Quality of Nutrient Data from Streams and Ground Water Sampled During Water Years 1992–2001

By David K. Mueller and Cindy J. Titus

National Water-Quality Assessment Program

Scientific Investigations Report 2005–5106

U.S. Department of the Interior
U.S. Geological Survey

Foreword

The U.S. Geological Survey (USGS) is committed to providing the Nation with accurate and timely scientific information that helps enhance and protect the overall quality of life and that facilitates effective management of water, biological, energy, and mineral resources (http://www.usgs.gov/). Information on the quality of the Nation's water resources is critical to assuring the long-term availability of water that is safe for drinking and recreation and suitable for industry, irrigation, and habitat for fish and wildlife. Population growth and increasing demands for multiple water uses make water availability, now measured in terms of quantity and quality, even more essential to the long-term sustainability of our communities and ecosystems.

The USGS implemented the National Water-Quality Assessment (NAWQA) Program in 1991 to support national, regional, and local information needs and decisions related to water-quality management and policy (http://water.usgs.gov/nawqa). Shaped by and coordinated with ongoing efforts of other Federal, State, and local agencies, the NAWQA Program is designed to answer: What is the condition of our Nation's streams and ground water? How are the conditions changing over time? How do natural features and human activities affect the quality of streams and ground water, and where are those effects most pronounced? By combining information on water chemistry, physical characteristics, stream habitat, and aquatic life, the NAWQA Program aims to provide science-based insights for current and emerging water issues and priorities.

From 1991–2001, the NAWQA Program completed interdisciplinary assessments in 51 of the Nation's major river basins and aquifer systems, referred to as Study Units (http://water.usgs.gov/nawqa/studyu.html). Baseline conditions were established for comparison to future assessments, and long-term monitoring was initiated in many of the basins. During the next decade, 42 of the 51 Study Units will be reassessed so that 10 years of comparable monitoring data will be available to determine trends at many of the Nation's streams and aquifers. The next 10 years of study also will fill in critical gaps in characterizing water-quality conditions, enhance understanding of factors that affect water quality, and establish links between sources of contaminants, the transport of those contaminants through the hydrologic system, and the potential effects of contaminants on humans and aquatic ecosystems.

The USGS aims to disseminate credible, timely, and relevant science information to inform practical and effective water-resource management and strategies that protect and restore water quality. We hope this NAWQA publication will provide you with insights and information to meet your needs, and will foster increased citizen awareness and involvement in the protection and restoration of our Nation's waters.

The USGS recognizes that a national assessment by a single program cannot address all water-resource issues of interest. External coordination at all levels is critical for a fully integrated understanding of watersheds and for cost-effective management, regulation, and conservation of our Nation's water resources. The NAWQA Program, therefore, depends on advice and information from other agencies—Federal, State, interstate, Tribal, and local—as well as nongovernmental organizations, industry, academia, and other stakeholder groups. Your assistance and suggestions are greatly appreciated.

Robert M. Hirsch

Associate Director for Water

U.S. Department of the Interior
Gale A. Norton, Secretary

U.S. Geological Survey
P. Patrick Leahy, Acting Director

U.S. Geological Survey, Reston, Virginia: 2005

This report supersedes U.S. Geological Survey Open-File Report 98–276: Quality of Nutrient Data from Streams and Ground Water Sampled during 1993–95—National Water-Quality Assessment program.

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

For more information about the USGS and its products:
Telephone: 1-888-ASK-USGS
World Wide Web: http://www.usgs.gov/

Contents

Figures

Tables

Abbreviations, Acronyms, and Converson Factors

LOWESS	Locally-weighted scatterplot smoothing
mg/L	Milligrams per liter
NAWQA	National Water-Quality Assessment
QC	Quality control
RSD	Relative standard deviation
UCL	Upper confidence limit
USEPA	United States Environmental Protection Agency
USGS	United States Geological Survey

Concentrations of chemical constituents in water are given in mg/L, which is approximately equal to parts per million.

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$°F = (1.8 \times °C) + 32$$

Quality of Nutrient Data from Streams and Ground Water Sampled During Water Years 1992–2001

By David K. Mueller and Cindy J. Titus

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, shipment, and analysis. Contamination can adversely affect interpretation of measured concentrations in comparison to standards or criteria. Variability can affect interpretation of small differences between individual measurements or mean concentrations. Contamination and variability are determined for nutrient data from quality-control samples (field blanks and replicates) collected as part of the National Water-Quality Assessment (NAWQA) Program during water years 1992–2001. Statistical methods are used to estimate the likelihood of contamination and variability in all samples. Results are presented for five nutrient analytes from stream samples and four nutrient analytes from ground-water samples. Ammonia contamination can add at least 0.04 milligram per liter in up to 5 percent of all samples. This could account for more than 22 percent of measured concentrations at the low range of aquatic-life criteria (0.18 milligram per liter). Orthophosphate contamination, at least 0.019 milligram per liter in up to 5 percent of all samples, could account for more than 38 percent of measured concentrations at the limit to avoid eutrophication (0.05 milligram per liter). Nitrite-plus-nitrate and Kjeldahl nitrogen contamination is less than 0.4 milligram per liter in 99 percent of all samples; thus there is no significant effect on measured concentrations of environmental significance. Sampling variability has little or no effect on reported concentrations of ammonia, nitrite-plus-nitrate, orthophosphate, or total phosphorus sampled after 1998. The potential errors due to sampling variability are greater for the Kjeldahl nitrogen analytes and for total phosphorus sampled before 1999. The uncertainty in a mean of 10 concentrations caused by sampling variability is within a small range (1 to 7 percent) for all nutrients. These results can be applied to interpretation of environmental data collected during water years 1992–2001 in 52 NAWQA study units.

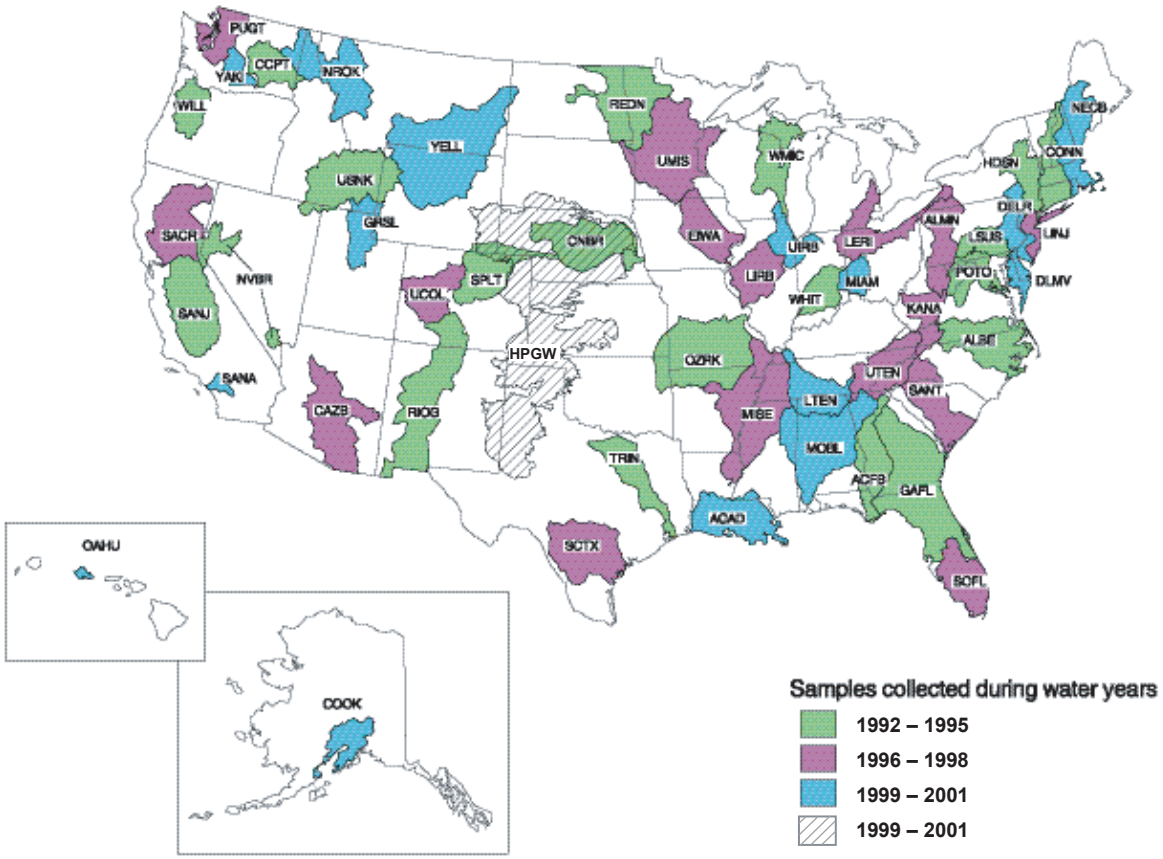
Introduction

To determine the extent of contamination in the Nation's streams and ground water, Congress appropriated funds beginning in 1991 for a National Water-Quality Assessment (NAWQA) Program, which is 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; and
- Improve understanding of the primary natural and human factors that affect water-quality conditions.

These objectives are being achieved through investigations in 52 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. During Cycle I of the NAWQA Program (1991–2001), the high-intensity phase of water sampling occurred over a period of 3 water years in each study unit. (Water year is the period from October through September and is identified by the year in which it ends.) In the first 20 study units, limited sampling began in water year 1992 and high-intensity sampling occurred during water years 1993–95. This phase was followed by low-intensity sampling at selected sites during water years 1996–2001. In the next group, consisting of 16 study units, high-intensity sampling occurred during water years 1996–1998, followed by low-intensity sampling at selected sites during water years 1999–2001. In the remaining 16 study units, high-intensity sampling occurred during water years 1999–2001. The locations of these three groups of study units are shown in figure 1.

