

Quality of Nutrient Data from Streams and Ground Water Sampled During 1993–95—National Water-Quality Assessment Program

By David K. Mueller

U.S. GEOLOGICAL SURVEY
Open-File Report 98–276

Denver, Colorado
1998

U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, Secretary

U.S. GEOLOGICAL SURVEY
Thomas J. Casadevall, Acting Director

The use of firm, trade, and brand names in this report is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey.

For additional information write to:

Chief, NAWQA Program
U.S. Geological Survey
Water Resources Division
12201 Sunrise Valley Drive, Mail Stop 413
Reston, VA 20192

Copies of this report can be purchased
from:

U.S. Geological Survey
Information Services
Box 25286
Federal Center
Denver, CO 80225

FOREWORD

The mission of the U.S. Geological Survey (USGS) is to assess the quantity and quality of the earth resources of the Nation and to provide information that will assist resource managers and policy-makers at Federal, State, and local levels in making sound decisions. Assessment of water-quality conditions and trends is an important part of this overall mission.

One of the greatest challenges faced by water-resources scientists is acquiring reliable information that will guide the use and protection of the Nation's water resources. That challenge is being addressed by Federal, State, interstate, and local water-resource agencies and by many academic institutions. These organizations are collecting water-quality data for a host of purposes that include: compliance with permits and water-supply standards; development of remediation plans for a specific contamination problem; operational decisions on industrial, wastewater, or water-supply facilities; and research on factors that affect water quality. An additional need for water-quality information is to provide a basis on which regional and national policy decisions can be based. Wise decisions must be based on sound information. As a society we need to know whether certain types of water-quality problems are isolated or ubiquitous, whether there are significant differences in conditions among regions, whether the conditions are changing over time, and why these conditions change from place to place and over time. The information can be used to help determine the efficacy of existing water-quality policies and to help analysts determine the need for, and likely consequences of, new policies.

To address these needs, the Congress appropriated funds in 1986 for the USGS to begin a pilot program in seven project areas to develop and refine the National Water-Quality Assessment (NAWQA) Program. In 1991, the USGS began full implementation of the program. The NAWQA Program builds upon an existing base of water-quality studies of the USGS, as well as those of other Federal, State, and local agencies. The objectives of the NAWQA Program are to:

- Describe current water-quality conditions for a large part of the Nation's freshwater streams, rivers, and aquifers.

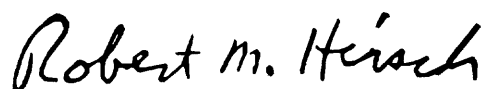
- Describe how water quality is changing over time.
- Improve understanding of the primary natural and human factors that affect water-quality conditions.

This information will help support the development and evaluation of management, regulatory, and monitoring decisions by other Federal, State, and local agencies to protect, use, and enhance water resources.

The goals of the NAWQA Program are being achieved through ongoing and proposed investigations of 59 of the Nation's most important river basins and aquifer systems, which are referred to as study units. These study units are distributed throughout the Nation and cover a diversity of hydrogeologic settings. More than two-thirds of the Nation's freshwater use occurs within the 59 study units and more than two-thirds of the people served by public water-supply systems live within their boundaries.

National synthesis of data analysis, based on aggregation of comparable information obtained from the study units, is a major component of the program. This effort focuses on selected water-quality topics using nationally consistent information. Comparative studies will explain differences and similarities in observed water-quality conditions among study areas and will identify changes and trends and their causes. The first topics addressed by the national synthesis are pesticides, nutrients, volatile organic compounds, and aquatic biology. Discussions on these and other water-quality topics will be published in periodic summaries of the quality of the Nation's ground and surface water as the information becomes available.

This report is an element of the comprehensive body of information developed as part of the NAWQA Program. The program depends heavily on the advice, cooperation, and information from many Federal, State, interstate, Tribal, and local agencies and the public. The assistance and suggestions of all are greatly appreciated.



Robert M. Hirsch
Chief Hydrologist

CONTENTS

Foreword	III
Glossary	VII
Abstract	1
Introduction.....	1
Purpose and Scope	3
Acknowledgments	3
Nutrients in Streams and Ground Water	3
Types of Quality-Control Samples	4
Compilation of Environmental and Quality-Control Data.....	4
Methods of Data Analysis	4
Methods Used to Determine Bias	5
Methods Used to Determine Variability	6
Quality of the Nutrient Data	8
Contamination Bias	9
Sampling Variability	11
Implications for Interpretation of Environmental Data	21
Potential Effects of Contamination.....	21
Potential Effects of Sampling Variability	22
Effects on Measured Concentrations	22
Effects on Comparisons Between Concentrations.....	23
Conclusions.....	25
References.....	25

FIGURES

1. Map showing location of the 20 National Water-Quality Assessment study units that contributed to the data analyzed in this report	2
2–11. Graphs showing:	
2. Example of a nonlinear model of replicate variability compared with the alternative, piecewise-linear model.....	7
3. Distribution of nutrient concentrations measured in field blanks from National Water-Quality Assessment stream and ground-water sites.....	9
4. Selected percentiles of the distributions of ammonia contamination in source-solution and field blanks.....	11
5. Nonlinear-model estimates of variability for nitrogen analytes in stream samples in comparison to smoothed curves through the center of the replicate data	14
6. Nonlinear-model estimates of variability for phosphorus analytes in stream samples in comparison to smoothed curves through the center of the replicate data	15
7. Residuals of standard deviation estimated by nonlinear models of variability for nitrogen analytes in stream samples	16
8. Residuals of standard deviation estimated by nonlinear models of variability for phosphorus analytes in stream samples	17
9. Nonlinear-model estimates of variability for selected nutrient analytes in ground-water samples in comparison to smoothed curves through the center of the replicate data	18
10. Residuals of standard deviation estimated by nonlinear models of variability for selected nutrient analytes in ground-water samples	19
11. Piecewise-linear model estimate of variability for dissolved phosphorus in ground-water samples in comparison to smoothed curves through the center of the replicate data	20

TABLES

1. Number of quality-control samples collected in each of the 20 National Water-Quality Assessment study units that contributed to the data analyzed in this report.....	4
2. Upper 90-percent tolerance bounds that exceed contamination by nutrient analytes in specified percentages of all samples based on data from field blanks prepared at stream and ground-water sampling sites	10
3. Parameter estimates for nonlinear models of sampling variability for selected nutrient analytes in streams and ground water.....	20
4. Estimates of standard deviation from piecewise-linear models of sampling variability for selected nutrient analytes in ground water	21
5. Maximum nutrient concentrations that are considered potentially affected by contamination, based on the 1991 National Water-Quality Assessment data, and selected critical values used to interpret environmental data	22
6. Estimated sampling variability and confidence intervals around measured concentrations of nutrient analytes at selected critical values used to interpret environmental data.....	23

GLOSSARY

quality control.—Data generated to estimate the magnitudes of bias and variability in the processes used for obtaining environmental data.

bias.—The systematic error inherent in a method. Bias may be either positive or negative.

contamination bias.—A common source of positive bias due to contaminants introduced into water samples during field processing, shipping, or laboratory analysis.

blank.—A water sample that is intended to be free of the analytes of interest. Blank samples are used to test for contamination bias.

field blank.—A blank sample prepared at a field location and exposed to all equipment normally used to obtain an environmental sample.

source-solution blank.—A sample of blank water taken directly from the source container without exposure to any sampling equipment.

variability.—The degree of random error in independent measurements of the same quantity.

sampling variability.—The variability introduced by sample collection, field processing, shipping, and laboratory analysis.

replicates.—Two or more samples collected or processed in a manner such that the samples are thought to be essentially identical in composition.

split replicates.—Replicates prepared by dividing a single volume of water into multiple samples.

concurrent replicates.—Multiple samples collected from an environmental matrix at the same location at the same time.

sequential replicates.—Multiple samples collected at the same location but at slightly different times, generally one right after the other.

Quality of Nutrient Data from Streams and Ground Water Sampled During 1993–95—National Water-Quality Assessment Program

By David K. Mueller

Abstract

Proper interpretation of water-quality data requires consideration of the effects that **bias** and **variability** might have on measured constituent concentrations. In this report, methods are described to estimate the bias due to contamination of samples in the field or laboratory and the variability due to sample collection, processing, and analysis. These methods are applied to **quality-control** data collected as part of the National Water-Quality Assessment (NAWQA) Program during 1993–95. Results are presented for seven nutrient analytes from stream samples and five nutrient analytes from ground-water samples. Contamination potentially affects measured concentrations of less than 0.4 milligram per liter for ammonia and less than 0.1–0.3 milligram per liter for various phosphorus analytes. Nitrite plus nitrate and the Kjeldahl nitrogen analytes are essentially unaffected by contamination at concentrations of environmental significance. **Sampling variability** can affect interpretation of small differences between individual measurements or mean concentrations. For the NAWQA data, there is essentially no effect of variability on measured concentrations of nitrite plus nitrate and orthophosphate, nor for ammonia at concentrations of about 0.2 milligram per liter or less. The potential errors due to variability are larger for dissolved and total Kjeldahl nitrogen and phosphorus, but the uncertainty in measured or mean concentrations caused by sampling

variability is within a small range for all nutrients. Differences in concentration within these ranges have limited, if any, environmental significance. These results can be applied to interpretation of environmental data collected during 1993–95 in 20 NAWQA study units.

INTRODUCTION

To determine the extent of contamination in the Nation's streams and ground water, Congress has appropriated funds for a National Water-Quality Assessment (NAWQA) Program, conducted by the U.S. Geological Survey (USGS). The objectives of the NAWQA Program are to:

- Describe current water-quality conditions for a large part of the Nation's freshwater streams, rivers, and aquifers.
- Describe how water quality is changing over time.
- Improve understanding of the primary natural and human factors that affect water-quality conditions.

These objectives are being achieved through investigations in 59 large river basins and aquifer systems, which are referred to as study units. Implementation of study-unit investigations are phased so that data are collected in about one-third of the study units at a time. Investigations in 20 study units began in 1991, and most water samples were collected during 1993–95. The location of these study units is shown in figure 1.

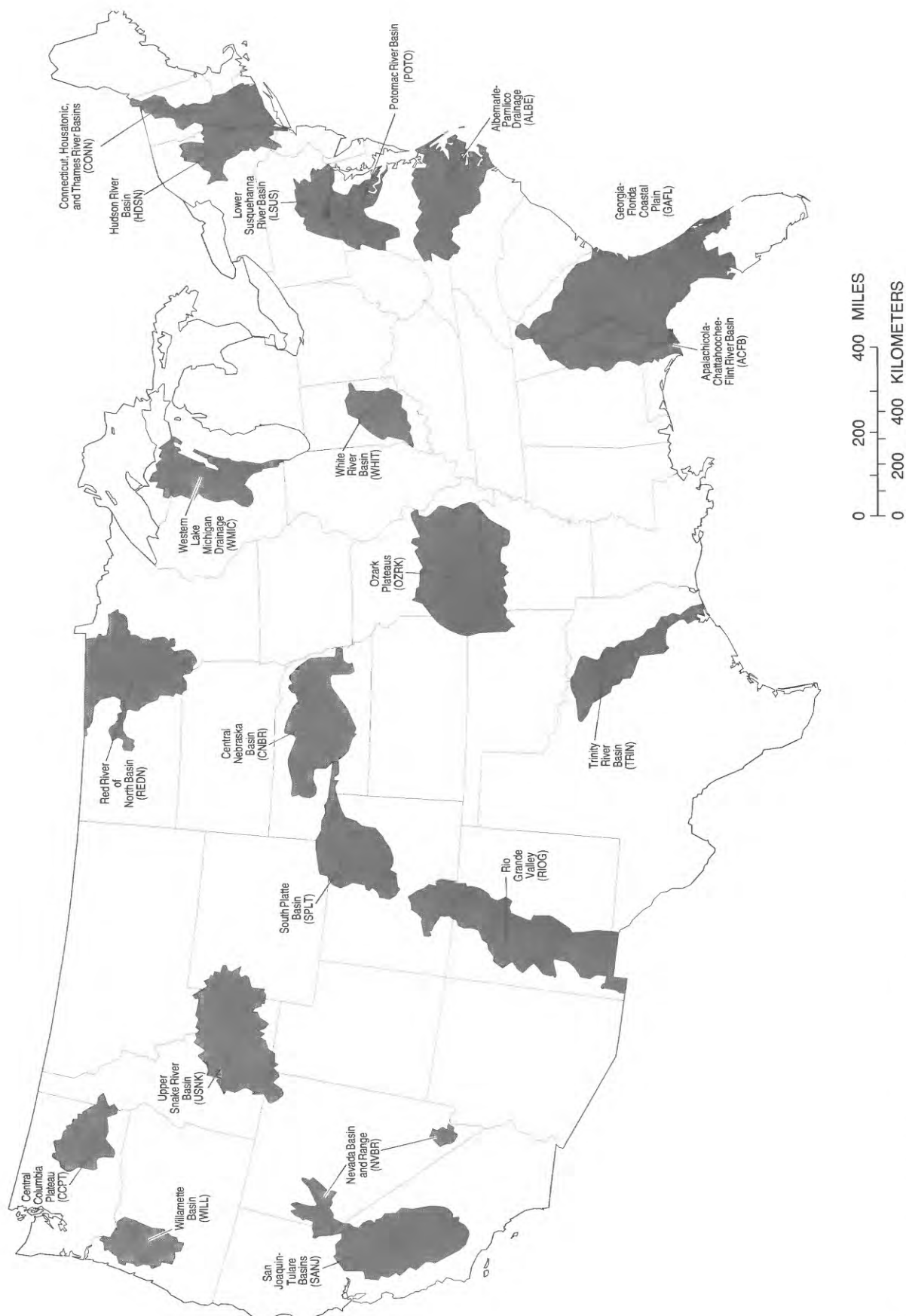


Figure 1. Location of the 20 National Water-Quality Assessment study units that contributed to the data analyzed in this report.

