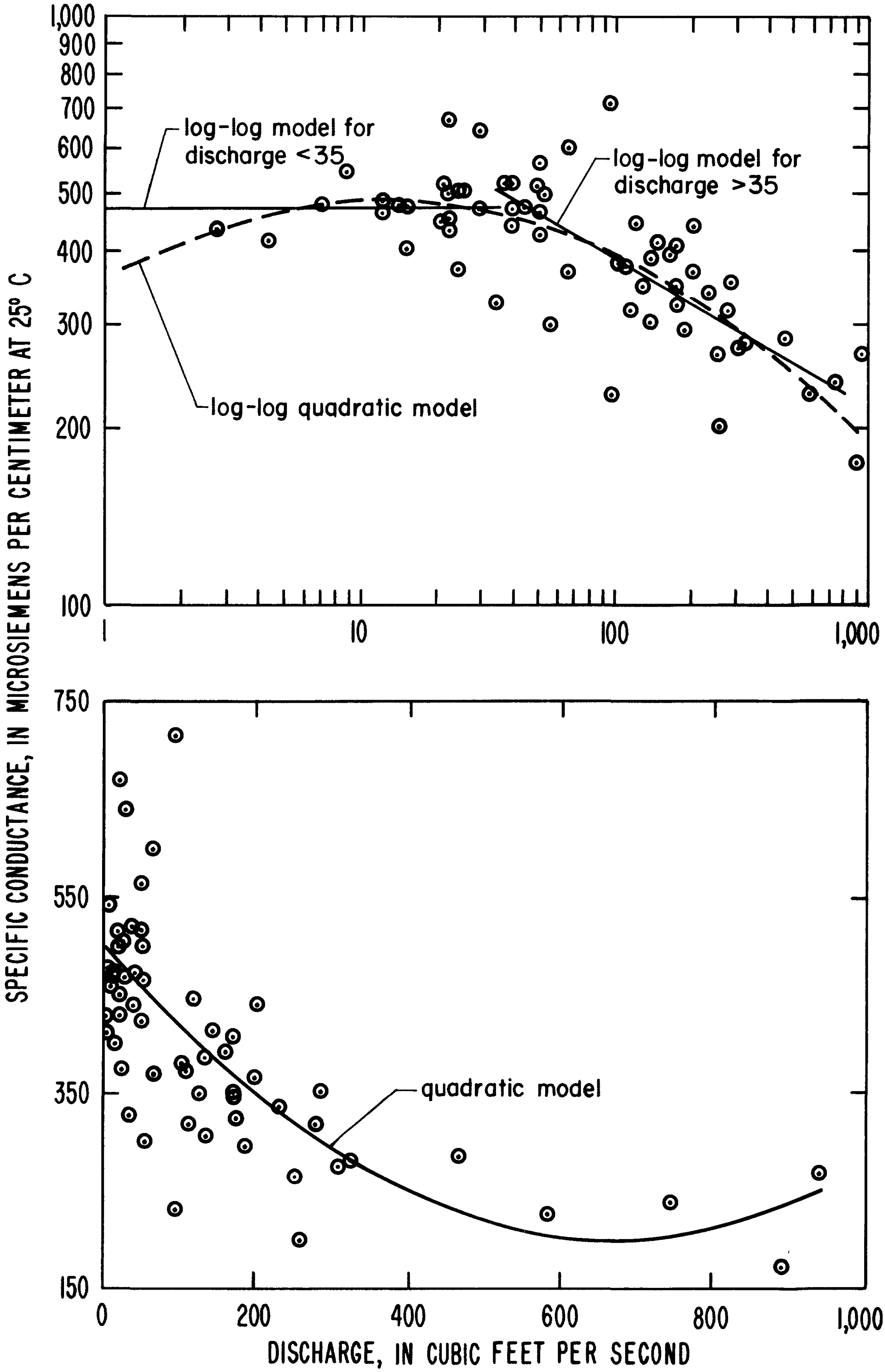


Figure 14.--Relations between specific conductance and stream discharge for Peace River at Fort Meade (site 6), October 1974 through September 1982.

Table 6.--Summary of best regression models for specific conductance as a function of stream discharge

[$R^2$ is the coefficient of determination; CV is the coefficient of variation, in percent]

| Site No. | Best regression model for specific conductance: Model type | $R^2$ | CV | Best regression model for log specific conductance: Model type | $R^2$ | CV |
|---|---|---|---|---|---|---|
| 3 | Quadratic | 0.17 | 21 | Log-log | 0.14 | 19 |
| 4 | Hyperbolic | .48 | 33 | Log-log quadratic | .52 | 32 |
| 6 | Quadratic | .54 | 19 | do. | .62 | 19 |
| 10 | do. | .74 | 9.2 | do. | .78 | 9.4 |
| 15 | Hyperbolic | .65 | 15 | do. | .66 | 15 |
| 21 | ([1]) | | | ([1]) | | |
| 22 | Hyperbolic | .60 | 27 | Log-log quadratic | .51 | 42 |
| 24 | do. | .64 | 21 | do. | .72 | 22 |

[1]No models were statistically significant at the 0.05 significance level.

The best-fit models for both log transformed and untransformed data were of the quadratic form. However, figure 14 reveals some problems with the quadratic model. The fit of nonlog data tends to overestimate specific conductance in the 100 to 200 $ft^3/s$ range and underestimate specific conductance at discharges exceeding about 350 $ft^3/s$. Another problem is that quadratic equations have a point of inflection, that is, where the slope of the relation changes sign. This change of slope direction is hard to justify in terms of basin hydrology--that is, there is no obvious reason for specific conductance to be at a minimum at a discharge of about 650 $ft^3/s$ and higher at either greater or lesser discharges, as predicted by the nonlog model. The log-log quadratic model seems to fit the data over the entire discharge range, but predicts a maximum specific conductance at about 10 $ft^3/s$ and lower specific conductance at either greater or lesser discharge. These hydrologic implications, probably untrue, may not affect use of the model as a predictive tool, but it seems more desirable to use a model which behaves in a manner analogous to stream hydrology. Because of these problems, the log-log models are probably better for relating water quality to stream discharge, and, in some cases, two log-log models are required, each to represent a selected range of discharge.

This use of two log-log models is shown in figure 14. The data were divided into two subsets, with the boundary determined by visual inspection of the plotted data. Then each subset was fitted separately by least squares regression. A statistical test indicated that the slope of the log-log relation for the low discharge subset was not significantly different from zero at a 5-percent significance level. Therefore, the mean specific conductance of the subset was used for the entire range of discharge of the subset.

Coefficients of the log-log regression models for specific conductance, total nitrogen, and total phosphorus are shown in tables 7, 8, and 9, respectively. Also shown are coefficients of determination and coefficients of variation.

The slope coefficient (b) of the regression models indicates whether the dependent variable increases (positive b), decreases (negative b), or is unaffected by increasing discharge. Table 7 shows that for high discharge, specific conductance decreases with increasing discharge at most of the stations. Table 8 shows that total nitrogen concentrations are relatively constant at high discharges (b = 0), but tend to increase with discharge at relatively low discharges. Total phosphorus (table 9) generally decreases with increasing discharge or is unaffected by discharge.

Table 7.--Logarithmic regression models for specific conductance as a function of discharge

[Specific conductance = $e^x$, where x = a + b log (discharge); $R^2$ is the coefficient of determination; CV is the coefficient of variation in percent; and logs are base e]

| Site No. | Discharge boundary (log units) | High discharge | | | | Low discharge | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | a | b[1] | $R^2$ | CV | a | b[1] | $R^2$ | CV |
| 3 | 4.0 | 5.64 | 0 | -- | 44 | 5.80 | 0 | -- | 48 |
| 4 | None | 7.28 | -.35 | .66 | 53 | Same as for high discharge | | | |
| 6 | 3.6 | 7.12 | -.25 | .56 | 46 | 6.17 | 0 | -- | 38 |
| 10 | 3.7 | 7.09 | -.34 | .77 | 31 | 5.82 | 0 | -- | 29 |
| 15 | 5.0 | 7.44 | -.26 | .68 | 36 | 6.22 | 0 | -- | 46 |
| 21 | None | 4.41 | 0 | -- | 57 | Same as for high discharge | | | |
| 22 | 2.0 | 5.99 | -.21 | .68 | 49 | 5.61 | 0 | -- | 52 |
| 24 | 5.0 | 8.05 | -.37 | .86 | 39 | 6.11 | 0 | -- | 68 |

[1]A zero value indicates the b coefficient (slope) is not statistically different from 0 at a significance level of 5 percent.

## Loads and Basin Yields of Selected Chemical Constituents

Computation of mean chemical constituent loading requires daily stream discharge record as well as sufficient water-quality data to define the relation between chemical constituent and stream discharge. Seven sites in the Peace River basin have records of daily discharge for the 1974-82 water-year period. These sites all have records for at least 20 samples of specific conductance, probably sufficient to define the relation between specific conductance and stream discharge. Data for nitrogen and phosphorus species are not as numerous, and there were sufficient data for computation of mean load at only four of the seven sites.

Table 8.--Logarithmic regression models for total nitrogen as a function of discharge

[Total nitrogen (mg/L) = $e^X$, where X = a + b log (discharge); $R^2$ is the coefficient of determination; CV is the coefficient of variation in percent; and logs are base e]

| Site No. | Discharge boundary (log units) | High discharge | | | | Low discharge | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | a | b[1] | $R^2$ | CV | a | b[1] | $R^2$ | CV |
| 4 | 4.9 | 1.51 | 0 | -- | 82 | -1.08 | 0.55 | 0.40 | 72 |
| 15 | 4.7 | .79 | 0 | -- | 50 | .67 | 0 | -- | 71 |
| 22 | 4 | .70 | 0 | -- | 74 | -1.53 | .57 | .77 | 52 |
| 24 | 6 | .66 | 0 | -- | 65 | -1.48 | .35 | .26 | 73 |

[1]A zero value indicates the b coefficient (slope) is not statistically different from 0 at a significance level of 5 percent.

Table 9.--Logarithmic regression models for total phosphorus as a function of discharge

[Total phosphorus, mg/L, = $e^X$, where X = a + b log (discharge); $R^2$ is the coefficient of determination; CV is the coefficient of variation in percent; and logs are base e]

| Site No. | Discharge boundary (log units) | High discharge | | | | Low discharge | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | a | b[1] | $R^2$ | CV | a | b[1] | $R^2$ | CV |
| 4 | 3.5 | 4.64 | -0.76 | 0.41 | 104 | 1.70 | 0 | -- | 141 |
| 15 | 5.1 | 3.16 | -.35 | .53 | 50 | 2.60 | -.24 | .44 | 33 |
| 22 | None | -.65 | 0 | -- | 63 | Same as for high discharge | | | |
| 24 | 5 | 3.49 | -.42 | .72 | 52 | .11 | .28 | .32 | 58 |

[1]A zero value indicates the b coefficient (slope) is not statistically different from 0 at a significance level of 5 percent.

Calculation method.--Computed discharge duration curves for each station (fig. 15) give the discharge which was exceeded for any selected percentage of the days of record. Figure 15 shows flow duration curves for the seven Peace River basin sites over the 1 to 99 percent range, and in terms of unit discharge. For example, at site 15, daily discharge was higher than about 0.32 $ft^3/s/mi^2$ for 50 percent of the days. These discharge duration curves were computed for the period October 1974 through September 1982.

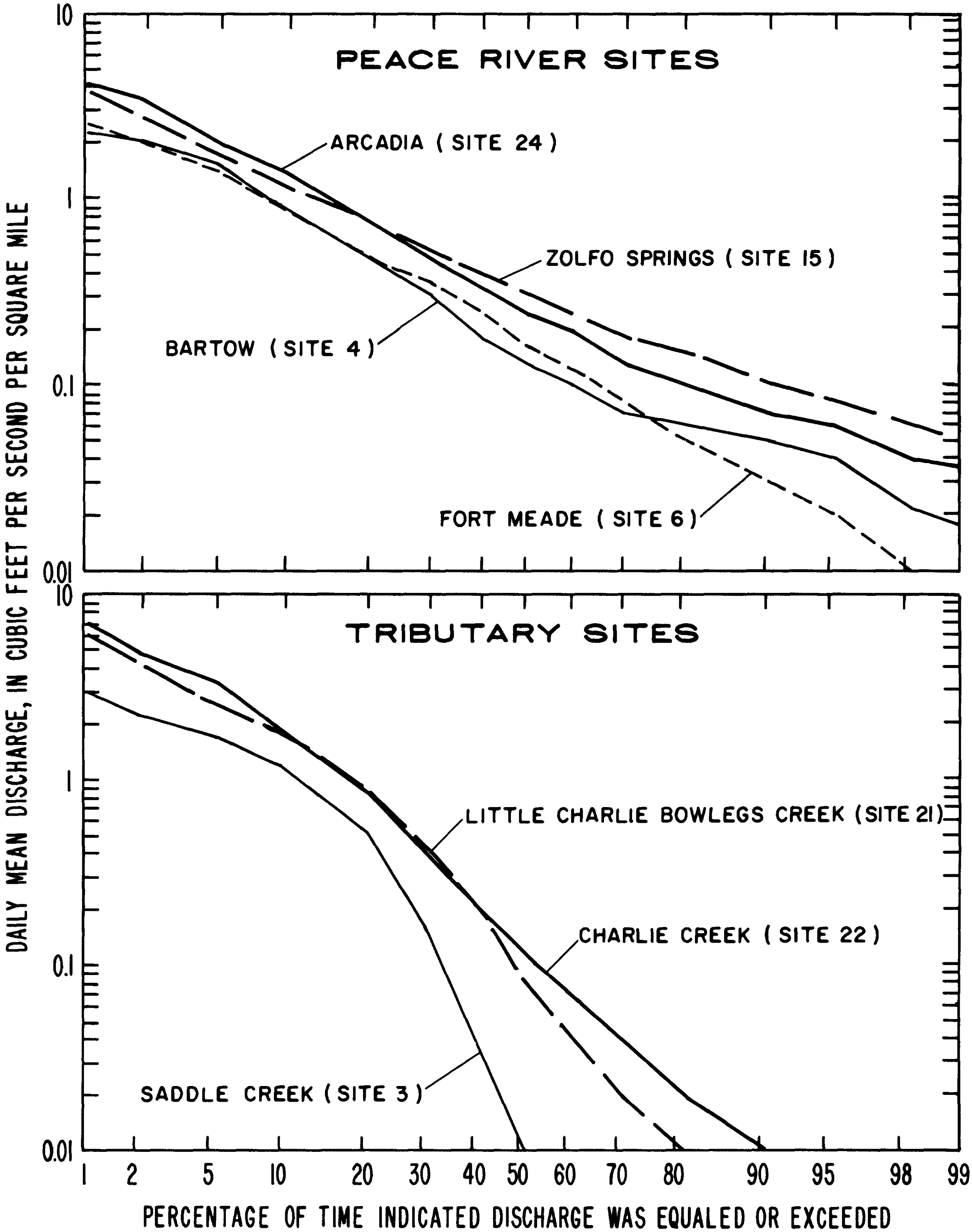


Figure 15.--Flow duration curves for the Peace River basin, October 1974 through September 1982.

Concentration-duration curves were computed from the discharge duration curves, using the regression relations to compute concentrations from discharges. Because the regression relations were in terms of logs of the dependent variables, it is necessary to correct for bias in the predicted values by adding in one-half the variance (in log units) before taking the antilog to obtain concentration values (Heien, 1968). Load duration curves are obtained by multiplying discharge by the computed concentration and a units-conversion factor.

The area under the load duration curve, which was calculated by using the trapezoidal method (Stark, 1970, p. 194), represents the mean load for the period of record used to develop the curve. Mean basin yield was computed by dividing mean load by drainage area.

Load-discharge relation for the Peace River.--The load-carrying characteristics at two Peace River sites are shown in figures 16 and 17 for dissolved solids, total nitrogen, and total phosphorus. Figure 16 shows the transport-duration curves for Peace River at Bartow (site 4), representing the upstream part of the basin, and figure 17 shows the curves for Peace River at Arcadia, the NASQAN station, (site 24) at the downstream end of the study area.

These curves show the percentage of the loads that are transported by discharges greater than the indicated discharge percentile. For example, figure 16 shows that discharges which were equaled or exceeded 50 percent of the time at Bartow (0.13 $ft^3/s/mi^2$ or greater, from fig. 15) transported about 79 percent of the dissolved-solids load.

Comparison of the curves for the three water-quality characteristics shows the relative effects of discharge on load. Relatively high discharges transport most of the loads. Half of the dissolved solids, phosphorus, and nitrogen loads are transported by discharges that are exceeded 31 or less percent of the time at Bartow (fig. 16), and 22 or less percent of the time at Arcadia (fig. 17). Of the three water-quality characteristics, total nitrogen has the highest percentage of the load carried by high discharge. Half of the nitrogen load is carried by discharges exceeded only 10 percent of the time at Bartow, and 11 percent of the time at Arcadia. This implies that direct runoff contributes relatively greater amounts of nitrogen than dissolved solids and phosphorus.

Mean unit stream discharge and basin yields of dissolved solids, total nitrogen, and total phosphorus.--Discharge and water-quality data from October 1974 through September 1982 were used to compute mean basin yields at sites in the Peace River basin. Dissolved solids concentrations were estimated by multiplying specific conductance by 0.68, the median ratio of dissolved solids to specific conductance for all samples from the seven sites for which dissolved solids yields were computed. Total nitrogen and phosphorus yields were computed for four sites; the other three sites have insufficient data (five or fewer samples) for definition of the concentration-discharge relation. Unit discharge (mean discharge divided by drainage area) and mean basin yields of dissolved solids, total nitrogen, and total phosphorus are shown in figure 18.

