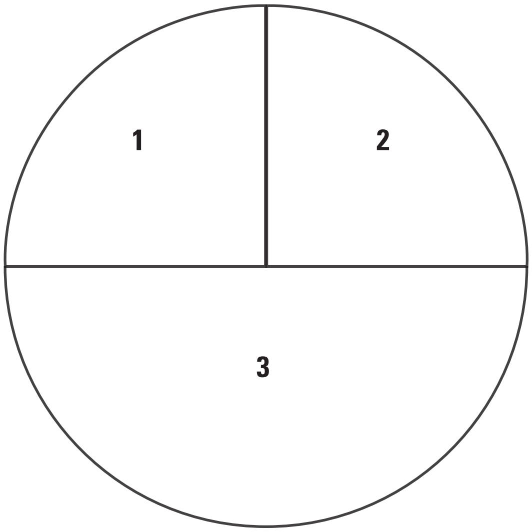
National Water-Quality Assessment Project of the U.S. Geological Survey's
National Water-Quality Program

Water-Quality Trends in the Nation's Rivers and Streams, 1972–2012—Data Preparation, Statistical Methods, and Trend Results

Scientific Investigations Report 2017–5006

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



Cover.

Background. Dismal River near Thedford, Nebraska (site 06775900), August 16, 2016. Photograph by Robert E. Zuellig, U.S. Geological Survey (USGS).

1. Water-quality sampling from the Colorado River near the Colorado-Utah State line. Photograph by Nancy Bauch, USGS.
2. Vallecito Creek near Bayfield, Colorado (site 09352900), August 25, 2016. Photograph by Robert E. Zuellig, USGS.
3. Riverweed darter (*Etheostoma podostemone*) found in the Roanoke River, Roanoke County, Virginia. Photograph by Robert E. Zuellig, USGS.

Water-Quality Trends in the Nation's Rivers and Streams, 1972–2012—Data Preparation, Statistical Methods, and Trend Results

By Gretchen P. Oelsner, Lori A. Sprague, Jennifer C. Murphy,
Robert E. Zuellig, Henry M. Johnson, Karen R. Ryberg, James A. Falcone,
Edward G. Stets, Aldo V. Vecchia, Melissa L. Riskin, Laura A. De Cicco,
Taylor J. Mills, and William H. Farmer

National Water-Quality Assessment Project of the U.S. Geological Survey's
National Water-Quality Program

Scientific Investigations Report 2017–5006

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

U.S. Department of the Interior
RYAN K. ZINKE, Secretary

U.S. Geological Survey
William H. Werkheiser, Acting Director

U.S. Geological Survey, Reston, Virginia: 2017

Suggested citation:
Oelsner, G.P., Sprague, L.A., Murphy, J.C., Zuellig, R.E., Johnson, H.M., Ryberg, K.R., Falcone, J.A., Stets, E.G., Vecchia, A.V., Riskin, M.L., De Cicco, L.A., Mills, T.J., and Farmer, W.H., 2017, Water-quality trends in the Nation's rivers and streams, 1972–2012—Data preparation, statistical methods, and trend results: U.S. Geological Survey Scientific Investigations Report 2017–5006, 136 p., https://doi.org/10.3133/sir20175006.

ISSN 2328-0328 (online)

Foreword

Sustaining the quality of the Nation’s water resources and the health of our diverse ecosystems depends on the availability of sound water-resources data and information to develop effective, science-based policies. Effective management of water resources also brings more certainty and efficiency to important economic sectors. Taken together, these actions lead to immediate and long-term economic, social, and environmental benefits that make a difference to the lives of the almost 400 million people projected to live in the United States by 2050. (http://water.usgs.gov/nawqa/applications/).

In 1991, Congress established the National Water-Quality Assessment (NAWQA) to address where, when, why, and how the Nation’s water quality has changed, or is likely to change in the future, in response to human activities and natural factors. Since then, NAWQA has been a leading source of scientific data and knowledge used by national, regional, State, and local agencies to develop science-based policies and management strategies to improve and protect water resources used for drinking water, recreation, irrigation, energy development, and ecosystem needs. Plans for the third decade of NAWQA (2013–23) address priority water-quality issues and science needs identified by NAWQA stakeholders, such as the Advisory Committee on Water Information, and the National Research Council as the Nation faces increasing challenges related to population growth, increasing needs for clean water, and changing land-use and weather patterns.

Federal, State, and local agencies have invested billions of dollars to reduce the amount of pollution entering rivers and streams that millions of Americans rely on for drinking water, recreation, and irrigation. Tracking changes in the quality of these waterways over multiple decades is crucial for evaluating the effectiveness of pollution control efforts and protecting the Nation’s water resources into the future. This report combines monitoring data collected by 74 organizations at almost 1,400 sites to provide a nationwide look at changes in the quality of our rivers and streams from 1972 to 2012—the 40 years since passage of the Clean Water Act. All NAWQA reports are available online (https://water.usgs.gov/nawqa/bib/).

We hope this publication will provide you with insights and information to meet your water resource needs and will foster increased citizen awareness and involvement in the protection and restoration of our Nation’s waters. The information in this report is intended primarily for those interested or involved in resource management and protection, conservation, regulation, and policymaking at the regional and national level.

Dr. Gary L. Rowe
Program Coordinator, National Water Quality Program
U.S. Geological Survey

Contents

Appendixes [available at https://doi.org/10.3133/sir20175006]

Figures

Tables

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
	Length	
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
yard (yd)	0.9144	meter (m)
	Area	
acre	4,047	square meter (m^2)
acre	0.004047	square kilometer (km^2)
square foot (ft^2)	0.09290	square meter (m^2)
square mile (mi^2)	2.590	square kilometer (km^2)
	Volume	
gallon (gal)	3.785	liter (L)
gallon (gal)	0.003785	cubic meter (m^3)
million gallons (Mgal)	3,785	cubic meter (m^3)
cubic inch (in^3)	16.39	cubic centimeter (cm^3)
cubic inch (in^3)	0.01639	liter (L)
cubic foot (ft^3)	0.02832	cubic meter (m^3)
cubic yard (yd^3)	0.7646	cubic meter (m^3)
acre–foot (acre–ft)	1,233	cubic meter (m^3)
	Flow rate	
acre–foot per day (acre–ft/d)	0.01427	cubic meter per second (m^3/s)
foot per second (ft/s)	0.3048	meter per second (m/s)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
gallon per minute (gal/min)	0.06309	liter per second (L/s)
gallon per day (gal/d)	0.003785	cubic meter per day (m^3/d)
million gallons per day (Mgal/d)	0.04381	cubic meter per second (m^3/s)
	Mass	
ounce, avoirdupois (oz)	28.35	gram (g)
pound, avoirdupois (lb)	0.4536	kilogram (kg)
ton per day (ton/d)	0.9072	metric ton (t)
ton per year (ton/yr)	0.9072	metric ton (t)

International System of Units to U.S. customary units

Multiply	By	To obtain
	Length	
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
meter (m)	1.094	yard (yd)
	Area	
square meter (m2)	0.0002471	acre
square kilometer (km2)	247.1	acre
square meter (m2)	10.76	square foot (ft2)
square kilometer (km2)	0.3861	square mile (mi2)
	Volume	
liter (L)	0.2642	gallon (gal)
cubic meter (m3)	264.2	gallon (gal)
cubic meter (m3)	0.0002642	million gallons (Mgal)
cubic meter (m3)	35.31	cubic foot (ft3)
cubic meter (m3)	1.308	cubic yard (yd3)
cubic kilometer (km3)	0.2399	cubic mile (mi3)
cubic meter (m3)	0.0008107	acre–foot (acre–ft)
	Flow rate	
cubic meter per second (m3/s)	70.07	acre–foot per day (acre–ft/d)
meter per second (m/s)	3.281	foot per second (ft/s)
cubic meter per second (m3/s)	35.31	cubic foot per second (ft3/s)
liter per second (L/s)	15.85	gallon per minute (gal/min)
cubic meter per day (m3/d)	264.2	gallon per day (gal/d)
millimeter per year (mm/yr)	0.03937	inch per year (in/yr)
	Mass	
gram (g)	0.03527	ounce, avoirdupois (oz)
kilogram (kg)	2.205	pound avoirdupois (lb)
metric ton per day	1.102	ton per day (ton/d)
metric ton per year	1.102	ton per year (ton/yr)

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

$$°F = (1.8 \times °C) + 32.$$

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

$$°C = (°F - 32) / 1.8$$

Temperature in degrees Celsius (°C) may be converted to degrees Kelvin (K) as follows:

$$K = (°C) + 273.15$$

Temperature in degrees Kelvin (K) may be converted to degrees Celsius (°C) as follows:

$$°C = K - 273.15$$

Datum

Vertical coordinate information is referenced to the North American Vertical Datum of 1988 (NAVD 88).

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).

Supplemental Information

Specific conductance is given in microsiemens per centimeter at 25 degrees Celsius (µS/cm at 25 °C).

Concentrations of chemical constituents in water are given either in milligrams per liter (mg/L) or micrograms per liter (µg/L).

A water year is the 12–month period October 1 through September 30 designated by the calendar year in which it ends.

Abbreviations

BioData	Aquatic Bioassessment database
CBP	Chesapeake Bay Program
EGRET	Exploration and Graphics for RivEr Trends
EPA	U.S. Environmental Protection Agency
EPT	Ephemeroptera, Plecoptera, and Trichoptera
GCMS	gas chromatography/mass spectrometry
LRL	laboratory reporting level
LT–MDL	long–term method detection level
maxLT–MDL	maximum long–term method detection level
MDL	method detection level
MMI	multimetric type indices
MRL	method reporting level
NASQAN	National Stream Quality Accounting Network
NAWQA	National Water–Quality Assessment
NHD	National Hydrography Dataset
NPDES	National Pollutant Discharge Elimination System
NWIS	National Water Information System database
NWQL	National Water–Quality Laboratory
O/E	observed/expected
OWQ	USGS Office of Water Quality
QAQC	quality assurance quality control
SC	specific conductance
SPARROW	SPAtially Referenced Regressions on Watershed attributes
SRP	soluble reactive phosphorus
SSC	suspended sediment concentration
STORET	STORage and RETrieval Data Warehouse
TDS	total dissolved solids
TIVs	tolerance indicator values
TSS	total suspended solids
USGS	U.S. Geological Survey
WQP	Water Quality Portal
WRTDS	Weighted Regressions on Time, Discharge, and Season

Water-Quality Trends in the Nation’s Rivers and Streams, 1972–2012—Data Preparation, Statistical Methods, and Trend Results

By Gretchen P. Oelsner, Lori A. Sprague, Jennifer C. Murphy, Robert E. Zuellig, Henry M. Johnson, Karen R. Ryberg, James A. Falcone, Edward G. Stets, Aldo V. Vecchia, Melissa L. Riskin, Laura A. De Cicco, Taylor J. Mills, and William H. Farmer

Abstract

Since passage of the Clean Water Act in 1972, Federal, State, and local governments have invested billions of dollars to reduce pollution entering rivers and streams. To understand the return on these investments and to effectively manage and protect the Nation’s water resources in the future, we need to know how and why water quality has been changing over time. As part of the National Water-Quality Assessment Project, of the U.S. Geological Survey’s National Water-Quality Program, data from the U.S. Geological Survey, along with multiple other Federal, State, Tribal, regional, and local agencies, have been used to support the most comprehensive assessment conducted to date of surface-water-quality trends in the United States. This report documents the methods used to determine trends in water quality and ecology because these methods are vital to ensuring the quality of the results. Specific objectives are to document (1) the data compilation and processing steps used to identify river and stream sites throughout the Nation suitable for water-quality, pesticide, and ecology trend analysis, (2) the statistical methods used to determine trends in target parameters, (3) considerations for water-quality, pesticide, and ecology data and streamflow data when modeling trends, (4) sensitivity analyses for selecting data and interpreting trend results with the Weighted Regressions on Time, Discharge, and Season method, and (5) the final trend results at each site. The scope of this study includes trends in water-quality concentrations and loads (nutrient, sediment, major ion, salinity, and carbon), pesticide concentrations and loads, and metrics for aquatic ecology (fish, invertebrates, and algae) for four time periods: (1) 1972–2012, (2) 1982–2012, (3) 1992–2012, and (4) 2002–12. In total, nearly 12,000 trends in concentration, load, and ecology metrics were evaluated in this study; there were 11,893 combinations of sites, parameters, and trend periods. The final trend results are presented with examples of how to interpret the results from each trend model. Interpretation of the trend results, such as causal analysis, is not included.

Introduction

The Nation’s rivers and streams are a valuable resource, providing drinking water for a growing population, irrigation for crops, habitat for aquatic life, and many recreational opportunities; however, pollution from urban and agricultural areas poses a continuing threat to water quality. Since passage of the Clean Water Act in 1972, Federal, State, Tribal, and local governments have invested billions of dollars to reduce pollution entering rivers and streams (U.S. Environmental Protection Agency, 2010; U.S. Department of Agriculture, 2014), yet recently, the U.S. Environmental Protection Agency (EPA) reported that nearly half of the Nation’s streams have poor biological quality (U.S. Environmental Protection Agency, 2016). To understand the return on these investments and to more effectively manage and protect the Nation’s water resources in the future, we need to know how and why water quality has been changing over time.

In 1991, the U.S. Geological Survey (USGS) began a study of more than 50 major river basins across the Nation as part of the National Water-Quality Assessment (NAWQA) Project of the USGS’s National Water-Quality Program. A major goal of NAWQA is to determine how water quality changes over time. To support that goal, long-term consistent and comparable monitoring has been conducted on rivers and streams throughout the Nation. Other USGS projects, as well as projects conducted by many other Federal, State, Tribal, regional, and local agencies have collected long-term water-quality data to support their own assessments of changing water-quality conditions. For the first time, in an effort meant to be as inclusive as possible, data from multiple sources have been aggregated, screened, standardized, and used to support a comprehensive assessment of surface-water-quality trends in the United States. Collectively, these trend results will be used to provide insight into how natural variation and human activities have contributed to water-quality changes over time in the Nation’s rivers and streams.

Objectives and Scope

The objectives of this report are to document (1) the data compilation and processing steps used to identify river and stream sites throughout the Nation suitable for water-quality, pesticide, and ecology trend analysis, (2) the statistical methods used to determine trends in target parameters, (3) considerations for water-quality, pesticide, and ecology data and streamflow data when modeling trends, (4) sensitivity analyses for selecting data and interpreting trend results with the Weighted Regressions on Time, Discharge, and Season method, and (5) a brief summary of the trend results presented as the presence or absence of a significant trend at each site for each parameter and trend period for which data are available. This report is intended to be used as a reference document for the multiple aspects of trend analysis described in the objectives above. The report is organized so that the data processing steps are detailed for each parameter group because most readers will not be interested in all of the parameters included in this study. The scope of this study includes trends in water-quality concentrations and loads (nutrient, sediment, major ion, salinity, and carbon), pesticide concentrations and loads, and metrics for aquatic ecology (fish, invertebrates, and algae) for four time periods: (1) 1972–2012, (2) 1982–2012, (3) 1992–2012, and (4) 2002–12. Pesticide trends were determined based on calendar year, and all other trends were determined based on water year, which is defined as the 12-month period beginning October 1 for any given year through September 30 of the following year and is designated by the calendar year in which it ends (see "Trend Analysis Methods" section). This report documents the methods used to determine trends in water quality and ecology because they are vital to ensuring the quality and interpretability of the results. The final trend results are presented, along with an example of how to interpret the results from each trend model. An overall or in-depth interpretation of the trend results, such as causal analysis, is not included in this report.

Five datasets (including all trend model inputs and results) were used to support this study:

1. De Cicco, L.A., Sprague, L.A., Murphy, J.C., Riskin, M.L., Falcone, J.A., Stets, E.G., Oelsner, G.P., and Johnson, H.M., 2017, Water-quality and streamflow datasets used in the Weighted Regressions on Time, Discharge, and Season (WRTDS) models to determine trends in the Nation's rivers and streams, 1972–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7KW5D4H.

2. Ryberg, K.R., Murphy, J.C., Falcone, J.A., Riskin, M.L., Wieben, C.M., and Vecchia, A.V., 2017, Pesticide concentration and streamflow datasets used to evaluate pesticide trends in the Nation's rivers and streams, 1992–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7BC3WPC.

3. Zuellig, R.E. and Riskin, M.L., 2017, Ecological community datasets used to evaluate the presence of trends in ecological communities in selected rivers and streams across the United States, 1992–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7G44ND3.

4. Farmer, W.H., Murphy, J.C., Riskin, M.L., Ryberg, K.R., and Zuellig, R.E., 2017, Daily streamflow datasets used to analyze trends in streamflow at sites also analyzed for trends in water quality and ecological condition in the Nation's rivers and streams: U.S. Geological Survey data release, http://www.dx.doi.org/10.5066/F7D798JN.

5. Mills, T.J., Sprague, L.A., Murphy, J.C., Riskin, M.L., Falcone, J.A., Stets, E.G., Oelsner, G.P, and Johnson, H.M., 2017, Water-quality and streamflow datasets used in Seasonal Kendall trend tests for the Nation's rivers and streams, 1972–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7QN64VT.