In order to properly interpret water-quality data, information is needed to estimate the bias and variability that result from sample collection, processing, shipment, and chemical analysis. Bias is the systematic error inherent in a method and may be either positive or negative. A common source of positive bias that can affect water-quality data is contamination of samples. Contaminants can be introduced into water samples during field processing, shipping, or laboratory analysis through exposure to airborne gases and particulates or to inadequately cleaned sampling or analytic equipment. Variability is the degree of random error in independent measurements of the same quantity. In water-quality data, variability results from the error inherent in laboratory analytic procedures and in collecting representative samples in the field. Sampling variability includes analytic variability plus the variability introduced by sample collection, field processing, and shipping. **Contamination bias** and sampling variability are evaluated by collecting and analyzing quality-control (QC) samples in addition to the environmental samples collected as part of a water-quality assessment. A glossary at the beginning of this report defines QC terms that might be unfamiliar to some readers.

Purpose and Scope

This report describes the analysis of contamination bias and sampling variability for nutrient constituents in stream and ground-water samples collected in 20 NAWQA study units during 1993–95. The results of the QC data analysis are compared to characteristics of the environmental data and to national water-quality standards and criteria to assess the potential effects of bias and variability on interpretation of the environmental data.

The NAWQA study units that provided the QC data represent a broad array of hydrologic conditions in the 48 contiguous States. The data were aggregated into a national data set in the summer of 1995, so not all samples collected during the high-intensity phase of sampling in each study unit are included.

Acknowledgments

The philosophy of QC data interpretation followed in this report was developed by the NAWQA QC work group: Michael Koterba, Terry Schertz,

Jeff Martin, Greg Delzer, and Jon Scott. In addition, Ed Gilroy (USGS, retired), Jeff Pritt (National Water-Quality Laboratory), and Mark Brigham (Red River of the North NAWQA study unit) provided invaluable advice and review of the statistical methods used for data analysis. Finally, this report relies on data that were collected by hydrologists and hydrologic technicians in the first 20 NAWQA study units. Without their diligent efforts, none of this analysis would have been possible.

NUTRIENTS IN STREAMS AND GROUND WATER

Nutrients are chemical elements that are essential to plant and animal nutrition. Nitrogen and phosphorus are nutrients that are important to aquatic life, but in high concentrations, they can be contaminants in water. These nutrients occur in a variety of forms. Both are affected by chemical and biological processes that can change their form and can transfer them to or from water, soil, biological organisms, and the atmosphere. Nutrient concentrations in water are generally reported in milligrams per liter (mg/L) as nitrogen or phosphorus.

Ammonia, a compound of nitrogen and hydrogen, is one of the primary forms of dissolved nitrogen in natural water. Depending on the number of hydrogen atoms in the compound, ammonia in water may be ionic (having an electrical charge) or un-ionized (having no charge). The un-ionized form is more toxic to fish. Ammonia is soluble in water but is not stable in most environments. It usually is transformed biologically to nitrate in water that contains oxygen and can be transformed to nitrogen gas in water that is low in oxygen.

Nitrate, a compound of nitrogen and oxygen, is another primary form of dissolved nitrogen in natural water. Nitrate is highly soluble in water and is stable over a wide range of environmental conditions. It is readily transported in ground water and streams.

Phosphates, including orthophosphate, are the only significant form of dissolved phosphorus in natural water. They are compounds of phosphorus, oxygen, and hydrogen. Phosphates are only moderately soluble and tend to adhere to soil particles. Relative to nitrate, phosphates are not very mobile in soil and ground water; however, erosion can transport considerable amounts of phosphate-laden particulates to streams and lakes.

TYPES OF QUALITY-CONTROL SAMPLES

A **blank** is a water sample that is intended to be free of the analytes of interest. Blank samples are used to test for bias that could result from contamination during any stage of the sample collection and analysis process. A **field blank** is a specific type of blank sample used to demonstrate that: (1) Equipment has been adequately cleaned to remove contamination introduced by samples obtained at previous sites; (2) sample collection and processing have not resulted in contamination; and (3) sample handling, transport, and laboratory analysis have not introduced contamination.

Replicates are two or more samples collected or processed in a manner such that the samples are thought to be essentially identical in composition. **Split replicates** are prepared by dividing a single volume of water into multiple samples. They provide a measure of the variability introduced during sample processing and analysis. **Concurrent replicates** are multiple samples collected from an environmental matrix at the same location at the same time. They include the variability measured by split replicates and also the variability introduced by sample collection. Depending on sampling procedures, concurrent replicates also might include an unknown amount of short-term environmental variability. **Sequential replicates** are multiple samples collected at the same location but at slightly different times, generally one right after the other. They provide a measure of the same sources of variability as concurrent replicates, including environmental variability.

COMPILATION OF ENVIRONMENTAL AND QUALITY-CONTROL DATA

This report is based on data from samples collected during high-intensity sampling (1993–95) at locations within the 20 NAWQA study units shown in figure 1. The data were compiled before analyses were available for some samples; therefore, not all samples collected in each study unit during 1993–95 are included. The numbers of field blanks and replicate-sample sets used in this report from each study unit are listed in table 1.

Table 1. Number of quality-control samples collected in each of the 20 National Water-Quality Assessment study units that contributed to the data analyzed in this report

Study unit (see fig. 1 for full name and location)	Number of field blanks		Number of replicate-sample sets	
	Stream sites	Ground- water sites	Streams	Ground water
ACFB	27	6	14	26
ALBE	32	10	2	5
CCPT	28	26	23	20
CNBR	17	2	14	0
CONN	16	11	2	4
GAFL	22	5	18	5
HDSN	46	10	18	6
LSUS	16	10	11	11
NVBR	0	0	7	9
OZRK	20	10	10	10
POTO	16	7	47	10
REDN	4	17	11	16
RIOG	21	8	31	9
SANJ	4	0	8	11
SPLT	38	13	30	15
TRIN	7	9	9	6
USNK	19	13	8	9
WHIT	26	4	16	3
WILL	16	6	10	3
WMIC	17	5	24	5
Total	392	172	313	183

The results of chemical analyses on QC samples and associated environmental samples were provided by each study unit to a national data base. The nutrient data used in this report were retrieved from this data base and were subjected to a series of tests to identify potential errors. In some cases, corrections were provided by the study units.

METHODS OF DATA ANALYSIS

Bias and variability are determined by statistical analysis of blanks and replicate samples. For laboratory data, acceptable limits of bias and variability are routinely defined, and specific analytic results are

compared to these limits to ensure that processes remain in control. When results are outside control limits, corrective actions are taken. Methods of evaluating bias and variability in field data are not as well defined. Some statistical techniques used for laboratory data also are applicable to field data, but rounding and censoring of analytic values reported for field QC samples can limit the utility of some techniques. Also, the objective of field QC data analysis usually is different from the laboratory objective. Rather than ensuring that bias and variability are within acceptable limits, field QC data generally are used to determine the extent to which bias and variability might affect interpretation of existing environmental data. The methods used to evaluate bias and variability in this report were developed considering the characteristics of field data and the objectives of field QC analysis.

Methods Used to Determine Bias

Ideally, the bias introduced by contamination would be so small that concentrations in field blanks are less than the detection limit. In practice, concentrations typically are less than detection in many blanks, but some blanks can contain concentrations much greater than the detection limit. The objective in analyzing data from blanks is to determine how great the contamination might be in a large percentage of the water samples represented by the blanks. This objective can be achieved by constructing an upper tolerance bound, which is the concentration expected to exceed contamination in a specified percentage of water samples with an acceptable degree of confidence. For example, if the upper 90-percent tolerance bound that exceeds 95 percent of the population values is determined to be X mg/L, using a particular set of blank data, this implies that, with 90-percent confidence, contamination is expected to be less than X mg/L in 95 percent of all samples (including environmental samples) that were collected, processed, and analyzed in the same manner as the blanks.

Because the distribution of concentrations in blanks can be highly skewed, statistical techniques that rely on assumptions of normality are not applicable. Hahn and Meeker (1991) described a method for determining a one-sided distribution-free tolerance bound, which is appropriate for skewed data. This method uses order statistics, based on ranking the data

values from small to large, and binomial probability to determine the tolerance bound. The binomial function is used to calculate the probability that m observed values from a total of n observations are less than or equal to the 100 p th percentile of the sampled population. This probability is the actual confidence level (cl) of the tolerance bound:

$$cl = Prob(p, n, m) \quad (1)$$

The 100 cl percent upper tolerance bound to exceed values in at least 100 p percent of the population is then determined by the value of the $m+1$ ranked observation.

For example, in a group of 200 blanks, the probability that the contamination in 190 of those blanks is less than or equal to the 95th percentile of contamination within the population of all samples is calculated:

$$Prob(0.95, 200, 190) = 0.545 \quad (2)$$

Thus, the 54.5-percent (confidence-level) upper tolerance bound on contamination in at least 95 percent of all samples would be the 191st-ranked concentration in the 200 blanks.

Usually, determination of exact-integer confidence levels for tolerance bounds is not possible by this distribution-free method. This difficulty arises because the confidence level is calculated using discrete values of n and m . In order to achieve an acceptable confidence level, such as 90 or 95 percent, equation 1 is solved by iteratively increasing m until the calculated confidence level is greater than or equal to the desired level. For the previous example, if the desired confidence level is 90 percent, m must be increased to 194. The 90-percent upper tolerance bound is then the value of the 195th-ranked concentration.

Contamination bias in the environmental samples is estimated from the upper tolerance bound calculated using blank data. In the previous example, if the 195th-ranked concentration of compound Y in the blanks was 0.1 mg/L, contamination bias can be described as follows:

Contamination by Y is estimated, with at least 90-percent confidence, to exceed 0.1 mg/L in fewer than 5 percent of all samples.

This level of contamination can then be compared to environmentally critical concentrations of compound Y to determine the likelihood that contamination has affected interpretation of the environmental data. Critical concentrations might be defined by background concentrations at undisturbed sites, minimum concentrations that indicate human impacts, or concentrations used for water-quality standards and criteria. Continuing with the previous example, assume compound Y has a drinking-water standard of 10 mg/L. Contamination as great as 0.1 mg/L is unlikely to affect a measurement that exceeds this critical value, particularly since concentrations generally are reported to only two significant figures. However, if the standard for compound Y is 0.2 mg/L, contamination might account for almost one-half of a measured exceedance. In this case, potential contamination limits the utility of the data.

Methods Used to Determine Variability

Sampling variability is equivalent to the standard deviation determined from field replicates. If only one set of a large number of replicates was collected, the standard deviation could be calculated directly. However, the general practice is to collect many sets of a small number of replicates. Even in this case, if the standard deviations of the individual replicate sets were about the same, variability could be determined by a pooled estimate of standard deviation (Snedecor and Cochran, 1980). But this is not the usual circumstance for chemical constituents in water. For many constituents, standard deviation within replicate sets is related to the mean concentration of that constituent in the replicate samples. This relation is generally monotonic, standard deviation increasing with concentration, but it is not linear. At low concentrations, less than about 10 times the reporting limit, standard deviation of replicates is approximately constant. At increasingly larger concentrations, standard deviations are higher and often are approximated as a percentage of concentration. At high concentrations, about 100 or more times the reporting limit, standard deviation tends to become constant again. This sigmoidal relation can be modeled by:

$$s = B_1 + e^{\left(-B_2 - \frac{B_3}{\bar{C}}\right)} \quad (3)$$

where

s = the standard deviation of a set of replicates,

\bar{C} = the mean concentration of the replicates,

B_1 , B_2 , and B_3 = model parameters that are fit to specific sets of replicate data, and

e = the base of natural logarithms.

At low concentrations, the exponent becomes very negative, so the exponential term approaches zero, and the standard deviation is approximately constant at a value of B_1 . At high concentrations, the exponent approaches a constant value of $-B_2$; therefore, standard deviation approaches a constant of $B_1 + e^{-B_2}$. Between these two limits, standard deviation increases monotonically as a function of replicate concentration.