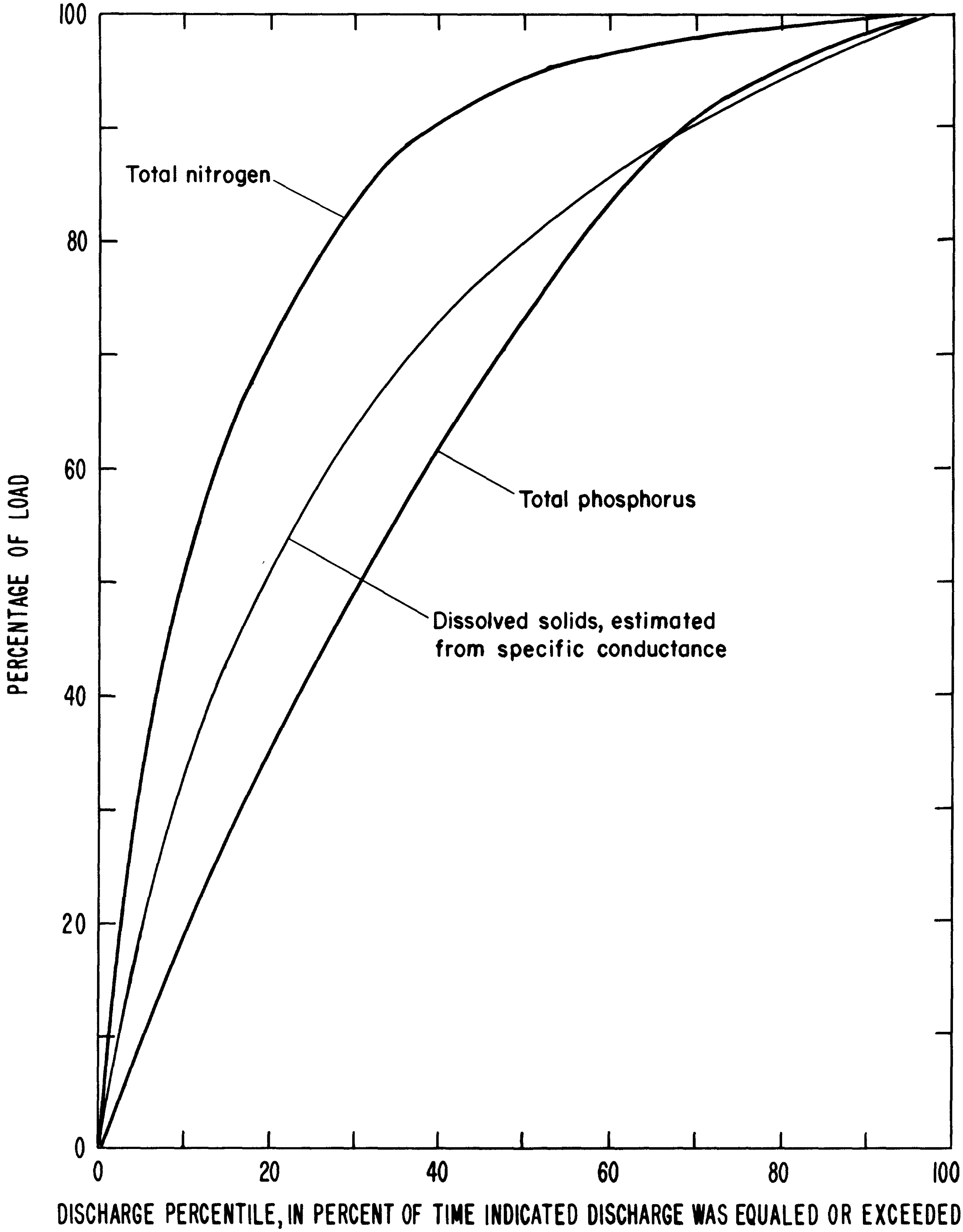


Figure 16.--Transport duration curves for loads of dissolved solids, total nitrogen, and total phosphorus, Peace River at Bartow (site 4), October 1974 through September 1982.

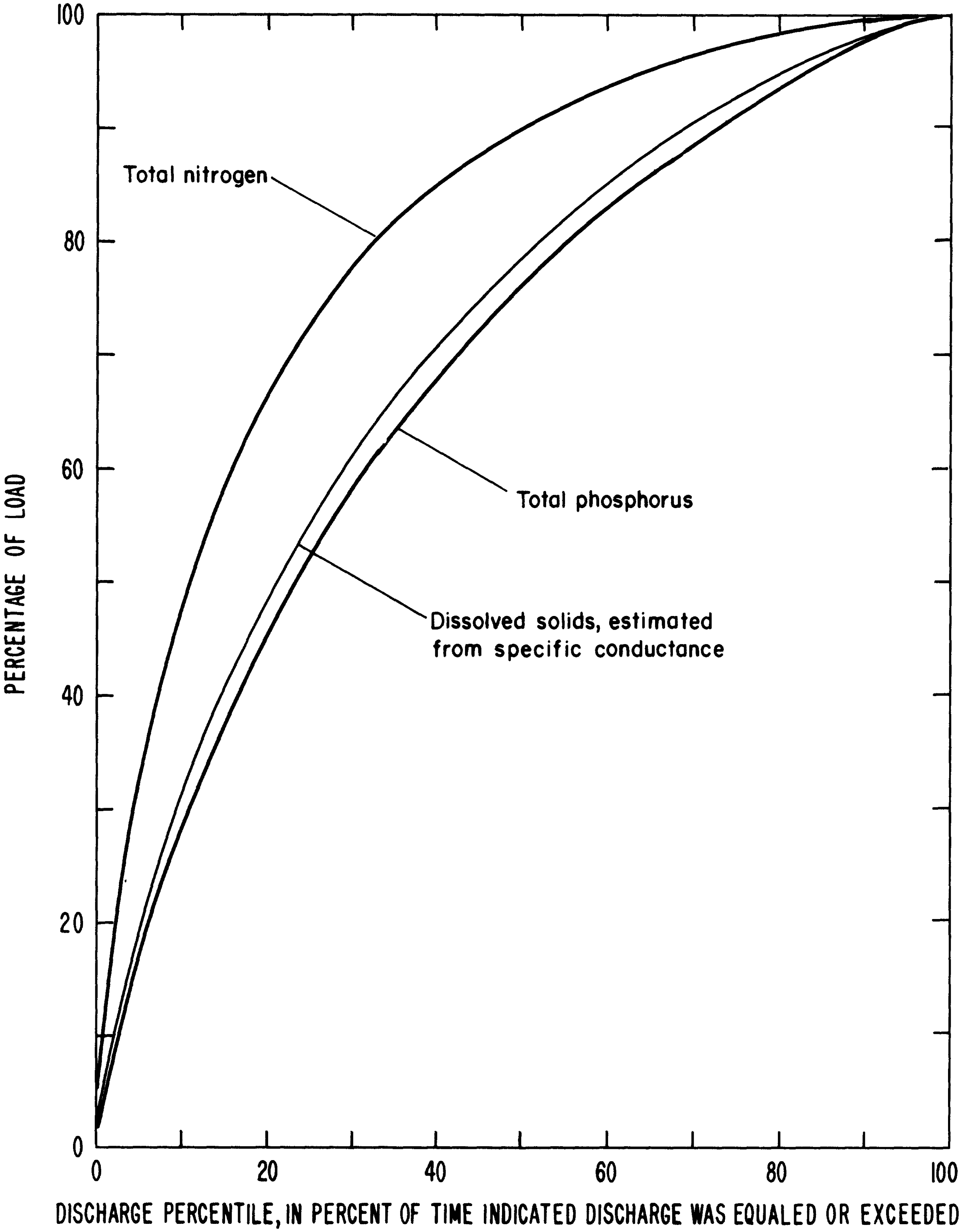


Figure 17.--Transport duration curves for loads of dissolved solids, total nitrogen, and total phosphorus, Peace River at Arcadia (site 24), October 1974 through September 1982.

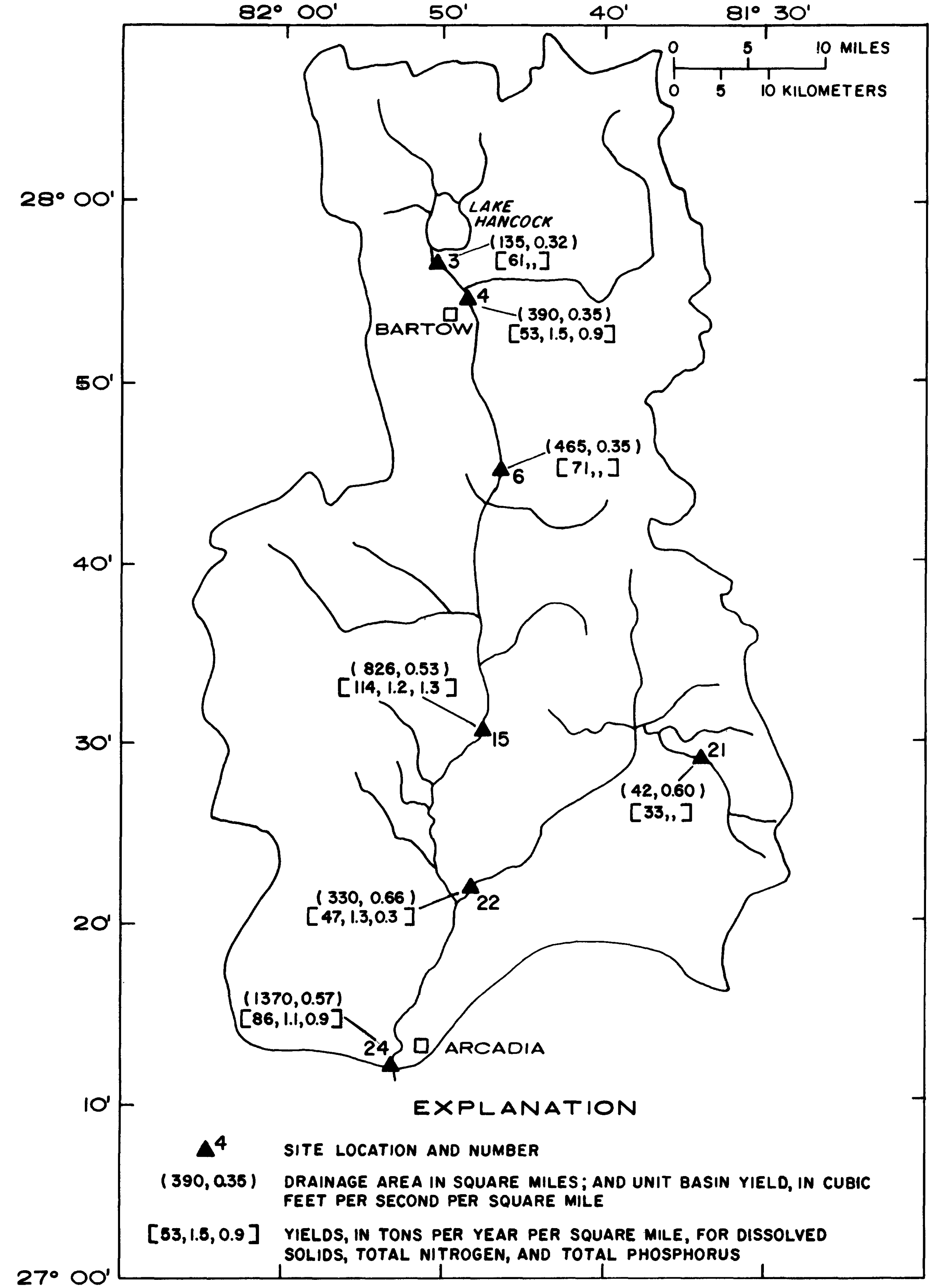


Figure 18.--Mean unit discharge and yields of dissolved solids, total nitrogen, and total phosphorus at selected sites, October 1974 through September 1982.

Unit discharge increases from the upper parts of the basin to the lower parts. The lowest unit discharge is at Saddle Creek (site 3) and is 0.32 $(ft^3/s)/mi^2$. Unit discharge at Arcadia (site 24) increased to 0.57 $(ft^3/s)/mi^2$. The highest unit discharge was for the Charlie Creek basin (site 22), and was 0.66 $(ft^3/s)/mi^2$. The higher unit discharge in the lower parts of the basin is probably because of increased runoff from areas of poor ground-water recharge.

Basin yields of dissolved solids at the seven sites ranged from 33 $(ton/yr)/mi^2$ at Little Charlie Bowlegs Creek (site 21), draining a 42 $mi^2$ area in the east-central part of the basin, to 114 $(ton/yr)/mi^2$ at Peace River at Zolfo Springs (site 15). Dissolved-solids yield along the Peace River was lowest (53 $(ton/yr)/mi^2$) at the most upstream site (Bartow, site 4), increased to 71 $(ton/yr)/mi^2$ at Fort Meade (site 6), and to 114 $(ton/yr)/mi^2$ at Zolfo Springs (site 15), and then decreased to 86 $(ton/yr)/mi^2$ at Arcadia (site 24). This indicates an increase of dissolved-solids loading between sites 4 and 6, and between sites 6 and 15.

Mean basin yields of total nitrogen at the four sites were highest (1.5 $(ton/yr)/mi^2$) at Bartow (site 4), lowest (1.1 $(ton/yr)/mi^2$) at Arcadia (site 24), and decreased from upstream to downstream along the Peace River.

Mean basin yields of total phosphorus along the Peace River had a pattern similar to that of dissolved solids--the yield was lowest (0.9 $(ton/yr)/mi^2$) at Bartow (site 4), was highest (1.3 $(ton/yr)/mi^2$) at Zolfo Springs (site 15), and then decreased to 0.9 at Arcadia (site 24). This indicates a major source of phosphorus to the Peace River between sites 4 and 15. Phosphorus yields of the Charlie Creek basin (0.3 $(ton/yr)/mi^2$ at site 22) are quite low in comparison to other parts of the Peace River basin.

## Water-Quality Trends

A trend in water quality is said to exist if values or concentrations have shown a tendency to either increase or decrease with time. Eight sites in the Peace River basin have periodic specific conductance data over a period of years, sufficient for trend testing. Of these eight sites, four also have total nitrogen and phosphorus data, at least since 1974. In addition, the NASQAN station has specific conductance and dissolved orthophosphate data beginning in 1959. These data were used to determine if water quality has changed with time.

Trend testing was done for two different periods of record. Short-term trend testing was done for 1974 through 1982, corresponding to the period of NASQAN data at site 24. Long-term trend testing was done using all specific conductance data for the eight stations, and dissolved orthophosphate data for site 24. In addition to the testing of trends in water quality, discharges at sampling time were also tested for trends.

### Seasonal Kendall Test

Many methods can be used for water-quality trend testing; these can be broadly classified into two types: parametric methods usually using time as an independent variable in a least-square regression model (which may also include other variables such as discharge), and nonparametric methods which examine the pattern of ranks (relative values) of the water-quality variable with time. Parametric tests are more powerful in trend detection than are nonparametric methods, but often the assumptions of regression are not met. These assumptions require data that do not fluctuate seasonally, are not serially correlated, and are normally distributed with constant variance over the entire discharge range. Because water-quality data generally do not meet these requirements, a nonparametric trend test was used in this study.

The Kendall's-tau test procedure (Kendall, 1975) examines all possible pairs of data values (water-quality measurements) and determines the number of concordant pairs (measurement latest in time has highest value) and the number of discordant pairs (measurement latest in time has lowest value). No trend exists if the number of concordant and discordant pairs are not statistically different. An upward trend exists if the number of concordant pairs is significantly greater than the number of discordant pairs, and the opposite situation indicates a downward trend.

Water-quality data commonly vary seasonally as a function of water temperature, biological activity, and runoff. For example, specific conductance (and major ion concentrations) generally is lowest during high-discharge months and highest during low-discharge months. Therefore, samples in any selected season should only be compared with samples in the same season of other years. This implementation of the Kendall's tau test is referred to as the Seasonal Kendall test (Hirsch and others, 1982). This procedure was used originally for testing for trends in phosphorus data collected at NASQAN stations nationwide (Smith and others, 1982). The procedure first determines median values for each seasonal grouping of data, and then does the Kendall's-tau test on the seasonal medians. A computer procedure for the trend analysis is included in the U.S. Geological Survey WATSTORE system, and is described by Crawford and others (1983).

The Kendall's-tau test by itself provides no information as to the rate of change of water quality with time, because concordance and discordance are determined from relative values, and thus, are not a function of either the time difference of the comparisons, or the magnitude of differences in the water-quality variable. Hirsch and others (1982) defined a measure of the time rate of change of the water-quality variable to be the median change per year for all possible pairings of samples. This is referred to as the Seasonal Kendall Slope Estimator, or the estimated change in the water-quality variable per year.

Flow adjustment.--It is well known that water quality generally varies with discharge. Thus, trends in water quality could be, in some cases, the result of trends in discharge, if, for example, samples from early in the data-analysis period were for high discharges, and samples near the end of the period were for low discharges. In such cases, trends could be occurring which have nothing to do with changes in basin development, land use, or waste disposal practices. These trends would ordinarily not be of as much interest as those caused by man's activities. Conversely, changes in water quality as a result of basin development could be opposed by changes related to discharge changes, so that no change in water quality is observed. For these reasons, flow-adjusted water-quality measurements, as well as the raw values, were used in trend testing.

Flow-adjusted specific conductance, nitrogen, and phosphorus concentrations were computed from the water quality-discharge regression relations presented in a previous section of this report. The flow adjusted value is the residual, or actual value minus the regression prediction (from discharge). The Seasonal Kendall test was applied to these flow adjusted values (residuals) as well as to the raw value. Trends in these residuals indicate a change in water quality not related to changes in discharge. However, the Seasonal Kendall Slope Estimator for the residual values is not equivalent to the change in actual concentrations per year.

Seasonal grouping of data.--Median values for seasons of relatively uniform discharge and climatic conditions were tested for trends. Three seasonal periods of four months each were used, as follows:

| Months | Characteristic conditions |
| --- | --- |
| March through June | Low to medium discharge, warming weather |
| July through October | High discharge, hot weather |
| November through February | Low to medium discharge, cooler weather |

Monthly means for selected water-quality variables and sampled discharges are shown in figure 19 for the NASQAN station at Arcadia (site 24) to illustrate seasonal variations in water quality and discharge. This figure shows the seasonal variation in specific conductance to be inversely related to the discharge. The seasonal pattern of total phosphorus concentration is similar to that of specific conductance. Total nitrogen, however, tends to be highest in the high-discharge months, but does not correspond closely with the pattern of discharge.