To quantify how much variability in water-quality measurements can be explained by field and laboratory methods as compared to environmental factors, estimates are needed of the bias and variability that result from sample collection, processing, shipment, and laboratory analysis. Bias is the systematic error inherent in sampling and laboratory methods, and can be either positive or negative. A common source of positive bias that can affect water-quality data is contamination of



ACAD	Acadian-Pontchartrain
ACFB	Apalachicola-Chattahoochee-Flint River Basin
ALBE	Albemarle-Pamlico Drainage
ALMN	Allegheny/Monongahela
CAZB	Central Arizona Basins
CCPT	Central Columbia Plateau
CNBR	Central Nebraska Basins
CONN	Connecticut, Housatonic, and Thames River Basins
COOK	Cook Inlet Basin
DELR	Delaware River Basin
DLMV	Delmarva Penninsula
EIWA	Eastern Iowa Basins
GAFL	Georgia-Florida Coastal Plain
GRSL	Great Salt Lake Basins
HDSN	Hudson River Basin
HPGW	High Plains Regional Ground Water Study
KANA	Kanawha-New River Basin
LERI	Lake Erie-Lake St. Clair Drainage
LINJ	Long Island-New Jersey Coastal Drainages
LIRB	Lower Illinois River Basin
LSUS	Lower Susquehanna River Basin
LTEN	Lower Tennessee River Basin
MIAM	Great and Little Miami River Basins
MISE	Mississippi Embayment
MOBL	Mobile River and Tributaries
NECB	New England Coastal Basins
NROK	Northern Rockies Intermontane Basins
NVBR	Nevada Basin and Range
OAHU	Oahu
OZRK	Ozark Plateaus
POTO	Potomac River Basin
PUGT	Puget Sound Basin
REDN	Red River of the North Basin
RIOG	Rio Grande Valley
SACR	Sacramento Basin
SANA	Santa Ana Basin
SANJ	San Joaquin-Tulare Basins
SANT	Santee Basin and Coastal Drainages
SCTX	South Central Texas
SOFL	Southern Florida
SPLT	South Platte River Basin
TRIN	Trinity River Basin
UCOL	Upper Colorado River Basin
UIRB	Upper Illinois River Basin
UMIS	Upper Mississippi River Basin
USNK	Upper Snake River Basin
UTEN	Upper Tennessee River Basin
WHIT	White River Basin
WILL	Willamette Basin
WMIC	Western Lake Michigan Drainages
YAKI	Yakima River Basin
YELL	Yellowstone Basin

Figure 1. Locations of the 52 NAWQA study units that contributed the data analyzed in this report.

samples. Contaminants can be introduced into water samples during sample collection, processing, shipping, or laboratory analysis by exposure to airborne gases and particulates or from inadequately cleaned sampling or analytic equipment. Variability is the degree of random error in sampling and laboratory methods and can be estimated from repeated, independent measurements of the same water sample. In water-quality data, variability results from the nonsystematic 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 relation to measurements of the environmental samples collected as part of a water-quality assessment. A glossary at the back of this report defines QC terms that might be unfamiliar to some readers.

Purpose and Scope

This report describes the quality of nutrient data in stream and ground-water samples collected in 52 NAWQA study units during water years 1992–2001 based on an analysis of contamination bias and sampling variability. The results of this QC 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 and Alaska and Hawaii. The data were aggregated into a national data set during the spring of 2002 and were extensively reviewed. This data set has been posted on the NAWQA Nutrient Synthesis web page (http://water.usgs.gov/nawqa-only/nutrient.html). Essentially all nutrient QC samples collected for the NAWQA Program during October 1991 through September 2001 study units are included.

Acknowledgments

The philosophy of QC data interpretation followed in this report was developed by the NAWQA QC work group: Michael Koterba (Chairman), Greg Delzer, Jeff Martin, Dave Mueller, Terry Schertz, and Jon Scott. In addition, Ed Gilroy (USGS, retired) and Mark Brigham (NAWQA Program, USGS Minnesota Water Science Center) 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 all 52 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. Chemical and biological processes can alter these forms and can transfer nitrogen and phosphorus to or from water, soil, biological organisms, and the atmosphere. Nutrient concentrations in water generally are 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 important 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, shipping, 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 vari-

ability 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 water years 1992–2001 at locations within the 52 NAWQA study units shown in figure 1. The numbers of field blanks and replicate-sample sets used in this report from each study unit are listed in table 1.

The results of chemical analyses on quality-control samples and associated environmental samples were provided by each study unit to a national database. The nutrient data used in this report were retrieved from that database and reviewed to identify potential errors. Study-unit personnel provided corrections where appropriate.

Methods of Data Analysis

Bias and variability are determined by statistical analysis of blanks and replicate samples, respectively. 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

In order to avoid false-positive quantification of a constituent, very low concentrations are censored, reported as a "less than" value by the laboratory. The USGS National Water-Quality Laboratory uses several types of censoring levels (Childress and others, 1999). The most basic is the method detection limit (hereinafter referred to as "detection limit"), defined as the minimum concentration that can be measured, with 99-percent confidence, to be significantly greater than zero. 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 the amount of contamination that is not likely to be exceeded in a large percentage of the water samples represented by the blanks. This objective can be achieved by constructing an upper confidence limit (UCL) for a high percentile of contamination in the population of water samples that includes environmental samples and blanks. This UCL is the maximum contamination expected in the specified percentage of water samples. For example, the 95-percent UCL for the 90th percentile of concentrations in blanks is the maximum contamination expected in 90 percent of all water samples. The 95-percent confidence level indicates there is only a 5-percent chance that this contamination has been underestimated. Another way to express this is that we are 95-percent confident that this amount of contamination would be exceeded in no more than 10 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) describe a method for determining a distribution-free UCL for a percentile, 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 UCL. The binomial function (B) is used to calculate the probability that no more than $n - u$ values from a total of n observations exceed the 100pth percentile of the sampled population. The rank (u) is chosen as the smallest integer such that:

$$B(u - 1, n, p) \geq 1 - \alpha \qquad (1)$$

The value of the 100(1–α) percent UCL for the 100pth percentile of contamination in the population then is determined by the measured value of the u ranked observation. For example, in a group of 100 blanks, the 95-percent UCL for the 90th percentile can be determined as follows. First find the smallest value of u that meets the criterion:

$$B(u - 1,100, 0.90) \geq 0.95 \qquad (2)$$

For $u = 95$, $B = 0.942$, which is less than the criterion of 0.95, but for $u = 96$, $B = 0.976$, which meets the criterion. Thus, the value of the 95-percent UCL is determined by the concentration of the 96th ranked blank.

Contamination bias in the environmental samples is estimated from the UCL calculated using blank data. In the previous example, if the concentration of compound Y in the

Table 1. Number of quality-control samples collected in each of the 52 National Water-Quality Assessment study units that were used for the data analysis in this report.

Study Unit (see fig. 1 for location)	Number of Field Blanks		Number of Replicate-Sample Sets	
	Streams	Ground Water	Streams	Ground Water
ACAD	26	3	51	4
ACFB	36	10	49	34
ALBE	64	14	13	10
ALMN	15	12	20	12
CAZB	27	19	10	22
CCPT	54	27	38	21
CNBR	39	16	30	5
CONN	34	20	16	8
COOK	11	2	13	2
DELR	24	4	16	3
DLMV	8	2	6	4
EIWA	12	10	36	14
GAFL	31	4	39	14
GRSL	17	13	15	11
HDSN	70	9	35	6
HPGW	0	18	0	21
KANA	19	8	19	1
LERI	9	8	11	7
LINJ	18	22	20	19
LIRB	12	16	11	11
LSUS	34	17	32	16
LTEN	23	7	26	7
MIAM	21	8	30	9
MISE	34	11	27	9
MOBL	24	5	25	5
NECB	28	13	28	5
NROK	3	3	3	3
NVBR	25	13	17	7
OAHU	6	2	2	2
OZRK	33	6	23	9
POTO	33	8	65	13
PUGT	38	8	31	4
REDN	11	23	16	29
RIOG	37	12	44	12
SACR	18	10	21	10

Table 1. Number of quality-control samples collected in each of the 52 National Water-Quality Assessment study units that were used for the data analysis in this report.—Continued

Study Unit (see fig. 1 for location)	Number of Field Blanks		Number of Replicate-Sample Sets	
	Streams	Ground Water	Streams	Ground Water
SANA	16	10	22	12
SANJ	15	1	27	18
SANT	85	26	8	9
SCTX	28	10	30	10
SOFL	19	12	26	3
SPLT	73	19	67	22
TRIN	36	9	45	7
UCOL	59	20	63	14
UIRB	22	7	31	5
UMIS	26	1	28	10
USNK	51	12	23	13
UTEN	24	5	19	3
WHIT	45	10	40	10
WILL	44	8	37	3
WMIC	49	4	48	8
YAKI	10	0	13	0
YELL	19	4	20	4
Total	1,515	541	1,385	520

96th ranked blank was 0.1 mg/L, contamination bias can be described as follows:

"Contamination by Y is estimated, with at least 95-percent confidence, to exceed 0.1 mg/L in no more than 10 percent of all samples."

This amount of contamination then can be compared to environmentally important concentrations of compound Y to determine the likelihood that contamination has affected interpretation of the environmental data. Important concentrations include background concentrations at undisturbed sites, minimum concentrations that indicate anthropogenic effects, and water-quality standards and criteria. Continuing with the previous example, suppose compound Y has a drinking-water standard of 10 mg/L. Contamination as great as 0.1 mg/L in 90 percent of all samples is unlikely to affect a measurement that exceeds this standard. The true environmental concentration would have to be more than 9.9 mg/L (but less than 10 mg/L) for this amount of contamination to produce a false exceedance. However, if the standard for compound Y was 0.2 mg/L, contamination in as many as 10 percent of all samples might account for one-half or more of a measured exceedance. In this instance, false exceedances would be likely enough that utility of the data would be compromised.

Methods Used to Determine Variability

Sampling variability can be estimated by using some measure of the dispersion of repeated measurements, such as the standard deviation of 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 instance, 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 a replicate set is correlated with the mean concentration of that constituent in the replicates. Over a low range of concentrations, standard deviation of replicates generally is uniform, but at higher concentrations, standard deviation tends to increase in proportion to concentration. Within this high range, the relative standard deviation (RSD), defined as the standard deviation divided by the mean concentration, is generally uniform. RSD also is referred to as the "coefficient of variation."

Variability over a large range of concentrations can be approximated by dividing this range into segments over which either the standard deviation or the RSD are reasonably

constant (Anderson, 1987). This method commonly is used to define laboratory analytic precision. Over the low-concentration range, variability is estimated as the average standard deviation of replicates; over the high-concentration range, variability is estimated as the average RSD.

Concentration ranges for this method can be selected by graphical analysis of standard deviation and 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. Some adjustment to the boundary concentration might be necessary if the average low-range standard deviation and high-range RSD do not intersect at the boundary.

A number of problems can interfere with making a good estimate of variability from 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 are predominant in field-replicate data, and few or no data might be available at high concentrations. In this instance, variability within the high range of concentrations might be impossible to define. Another issue 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, determination of variability for the high range of concentrations might be adversely affected.

After sampling variability has been estimated from field-replicate data, it can be used to construct confidence intervals. For any concentration (C) measured in a single sample, the confidence interval for the true concentration is:

$$[C_L, C_U] = C \pm Z_{(1-\alpha/2)}\,\sigma \qquad (3)$$

where

C_L, C_U is the lower and upper limits of concentration for the $100(1-\alpha)$ percent confidence interval;

Z is the ordinate of the normal curve (Z-value) that contains $100(1-\alpha)$ percent of the distribution;

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

σ is the sampling variability for the measured concentration.

If the measured concentration is in the low range, σ is the average standard deviation of replicates within that range. If the measured concentration is in the high range, $\sigma = C\,(\mathrm{RSD}/100)$.

The second term, $Z_{(1-\alpha/2)}\,\sigma$, in equation 3 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 $100(1-\alpha)$ percent confidence) to determine whether the concentration in the sample exceeds the standard.