A number of problems can interfere with achieving a good model fit to field-replicate data. The distribution of constituent concentrations among sets of field replicates is not likely to be uniform because frequency of occurrence typically is inverse to concentration. Thus, low concentrations generally dominate field-replicate data, and few or no data might be available at high concentrations. In this case, the upper part of the curve could be impossible to define. Another problem results from laboratory rounding of the analyzed concentrations. The possible differences among rounded concentration values are not continuous, but occur at discrete intervals that change with the order of magnitude of concentration. Thus, standard deviations can be defined with better resolution for low concentration replicates than for high concentration replicates. Again, fitting the upper part of the model curve might be adversely affected.

If the nonlinear model cannot be adequately fit, an alternative method is required to estimate variability. The method that most closely approximates the nonlinear model results is to compute linear estimates of variability for specific ranges of mean replicate concentration. In the low concentration range, standard deviation of replicates is approximately constant; therefore, variability is estimated simply as the mean of the standard deviations. In the middle concentration range, standard deviation generally increases with concentration; therefore, variability is estimated by the mean of the relative standard deviation (*RSD*), which is defined:

$$RSD = \frac{s}{\bar{C}} \quad (4)$$

In the high concentration range, standard deviations also have a broad range, but generally have no strong relation with concentration; therefore, variability is again estimated as the mean of the standard deviations.

An example comparison between the nonlinear model and the alternative method for estimating replicate variability is shown in figure 2. The data used in this example are for concentrations of nitrite plus nitrate in ground-water replicates from the NAWQA data base. In general, the piecewise-linear model, produced by the alternative method, follows the curve of the nonlinear model.

Concentration ranges for the alternative method are selected by graphical analysis of standard deviation or *RSD* in relation to mean concentration. Appropriate boundary values between ranges are determined by a change in slope of a curve, such as a spline

smooth (SAS Institute, 1990) or a locally weighted scatterplot (LOWESS) smooth (Chambers and others, 1983), through the center of the data. For the NAWQA nutrient data, boundaries between the low and middle concentration ranges typically were about 10 times the reporting limit, and boundaries between the middle and high concentration ranges were about 100 times the reporting limit. Usually the boundaries seemed to occur at order-of-magnitude changes in concentration. This characteristic of the relation between standard deviation and concentration probably is a result of data rounding. For the nitrite plus nitrate data used in figure 2, the reporting limit was 0.05 mg/L, the boundary between the low and middle concentration ranges was set at 1 mg/L, and the boundary between the middle and high concentration ranges was set at 10 mg/L. For some constituents, the overall extent of the measured concentrations was small, so one or two of the ranges could not be defined. Also, discontinuities can occur in the linear model if the fitted lines do not intersect precisely at the boundary concentrations.

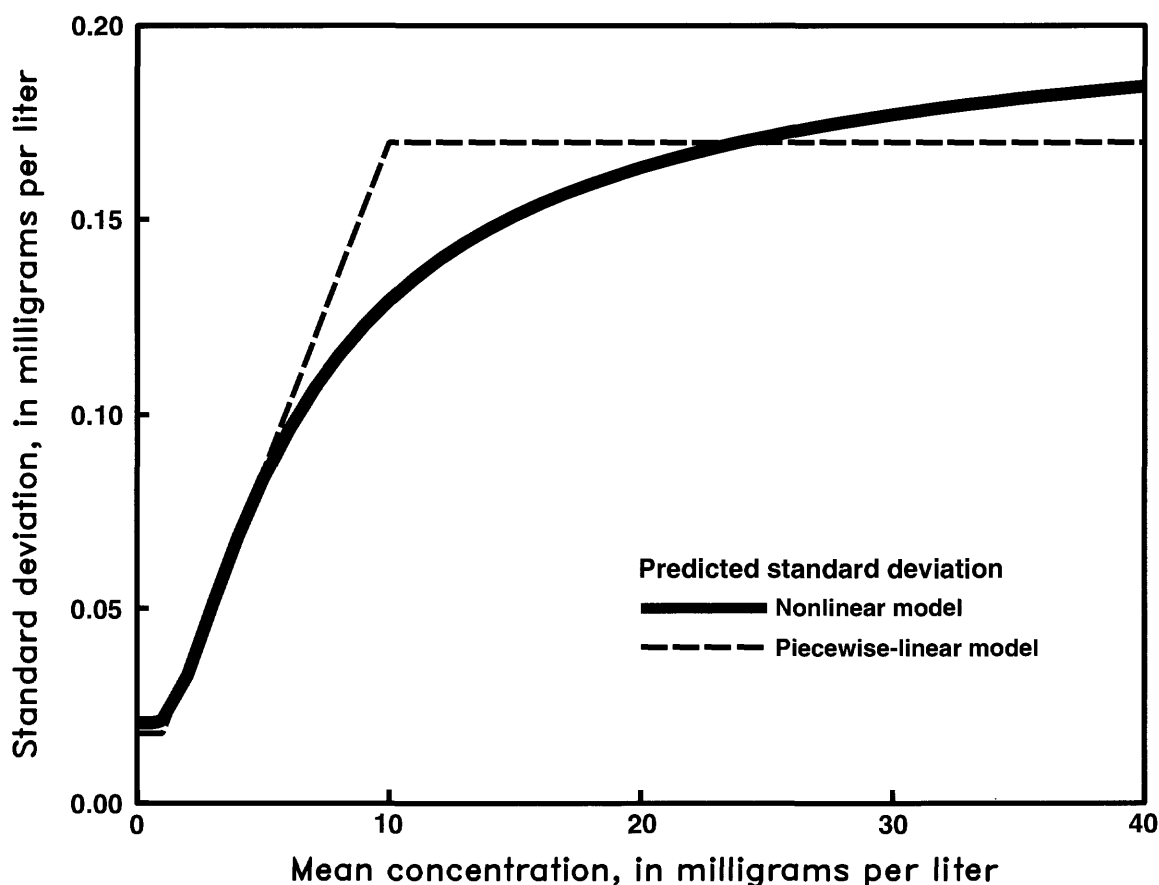


Figure 2. Example of a nonlinear model of replicate variability compared with the alternative, piecewise-linear model. Data used to fit the models were for nitrite plus nitrate as nitrogen in ground-water samples.

Once an appropriate model has been selected and fit to field-replicate data, sampling variability can be estimated for specific analyte concentrations and then used to construct confidence intervals. Based on a concentration measured in a single sample, the confidence interval for the true concentration is:

$$[C_L, C_U] = C \pm Z_{(1-\alpha/2)} \sigma \quad (5)$$

where

C_L, C_U = the lower and upper limits of concentration for the 100(1- α) percent confidence interval;

α = the probability that the confidence interval does not include the true concentration;

C = the concentration measured in the sample;

$\pm Z$ = the ordinate of the normal curve (Z-value) that contains 100(1- α) percent of the distribution; and

σ = sampling variability, estimated as standard deviation of field replicates using the nonlinear or the piecewise-linear model.

The second term ($\pm Z_{(1-\alpha/2)} \sigma$) in equation 5 represents the error inherent in a single measurement of concentration due to sampling variability. If a single measurement differs from a standard by less than this error, it is not possible (with 100(1- α) percent confidence) to determine whether the concentration in the sample exceeded the standard.

For a mean concentration (\bar{C}) from multiple samples, the confidence interval for the true mean is calculated:

$$[C_L, C_U] = \bar{C} \pm Z_{(1-\alpha/2)} \frac{\sigma}{\sqrt{n}} \quad (6)$$

where

n = the number of samples, and the other variables are as previously defined.

Again, the error due to sampling variability is represented by the second term of equation 6, but in this case, it includes the number of samples as well as the standard deviation. Thus, the error inherent in a mean concentration due to sampling variability can be decreased by collecting more samples. This error can be considered the minimum

that is typically achievable for determining a mean concentration in the absence of environmental variability. Statistical significance is unlikely for a difference between two mean concentrations that is less than the sum of their inherent errors; therefore, small but true environmental differences might not be detected.

QUALITY OF THE NUTRIENT DATA

Filtered stream and ground-water samples collected for the NAWQA Program were analyzed for the following nutrient analytes:

- Ammonia as nitrogen (includes dissolved ammonium ion and un-ionized ammonia, hereinafter referred to as ammonia)
- Nitrite as nitrogen
- Nitrite plus nitrate as nitrogen (hereinafter referred to as nitrite plus nitrate)
- Dissolved Kjeldahl nitrogen (the Kjeldahl analysis includes ammonia plus organic nitrogen)
- Orthophosphate as phosphorus (hereinafter referred to as orthophosphate)
- Dissolved phosphorus

In addition, the following nutrient analytes were determined for unfiltered stream samples:

- Total Kjeldahl nitrogen
- Total phosphorus

Nitrite concentrations are generally less than the reporting limit in streams and in oxygenated ground water; therefore, nitrite was not included in the QC analysis.

The QC samples collected to evaluate the quality of the NAWQA nutrient data included blanks prepared using stream and ground-water sampling equipment and various types of replicates collected at stream and ground-water sampling sites. The stream-sample replicates were a combination of split, concurrent, and sequential replicates. Information about replicate type was included in the data base for many, but not all, replicates. Ground-water replicates were collected sequentially as water was pumped from the well.

Contamination Bias

Potential contamination bias was determined separately for stream and ground-water samples on the basis of blanks prepared at sampling sites. Contamination was evaluated for seven nutrient analytes in blanks prepared at stream sites and five nutrient analytes in blanks prepared at ground-water sites.

The first step in evaluating the QC data for each analyte was to plot the relation between analyte concentrations and dates of blank preparation to identify possible temporal trends in contamination. No trends were obvious in these relations for any analyte. Thus, potential contamination was considered constant throughout the period of record.

Distributions of nutrient concentrations in stream and ground-water field blanks are shown in figure 3. For all analytes, the 95th percentile of measured values was at or near the detection

limit. Only six measurements (all nitrite plus nitrate) exceeded three times the 95th percentile value. Thus, the 95th percentile represents a reasonable upper limit for general evaluation of contamination.

Data from the blanks were used to calculate upper tolerance bounds on contamination for all samples collected during NAWQA high-intensity sampling (1993–95). A 90-percent minimum level of confidence was selected for the tolerance bounds, and separate calculations were made for bounds to exceed contamination in 75, 90, and 95 percent of all samples (table 2). Potential contamination in at least 75 percent of all samples is estimated to be no greater than the reporting limit for all nutrient analytes. For dissolved and total Kjeldahl nitrogen, potential contamination is estimated to be less than the reporting level in at least 95 percent of stream and ground-water samples.

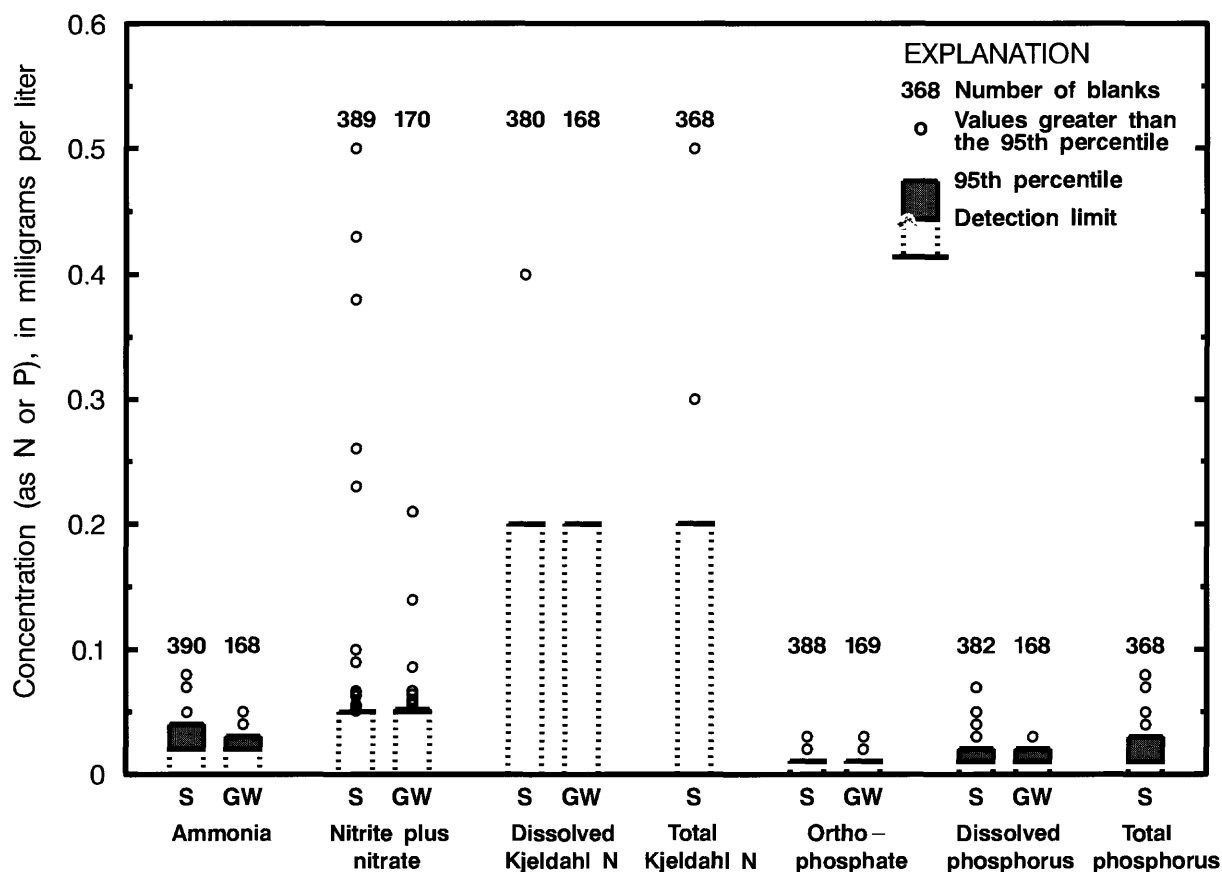


Figure 3. Distribution of nutrient concentrations measured in field blanks from National Water-Quality Assessment stream (S) and ground-water (GW) sites.