Methods

Water-Quality, Pesticide, and Ecology Data Compilation

Water-quality monitoring and some aquatic-ecology data (invertebrates) were compiled from multiple sources for this study and included ambient monitoring data that were readily accessible from Federal, State, regional, and local government agencies and nongovernmental organizations. The primary sources of data were the USGS National Water Information System (NWIS) database, the USGS Aquatic Bioassessment (BioData) database, the EPA STOrage and RETrieval Data Warehouse, Water Quality Exchange (STORET) database, and additional repositories for monitoring data that are not included in STORET. STORET serves water-quality data from multiple Federal, State, regional, and local organizations; NWIS primarily serves data from USGS but also contains a small amount of data collected by other organizations. All available water-quality data for nutrient, sediment, major ion, salinity, carbon, and associated physical parameters from the United States and Puerto Rico for all possible dates were retrieved from NWIS (U.S. Geological Survey, 2016) in May 2013 and from STORET (https://www.epa.gov/waterdata/storage-and-retrieval-and-water-quality-exchange) in October 2013. Only public data from NWIS, those with a Data Quality Indicator code of A (Historical Data), S (Provisional), and R (Reviewed and approved), were used in this study. Replicate samples collected by the NAWQA Project were retrieved in 2016 as described in the "Replicate Evaluation" section of this report. When the data were retrieved in 2013, some water-quality records did not extend into 2012. To ensure that recent data were included, data from NWIS and STORET collected from 2008 to 2014 were retrieved from the Water Quality Portal (WQP, http://www.waterqualitydata.us), a public clearinghouse for both NWIS and STORET data, in three batches between January and March 2015. The additional data from the WQP were added to the original NWIS and STORET data. Duplicate samples that resulted from this addition were removed according to the procedures described in the "Duplicate Samples" section of this report. Publicly-available fish, invertebrate, and algae data were downloaded in December 2013 from BioData (MacCoy, 2011). Between 2011 and 2015, Federal, State, and local agencies in each of the 48 contiguous States were contacted to acquire suitable water-quality data that were available in electronic format and not in NWIS or STORET. Regional groups, such as the Chesapeake Bay Program (CBP), were also contacted to request suitable water-quality monitoring data. The acquired data were provided in multiple formats and were processed into a standardized format with common data elements and terminologies for site types, parameters, fractions, and reporting units. If data from the same agency were received from multiple sources (directly from an agency as well as through STORET), then those data were combined under the same site; however, if two different agencies were collecting data at the same location, the data were not automatically combined. Data from multiple agencies were not combined until after harmonization and only in the instances described in the "Merging Colocated or Neighboring Water-Quality Sites" sections of this report.

More than 185 million water-quality and ecology records from approximately 480,000 sites and over 600 agencies were initially evaluated for this trend study. A list of all of the agencies that supplied data for the data compilation effort including the 74 agencies that supplied data used in this trend study is provided in table 1. Overall, approximately 60 percent of the final records came from NWIS, BioData, and STORET databases. The 73 non-USGS organizations across the United States included in this trend study were composed of 3 Federal agencies; 2 regional (multiple State) organizations; 40 State water, natural resources or environmental protection agencies; 1 tribal organization; 19 county or subcounty organizations; 3 academic organizations; 2 non-governmental organizations; 2 volunteer organizations; and 1 private organization (table 1). Approximately one-half of the data used in this trend study came from State water, natural resources, or environmental protection agencies; one-third of the data came from Federal agencies.

Streamflow Data Compilation

For many rivers, water quality and streamflow are correlated, and properly accounting for this relation during trend analyses allows for a better determination of general changes in water quality over time (Helsel and Hirsch, 2002). This is important for parameters that are strongly related to streamflow and sites where streamflow is highly variable; furthermore, if trends in load are of interest, streamflow data are essential to calculate parameter load (mass), which is the product of water-quality concentration (a mass per volume) and an associated streamflow rate (volume per time). Streamflow data were used in this analyses as an explanatory variable in each of three trend methods and to calculate water-quality loads. The USGS streamflow gages (henceforth, referred to as "streamgages") were the primary source of streamflow data. Most daily mean streamflow values for USGS streamgages were retrieved from NWIS (http://dx.doi.org/10.5066/F7P55KJN) using the freely-available statistical software R (R Development Core Team, 2014) and the function "importDVs" available in the R package "waterData" (Ryberg and Vecchia, 2012). Daily mean streamflow values for four USGS streamgages were not available or were incomplete in the public NWIS database because of database coding issues and were acquired directly from the USGS Water Science Centers that collected the data. The streamflow records for these four sites were all approved for public distribution. Daily mean streamflow values for 17 non-USGS streamgages were retrieved from public databases hosted by the source agencies when available. Recent streamflow data (2011–13) for one of these non-USGS streamgages were not publically available, and this information was retrieved directly from the source agency. Streamflow data from 1,134 unique streamgages were used for this study and included data from 1,117 USGS streamgages and 17 non-USGS streamgages (appendix 1, tables 1–1 through 1–5). Streamgages used in the nutrient, sediment, major ion, salinity, and carbon trend anlalyses are listed in table 1–1; streamgages used in the pesticide trend analysis are listed in table 1–2; and streamgages used in the ecology trend analysis are listed in tables 1–3, 1–4, and 1–5.

Streamgage Matching

Because streamflow data are required for water-quality, pesticide, and ecology trend analysis, sites with water-quality, pesticide, and (or) ecology data had to be assessed to determine whether there was a suitable nearby streamgage. All water-quality sites with at least 3 consecutive years of quarterly samples for any parameter were assessed for their proximity to a suitable streamgage as well as all pesticide and ecology sites. The process of pairing water-quality sites and streamgages is referred to as "streamgage matching" and the generalized process is presented in figure 1.

Because of the large scale of this effort, there are numerous potential sources of error in the site location data, the NHDPlus dataset, and ancillary datasets used to generate site-streamgage combinations. Reasonable efforts were made to identify and correct errors; nevertheless, a small percentage of incorrect site-streamgage combinations may exist in the final combinations because of undetected issues from one or more datasets used in the streamgage-matching process.

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Alabama Department of Environmental Management	AL DEM	Alabama	State	Yes
Alabama Water Watch Association	AL AWWA	Alabama	Volunteer	No
Poarch Band of Creek Indians	AL TPBCI	Alabama	Tribe	No
Alaska Department of Environmental Conservation	AK DEC	Alaska	State	No
Alaska Soil and Water Conservation District	AK SWCD	Alaska	State	No
Kenai Watershed Forum	AK TKWF	Alaska	Tribe	No
Salt Chuck Mine	AK SCM	Alaska	County or subcounty	No
AK-Chin Indian Community	AZ TAKIC	Arizona	Tribe	No
Arizona Department of Environmental Quality	AZ DEQ	Arizona	State	No
Arizona Southwest Watershed Research Center	AZ SWRC	Arizona	Federal	No
Cocopah Indian Tribe Environmental Protection Office	AZ TCOCOPAH	Arizona	Tribe	No
Fort McDowell Yavapai Nation	AZ TFMYN	Arizona	Tribe	No
Hopi Tribe Water Resources Program	AR HOPI	Arizona	Tribe	No
Hualapai Tribe	AZ THUALAPAI	Arizona	Tribe	No
Salt River Pima-Maricopa Indian Community	AZ TSRPMIC	Arizona	Tribe	No
White Mountain Apache Tribe Watershed Program	AZ TWMAWP	Arizona	Tribe	No
Yavapai-Apache Nation	AZ TYAN	Arizona	Tribe	No
Arkansas Department of Environmental Quality	AR DEQ	Arkansas	State	Yes
Arkansas State University Ecotoxicology Research Facility	AR ASUERF	Arkansas	Academic	No
Equilibrium	AR EQUIL	Arkansas	Private	No
GBMc and Associates	AR GBMC	Arkansas	Private	No
Bay Area Clean Water Agencies	CA BACWA	California	County or subcounty	No
Bay Area Stormwater Management Agencies Association	CA BASMAA	California	County or subcounty	No
Blue Lake Rancheria	CA TBLR	California	Tribe	No
California Department of Water Resources	CA DWR	California	State	Yes
California Environmental Data Exchange Network	CA CEDEN	California	State	No
California Marine Research	CA MR	California	Private	No
California Pauma Band of Mission Indians	CA TPBMI	California	Tribe	No
California Regional Water Board, Region 1	CA RWB1	California	State	No
California Regional Water Board, Region 2	CA RWB2	California	State	No
California Regional Water Board, Region 4	CA RWB4	California	State	No
California Regional Water Board, Region 5	CA RWB5	California	State	Yes
California Regional Water Board, Region 6	CA RWB6	California	State	No
California Regional Water Board, Region 7	CA RWB7	California	State	No
California Regional Water Board, Region 9	CA RWB9	California	State	No
California Regional Water Quality Control Board	CA RWQCB	California	State	Yes
California Rice Commission	CA CRC	California	NGO	No
Central Coast Ambient Monitoring Program	CA CCAMP	California	State	No
Coalition for Urban and Rural Environmental Stewardship	CA CURES	California	NGO	No
Coyote Valley Tribal Environmental Protection Department	CA TCV	California	Tribe	No
East San Joaquin Water Quality Coalition	CA ESJWQC	California	NGO	No
Elk Valley Rancheria	CA TEVR	California	Tribe	No
Feather River Coordinated Resource Management	CA FRCRM	California	NGO	No
Fresno River Water Quality Monitoring	CA FRWQM	California	State	No
Goose Lake Coalition	CA GLC	California	NGO	No
Hoopa Valley Tribal Environmental Protection Agency	CA THV	California	Tribe	No
Hopland Band of Pomo Indians Tribal Environmental Protection Agency	CA THBPI	California	Tribe	No
Karuk Tribe, Department of Natural Resources	CA TKARUK	California	Tribe	No
La Posta Band of Mission Indians	CA TLP	California	Tribe	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Lake Elsinore and San Jacinto Watersheds Authority	CA LESJWA	California	NGO	No
Larry Walker Associates	CA LWA	California	NGO	No
Livestock and Land Area of Special Biological Significance	CA LL ASBS	California	State	No
Los Coyotes Band of Cahuilla Cupeno Indians	CA TLCBCCI	California	Tribe	No
Mattole Salmon Group	CA TMSG	California	Tribe	No
Modesto Irrigation District	CA MID	California	County or subcounty	No
Morongo Band of Mission Indians	CA TMBMI	California	Tribe	No
Owens Valley Indian Water Commission	CA TOVIWC	California	Tribe	No
Pit River Tribe	CA TPR	California	Tribe	No
Quartz Valley Indian Reservation	CA TQVIR	California	Tribe	No
Quechan Indian Tribe	CA TQI	California	Tribe	No
Resighini Rancheria Environmental Protection	CA RREP	California	Tribe	No
Robinson Rancheria Environmental Center	CA RREC	California	Tribe	No
Sacramento Valley Water Quality Coalition	CA SVWQC	California	NGO	No
San Francisco Estuary Institute	CA SFEI	California	NGO	No
San Joaquin County and Delta Water Quality Coalition	CA SJCDWQC	California	NGO	No
San Joaquin River, Total Maximum Daily Load	CA SJR	California	State	No
San Manuel Band of Mission Indians	CA TSMBMI	California	Tribe	No
San Pasqual Band of Diegueno Indian	CA TSPAS	California	Tribe	No
Santa Ana Watershed Project Authority	CA SAWPA	California	State	No
Santa Ynez Chumash Environmental Office	CA TSYCEO	California	Tribe	No
Smith River Rancheria	CA TSRR	California	Tribe	No
Soboba Band of Luiseno Indians	CA TSBLI	California	Tribe	No
Southern California Stormwater Monitoring Coalition	CA SMC	California	County or subcounty	No
Southern San Joaquin Water Quality Coalition	CA SSJWQC	California	NGO	No
State Water Resources Control Board	CA SWRCB	California	State	No
Surface Water Ambient Monitoring Program	CA SWAMP	California	State	No
Sutter County Resource Conservation District	CA SCRCD	California	County or subcounty	No
Table Mountain Rancheria	CA TTMR	California	Tribe	No
Twenty-Nine Palms	CA T29P	California	Tribe	No
University of California	CA UC	California	State	No
Volunteer Monitoring Groups, US Environmental Protection Agency Region 9	CA VMG	California	Volunteer	No
Westlands Stormwater Coalition	CA WSWC	California	NGO	No
Westside San Joaquin River Watershed Coalition	CA WSSJRWC	California	NGO	Yes
Wiyot Tribe	CA TWIYOT	California	Tribe	No
Yurok Tribe	CA TYUROKTEP	California	Tribe	No
Adams Rib Ranch	CO ARR	Colorado	Volunteer	No
Adrian Brown Consultants Incorporated	CO ABC	Colorado	Private	No
Advanced Sciences Incorporated	CO ASI	Colorado	Private	No
American Geological Services Incorporated	CO AGS	Colorado	Private	No
Animas River Stakeholders Group	CO ARSG	Colorado	NGO	No
Arapahoe Basin Ski Resort	CO ABSR	Colorado	Private	No
Barr Lake Milton Reservoir Watershed Association	CO BLMRW	Colorado	NGO	No
Bear Creek Reservoir	CO BCR	Colorado	County or subcounty	No
Berry Petroleum Company	CO BPG	Colorado	Private	No
Big Thompson Watershed Forum	CO BTWF	Colorado	NGO	No
Breckenridge Sanitation District	CO BSD	Colorado	County or subcounty	No
CBS Operations Incorporated	CO CBS	Colorado	Private	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Centennial Water and Sanitation District	CO CWSD	Colorado	County or subcounty	No
Chatfield Reservoir	CO CR	Colorado	County or subcounty	No
City of Aurora	CO AUR	Colorado	County or subcounty	No
City of Brighton	CO BRIGHT	Colorado	County or subcounty	No
City of Fort Collins	CO FORT	Colorado	County or subcounty	No
City of Glendale	CO GLEN	Colorado	County or subcounty	No
City of Grand Junction	CO GRJ	Colorado	County or subcounty	No
City of Palisades	CO PALI	Colorado	County or subcounty	No
City of Pueblo	CO PUEB	Colorado	County or subcounty	No
City of Rifle	CO RIFLE	Colorado	County or subcounty	No
City of Silt	CO SILT	Colorado	County or subcounty	No
City of Steamboat Springs	CO STEAMBOAT	Colorado	County or subcounty	No
City of Thornton	CO THORN	Colorado	County or subcounty	Yes
Clear Creek Watershed Foundation	CO CCWF	Colorado	NGO	No
Coal Creek Watershed Coalition	CO CCWC	Colorado	NGO	No
Colorado 319 Project	CO 319	Colorado	State	No
Colorado Department of Natural Resources	CO DNR	Colorado	State	No
Colorado Department of Public Health and Environment	CO DPHE	Colorado	State	Yes
Colorado Department of Transportation	CO DOT	Colorado	State	No
Colorado Division of Reclamation, Mining and Safety	CO DRMS	Colorado	State	No
Colorado Geological Survey	CO GS	Colorado	State	No
Colorado Mountain College	CO CMC	Colorado	Academic	No
Colorado Oil and Gas Conservation Commission	CO OGCC	Colorado	County or subcounty	No
Colorado Selenium Studies (USGS)	CO SS USGS	Colorado	Federal	No
Colorado Springs Utilities	CO CSUTL	Colorado	County or subcounty	No
Colorado State University	CO CSU	Colorado	Academic	No
Colorado Wildlife Division	CO WD	Colorado	State	No
Denver Department Of Environmental Health	CO DDEH	Colorado	County or subcounty	No
Denver Water Department	CO DWD	Colorado	County or subcounty	No
Eagle Mine Superfund Site Assessment	CO EMSSA	Colorado	Federal	No
Eagle River Water and Sanitation District	CO ERWSD	Colorado	County or subcounty	No
ENCANA Corporartion	CO ENCANA	Colorado	Private	No
Grand River Consulting	CO GRC	Colorado	Private	No
Gunnison Basin & Grand Valley Selenium Task Force	CO GBGV STF	Colorado	Volunteer	No
Hydrosphere Resource Consultants Incorporated	CO HRC	Colorado	Private	No
Lake Fork Watershed Stakeholders	CO LFWS	Colorado	NGO	No
Littleton and Englewood Wastewater Treatment Plant	CO LE WWTP	Colorado	County or subcounty	Yes
Metro Wastewater Reclamation District	CO MWRD	Colorado	County or subcounty	Yes
Mining Remedial Recovery Company	CO MRRC	Colorado	Private	No
North Fork River Improvement Association	CO NFRIA	Colorado	NGO	No
Northern Colorado Water Conservancy District	CO NCWCD	Colorado	Federal	No
Northwest Colarado Council of Government	CO NWCG	Colorado	County or subcounty	No
Pueblo of Acoma	CO TPA	Colorado	Tribe	No
San Juan Citizens Alliance	CO SJCA	Colorado	NGO	No
Silverthorne Dillon Joint Sewer Authority	CO SDJSA	Colorado	County or subcounty	No
South Adams County Water and Sanitation District	CO SACWSD	Colorado	County or subcounty	No
Southern Ute Tribe	CO TSU	Colorado	Tribe	No
Standley Lake Watershed Group	CO STAND	Colorado	NGO	No
Summit Water Quality Committee	CO SWQC	Colorado	County or subcounty	Yes
Sunnyside Gold Corporation	CO SGC	Colorado	Private	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
The Rivers of Colorado Water Watch Network	CO CRWWN	Colorado	Volunteer	Yes
Town of Telluride Public Works	CO TPW	Colorado	County or subcounty	No
University of Colorado Institute of Arctic and Alpine Research	CO INSTAAR	Colorado	Academic	No
Ute Mountain Utes Tribe	CO TUMU	Colorado	Tribe	No
Wright Water Engineers Incorporated	CO WWE	Colorado	Private	No
Connecticut Department of Energy and Environmental Protection	CT DEEP	Connecticut	State	No
Delaware Department of Natural Resources and Environmental Control	DE DNREC	Delaware	State	Yes
Delaware Geological Survey	DE DGS	Delaware	State	No
District Department of the Environment	DC DDOE	District of Columbia	State	Yes
Alachua County Environmental Protection Department	FL ACEPD	Florida	County or subcounty	No
AMEC Foster Wheeler	FL AMEC	Florida	Private	No
Avon Park Air Force Range	FL APAFR	Florida	Federal	No
Babcock Ranch	FL BR	Florida	Private	No
Biological Research Associates	FL BRA	Florida	Private	No
Brevard County Stormwater Utility Department	FL BCSUD	Florida	County or subcounty	No
Broward County Environmental Protection Department	FL BCEPD	Florida	County or subcounty	No
Charlotte County	FL CC	Florida	County or subcounty	No
Choctawhatchee Basin Alliance	FL CBA	Florida	Volunteer	No
City of Atlantic Beach	FL CAB	Florida	County or subcounty	No
City of Cape Coral	FL CCC	Florida	County or subcounty	No
City of Deltona	FL CD	Florida	County or subcounty	No
City of Jacksonville	FL CJ	Florida	County or subcounty	No
City of Jacksonville Beach	FL JB	Florida	County or subcounty	No
City of Naples	FL NAPLES	Florida	County or subcounty	No
City of Neptune Beach	FL CNB	Florida	County or subcounty	No
City of Port Saint Joe Wastewater Treatment Plant	FL CPSJ	Florida	County or subcounty	No
City of Punta Gorda	FL CPG	Florida	County or subcounty	No
City of Sanibel, Natural Resources Department	FL CSNRD	Florida	County or subcounty	No
City of St Petersburg	FL CSP	Florida	County or subcounty	No
City of Tallahassee Stormwater	FL CTS	Florida	County or subcounty	No
City of Tampa Bay Study Group	FL TAMP	Florida	County or subcounty	No
Collier County Coastal Zone Management Department	FL CCCZMD	Florida	County or subcounty	No
Collier County Pollution Control	FL CCPC	Florida	County or subcounty	No
Dade Environmental Resource Management	FL DERM	Florida	County or subcounty	No
Environmental Protection Commission of Hillsborough County	FL HLCE	Florida	County or subcounty	Yes
Florida Department of Environmental Protection	FL DEP	Florida	State	Yes
Florida Department of Regulatory and Economic Resources	FL DRER	Florida	County or subcounty	No
Florida Fish and Wildlife	FL FW	Florida	State	No
Florida Fish and Wildlife Conservation Commission	FL FWCC	Florida	State	Yes
Florida Keys National Marine Sanctuary	FL NMS	Florida	State	No
Florida Lakewatch	FL LW	Florida	Volunteer	No
Gulf Power Company	FL GPC	Florida	Private	No
IMC Agrico	FL IMCA	Florida	Private	No
Jacksonville Electric Authority	FL JEA	Florida	Private	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Lake County Water Resource Management	FL LCRM	Florida	County or subcounty	No
Lee County	FL LEECO	Florida	County or subcounty	No
Lee County Hyacinth Control District	FL LCHCD	Florida	County or subcounty	No
Leon County Public Works	FL LCPW	Florida	County or subcounty	Yes
Loxahatchee River District	FL LRD	Florida	County or subcounty	Yes
Manatee County Environmental Management Department	FL MCEMD	Florida	County or subcounty	Yes
Marine Resources Council of East Florida	FL MRCEF	Florida	Private	No
McGlynn Laboratories Incorporated	FL MCGL	Florida	Private	Yes
Naval Station Mayport	FL NSM	Florida	Federal	No
Northwest Florida Water Management District	FL NWFWMD	Florida	County or subcounty	Yes
Orange County Environmental Protection	FL OCEP	Florida	County or subcounty	Yes
Palm Beach County Environmental Resources Management Department	FL PBCERMD	Florida	County or subcounty	No
Pasco County Stormwater Management Division	FL PCSMD	Florida	County or subcounty	No
Peace River Manasota Regional Water Supply Authority	FL PRMRWS	Florida	County or subcounty	No
Pinellas County Department of Environmental Management	FL PCDEM	Florida	County or subcounty	Yes
Polk County Water Resources	FL PCWR	Florida	County or subcounty	No
Reedy Creek Improvement District - Environmental Services	FL RCID	Florida	County or subcounty	No
Sanibel Captiva Conservation Foundation	FL SCCF	Florida	County or subcounty	No
Sarasota County Environmental Services	FL SCES	Florida	County or subcounty	No
Schroeder-Manatee Ranch Communities Incorporated	FL SMRC	Florida	Private	No
Seminole County	FL SEMC	Florida	County or subcounty	Yes
Seminole Tribe of Florida	FL TSF	Florida	Tribe	No
South Florida Water Management District	FL SFWMD	Florida	County or subcounty	No
Southwest Florida Water Management District	FL SWFWMD	Florida	County or subcounty	Yes
St. Johns Water Management District	FL SJWMD	Florida	County or subcounty	Yes
Suwannee River Water Management District	FL SRWMD	Florida	County or subcounty	Yes
Tampa Bay Water	FL TBW	Florida	County or subcounty	No
The Conservancy of Southwest Florida	FL CSW	Florida	NGO	No
Volusia County Environmental Health Lab	FL VCEHL	Florida	County or subcounty	No
Dial Cordy and Associates	GA DCA	Georgia	Private	No
Georgia Department of Natural Resources	GA DNR	Georgia	State	Yes
Georgia Department of Natural Resources, Coastal Resources Division	GA DNRCRD	Georgia	State	No
Southeastern Natural Sciences Academy	GA SENSA	Georgia	Private	No
Hawaii Department of Health	HI DOH	Hawaii	State	No
Coeur dAlene Tribe	ID TCD	Idaho	Tribe	No
Idaho Department of Environmental Quality	ID DEQ	Idaho	State	Yes
Idaho Environmental Data Management System	ID EDMS	Idaho	State	No
Nez Perce Tribe	ID TNP	Idaho	Tribe	No
Shoshone Bannock Tribe	ID TSB	Idaho	Tribe	No
American Water Company	IL AWC	Illinois	County or subcounty	No
Illinois Environmental Protection Agency	IL EPA	Illinois	State	Yes
Illinois State Water Survey	IL SWS	Illinois	State	No
Metropolitan Water Reclamation District of Greater Chicago	IL MWRDGC	Illinois	County or subcounty	No
Park District of Highland Park	IL PDHP	Illinois	County or subcounty	No
City of Elkhart Public Works and Utilities	IN EPWU	Indiana	County or subcounty	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Clifty Creek Watershed Project	IN CCWP	Indiana	Volunteer	No
Clinton County Soil Water and Conservation District	IN CSWCD	Indiana	County or subcounty	No
Dearborn County Soil Water and Conservation District	IN DSWCD	Indiana	County or subcounty	No
Hoosier Riverwatch	IN HRW	Indiana	Volunteer	No
Indiana Department of Environmental Management	IN DEM	Indiana	State	Yes
Indiana Office of State Chemist	IN OSC	Indiana	State	No
Indianapolis Citizen Energy Group	IN CEG	Indiana	County or subcounty	No
Marion County Public Health Department	IN MCHD	Indiana	County or subcounty	Yes
Muncie Bureau of Water Quality	IN MBWQ	Indiana	County or subcounty	No
Save the Dunes Conservation Fund	IN STDCF	Indiana	Volunteer	No
Tipton County Soil Water and Conservation District	IN TSWCD	Indiana	County or subcounty	No
Wabash River Enhancement Corporation	IN WREC	Indiana	Volunteer	Yes
Wayne County Soil Water and Conservation District	IN WSWCD	Indiana	County or subcounty	No
Des Moines River Water Quality Network	IA ISUCOE	Iowa	Academic/Federal	Yes
Des Moines Water Works	IA DMWW	Iowa	County or subcounty	Yes
Iowa Department of Natural Resources	IA DNR	Iowa	State	Yes
Iowa Geological Survey Watershed Snapshots	IA GSWSP	Iowa	State	No
Iowa Volunteer Water Monitoring Program	IA VWMP	Iowa	Volunteer	No
Meskwaki Nation, Department of Natural Resources	IA TMN	Iowa	Tribe	No
Kansas Biological Survey	KS BS	Kansas	State	No
Kansas Department of Health and Environment	KS DHE	Kansas	State	Yes
Kaw Nation	KS TKN	Kansas	Tribe	No
Kickapoo Tribe in Kansas	KS TKTK	Kansas	Tribe	No
Prairie Band Potawatomi Nation	KS TPBPN	Kansas	Tribe	No
Quapaw Tribe of Oklahoma	KS TQTO	Kansas	Tribe	No
Kentucky Department for Natural Resources and Environmental Protection	KY DEP	Kentucky	State	Yes
Northern Kentucky University	KY NKU	Kentucky	Academic	No
Watershed Watch	KY WW	Kentucky	Volunteer	No
Louisiana Department of Environmental Quality	LA DEQ	Louisiana	State	Yes
Houlton Band of Maliseet Indians	ME THBMI	Maine	Tribe	No
Maine Department of Environmental Protection	ME DEP	Maine	State	No
Passamaquoddy Tribe at Pleasant Point	ME TPTPP	Maine	Tribe	No
Chesapeake Bay Program	MD CBP	Maryland	Regional	No
Friends of Sligo Creek	MD FSC	Maryland	Volunteer	No
Maryland Department of Natural Resources	MD DNR	Maryland	State	Yes
Maryland Department of the Environment	MD DOE	Maryland	State	No
Charles River Watershed Association	MA CRWA	Massachusetts	Volunteer	No
Massachusetts Department of Environmental Protection	MA DEP	Massachusetts	State	No
Massachusetts Westfield River Water Monitoring Program	MA WRWMP	Massachusetts	Volunteer	No
Merrimack River Watershed Council	MA MRWC	Massachusetts	Volunteer	No
Mystic River Watershed Association	MA MRWA	Massachusetts	Volunteer	No
Wampanoag Tribe of Gay Head	MA TWTGH	Massachusetts	Tribe	No
Bay Mills Indian Community	MI TBMIC	Michigan	Tribe	No
Grand Traverse Band of Ottawa and Chippewa Indians	MI TGTBOCI	Michigan	Tribe	No
Hannaville Tribal Community	MI THTC	Michigan	Tribe	No
Little River Band of Ottawa Indians	MI TLRBOI	Michigan	Tribe	No
Little Traverse Bay Bands of Odawa Indians	MI TLTBBOI	Michigan	Tribe	No
Match-e-be-nash-she-wish Band of Potawatomi	MI TMBP	Michigan	Tribe	No
Michigan Department of Environmental Quality	MI DEQ	Michigan	State	Yes