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

$$[C_L, C_U] = \overline{C} \pm Z_{(1-\alpha/2)}\frac{\sigma}{\sqrt{n}} \qquad (4)$$

where

n is the number of samples;

$\overline{C}$ is the mean concentration in these samples;

and the other variables are as previously defined.

Again, the error due to sampling variability is represented by the second term of equation 4, but in this instance 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. A determination of 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

Water samples, including field blanks and replicates, collected for the NAWQA Program were obtained following standard protocols (Shelton, 1994; U.S. Geological Survey, variously dated). Samples for analysis of dissolved constituents were filtered in the field through either a nitrocellulose filter or a polyether-sulfone medium with a pore size of 0.45 micrometers. All samples were analyzed at the USGS National Water-Quality Laboratory in Colorado. Filtered stream and ground-water samples were analyzed for the following nutrient constituents:

- 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”)
- Orthophosphate as phosphorus (hereinafter referred to as “orthophosphate”)

Ground-water samples also were analyzed for:

- Dissolved Kjeldahl nitrogen (the Kjeldahl analysis includes ammonia plus organic nitrogen)

Two additional nutrient analytes were determined for

unfiltered stream samples:

- Total Kjeldahl nitrogen
- Total phosphorus

Nitrite concentrations generally are less than the detection limit in streams and in oxygenated ground water; therefore, the analyte "nitrite as nitrogen" was not included in the QC analysis for this report.

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 database 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. The first step in evaluating the QC data for each analyte was to evaluate the relation between analyte concentrations and dates of blank preparation to identify possible temporal trends in contamination. Because the blank data include many censored values, trends were determined by maximum-likelihood regression (Helsel, 2005, p. 201–209). Statistically significant trends ($p < 0.10$) were identified for all nutrient constituents except nitrite-plus-nitrate (table 2). A lognormal model was used, so trends are expressed in percent change in concentration per year. All these trends were associated with changes in laboratory detection limits, and all trend lines (fitted concentrations) were contained entirely below the most common detection limit for each constituent. The distribution of concentrations greater than the most common detection limit seemed more affected by changes in sampling frequency and reporting level (number of significant decimal places) than by any overall increase or decrease during the period of record. For consistency over time, concentrations in blanks were re-censored to the most common detection limit for all subsequent analyses. This censoring had no effect on measured concentrations at the higher percentiles of contamination.

Distributions of nutrient concentrations in stream and ground-water field blanks are shown in figure 2. Dashed lines on the bars represent data below the most common detection limit for each analyte. For all analytes, the 95th percentile of measured values was at or near the detection limit.

Data from the blanks were used to calculate UCLs for selected percentiles of contamination in all samples collected during NAWQA Cycle I (water years 1992–2001). A 99-percent level of confidence was selected, and separate calculations were made for the 80th, 90th, 95th, and 99th percentiles (table 3). Potential contamination in at least 80 percent of all samples is estimated to be no greater than the detection limit for all nutrient analytes except ammonia in ground water. Plots of the 99-percent UCLs in figure 3 show that potential contamination remains low through at least the 95th percentiles for all nutrient species; however, contamination can be much higher in a small percentage of samples.

Table 2. Trend-analysis results for data from field blanks prepared at stream and ground-water sampling sites.

[mg/L, milligrams per liter; %, percent; <, less than; N, nitrogen; P, phosphorus]

Nutrient analyte	Number of blanks	Trend slope (%/year)	p-value	Fitted concentration (mg/L)			Most common detection limit (mg/L)
				1992	2001	Change	
Streams							
Ammonia, as N	1,507	-14.7	<0.0001	0.0176	0.0039	-0.0137	0.02
Nitrite + Nitrate, as N	1,511	2.1	0.4220	0.0053	0.0065	0.0012	0.05
Total Kjeldahl Nitrogen	1,445	-7.5	<0.0001	0.0542	0.0262	-0.0280	0.2
Orthophosphate, as P	1,507	4.3	0.0639	0.0015	0.0022	0.0007	0.01
Total Phosphorus	1,397	-24.9	<0.0001	0.0061	0.0004	-0.0057	0.01
Ground water							
Ammonia, as N	537	-5.4	0.0002	0.0188	0.0121	-0.0067	0.02
Nitrite + Nitrate, as N	541	-0.9	0.7510	0.0190	0.0177	-0.0013	0.05
Dissolved Kjeldahl Nitrogen	536	-12.5	<0.0001	0.1479	0.0518	-0.0961	0.2
Orthophosphate, as P	537	8.8	0.0002	0.0034	0.0066	0.0032	0.01

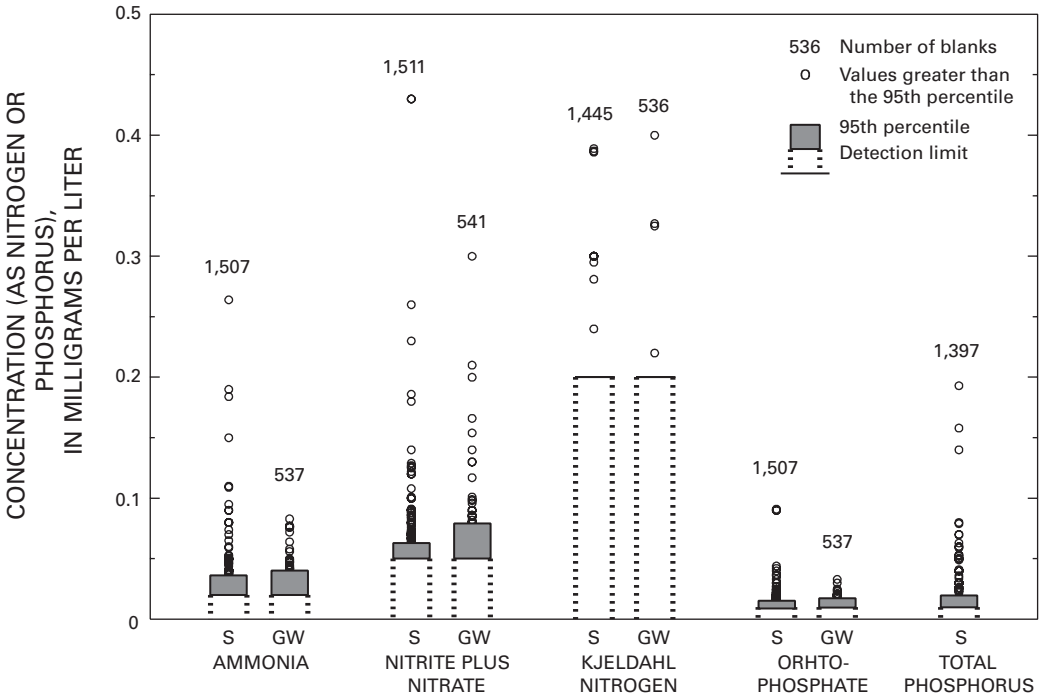


Figure 2. Distribution of nutrient concentrations measured in field blanks from stream(S) and ground-water (GW) sampling sites collected during 1992 through 2001. [Kjeldahl nitrogen is total in stream samples and dissolved in ground-water samples.]

Table 3. Upper 99-percent confidence limits for contamination by nutrients in specified percentiles of all samples based on data from field blanks prepared at stream and ground-water sampling sites.

[mg/L, milligrams per liter; <, less than; N, nitrogen; P, phosphorus]

Nutrient analyte	Number of blanks	Most common detection limit (mg/L)	Upper 99-percent confidence limit (mg/L)			
			80th percentile	90th percentile	95th percentile	99th percentile
Streams						
Ammonia, as N	1,507	0.02	0.02	0.03	0.04	0.095
Nitrite + Nitrate, as N	1,511	0.05	<0.05	0.054	0.07	0.23
Total Kjeldahl Nitrogen	1,445	0.2	<0.2	<0.2	<0.2	0.39
Orthophosphate, as P	1,507	0.01	<0.01	0.011	0.019	0.037
Total Phosphorus	1,397	0.01	<0.01	0.02	0.03	0.079
Ground water						
Ammonia, as N	537	0.02	0.03	0.037	0.044	0.083
Nitrite + Nitrate, as N	541	0.05	<0.05	0.07	0.09	0.3
Dissolved Kjeldahl Nitrogen	536	0.2	<0.2	<0.2	<0.2	0.4
Orthophosphate, as P	537	0.01	0.01	0.015	0.02	0.033

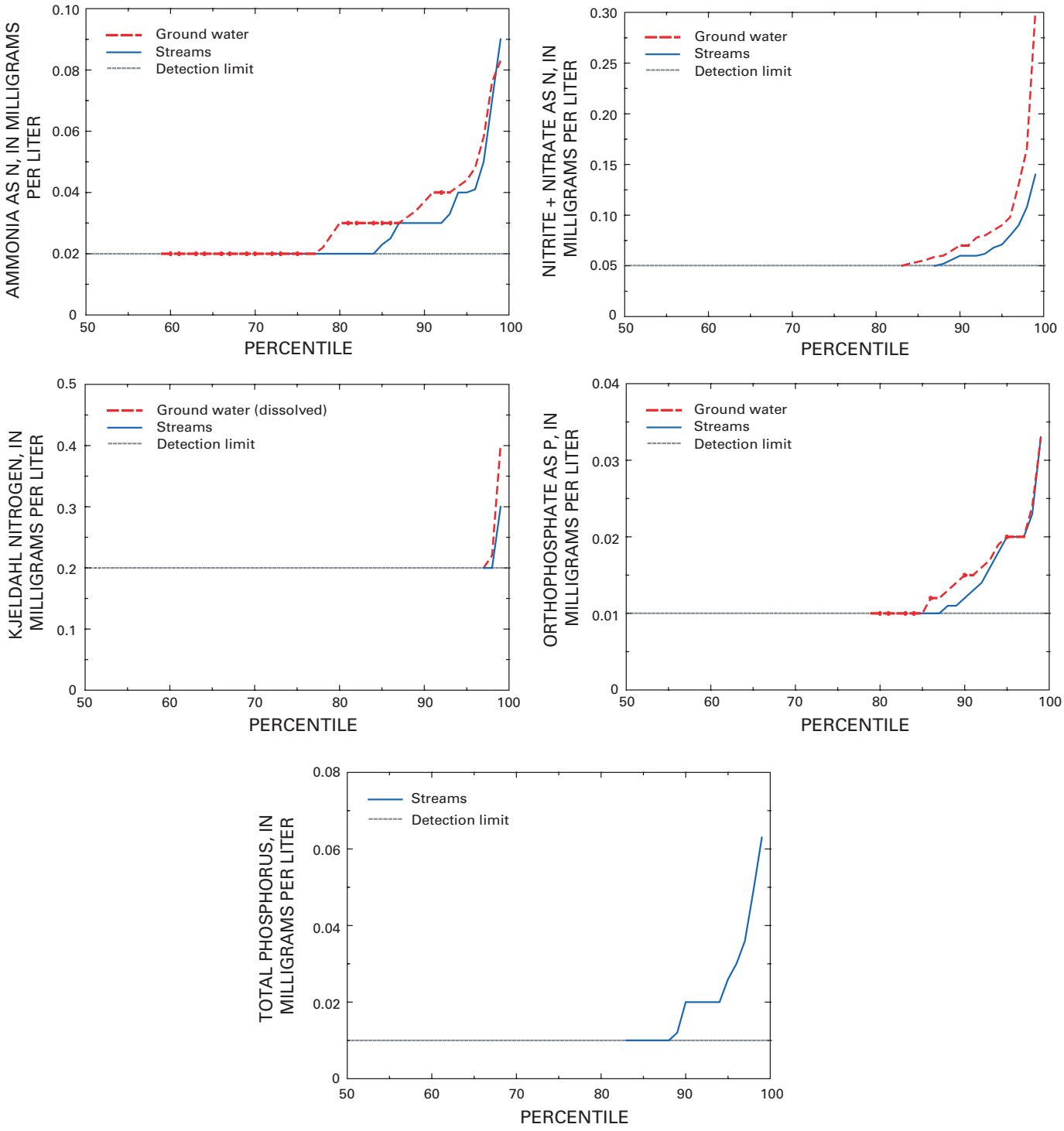


Figure 3. Upper 99-percent confidence limits for contamination by nutrients in all samples based on data from field blanks.