Table 2. Upper 90-percent tolerance bounds that exceed contamination by nutrient analytes in specified percentages of all samples based on data from field blanks prepared at stream and ground-water sampling sites

[mg/L, milligrams per liter; <, less than]

Nutrient analyte	Number of blanks	Reporting limit (mg/L)	Upper 90-percent tolerance bound (mg/L)		
			75 percent of samples	90 percent of samples	95 percent of samples
Streams					
Ammonia	392	0.02	0.02	0.03	0.04
Nitrite plus nitrate	389	0.05	<0.05	<0.05	0.063
Dissolved Kjeldahl nitrogen	380	0.2	<0.2	<0.2	<0.2
Total Kjeldahl nitrogen	368	0.2	<0.2	<0.2	<0.2
Orthophosphate	388	0.01	<0.01	<0.01	0.01
Dissolved phosphorus	382	0.01	<0.01	0.02	0.02
Total phosphorus	368	0.01	0.01	0.02	0.03
Ground water					
Ammonia	172	0.02	0.02	0.03	0.03
Nitrite plus nitrate	170	0.05	<0.05	<0.05	0.067
Dissolved Kjeldahl nitrogen	168	0.2	<0.2	<0.2	<0.2
Orthophosphate	169	0.01	<0.01	0.01	0.02
Dissolved phosphorus	168	0.01	<0.01	0.02	0.02

The rightmost column in table 2 lists the contamination that is likely to be exceeded in less than 5 percent of all samples. These concentrations can be compared to critical values to identify potential problems with interpretation of nutrient data. For example, the drinking-water standard for nitrate is 10 mg/L (U.S. Environmental Protection Agency, 1986), which is more than 100 times greater than the potential contamination by nitrite plus nitrate in at least 95 percent of all samples. Thus, contamination is unlikely to cause problems with identifying exceedances of this standard. For ammonia, the results are not as encouraging. The aquatic-life criterion for protection of salmonids (such as trout) can be as low as 0.07 mg/L, depending on water temperature and pH (U.S. Environmental Protection Agency, 1986). At this level, more than one-half the reported ammonia in a stream sample might result from contamination. A similar concern might be raised about total phosphorus. For prevention of nuisance plant growth, the U.S. Environmental Protection Agency (USEPA) recommends a maximum concentration of 0.1 mg/L total phosphorus in streams. Contamination in at least 5 percent of stream samples might account for more than 30 percent of this value. Because of these

contamination problems, data for ammonia and total phosphorus at low concentrations may be of limited use for comparison to standards. At concentrations more than 10 times the potential contamination listed in table 2, these problems become practically insignificant.

The source of ammonia and phosphorus contamination was investigated by comparing concentrations measured in field blanks to concentrations measured in the source solutions used to prepare the blanks. The field blanks were exposed to sampling equipment and environmental conditions during preparation, but the **source-solution blanks** were exposed only to shipping conditions and laboratory analyses. Selected percentiles of the distribution of ammonia concentrations in field and source-solution blanks are plotted in figure 4. In general, contamination of the source-solution blanks is equivalent to that of the field blanks. This result indicates that either the source solutions were contaminated prior to preparation of the field blanks or that contamination occurred during shipping or in the laboratory. If the source solutions were contaminated, then environmental samples would not be affected, and the calculated potential contamination by ammonia is too large. However, if

contamination occurred during shipping or in the laboratory, then environmental samples would be affected the same as the blanks, and the calculated potential contamination remains valid. In either case, field procedures are not implicated and no revisions to environmental sampling protocols are necessary. Similar results were obtained for comparison of total phosphorus in the three types of blanks.

Sampling Variability

Sampling **variability** was estimated using the standard deviations within sets of replicate samples collected at stream and ground-water sites. Separate estimates were made for seven nutrient analytes in stream samples and five nutrient analytes in ground-water samples.

A preliminary evaluation was made to determine whether the standard deviation of stream-sample replicates was related to the type of replicate preparation (split, concurrent, or sequential). Based on available data, type could only be determined consistently for split replicates. Thus, the data were divided into “split” and “other” replicate types. The data also were divided into ranges of concentration based on the order of magnitude of the mean concentration of the replicate set. This division was made for two reasons. First, variability is typically related to concentration, and differences in standard deviation over the full range of concentration might mask differences among replicate types. Also, data rounding from the laboratory causes the precision of reported concentrations to vary with order of magnitude. An effect of this variation is that the potential differences in concentration among replicates, and thus, their standard deviations, change with order of magnitude. Analysis of variance

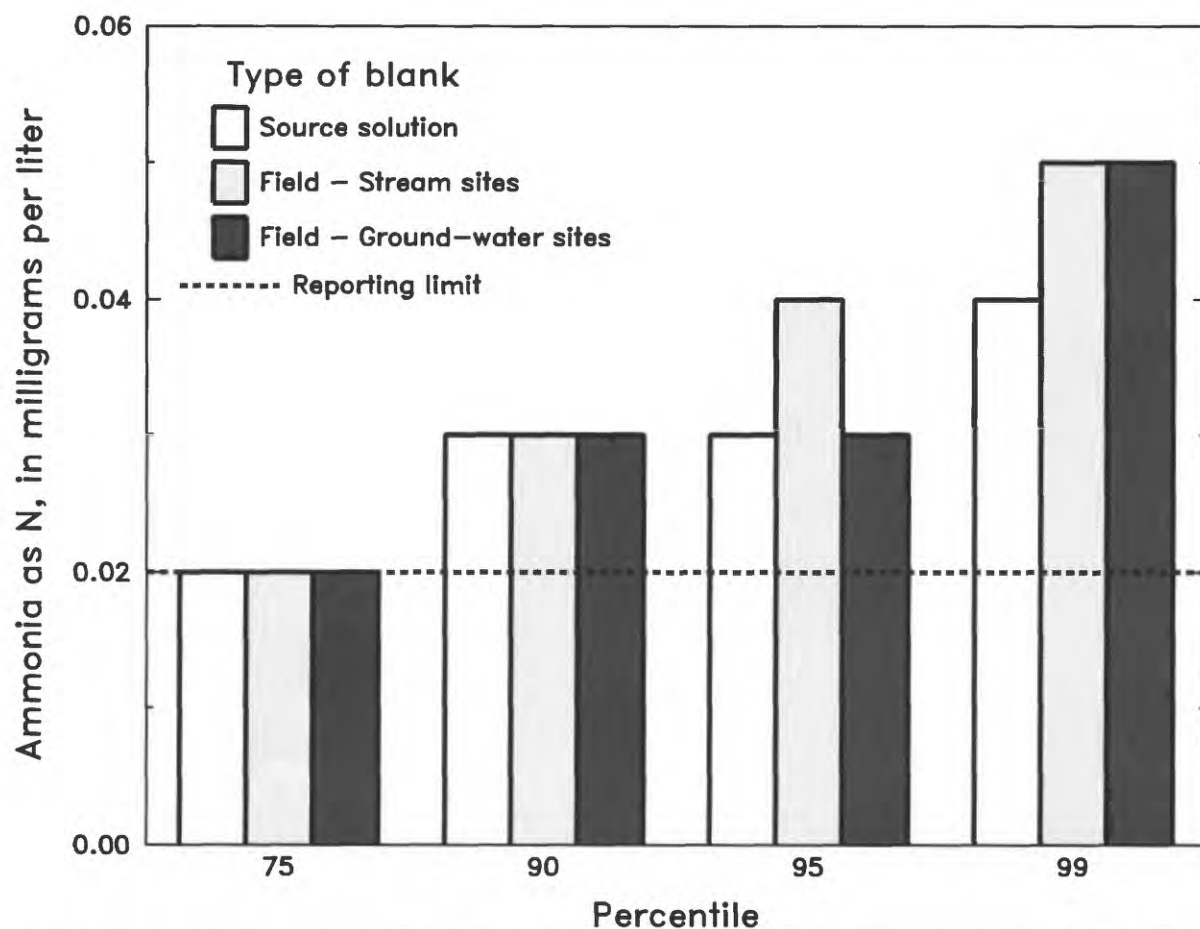


Figure 4. Selected percentiles of the distributions of ammonia contamination in source-solution and field blanks.

(ANOVA) was used to determine whether standard deviations within order-of-magnitude ranges for mean replicate concentration were significantly different between “split” and “other” replicate types. The ANOVA was performed on ranks of the data, analogous to the nonparametric Kruskal-Wallis test (Iman and Conover, 1983). Differences in standard deviation of replicates were identified between ranges of concentration, but no differences occurred between replicate types within a single concentration range. Therefore, all types of replicates were combined in subsequent analysis of sampling variability.

The nonlinear model was fit to the replicate data for each nutrient analyte. Results were scrutinized to determine whether the model was appropriate for individual analytes. Because of the typically large range in standard deviation throughout the range of mean concentration for most analytes, simply plotting the model curve on a scatter diagram of the replicate data did not provide a good visual indication of model fit. In addition, multiple occurrences of the same combinations of standard deviation and mean concentration do not appear on the scatter diagram; therefore, the influence of single occurrences might be overemphasized. These problems were avoided by including spline and LOWESS smooth curves through the center of mass of the replicate data on the scatter diagrams. Model results could then be compared to these curves.

Results of nonlinear models fit to nutrient analytes in stream-sample replicates are shown in comparison to spline and LOWESS smooth curves in figure 5, for nitrogen analytes, and figure 6, for phosphorus analytes. (Note that comparisons among analytes must consider differences in scale among the graphs.) In general, model estimates of standard deviation for each analyte are similar to both data smooths at low concentrations and similar to the spline smooth but greater than the LOWESS smooth in the middle of the concentration range. The major differences occur at high concentrations. In this range, standard deviation is not always well defined, as shown by divergence of the two data smooths for total Kjeldahl nitrogen (fig. 5) and total phosphorus (fig. 6). This problem arises because the reported values at high concentrations are less precise than at low concentrations due to rounding. Differences between high replicate concentrations were generally restricted to a few values and were primarily zero or one. In addition, fewer replicates were collected at high concentration, so a small number of replicates that happen to be

rounded to the same difference can have a large influence on the smooth curves. In cases where the smooths diverge from the model estimates at high concentrations, the model is generally conservative, overestimating standard deviation. Model estimates are less than 90 percent of either smooth only for very high concentrations of nitrite plus nitrate, dissolved phosphorus, and total phosphorus, and in each case only for one of the smooth curves. For all other conditions, the model estimates of standard deviation seem reasonable or perhaps too high. Thus, standard deviations estimated by the models can be used to provide a generally conservative estimate of sampling variability. However, as shown by the individually plotted values, standard deviations for some replicate sets were much greater than model estimates.

Another indication of model adequacy is shown by plotting the residuals, the difference between the standard deviation of each replicate set and the estimated standard deviation (figs. 7 and 8). Residuals less than zero indicate the model has overestimated standard deviation; residuals greater than zero indicate an underestimation. Many plotted points represent more than one replicate set, particularly for model overestimations at low concentrations. There, the few high residuals are offset, more than is obvious, by many low residuals. The apparent parallel curves in these plots are a result of the discrete values of standard deviation possible for duplicate samples with rounded data. The lowest “curve” on each graph is for measured replicate standard deviations of zero, and thus, represents the lower limit of residual values. There is no upper limit for residuals because the difference between replicate concentrations has no bound. Therefore, the magnitude of positive residuals is typically greater than the magnitude of negative residuals, as is shown in figures 7 and 8. However, the distribution of residuals is centered at about zero, and the range generally appears unrelated to concentration. For three analytes (nitrite plus nitrate, total Kjeldahl nitrogen, and orthophosphate), residuals at high concentrations are predominantly negative. For nitrite plus nitrate and orthophosphate, this situation might result primarily from the order of magnitude of the reported replicate concentrations. At nitrite plus nitrate concentrations greater than 10 and orthophosphate concentrations greater than 1, sampling variability might be less than the precision due to data rounding, and the difference between replicate concentrations

would likely round to zero. Thus, the low bias in the distribution of residuals is an artifact of data reporting and does not necessarily indicate a problem with the model fit.