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Nottawaseppi Huron Band of the Potawatomi	MI TNHBP	Michigan	Tribe	No
Pokagon Band of Potawatomi Indians, Department of Natural Resources	MI TPBPI	Michigan	Tribe	No
Saginaw Chippewa Planning Department	MI TSCPD	Michigan	Tribe	No
University of Michigan	MI UMI	Michigan	Academic	No
Emmons and Oliver Resources	MN EOR	Minnesota	Private	No
Fond du Lac Band of Chippewa	MN TFDL	Minnesota	Tribe	No
Grand Portage Band of Chippewa Indians	MN TGPR	Minnesota	Tribe	No
Lower Sioux Indian Community	MN TLSIC	Minnesota	Tribe	No
Metropolitan Council Environmental Services	MN MCES	Minnesota	County or subcounty	Yes
Minnesota Department of Agriculture	MN DA	Minnesota	State	Yes
Minnesota Pollution Control Agency	MN PCA	Minnesota	State	Yes
Nett Lake Band of Chippawa Indians	MN TBFNL	Minnesota	Tribe	No
Prairie Island Community	MN TPIC	Minnesota	Tribe	No
Red Lake Division of Natural Resources	MN TRLDNR	Minnesota	Tribe	No
Shakopee Mdewakanton	MN TSMSC	Minnesota	Tribe	No
Upper Sioux Community	MN TUSIC	Minnesota	Tribe	No
Mississippi Department of Environmental Quality	MS DEQ	Mississippi	State	No
MS Band of Choctaw Indians	MS TBCI	Mississippi	Tribe	No
Advanced Aquatic Technology Associates Incorporated	MO AATA	Missouri	Private	No
Ameren Union Electric Corporation	MO AUEC	Missouri	Private	No
Archer Engineering Incorporated	MO ARCH	Missouri	Private	No
Associated Electric Cooperative Incorporated	MO AECI	Missouri	Private	No
Black and Veatch Incorporated	MO BV	Missouri	Private	No
Blue River Watershed Association	MO BRWA	Missouri	Volunteer	No
Brookside Environmental Services	MO BRKSD	Missouri	Private	No
Camp Dresser McKee Incorporated	MO CDM	Missouri	Private	No
Carthage High School	MO CHS	Missouri	Academic	No
Certainteed Incorporated	MO CT	Missouri	Private	No
Christian County Health Department	MO CCHD	Missouri	County or subcounty	No
City of Sprinfield Department of Public Works	MO SPFDPW	Missouri	County or subcounty	No
City Utilities of Springfield	MO CUSF	Missouri	County or subcounty	No
Consulting Analytical Services Incorporated	MO CASI	Missouri	Private	No
Continental Grain Corporation	MO CNGRN	Missouri	Private	No
Crowder College, Neosho, Missouri	MO CC	Missouri	Academic	Yes
Doe Run Corporation	MO DOERN	Missouri	Private	No
Ecology and Environment Incorporated	MO EE	Missouri	Private	No
Empire Electric Cooperative Incorporated	MO EEC	Missouri	Private	No
Entrix Incorporated	MO ENTX	Missouri	Private	No
Envirodyne Incorporated	MO ENVDYN	Missouri	Private	No
Environ International	MO ENVRN	Missouri	Private	No
Environmental Science and Engineering	MO ESE	Missouri	Private	No
Environmental Systems Corporation	MO ESC	Missouri	Private	No
Forrester Group Incorporated	MO FORS	Missouri	Private	No
Hydrometrics Incorporated	MO HYDR	Missouri	Private	No
Jasper County Health Department	MO JCHD	Missouri	County or subcounty	No
Johnson County Egg Farm	MO JCEF	Missouri	Private	No
Kansas City Public Works	MO KCPW	Missouri	County or subcounty	No
Lake Ozark Council of Government	MO LOCOG	Missouri	County or subcounty	No
Lake Saint Louis Community Association	MO LSLCA	Missouri	County or subcounty	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Lawrence County Health Department	MO LCHD	Missouri	County or subcounty	No
Mactec Incorporated	MO MACT	Missouri	Private	No
Meek Farms Incorporated	MO MEEK	Missouri	Private	No
Metropolitan Sewer District of Saint Louis	MO MSD	Missouri	County or subcounty	No
Midwest Environmental Consultants	MO MEC	Missouri	Private	No
Missouri American Water Company	MO AMRWC	Missouri	County or subcounty	No
Missouri Department of Conservation	MO DOC	Missouri	State	No
Missouri Department of Health and Senior Services	MO DHSS	Missouri	State	No
Missouri Department of Natural Resources	MO DNR	Missouri	State	No
Missouri State University	MO MSU	Missouri	Academic	No
Missouri Water Pollution Control Board	MO WPCB	Missouri	State	No
Murphy Family Farms Incorporated	MO MRPH	Missouri	Private	No
Newfields Incorporated	MO NWFD	Missouri	Private	No
Newton County Health Department	MO NCHD	Missouri	County or subcounty	No
Novartis Incorporated	MO NVRT	Missouri	Private	No
Peabody Coal Company	MO PCC	Missouri	Private	No
Perry County Soil and Water District	MO PCSWD	Missouri	County or subcounty	No
Premium Standard Farms Incorporated	MO PSF	Missouri	Private	No
Sac and Fox Nation of Missouri in Kansas and Nebraska	MO TSACFOX	Missouri	Tribe	No
Saint Joseph City Water Company	MO SJCWC	Missouri	County or subcounty	No
Saint Louis City Water Company	MO STLWC	Missouri	County or subcounty	No
Science Applications International Corporation	MO SAIC	Missouri	Private	No
Sharp Land and Cattle Company	MO SLCC	Missouri	Private	No
Southeast Missouri State University	MO SEMSU	Missouri	Academic	No
Stone County Health Department	MO SCHD	Missouri	County or subcounty	No
Sverdrup and Parcel Incorporated	MO SVDRP	Missouri	Private	No
Syngenta Crop Science Incorporated	MO SNGNTA	Missouri	Private	No
Terranext Corporation	MO TRNX	Missouri	Private	No
Tetratech Incorporated	MO TTRT	Missouri	Private	No
Tyson Foods	MO TYSN	Missouri	Private	No
University of Missouri, Columbia	MO UMC	Missouri	Academic	No
University of Missouri, Rolla	MO UMR	Missouri	Academic	No
Versar Incorporated	MO VRSR	Missouri	Private	No
Washington University, Saint Louis	MO WUSL	Missouri	Academic	No
Water Sentinels Sierra Club	MO WSSC	Missouri	Volunteer	No
West Incorporated	MO WEST	Missouri	Private	No
William Jud, volunteer	MO JUD	Missouri	Volunteer	No
Eastern Shawnee Tribe of Oklahoma	OK ESTO	MO	Tribe	No
Assiniboine and Sioux Tribes Fort Peck Indian Reservation	MT TFPIR	Montana	Tribe	No
Blackfeet Nation (Montana)	MT TBN	Montana	Tribe	No
Chippewa Cree Tribe	MT TCC	Montana	Tribe	No
Gros Ventre and Assiniboine Tribe	MT TGVAT	Montana	Tribe	No
Montana Bureau of Mines and Geology	MT BMG	Montana	State	No
Montana Department of Environmental Quality	MT DEQ	Montana	State	Yes
Montana PPL Corporation	MT PPLC	Montana	Private	No
Montana Volunteer Water Quality Monitoring	MT VWQM	Montana	Volunteer	No
Montana Watershed Data	MT WD	Montana	State	No
Northen Cheyenne Tribe	MT TNC	Montana	Tribe	No
Otter Creek Coal	MT OCC	Montana	Private	No
Nebraska Department of Environmental Quality	NE DEQ	Nebraska	State	Yes