The rightmost column in table 3 lists the contamination that is likely to be exceeded in no more than 1 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 as N (U.S. Environmental Protection Agency, 1986), which is more than 40 times greater than the maximum contamination expected in 99 percent of all samples. Thus, contamination is unlikely to cause problems with identifying exceedances of the nitrate 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 0.18 mg/L, depending on water temperature and pH (U.S. Environmental Protection Agency, 1999). At this concentration, there is a 1-percent chance that 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) has recommended a maximum concentration of 0.1 mg/L total phosphorus in streams (U.S. Environmental Protection Agency, 1986). Contamination in 1 percent of stream samples might account for almost 80 percent of this value. Because of these issues, data for ammonia and total phosphorus at low concentrations can be of limited use for comparison to standards. However, for reported concentrations more than 10 times the potential contamination listed in table 3, contamination issues 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 distributions 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 instance, 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. An evaluation previously made using

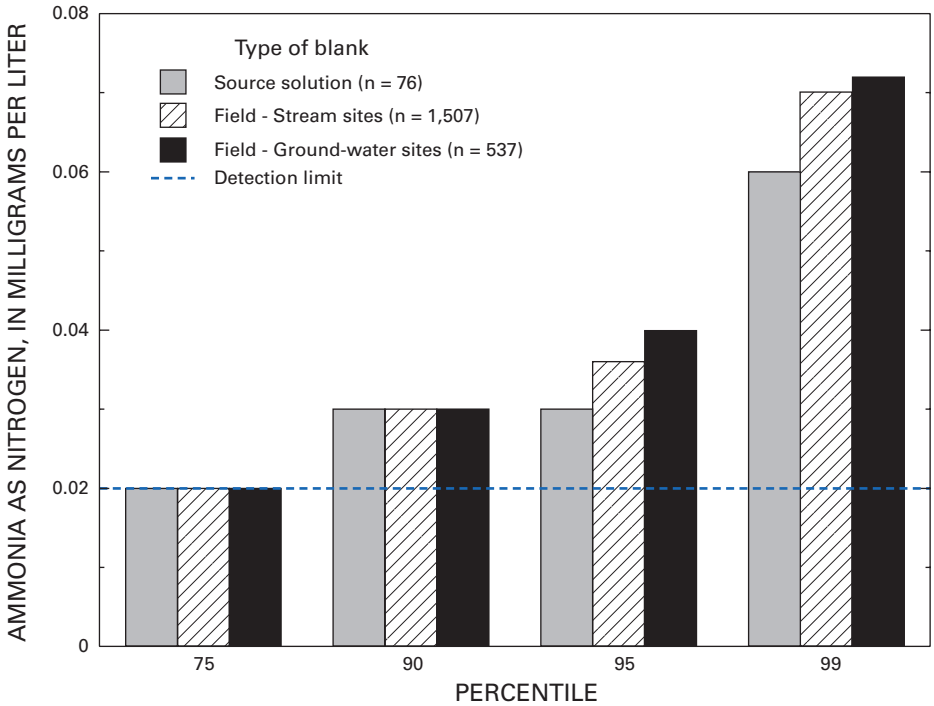


Figure 4. Selected percentiles of the distributions of ammonia contamination in source-solution and field blanks.

data from the first 20 study units indicated little difference in variability among replicate types: split, concurrent, or sequential (Mueller, 1998). Therefore, all these types of replicates were combined in subsequent analyses. However, the NAWQA samples were carefully scrutinized to ensure that only true field replicates were included. Topical replicates, collected to compare different field or laboratory methods, were identified and excluded.

For each nutrient analyte, sampling variability was estimated using standard deviation over a low range of concentrations and relative standard deviation over a high range. Separate analyses were made for nutrients in streams and nutrients in ground water. First, replicate standard deviation and RSD were plotted for each analyte to determine the division between the low-range and high-range concentrations (figs. 5 and 6). In the plots of standard deviation, concentration is shown on a logarithmic scale to emphasize the lower range. All plots include a smooth curve through the data points to show the general relation between concentration and standard deviation or RSD. Where this curve is horizontal, there is no relation between standard deviation or RSD and concentration, so variability can be considered constant. The boundary concentration was determined by finding a point on the x-axis below which the curve for standard deviation was essentially horizontal and above which the curve for RSD was essentially horizontal. The division between low and high concentrations is represented on each plot with a vertical dashed line. After the low-range and high-range concentrations for each nutrient analyte were determined, sampling variability was estimated as the average standard deviation for the low range or the average RSD for the high range (table 4).

Several interesting comparisons can be made among the variability estimates in table 4. Standard deviations over the low range of mean concentrations and RSD over the high range are about the same for stream and ground-water samples, with two exceptions. The standard deviation for nitrite-plus-nitrate and the RSD for orthophosphate are more than three times greater in ground-water samples than in stream samples. These anomalies could be a result of the relatively small number of ground-water replicates within the low range of nitrite-plus-nitrate and the high range of orthophosphate. Other than the seemingly anomalous value for orthophosphate in ground water, sampling variabilities within the high-concentration range are largest for Kjeldahl nitrogen (total in stream samples and dissolved in ground-water samples) and for total phosphorus. This large variability could be an effect of laboratory methods. Analysis of Kjeldahl nitrogen and total phosphorus potentially can introduce more measurement error than the methods used to analyze ammonia, nitrite-plus-nitrate, and orthophosphate. In addition, the "total" constituent analyses are subject to errors in obtaining a representative laboratory subsample from unfiltered field samples.

Trends in Variability

The stability of each variability value over the period of record was evaluated by analyzing temporal trends in the standard deviation and RSD for concentrations of each constituent. Separate analyses were made for stream and ground-water samples. Trends in standard deviation were computed only for samples in the low range of mean concentration; trends in RSD were computed only for samples in the high range. Statistically significant trends ($p < 0.05$) were identified for only 6 of the 18 measures of variability listed in table 4. Two of these trends, for the standard deviations of nitrite-plus-nitrate and orthophosphate in stream samples, were significant only because of a single large value early in the time period. Neither of these trends indicates an overall shift in variability. The other four significant trends were for both standard deviation and RSD of total Kjeldahl nitrogen and total phosphorus in streams. These trends are associated with real changes during the time period.

The laboratory method used for analysis of total phosphorus changed on January 1, 1999 (U.S. Geological Survey National Water Quality Laboratory, 1999). The trend in the variability of reported total phosphorus concentrations is entirely a result of this change. There are no trends in either standard deviation or RSD during 1992–98 or during 1999–2001, but there is a decrease in both these measures of variability after the change in analytic method. These measures are listed in table 5 for the time periods before and after the method change and for the overall time period for comparison.

For total Kjeldahl nitrogen, there was no change in analytic method during 1992–2001, but the number of significant decimal places in the reported concentrations was increased in April 1997. The effect of this change is shown in figure 7. Before the change, the magnitude of differences between pairs of replicate concentrations was limited to a small number of values. For example, the difference between two rounded concentrations could be 0.0 or 0.1, but not 0.005. This limitation produces the linear patterns in standard deviations before 1997 in figure 7. After 1997, rounding is less severe and the linear pattern disappears. This change is enough to create a significant trend over the entire time period, but there are no trends in variability for the individual time periods before or after the change in rounding.

If total Kjeldahl nitrogen replicates are restricted to those after the change in rounding, the division between the low-