Groupings of 1 year were also tested. Comparison of results of the two trend tests may indicate trends in one or more seasons which do not show up on an annual basis. This is an alternate to testing each season by itself for trends, and was chosen to minimize the number of statistical tests being done. The risk of rejecting a null hypothesis of no trend, even when there is no trend, increases as the number of tests increases. This is because of the probabilistic nature of statistical testing, which carries with each statistical decision a risk (the significance level) of a wrong decision.

Limitations.--The Seasonal Kendall test as used here should be regarded as a screening tool for helping to sift through large quantities of data, rather than an absolute yes or no answer. Though the test is nonparametric and thus largely free of restrictions on data distribution, it still relies on definition of the water-quality discharge relation for the flow adjusted values and is indicative only of general tendencies over the time period being tested. Thus, a trend which reverses in direction during the test period may not be detected, nor might one which shows up only early or late in the test period. Also, with each test there is a 5 percent chance (the significance level) that the test might indicate a trend even though, in fact, there was none. However, it is likely that large, ongoing trends would be detected, and thus called to the attention of the investigator for further scrutiny.

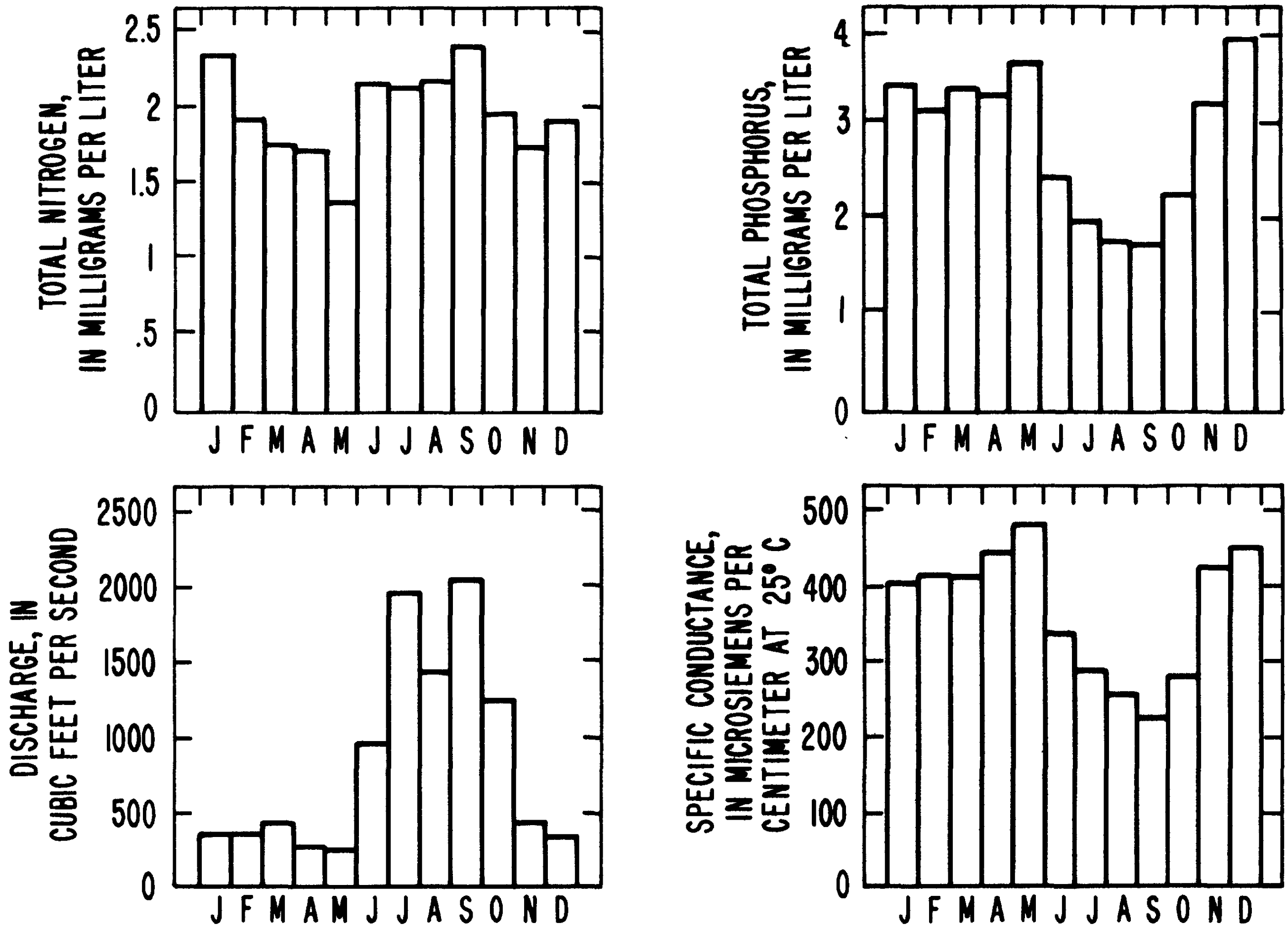


Figure 19.--Mean monthly concentrations of total nitrogen and total phosphorus, sampled discharge, and specific conductance, Peace River at Arcadia (site 24), October 1974 through September 1982.

## Results of Trend Analysis

Short-term trends.--Results of the short-term trend tests are given in table 10. This table shows the estimated change per year in specific conductance, total nitrogen, and total phosphorus for the period 1974 through 1982. The change estimates are given only for cases with statistical evidence of a trend at the 5 percent significance level.

Only two of the eight sites have statistical evidence of a trend in specific conductance. Little Charlie Bowlegs Creek (site 21) had an estimated increase in specific conductance of 6.2 $\mu$S/cm per year using seasonal comparisons of data, and 5.5 $\mu$S/cm using annual comparisons. The trends in specific conductance values and specific conductance residuals are identical because the specific conductance did not vary significantly with discharge,

Table 10.--Trends in water quality and sampled discharge, October 1974 through September 1982

[Estimated rates of change in water quality are given for cases where there is statistical evidence of a trend at a 5-percent significance level]

| Site No. | Estimated annual change in water quality: Seasonal test Value | Seasonal test Residual | Annual test Value | Annual test Residual | Sampled discharge (cubic foot per second per year): Seasonal test | Sampled discharge (cubic foot per second per year): Annual test |
|---|---|---|---|---|---|---|
| | Specific conductance, in microsiemens per centimeter per year | | | | | |
| 3 | -- | -- | -- | -- | 7.6 | 12 |
| 4 | -- | -- | -- | -- | -- | -- |
| 6 | -- | -- | -- | -- | -- | -- |
| 10 | -- | -- | -- | -- | -- | -- |
| 15 | -- | -- | -- | -- | -- | -- |
| 21 | 6.2 | 6.2 | 5.5 | 5.5 | -- | -- |
| 22 | -- | 5.1 | -- | -- | -- | -- |
| 24 | -- | -- | -- | -- | -- | -- |
| | Total nitrogen, in milligrams per liter per year | | | | | |
| 4 | 0.25 | -- | 0.37 | -- | 15 | -- |
| 15 | -- | -- | -- | -- | -- | -- |
| 22 | -- | -- | -- | -- | -- | -- |
| 24 | -- | -- | -- | -- | -- | -- |
| | Total phosphorus, in milligrams per liter per year | | | | | |
| 4 | -- | -- | -- | -- | 15 | -- |
| 15 | -- | -- | -- | -- | -- | -- |
| 22 | -- | -- | -- | -- | -- | -- |
| 24 | -- | -- | -- | -- | -- | -- |

at least for the data available. Charlie Creek (site 22) had a significant trend in seasonally compared specific conductance residuals, but not in specific conductance values. This is an indication that the specific conductance-discharge relation may have changed.

Plots of seasonal median specific conductance and residuals at Little Charlie Bowlegs Creek are shown in figure 20. This plot verifies the trend detected by the Kendall's-tau test, and shows that specific conductance was generally greater than 100 μS/cm after 1979, and lower than 100 μS/cm in earlier years.

A plot of specific conductance and residuals for Charlie Creek is shown in figure 21. This figure shows no noticeable overall trend in specific conductance or in residuals, as indicated by the Kendall's-tau test. However, a comparison of median residuals for November through February samples does show a tendency for increasing residuals during these months.

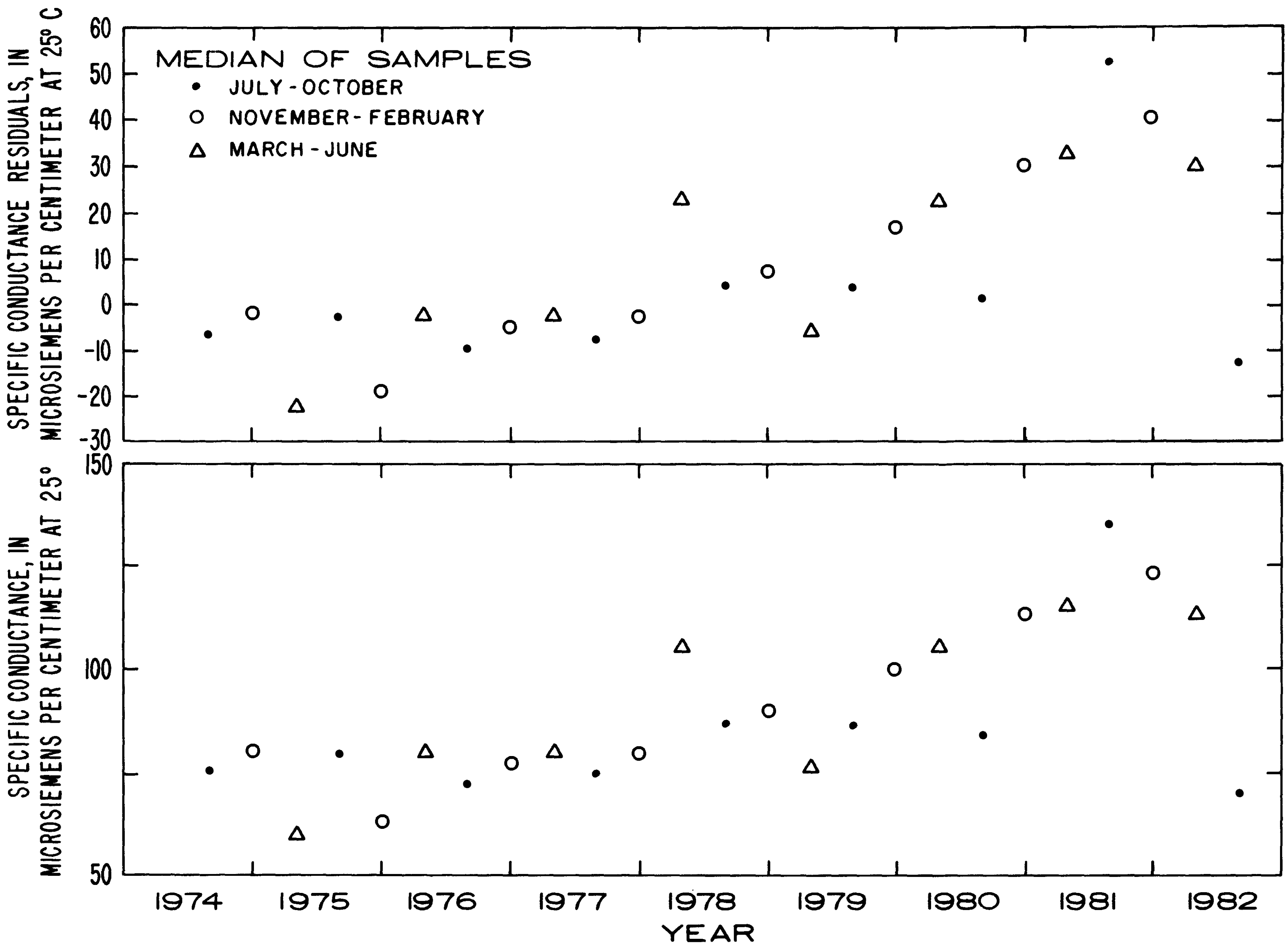


Figure 20.--Seasonal specific conductance residuals (observed minus predicted) and values at Little Charlie Bowlegs Creek (site 21), October 1974 through September 1982.

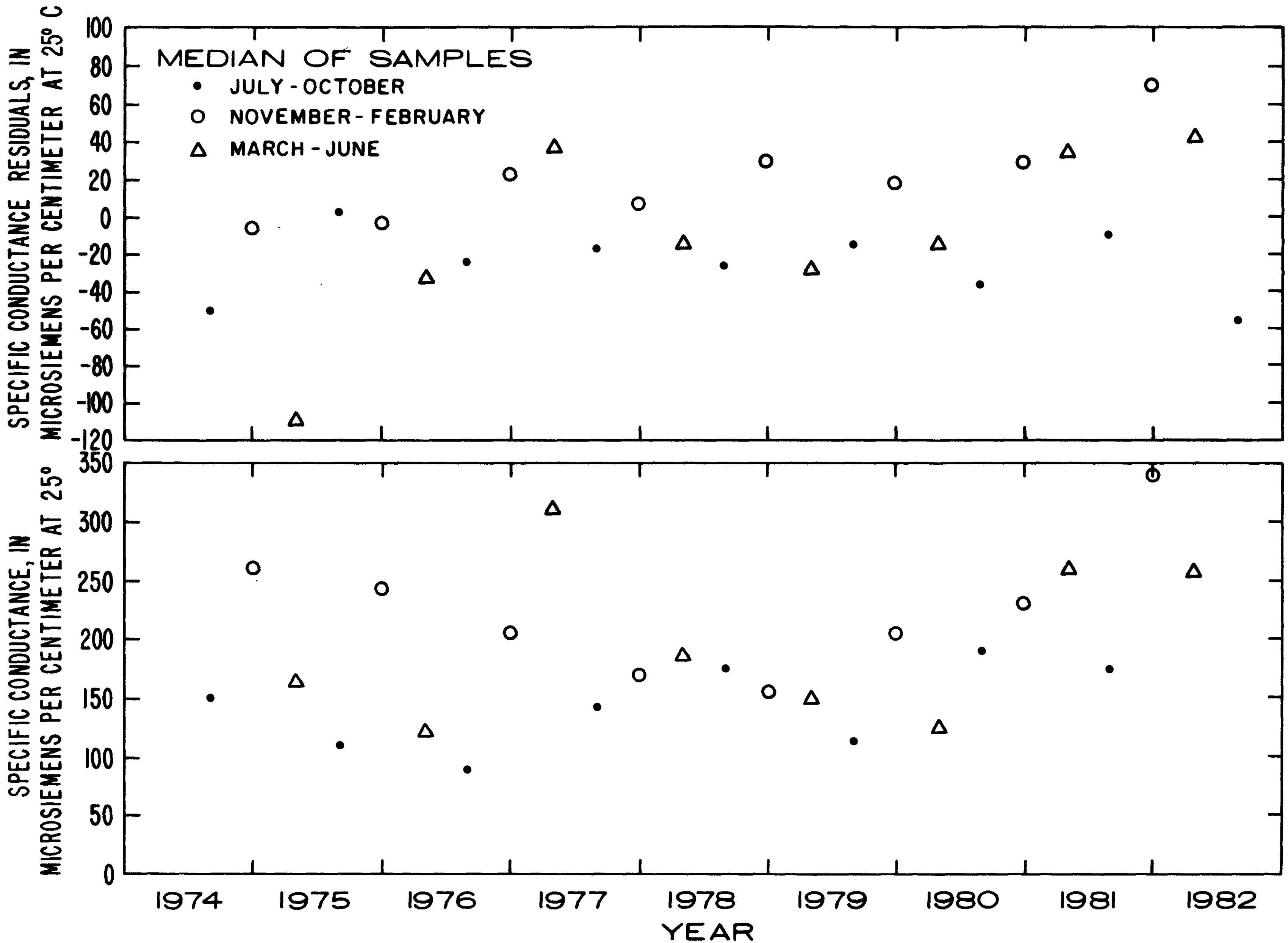


Figure 21.--Seasonal specific conductance residuals (observed minus predicted) and values at Charlie Creek (site 22), October 1974 through September 1982.

A trend in total nitrogen concentrations was detected at Peace River at Bartow (site 4). However, there was no detectable trend in residuals (fig. 22), and a test of sampled discharges, using the seasonal test, showed a significant upward trend. Also, total nitrogen increases with discharge during relatively low discharges (table 8). Therefore, the apparent trend in nitrogen can probably be accounted for by the increased discharges, rather than basin changes.

There were no statistically significant trends observed in total phosphorus among the four stations tested.

This analysis of data for trends indicates that although water-quality trends are not widespread in the Peace River basin, some changes probably are taking place but are insufficient to affect the water quality at the NASQAN station. This emphasizes the fact that the NASQAN data by itself cannot describe all aspects of water quality in the basin upstream.

Long-term trends.--The results of long-term trend tests for specific conductance are given in table 11. These tests are not specific to a selected interval of time since the period of data is generally different for each site, as indicated.

Four of the eight sites showed significant trends in specific conductance or residuals and in all cases these trends were for increasing specific conductance with time.

Plots of specific conductance for the three sites with trends in both specific conductance values and residuals are shown in figure 23. These plots tend to confirm the statistical tests. Specific conductance was generally less than 250 $\mu$S/cm before 1970 (and greater than 250 $\mu$S/cm from 1979 through 1982) at Payne Creek (site 10), was frequently less than 300 $\mu$S/cm before 1975 at Peace River at Zolfo Springs (site 15), and was frequently less than 250 $\mu$S/cm before 1973 at Arcadia. Interestingly, Little Charley Bowlegs Creek did not appear to have a specific conductance trend using all data (1966 through 1982) but did have a statistically-significant short-term trend (1974 through 1982).

Trend tests of dissolved orthophosphate concentrations and residuals at Arcadia for 1959-82 indicated no significant trend in seasonal or annual median concentrations or sampled discharges, but did indicate an increasing trend in seasonal residuals. This could be an indication that the dissolved orthophosphate-discharge relation has changed.

## SAMPLING FREQUENCY FOR DESCRIBING WATER-QUALITY CONDITIONS

Frequency of water-quality sampling needs to be based on intended uses of the data, balanced against economics of sample collection and analysis. For example, if it is desired to determine accurate annual chemical constituent loads at a station, analyses of composites of daily samples probably is the best method. However, it is expensive and not practical for all water-quality variables. Frequency of sampling also depends on variability of the water-quality measurements; highly variable constituents will require more sampling to define the distribution than will constituents that are always at about the same concentration.