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Winnebago Tribe	NE TWIN	Nebraska	Tribe	No
Moapa Band of Pautes	NV TMBP	Nevada	Tribe	No
Nevada Department of Agriculture	NV DA	Nevada	State	No
Nevada Department of Conservation and Natural Resources	NV DECNR	Nevada	State	Yes
Pyramid Lake Paiute Tribe	NV TPLP	Nevada	Tribe	No
Shoshone-Paiute Tribes	NV TSP	Nevada	Tribe	No
Southern Nevada Water Authority (Environmental Monitoring and Management)	NV SNWA	Nevada	State	No
Washoe Tribe of NV and CA	NV TW	Nevada	Tribe	No
New Hampshire Department of Environmental Services	NH DES	New Hampshire	State	No
University of New Hampshire Center for Freshwater Biology	NH UCFB	New Hampshire	Academic	No
Upper Merrimack River Local Advisory Committee	NH UMRLAC	New Hampshire	Volunteer	No
Alliance For A Living Ocean	NJ ALO	New Jersey	Volunteer	No
Brick Utilities	NJ BTMUA	New Jersey	Private	No
Great Swamp Watershed Association	NJ GSWA	New Jersey	Volunteer	No
Monmouth County Health Department	NJ MCHD	New Jersey	County or subcounty	No
Musconetcong River Watershed Association	NJ MRWA	New Jersey	Volunteer	No
New Jersey Department of Environmental Protection	NJ DEP	New Jersey	State	Yes
New Jersey Harbor Dischargers Group	NJ HDG	New Jersey	Volunteer	No
New Jersey Pinelands Commission	NJ PC	New Jersey	Volunteer	No
Pequannock River Coalition	NJ PRC	New Jersey	Volunteer	No
Rutgers Cooperative Extension Water Resource Program	NJ RCEWRP	New Jersey	Academic	No
Stony Brook-Millstone Watershed Association	NJ SBMWA	New Jersey	Volunteer	No
Daniel B. Stephens and Associates	NM DBSA	New Mexico	Private	No
LV Wilcox and J Williams	NM LVW JW	New Mexico	Academic	No
Mescalero Apache Tribe	NM TMAT	New Mexico	Tribe	No
Nambe Pueblo	NM TNP	New Mexico	Tribe	No
New Mexico Environment Department	NM ED	New Mexico	State	No
New Mexico Interstate Stream Commission Lower Rio Grande Compendium	NM ISC LRG COMP	New Mexico	Regional	No
New Mexico San Ildefonso Pueblo	NM SIP	New Mexico	State	No
New Mexico State University	NM SU	New Mexico	Academic	No
Pueblo of Isleta	NM TPI	New Mexico	Tribe	No
Pueblo of Jemez	NM TPJ	New Mexico	Tribe	No
Pueblo of Laguna	NM TPL	New Mexico	Tribe	No
Pueblo of Pojoaque	NM TPP	New Mexico	Tribe	No
Pueblo of Sandia Water Quality Program	NM TPSWQP	New Mexico	Tribe	No
Pueblo of Santa Ana	NM TPSA	New Mexico	Tribe	No
Pueblo of Santa Clara	NM TPSC	New Mexico	Tribe	No
Pueblo of Taos	NM TPT	New Mexico	Tribe	No
Pueblo of Tesuque Environment Department	NM TPTED	New Mexico	Tribe	No
San Felipe Pueblo	NM TSFP	New Mexico	Tribe	No
SS Papadopulos and Associates	NM SSPA	New Mexico	Private	No
Sustainability of semi-Arid Hydrology and Riparian Areas	NM SAHRA	New Mexico	Academic	No
University of New Mexico	NM UNM	New Mexico	Academic	No
Community Science Institute	NY COMSI	New York	NGO	Yes

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Cortland County Soil and Water Conservation District	NY CCSWCD	New York	County or subcounty	No
New York Department of Environmental Conservation	NY DEC	New York	State	No
New York tribe unspecified	NY TMAMIN	New York	Tribe	No
Otsego County Soil and Water Conservation District	NY OCSWCD	New York	County or subcounty	No
Saint Regis Mohawk Tribe	NY TSRM	New York	Tribe	No
Sparkill Creek Watershed Alliance	NY SCWA	New York	Volunteer	No
State University of New York, Oneonta	NY SUNYO	New York	Academic	No
Eastern Band of Cherokee Indians	NC TEBCI	North Carolina	Tribe	No
Haywood Waterways Association, Incorporated	NC HWA	North Carolina	Private	No
North Carolina Department of Envrionment and Natural Resources	NC DENR	North Carolina	State	Yes
North Dakota Department of Health	ND DOH	North Dakota	State	Yes
North Dakota State Water Commission	ND SWC	North Dakota	State	No
Spirit Lake Tribal	ND TSL	North Dakota	Tribe	No
Three Affiliated Tribes: Mandan, Hidatsa, and Arikara Nation	ND TMHAN	North Dakota	Tribe	No
Turtle Mountain Environmental Office	ND TMEO	North Dakota	Tribe	No
Heidelberg University, National Center for Water Quality Research	OH HDLBG	Ohio	Academic	Yes
Miami Conservancy District	OH MCD	Ohio	County or subcounty	No
Northeast Ohio Regional Sewer District	OH NEORSD	Ohio	State	No
Ohio Environmental Protection Agency	OH EPA	Ohio	State	Yes
Absentee Shawnee Tribe of Oklahoma	OK TAST	Oklahoma	Tribe	No
Cherokee Nation	OK TCN	Oklahoma	Tribe	No
Chickasaw Nation	OK TCKN	Oklahoma	Tribe	No
Choctaw Nation	OK TCHN	Oklahoma	Tribe	No
Citizen Potawatomi Nation	OK TCPN	Oklahoma	Tribe	No
Delaware Nation	OK TDN	Oklahoma	Tribe	No
Iowa Tribe of Oklahoma	OK TI	Oklahoma	Tribe	No
Kaw Nation of Oklahoma	OK TKN	Oklahoma	Tribe	No
Kickapoo Tribe of Oklahoma	OK TKCN	Oklahoma	Tribe	No
Kiowa Tribe of Oklahoma	OK TK	Oklahoma	Tribe	No
Miami Tribe of Oklahoma	OK TMI	Oklahoma	Tribe	No
Muscogee Creek Nation	OK TMCN	Oklahoma	Tribe	No
Oklahoma Conservation Commission	OK CONC	Oklahoma	State	No
Oklahoma Corporation Commission	OK CORC	Oklahoma	State	No
Oklahoma Department of Environmental Quality	OK DEQ	Oklahoma	State	No
Oklahoma Environmental Protection Office	OK TEPO	Oklahoma	Tribe	No
Oklahoma Water Resources Board	OK WRB	Oklahoma	State	Yes
Osage Nation Environmental and Natural Resources Department	OK ONENRD	Oklahoma	Tribe	No
Otoe Missouria Tribe of Oklahoma	OK TOM	Oklahoma	Tribe	No
Ottawa Tribe of Oklahoma	OK TOT	Oklahoma	Tribe	No
Pawnee Nation	OK TPN	Oklahoma	Tribe	No
Peoria Tribe of Indians	OK TPI	Oklahoma	Tribe	No
Quapaw Tribe Environmental Office	OK TQU	Oklahoma	Tribe	No
Sac and Fox Nation	OK TSFN	Oklahoma	Tribe	No
Seneca Cayuga Tribe	OK TSC	Oklahoma	Tribe	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Tonkawa Tribe of Oklahoma	OK TT	Oklahoma	Tribe	No
Wichita Department of Environmental Programs	OK WDEP	Oklahoma	State	No
Wyandotte Nation (Oklahoma)	OK TWN	Oklahoma	Tribe	No
Burns Paiute Tribe	OR TBP	Oregon	Tribe	No
City of Hillsboro, Oregon Water Department	OR HWD	Oregon	County or subcounty	No
Confederated Tribes of the Coos, Lower Umpqua and Siuslaw Indians	OR TCTC LUSI	Oregon	Tribe	No
Cow Creek Umpqua Indians	OR TCCU	Oregon	Tribe	No
Klamath Tribes Research Station	OR KTRS	Oregon	Tribe	Yes
Oregon Clackamas River Water	OR CRW	Oregon	State	No
Oregon Department of Agriculture	OR DA	Oregon	State	No
Oregon Department of Environmental Quality	OR DEQ	Oregon	State	Yes
Oregon Klamath Basin Monitoring Program	OR KBMP	Oregon	NGO	No
Oregon Public Works, City of Salem	OR SPW	Oregon	County or subcounty	No
Oregon, Portland Water Bureau	OR PWB	Oregon	County or subcounty	Yes
The Confederated Tribes of Grand Ronde	OR TCTGR	Oregon	Tribe	No
Upper Deschutes Watershed Council	OR UDWC	Oregon	State	No
Antietam Watershed Association	PA AWA	Pennsylvania	Volunteer/Academic	No
Appalachia Hydrogeologic and Environmental Consulting, Inc.	PA AHEC	Pennsylvania	Private	No
Centre County Senior Environmental Corps	PA CCSEC	Pennsylvania	Volunteer	No
ClearWater Conservancy	PA CWCY	Pennsylvania	Volunteer	No
Clinton County Conservation District	PA CCCD	Pennsylvania	County or subcounty	No
Conodoguinet Creek Watershed Association	PA CDGTCWA	Pennsylvania	Volunteer	No
Countryside Conservancy	PA CSCY	Pennsylvania	Volunteer	No
Dauphin County Conservation District	PA DCCD	Pennsylvania	County or subcounty	No
Dickinson College - ALLARM environmental organization	PA DCALL	Pennsylvania	Academic	No
God’s Country Trout Unlimited	PA GCTU	Pennsylvania	Volunteer	No
Keystone Watershed Montioring Network	PA KWMN	Pennsylvania	Volunteer	No
Lancaster County Conservation District	PA LCCD	Pennsylvania	County or subcounty	No
Northumberland County Conservation District	PA NCCD	Pennsylvania	County or subcounty	No
Pennsylvania Department of Environmental Protection	PA DEP	Pennsylvania	State	Yes
Pennsylvania Fish and Boat Commission	PA FBC	Pennsylvania	State	No
Pennsylvania Senior Environment Corps and Nature Abounds	PA SECSK	Pennsylvania	Volunteer	No
Pennsylvania State University	PA PSU	Pennsylvania	Academic	No
Schuylkill Action Network	PA SAN	Pennsylvania	Volunteer	No
Shermans Creek Conservation Association	PA SHMCCA	Pennsylvania	Volunteer	No
Stroud Water Research Center	PA SWRC	Pennsylvania	Academic	No
The E.L. Rose Conservancy of Susquehanna County	PA ELRCSC	Pennsylvania	Volunteer	No
Puerto Rico Department of Natural and Environmental Resources	PR DNER	Puerto Rico	State	No
Rhode Island Department of Health	RI DH	Rhode Island	State	No
Environmental Quality Laboratory Coastal Carolina University	SC EQLCCU	South Carolina	Academic	No
South Carolina Department of Health and Environmental Control	SC DHEC	South Carolina	State	Yes
South Carolina Santee Cooper Public Service Authority	SC SCPSA	South Carolina	State	No
Crow Creek Sioux Tribe	SD TCCS	South Dakota	Tribe	No
Flandreau Santee Sioux Tribe	SD TFSS	South Dakota	Tribe	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Lower Brule Sioux Tribe	SD TLBS	South Dakota	Tribe	No
Oglala Sioux Tribe	SD TOS	South Dakota	Tribe	No
South Dakota Department of Environmental and Natural Resources	SD DENR	South Dakota	State	Yes
South Dakota Geological Survey	SD GS	South Dakota	State	No
Tennessee Department of Agriculture	TN DA	Tennessee	State	No
Tennessee Department of Environment and Conservation	TN DEC	Tennessee	State	No
Alabama-Coushatta Tribe of Texas	TX TAC	Texas	Tribe	No
Meadows Center for Water and the Environment	TX MCWE	Texas	Volunteer	No
Texas Commission on Environmental Quality	TX CEQ	Texas	State	Yes
LT Environmental Incorporated	UT LTE	Utah	Private	No
Utah Department of Environmental Quality	UT DEQ	Utah	State	Yes
Utah State University	UT USU	Utah	Academic	No
Vermont Department of Environmental Conservation	VT DEC	Vermont	State	No
Virginia Department of Environmental Quality	VA DEQ	Virginia	State	Yes
Virginia Department of Health	VA DH	Virginia	State	No
Clallam County Department of Community Development	WA CC DCD	Washington	County or subcounty	No
Colville Confederated Tribes	WA TCC	Washington	Tribe	No
Jamestown Sklallam Tribe	WA TJS	Washington	Tribe	No
Kalispel Natural Resources Department Tribal	WA TK NRD	Washington	Tribe	No
Lower Elwha Klallam Tribe	WA TLEK	Washington	Tribe	No
LummiNation	WA TLN	Washington	Tribe	No
Makah Nation	WA TMN	Washington	Tribe	No
Makah Tribe	WA TMT	Washington	Tribe	No
Midnite Mine Environmental Data	WA MMED	Washington	Federal	No
Muckleshoot Indian Tribe	WA TMIT	Washington	Tribe	No
Nisqually Water Quality	WA TNQ	Washington	Tribe	No
Nooksack Indian Tribe	WA TN	Washington	Tribe	No
Port Gamble S'Klallam Tribe	WA TPGS	Washington	Tribe	No
Puyallup Tribe of Indians	WA TP	Washington	Tribe	No
Quileute Natural Resources	WA TQNR	Washington	Tribe	No
Quinault Indian Nation	WA TQIN	Washington	Tribe	No
Samish Indian Nation	WA TSIN	Washington	Tribe	No
Shoaltwater Bay Tribe	WA TSB	Washington	Tribe	No
Skagit River Watershed Grant	WA TSRWG	Washington	Tribe	No
Skokomish Indian Tribe	WA TSKO	Washington	Tribe	No
Snoqualmie Tribe Environmental and Natural Resources Department	WA TS	Washington	Tribe	No
Spokane Tribe of Indians	WA TSPO	Washington	Tribe	No
Squaxin Island Tribe	WA TSQU	Washington	Tribe	No
Stillaguamish Tribe	WA TSTILL	Washington	Tribe	No
Suquamish Tribe	WA TSUQUA	Washington	Tribe	No
Swinomish Indian Tribal Community	WA TSWIN	Washington	Tribe	No
Tacoma-Pierce County Health Department	WA TPDH	Washington	County or subcounty	No
TerraGraphics Environmental Engineering, Inc.	WA TEE	Washington	Private	No
Thurston County Environmental Health	WA TCEH	Washington	Private	No
Tulalip Tribes	WA TT	Washington	Tribe	No
Upper Skagit Tribe	WA TUS	Washington	Tribe	No
Washington State Department of Ecology	WA SDE	Washington	State	Yes
Whatcom County Public Works	WA WCPW	Washington	State	No

Table 1. Sources of water-quality data included in the initial data compilation and used in the trend analysis.—Continued

[NGO, non-governmental organization; --, not applicable]

Agency name	Agency abbreviation	State	Agency type	Included in the trend analysis[1]
Yakama Nation Environmental Management Program	WA TYNEMP	Washington	Tribe	No
West Virginia Department of Environmental Protection	WV DEP	West Virginia	State	Yes
Green Bay Metropolitan Sewerage District	WI GBMSD	Wisconsin	County or subcounty	No
Bad River Tribe	WI TBR	Wisconsin	Tribe	No
Forest County Potawatomi Community	WI TFCPC	Wisconsin	Tribe	No
Ho-Chunk Nation	WI THCN	Wisconsin	Tribe	No
Lac Courte Oreilles Band of Lake Superior Chippewa Indians of Wisconsin	WI TLCORBLSCI	Wisconsin	Tribe	No
Lac du Flambeau Band of Lake Superior Chippewa Indiana Water Program	WI LDF	Wisconsin	Tribe	No
Oneida Tribe of Wisconsin	WI TON	Wisconsin	Tribe	No
Saint Croix Chippewa Indians of Wisconsin	WI TSCCI	Wisconsin	Tribe	No
Sokaogon Chippewa Community	WI TSKCC	Wisconsin	Tribe	No
Stockbridge-Munsee Community	WI TSMC	Wisconsin	Tribe	No
Wisconsin Department of Natural Resources	WI DNR	Wisconsin	State	Yes
Wisconsin Volunteer Stream Monitoring	WI VOL	Wisconsin	Volunteer	No
Wind River Environmental Quality Commission	WY WREQC	Wyoming	County or subcounty	No
Wyoming Department of Environmental Quality	WY DEQ	Wyoming	State	No
Wyoming Popo Agie Conservation District	WY PACD	Wyoming	State	No
Delaware River Basin Commission	RG4 DRBC	--	Regional	No
Guam Environmental Protection Agency	US GUAM	--	State	No
International Boundary and Water Commission - Water Quality	IBWC	--	International	No
Katrina Response Special Projects	RG1 KRSP	--	Federal	No
Missouri River Public Water Supplies Association	RG4 MRPWSA	--	Regional	No
New York Interstate Environmental Commission	RG3 NYIEC	--	Regional	No
Ohio River Valley Water Sanitation Commission (ORSANCO)	RG8 ORSANCO	--	Regional	Yes
Potomac Appalachian Trail Club	RG2 PATCM	--	Volunteer	No
Saint Joseph River Watershed Initiative	RG3 SJRWI	--	Regional	No
Susquehanna River Basin Commission	SRBC	--	Regional	Yes
Tri-State Water Quality Council	RG3 TSWQC	--	State	No
U.S. Army Corps of Engineers	US ACE	--	Federal	No
U.S. Bureau of Land Management	US BLM	--	Federal	No
U.S. Bureau of Reclamation	US BR	--	Federal	Yes
U.S. Department of Agriculture	US DA	--	Federal	No
U.S. Department of Agriculture, Agriculture Research Service	US DA ARS	--	Federal	No
U.S. Department of Defense	US DOD	--	Federal	No
U.S. Department of Energy	US DOE	--	Federal	No
U.S. Environmental Protection Agency	US EPA	--	Federal	Yes
U.S. Fish and Wildlife Service	US FWS	--	Federal	No
U.S. Forest Service	US FS	--	Federal	No
U.S. Geological Survey	USGS	--	Federal	Yes
U.S. National Estuarine Research Reserve System	US NERRS	--	Federal	No
U.S. National Park Service	US NPS	--	Federal	Yes
U.S. Office of Surface Mining	US OSM	--	Federal	No
Upper Midwest Environmental Sciences Center	RG5 UMESC	--	Federal	No

[1] All pesticide and ecology data used in the trend analysis were from the U.S. Geological Survey.