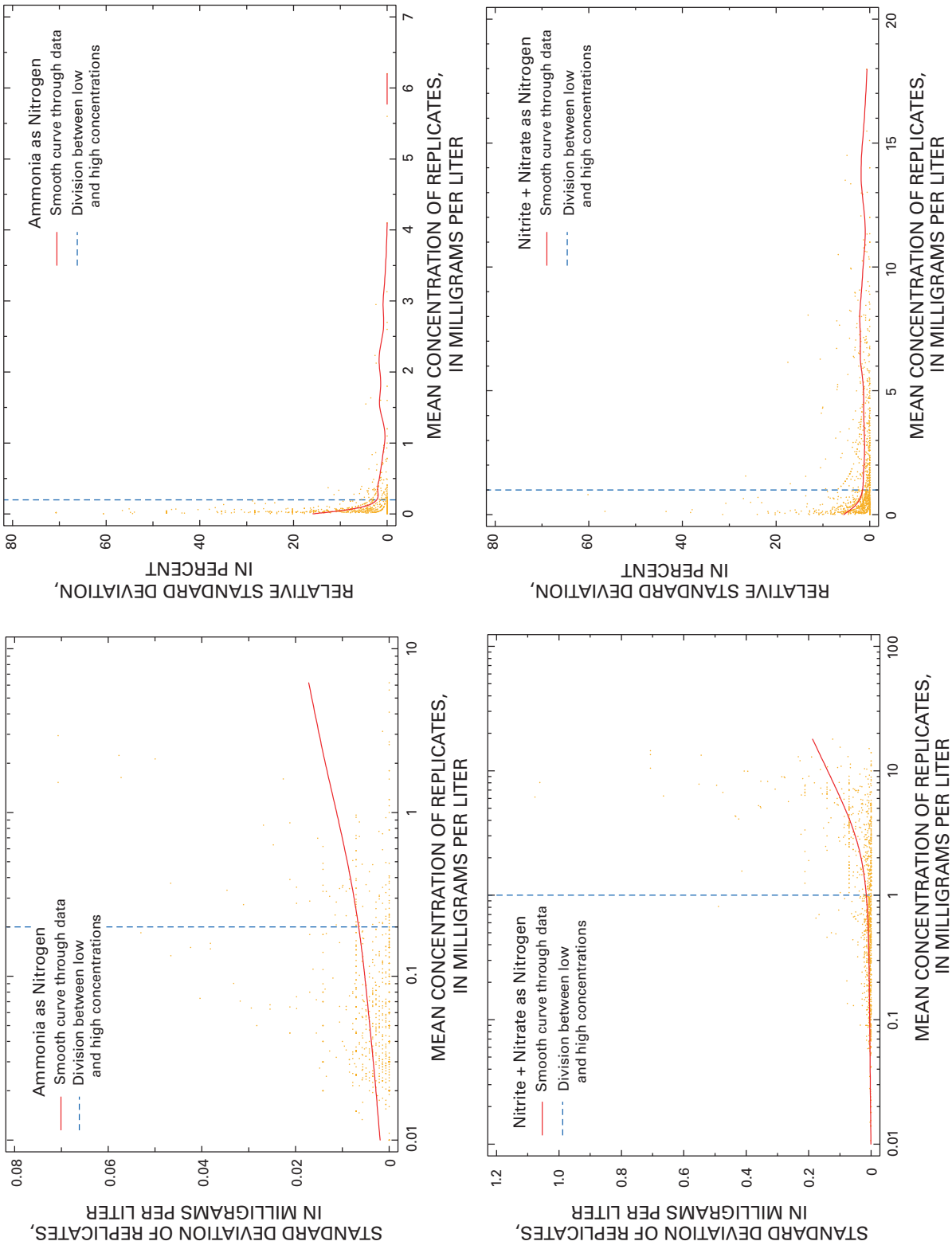

Ammonia as Nitrogen
Smooth curve through data
Division between low and high concentrations
STANDARD DEVIATION OF REPLICATES, IN MILLIGRAMS PER LITER
MEAN CONCENTRATION OF REPLICATES, IN MILLIGRAMS PER LITER
RELATIVE STANDARD DEVIATION, IN PERCENT
Nitrite + Nitrate as Nitrogen
Smooth curve through data
Division between low and high concentrations
STANDARD DEVIATION OF REPLICATES, IN MILLIGRAMS PER LITER
MEAN CONCENTRATION OF REPLICATES, IN MILLIGRAMS PER LITER
RELATIVE STANDARD DEVIATION, IN PERCENT