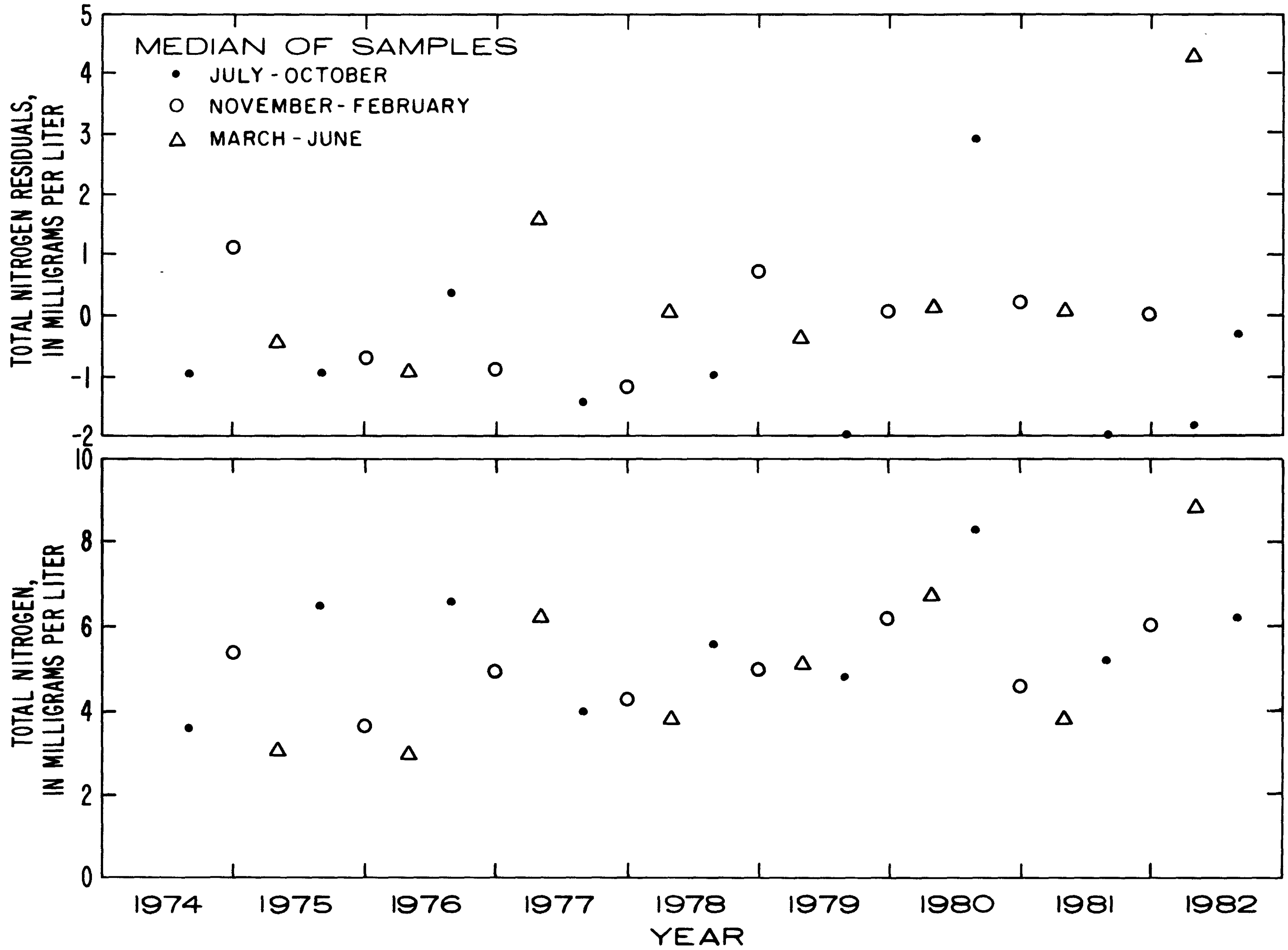


Figure 22.--Seasonal total nitrogen residuals and values, Peace River at Bartow (site 4), October 1974 through September 1982.

Table 11.--Trends in specific conductance and sampled discharge, period of record

[Estimated rates of change in specific conductance are given for cases where there is statistical evidence of a trend at a 5-percent significance level]

| Site No. | Period of record | Estimated annual change in specific conductance, in microsiemens per centimeter per year: Seasonal test Value | Seasonal test Residual | Annual test Value | Annual test Residual | Sampled discharge (cubic feet per second per year): Seasonal test | Annual test |
|---|---|---|---|---|---|---|---|
| 3 | 1965-82 | -- | -- | -- | -- | -- | -- |
| 4 | 1963-82 | -- | -- | -- | -- | -- | -- |
| 6 | 1965-82 | -- | -- | -- | -- | -- | -- |
| 10 | 1956-82 | 9.1 | 7.4 | 9.1 | -- | -- | -- |
| 15 | 1951-82 | 6.0 | 2.7 | 5.5 | 2.9 | -- | -- |
| 21 | 1966-82 | -- | -- | -- | -- | -- | 0.2 |
| 22 | 1965-82 | -- | 3.2 | -- | -- | -- | -- |
| 24 | 1957-82 | 6.0 | 2.4 | -- | 2.0 | -12 | -- |

In this section, sampling coverage of the range of discharge conditions and variation of the estimated annual mean concentration or value as a function of sampling frequency at the NASQAN station (site 24) is discussed. Although this treatment is not exhaustive, it is intended to be a preliminary step in developing methods for determining adequacy of various sampling frequencies to describe water quality.

## Range of Stream Discharge Conditions

Water quality may be profoundly affected by discharge. For this reason, it is desirable to sample a wide range of discharge conditions, making sure that low discharges as well as flood events are sampled. Generally, sampling trips are scheduled at regular intervals, so that in the long run, all magnitudes of discharge will be covered. However, unless specific efforts are made to sample extreme flow conditions, other data-collection activities (flood measurements, low-flow measurements) could disrupt the routine sampling schedules.

The number of trace metals, nutrients, and specific conductance samples at the NASQAN station, in selected discharge categories, was examined to determine if the fixed frequency of sampling had been adequate to represent the discharge range. A nutrient sample and a metal sample generally consist of several constituents. For this tabulation, a total nitrogen analysis was assumed to represent a nutrient analysis, and a total recoverable or dissolved lead or zinc analysis was assumed to represent a metals sample. Figure 24 shows sampling coverage according to discharge, using discharge intervals of 5 percent duration. For example, figure 24 shows that samples of specific

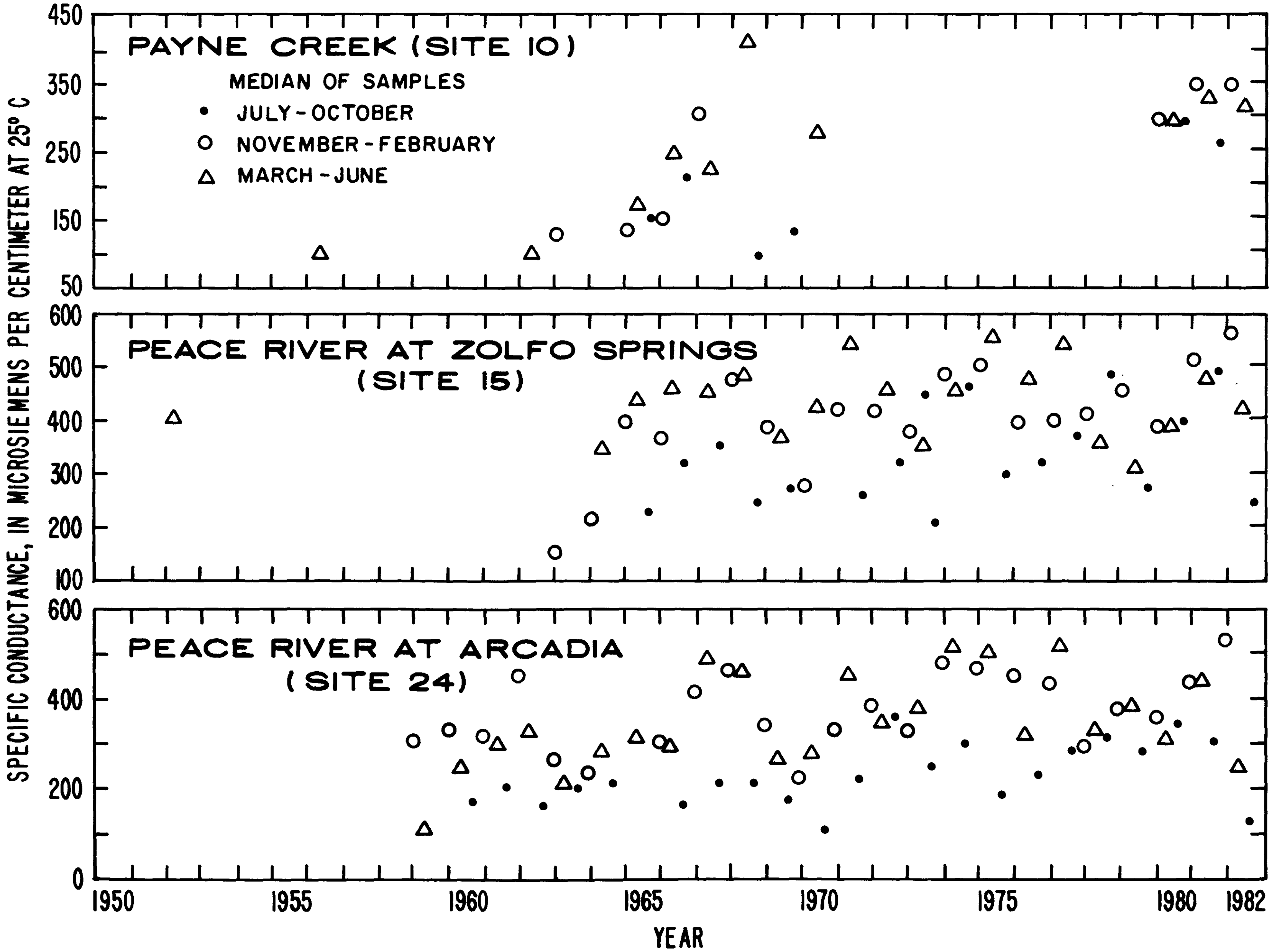


Figure 23.--Seasonal specific conductance for periods of record at sites 10, 15, and 24.

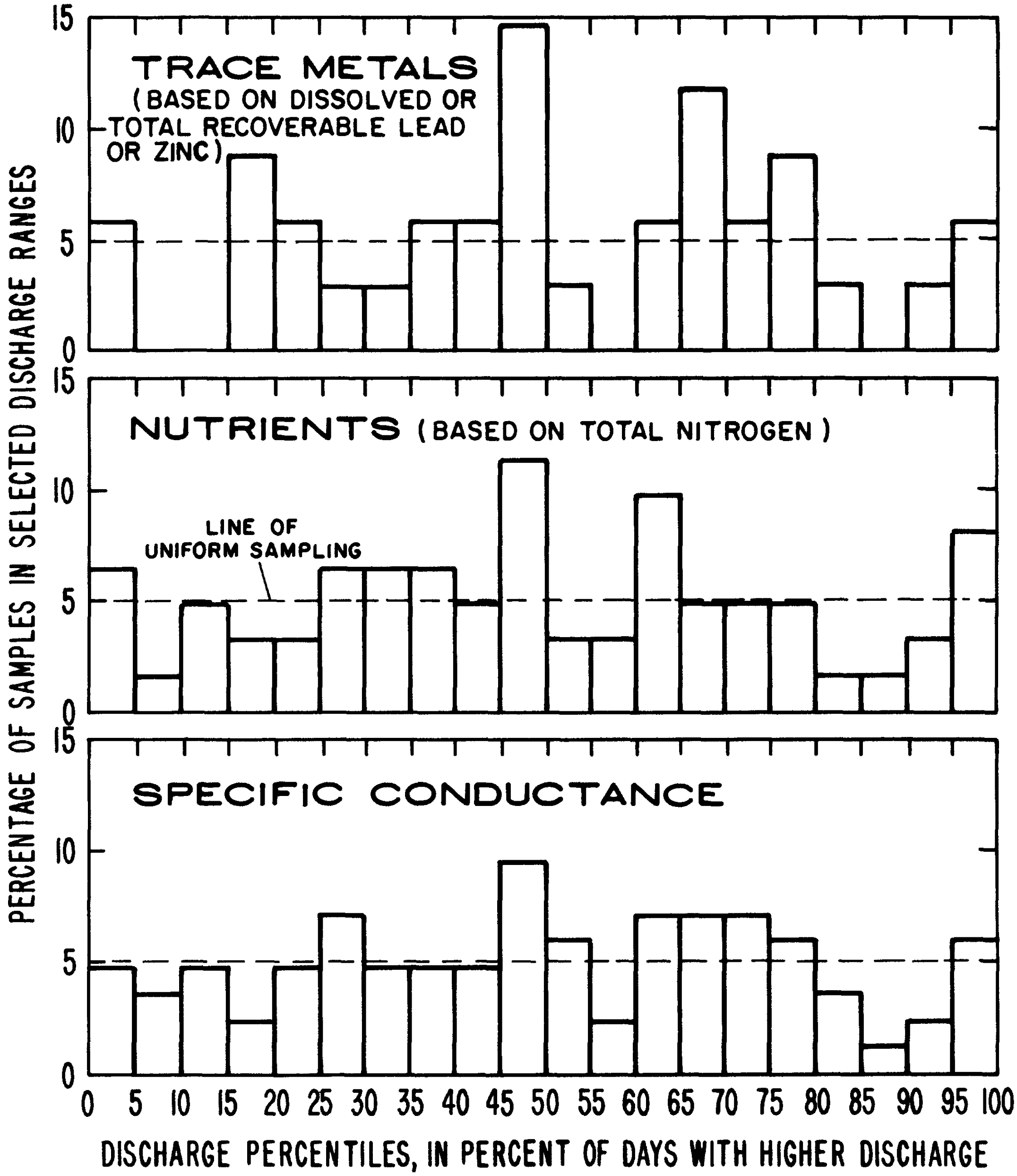


Figure 24.--Distribution of trace metals, nutrients, and specific conductance samples according to discharge, Peace River at Arcadia (site 24), October 1974 through September 1982.

conductance within the range of discharges which were exceeded from 0 to 5 percent of the time accounted for just less than 5 percent of all the samples. If uniform sampling coverage over the range of discharges is desirable, each 5 percent interval of discharge percentiles should contain about 5 percent of the samples. Completely uniform sampling probably is not a necessary criteria for a representative set of data; however, samples need to be collected to include extremes of discharge.

Figure 24 indicates that, for the period October 1974 through September 1982, sampling of trace metals, nutrients, and specific conductance was adequate to represent the extremes of discharge. Indeed, sampling coverage throughout the range of discharge was generally uniform, except for a relatively small number of samples in the discharge range exceeded from 80 to 95 percent of time, and for trace metals, lack of any samples in the range of discharges exceeded from 5 to 15 percent of the time.

## Simulation of Selected Sampling Frequencies

A Monte Carlo simulation of specific conductance, total nitrogen and phosphorus, total-recoverable lead, and total-recoverable zinc was done to evaluate adequacy of selected sampling frequencies for determining annual mean concentrations at the NASQAN station (site 24). Discharge was also included to represent a water-quality pseudovariable with a highly skewed distribution. The simulation involved computation of an annual mean for each of 1,000 years from "samples" picked at random according to the distribution function for each water-quality measure.

The procedure was to first determine the cumulative frequency distribution of the water-quality variables. Then, a random number in the range 0 to 100 was picked, and a value corresponding to the random number was taken from the cumulative distribution function for each water-quality variable. These values represent random selections with a distribution identical to the actual distribution of the data. For each year of the simulated 1,000 years of sampling, means were calculated for a "sample" size of 2 (twice annually), 4 (quarterly), 6 (bimonthly), 12 (monthly), 26 (biweekly), 52 (weekly), and 365 (daily). A comparison of the ranges of the annual means for these simulated sampling frequencies gives some insight into accuracy of the mean as a function of sampling frequency.

This simulation implies that: (1) water quality each year varies according to the frequency distributions for the period 1974-82, and (2) that water quality is not serially correlated (that is, that each sample is independent of the last sample). The first assumption is definitely not met each year, but should be appropriate in the long run. The second assumption is probably valid only for the relatively low sampling frequencies (4 per year, or less). At higher sampling frequencies, the distribution of the simulated values will have more scatter than would the distribution of the serially correlated (usually) samples. But the annual means of simulated samples will probably be unaffected by the lack of consideration of serial correlation because, in the long run, the excess scatter will have an average of zero.