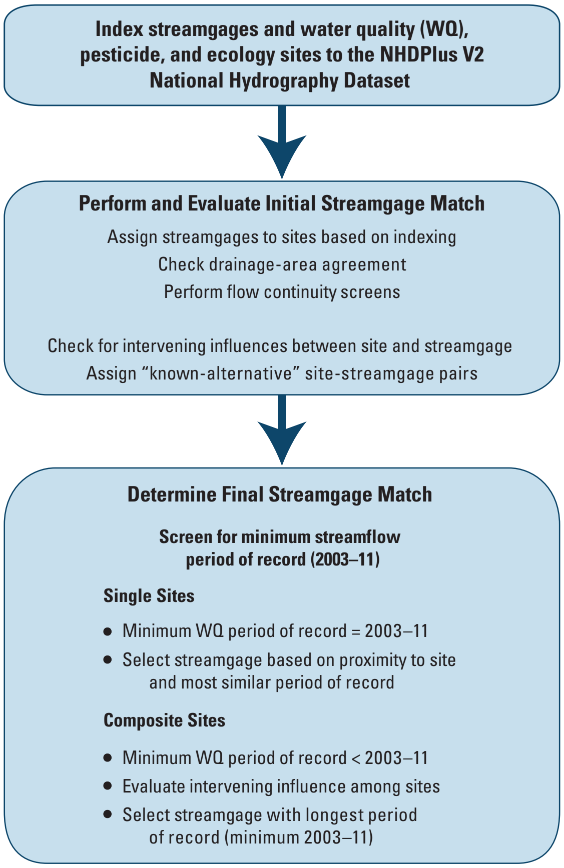


Figure 1. General steps for streamgage matching—Assigning water-quality, pesticide, and ecology sites to streamgages for use in trend analysis.

Indexing Water-Quality Sites and Streamgages to the National Hydrography Dataset Network

The medium resolution National Hydrography Dataset (NHDPlus)V2 (McKay and others, 2012) was used as the geospatial framework to pair water-quality sites and streamgages and to identify groups of water-quality sites that were geographically close enough to consider combining their water-quality records (referred to as "compositing"). In the NHDPlus, line segments termed "flowlines" represent portions of streams. These are in concept "stream reaches," and typically span the distance between two inflowing tributaries or between a tributary and a confluence with a larger stream. Each flowline is associated with a unique identifier (ComID). The ComID serves as the primary key in the streamgage matching and compositing processes.

The term "indexing" refers to the process of matching water-quality sites and streamgages to a stream flowline in the NHDPlus. Indexing streamgage and water-quality sites to flowlines in the NHDPlus is the critical first step of the streamgage-matching and compositing processes. Because USGS streamgage locations are already a part of the NHDPlus dataset, no additional work was necessary to index USGS streamgages to a ComID.

As part of the data compilation for this study, a total of 2,911 non-USGS streamgages were identified, which spanned 15 States and were operated by six different agencies including the U.S. Army Corps of Engineers, the International Boundary and Water Commission, the Colorado Division of Water Resources, the Oregon Water Resources Department, the Tennessee Valley Authority, and the Bureau of Reclamation. The non-USGS streamgages were identified based on previous USGS studies (Hooper and others, 2001; Ryberg and others, 2014; Chanat and others, 2016) and information from USGS Water Science Centers about the operation of streamgages by other agencies in their State. The locations of non-USGS streamgages were obtained from their respective source agencies, manually verified for accuracy, corrected if necessary, and indexed to the NHDPlus hydrologic network, thereby assigning the streamgage a ComID.

The water-quality site-location information—latitude, longitude, and name—was acquired from a large number of sources and agencies and varied in precision and accuracy. Inspections of the original location information revealed an unacceptable number of water-quality sites that were located on or near the wrong stream flowline. To ensure water-quality sites were indexed with the correct stream flowline, an ARCGIS NHDPlus version 2 indexing tool "Reference Points to Flowlines" (Laura Hayes and C.M. Johnston, U.S. Geological Survey, written commun., 2015) was executed for all water-quality sites that identified the most likely flowline for each water-quality site, based on the distance between the point (water-quality site) and a flowline and the stream name when available. If needed, the indexing algorithm assigned a new latitude/longitude, moved the point to a new, more probable location on the NHDPlus flowline, and reported a quality assurance/quality control score referred to as the "qaqc_score." A high qaqc_score indicated a greater likelihood of an indexing problem; a score of 0 indicated a high degree of confidence in the final indexing result. Possible causes of a high qaqc_score include: a mismatch in stream name in the site name and the stream name for the flowline, if the distance from the site to the flowline was greater than 250 meters, if there was a mismatch in NHDPlus ComID between the site location's catchment and the flowline's catchment, if the flowline was a Coastline, or if the site was within a 2-dimensional polygon river area in the NHDPlus and a higher order stream existed within a buffer area of the flowline location, indicating a possible erroneous match to a tributary stream.

Once a ComID was associated with a water-quality site, other attributes of interest from the NHDPlus were associated with the water-quality site, such as drainage area and LevelPathID, which identifies the collection of flowlines that compose the mainstem of a stream and is used for streamgage matching.

After screening water-quality sites for a minimum level of data availability (3 consecutive years of quarterly data), an attempt was made to index 21,692 water-quality sites to the NHDPlus. Approximately 5 percent could not be associated with a NHDPlus flowline (no index), 15 percent had a qaqc_score greater than 0, and 80 percent had a qaqc_score of 0. To evaluate the success of the indexing algorithm and the reliability of the qaqc_score, a random sample of approximately 300 indexed water-quality sites was manually checked. The sample included water-quality sites from each of the previous three groups: (1) no index, (2) qaqc_score greater than 0, and (3) qaqc_score of 0. Among the no-index water-quality sites in the random sample, none were meaningfully related to the NHDPlus. All either had incorrect latitude/longitude coordinates or were not suitable sites for determining water-quality trends in rivers and streams (for example, located on buoys in bays or estuaries or on intermittent streams). All water-quality sites that failed to be associated with the NHDPlus during the initial run of the indexing algorithm were eliminated from further consideration. Among the remaining randomly checked sites, the quality of the index process was directly related to the qaqc_score: (1) those with scores of 0 were virtually always correct, (2) those with scores of 1 or 2 were sometimes correct, and (3) those with scores greater than 2 were almost always incorrect. Based on these scores, it was decided to accept the index result for water-quality sites when the qaqc_score was 0. Of the remaining sites (qaqc_score greater than 0), approximately 2,000 were manually checked and approximately 1,400 of these were identified as being correctly indexed or were manually moved to a new location and reindexed to the correct ComID. The result of this entire process was a set of 18,786 water-quality sites assigned to an NHDPlus flowline and ComID.

Identifying Groups of Near-Colocated Sites

For sites that had an insufficient record of water-quality data for trend analysis (see the “Final Screening for Data Coverage” section), an effort was made to combine the water-quality data from two sites to increase the period of record and (or) sampling frequency to meet the screening criteria (compositing). In order to identify sites that could be composited, all sites located on the same flowline (ComID) of the NHDPlus network and matched with the same streamgage were grouped together. All of these water-quality sites sharing a common ComID and streamgage match were considered for possible combining and were evaluated for intervening influences. While no more than two sites were combined for any parameter, a group of potential composite sites could contain more than two sites. (See the sections “Merging Colocated or Neighboring Water-Quality Sites” for specific parameter groups for details about the selection of composite pairs.)

Streamgage Selection

The trend analysis required that every water-quality, pesticide, and ecology site have an associated continuous streamgage; thus, streamgage matching was performed that paired each water-quality site or group of sites with a streamgage. The goal was to identify the nearest streamgage having a streamflow record that encompassed as much of the water-quality period of record as possible and that did not have intervening influences between the two locations that would cause streamflow to be substantially different. In many cases, a water-quality site could not be matched to a streamgage, and the water-quality site was dropped from further consideration. There were variations in streamgage selection process for pesticide water-quality sites and ecology sites.

The streamgage-matching process consisted of the following broad steps:

1. Initial streamgage match and compilation of water-quality site-streamgage combinations,
2. Evaluating intervening influences between water-quality site-streamgage combinations,
3. Evaluating water-quality site-streamgage combinations using screens for streamflow continuity,
4. Incorporating manual edits, and
5. Selecting the preferred water-quality site-streamgage combinations.

Initial Streamgage Match and Compilation of Water-Quality Site-Streamgage Combinations

The 18,786 indexed water-quality sites were joined with the indexed streamgages (using the NHDPlus LevelPathID) to identify water-quality sites and streamgages that shared the same stream in the NHDPlus network. Joining sites and streamgages by stream (LevelPathID), as opposed to flowline (ComID), allowed matches beyond the immediate reach that could later be filtered by differences in drainage area (described in the following paragraph). The joint result was potentially a one-to-many relation, meaning a single water-quality site could be associated with zero, one, or many streamgages. Of the 18,786 water-quality sites, 15,255 joined to one or more streamgages, resulting in 169,498 potential water-quality site-streamgage combinations. The drainage area associated with each water-quality site and streamgage was also captured from the NHDPlus during this process.

To eliminate combinations not useful for this study, each water-quality site-streamgage combination was evaluated using four screens that assessed drainage area agreement between the site and streamgage and basic streamflow requirements at the matched streamgage. Water-quality site-streamgage combinations meeting all of the following criteria were retained: (1) there was an absolute drainage area difference of less than 10 percent between the water-quality site and the streamgage, (2) streamflow was available beginning on or before December 31, 2002 (the latest start year of a trend period), (3) streamflow was available on or after January 1, 2011 (the earliest end year of a trend period), and (4) the streamflow record length was at least 2,700 days (equivalent to 9 months of streamflow per year over 10 years). Of the 15,255 water-quality sites that joined to one or more USGS streamgages, 6,647 sites passed all four of these criteria, giving 10,518 water-quality site-streamgage combinations.

Other programs and projects, including NAWQA, National Stream Quality Accounting Network (NASQAN), various USGS discrete studies (Hooper and others, 2001; Ryberg and others, 2014; and Chanat and others, 2016), and CBP, have established long-standing water-quality site-streamgage combinations. Some of these known alternative combinations are identical to those derived using the above process; however, other combinations differed or failed one or more of the basic screening criteria described above and have been filtered out. For comparability and consistency with previous trend studies, these known alternative combinations were added to the dataset of water-quality site-streamgage combinations assembled for this study and used as the preferred combination when appropriate. A known alternative combination could still be eliminated from the final dataset for this study if it failed subsequent screenings, such as requirements related to streamflow continuity or number and (or) type of water-quality samples.

Evaluating Intervening Influences Between Water-Quality Site-Streamgage Combinations

Water-quality site-streamgage combinations were manually checked for major intervening influences that might cause the streamflow at the water-quality site to be substantially different from the streamflow at the streamgage location. Please see the “Indentifying Groups of Near-Colocated Sites” section for a description of the intervening influence checking procedure.

Screening Streamflow Record for Streamflow Continuity

The WRTDS model (Hirsch and others, 2010; Hirsch and De Cicco, 2015) used for this trend analysis for nutrients, sediment, major ions, and carbon requires a continuous record of daily streamflow values that are positive and nonzero; therefore, the streamflow records for all streamgages that were matched to water-quality sites were screened for continuity, magnitude, and sign. For each streamflow record, the following information was calculated for each water year: the number of days with missing streamflow, the longest sequence of consecutive days with missing streamflow, the number of days with negative streamflow, and the number of days each year with zero streamflow. Each streamflow record was required to meet all of the following criteria:

- No more than 30 days with missing streamflow per year,
- No more than 3 consecutive days with missing streamflow per year,
- No days with negative streamflow,
- No more than 30 days with zero streamflow per year,
- First complete water year is 2003 or earlier, and
- Last complete water year is 2011 or later.

In some cases, a streamflow record failed to meet all criteria; however, it could be made usable by deleting problematic parts in the older period of record. In those cases, the longest usable record was maintained for the site. For example, if a streamflow record was complete from water years 1972 through 2012, except for the years 1983–85, the streamflow prior to water year 1986 would not be used.

Because the trend method in this study uses the natural log of streamflow, there could not be any zero streamflow days in the record (Hirsch and De Cicco, 2015). When zero streamflow values were present, a very small constant equivalent to 0.1 percent of the mean streamflow for the site was added to the streamflow value for every day of the record (even when the streamflow was not zero). Final numerical results from the trend models did not have this small streamflow increment subtracted out of streamflow and load results because typically these adjustments would be too small to have any consequence (Hirsch and De Cicco, 2015). To avoid bias in the trend results stemming from this addition, a streamflow was excluded when the daily streamflow record had more than 30 zero streamflow values in any year of the period of record.

Use of the natural log of streamflow is also problematic with negative streamflow values, which can occur in rivers and streams with backwater conditions and streamflow reversals resulting from tides or wind. Negative streamflow values also add complexity to the calculation of loads at a site. For these reasons, water-quality sites were not matched to streamgages with any negative streamflow values. If another suitable nearby streamgage was not available, the water-quality site was excluded.

At some of the streamgages, there were periods of time with missing daily streamflow values. The WRTDS model will not execute if a streamflow record contains missing values; missing values will cause the model to fail and need to be addressed outside of the model. Missing values occur for a variety of reasons. Small gaps, such as those permitted by the above criteria, are usually caused by instrument failure and can be reasonably estimated from nearby (in time) values. Missing streamflow values were estimated using the R function “fillMiss” available in the R package “waterData” (Ryberg and Vecchia, 2012). The fillMiss function uses a structural time series to estimate missing streamflow values. Structural time series models decompose a time series in linear combinations of components and are useful for simulating datasets having superimposed data of varying frequencies. Following estimation, the filled data are smoothed using a state-space model. Default values for the function were used as described in Ryberg and Vecchia (2012). The function was tested on streamflow records with artificially deleted values and found to produce reasonable results. Streamflow screens for gages used with the pesticide water-quality sites and the ecology sites were different from those used with water-quality sites being analyzed with WRTDS. See the section "Variations in the Streamgage Selection Process for Pesticide Water-Quality Sites and Ecology Sites" for details on streamflow screens used for these sites.

Selecting the Preferred Water-Quality Site-Streamgage Combinations

For water-quality sites being analyzed with WRTDS, lists of candidate water-quality site-streamgage combinations were developed from the screening process described in steps 1–4 listed at the beginning of the "Streamgage Selection" section. For all water-quality site-streamgage combinations, the streamgage must have screened streamflow records from at least 2003 to 2011 and no intervening influences between the water-quality sites and streamgage. The individual water-quality sites in a composite group also must have the same ComID and be matched to the same streamgage, and there must be no intervening influences between the individual water-quality sites that make up the group. These additional screens reduced the 10,518 possible water-quality site-streamgage combinations to 3,247. There were 3,247 candidate single water-quality site-streamgage combinations and 1,589 candidate composite water-quality site-streamgage combinations. These combinations included 2,793 unique single water-quality sites and 1,335 potential composite water-quality site groups. These candidate water-quality site-streamgage combinations were also used for selecting gages for the pesticide water-quality sites and the ecology sites, though there were different considerations for streamgage selection at these sites (see the "Variations in the Streamgage Selection Process for Pesticide Water-Quality Sites and Ecology Sites" section).

Inconsistencies and errors in the information about a streamgage, water-quality site, or water-quality site-streamgage combination were periodically encountered, which resulted in multiple iterations through the streamgage-matching process. As part of this process, many of the problematic water-quality site-streamgage combinations were individually investigated and information about the combination was updated. Common issues that triggered investigations included (1) incorrectly indexed water-quality sites and streamgages resulting in incorrect drainage areas and ComIDs, (2) identification of a previously unrecognized intervening influence between a water-quality site and streamgage or between individual water-quality sites in a group, and (3) incorrect streamgage location (uncommon). The results of these individual site investigations were used to update the list of candidate water-quality site-streamgage combinations.

For single water-quality sites matched with only one candidate streamgage, that combination was selected. If a water-quality site was matched to more than one candidate streamgage, then the closest streamgage (smallest difference in drainage areas) having a period of record that most completely encompassed the period of record of the water-quality site was selected. For example, if a water-quality site had data from 1989 to 2012 and was paired with two candidate streamgages, the closer streamgage having a streamflow record from 2000 to 2012 and the more distant streamgage having a streamflow period from 1987 to 2012, then the more distant streamgage would be selected. If the streamflow record at both streamgages encompassed the water-quality record, then the closer streamgage would be selected. This list of screened single water-quality site-streamgage combinations (number [n] equals [=] 2,793) was then used to pair single water-quality sites with a streamgage after data processing for final screening of data coverage for WRTDS analyses.

The list of screened potential composite water-quality site groups-streamgage combinations was developed separately from the single water-quality site-streamgage combinations because of the slightly different method used for streamgage selection. For potential composite water-quality site groups matched with only one candidate streamgage, that combination was selected. If a potential composite water-quality site group was matched to more than one candidate streamgage, then the streamgage with the longest period of record was selected, regardless of proximity. This list of screened potential composite water-quality site group-streamgage combinations (n=1,335 potential composite water-quality site groups, which includes 3,191 individual sites) was then used to pair potential composite water-quality sites with a streamgage for final screening of data coverage for WRTDS analyses. Any water-quality site (single or composite) not paired with a streamgage was not considered for trend analysis.