Figure 5. Gra

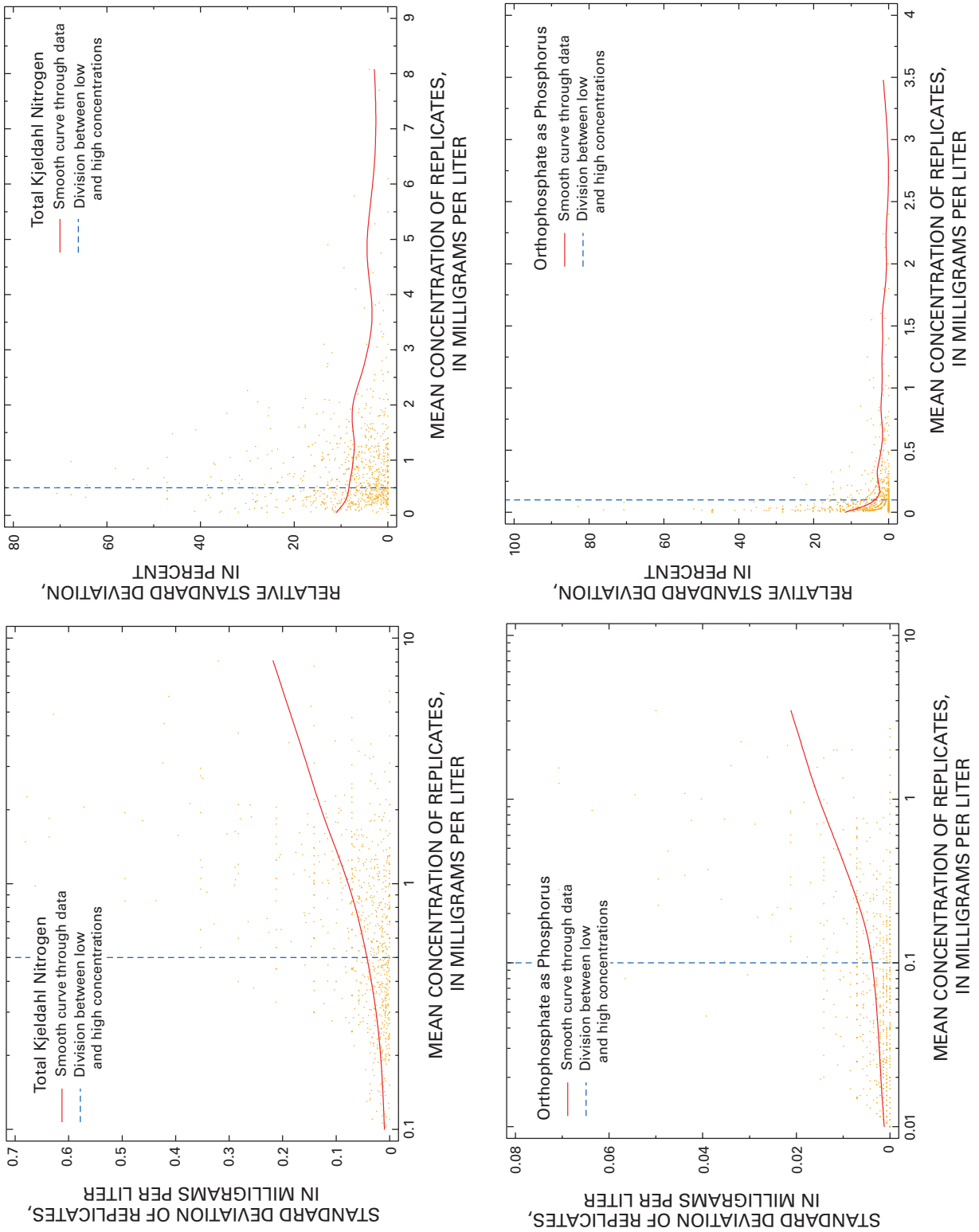


Figure 5. Graphs used t

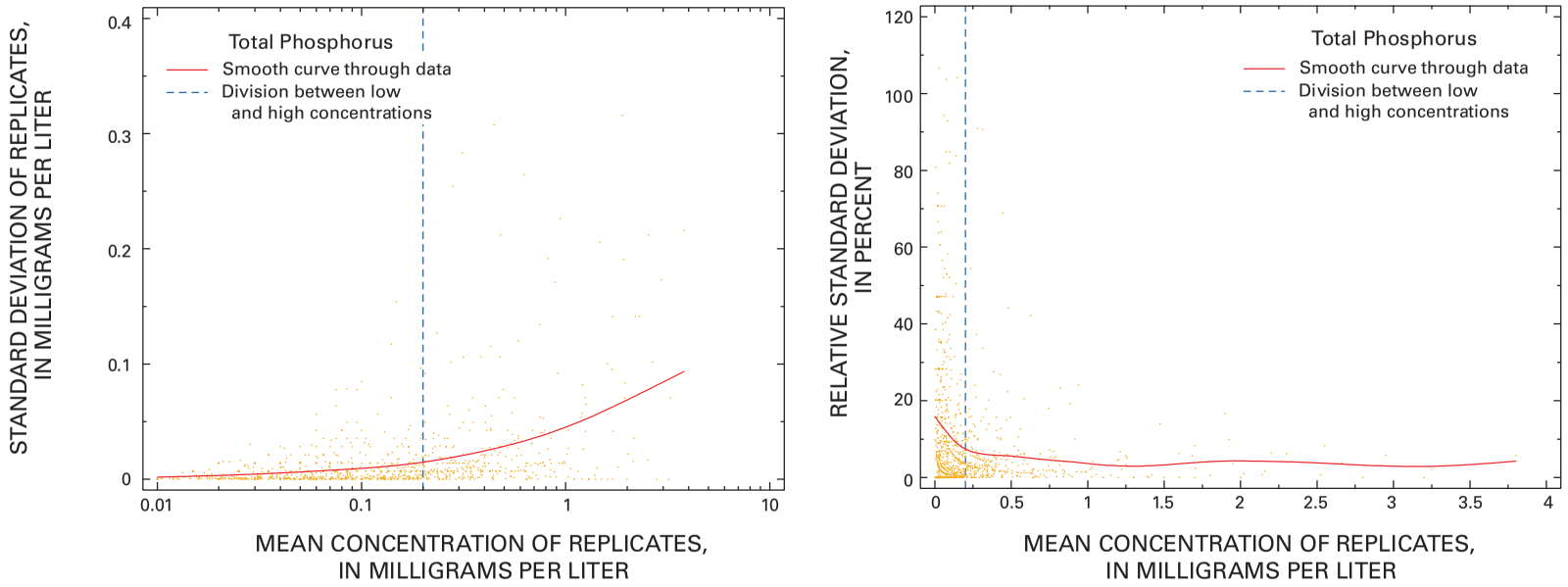


Figure 5. Graphs used t

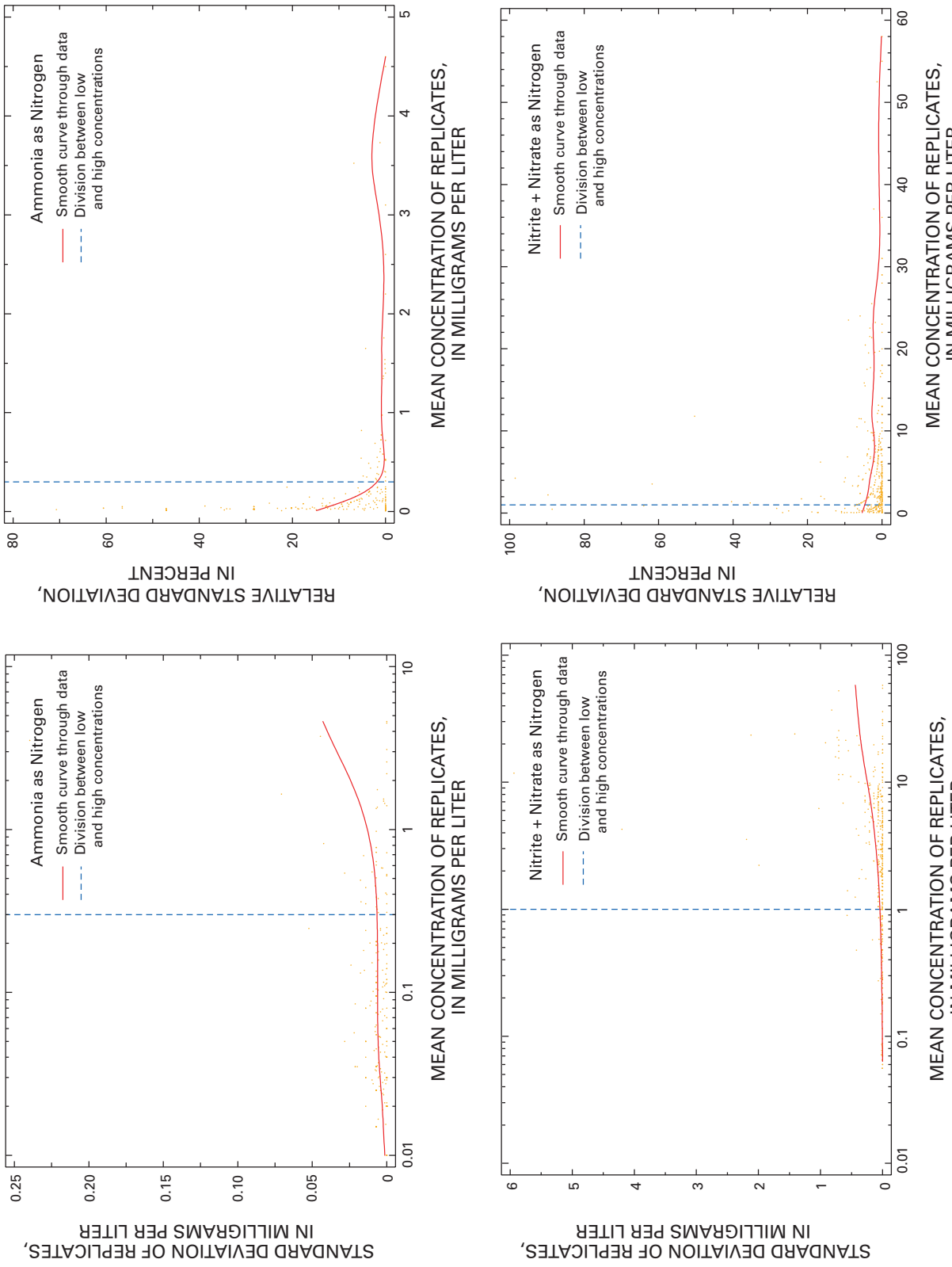


Figure 6. Graphs us

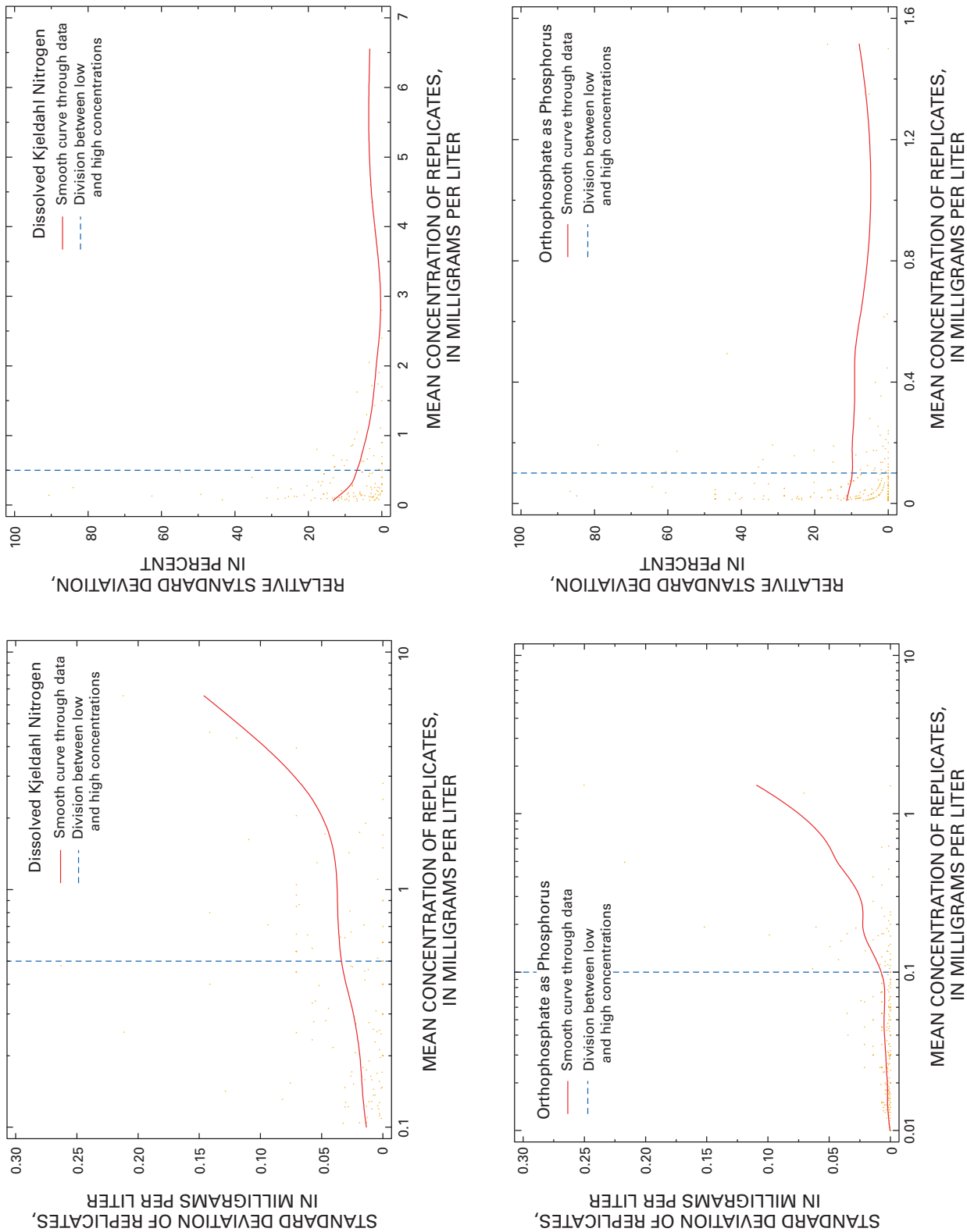


Figure 6. G
samples.—Continued

Table 4. Estimates of variability for nutrient analytes in stream and ground-water samples.

[mg/L, milligrams per liter; N, nitrogen; P, phosphorus]

Nutrient analyte	Low Concentrations			High Concentrations		
	Range (mg/L)	Number of replicate sets	Variability (standard deviation, in mg/L)	Range (mg/L)	Number of replicate sets	Variability (relative standard deviation, in percent)
Streams						
Ammonia, as N	0.003–0.2	800	0.0045	0.2–6.2	126	1.9
Nitrite + Nitrate, as N	0.004–1.0	691	0.012	1.0–41	541	2.2
Total Kjeldahl Nitrogen	0.05–0.5	523	0.027	0.5–40	591	7.6
Orthophosphate, as P	0.001–0.1	650	0.0027	0.1–3.5	306	2.8
Total Phosphorus	0.002–0.2	810	0.0072	0.2–25	338	6.5
Ground water						
Ammonia, as N	0.02–0.3	238	0.0047	0.3–23	48	1.3
Nitrite + Nitrate, as N	0.05–1.0	99	0.043	1.0–58	289	2.9
Dissolved Kjeldahl Nitrogen	0.2–0.5	62	0.022	0.5–29	55	7.8
Orthophosphate, as P	0.01–0.1	240	0.0039	0.1–2.4	52	10

Table 5. Changes in the variability of total Kjeldahl nitrogen and total phosphorus in stream samples during water years 1992–2001.

[mg/L, milligrams per liter]

Nutrient analyte	Low Concentrations				High Concentrations		
	Time period (water years)	Range (mg/L)	Number of replicate sets	Variability (standard deviation, in mg/L)	Range (mg/L)	Number of replicate sets	Variability (relative standard deviation, in percent)
Total Kjeldahl Nitrogen	1992–2001	0.05–0.5	523	0.027	0.5–40	591	7.6
	1992–1996	0.10–1.0	346	0.045	1.0–23	119	8.3
	1997–2001	0.05–1.0	476	0.034	1.0–40	162	6.1
Total Phosphorus	1992–2001	0.002–0.2	810	0.0072	0.2–25	338	6.5
	1992–1998	0.004–0.2	529	0.0094	0.2–4	207	8.1
	1999–2001	0.002–0.2	281	0.0032	0.2–25	131	4.0

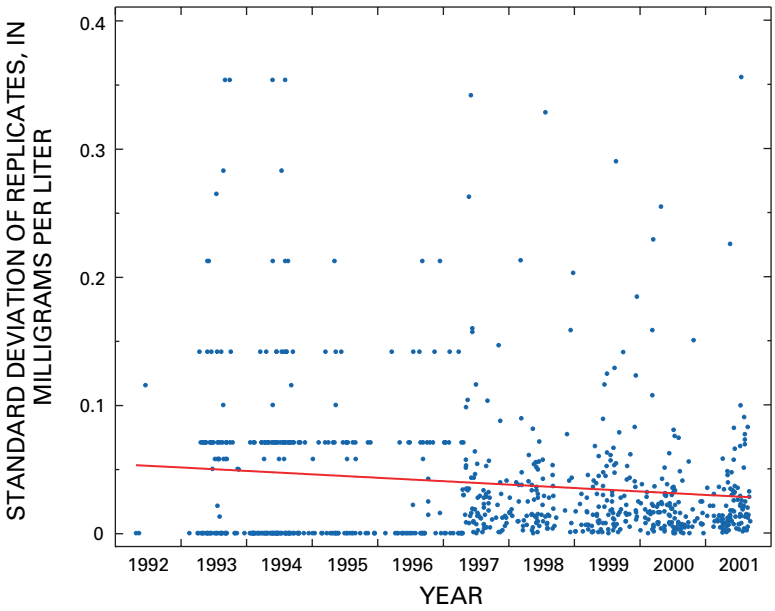


Figure 7. Time series plot showing a linear trend in the standard deviation of total Kjeldahl nitrogen replicates that had a mean concentration less than 0.5 mg/L.

and high-concentration ranges is better defined than it was using data for the entire time period. Plots of standard deviation and RSD in relation to mean concentration for the April 1997 through September 2001 replicates are shown in figure 8. Not only is the division (at 1.0 mg/L) more distinct, but also it is at a slightly different concentration than was selected for the entire time period (0.5 mg/L, fig. 5). Variability measures for total Kjeldahl nitrogen, computed using this division at 1.0 mg/L, are listed in table 5 for the time periods before and after the change in rounding. For comparison, the table includes variability measures previously computed for the entire time period, using a division at 0.5 mg/L. Variability over both concentration ranges is smaller for samples collected during 1997–2001. Because the analytic method was not changed, it seems likely that if more significant figures had been reported during the earlier time period (1992–97), variabilities would have been similar to the 1997–2001 values. Thus, variabilities computed for the 1997–2001 replicates are considered appropriate for application to environmental data collected during the entire time period.

In addition to temporal trends, sediment concentration was tested as a factor that might be related to variability of total Kjeldahl nitrogen and total phosphorus, constituents that are analyzed from unfiltered samples. Particulates, such as sediment, in samples can complicate the process of replication; however, no statistically significant relations were found between sediment concentration and the standard deviation or RSD for either of these constituents.