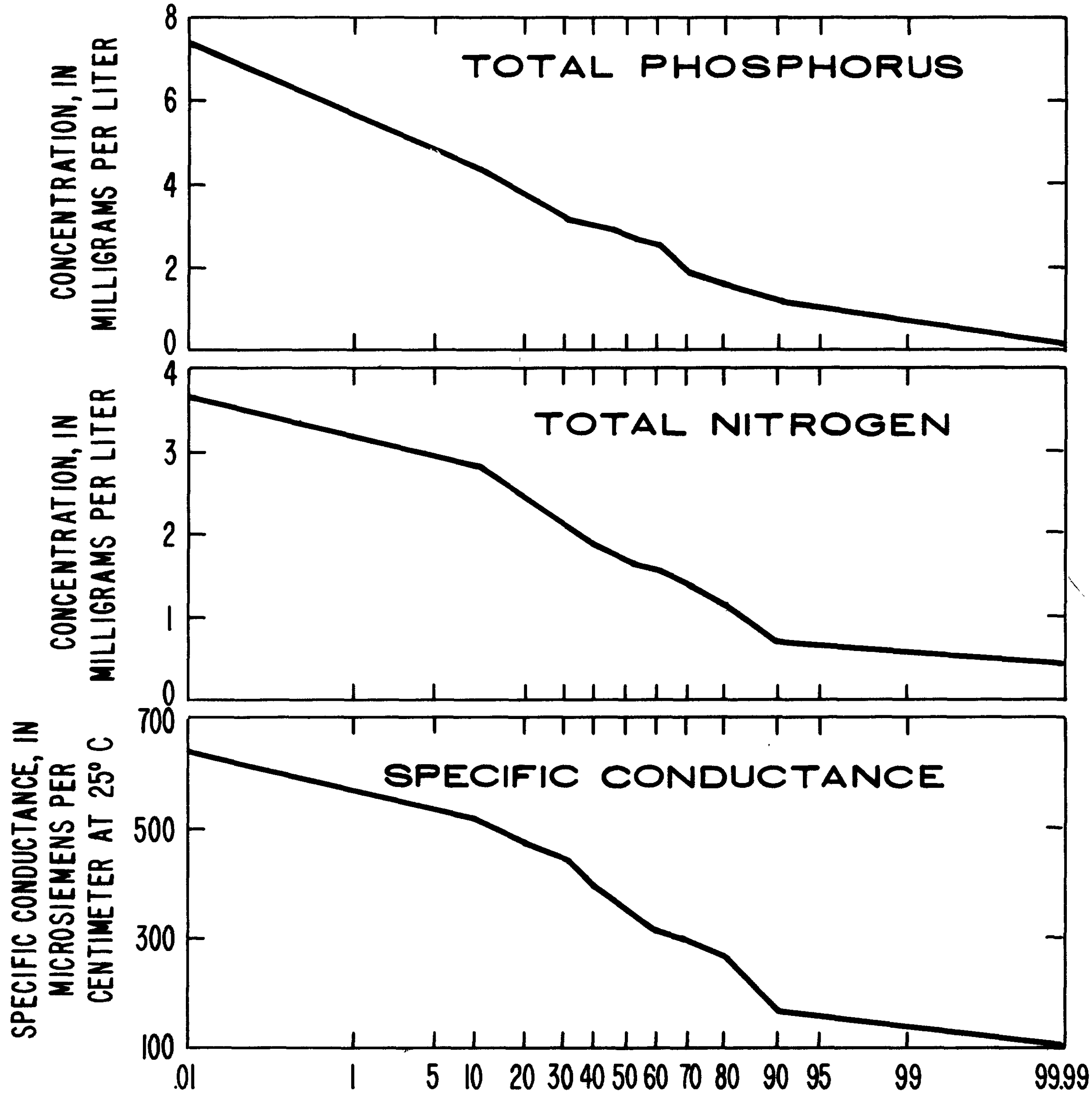


Figure 25.--Cumulative distribution of total phosphorus, total nitrogen, and specific conductance, Peace River at Arcadia (site 24), October 1974 through September 1982.

The sampling simulation may be atypical of some fixed-interval sampling programs because of the effects of seasonality. This lack of representation is probably most pronounced for small annual numbers of samples, and is the result of the efforts in most sampling programs to avoid repetitive sampling of a particular seasonal condition. For example, a fixed-frequency sampling program of two samples per year would probably arrange for the sampling to cover contrasting seasonal conditions (for example, a low-flow season sample and a high-flow season sample), rather than to collect both samples during the season of low or high discharge. Thus, there would be a tendency for a fixed-interval sampling program to sample extreme conditions more often than simulated by the procedure used here. The effect of seasonal-oriented sampling on the simulated sampling has not been investigated, but could be addressed using a more sophisticated simulation algorithm.

Cumulative frequency distribution curves for specific conductance, total nitrogen, and total phosphorus are shown in figure 25, and for discharge, total-recoverable lead, and total-recoverable zinc in figure 26. The curves for discharge and lead (fig. 26) show a high degree of skewness, with only a small percentage of high values. Therefore, for these variables, high values will be selected by the random simulation process only for the appropriately small percentage of the samples.

Interquartile ranges for annual means of 1,000 simulated years of sampling at the selected annual frequencies, and the actual mean of the data defining the distribution curves, are shown in figure 27. The interquartile range is the interval between the 25th percentile of the 1,000 annual means and the 75th percentile. It therefore, includes 50 percent of the calculated means, and of the remaining means, 25 percent were higher than the limits of the interquartile range. The magnitude of this interquartile range is an indication of effect of sampling frequency on accuracy of the computed mean, because the smaller this range is, the more likely that, in any given year, a mean determined by sampling will represent the true mean.

Higher sampling frequencies generally produce a more accurate estimate of the annual mean (fig. 27). Parameters with skewed distributions, such as lead and discharge (fig. 26), need to be sampled more frequently than parameters with less skew, or annual means may be biased to the low side. This is because high concentrations occur relatively rarely for lead and discharge, and will likely be missed if sampling is infrequent. Two samples per year in the simulated sampling for lead and discharge produced an interquartile range which did not include the actual mean of the distributions.

The effect of annual sample frequency on coefficient of variation (standard deviation, in percent of mean, of the annual mean concentrations) for the simulated sampling is shown by figure 28. The coefficients of variation for the six parameters analyzed decrease sharply with increasing sample frequency up to a frequency of 30 to 60 samples per year, and decrease more gradually at higher sampling frequencies. This implies an optimum sampling frequency of weekly to biweekly, with increased sampling beyond this point yielding smaller and smaller decreases in annual mean errors. This frequency of sampling would be much more expensive than the present NASQAN sampling schedule at the Peace River station (quarterly for metals, bimonthly for nutrients and some other variables), and probably would not be practical for all sites within the NASQAN network.

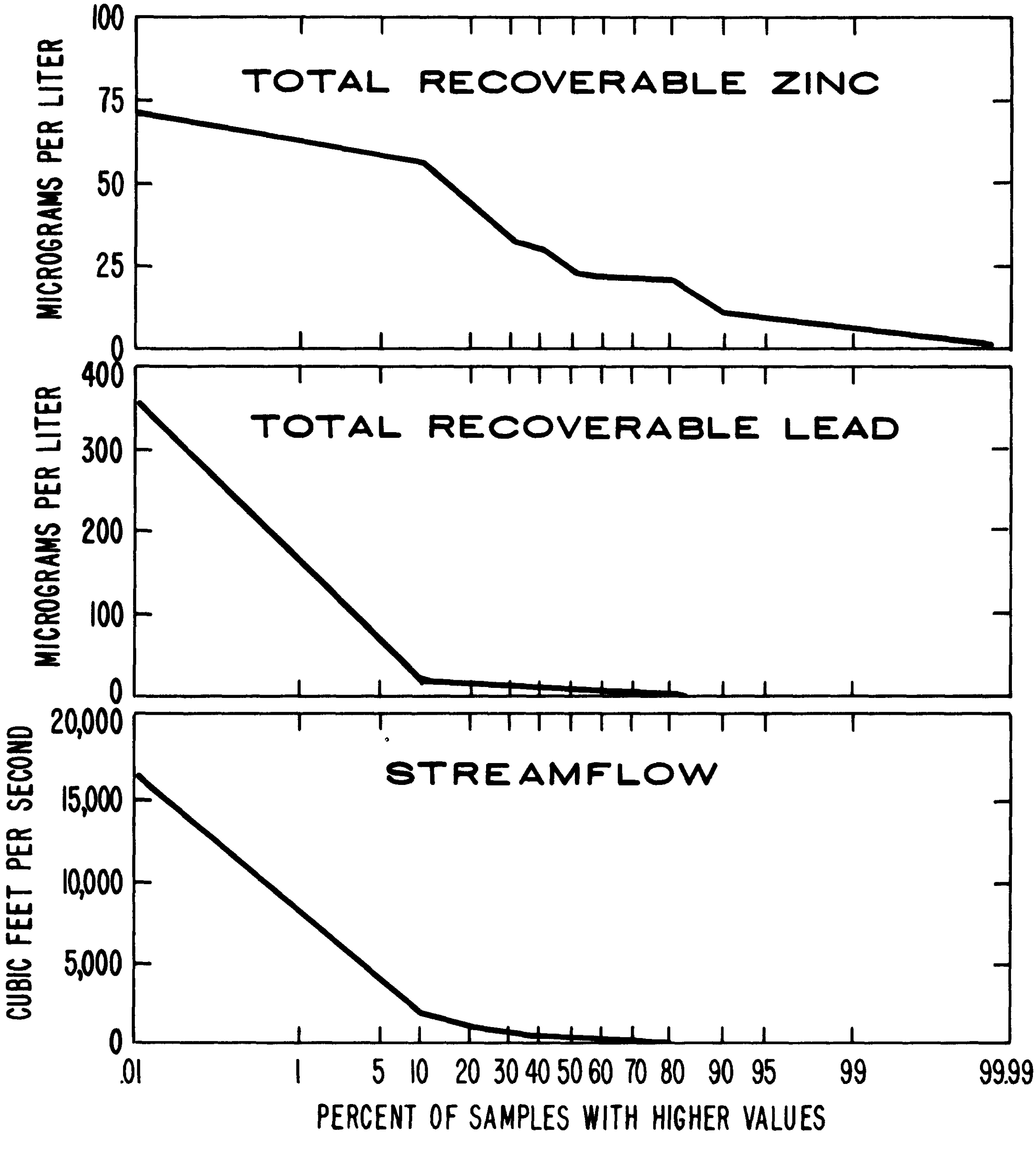


Figure 26.--Cumulative distribution of total recoverable zinc, total recoverable lead, and streamflow, Peace River at Arcadia (site 24), September 1974 through October 1982.

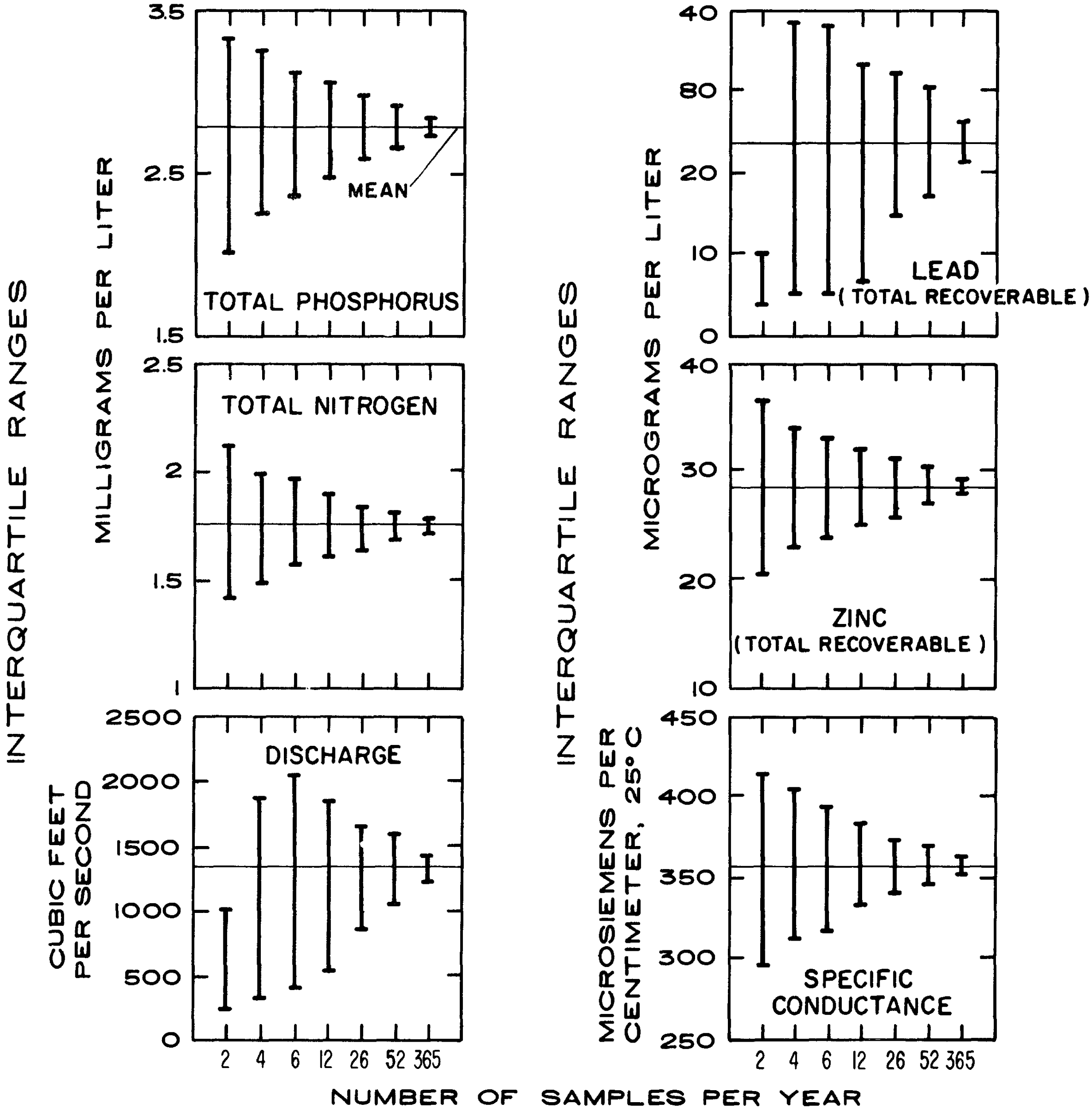


Figure 27.--Interquartile ranges in annual means of selected parameters for 1,000 years of simulated sampling, Peace River at Arcadia (site 24).

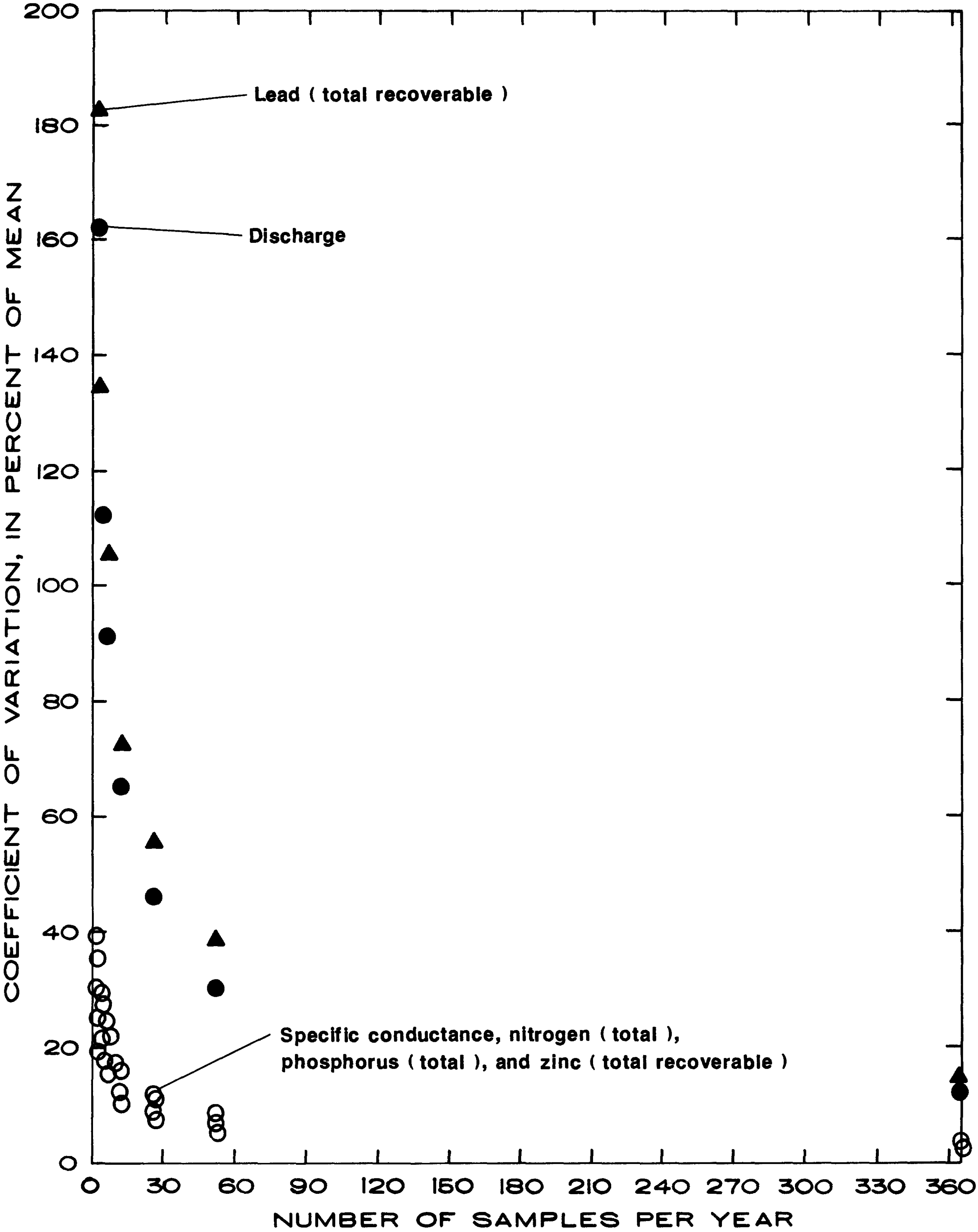


Figure 28.--Uncertainty in annual mean of selected parameters for 1,000 years of simulated sampling, Peace River at Arcadia (site 24).

The coefficient of variation of the means of any set of normally distributed data is inversely proportional to the square root of the number of samples used to compute the mean. For example, the set of data created by taking means of four samples will have only half the coefficient of variation of the original data, and increasing the sample size to 16 will give a coefficient of variation half that of the mean of four samples. This relation of coefficient of variation to sample size for normally distributed data fits very closely all relations shown in figure 28, even though not all the data are normally distributed.

## Sampling Frequency for Water-Quality Variables Related to Stream Discharge

The preceding section discussed sampling frequency without using independent variables that could be used to enhance estimation of water quality provided by periodic sampling. Many water-quality variables are related to discharge, and discharge is computed daily for many stations through use of discharge-stage relations. Thus, through use of discharge data and water-quality data jointly, it should be possible to obtain more accurate estimates of annual mean water quality than could be derived from the periodic sampling alone.

A procedure developed by Moss and Gilroy (1980) to optimize the number of annual discharge measurements necessary to define the discharge-stage relations was used to estimate effect of sampling frequency on accuracy of annual mean water quality when the concentration-discharge relation is used with daily discharge to compute the annual mean values. In the procedure used (analogous to that used by Moss and Gilroy (1980)), water quality was the predicted variable (instead of discharge), and discharge was the independent variable (instead of stage).

Water-quality variables selected were specific conductance, total nitrogen, and total phosphorus at site 24. These variables are related to discharge, and have been sampled frequently at many stations. No attempt was made to study estimations of annual trace-constituent means, because these are commonly present in concentrations at or below laboratory detection limits and are not sampled as frequently as other variables; this makes the concentration-discharge relations difficult to determine.