Variations in Streamgage Selection Process for Pesticide and Ecology Sites

Trend analysis of pesticide and ecology data did not use the WRTDS model; thus, there were different streamflow requirements for pairing these data with streamgages.
For the water-quality sites with pesticide data, potentially useable streamgages were identified using steps 1-2 in the "Streamgage Selection" section. The resulting combinations included streamgages with complete streamflow records and others with potentially problematic streamflow records (streamgage did not pass streamflow continuity screens).
The streamflow data at these potentially useable streamgages (n=98) were reassessed because the method for pesticide concentration trend analysis, SEAWAVE–Q (Vecchia and others, 2008; Ryberg and Vecchia, 2013), requires a complete set of daily streamflow observations as well as a leading year

of observations to calculate long-term streamflow anomalies; the method can allow slightly more consecutive days with missing flows than WRTDS. If the streamgage was missing more than 0.5 percent of the daily values, then the streamgage was dropped from further consideration. If a streamgage was missing less than 0.5 percent of the daily values, then the missing values were estimated using the fillMiss function of the R package waterData (Ryberg and Vecchia, 2012). Also, because SEAWAVE–Q uses logarithms during model estimation, streamgages with more than 0.5 percent of the observed daily streamflows recorded as zeroes were excluded. For streamgages with less than 0.5 percent zero flows, on days when streamflow was reported as zero, the streamflow was set to the minimum recorded value for that particular stream (0.01 ft^3/s in all cases). Streamgages used in the pesticide trend assessment are listed in appendix 1, table 1–2.

For sites with ecology data, the sites and streamgages were indexed as described above, but the streamgage-matching process was primarily manual. Similar to the pesticide sites, the method to evaluate trends for ecology sites required streamflow records that began prior to the start of the ecology period of record to determine antecedent streamflow conditions. All but four of the ecology sites have a colocated streamgage. The four sites without colocated streamgages use data from a nearby streamgage that is within 1 mile upstream or downstream from the ecology site (appendix 1, table 1–3). Flows were estimated for streamgages with missing streamflow data using streamflow information from nearby streamgages (appendix 1, table 1–4). For specifics about the streamgage selection process and streamflow estimation at ecology sites see the “Antecedent Hydrology” section of the “Ecology Data Preparation” section of this report. Ecology sites with colocated streamgages (with complete streamflow records) are listed in appendix 1, table 1–5.

Scaling Streamflow Records

The streamgage-matching process for this study was designed to identify colocated or nearby representative streamgages for each water-quality monitoring site by optimizing the physical proximity of the water-quality site and streamgage and the temporal overlap between the water-quality data and streamflow data (see the “Streamgage Matching” section). The streamgage-matching process only allowed water-quality site-streamgage combinations when the difference in drainage area between the water-quality site and streamgages was less than 10 percent. This process resulted in some water-quality site-streamgage combinations that were colocated (water-quality and streamflow data collected at the same location) and other combinations when the water-quality data and streamflow data were collected at different but representative locations. As discussed in the “Initial Streamgage Match and Compilation of Water-Quality Site-Streamgage Combinations” and “Evaluating Intervening Influences Between Water-Quality Site-Streamgage Combinations” sections, if this difference in drainage area was less than 10 percent, and there were no major intervening influences affecting the variability of streamflow between the locations, it was assumed that the relative, day-to-day variability in streamflow was the same at the water-quality site and streamgage. When the site and streamgage are not colocated, the streamgage drains a land area that differs in size from the water-quality site, and the magnitude of the streamflow would be expected to be different between the two locations, and the load estimates from WRTDS would likely be slightly larger or smaller than anticipated because of this difference. To acquire more accurate estimations of load from WRTDS, daily streamflow data were scaled prior to model calibration for water-quality sites that were not colocated with their paired streamgage.

If any of the following criteria were met, a water-quality site-streamgage combination was considered to be colocated and streamflow data were not scaled:

- Water-quality site and streamgage had the same identification number,
- Water-quality site and streamgage shared an Umbrella Site Identifier (ID),
- Water-quality site and streamgage had the same comID, and the distance between locations was less than or equal to 1,000 meters (m),
- Percentage difference in drainage area between the water-quality site and streamgage was less than 1 percent, or
- Water-quality site-streamgage combination was a known alternative combination used by other USGS programs and projects.

Flow scaling occurred for water-quality site-streamgage combinations that did not meet any of these criteria but still had a difference in drainage area of less than 10 percent. Streamflows were scaled by multiplying each daily streamflow value by the ratio of the water-quality site-drainage area to the streamgage-drainage area.

For the most part, the water-quality site-drainage area was used for the streamgage-drainage area, except when a water-quality site-streamgage combination had different identification numbers and were on different comIDs or had different identification numbers and were on the same comID but more than 1,000 m apart. For these instances, the drainage area of the streamgage was calculated as the drainage area of the water-quality site plus or minus the difference in NHDPlus streamflow accumulation grids (McKay and others, 2012) between the two locations. These updated streamgage-drainage areas were used to assess if the water-quality site and streamgage were colocated and were appropriate for streamflow scaling. For details on the delineation of basin boundaries and determination of drainage areas for the water-quality sites and streamgages see Falcone and others (2017).

Water-Quality Data Preparation

The initial pool of sites for this study included over half a million sites. Because of the large number of sites, it was not possible to tailor data-processing choices to each individual site; instead, data-processing decisions were made at a high level and applied consistently to all sites. This approach avoided introducing differences in methodology that could invalidate comparisons of the final trend results across nationwide sites and parameters. A single set of decisions appropriate for all sites was not possible because of the wide variation in data characteristics and quality across sites. In the end, some sites that could have been salvaged by making individual site-based data processing and screening choices were excluded from the study.

The high-level decisions made in this study provide a robust approach to data processing and screening when the goal is to compare trends regionally and nationally at a large number of sites and for a large number of parameters. When the goal is to compare trends at a smaller subset of sites and for a smaller number of parameters, different choices could be possible—either because consistency can be achieved in a different manner or because it is reasonable in scope to tailor data-processing choices to each site. In these situations, trend results that are slightly different in magnitude and (or) uncertainty than those from this study could be obtained for the same site, parameter, and trend period, but any differences should be minor when the underlying data describe the full range of conditions at a site and when similar and appropriate trend methods are used. The USGS conducts numerous trend studies throughout the Nation using smaller subsets of sites. Some examples include studies of nutrients and sediment trends in tributaries to Chesapeake Bay (Chanat and others, 2016); nutrient, organic carbon, and chloride trends in tributaries to Long Island Sound (Mullaney, 2016); and chloride trends in snow-affected urban streams of the United States (Corsi and others, 2015). The generalized process for preparing and screening nutrient, sediment, major ion, salinity, and carbon data is presented in figure 2*A*.

A

Nutrient, Sediment, Major Ion, Salinity, and Carbon Parameters

Paired Sample Evaluation

Evaluate the suitability of
combining fractions and parameters

Data Harmonization

- Select and (or) combine parameters of interest
- Select and (or) combine fractions of interest
- Remove non-discrete samples
- Remove samples analyzed in the field
- Harmonize remark codes and remove rejected/poor-quality values
- Adjust data based on historical memoranda
- Recensor U.S. Geologcial Survey data
- Remove duplicate samples
- Evaluate zero, negative, and missing concentration values
- Calculate target parameters

Trend Screening

- Streamgage selection
- Evaluate coverage throughout target trend periods
- Merge suitable co-located or neighboring water-quality sites
- Evaluate coverage of all seasons
- Evaluate coverage of all streamflows
- Evaluate the amount of censored data
- Remove outliers likely to be typographical errors
- Thin data to 1 sample per day

B

Pesticide Parameters

Data Preparation

- Select parameters of interest
- Adjust data based on recovery
- Recensor data
- Thin data to 1 sample per week
- Evaluate the amount of censored data

Trend Screening

- Evaluate coverage throughout target trend periods
- Evaluate coverage of all seasons
- Evaluate the availability of streamflow data and select streamgage

C

Ecology Parameters

Data and Trend Screening for Site and Sample Selection

- Evaluate coverage of fish, invertebrate, and algae data throughout target trend periods
- Select wadeable sites
- Streamgage selection
- Harmonize taxonomy and resolve ambiguous taxa
- Calculate ecological endpoints
- Calculate associated environmental variables

Figure 2. General steps for data processing and trend screening for *A*, nutrient, sediment, major ion, and carbon parameters; *B*, pesticide parameters; and *C*, ecology parameters.

Paired Sample Evaluation

Similar water-quality parameters that were categorized separately in the original data sources were evaluated for their potential to be combined into a single dataset for use in trend analysis. The goals of combining parameters were to (1) increase the number of sites for which trends can be calculated, and (2) increase the length of time over which trends can be calculated. These goals had to be balanced against the possibility of increasing variability in the data. Trend analysis evaluates whether there is a pattern underlying the variability of the data over time, and increasing data variability by combining different fractions of the same parameter or different parameters can hinder trend determination. Generally, the filtered fraction and the unfiltered or unknown fraction of the same parameter were evaluated for combination, for example, filtered ammonia and unfiltered ammonia; however, some comparisons were made between different parameters, for example, filtered nitrite plus nitrate and filtered nitrate. Whenever the preferred fraction comprised less than 80 percent of the total number of samples available for that parameter, the suitability of combining one or more nonpreferred fractions with the preferred fraction was evaluated. A complete list of evaluated parameters and fractions is provided in table 2.

When parameters were evaluated, one member was designated the preferred parameter and the other designated the nonpreferred parameter. The filtered fraction was generally preferred over total, unknown, or unfiltered fractions, except for total phosphorus and Kjeldahl nitrogen, when the unfiltered fraction was preferred. If two parameters were deemed sufficiently similar to be combined, and both the preferred and nonpreferred parameters were available for a sample day, the value for the preferred parameter was selected during trend analysis

An analysis of paired samples was conducted on the parameters being evaluated. Paired samples are those when the preferred and nonpreferred parameters were both analyzed in a sample collected at the same site on the same date and time. If necessary, the two values in each pair of samples were rounded to have the same number of significant digits, and all paired samples within a parameter group were recensored to a common value. The two evaluated parameters often contained a number of different censoring levels; the common censoring level was chosen as the 90th percentile of all censoring levels in a parameter group. Sample pairs when one or both values were censored at a value greater than the common censoring level were excluded from further analysis. The remaining censored values were recoded to the common censoring level, and uncensored values less than the common censoring level were recoded as censored values at the common censoring level.

The paired samples were evaluated using a weight-of-evidence approach that utilized a quantitative statistical test and a qualitative evaluation of the percentage difference between paired samples in the group. The parameter comparisons were evaluated using all paired data from all sites and on a site-by-site basis. All evaluations were conducted on the consistently rounded, recensored paired data. A summary of the evaluation for each parameter, following the process described here, is presented in table 2:

- Perform the nonparametric sign test on paired data across all sites. The sign test is used to test for consistent differences between pairs of observations. The sign test was implemented as described in Helsel (2005).
- Perform the sign test for each site that had 10 or more pairs of data (not shown table 2).
- Generate a subset of the paired data containing only pairs when both values are uncensored.
- Perform the sign test on the uncensored paired data across all sites.
- Perform sign test for each site that had at least 10 uncensored pairs of data (not shown table 2).
- Calculate the percentage difference (pctDiff) between the preferred and nonpreferred values for uncensored paired data as:

$$pctDiff = 100 \times \left(\frac{C_{nonpref} - C_{pref}}{C_{pref}} \right) \quad (1)$$

where

$C_{nonpref}$ is the concentration of the nonpreferred parameter, and

C_{pref} is the concentration of the preferred parameter.

- Determine the percentage of paired data values that had (1) pctDiff values less than or equal to (≤) -20 percent, (2) pctDiff values greater than or equal to (≥) 20 percent, and (3) pctDiff values between -20 percent and 20 percent. It was assumed that laboratory and environmental variability typically accounts for less than 20 percent of the observed intersample variability, a threshold supported by several studies of replicate sample quality assurance data collected across the country (Rinella and Janet, 1998; Maluk, 2000; Kent and Belitz, 2004).

The data available for paired-sample evaluations are limited in time, space, and by the collecting agency. Paired samples generally were collected for a limited duration, often during times when analytical methods were changing or new methods being developed. Paired samples might have been collected to answer a specific question being posed at a limited number of sites or by questions arising within a specific agency. For these reasons, and likely others not mentioned, the evaluations represent an imperfect window on the data. For example, it was not possible to evaluate the temporal continuity in the relation between the paired data at a site, and

Table 2. Preferred and nonpreferred parameters, fractions, and summary of the paired sample evaluation metrics.

[n, number; pctDiff, percentage difference; ≥, greater than or equal to; %, percent; ≤, less than or equal to]

				All data					
Preferred parameter	**Preferred fraction**	**Nonpreferred parameter**	**Nonpreferred fraction**	**Number of agencies**	**Number of sites**	**Number of pairs**	**Sign test p-value**	**Number of sites with sign test (n pairs ≥ 10)**	**Number of sites with sign test p-value ≤ 0.1**
Ammonia	Filtered	Ammonia	Unfiltered	13	1,927	63,200	1	1,212	52
Ammonia	Filtered	Ammonia	Total	12	163	8,960	1	127	17
Ammonia	Filtered	Ammonia	Unknown	4	123	6,878	1	42	4
Kjeldahl nitrogen	Unfiltered	Kjeldahl nitrogen	Filtered	13	2,401	109,293	0	1,683	998
Nitrate	Filtered	Nitrate	Unfiltered	5	692	6,324	1	165	27
Nitrate	Filtered	Nitrate	Total	5	85	1,928	1	35	3
Nitrite	Filtered	Nitrite	Unfiltered	8	1,181	20,488	1	623	40
Nitrite	Filtered	Nitrite	Total	5	84	1,561	1	32	8
Nitrite	Filtered	Nitrite	Unknown	1	17	82	1	1	0
Nitrite plus nitrate	Filtered	Nitrite plus nitrate	Unfiltered	11	1,900	60,472	1	1,328	15
Nitrite plus nitrate	Filtered	Nitrite plus nitrate	Total	9	279	3,983	1	43	11
Nitrite plus nitrate	Filtered	Nitrate	Filtered	20	2,297	97,223	0.0035	1,418	456
Orthophosphate	Filtered	Orthophosphate	Unfiltered	8	2,562	51,944	1	1,420	381
Total phosphorus	Unfiltered	Total dissolved phosphorus	Filtered	20	3,557	270,938	0	2,880	2,016
Suspended sediment concentration	Suspended	Total suspended solids	Nonfilterable	1	160	5,877	1	99	28
Suspended sediment concentration	Suspended	Total suspended solids	Suspended	1	5	46	0.9605	3	0
Suspended sediment concentration	Suspended	Total suspended solids	Unknown	3	777	17,601	0	336	156
Chloride	Filtered	Chloride	Unfiltered	1	22	1,473	1	13	0
Chloride	Filtered	Chloride	Total	7	84	4,216	1	60	0
Chloride	Filtered	Chloride	Unknown	1	369	2,715	0	114	13
Sulfate	Filtered	Sulfate	Unfiltered	1	40	689	0.2231	23	5
Sulfate	Filtered	Sulfate	Total	5	22	96	1	3	0
Sulfate	Filtered	Sulfate	Unknown	1	1	8	--	0	0
Alkalinity	Filtered	Alkalinity	Unfiltered	2	1,300	49,160	0	859	359
Alkalinity	Filtered	Alkalinity	Total	3	208	3,647	1	70	3
Alkalinity	Filtered	Alkalinity	Unknown	1	31	5,751	1	31	0

				Uncensored pairs of data									
Preferred parameter	Preferred fraction	Nonpreferred parameter	Nonpreferred fraction	Number of pairs	Percent of uncensored pairs of data	Sign test p-value	Number of sites with sign test (n pairs ≥ 10)	Number of sites with sign test p-value ≤ 0.1	Number of samples with pctDiff ≥ 20%	Number of samples with pctDiff ≤ -20%	Percentage of uncensored pairs with pctDiff within ≥ 20%	Percentage of uncensored pairs with pctDiff ≤ -20%	Percentage of uncensored pairs with pctDiff within +/- 20%
Ammonia	Filtered	Ammonia	Unfiltered	38,289	60.6	1	954	60	9,616	4,034	25.1	10.5	64.4
Ammonia	Filtered	Ammonia	Total	4,925	55.0	1	107	28	826	360	16.8	7.3	75.9
Ammonia	Filtered	Ammonia	Unknown	274	4.0	1	8	3	5	60	1.8	21.9	76.3
Kjeldahl nitrogen	Unfiltered	Kjeldahl nitrogen	Filtered	74,705	68.4	0	1,365	1,067	2,544	41,256	3.4	55.2	41.4
Nitrate	Filtered	Nitrate	Unfiltered	2,751	43.5	0.0197	74	32	73	9	2.7	0.3	97.0
Nitrate	Filtered	Nitrate	Total	1,695	87.9	1	35	3	61	41	3.6	2.4	94.0
Nitrite	Filtered	Nitrite	Unfiltered	7,472	36.5	0	211	55	3,214	271	43.0	3.6	53.4
Nitrite	Filtered	Nitrite	Total	33	2.1	0.0191	1	0	8	0	24.2	0.0	75.8
Nitrite	Filtered	Nitrite	Unknown	18	22.0	1	1	0	0	0	0.0	0.0	100.0
Nitrite plus nitrate	Filtered	Nitrite plus nitrate	Unfiltered	48,376	80.0	1	1,082	16	2,587	1,027	5.3	2.1	92.5
Nitrite plus nitrate	Filtered	Nitrite plus nitrate	Total	3,175	79.7	0.0105	42	21	219	67	6.9	2.1	91.0
Nitrite plus nitrate	Filtered	Nitrate	Filtered	76,872	79.1	0	1,267	562	168	4,342	0.2	5.6	94.1
Orthophosphate	Filtered	Orthophosphate	Unfiltered	38,879	74.8	1	1,076	311	9,474	1,558	24.4	4.0	71.6
Total phosphorus	Unfiltered	Total dissolved phosphorus	Filtered	149,567	55.2	0	2,115	1,950	2,359	110,483	1.6	73.9	24.6
Suspended sediment concentration	Suspended	Total suspended solids	Nonfilterable	2,932	49.9	0	64	45	157	1,430	5.4	48.8	45.9
Suspended sediment concentration	Suspended	Total suspended solids	Suspended	46	100.0	0.9605	3	0	6	3	13.0	6.5	80.4
Suspended sediment concentration	Suspended	Total suspended solids	Unknown	10,854	61.7	0	226	160	1,095	5,550	10.1	51.1	38.8
Chloride	Filtered	Chloride	Unfiltered	1,473	100.0	1	13	0	0	0	0.0	0.0	100.0
Chloride	Filtered	Chloride	Total	3,466	82.2	1	53	0	62	15	1.8	0.4	97.8
Chloride	Filtered	Chloride	Unknown	2,715	100.0	0	114	13	215	584	7.9	21.5	70.6
Sulfate	Filtered	Sulfate	Unfiltered	671	97.4	0.0823	20	5	2	3	0.3	0.4	99.3
Sulfate	Filtered	Sulfate	Total	91	94.8	1	3	0	1	0	1.1	0.0	98.9
Sulfate	Filtered	Sulfate	Unknown	8	100.0	--	0	0	--	--	--	--	--
Alkalinity	Filtered	Alkalinity	Unfiltered	49,160	100.0	0	859	359	3,604	1,076	7.3	2.2	90.5
Alkalinity	Filtered	Alkalinity	Total	3,647	100.0	1	70	3	23	10	0.6	0.3	99.1
Alkalinity	Filtered	Alkalinity	Unknown	5,751	100.0	1	31	0	0	0	0.0	0.0	100.0

it was not possible to confirm the validity of combining data for each agency from whom data were received. Nevertheless, most parameters had tens of thousands of paired samples (table 2); collectively, these data provide excellent spatial, temporal, and agency coverage. Occasionally, a site, region, or agency was found to skew the results of an evaluation, and the results were dismissed because of the overwhelming lack of problems at the other sites. This logic is consistent with other decisions made in other steps of processing and filtering the entire water-quality dataset (see section "Nutrient Parameters of Interest and Harmonization"). When the paired sample evaluations did not show a large magnitude of bias, the data were combined even when there was a temporal pattern in the data availability for each individual parameter. On the other hand, when there was concerning amounts of bias in the paired sample comparison, the data were not combined. It was expected that egregious issues with combining parameters would be caught when evaluating the WRTDS models.