Confidence Intervals

Confidence intervals around measured concentrations can be calculated for any nutrient analyte as the product of sampling variabilities in tables 4 or 5 and an appropriate statistic (Z) from a table of normal deviates. For a 95-percent confidence interval, $\alpha = 0.05$ and $Z_{(1-\sigma/2)} = 1.96$. If ammonia in a stream sample is measured as 0.15 mg/L, the estimated sampling variability from table 4 is 0.0045 mg/L. A 95-percent confidence interval for the true concentration, based on this measurement, can be determined using equation 3:

$$[C_L, C_U] = 0.15 \pm 1.96(0.0045) \quad (5)$$

The second term of this equation, which is approximately 0.0088 mg/L (±5.9 percent of the measured concentration), represents the inherent error of the measurement.

For a measured ammonia concentration in the high range, such as 0.3 mg/L, sampling variability is 1.9 percent, and the 95-percent confidence interval is:

$$[C_L, C_U] = 0.3 \pm 1.96\left(0.3\,\frac{1.9}{100}\right) \quad (6)$$

Here, the inherent error is 0.011 mg/L, or 3.7 percent.

If 0.15 mg/L was the mean concentration of ammonia in 10 stream samples, the 95-percent confidence interval, based solely on sampling variability, is determined using equation 4:

$$[C_L, C_U] = 0.15 \pm 1.96\,\frac{0.0045}{\sqrt{10}} \quad (7)$$

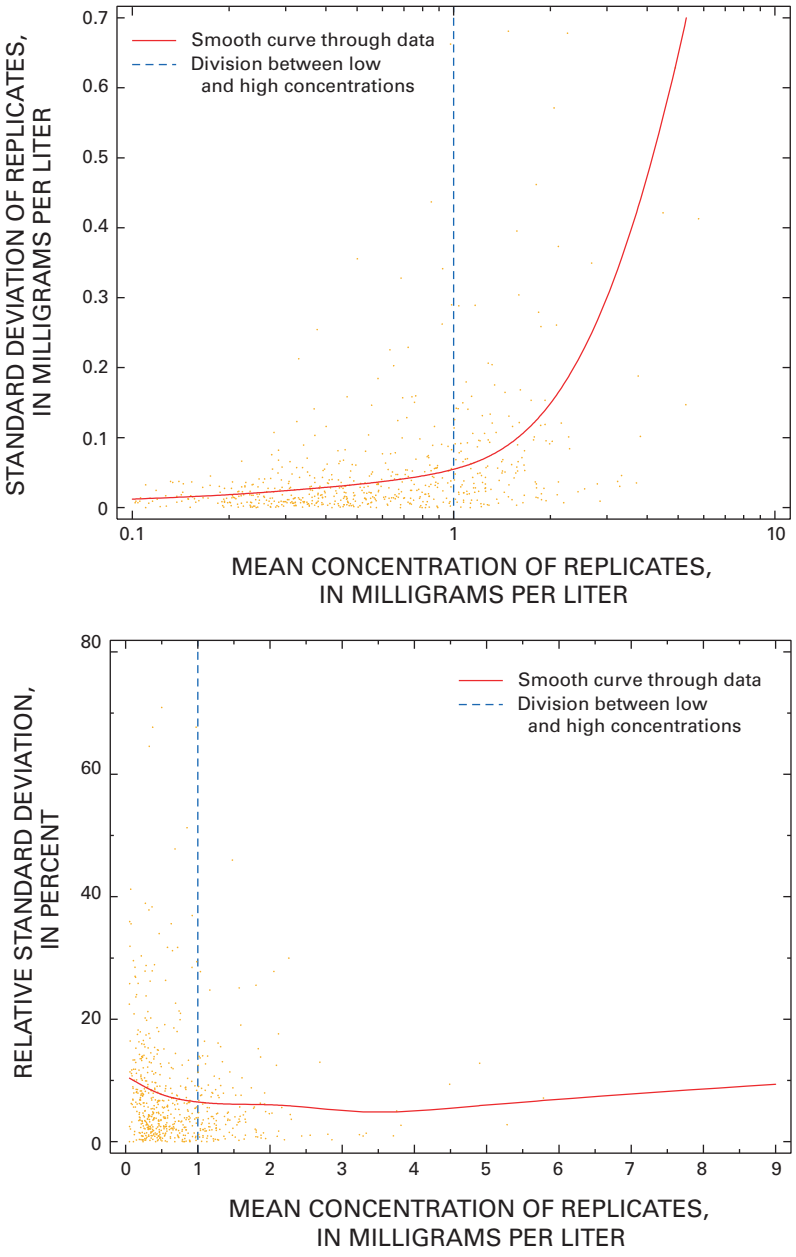


Figure 8. Graphs showing low and high ranges of replicate concentrations used to determine variability of total Kjeldahl nitrogen in stream samples collected during April 1997 through September 2001.

In this example, the inherent error is estimated to be 0.0028 mg/L, which is about ±1.9 percent of the mean concentration. This is the unavoidable measurement error expected if ammonia concentrations were truly identical in all 10 samples. Other factors, such as environmental variability among samples, would likely increase the actual standard deviation. Thus, sampling variability for a mean represents the expected lower limit of overall variability.

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 instances, 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 UCLs for the 95th percentile of contamination determined from NAWQA field blanks (table 3), maximum affected concentrations were calculated from selected nutrient analytes and are listed in table 6 in addition to the 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 affected 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, 1999). At the low end of the range for this criterion (0.18 mg/L), contamination greater than the 95th percentile value of 0.04 mg/L (table 3) could account for more than 22 percent of the measured ammonia concentration. However, in the range of pH and temperature for most streams, the criterion generally is greater than 0.4 mg/L. Based on the average pH (7.7) and temperature (14°C) for all NAWQA sampling sites during water years 1992–2001, the most restrictive ammonia criterion would be approximately 3.6 mg/L, which is 9 times larger than the maximum contamination expected in at least 95 percent of all samples. Thus, although contamination could limit the use of ammonia data in comparison to criteria under some conditions, the majority of measurements that exceed a criterion probably would not be adversely affected.

Contamination also might affect interpretation of phosphorus data. The maximum affected concentrations listed in table 6 exceed USEPA recommendations for protection of streams (U.S. Environmental Protection Agency, 1986) and national average background concentrations in streams (Mueller and others, 1995). Unqualified determination of adverse environmental effects due to phosphorus is limited to sites where measured concentrations are greater than about 0.2 mg/L for orthophosphate and 0.3 mg/L for total phosphorus. At the limit to avoid eutrophication (0.05 mg/L), contamination greater than the 95th percentile value of 0.019 mg/L (table 3) could account for more than 38 percent of the measured orthophosphate concentration.

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 0.71 mg/L (table 6). 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 most common detection limit (table 3); therefore, any measured concentration in excess of 10 times the detection 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:

- It determines the potential error in an individual measurement; and
- 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 using estimates of sampling variability. Intervals around selected critical concentrations of nutrient analytes are listed in table 7. The sampling variabilities used to construct these intervals were estimated by using the replicate-analysis results in table 4.

Effects on Measured Concentrations

The first measured value listed for ammonia in table 7 is 0.1 mg/L, which is the background concentration in streams estimated by Mueller and others (1995). At this concentration, the sampling variability is estimated to be approximately 0.0045 mg/L. The range of the 95-percent confidence interval for an individual measurement of 0.1 mg/L is 0.091 to 0.109 mg/L. The potential error in that measurement due to sampling variability is 0.009 mg/L, which is a relative error of ±9 percent of the measured concentration. Although this relative error might be considered somewhat large, the absolute error is small. For a measured ammonia concentration of

Table 6. Maximum nutrient concentrations that are considered potentially affected by contamination, and selected critical values used to interpret environmental data.

[mg/L, milligrams per liter; N, nitrogen; P, phosphorus; <, less than; --, not applicable]

Nutrient analyte	Maximum affected concentration[1] (mg/L)	Critical value	
		Description	Concentration (mg/L)
Streams			
Ammonia, as N	0.4	Background[2]	0.1
		Aquatic-life criterion[3]	0.18–6.7
		Median downstream from urban areas[2]	0.2
Nitrite + Nitrate, as N	0.7	Background[2]	0.6
		Drinking-water standard[4]	10
Total Kjeldahl Nitrogen	< 2.0		--
Orthophosphate, as P	0.19	Recommended to avoid eutrophication[4]	0.05
Total Phosphorus	0.3	Background[2]	0.1
		Recommended to avoid eutrophication[4]	0.1
Ground water			
Ammonia, as N	0.44		--
Nitrite + Nitrate, as N	0.9	Background[5]	1.1
		Drinking-water standard[4]	10
Dissolved Kjeldahl Nitrogen	< 2.0		--
Orthophosphate, as P	0.2		--

[1] Estimated at 10 times the 99-percent upper confidence limit for the 95th percentile, listed in table 3.

[2] Mueller and others, 1995.

[3] Criterion varies depending on water temperature and pH (U.S. Environmental Protection Agency, 1999).

[4] U.S. Environmental Protection Agency, 1986.

[5] Nolan and Hitt, 2003.

Table 7. 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; N, nitrogen; P, phosphorus]

Nutrient analyte	Measured value[1] (mg/L)	Time period (water years)	Estimated sampling variability[2] (mg/L)	95-percent confidence interval (mg/L)	
				Individual measurements	Mean of 10 measurements
Streams					
Ammonia, as N	0.1	1992–2001	0.0045	0.091–0.109	0.097–0.103
	0.18	1992–2001	0.0045	0.171–0.189	0.177–0.183
	6.7	1992–2001	0.13	6.45–6.95	6.62–6.78
Nitrite + Nitrate, as N	0.6	1992–2001	0.012	0.576–0.624	0.593–0.607
	10	1992–2001	0.22	9.56–10.4	9.86–10.1
Total Kjeldahl Nitrogen[3]	6.7	1997–2001	0.41	5.90–7.50	6.45–6.95
Orthophosphate, as P	0.05	1992–2001	0.0027	0.045–0.055	0.048–0.052
	0.1	1992–2001	0.0028	0.095–0.105	0.098–0.102
Total Phosphorus	0.1	1992–1998	0.0094	0.082–0.118	0.094–0.106
	0.1	1999–2001	0.0032	0.094–0.106	0.098–0.102
Ground water					
Ammonia, as N[3]	0.2	1992–2001	0.0047	0.191–0.209	0.197–0.203
Nitrite + Nitrate, as N	1.1	1992–2001	0.03	1.04–1.16	1.08–1.12
	10	1992–2001	0.29	9.43–10.6	9.82–10.2
Dissolved Kjeldahl Nitrogen[3]	0.2	1992–2001	0.022	0.157–0.243	0.186–0.214
Orthophosphate, as P[3]	0.05	1992–2001	00039	0.042–0.058	0.048–0.052

[1] Based on "critical values" identified in table 6.

[2] From tables 4 and 5. For concentrations in the high range, sampling variability is: C (RSD/100).

[3] No critical values were identified in table 6 for total or dissolved Kjeldahl nitrogen, or for ammonia or orthophosphate in ground water. The values used here are for comparison to other nutrient analytes.