Development of the procedure is complex and lengthy, and will not be repeated here. Basically, the procedure follows three steps. First, the relation between chemical constituents and stream discharge is determined, and residuals (actual minus predicted values) were computed for each sample, using discharge at sampling time to obtain the predicted value. Next, the residuals are analyzed by a Kalman-filter technique (Gelb, 1974) requiring as input the residuals, and an estimation of average measurement error, in percent. Errors were estimated to be 5 percent for specific conductance (based on author's field experience), 10 percent for total nitrogen, and 15 percent for total phosphorus (based on data from the U.S. Geological Survey National Water Quality Laboratory). Output of the Kalman-filter procedure was the estimated 1-day lagged autocorrelation coefficient (correlation of water quality on a given day with that for the day before), and the estimated process variance, or the variance in water quality not associated with measurement error. The final step is to compute the coefficient of variation of the estimated annual mean water-quality values for selected annual sample frequencies based on the process variance, the 1-day autocorrelation coefficient, and the estimated measurement error. These estimates are then plotted with annual sample frequency, and the plots are inspected to determine optimum sampling frequency.

Results of this analysis are shown in figure 29. Total nitrogen and phosphorus errors decrease rapidly with increased sampling frequencies up to about 60 samples per year, and at higher sampling frequencies the return in decreased error is much lower. This implies that weekly sampling of these parameters would be optimum, and that more frequent sampling would produce only small decreases in error in the estimated annual mean concentrations. For specific conductance, error is relatively small even with no sampling and is not greatly reduced even with a high sampling frequency.

The effect of using concentration-discharge relations on errors in estimated annual mean values is indicated by comparing uncertainty plots shown in figure 28 with those shown in figure 29. Biweekly sampling of specific conductance, total nitrogen, and total phosphorus, yielding 26 samples per year, would result in coefficients of variation of about 8 percent for the estimated annual means if daily discharge record is not used to estimate water quality for unsampled days. The same magnitude of error could be obtained with only about 5 phosphorus samples per year, and errors of less than 8 percent would result even if no samples of nitrogen or specific conductance were taken, if the concentration-discharge relations were used to estimate water quality as a function of discharge. It would be necessary to obtain some samples for all variables to ensure that the concentration-discharge relations had not changed.

Further work is needed to determine optimum or minimum sampling frequencies necessary to achieve monitoring objectives. The use of concentration-discharge relations could reduce sampling requirements for estimation of annual means, however, more work is needed to verify this and to develop methods for detecting changes in concentration-discharge relations. Furthermore, the effect of sampling frequency on trend-detection techniques and characterization of extreme values need to be considered before any reduction of the sampling frequency, based only on error of estimated annual mean, is done.

## INTENSIVE RECONNAISSANCE OF THE PEACE RIVER BASIN

During 1983, a network of 23 stream sites in the upper Peace River basin was sampled twice as part of a cooperative study by the U.S. Geological Survey and the Florida Department of Environmental Regulation. The overall objective of the sampling was to describe water quality of the river for different discharge conditions, and to relate water quality of the river to mining and other activities that could contribute to constituent loading of the river.

Data from these intensive sampling surveys, though limited in number of samples per site, are useful in describing water-quality changes in the basin for relatively uniform discharge conditions during a short interval of time--generally 2 to 3 days. This system of data collection is analogous to a

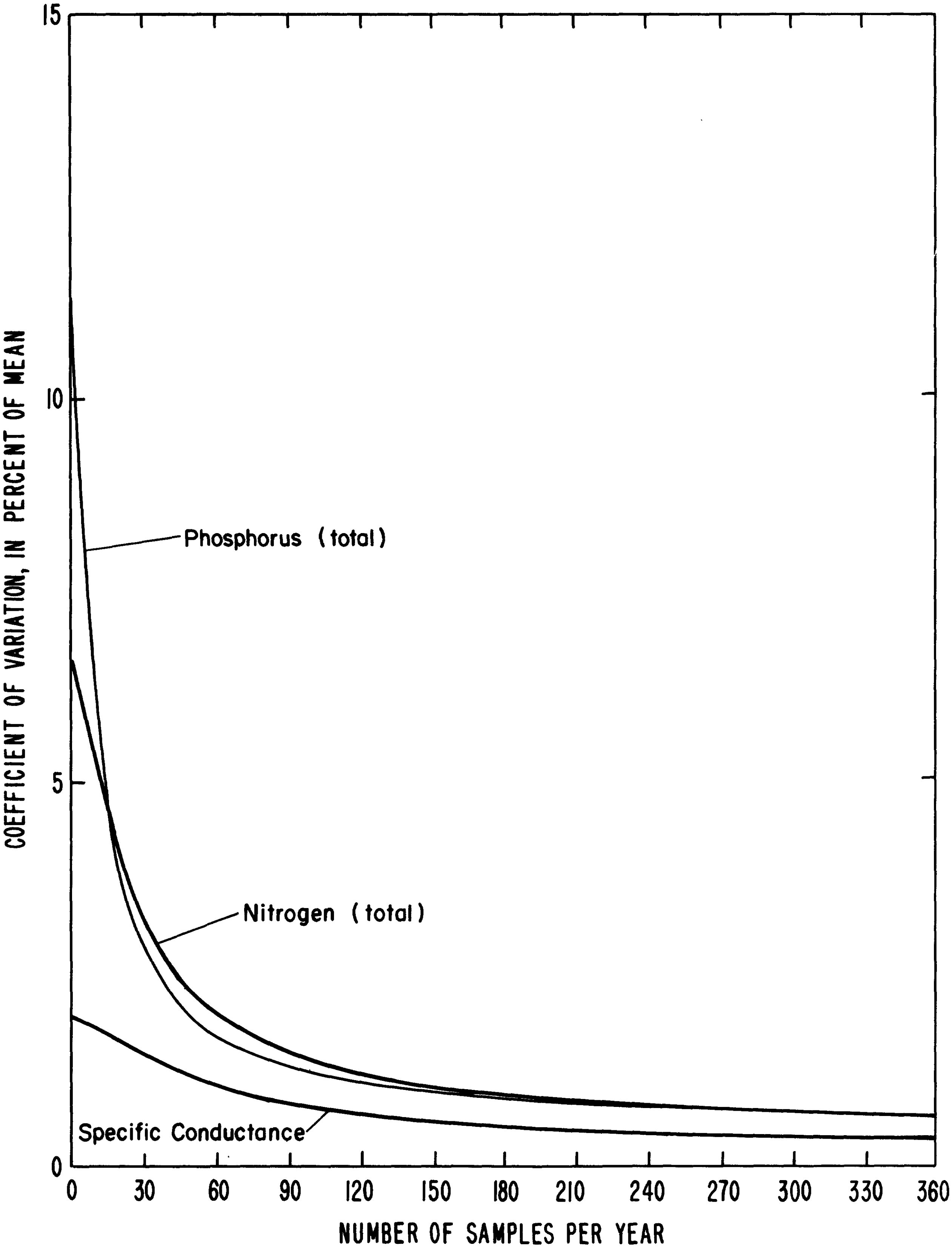


Figure 29.--Uncertainty in annual mean specific conductance, total nitrogen, and total phosphorus based on discharge relations, Peace River at Arcadia (site 24).

"snapshot" in that the coverage of data is large areally but small temporally and is an alternative to the continuing fixed station fixed frequency system in which relatively large numbers of samples are collected at a few (or one) sites.

Data from 23 stream sites (listed in table 12) were used to compute basin yields of dissolved solids, total nitrogen, and total phosphorus, for two sets of samples. The 23 stream sites are indicated by circles in figure 2. The two sample sets used were for February 14-15, 1983, when discharges were high, and May 24-25, 1983, when discharges were lower. The high-discharge samples generally reflect water-quality effects from nonpoint sources, and the lower discharge samples are more influenced by point sources and geohydrologic factors.

Table 12.--List of intensive reconnaissance sites

| USGS station No. | Map and site No. | Station name | Latitude | Longitude |
|---|---|---|---|---|
| 02293986 | 1 | Peace Creek drainage canal | 27°55'23" | 81°42'28" |
| 02294290 | 2 | Saddle Creek at State Hwy 540 near Eaton Park | 28°00'15" | 81°51'08" |
| 02294491 | 3 | Saddle Creek at structure P-11 near Bartow | 27°56'17" | 81°51'05" |
| 02294650 | 4 | Peace River at Bartow | 27°54'07" | 81°49'03" |
| 02294781 | 5 | Peace River near Homeland | 27°49'13" | 81°47'57" |
| 02294898 | 6 | Peace River at Fort Meade | 27°45'04" | 81°46'56" |
| 02295067 | 7 | Bowlegs Creek | 27°43'15" | 81°47'20" |
| 02295163 | 8 | Whidden Creek | 27°42'35" | 81°48'28" |
| 02295194 | 9 | Peace River at Bowling Green | 27°38'45" | 81°48'09" |
| 02295420 | 10 | Payne Creek | 27°37'13" | 81°49'33" |
| 02295440 | 11 | Peace River at State Hwy 664A near Wauchula | 27°34'32" | 81°48'17" |
| 02295557 | 12 | Little Charlie Creek near Wauchula | 27°35'15" | 81°46'17" |
| 02295607 | 13 | Peace River at Wauchula | 27°33'01" | 81°47'38" |
| 02295614 | 14 | Peace River near Wauchula | 27°32'24" | 81°47'33" |
| 02295637 | 15 | Peace River at Zolfo Springs | 27°30'15" | 81°48'04" |
| 02295642 | 16 | Peace River at State Hwy 64 at Zolfo Springs | 27°29'59" | 81°48'39" |
| 02295735 | 17 | Troublesome Creek | 27°28'53" | 81°51'57" |
| 02295760 | 18 | Hickory Creek | 27°26'35" | 81°52'29" |
| 02295800 | 19 | Peace River near Limestone | 27°24'52" | 81°50'53" |
| 02295870 | 20 | Oak Creek | 27°24'53" | 81°52'54" |
| 02296500 | 22 | Charlie Creek | 27°22'29" | 81°47'48" |
| 02296600 | 23 | Peace River at Brownville | 27°18'09" | 81°50'48" |
| 02296750 | 24 | Peace River at Arcadia | 27°13'19" | 81°52'34" |

## Streamflow Conditions

Discharges during the February run were high and increasing at all three of the Peace River sites for which continuous discharge data are recorded (fig. 30), in response to heavy rainfall on February 13. Rainfall on February 13 was 2.10 inches at Bartow and 2.16 inches at Arcadia, according to records of the National Oceanic and Atmospheric Administration (U.S. Department of Commerce, 1983). Daily mean discharges on sampling date (February 14 at Bartow and Fort Meade, and February 15 at Arcadia) for these three stations ranged from 801 $ft^3/s$ at Fort Meade to 4,100 $ft^3/s$ at Arcadia, and were exceeded about 2 to 4 percent of the days from October 1974 through September 1982 (fig. 15). These data indicate that discharges were relatively high at all of the stations sampled on February 14-15, 1983.

Discharges in May 1983 were much lower than in February and the sampling on May 24 and 25 was preceeded by dry weather and decreasing discharge (fig. 30). Rainfall near Fort Meade in late afternoon of May 25 caused the increase in discharge indicated in figure 30, but all sampling was completed before this occurred. Daily discharges on May 24 at the three Peace River sites with continuous discharge record ranged from 27 $ft^3/s$ at Bartow to 161 $ft^3/s$ at Arcadia, and were exceeded about 70 percent of the days from October 1974 through September 1982 (fig. 15). These data indicate that discharges throughout the basin were relatively low, but not extremely low, during the sampling on May 24-25, 1983.

## Basin Yields of Chemical Constituents

Basin yields of dissolved solids, total nitrogen, and total phosphorus were computed for the February and May 1983 sampling trips, and are shown in figures 31 through 36. Basin yields shown are for tributaries to the Peace River, and for selected sites on the Peace River. These yields are for the entire basins upstream of the stations. Also shown are selected subbasin yields, for the incremental area of the Peace River basin between two stations. These subbasin yields are useful in assessing input to the main stem of the river from tributaries and interbasin areas.

Subbasin yields were calculated by subtracting load at the upstream site from load at the downstream site and dividing by the drainage area between the two sites. Loads are the product of concentration and discharge, and thus combine the errors in the determination of both parameters. Total error in the two types of measurements could be about 10 percent. Therefore, subbasin yields are only approximate and are not shown unless the difference in upstream and downstream loads was greater than 20 percent of the average of the two site loads.

### Dissolved Solids

Basin and subbasin yields of dissolved solids for the February 1983 samples are shown in figure 31, in $(ton/d)/mi^2$. Dissolved-solids yields were estimated, using a factor of 0.68 times specific conductance in $\mu S/cm$ to give dissolved-solids concentrations in mg/L, and then converting to yields.

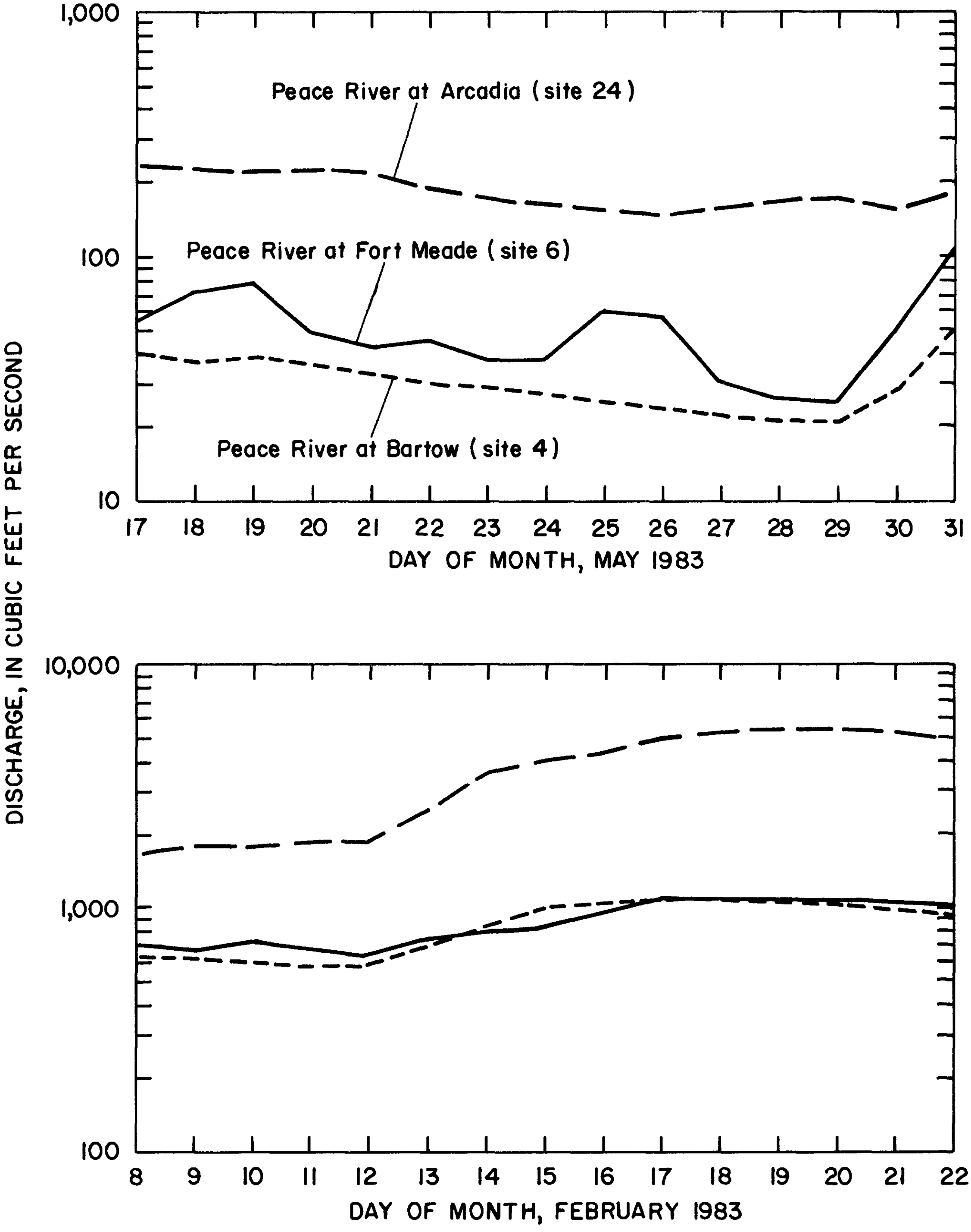


Figure 30.--Daily discharge for selected days in February and May 1983 at Peace River sites.