Results and decisions for each parameter group are summarized in the following subsections. Within all the paired-sample discussions, statistical significance for the sign test was defined as a p-value ≤0.1.

Ammonia

Ammonia values with a fraction assigned to filtered were compared separately with ammonia values with fractions assigned to unfiltered, total, and unknown (table 2). The largest number of paired samples was in the parameter group composed of filtered and unfiltered ammonia fractions (n=63,200), followed by the group composed of filtered and total ammonia (n=8,960), and lastly the group composed of filtered and unknown ammonia (n=6,878). Relative numbers of unique sites and agencies follow the same order, with the largest number of each being represented in the group composed of filtered and unfiltered fractions.

About half of all the paired samples in the group composed of filtered and unfiltered fractions and the group composed of filtered and total fractions were uncensored and had percentage differences calculated; only 4 percent of the pairs in the filtered-unknown comparison were uncensored. For all three parameter groups, at least 64 percent of percentage differences of uncensored pairs were within plus or minus (+/-) 20 percent. For the filtered-unfiltered and filtered-total parameter groups, the number of percentage differences greater than 20 percent exceeded the number less than -20 percent by about 2 to 1, which indicates a slight bias toward higher unfiltered and total ammonia values relative to filtered ammonia values. The sample pairs in the filtered-unknown parameter group were heavily biased in the opposite direction with a 12- to 1-bias favoring pctDiff values less than -20 percent. A review of the data indicates many of the biased values are driven by individual sites or agencies and do not reflect the data as a whole.

The sign test on all data pairs and on the uncensored data pairs indicated no statistically significant difference among values within any of the comparisons (table 2). Site-by-site sign tests were not statistically significant at most sites. Most pctDiff values were within +/- 20 percent, and the distribution of pctDiff values for the two most populous parameter groups and at most sites was relatively symmetrical about zero. These results demonstrated that these four fractions could be combined when a longer or denser record was needed.

Kjeldahl Nitrogen

Kjeldahl nitrogen values with a fraction assigned to unfiltered were compared with Kjeldahl nitrogen values with a fraction assigned to filtered (table 2). Total and unknown Kjeldahl nitrogen fractions were not considered because of their ambiguity. The evaluation used 109,293 pairs of data from 2,401 sites and 13 agencies.

About 68 percent of all the paired samples were uncensored and had percentage differences calculated for them. Less than 42 percent of the pctDiff values were within +/- 20 percent and also indicated a significant bias toward lower values for the filtered fraction, with 55 percent of the pairs having a pctDiff value ≤-20 percent and only 3 percent of the pairs having a pctDiff ≥20 percent. A review of individual sites demonstrates a widespread, systematic bias spanning multiple agencies and affecting most sites.

The sign test on all data pairs and on the uncensored data pairs indicated a highly statistically significant difference among values within the Kjeldahl nitrogen parameter group composed of filtered and unfiltered fractions (p=0.0000 for tests on both groups). Highly significant differences were also observed for individual sites. These results demonstrated that these two fractions should not be combined to create a longer or denser record.

Nitrate

Nitrate values with a fraction assigned to filtered were compared separately with nitrate values with fractions assigned to unfiltered and total (table 2). The largest number of paired samples was in the parameter group composed of filtered and unfiltered nitrate fractions (n=6,324), followed by the group composed of filtered and total nitrate (n=1,928). Relative numbers of unique sites follow the same order, and both parameter groups were represented by five agencies.

About 43 percent of the pairs in the parameter group containing unfiltered nitrate and 88 percent of the pairs in the group containing total nitrate were uncensored and had percentage differences calculated for them. Both parameter groups had 94 percent or more of their percentage differences within +/- 20 percent, and neither was substantially biased in one direction.

The sign test on all data for both parameter groups indicated no statistically significant difference among values within their respective group. The uncensored values for the parameter group containing total nitrate also showed no statistically significant difference. The sign test p-value was statistically significant (p=0.0197) for the group containing unfiltered nitrate, indicating a bias in the uncensored nitrate values for this parameter group. Although pervasive, this bias is small; 97 percent of the pairs are within +/- 20 percent of one another. Further, the vast majority of the biased sites are from one agency.

The sign test on all data pairs indicated no statistically significant difference among values within the parameter groups. The sign test on uncensored values in the parameter group containing unfiltered nitrate indicates a small bias, but site-by-site sign tests indicate the issue is largely limited to one agency. Nearly all pctDiff values were within +/- 20 percent for both parameter groups. These results demonstrated that these three fractions could be combined when a longer or denser record was needed.

Nitrite

Nitrite values with a fraction assigned to filtered were compared separately with nitrite values with fractions assigned to unfiltered, total, and unknown (table 2). The largest number of paired samples was in the parameter group composed of filtered and unfiltered nitrite (n=20,488), followed by the group composed of filtered and total nitrite (n=1,561), and lastly the group composed of filtered and unknown nitrite fractions (n=82). Relative numbers of unique sites and agencies follow the same order.

About 36 percent of the pairs in the parameter group containing unfiltered nitrite fractions, 22 percent of the pairs in the parameter group containing unknown nitrite fractions, and 2 percent of the pairs in the group containing total nitrite fractions were uncensored and had percentage differences calculated for them. All three parameter groups had at least half of their percentage differences within +/- 20 percent; however, the filtered-unfiltered and the filtered-unknown groups were both biased toward larger values for the unfiltered fraction as indicated by a large number of pairs having pctDiff values ≥20 percent.

The sign test on all data pairs indicated no statistically significant difference among values within the parameter groups. The uncensored values for the parameter group containing nitrite of unknown fraction also showed no statistically significant difference. The sign test p-value was statistically significant for the parameter groups containing unfiltered nitrite (p=0.0000) and total nitrite (p=0.0191), indicating a bias in the uncensored nitrite values for these parameter groups. In both groups, the pctDiff values ≥20 percent exceed the number of pctDiff values ≤-20 percent by a large amount and indicate the uncensored nitrite values are biased toward higher unfiltered and total fractions. The magnitude of the bias is small, and for the purposes of this study, negligible because trends were not determined for nitrite alone, but only in conjunction with nitrate (nitrite + nitrate); in most samples, the concentration of nitrate exceeds that of nitrite by at least one order of magnitude. Additionally, many of the biased sites are from one agency.

The sign test on all data pairs indicated no statistically significant difference among values within the parameter groups. The sign test on uncensored values in the parameter groups containing unfiltered and total nitrite indicate a large percent bias, but the magnitude of this bias is small given the generally small concentrations of nitrite detected. This represents less than 60 percent of all data pairs due to the highly censored nature of the nitrite data. Considering just the uncensored values, more than half of the pctDiff values were within +/- 20 percent for all three parameter groups. These results demonstrated that these four fractions could be combined when a longer or denser record was needed.

Nitrite Plus Nitrate

Nitrite plus nitrate values with a fraction assigned to filtered were compared separately with nitrite plus nitrate values with fractions assigned to unfiltered and total (table 2); filtered nitrite plus nitrate values also were compared with filtered nitrate values (table 2). The largest number of paired samples was in the parameter group composed of filtered nitrite plus nitrate and filtered nitrate (n=97,223), followed by the group composed of filtered and unfiltered nitrite plus nitrate fractions (n=60,472), and lastly by the group composed of filtered and total nitrite plus nitrate (n=3,983). Relative numbers of unique sites and agencies follow the same order.

About 79 percent of the pairs in all three parameter groups were uncensored and had percentage differences calculated for them. All three parameter groups had 91 percent or more of their percentage differences within +/- 20 percent.

The sign test for the parameter group containing filtered nitrite plus nitrate and filtered nitrate indicated statistically significant differences for pairs containing all the values (p=0.0035) and pairs containing only uncensored values (p=0.0000) and is due to the presence of nitrite in one and not the other of the pair. Nitrite constitutes a small percentage of the filtered nitrite plus nitrate value but introduces a measureable difference between the pairs of values.

The sign test for the parameter group containing filtered and unfiltered nitrite plus nitrate was highly insignificant for pairs containing all the values (p=1.0000) and pairs containing only uncensored values (p=1.0000). The sign test for the parameter group containing filtered and total nitrite plus nitrate was not significant for pairs containing all the values (p=1.0000) but was significant for pairs containing only uncensored values (p=0.0105). Sites from one agency appear to be driving the significant sign test results.

The sign test on all nitrite plus nitrate data pairs indicated no statistically significant difference among values within the parameter groups. However, the sign test comparing the parameter group containing filtered nitrite plus nitrate and filtered nitrate indicated a significant difference. This difference is small because of the presence of nitrite and is considered negligible considering the relative magnitude of nitrite to the total nitrite plus nitrate value in most cases. The sign test on uncensored values in the parameter group containing filtered and total nitrite plus nitrate indicates a small bias, but site-by-site sign tests indicate the issue is largely limited to one agency. Nearly all pctDiff values were within +/- 20 percent for all four parameter groups. These results demonstrated that these four fractions could be combined when a longer or denser record was needed.

Orthophosphate

Orthophosphate values with a fraction assigned to filtered were compared with orthophosphate values with a fraction assigned to unfiltered (table 2). No total or unknown orthophosphate fractions were considered because of their ambiguity. The evaluation used 51,944 pairs of data from 2,562 sites and eight agencies.

About 75 percent of all the paired samples were uncensored and had percentage differences calculated for them. Although more than 71 percent of the pctDiff values were within +/- 20 percent, the remaining values indicated a significant bias toward higher values for the unfiltered fraction, with 24 percent of the pairs having a pctDiff value ≥20 percent and only 4 percent of the pairs having a pctDiff ≤-20 percent. A review of individual sites demonstrates a widespread, systematic bias spanning multiple agencies and affecting most sites. The sign test on all data pairs and on the uncensored data pairs indicated no statistically significant difference among values within the orthophosphate parameter group composed of filtered and unfiltered fractions; however, site-by-site sign tests were statistically significant at about 30 percent of the sites.

Approximately 30 percent of individual sites fail the sign test, and nearly 25 percent of pctDiff values are biased toward higher values for the unfiltered fraction (≥20 percent). These results demonstrated that these two fractions should not be combined to create a longer or denser record.

Total Phosphorus

Total phosphorus values with a fraction assigned to unfiltered were compared with total dissolved phosphorus values with a fraction assigned to filtered (table 2). Total and unknown total phosphorus fractions were not considered because of their ambiguity. The evaluation used 270,938 pairs of data from 3,557 sites and 20 agencies.

About 55 percent of all the paired samples were uncensored and had percentage differences calculated for them. Less than 25 percent of the pctDiff values were within +/- 20 percent, and they also indicated a significant bias toward lower values for the filtered total phosphorus fraction, with about 74 percent of the pairs having a pctDiff value ≤-20 percent and only 2 percent of the pairs having a pctDiff ≥20 percent. A review of individual sites demonstrates a widespread, systematic bias spanning multiple agencies and affecting most sites.

The sign test on all data pairs and on the uncensored data pairs indicated a highly statistically significant difference among values within the total phosphorus parameter group (p=0.0000 for tests on both groups). Highly significant differences were also observed for individual sites. These results demonstrated that these two fractions should not be combined to create a longer or denser record.

Suspended Sediment

Suspended sediment concentration (SSC) values with a fraction assigned to suspended were compared separately with total suspended solids (TSS) values with fraction assigned to nonfilterable, suspended, and unknown (table 2). The largest number of paired samples was in the parameter group composed of SSC and unknown TSS (n=17,601), followed by the group composed of SSC and nonfilterable TSS (n=5,877), and lastly by the group composed of SSC and TSS (n=46). Relative numbers of unique sites and agencies follow the same order. The latter group consists of just five sites from one agency and will not be considered further because of the small sample size; SSC and nonfilterable TSS were not combined.

More than half of all the paired samples were uncensored and had percentage differences calculated for them. The two parameter groups had less than 46 percent of their percentage differences within +/- 20 percent, indicating a high degree of incompatibility between the fractions in both parameter groups. At least 48 percent of the pctDiff values were ≤-20 percent, which also points to a large bias in the paired data.

The sign test on all data pairs for the unfilterable TSS parameter group indicated no statistically significant difference among values within the parameter group (p=1.0000); the sign test on the uncensored values was highly significant (p=0.0000), and 70 percent of the individual sites failed the sign test at alpha [α]=0.1. The sign test for all data and uncensored data pairs in the parameter group containing the unknown TSS fraction was highly significant (p=0.0000).

Most results from the sign test indicate statistically significant differences between the evaluated parameter groups. In addition, fewer than 46 percent of all the pctDiff values were within +/- 20 percent and more than 48 percent were ≤-20 percent, indicating a substantial bias in the paired data. These results demonstrated that these two fractions should not be combined to create a longer or denser record, and that any other TSS values should be evaluated separately from SSC values. This is consistent with the findings of the study by Gray and others (2000) that concentrations of suspended sediment and TSS are incompatible, particularly in samples having a large fraction of sand-sized material.

Chloride

Chloride values with a fraction assigned to filtered were compared separately with chloride values with fractions assigned to unfiltered, total, and unknown (table 2). The largest number of paired samples was in the parameter group composed of filtered and total chloride (n=4,216), followed by the group composed of filtered and unknown chloride (n=2,715), and lastly the group composed of filtered and unfiltered chloride fractions (n=1,473). The largest number of agencies was in the parameter group containing total chloride, and the largest number of sites was in the parameter group containing chloride with an unknown fraction; all sites in the latter group were from one agency.

Almost all of the paired samples were uncensored and had percentage differences calculated for them. The parameter group containing the unknown chloride fraction was biased based on sign-test results, but more than 70 percent of the pctDiff values were still within +/- 20 percent; all of these data were from one agency. A site-by-site evaluation of the pairs containing unknown chloride fractions indicates most sites were not biased, but about 10 percent of the sites exhibited a bias. Most sites exhibiting a bias have fewer than 16 samples, and it is difficult to determine if the bias is real or the result of a small sample size. Sign tests at most of the sites exhibiting bias are not significant. The other two parameter groups (total and unfiltered) had 97 percent or more of their percentage differences within +/- 20 percent, indicating a high degree of compatibility between the fractions in the parameter groups.

The sign test on all data pairs and on the uncensored data pairs for groups containing unfiltered and total chloride indicated no statistically significant difference among values within the parameter groups. The sign test was statistically significant for the parameter group containing the unknown fraction but was driven by data from one agency. Site-by-site sign tests were not statistically significant at most sites. Nearly all pctDiff values were within +/- 20 percent for the unfiltered and total parameter groups. These results demonstrated that these four fractions could be combined when a longer or denser record was needed.

Sulfate

Sulfate values with a fraction assigned to filtered were compared separately with sulfate values with fractions assigned to unfiltered, total, and unknown (table 2). The largest number of paired samples was in the parameter group composed of filtered and unfiltered sulfate (n=689), followed by the group composed of filtered and total sulfate (n=96). The group containing filtered and unknown sulfate did not have enough data (n=8) to run the sign test or compute pctDiff values. The largest number of agencies was in the parameter group containing total sulfate, while the largest number of sites was in the parameter group containing sulfate with an unfiltered fraction; all sites in the latter group were from one agency.

Almost all of the paired samples were uncensored and had percentage differences calculated for them. The two parameter groups had 98 percent or more of their percentage differences within +/- 20 percent, indicating a high degree of compatibility between the fractions in the parameter groups.

The sign test on all data pairs for both parameter groups indicated no statistically significant difference among values within the parameter groups (unfiltered group p=0.2231; total group p=1.0000). The sign test for uncensored data pairs in the parameter group containing unfiltered sulfate was significant (p=0.0823), but the test on the group containing total sulfate was not (p=1.0000). Site-by-site p-values greater than 0.1 at 16 of 20 sites indicate the issue of statistically significant differences in values is not widespread; however, the small number of sites is not a robust sample to evaluate the importance of the sign test results.