Results of nonlinear models fit to three of five nutrient analytes in ground-water replicates are shown in comparison to spline and LOWESS smooth curves in figure 9. Models could be adequately fit to data for ammonia, nitrite plus nitrate, and orthophosphate, but not for dissolved Kjeldahl nitrogen or dissolved phosphorus. The curves for the three models compare well with the spline curves through the data, but are similar to the LOWESS curves only in the low concentration ranges. The two data smooths diverge at high concentrations of nitrite plus nitrate and orthophosphate, possibly for the same reasons discussed for stream-sample replicates: a decrease in the precision of reported concentrations and a paucity of samples. Residuals for the three models are plotted in figure 10. The patterns are similar to those for stream-sample replicates except for a few large positive residuals for nitrite plus nitrate.

Because data for replicates of dissolved Kjeldahl nitrogen and dissolved phosphorus could not be adequately fit with a nonlinear model, the alternative piecewise-linear model was used for these analytes. Only 35 sets of replicate samples were analyzed for dissolved Kjeldahl nitrogen in ground water, and all concentrations were low; therefore, only the low range of the piecewise-model could be applied. Thus, the best estimate of sampling variability for dissolved Kjeldahl nitrogen is simply the mean of all replicate standard deviations. For dissolved phosphorus, the low and middle ranges of concentrations had adequate data. In the low range, sampling variability is estimated by the mean of the replicate standard deviations, and in the middle range, it is estimated by the mean of the replicate RSDs. These estimates are plotted with spline and LOWESS curves through the data in figure 11. The estimates are similar to both smooth curves for concentration less than about 0.5 mg/L, and they are similar to, though slightly higher than, the spline curve over the remaining range of the replicate data. The two smooths diverge like those for some analytes in figures 5, 6, and 9, probably for the same reasons.

Several interesting comparisons can be made among the models shown in figures 5, 6, 9, and 11. Among the nitrogen species, the smallest sampling

variabilities, based on model estimates of standard deviation, were for nitrite plus nitrate in streams and ammonia in ground water. However, the range of concentrations for ammonia in ground water was very small, covering primarily the lower part of the range in streams. Within this range, the sampling variabilities of ammonia in streams and ground water were about the same. Also, the sampling variability at the highest observed concentrations for ammonia in streams was only slightly greater than sampling variability at high concentrations of nitrite plus nitrate. Much higher sampling variabilities were estimated within the high range of observed concentrations for dissolved and total Kjeldahl nitrogen. The Kjeldahl method might introduce more variability to analytic results than the methods used to analyze ammonia and nitrite plus nitrate. Among phosphorus species, sampling variabilities are smallest for orthophosphate and largest for total phosphorus, and larger for ground water than for streams. With the exception of the high concentration range, the piecewise-linear model results for dissolved phosphorus in ground water are similar to the nonlinear model results for dissolved phosphorus in streams.

Parameter values for the nonlinear models fit to nutrient analytes in stream and ground-water replicates are listed in table 3. The parameter values for an analyte can be used with equation 3 to estimate the sampling variability for a selected analyte concentration. A confidence interval around the mean concentration can then be calculated as the product of this sampling variability and an appropriate Z-value. For example, from table 3 and equation 3, the sampling variability of ammonia in streams is estimated by:

$$\sigma = 0.003889 + e^{\left(-1.302 - \frac{3.151}{C}\right)} \quad (7)$$

For a measured ammonia concentration of 0.5 mg/L, the estimated sampling variability is about 0.0044. The 95-percent confidence interval for the true concentration, based on this measurement, can then be determined using equation 5 with a Z-value of 1.96 (from a table of standard normal deviates):

$$[C_L, C_U] = 0.5 \pm 1.96(0.0044) \quad (8)$$

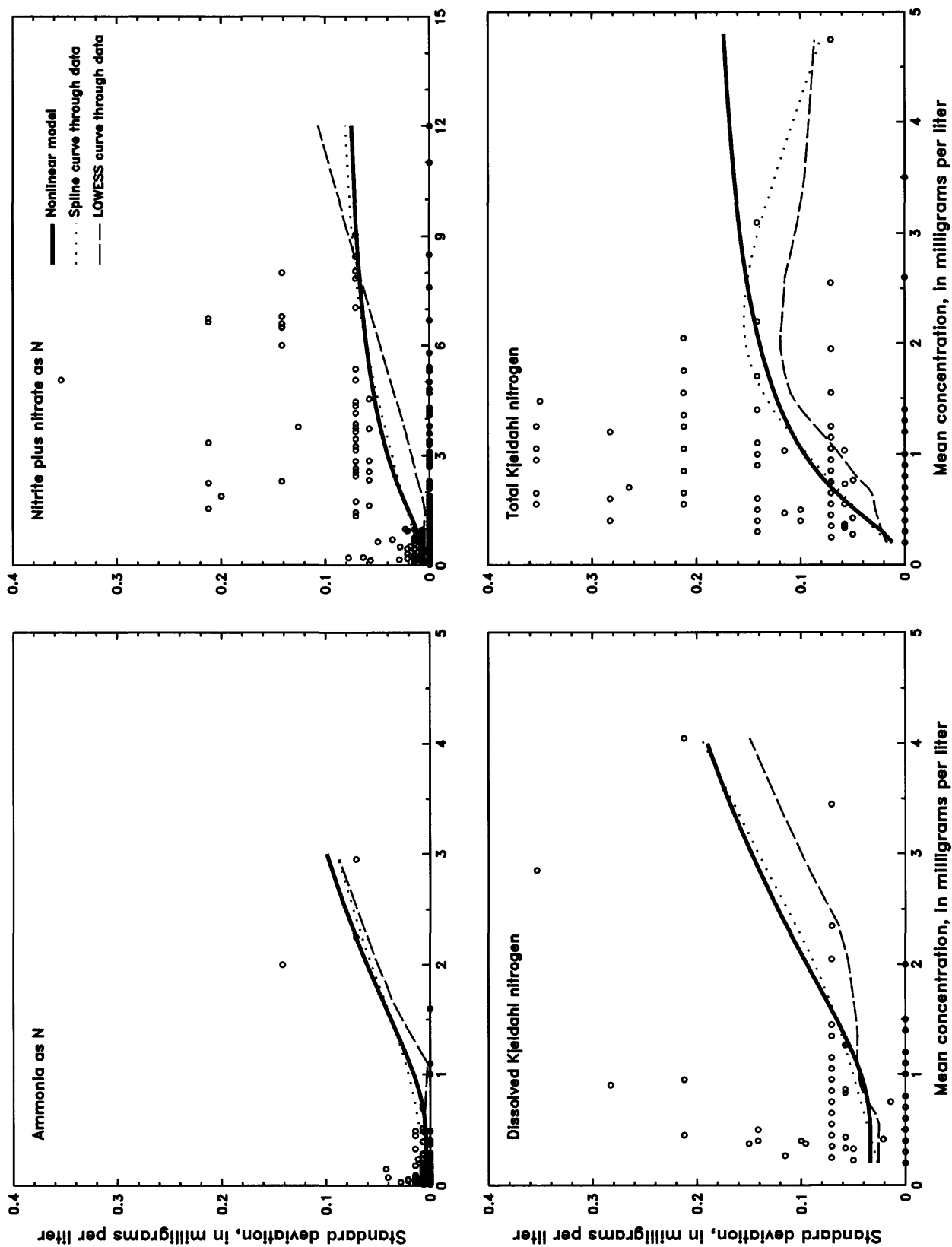


Figure 5. Nonlinear-model estimates of variability for nitrogen analytes in stream samples in comparison to smoothed curves through the center of the replicate data.

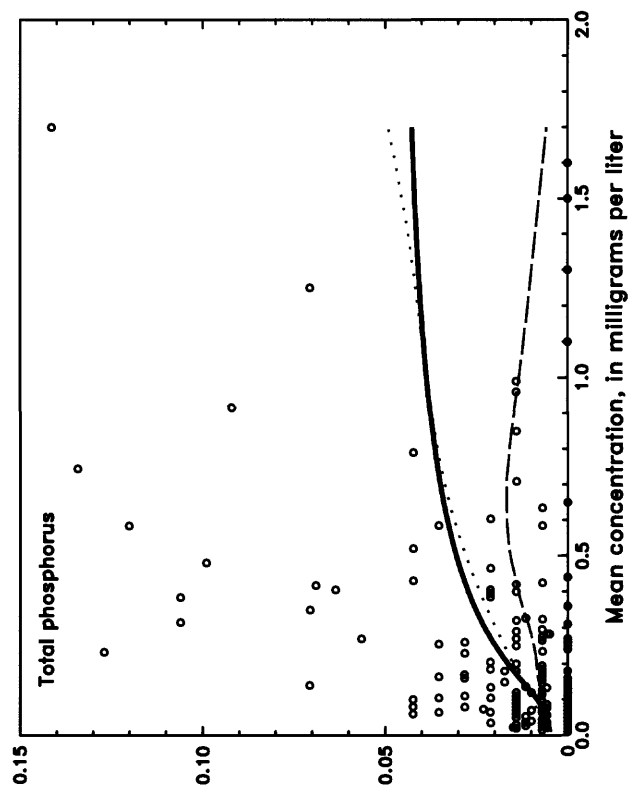
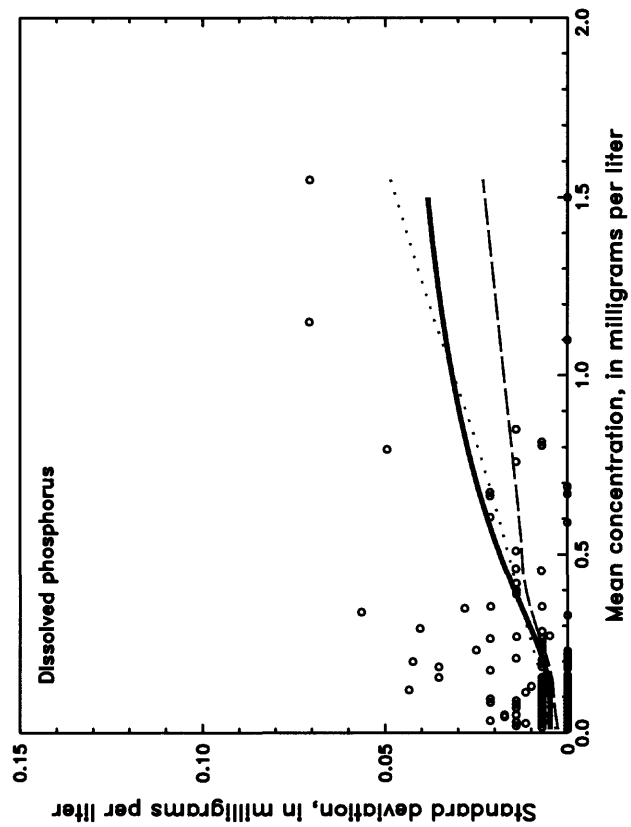
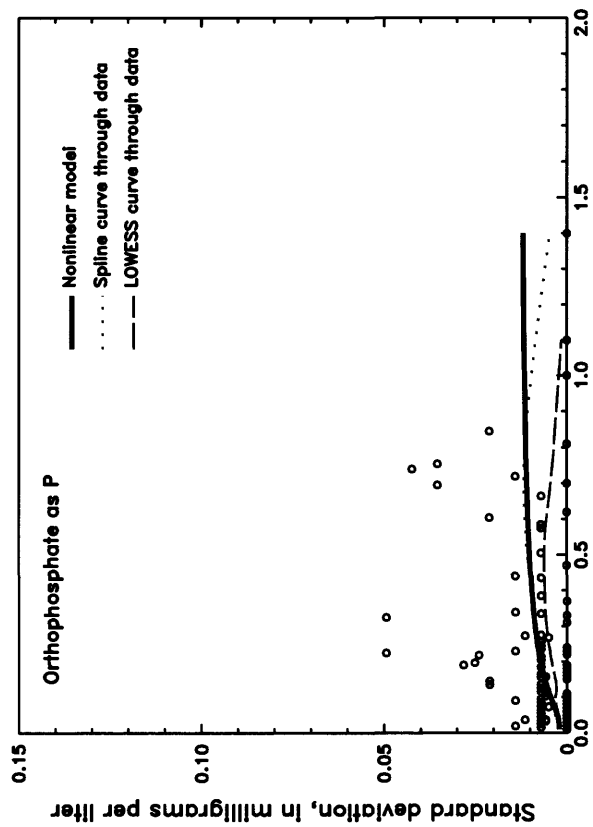


Figure 6. Nonlinear-model estimates of variability for phosphorus analytes in stream samples in comparison to smoothed curves through the center of the replicate data.