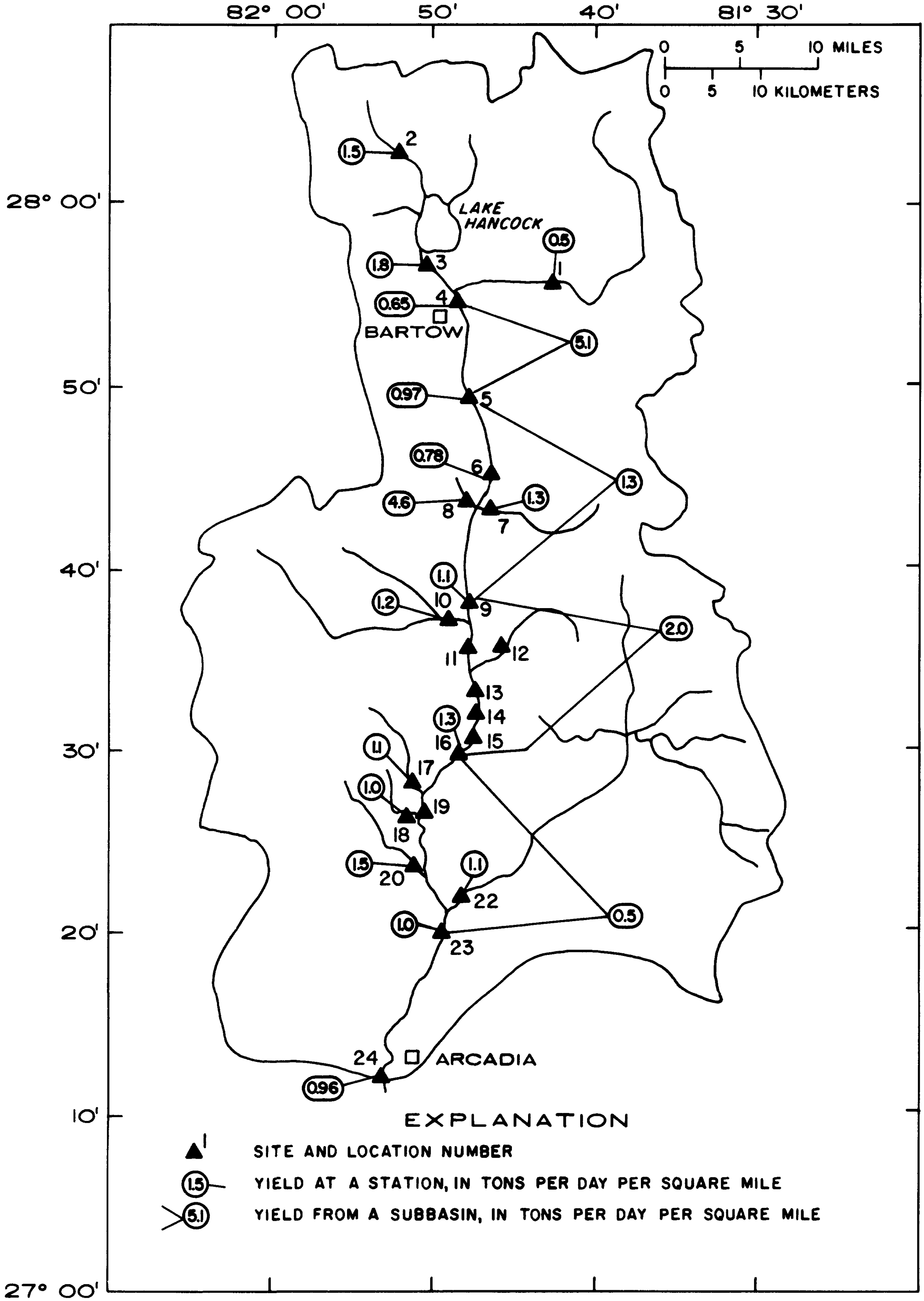


Figure 31.--Dissolved solids for Peace River subbasins, February 14-15, 1983.

Yields in the upper part of the Peace River basin, upstream from Bartow, were 1.5 (ton/d)/$mi^2$ upstream from Lake Hancock (site 2) and 1.8 (ton/d)/$mi^2$ immediately downstream from the lake (site 3). These yields were relatively high in comparison with the stations on the Peace River, but not as high as for some subbasins along the river. For example, the yield between sites 4 and 5 downstream from Bartow, was 5.1 (ton/d)/$mi^2$. This subbasin contains considerable phosphate mining area (fig. 1) and it seems likely that the high yield of dissolved solids was related to runoff from this activity. The Whidden Creek basin (site 8) also had a high dissolved solids yield (4.6 (ton/d)/$mi^2$), probably also related to the high density of phosphate mining in the basin. The basin yield at the NASQAN station (site 24) was 0.96 (ton/d)/$mi^2$ and was low in comparison to yields from part of the basin upstream.

Dissolved-solids yields for the May sampling trip, shown in figure 32, were 0.11 or less except at site 8 (Whidden Creek) where the yield was 0.7 (ton/d)/$mi^2$. Another area of high dissolved solids yield is the subbasin between sites 4 and 5 where the computed yield was 1.1 (ton/d)/$mi^2$. The input of dissolved solids between sites 4 and 5 in May was probably not from mined areas, because phosphorus yield from this area was not high, but may be due to discharge from industrial operations located between the two sites. Basin yield at the NASQAN station (site 24) was 0.09 (ton/d)/$mi^2$.

## Total Nitrogen

Total nitrogen yields, in (lb/d)/$mi^2$, are shown in figure 33 for the February sampling trip. The highest nitrogen yield is at site 3 (75 (lb/d)/$mi^2$), and probably is the result of outflow of water high in nitrogen from Lake Hancock. This lake may act as a nitrogen sink during dry periods, and water flushed from the lake during wet periods may convey the accumulated nitrogen downstream. Two other areas of relatively high nitrogen yield were located in the lower part of the basin, and had yields of 60 (lb/d)/$mi^2$, (site 20, Oak Creek), and 50 (lb/d)/$mi^2$ (site 17, Troublesome Creek). These nitrogen yields may be associated with runoff from agricultural areas. The basin yield at the NASQAN station (site 24) was 31 (lb/d)/$mi^2$.

In May, the area of highest nitrogen input was different than in February. Figure 34 shows that the subbasin between sites 4 and 5 contributed about 76 (lb/d)/$mi^2$ of nitrogen, and increased the basin yield from 0.6 (lb/d)/$mi^2$ at site 4 to 6 (lb/d)/$mi^2$ at site 5. The source of the nitrogen may be industrial operations between the two sites. The mined areas, or effluent from the Bartow sewage treatment plant, are unlikely as sources for much of the nitrogen, because these origins would also have contributed phosphorus. Nitrogen was lost from the river between sites 5 and 6, and to a lesser degree, between sites 6 and 9. This loss could be due to settling of suspended nitrogen-containing material, uptake of nitrogen by aquatic plants, or loss of nitrogen to the air through denitrification processes. The Troublesome Creek basin (site 17) had a high nitrogen yield in May (5 (lb/d)/$mi^2$) compared to other tributaries. The basin yield at the NASQAN station (site 24) was 2 (lb/d)/$mi^2$.

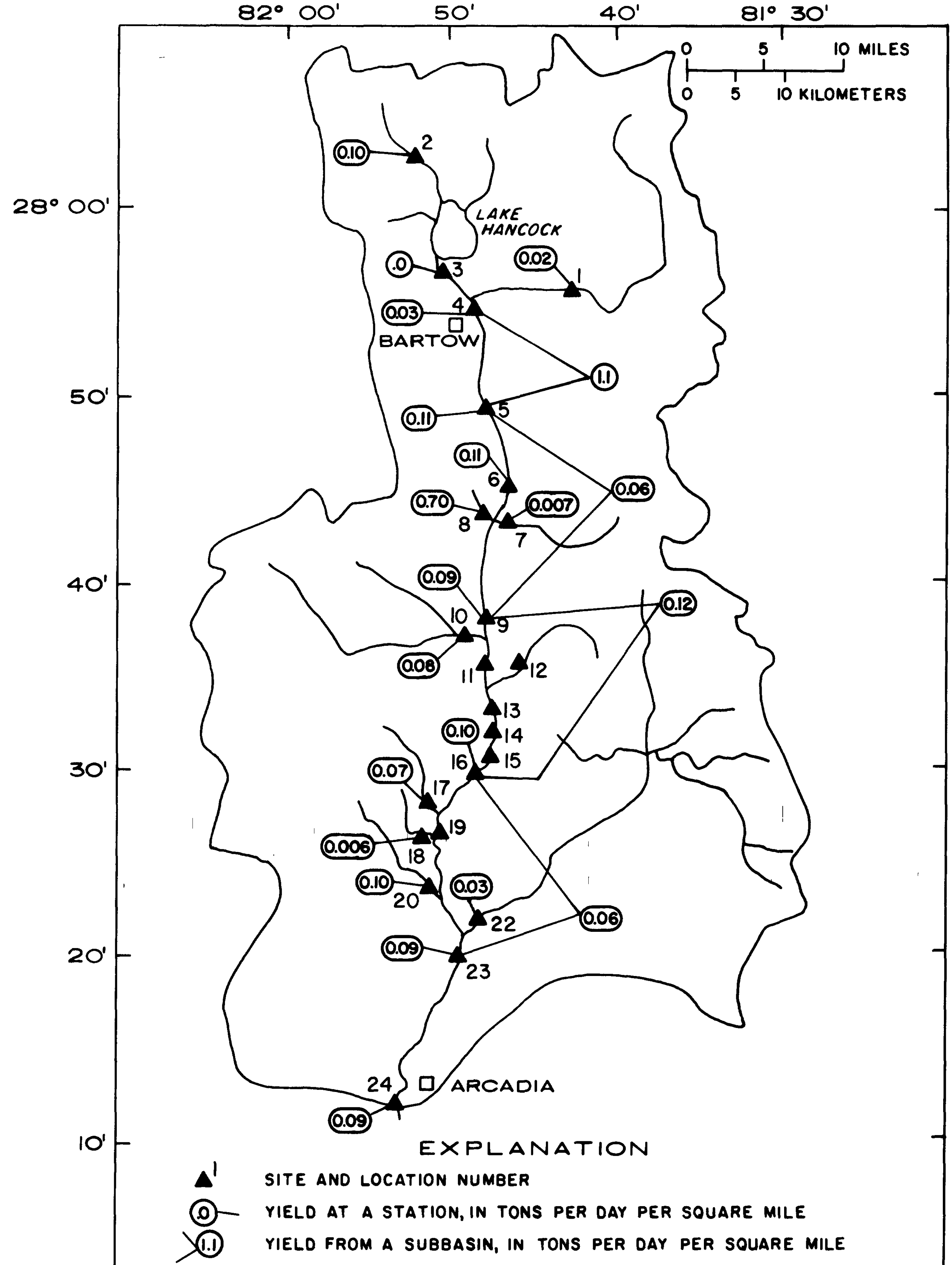


Figure 32.--Yields of dissolved solids for Peace River subbasins, May 24-25, 1983.

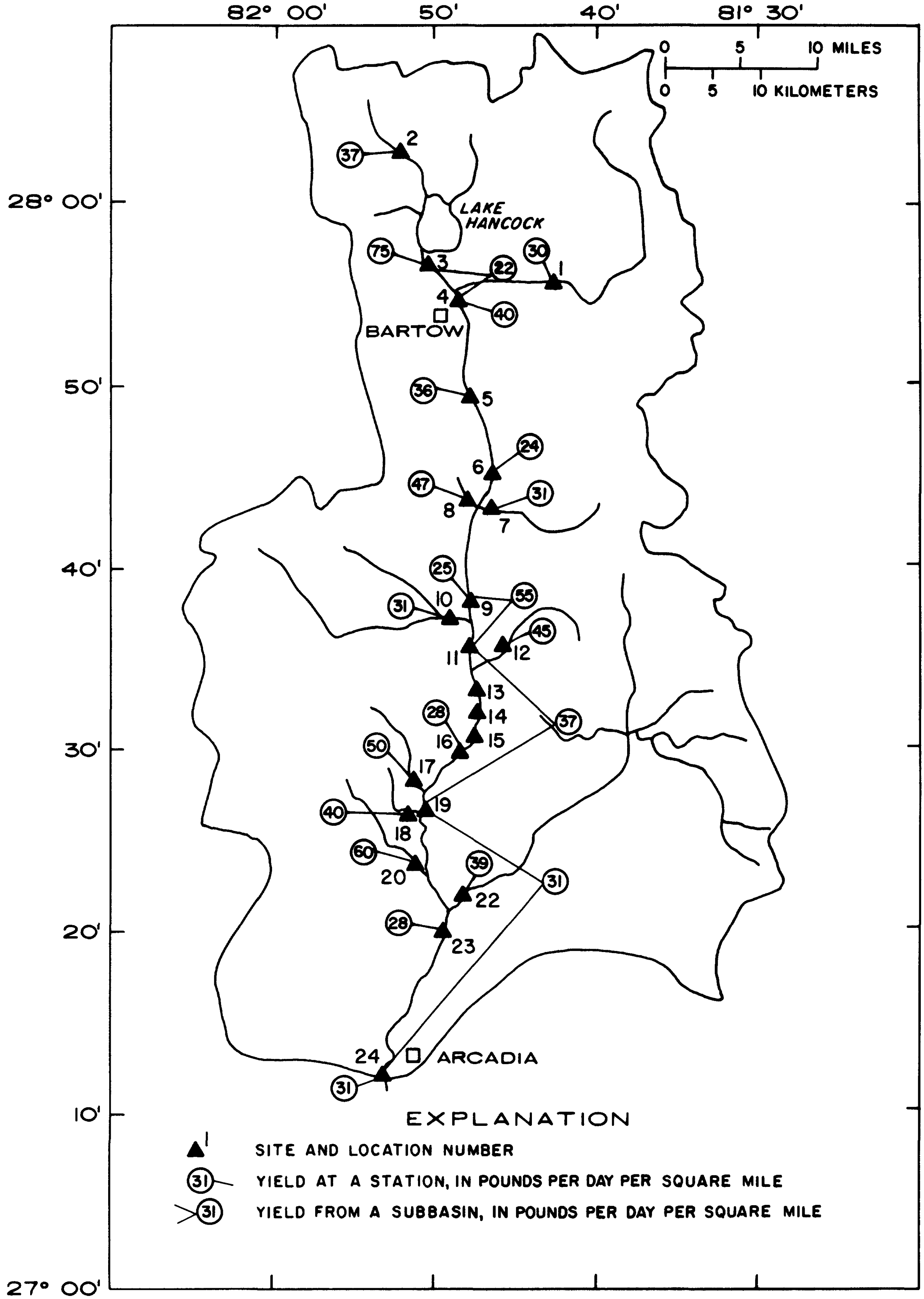


Figure 33.--Yields of total nitrogen for Peace River subbasins, February 14-15, 1983.

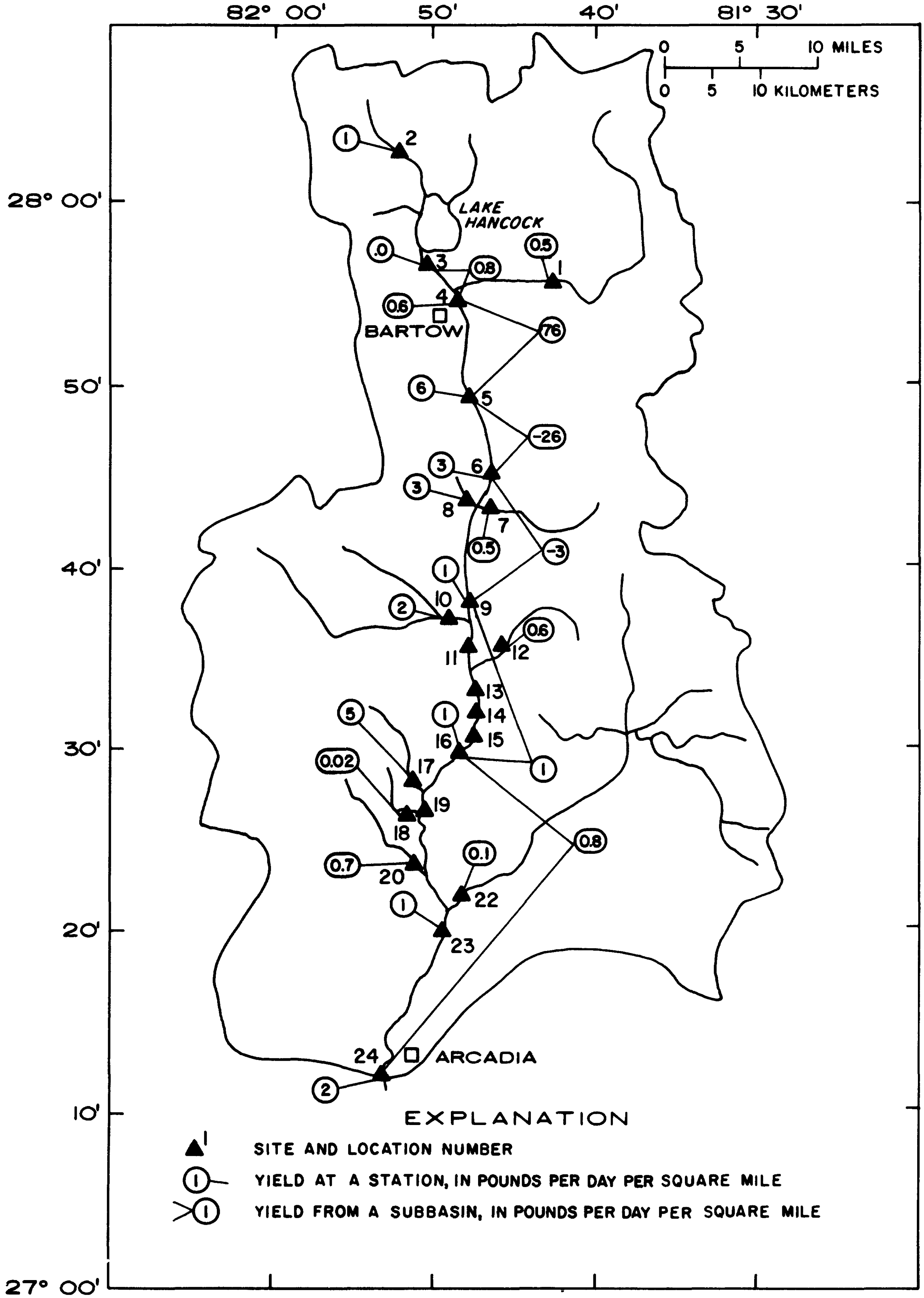


Figure 34.--Yields of total nitrogen for Peace River subbasins, May 24-25, 1983.

### Total Phosphorus

Total-phosphorus yields for February, shown in figure 35, show an area of extremely high phosphorus input (570 (lb/d)/mi$^2$) between sites 4 and 5. This high yield is probably due to runoff from the phosphate mines (fig. 1). Phosphorus was lost between sites 5 and 6, perhaps because of settling of phosphorus-bearing sediments. The Whidden Creek basin (site 8), and the area between sites 11 and 14, near Wauchula, were other areas of high phosphorus yield, 71 and 120 (lb/d)/mi$^2$, respectively. A phosphorus sink is indicated between sites 14 and 16. These data indicate that three areas of the basin were contributing relatively high phosphorus loads, and that two areas were removing substantial amounts of phosphorus from the river. However, it should be pointed out that the yields are based on only one sample at each site, and that more samples for high discharge conditions would be necessary to confirm this pattern of alternating high phosphorus input and phosphorus sink areas. The phosphorus yield at the NASQAN station (site 24) was relatively low (13 (lb/d)/mi$^2$).