6.7 mg/L, which is the maximum (pH and temperature dependent) criterion for protection of aquatic life (U.S. Environmental Protection Agency, 1999), the absolute error is larger (±0.25 mg/L), but the relative error is smaller (0.25 divided by 6.7, or 4 percent). 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 0.25 mg/L or 9 percent of the measurement, whichever is smaller. In most circumstances, variability in this range has little effect on interpretation of ammonia data.

Variabilities for Kjeldahl nitrogen, which includes ammonia and organic nitrogen, are higher than for ammonia alone. This might occur because of larger errors inherent in the Kjeldahl analysis than in the method used to analyze ammonia. For a measured concentration of 6.7 mg/L, the relative error, with 95-percent confidence, is 4 percent (0.25 divided by 6.7) for ammonia and 12 percent (0.8 divided by 6.7) for total Kjeldahl nitrogen. A similar pattern can be seen in the phosphorus results. At the same measured concentrations, sampling variability is greater for total phosphorus than for orthophosphate, particularly for samples analyzed during 1992–98. The difference in analytic methods for these phosphorus analytes is similar to those for ammonia and Kjeldahl nitrogen, so method errors might be the primary cause of these differences in sampling variability.

In addition to total Kjeldahl nitrogen, variability of individual measurements exceeds 10 percent for dissolved Kjeldahl nitrogen in ground-water samples and total phosphorus in surface-water samples collected before 1999. Relative variability is also high for orthophosphate in surface-water and ground-water samples (about 11 to 15 percent); however, absolute variability is small (0.005 to 0.008 mg/L), and has little effect on interpretation of orthophosphate at environmentally significant concentrations. Relative variability in a mean of 10 measurements is low (1 to 7 percent) for all nutrients.

When individual measurements are compared to standards, sampling variability must be considered in the determination of exceedance or compliance. At the highest aquatic-life criterion for ammonia (6.7 mg/L), the 95-percent confidence interval is 0.25 mg/L. Therefore, measured concentrations as high as 6.95 mg/L do not indicate exceedance of the criterion, with 95-percent confidence. Similarly, measured concentrations as low as 6.45 mg/L do not necessarily indicate compliance. For nitrite-plus-nitrate measurements at the drinking-water standard (10 mg/L), the 95-percent confidence interval is approximately 0.4 mg/L for stream samples. If laboratory results are rounded to two significant figures, a reported concentration of at least 11 mg/L would indicate an exceedance of the standard with 95-percent confidence. In this instance, the uncertainty caused by sampling variability has no real effect, because it does not change the least significant figure of the rounded value.

Effects on Comparisons between Concentrations

Measurement errors that result from sampling variability can affect identification of significant differences between two measurements (or means). If a true difference falls within the inherent sampling variability, it is not likely to be distinguished. 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, the measured difference is not statistically distinguishable. Consider, for example, an individual measurement of 0.10 mg/L for total phosphorus from a stream sample collected during 1992–98. The 95-percent confidence interval is 0.082 to 0.118 mg/L (table 7). For another sample, collected during the same time period, with a measured concentration of 0.13 mg/L, the 95-percent confidence interval is 0.112 to 0.148 mg/L, computed using equation 3 with sampling variabilities from table 5 and Z=1.96. Because these intervals overlap, the two measurements cannot be considered significantly different. For a measurement of 0.14 mg/L, the confidence interval is 0.122 to 0.158 mg/L. Because this interval does not overlap with the interval for a measurement of 0.10 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 eq. 4). 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 7). Using the same example of total phosphorus analyzed during 1992–98, the confidence interval for a mean concentration of 0.10 mg/L is 0.094 to 0.106 mg/L. The 95-percent confidence interval for a mean concentration of 0.12 mg/L from 10 samples is 0.114 to 0.126 mg/L. Because these intervals do not overlap, the difference between the two means is statistically significant. An actual set of 10 samples also might be affected by environmental variability, so such small differences might not be identifiable as statistically significant. However, sampling variability is unlikely to interfere with identification of real differences in total phosphorus concentration of at least 0.02 mg/L.

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 analyzed during 1992–98, the potential error for an individual measurement of 0.10 mg/L is ±0.018 mg/L, based on the 95-percent confidence limit in table 7. Therefore, the minimum difference unaffected by this error is approximately 0.036 mg/L. Measurements within this minimum difference will not be significantly different. As previously shown, a measured value of 0.13 mg/L is not, but a measured value of 0.14 mg/L is, significantly greater than a measured value of 0.10 mg/L. Likewise, a measured value of 0.07 mg/L is not, but a measured value of 0.06 mg/L is, significantly less than a measured value of 0.10 mg/L.

The minimum differences unaffected by sampling variability for critical concentrations of most nutrient analytes are

small. For orthophosphate, total phosphorus sampled after 1998, and most concentrations of ammonia listed in table 7, differences of 0.02 mg/L would be considered significant for most individual measurements, and differences greater than 0.006 mg/L between means of 10 measurements would likely be unaffected by sampling variability. For the highest critical value for ammonia (6.7 mg/L), differences in individual measurements must exceed 0.5 mg/L to be considered significant. At this same concentration, a difference in individual measurements of total Kjeldahl nitrogen would need to be more than 3 times as large, at least 1.6 mg/L, to be considered significant.

Conclusions

Contamination bias, based on the upper 95th percentile of measurements in field blanks, potentially affects measured concentrations of less than 0.4 mg/L for ammonia, less than 0.2 mg/L for orthophosphate, and less than 0.3 mg/L for total phosphorus. These ranges include some environmentally significant concentrations, based on established water-quality criteria and standards. Nitrite-plus-nitrate at environmentally significant concentrations is mostly unaffected by contamination, which was less than 0.1 mg/L in 95 percent of all field blanks. Similarly, contamination in 95 percent of all field blanks was less than detection (0.2 mg/L) for the Kjeldahl nitrogen analytes.

The effects of sampling variability on interpretation of standard exceedance or differences between individual measurements or mean concentrations are limited to only a small range of a few hundredths to a few tenths of a milligram per liter, depending on analyte and concentration. For environmentally significant concentrations of some nutrient analytes, sampling variability can be less than the resolution of reported data because of rounding. For this reason, sampling variability has little or no effect on reported concentrations of ammonia, nitrite-plus-nitrate, orthophosphate, or total phosphorus sampled after 1998. The potential errors due to sampling variability are greater for the Kjeldahl nitrogen analytes and for total phosphorus sampled before 1999; uncertainty can be more than 10 percent of the individual measured values at environmentally significant concentrations. However, the uncertainty in a mean of 10 concentrations is within a small range (1 to 7 percent) for all nutrients. Differences in concentration within these ranges have limited, if any, environmental importance.

These results can be applied to interpretation of the environmental data collected during water years 1992–2001 in the 52 NAWQA study units. They also provide a basis for comparison with QC results for samples collected in subsequent years.

References

Anderson, R.L., 1987, Practical statistics for analytical chemists: New York, Van Norstrand Reinhold, 316 p.

Chambers, J.M., Cleveland, W.S., Kleiner, Beat, and Tukey, P.A., 1983, Graphical methods for data analysis: Boston, Mass., Duxbury Press, 395 p.

Childress, C.J.O., Foreman, W.T., Connor, B.F., and Maloney, T.J., 1999, New reporting procedures based on long-term method detection levels and some considerations for interpretations of water-quality data provided by the U.S. Geological Survey National Water Quality Laboratory: U.S. Geological Survey Open-File Report 99–193, 19 p.

Hahn, G.J., and Meeker, W.Q., 1991, Statistical intervals—A guide for practitioners: New York, John Wiley and Sons, 392 p.

Helsel, D.R., 2005, Nondetects and Data Analysis—Statistics for censored environmental data: New York, John Wiley and Sons, 250 p.

Mueller, D.K., 1998, Quality of nutrient data from streams and ground water sampled during 1993–95 —National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 98–276, 25 p.

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.

Nolan, B.T., and Hitt, K.J., 2003, Nutrients in shallow ground waters beneath relatively undeveloped areas in the conterminous United States: U.S. Geological Survey Water-Resources Investigations Report 02–4289, 17 p.

SAS Institute, 1990, SAS/GRAPH software: Version 8.1, SAS Institute Inc., Cary, NC.

Shelton, L.R., 1994, Field guide for collecting and processing stream-water samples for the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 94–455, 42 p.

Snedecor, G.W., and Cochran, W.G., 1980, Statistical methods, 7th ed.: Ames, Iowa, Iowa State University Press, 507 p.

U.S. Environmental Protection Agency, 1999, (1999 Update of ambient water quality criteria for ammonia: Washington, D.C., U.S. Environmental Protection Agency Report EPA–822–R–99–014, Office of Water, [variously paged].

U.S. Environmental Protection Agency, 1986, Quality criteria for water 1986: Washington, D.C., U.S. Environmental Protection Agency Report 440/5–86–001, Office of Water, [variously paged].

U.S. Geological Survey, variously dated, National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1–A9. Available online at: http://pubs.water.usgs.gov/twri9A

U.S. Geological Survey National Water Quality Laboratory, 1999, Validation of U.S. Environmental Protection Agency (USEPA) method 365.1 to replace U.S. Geological Survey (USGS) TWRI methods I–2607–90 and I–4607–90 for determination of phosphorus in the concentration range from 0.004 to 0.200 mg-P/L: National Water Quality Laboratory Technical Memorandum 99.05, accessed January 3, 2005, http://nwql.usgs.gov/Public/tech_memos/nwql.99-05.html

Glossary of Data-Quality Terms

Accuracy: The degree of agreement between a measured value and the true or expected value. Accuracy is affected by both bias and variability.

Bias: The systematic error inherent in a method; it can be either positive or negative.

Blank sample: A sample prepared from water that is free of the analyte(s) of interest for determining contamination.

Contamination bias: A positive bias due to the inadvertent introduction of analytes into water samples during sample collection, processing, shipment, or analysis.

Field blank: A blank sample that has been exposed in the field to all sampling equipment and conditions that normally are associated with the collection of an environmental sample.

Quality assessment: The overall process of assessing the quality of environmental data by reviewing the application of the quality-assurance elements and the analysis of the quality-control data.

Quality assurance (QA): Procedures used to control the nonquantifiable components of a project, such as sampling at the correct location with the proper equipment and using the appropriate methods.

Quality control (QC): Data generated to estimate the magnitude of the bias and variability in the process of obtaining environmental data.

Precision: The degree of mutual agreement among independent measurements from the repeated application of a measurement process under identical conditions. Precision is the inverse of variability.

Replicates: Two (duplicate) or more environmental samples collected and processed such that their compositions can be considered identical. Replicates are used to estimate sampling variability.

Sampling variability: The variability introduced into sample measurements because of field procedures (collection, processing, and shipment) plus laboratory analysis.

Source-solution blank: A sample of blank water taken directly from its source container without exposure to sampling equipment or conditions.

Variability: Random error in independent measurements as the result of repeated application of the measurement process under identical conditions.