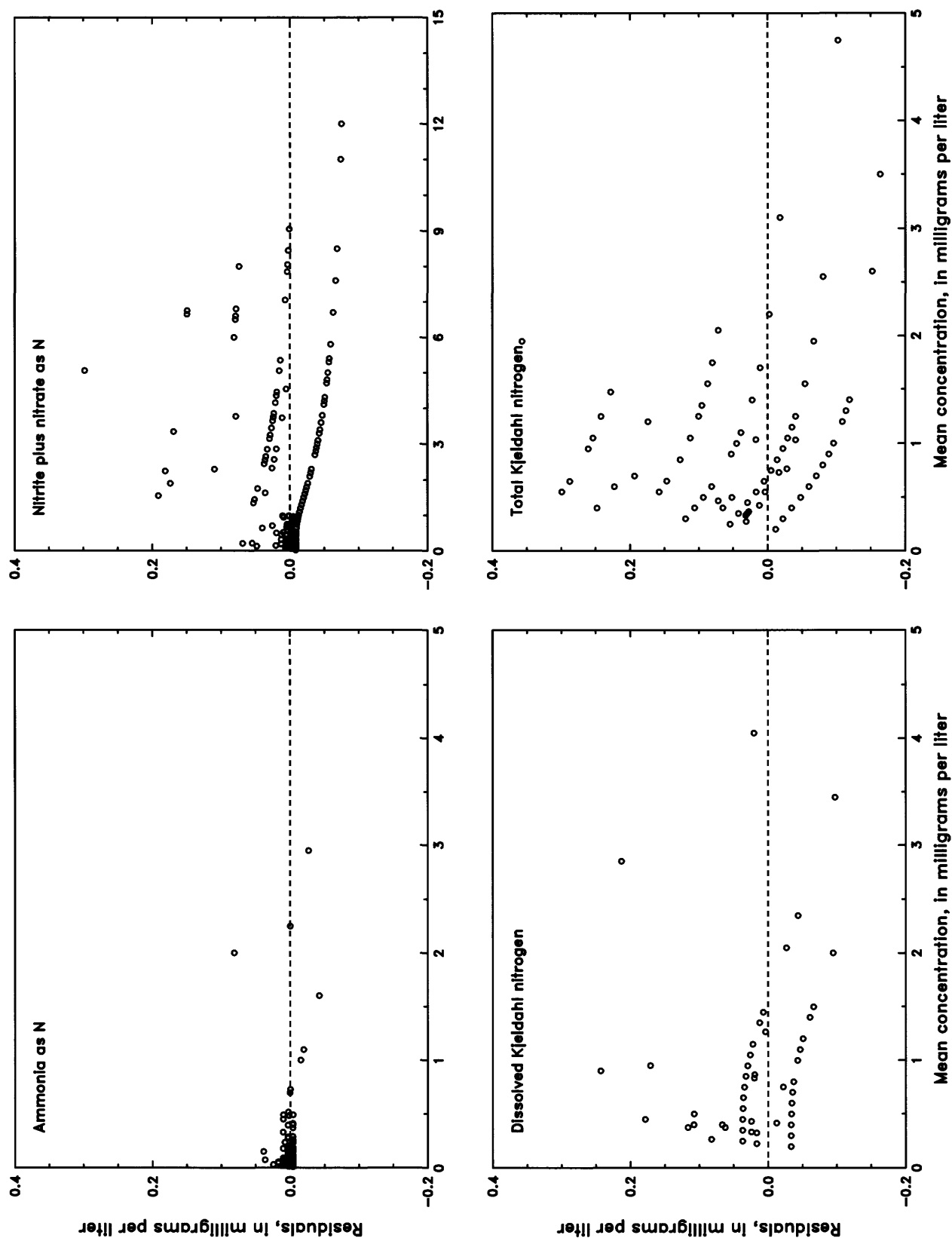


Figure 7. Residuals of standard deviation estimated by nonlinear models of variability for nitrogen analytes in stream samples.

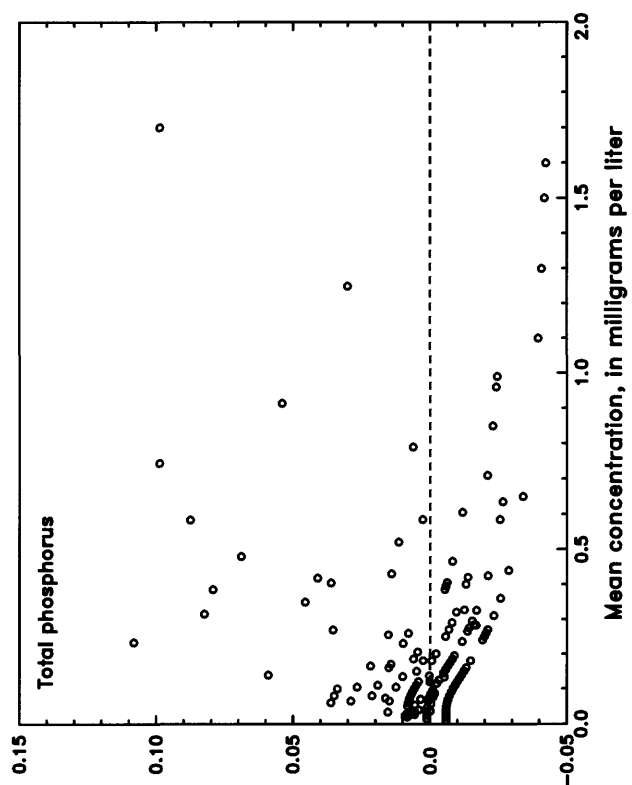
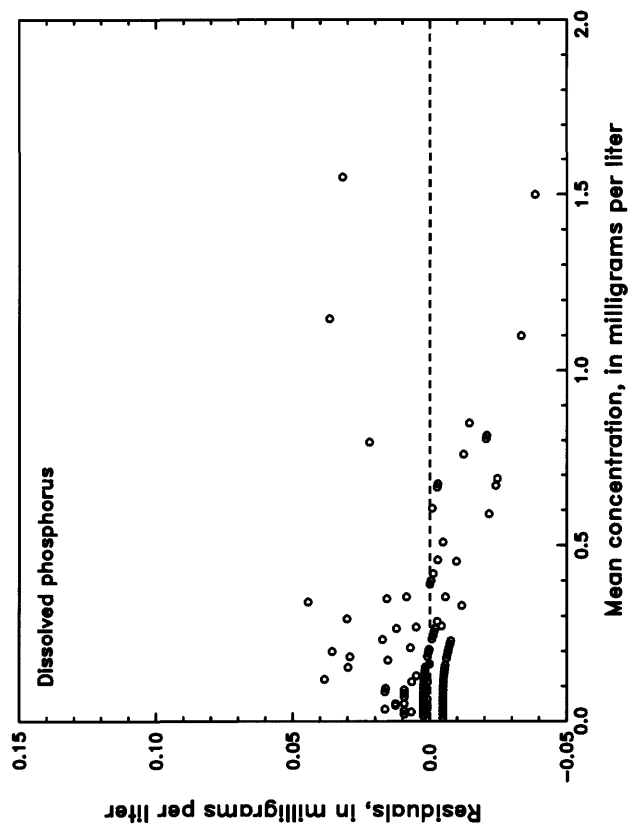
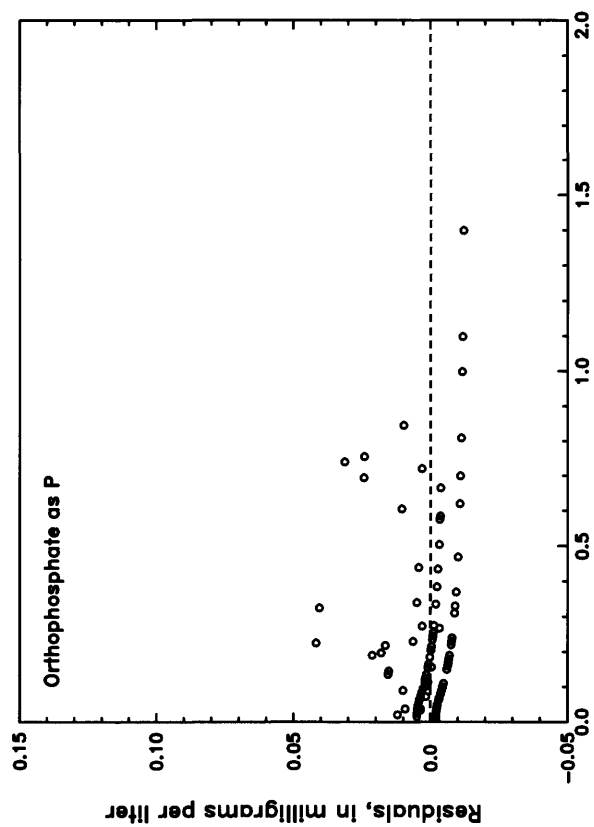


Figure 8. Residuals of standard deviation estimated by nonlinear models of variability for phosphorus analytes in stream samples.

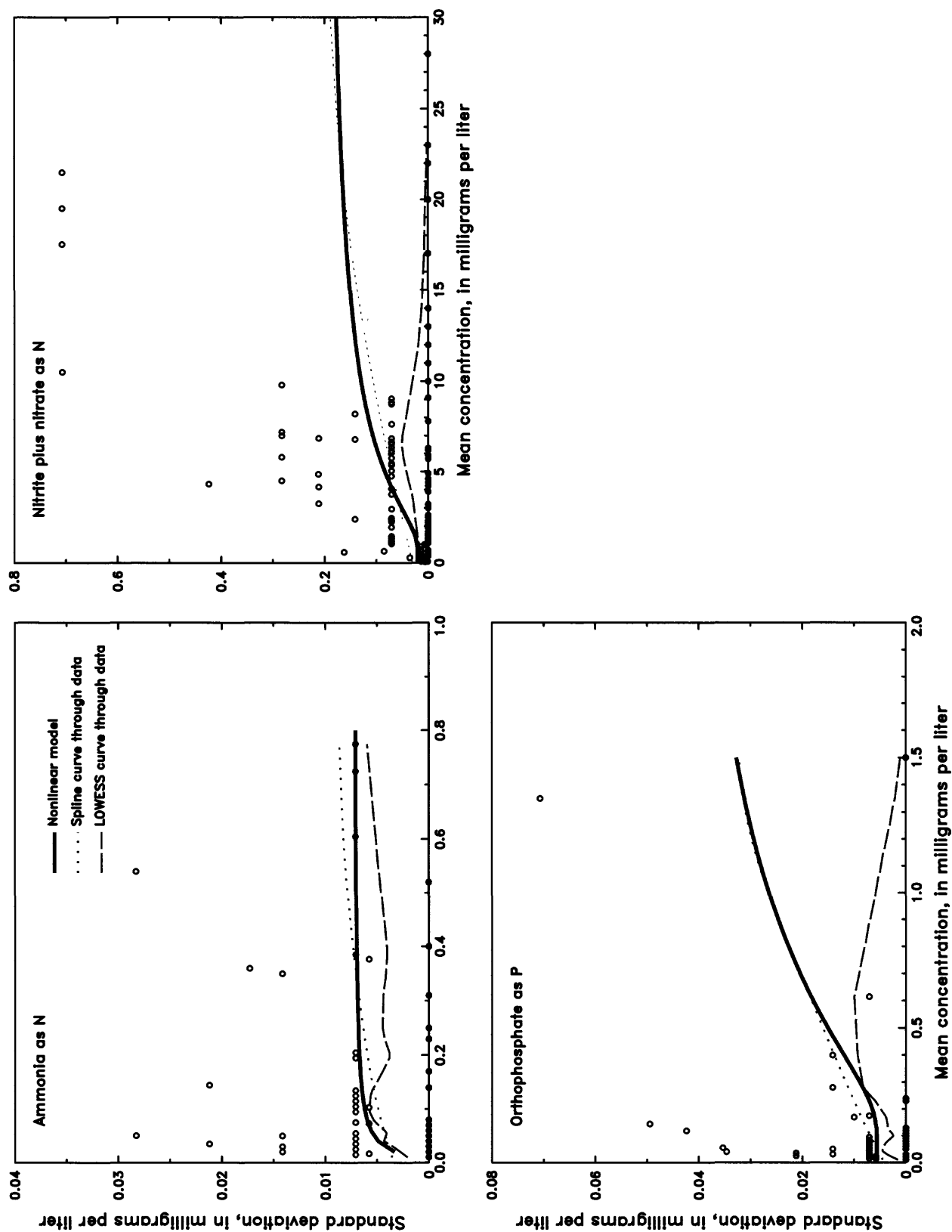


Figure 9. Nonlinear-model estimates of variability for selected nutrient analytes in ground-water samples in comparison to smoothed curves through the center of the replicate data. [Note: Axis scales vary.]