In May, the pattern of phosphorus yields (fig. 36) was not as variable as in February. A significant input of phosphorus is indicated upstream of Bartow (site 4) because no water was flowing at site 3, and the yield at site 1 (0.2 (lb/d)/mi$^2$) was much less than at site 4 (2 (lb/d)/mi$^2$). Another area of high phosphorus yield was Whidden Creek (site 8), with 19 (lb/d)/mi$^2$. Yield at the NASQAN station (site 24) was 2 (lb/d)/mi$^2$.

One of the purposes of this study was to determine if the NASQAN station at Arcadia (site 24), could adequately represent water quality throughout the basin. Data from these intensive reconnaissance sampling trips clearly indicate that it does not. Yields of dissolved solids, total nitrogen, and total phosphorus calculated at Arcadia do not represent the entire basin, either at high or at low discharge conditions. This lack of representation of the entire basin by data for the NASQAN station is most acute for phosphorus, especially at high discharge conditions. Sources of phosphorus, as well as sinks, may exist upstream from Arcadia, and sampling at the Arcadia station, no matter how frequent, could not hope to reveal the dynamics of the phosphorus (and other parameters) transport processes throughout the basin. However, a continuous network adequate to define basin water quality at all points would obviously be extremely expensive to operate, due to the complexity of the water-quality processes involved. One possible solution is to combine fixed frequency-fixed site sampling with periodic intensive sampling so that quality of water leaving an accounting unit could be assessed and monitored for changes, with insight into variability of water quality throughout the basin being provided by the periodic reconnaissance studies.

## SUMMARY AND CONCLUSIONS

Selected water-quality data, collected periodically in the Peace River basin at and upstream from Arcadia, Fla., were analyzed to determine water-quality conditions, basin chemical constituent yields, and stream loadings. The data were examined for statistical evidence of trends, and to determine if sampling at the NASQAN station at Arcadia adequately represents the range of discharges at this station and water quality throughout the basin. In

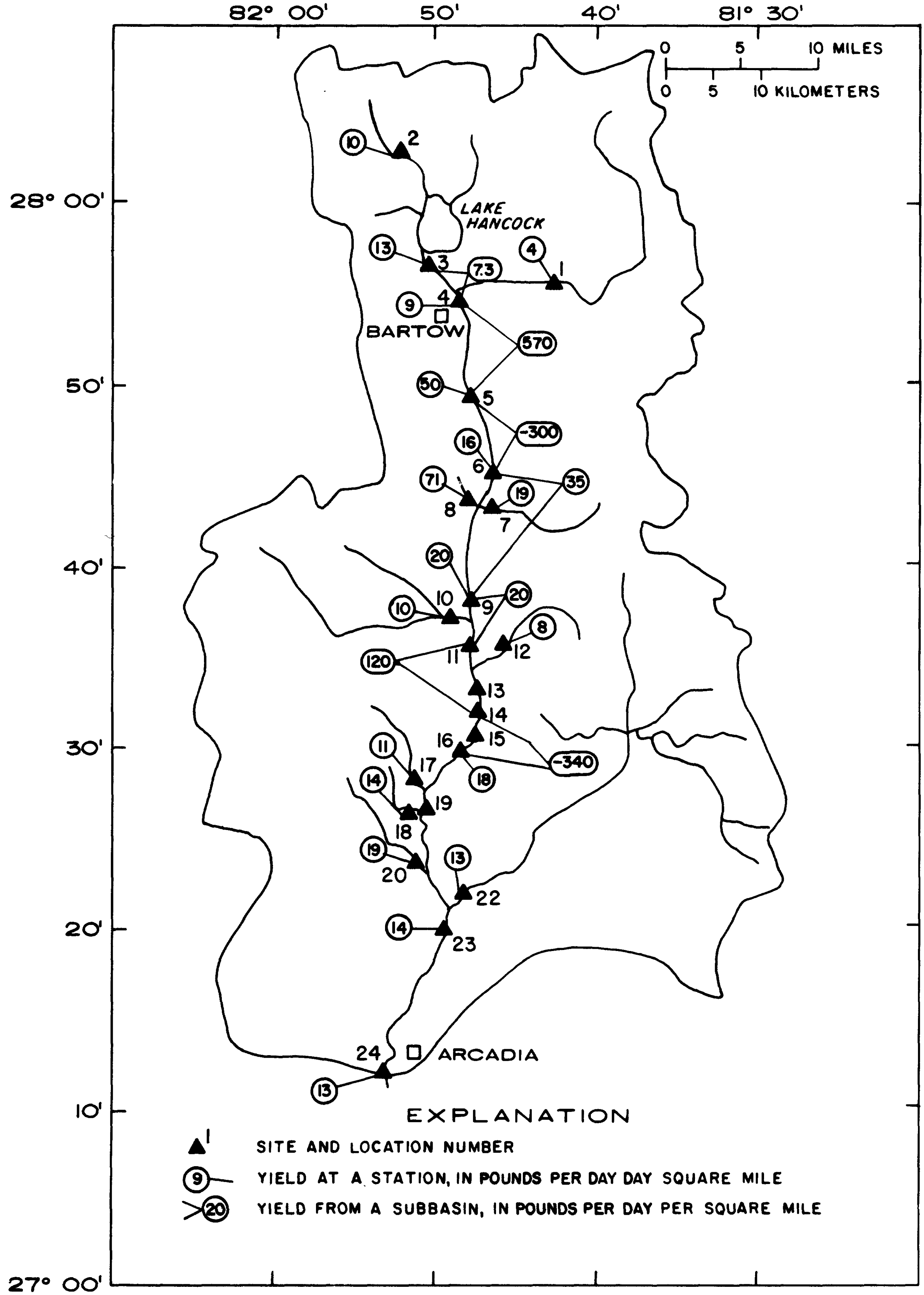


Figure 35.--Yields of total phosphorus for Peace River subbasins, February 14-15, 1983.

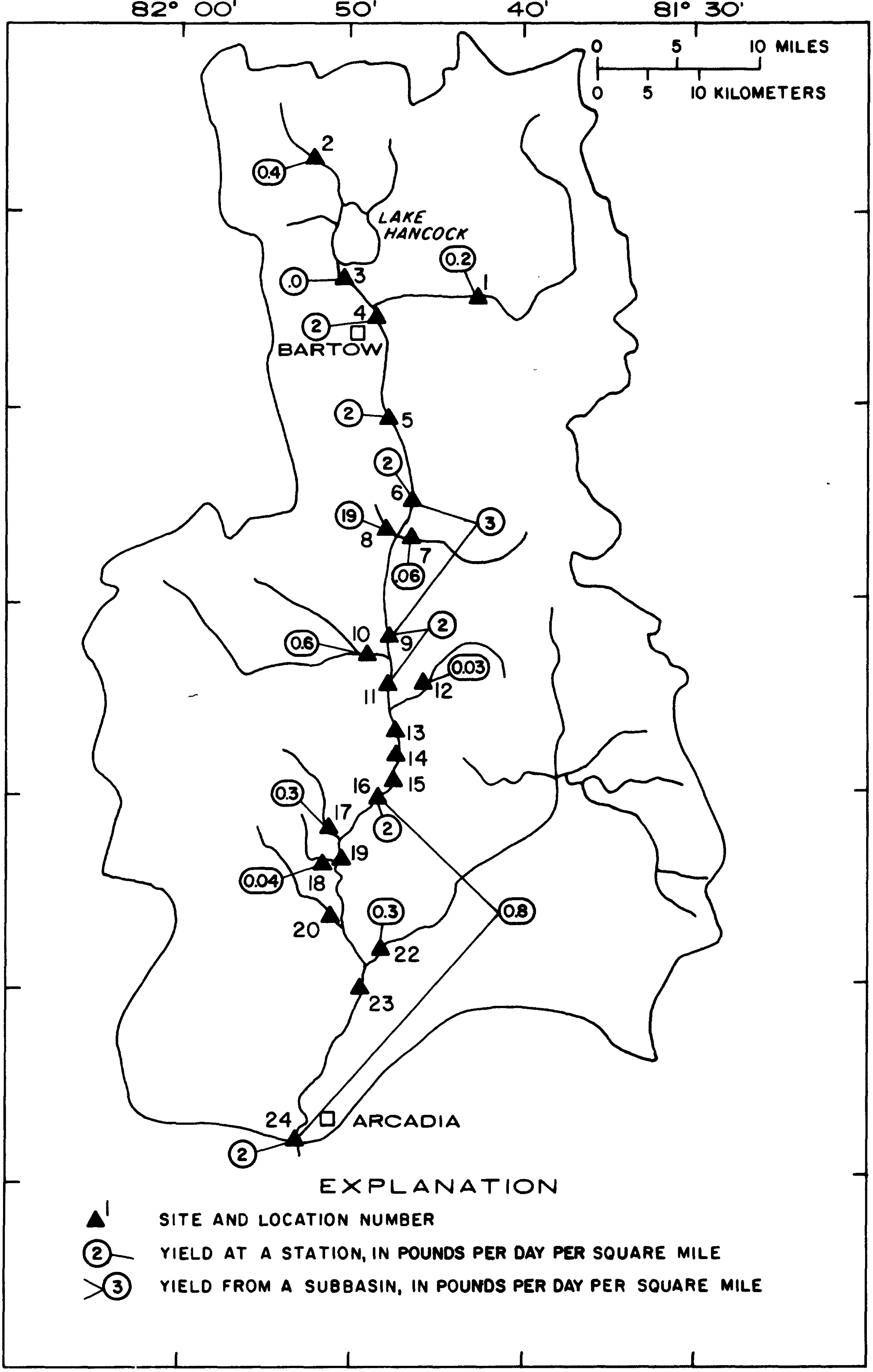


Figure 36.--Yields of total phosphorus for Peace River subbasins, May 24-25, 1983.

addition, data from two intensive reconnaissance studies, one in February 1983 during relatively high-discharge conditions, and the other in May 1983 during low-discharge conditions, were evaluated to determine locations of sources and sinks of selected water-quality variables. The major findings, conclusions, and suggestions for further work and an enhanced Peace River basin water-quality assessment network follow:

- Data collected during October 1974 to September 1982 indicate that water quality varies from station to station in the Peace River basin, and that data for the NASQAN station at Arcadia are not representative of the entire basin upstream from Arcadia. This finding is based on statistical testing of data from four to eight sites for specific conductance, total nitrogen, nitrate, and phosphorus, total recoverable iron and lead, and dissolved solids. Nonparametric statistical test results indicate that, except for concentrations of iron and lead, data probably are not from a single basinwide population. However, data are limited; for example, iron and lead concentration data are available only at four sites including Arcadia.

- Total phosphorus and nitrogen concentrations at eight sites commonly exceeded the desirable level of 0.1 mg/L (phosphorus) and 1 mg/L (nitrogen) during October 1974 through September 1982. According to these criteria, total phosphorus concentrations were excessive in nearly all samples from the eight sites, and total nitrogen concentrations were excessive in most samples at two Peace River sites.

- Analysis of mean basin chemical constituent yields at sites that have sufficient water-quality data for October 1974 through September 1982 indicates that input of some chemical constituents vary within the basin. A source of dissolved solids input to the river is located between Bartow (site 4) and Fort Meade (site 6), and between Fort Meade and Zolfo Springs (site 15). Total phosphorus yields increased from 0.9 (ton/yr)/mi$^2$ at Bartow to 1.3 (ton/yr)/mi$^2$ at Zolfo Springs.

- Trends in water quality were not detected at the Arcadia NASQAN station for October 1974 through September 1982, but were detected at other sites upstream. Statistical tests indicate that two of eight sites show an increase in specific conductance since 1974, although long-term trends in specific conductance were indicated at four of eight sites using data as early as 1951. One site shows a trend in total nitrogen since 1974. Dissolved orthophosphate concentration data since 1959 at the NASQAN station indicate an increase in orthophosphate residuals (that is, flow-adjusted concentrations), but not in orthophosphate concentrations. This is an indication that the relation between dissolved orthophosphate and stream discharge has changed.

- Samples for determination of trace metals, nutrients, and specific conductance were, in general, collected so as to be representative of the range of discharge at the NASQAN station for October 1974 through September 1982.

- A Monte Carlo analysis of 1,000 years of simulated sampling for lead, zinc, specific conductance, nitrogen, and phosphorus at the NASQAN station indicated that the optimum sampling frequency for predicting annual mean values is probably weekly to biweekly. However, this frequency of sampling would probably not be practical for the entire NASQAN network. The use of concentration-discharge relations could reduce the number of samples necessary for estimating mean annual concentrations.

- Water-quality data from two intensive reconnaissance sampling trips in 1983 indicate that basin chemical constituent yields were not uniform throughout the Peace River basin. This supports conclusions drawn from analysis of mean basin chemical constituent yields for the period 1974-82 at the few sites with periodic water-quality data in the basin. Yields for some parts of the basin upstream from the NASQAN station at Arcadia were much higher than the overall yields for the entire NASQAN basin. For example, dissolved-solids yield in the May 1983 sampling was 1.1 $(ton/d)/mi^2$ in a subbasin downstream from Bartow, and was only 0.09 $(ton/d)/mi^2$ for the entire basin upstream from the NASQAN station. The reconnaissance data indicate that some segments of the Peace River apparently serve as sinks for nitrogen and phosphorus, though repeated sampling would be necessary to confirm this.

- NASQAN station data can be used to assess quality of water leaving the accounting unit, but these data do not represent water-quality conditions throughout the basin. Sources and sinks of materials in the basin cannot be identified using only data collected at the NASQAN station. Factors affecting water quality in upper parts of the basin may not affect water quality at the NASQAN station to a measurable degree. Despite these shortcomings, the systematic collection of data at a fixed point (the NASQAN station) seems the most practical way to evaluate water quality and changes in water quality.

If funds were available, the NASQAN concept could be expanded to meet most water-quality monitoring objectives by expanding the network to include most basins and subbasins of any preselected size. This will not be possible in most cases, however. Some of the information unavailable to the NASQAN network could be provided by periodic intensive and short-term reconnaissance studies such as those summarized in this report. These can provide a nearly instantaneous picture of water quality under a variety of flow conditions. Some considerations in the usage of intensive reconnaissance studies are:

- Discharge and water-quality data need to be measured at each site;

- Studies need to be done under steady-state discharge conditions, insofar as possible;

- Studies could be repeated during low, medium, and high discharge conditions;

- Studies could be supplemented by continuous record of specific conductance (and perhaps other parameters) for several days before, during, and after the sampling, at selected points in the basin.

A disadvantage of intensive sampling studies is that interpretations are based on single samples at each location, so that errors in sampling or analysis could easily lead to erroneous conclusions. For this reason, it would be highly desirable for quality-control measures, such as duplicate sampling, to be designed into the studies. Unusual findings of the studies could be verified by resampling. The appropriate sampling frequency and areal coverage of these studies will not be the same for all basins, but consideration might be given to conducting two reconnaissance samplings of NASQAN accounting units, one under low-flow conditions and another under high-flow conditions during a 1-year or 2-year period.

## SELECTED REFERENCES


Anderson, J.R., Hardy, E.E., Roach, J.T., and Witmer, R.E., 1976, A land use and land cover classification system for use with remote sensor data: U.S. Geological Survey Professional Paper 964, 28 p.

Conover, W.J., 1971, Practical nonparametric statistics: New York, John Wiley & Sons, Inc., 462 p.

Crawford, C.G., Slack, J.R., and Hirsch, R.M., 1983, Nonparametric tests for trends in water-quality data using the Statistical Analysis System: U.S. Geological Survey Open-File Report 83-550, 102 p.

Daniel, W.N., 1978, Applied nonparametric statistics: Boston, Houghton Mifflin Company, 503 p.

Ficke, J.F., and Hawkinson, R.O., 1975, The National Stream Quality Accounting Network (NASQAN) - some questions and answers: U.S. Geological Survey Circular 719, 23 p.

Florida Department of Environmental Regulation, 1983, Water quality standards: Chapter 17-3 in Florida Administrative Code.

Gelb, A., 1974, Applied optimal estimation: Cambridge, Mass., The Massachusetts Institute of Technology Press, 374 p.

Gunst, R.F., and Mason, R.L., 1980, Regression analysis and its application: New York, Marcel Dekker, Inc., 402 p.

Heien, D.M., 1968, A note on log-linear regression: Journal of the American Statistical Association, v. 3, p. 1034-1038.

Hem, J.D., 1970, Study and interpretation of the chemical characteristics of natural water: U.S. Geological Survey Water-Supply Paper 1473, 363 p.

Hirsch, R.M., Slack, J.R., and Smith, R.A., 1982, Techniques of trend analysis for monthly water-quality data: Water Resources Research, v. 18, no. 1, p. 107-121.

Kendall, Maurice, 1975, Rank correlation methods: London, Charles Griffin and Co., Ltd., 202 p.

Mackenthun, K.M., 1973, Toward a cleaner aquatic environment: Washington, D.C., U.S. Environmental Protection Agency.

Moss, M.E., and Gilroy, E.J., 1980, Cost-effective stream-gaging strategies for the Lower Colorado River Basin: U.S. Geological Survey Open-File Report 80-1048, 111 p.

SAS Institute, Inc., 1980, SAS Supplemental Library User's Guide, 1980 edition: Cary, N.C., SAS Institute, Inc., 202 p.

----- 1982a, SAS User's Guide Basics, 1983 edition: Cary, N.C., SAS Institute, Inc., 924 p.

----- 1982b, SAS User's Guide Statistics, 1982 ed.: Cary, N.C., SAS Institute, Inc., 584 p.

Smith, R.A., Hirsch, R.M., and Slack, J.R., 1982, A study of trends in total phosphorus measurements at NASQAN stations: U.S. Geological Survey Water-Supply Paper 2190, 34 p.

Stark, P.A., 1970, Introduction to numerical methods: New York, MacMillan Publication Co., Inc., 334 p.

Tukey, J.W., 1977, Exploratory data analysis: Reading, Mass., Addison-Wesley Publishing Co., 688 p.

U.S. Department of Commerce, 1983, Climatological data: Florida, February 1983, v. 87, no. 2, p. 4.

Water Resources Council, 1970, Water resources regions and subregions for the national assessment of water and related land resources: Washington, D.C., U.S. Water Resources Council, 188 p., 6 plates.


☆U.S. GOVERNMENT PRINTING OFFICE: 1988-530-101/60018