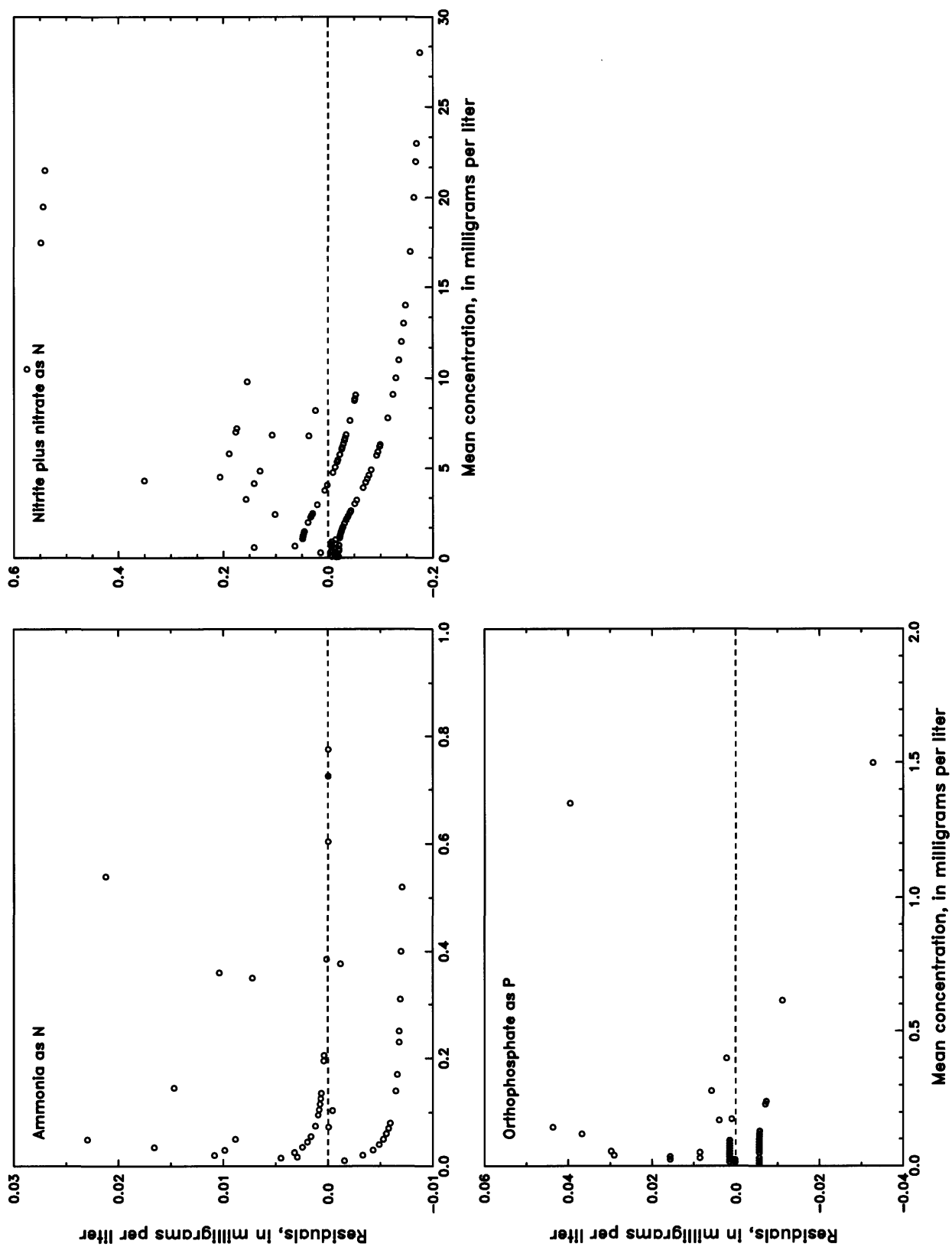


Figure 10. Residuals of standard deviation estimated by nonlinear models of variability for selected nutrient analytes in ground-water samples. [Note: Axis scales vary.]

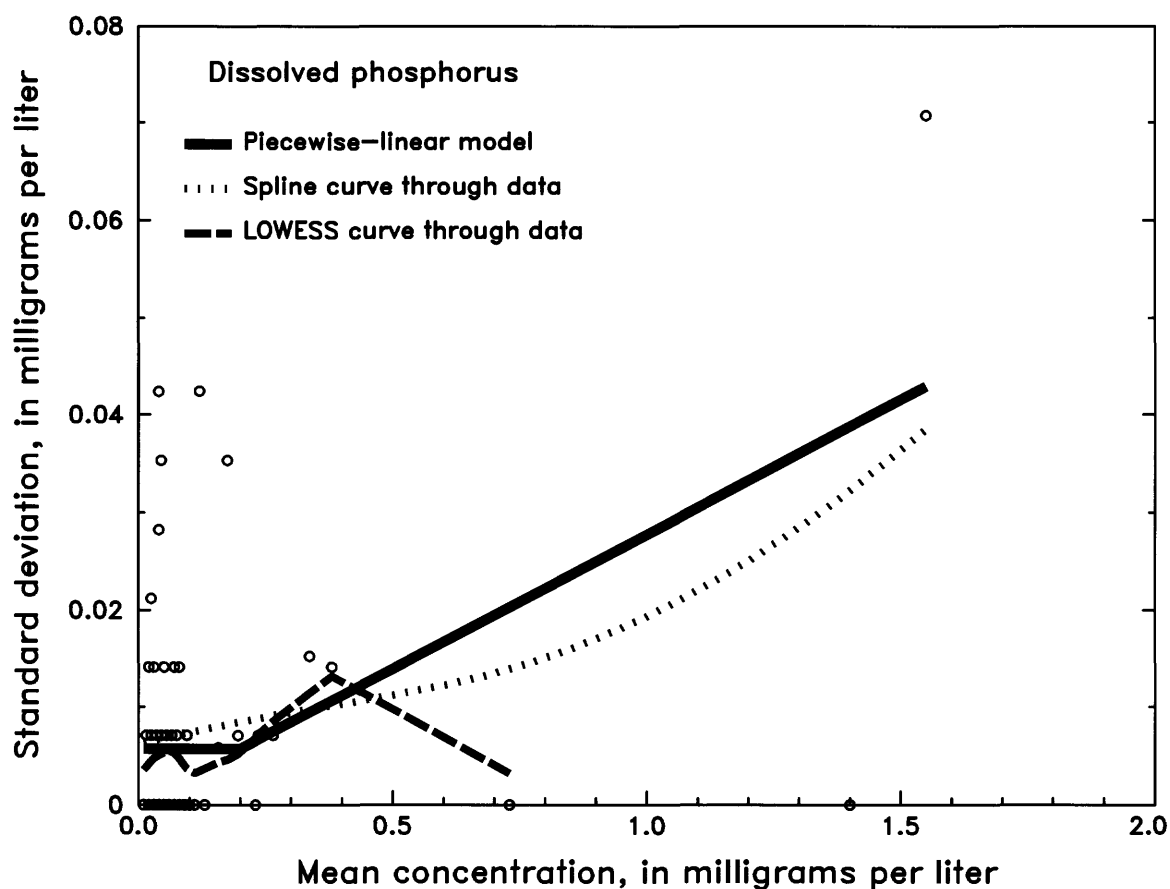


Figure 11. Piecewise-linear model estimate of variability for dissolved phosphorus in ground-water samples in comparison to smoothed curves through the center of the replicate data.

The second term of this equation, which is calculated to be about ± 0.009 mg/L, or ± 1.8 percent of the measured concentration, represents the inherent error of the measurement. If the concentration was a mean of 10 measurements, the 95-percent confidence interval for the true mean is determined using equation 6:

$$[C_L, C_U] = 0.5 \pm 1.96 \left(\frac{0.0044}{\sqrt{10}} \right) \quad (9)$$

In this case, the inherent error is estimated to be ± 0.003 mg/L, or ± 0.6 percent of the mean concentration. Other factors, such as environmental variability, would likely increase the actual standard deviation calculated from this hypothetical set of 10 samples. Sampling variability is the unavoidable error in measuring the 10 concentrations, even if the true environmental concentration is constant. Thus, sampling variability for a mean represents the likely lower limit of overall variability.

Table 3. Parameter estimates for nonlinear models of sampling variability for selected nutrient analytes in streams and ground water

Nutrient analyte	Number of replicate sets	Estimated parameter values for the nonlinear model (equation 3)		
		B ₁	B ₂	B ₃
Streams				
Ammonia	284	0.003889	1.302	3.151
Nitrite plus nitrate	264	0.008629	2.462	3.020
Dissolved Kjeldahl nitrogen	212	0.03346	0.9141	3.769
Total Kjeldahl nitrogen	233	0.007993	1.633	0.7950
Orthophosphate	212	0.002076	4.503	0.1549
Dissolved phosphorus	221	0.004816	2.949	0.6719
Total phosphorus	250	0.005855	3.133	0.2846
Ground water				
Ammonia	135	0.0001168	4.944	0.01611
Nitrite plus nitrate	126	0.02071	1.672	5.484
Orthophosphate	87	0.005589	3.077	0.7940

Sampling variability for the two nutrient analytes that had to be estimated using the piecewise-linear model are listed in table 4. The sampling variability over the entire range of concentrations measured for dissolved Kjeldahl nitrogen is estimated to be constant. Data were insufficient to estimate sampling variabilities for higher ranges. Data for dissolved phosphorus were sufficient over the low and middle ranges of concentration, but not for the high range. The effect of sampling variability on measured concentrations of these analytes is determined in the same way as was done for nonlinear model estimates. For example, sampling variability for dissolved phosphorus measurements between 0.2 and 1.5 mg/L is 2.31 percent of the concentration. If the measured concentration is 1.0 mg/L, the estimated sampling variability is 0.0231 mg/L. The error inherent in that measurement, determined from a 95-percent confidence interval, is $\pm 1.96 (0.0231)$, which equals about ± 0.045 mg/L or ± 4.5 percent of the measured concentration.

IMPLICATIONS FOR INTERPRETATION OF ENVIRONMENTAL DATA

Proper interpretation of the NAWQA data requires consideration of the effects that contamination bias and sampling variability might have on nutrient concentrations measured in stream and ground-water samples.

Potential Effects of Contamination

In general, if potential contamination is less than 10 percent of a measured value, the effect of contamination bias on that measured value can be

ignored. For concentration data rounded to two significant digits, a positive bias of less than 10 percent does not usually affect the first digit, and if it does, its maximum effect is only a single unit. In most cases, such an effect has no practical significance. Thus, the largest measured concentration that might be affected can be estimated as 10 times the potential contamination. Using the upper tolerance bounds on contamination determined from NAWQA field blanks (table 2), maximum affected concentrations were calculated from selected nutrient analytes and are listed in table 5 along with various critical concentrations for those analytes. For ammonia, measured concentrations less than or equal to 0.4 mg/L are potentially affected by contamination. This concentration exceeds the background level for ammonia in streams and the median concentration of ammonia downstream from urban areas (Mueller and others, 1995). Thus, contamination could affect use of some ammonia data in identifying streams that might be adversely impacted by urban development. Also, the aquatic-life criterion for ammonia is less than 0.4 mg/L for some combinations of water temperature and pH (U.S. Environmental Protection Agency, 1986); however, in the range of pH and temperature for most streams, the criterion is generally greater than 0.4 mg/L. Thus, although contamination could limit the use of ammonia data in comparison to criteria under some conditions, the majority of measurements that exceed the criterion would not be affected.

Contamination also might affect interpretation of phosphorus data. The maximum affected concentrations listed in table 5 exceed USEPA recommendations for protection of surface water (U.S. Environmental Protection Agency, 1986) and national average background concentrations in streams

Table 4. Estimates of standard deviation from piecewise-linear models of sampling variability for selected nutrient analytes in ground water¹

[mg/L, milligrams per liter; --, insufficient data]

Nutrient analyte	Low concentration range			Middle concentration range		
	Concentrations (mg/L)	Number of replicate sets	Standard deviation (mg/L)	Concentrations (mg/L)	Number of replicate sets	Relative standard deviation ² (percent)
Dissolved Kjeldahl nitrogen	0.2 to 1.05	35	0.019	--	--	--
Dissolved phosphorus	0.01 to 0.2	76	0.007	0.2 to 1.5	8	2.31

¹Neither analyte had sufficient data to estimate standard deviation for the high concentration range.

²See equation 4.

Table 5. Maximum nutrient concentrations that are considered potentially affected by contamination, based on the 1991 National Water-Quality Assessment data, and selected critical values used to interpret environmental data

[mg/L, milligrams per liter]

Nutrient analyte	Maximum affected concentration ¹ (mg/L)	Critical value	
		Description	Concentration (mg/L)
		Streams	
Ammonia	0.4	Background ²	0.1
		Aquatic-life criterion ³	0.07–2.1
		Median downstream from urban areas ²	0.2
Nitrite plus nitrate	0.63	Background ²	0.6
		Drinking-water standard ⁴	10
Orthophosphate	0.1	Recommended to avoid eutrophication ⁴	0.05
Total phosphorus	0.3	Background ²	0.1
		Recommended to avoid eutrophication ⁴	0.1
		Ground water	
Nitrite plus nitrate	0.67	Background ⁵	3
		Drinking-water standard ⁴	10

¹Estimated at 10 times the upper tolerance bound to include 95 percent of samples, listed in table 2.

²Mueller and others, 1995.

³Criterion varies depending on water temperature and pH (U.S. Environmental Protection Agency, 1986).

⁴U.S. Environmental Protection Agency, 1986.

⁵Madison and Brunett, 1985.

(Mueller and others, 1995). Unqualified determination of adverse environmental impacts due to phosphorus is limited to sites where measured concentrations are greater than 0.1 mg/L for orthophosphate and 0.3 mg/L for total phosphorus.

Nitrite plus nitrate and Kjeldahl nitrogen data are essentially unaffected by contamination. The maximum concentration of nitrite plus nitrate that would be of concern is 0.63 mg/L (table 5). Except for background levels in streams, critical concentrations for interpretation of nitrite plus nitrate data are much larger. For Kjeldahl nitrogen, the potential contamination in more than 95 percent of all samples was less than the reporting limit (table 2); therefore, any measured concentration in excess of 10 times the reporting limit is not likely to be affected.

Potential Effects of Sampling Variability

Sampling variability has two primary effects on the interpretation of water-quality data:

1. It determines the minimum potential error in an individual measurement, and
2. It affects the minimum difference or trend that is likely to be identified as statistically significant.

Both these effects can be evaluated using confidence intervals constructed with estimates of the sampling variability. Intervals around selected critical concentrations of nutrient analytes are listed in table 6. The sampling variabilities used to construct these intervals were estimated using the nonlinear model results in table 3 and the piecewise-linear model results in table 4.

Effects on Measured Concentrations

The first critical value listed for ammonia in table 6 is 0.07 mg/L, the minimum (pH and temperature dependent) criterion for protection of aquatic life (U.S. Environmental Protection Agency, 1986). At this concentration, the sampling variability is estimated to be about 0.004 mg/L. The range of the 95-percent confidence interval for an individual measurement of 0.07 mg/L is 0.062 to 0.078 mg/L. The potential error in that measurement due to sampling variability is ± 0.008 mg/L, which is a relative error of ± 11 percent of the measured concentration. Although this relative error might be considered large, the absolute error is small. For a measured ammonia concentration of 2.1 mg/L (the maximum aquatic-life criterion), the relative error is smaller (± 6 percent), but the absolute

error is larger (± 0.13 mg/L). Based on this analysis of the NAWQA data, 95 percent of all measured concentrations within the range of critical values identified for ammonia in streams are expected to differ from the actual concentrations by no more than 0.13 mg/L or 11 percent of the measurement, whichever is smaller.

In comparison to ammonia, the results for dissolved and total Kjeldahl nitrogen show increasing variability. This might be due to larger errors inherent in the Kjeldahl analysis than in the method used to analyze ammonia. In addition, the total Kjeldahl analysis is subject to errors in obtaining a representative laboratory subsample from the unfiltered field sample. For a measured concentration of 2.1 mg/L, the relative error increases from ± 6 percent for ammonia to ± 10 percent for dissolved Kjeldahl nitrogen and to ± 13 percent for total Kjeldahl nitrogen. A similar pattern can be seen in the phosphorus results. At the same measured concentrations, sampling variability is greater for dissolved phosphorus than for orthophosphate and greater for total

phosphorus than for dissolved phosphorus. The differences in analytic method among these three analytes are similar to the differences in methods for ammonia and Kjeldahl nitrogen, so method errors might be the primary cause of these differences in sampling variability.

Effects on Comparisons Between Concentrations

Measurement errors that result from sampling variability can affect identification of significant differences between two measurements (or means) if the differences are small. This effect can be determined by the size of the confidence intervals for the two measurements. If the confidence intervals do not overlap, the difference is considered statistically significant. If the intervals do overlap, significance cannot be determined. For example, the confidence interval for an individual measurement of 0.1 mg/L total phosphorus from a stream sample is 0.084–0.116 mg/L (table 6). The 95-percent

Table 6. Estimated sampling variability and confidence intervals around measured concentrations of nutrient analytes at selected critical values used to interpret environmental data

[mg/L, milligrams per liter]

Nutrient analyte	Critical value ¹ (mg/L)	Estimated sampling variability ² (mg/L)	95-percent confidence interval (mg/L)	
			Individual measurement	Mean of 10 measurements
Streams				
Ammonia	0.07	0.004	0.062–0.078	0.068–0.072
	0.1	0.004	0.092–0.108	0.098–0.102
	0.2	0.004	0.192–0.208	0.198–0.202
	2.1	0.065	1.97–2.23	2.06–2.14
Dissolved Kjeldahl nitrogen ³	2.1	0.10	1.90–2.30	2.04–2.16
Total Kjeldahl nitrogen ³	2.1	0.14	1.82–2.38	2.01–2.19
Nitrite plus nitrate	0.6	0.009	0.58–0.62	0.59–0.61
	10	0.072	9.86–10.14	9.96–10.04
Orthophosphate	0.05	0.003	0.045–0.055	0.048–0.052
Dissolved phosphorus ³	0.05	0.005	0.041– 0.059	0.047–0.053
	0.1	0.005	0.090–0.110	0.097–0.103
Total phosphorus	0.1	0.008	0.084–0.116	0.095–0.105
Ground water				
Nitrite plus nitrate	3	0.051	2.90–3.10	2.97–3.03
	10	0.13	9.75–10.25	9.92–10.08

¹See table 5 for description and source.

²From table 3 and equation 3.

³No critical values were identified in table 5 for dissolved or total Kjeldahl nitrogen or for dissolved phosphorus. The values used here are for comparison to other nutrient analytes.

confidence interval for a measurement of 0.13 mg/L, computed using equation 3 with parameter estimates from table 3 and equation 6 with $Z = 1.96$, is 0.109–0.151 mg/L. Because these intervals overlap, the two measurements cannot be considered significantly different. For a measurement of 0.14 mg/L, the confidence interval is 0.117–0.163 mg/L. Because this interval does not overlap the interval for a measurement of 0.1 mg/L, these two measurements can be considered significantly different.

For two mean values, the effect of sampling variability is moderated by the number of samples (see equation 6). The potential error due to sampling variability for a mean of 10 measurements is less than that for an individual measurement, and the confidence interval is smaller (table 6). Using the same example of total phosphorus, the confidence interval for a mean concentration of 0.1 mg/L is 0.095–0.105 mg/L. The 95-percent confidence interval for a mean of 0.13 mg/L is 0.123–0.137 mg/L. Because these intervals do not overlap, the difference between these two means is statistically significant. A mean of 0.12 mg/L also is significantly different from 0.1 mg/L, but a mean of 0.11 mg/L cannot be distinguished from 0.1 mg/L due to the potential error resulting from sampling variability. An actual set of 10 samples also might be affected by environmental variability, so even a larger difference might not be statistically significant. However, sampling variability determines the smallest difference that is likely to be identified as significant.

For any selected concentration, the minimum difference that is not likely to be affected by sampling variability can be estimated as twice the potential error indicated by the confidence interval. For total phosphorus, the potential error for an individual measurement of 0.1 mg/L is ± 0.016 mg/L, based on the 95-percent confidence limit in table 6. The minimum difference unaffected by this error is about 0.032 mg/L. Measurements within this minimum difference will not be significantly different from 0.1 mg/L. Thus, as previously shown, a measurement of 0.13 mg/L is not, but a measurement of 0.14 mg/L is, significantly different from 0.1 mg/L. Likewise, a measurement of 0.07 mg/L is not, but a measurement of 0.06 mg/L is, significantly different from 0.1 mg/L. Thus, a measurement of 0.13 mg/L does not indicate, with 95-percent confidence, that the USEPA recommended limit for total phosphorus has been exceeded. And a measurement of 0.07 mg/L does not indicate, with the same confidence, that the

limit has not been exceeded. Sampling variability causes uncertainty in identifying an exceedance for total phosphorus concentrations in the range of 0.07–0.13 mg/L.

The minimum differences unaffected by sampling variability for critical concentrations of most nutrient analytes are small. For orthophosphate and most of the critical values for ammonia listed in table 6, differences of 0.02 mg/L would be considered significant for individual measurements, and differences of 0.01 mg/L between means of 10 measurements would not likely be affected by sampling variability. The latter result is at the limit of resolution for data from the USGS National Water-Quality Laboratory (NWQL), so sampling variability has essentially no effect on reported differences in concentrations of these analytes. For the highest critical value for ammonia, differences in individual measurements of 0.3 mg/L would be considered significant. Therefore, measurements in the range of 1.9–2.3 mg/L do not indicate compliance with or exceedance of the 2.1 mg/L criterion at the 95-percent confidence level. However, any measurement greater than 2.3 mg/L exceeds the criterion at a confidence level of at least 95 percent. For nitrite plus nitrate measurements at the drinking-water standard (10 mg/L) in stream samples, minimum difference unaffected by sampling variability also is about 0.3 mg/L. Measurements in the range of 9.8–10.2 mg/L are not significantly different from this standard. However, the limit of resolution for NWQL data greater than 10 mg/L is 1 mg/L, so the nearest possible measurement greater than 10 mg/L is 11 mg/L, which would be considered significantly greater than the standard. Thus, sampling variability has no effect on measurements greater than the standard, but might introduce some uncertainty in measurements that are within 0.2 mg/L less than the standard. For nitrite plus nitrate measurements in ground-water samples, the potential error is slightly larger, so the range of uncertainty is broader. For a measurement at the standard (10 mg/L), the minimum difference unaffected by sampling variability is about 0.6 mg/L, and the range of uncertainty for exceedance of the standard is 9.5–10.5 mg/L, which would be rounded to 9.5–11 mg/L. Thus, 12 mg/L is the lowest measurement that would indicate an exceedance with at least 95-percent confidence.

CONCLUSIONS

Statistical methods were identified that are appropriate for analysis of contamination bias and sampling variability in a set of data represented by quality (QC) samples. These methods were applied to nutrient data collected during 1993–95 in 20 study units investigated as part of the NAWQA Program. The available QC samples included blanks and replicates.

Contamination potentially affects measured concentrations of less than 0.4 mg/L for ammonia and less than 0.1–0.3 mg/L for various phosphorus analytes. These ranges include some environmentally significant concentrations, based on established water-quality criteria and standards. Nitrite plus nitrate and the Kjeldahl nitrogen analytes are essentially unaffected by contamination.

Sampling variability can affect interpretation of small differences between individual measurements or mean concentrations. For the NAWQA data, these effects are limited to a range of a few hundredths to a few tenths of a milligram per liter, depending on analyte and concentration. For environmentally significant concentrations of several analytes, the effect of sampling variability is less than the resolution of the reported data (due to rounding). For this reason, there is essentially no effect of sampling variability on measured concentrations of nitrite plus nitrate and orthophosphate, nor for ammonia at concentrations of about 0.2 mg/L or less. The potential errors due to sampling variability are larger for dissolved Kjeldahl nitrogen and dissolved phosphorus, and even larger for total Kjeldahl nitrogen and total phosphorus. These increases seem related to differences in analytic methods. The uncertainty in measured or mean concentrations caused by sampling variability is of small range for all nutrients. Differences in concentration within these ranges have limited, if any, environmental significance.

These results can be applied to interpretation of the environmental data collected during 1993–95 in the 20 NAWQA study units. The results also provide a basis for comparison with QC results for samples collected in subsequent years by other study units.

REFERENCES

- Chambers, J.M., Cleveland, W.S., Kleiner, Beat, and Tukey, P.A., 1983, *Graphical methods for data analysis*: Boston, Mass., Duxbury Press, 395 p.
- Hahn, G.J., and Meeker, W.Q., 1991, *Statistical intervals—A guide for practitioners*: New York, John Wiley and Sons, 392 p.
- Iman, R.L., and Conover, W.J., 1983, *A modern approach to statistics*: New York, John Wiley and Sons, 497 p.
- Madison, R.J., and Brunett, J.O., 1985, Overview of the occurrence of nitrate in ground water of the United States, *in* National water summary 1984—Hydrologic events, selected water-quality trends, and ground-water resources: U.S. Geological Survey Water-Supply Paper 2275, p. 93–105.
- Mueller, D.K., Hamilton, P.A., Helsel, D.R., Hitt, K.J., and Ruddy, B.C., 1995, *Nutrients in ground water and surface water of the United States—An analysis of data through 1992*: U.S. Geological Survey Water-Resources Investigations Report 95–4031, 74 p.
- SAS Institute, 1990, *SAS/GRAPH software*: Cary, N.C., SAS Institute Inc., reference, version 6, 1st ed., v. 1, 794 p.
- Snedecor, G.W., and Cochran, W.G., 1980, *Statistical methods*, 7th ed.: Ames, Iowa State University Press, 507 p.
- U.S. Environmental Protection Agency, 1986, *Quality criteria for water 1986*: Washington, D.C., Office of Water, U.S. Environmental Protection Agency Report 440/5–86–001 [variously paged].