Site-by-site sign tests were not statistically significant at most sites, although the sign test conducted on uncensored data for the parameter group containing unfiltered sulfate indicates a bias may be present. However, 98 percent of all the pctDiff values were within +/- 20 percent for the unfiltered and total parameter groups, which indicates that if a bias does exist, it is small. These results demonstrated that these three fractions could be combined when a longer or denser record was needed.

Alkalinity

Alkalinity values with a fraction assigned to filtered were compared separately with alkalinity values with fractions assigned to unfiltered, total, and unknown (table 2). The largest number of paired samples was in the parameter group composed of filtered and unfiltered alkalinity (n=49,160), followed by the group composed of filtered and unknown alkalinity (n=5,751), and lastly the filtered and total group (n=3,647). The largest number of agencies was in the parameter group containing total alkalinity, while the largest number of sites was in the parameter group containing alkalinity with an unfiltered fraction.

All of the paired alkalinity samples were uncensored and had percentage differences calculated for them. The three parameter groups had 90 percent or more of their percentage differences within +/- 20 percent, indicating a high degree of compatibility between the fractions in the parameter groups. The unknown and total groups had more than 99 percent of their sites within +/- 20 percent. For the filtered-unfiltered group, sites having values of pctDiff ≥20 percent exceeded those with values ≤-20 percent by a factor of three. At 65 percent of the sites having a large number of pctDiff values ≥20 percent, the median concentration was less than 50 mg/L, which indicates that dilute sites are responsible for much of the observed bias.

Because there were no censored values, the sign test was identical for all data and uncensored data pairs. The sign test for the parameter groups containing total and unknown alkalinity samples indicated no statistically significant difference among values within the parameter groups (p=1.0000). The sign test on the parameter group comparing filtered and unfiltered alkalinity was significant (p=0.00000: p-values ≤0.1 occurred at 359 sites (28 percent of the total number of sites). Visual inspection of data from a subset of sites with significant p-values indicated that most of the paired alkalinity values were within 20 percent of one another. Sites were visually spot checked, and a small bias toward higher unfiltered values was noted at some sites. Obviously bad or misentered data rarely were identified as the cause of the bias.

The observed bias at sites with statistically significant p-values is likely related to the bias in high pctDiff values discussed previously—in diluted waters the bias results in pctDiff values ≥20 percent, but at higher alkalinity concentrations, the pctDiff is <20 percent but still results in a significant p-value when the sign test is applied. These results demonstrated that these four fractions could be combined when a longer or denser record was needed.

Replicate Evaluation

The criteria used to determine the validity of combining parameters of differing fractions were applied to environmental replicate data collected by NAWQA between 1992 and 2012 for the purposes of substantiating the decisions to combine or not combine parameters for trend analysis. Most of the environmental replicate samples collected by NAWQA were split replicates—two aliquots of water were taken from the same compositing vessel and submitted for analysis. A comparison of these values provides insight into the sample processing and analytical variability. The percentage difference was calculated for the replicate data in the same manner as it was for the data used to evaluate paired samples. Arbitrarily, the environmental sample was denoted the preferred sample, and the replicate sample was treated as the nonpreferred sample. The results of this replicate analysis are reported in table 3.

The variability of split replicate samples of the same parameter and fraction should be smaller than the variability observed in sample pairs of different fractions, such as those in the data under consideration for combining (table 3). If the variability of the paired data under consideration for combining is substantially greater than the variability observed for the NAWQA replicate pairs, then that observation lends support to a decision not to combine those data pairs. On the other hand, if the variability of the data pairs is similar to the variability of the replicates, that would lend support to a decision to combine those pairs. As an example, 97.1 percent of the NAWQA replicates for filtered nitrite plus nitrate had a percentage difference between -20 percent and plus (+)20 percent compared to 92.5 percent of the pairs of filtered nitrite plus nitrate and unfiltered nitrite plus nitrate had a percentage difference between -20 percent and +20 percent. As expected, a lower percentage, but not a substantially lower percentage, of the filtered-unfiltered pairs fell within the -/+ 20 percent window, which indicates that the variability associated with the filtered-unfiltered pair is only slightly greater and not enough to support keeping the separate datasets.

The USGS NAWQA replicates were evaluated for filtered ammonia, Kjeldahl nitrogen, filtered nitrite plus nitrate, total phosphorus, filtered orthophosphate, SSC, filtered chloride, and filtered sulfate. In all cases when comparisons could be made between the replicate data and the paired sample data, the conclusions reached in the paired sample evaluation were substantiated. Table 3 summarizes the similarity of the two methods for evaluating the percentage of paired samples having pctDiff values less than 20 percent and greater than -20 percent. The table is intended to broadly show the two methods corroborate one another but does not capture the interpretive nuances used to evaluate the paired samples for combining.

Table 3. Percentage of uncensored data pairs having percentage difference values less than 20 percent and greater than -20 percent.

[USGS, U.S. Geological Survey; pctDiff, percentage difference; Rep, replicate sample; Env, environmental sample; nonPref, nonpreferred fraction; Pref, preferred fraction]

Parameter	USGS replicate samples, 1992–2012	Paired samples, values are for the constituent group with largest number of uncensored pairs	Paired-sample decision
Ammonia	83.0	64.4	Combine
Kjeldahl nitrogen	81.0	41.4	Do not combine
Nitrite plus nitrate	97.1	92.5	Combine
Orthophosphate	90.3	71.6	Do not combine
Total phosphorus	87.8	24.6	Do not combine
Suspended sediment concentration	72.2	38.8	Do not combine
Chloride	99.3	97.8	Combine
Sulfate	98.8	99.3	Combine

pctDiff calculated as
(1) USGS replicate samples: 100*(Rep-Env)/Env
(2) Paired samples: 100*(nonPref-Pref)/Pref

Nutrient Parameters of Interest and Harmonization

The target nutrient parameters for trend analysis were ammonia, nitrate (nitrate and nitrite plus nitrate), orthophosphate, total nitrogen (N), and total phosphorus. Because these were sometimes reported directly and sometimes calculated from other parameters, all available data initially were retained for ammonia, dissolved nitrogen, dissolved phosphorus, Kjeldahl nitrogen, nitrate, nitrite, nitrite plus nitrate, organic nitrogen, orthophosphate, particulate nitrogen, particulate phosphorus, total nitrogen, and total phosphorus. Parameters that clearly were not a variation on the target parameters (for example "Ammonia Nitrogen Transport" and "Nitrogen Gas Saturation") were excluded. Numerous parameters were ambiguous and were used only when a data source specified their interpretation of the parameter (for example, some sources specified that they used “Nitrogen Ion” to mean total nitrogen,” while others used “Nitrogen Ion” but did not specify their interpretation). Challenges with combining nutrient data from multiple sources for secondary uses, like this study, are described in Sprague and others (2017).

Naming conventions for a single parameter varied widely among data sources. For example, all of the variations on nitrate are shown in appendix 2, table 2–1. Numerous steps were taken to harmonize different parameter names that represented the same parameter. In many cases, the chemical represented by the parameter name was not clear. Metadata critical to precisely identifying and quantifying each data value—such as filtration status (fraction), units, and chemical form (molecular or elemental)—were often incomplete, rendering the data value ambiguous. When critical metadata were missing, the data were not used. In select cases, agencies were contacted to fill in metadata gaps.

One primary point of ambiguity that often could not be resolved without supporting metadata specifying whether a sample was filtered or unfiltered was the use of “total” in nutrient parameter names. Total was used two ways among data sources: (1) to represent the inclusion of multiple species, such as when ammonia and organic nitrogen were summed to give total Kjeldahl nitrogen, or when all dissolved phosphorus species were included in the determination of total dissolved phosphorus; and (2) to represent an unfiltered sample. Some values were clearly identified as being both summed and unfiltered (for example, total Kjeldahl nitrogen, unfiltered); others were clearly identified as being summed and filtered (for example, total Kjeldahl nitrogen, filtered or dissolved Kjeldahl nitrogen); others could only have been unfiltered and not summed, because they could not have been derived as a sum (for example, total nitrite). Values with these descriptors were retained; however, many other values were ambiguously described. These ambiguous values potentially could have been unfiltered and (or) derived as a sum, but clarifying metadata were not provided (for example, total Kjeldahl nitrogen, filtration status unknown). In such cases, total as part of the parameter name was assumed to mean unfiltered when a given source reported all of their parameter fractions as either total or filtered. If a source reported some parameter fractions as total and some as unfiltered, or if they reported parameters as just total, no assumption was made and all data with total as the fraction were excluded. Whenever possible, assumptions were verified with the data source.

Any data values that were reported without units, such that it could not be unambiguously determined whether they were reported in milligrams per liter, micrograms per liter, or some other unit, were excluded. Any data values that were reported without form, such that it could not be unambiguously determined whether they were reported in elemental form (for example, nitrate as nitrogen [as N]) or in molecular form (for example, nitrate as NO_3^-) also were excluded. In both cases, the choice of units or form commonly could substantially alter the value. For example, a nitrate concentration reported as NO_3^- would be more than four times larger than the same nitrate concentration reported as N.

Parameters reported as variants on the names soluble reactive phosphorus (SRP), orthophosphate, and orthophosphate-plus-hydrolyzable-phosphate were treated as being equivalent, and all were renamed orthophosphate. Operationally, SRP is determined using a spectrophotometric (colorimetry) analysis that measures orthophosphate as well as a small amount of other polyphosphates that are unavoidably hydrolyzed during the analysis. Most laboratories employ the spectrophotometric method in the measurement of what is alternatively called orthophosphate; thus, orthophosphate and SRP are operationally equivalent when the spectrophotometric method is used (Jarvie and others, 2002). Operationally, orthophosphate is not the same as SRP when orthophosphate is determined using ion chromatography because there is no inadvertent hydrolysis of other polyphosphates during the ion chromatography analysis (Westland and Boisclair, 1974). Because the amount of inadvertent polyphosphate hydrolysis during the spectrophotometric method typically is very small, orthophosphate determined using the spectrophotometric method, orthophosphate determined using ion chromatography, and SRP determined using spectrophotometry are considered to be equivalent for this study. Operationally, orthophosphate-plus-hydrolyzable-phosphate is not the same as either SRP or orthophosphate because it includes an additional deliberate hydrolysis step that can hydrolyze condensed polyphosphates (Jarvie and others, 2002). However, the amount of condensed polyphosphates typically is so low in natural waters (Kornberg and others, 1999) that orthophosphate-plus-hydrolyzable-phosphate also was considered to be equivalent to SRP and orthophosphate in this study.

The parameter phosphate was more complicated. Variations on the name (such as phosphate, total phosphate, or phosphate-phosphorus) can be used to represent either orthophosphate or total phosphorus, depending on the way total phosphorus is analyzed in the laboratory. A water sample that is collected in the field for total phosphorus is analyzed in a laboratory by converting all forms of phosphorus in the sample to phosphate. As a result, some agencies report the value as total phosphorus (based on the field perspective), but other agencies report the same value as phosphate or total phosphate (based on the laboratory perspective). In STORET, phosphate-phosphorus is used to represent total phosphorus (U.S. Environmental Protection Agency, 2015). Outside of STORET, some agencies appeared to follow this convention and used a variation on the name phosphate to represent total phosphorus. Other agencies used a variation on the name phosphate to represent a variation on orthophosphate. In many cases, it was unclear what the parameter name phosphate was intended to represent. When the meaning of phosphate could not be determined, the associated data were not used.

When the parameter name phosphate could unambiguously be determined to mean a variation on orthophosphate, it often was unclear whether the analysis included polyphosphate in addition to orthophosphate. Because polyphosphate is highly reactive and concentrations in most streams are likely to be very small compared to orthophosphate concentrations (Kornberg and others, 1999), phosphate was assumed to be equivalent to orthophosphate.

Concentrations of nitrite in streams are likely to be very small compared to nitrate concentrations (Dubrovsky and others, 2010), because nitrite is unstable in aerated waters (Hem, 1984). Unlike orthophosphate and polyphosphate, nearly 100,000 paired nitrate and nitrite plus nitrate samples were available in the dataset; these paired samples were used to evaluate the relative contribution of nitrite (assumed to be the difference between nitrite plus nitrate and nitrate) to nitrite plus nitrate in a wide variety of streams across the United States. Detailed results of the paired statistical test between nitrite plus nitrate and nitrate are presented in table 2. In general, the median percentage difference between paired nitrate and nitrite plus nitrate samples was about 1 percent; 94 percent of uncensored nitrate and nitrite plus nitrate pairs had a percentage difference within 20 percent. Based on these results, nitrate and nitrite plus nitrate were considered to be indistinguishable in this study, and data for both parameters were combined. When nitrate and nitrite plus nitrate results were present in a single sample, the priority for use in trend analysis was nitrite plus nitrate, then nitrate. The resulting data are collectively referred to as "nitrate."

All remaining nutrient concentration data with complete and unambiguous metadata were converted to milligrams per liter as nitrogen or as phosphorus (as P). Converted concentrations were rounded to the number of decimal places in the original value, with the exception of values converted from micrograms per liter to milligrams per liter (because of the possibility of rounding a converted value to zero).

Nutrient Data Processing

In order to reduce the nutrient dataset to a more reasonable size and to eliminate sites with very limited data that were not suitable for trend analysis, only sites that had at least 3 years of consecutive nutrient data with at least quarterly sampling initially were retained. Details on how quarters were defined can be found in the "Coverage of Changing Seasons" section in this report.

Unfiltered and Filtered Fractions

The USGS Office of Water-Quality (OWQ) Technical Memorandum 93–04 (accessed April 1, 2015, at http://water.usgs.gov/admin/memo/QW/qw93.04.html) states that prior to December 2, 1992, concentrations of total (unfiltered) and dissolved (filtered) nitrite, nitrite plus nitrate, ammonia, and orthophosphate were determined simultaneously using a four-channel, continuous flow, colorimetric analyzer. No digestion was performed on unfiltered samples prior to analytical determinations, and thus analytical procedures were identical for dissolved and total determinations. Moreover, because unfiltered samples with large concentrations of solids could foul the four-channel analyzer, an unknown number of total samples were laboratory filtered at the discretion of the analyst. Records of those laboratory filtrations were not kept, and the results were reported as a total concentration prior to that date; therefore, no analytical basis exists for a distinction between dissolved and total fractions. A subsequent statistical comparison of paired total and dissolved samples collected during calendar year 1989 found that the median concentrations were not significantly different at the 95-percent confidence level. The conclusion in the memorandum was that unfiltered and filtered concentrations of nitrite, nitrite plus nitrate, ammonia, and orthophosphate previously reported by the National Water Quality Laboratory (NWQL) were statistically indistinguishable. (By extension, theoretically this would also apply to nitrate, which is calculated as the difference between nitrite plus nitrate and nitrite.)

Because some USGS data may have been analyzed with different laboratory methods during the period before December 2, 1992 (that did not use the four-channel analyzer), and because a similar evaluation of paired filtered and unfiltered nutrient data was not available from all non-USGS laboratories, paired filtered and unfiltered nitrite, nitrate, nitrite plus nitrate, ammonia, Kjeldahl nitrogen, and orthophosphate data from all sources represented in the dataset were further examined as part of the data processing for this study. Paired filtered and unfiltered concentrations for each of the above parameters were compared collectively (all sources pooled together). Results are presented in the "Paired Sample Evaluation" section of this report and in table 2.

Based on the results in table 2 and in the OWQ memorandum 93–04, filtered and unfiltered nitrite, nitrate plus nitrite, nitrate, and ammonia concentrations were considered to be indistinguishable in this study, and data for the fractions were combined for each parameter. Data for these parameters also were accepted from agencies that did not specify a fraction in the metadata because either fraction was acceptable. When filtered, unfiltered, and (or) missing fraction results were present in a single sample, the preferred order for use in trend analysis was filtered, then unfiltered, and missing.

Based on the results in table 2, filtered and unfiltered orthophosphate concentrations were not treated as indistinguishable. The difference between paired filtered and unfiltered orthophosphate concentrations may have resulted from the inadvertent inclusion of organic phosphorus during the laboratory analysis of orthophosphate (Tarapchak and others, 1982). Because some agencies collected only filtered orthophosphate samples and others collected only unfiltered orthophosphate samples, both types of samples were retained as separate parameters. They were not combined at an individual site because switching between the two over time could induce an artificial trend. Orthophosphate data for which no fraction was reported were excluded.

For nutrient parameters that potentially have some particulate component—including total nitrogen, Kjeldahl nitrogen, organic nitrogen, and total phosphorus—filtered and unfiltered data were not treated as indistinguishable and were retained as separate parameters. Data without a reported fraction were excluded. For other nutrient parameters that have only a particulate component—particulate nitrogen and particulate phosphorus—only data with particulate or nonfilterable in the parameter name or given as the fraction were used. Data reported as filtered or for which no fraction was reported (or could be deduced from the name) were excluded.

Paired filtered and unfiltered nutrient values also were used for a more general quality screen. All paired filtered and unfiltered data from a given source were pooled together by parameter. If filtered values exceeded unfiltered values by more than 25 percent in more than 50 percent of paired samples, there was a high likelihood of laboratory performance issues; therefore, all filtered and unfiltered data from that source were excluded for that parameter. This comparison was made for all target nutrient parameters, as well as all other nutrient parameters used in the calculation of the target nutrient parameters. Censored data were excluded from this comparison to avoid comparisons of data with different detection limits.

Discrete Samples

Discrete samples were used in the study to avoid introducing sampling inconsistency. Composite samples—multiple discrete samples collected over an extended period of time (often more than 24 hours or throughout a storm event) that were composited together into a single sample— were excluded from the study.

Samples Analyzed in the Field

To maintain basic comparability in precision and accuracy of analytical determinations, data were restricted to samples analyzed in laboratories. Samples that were known to be analyzed using field chemistry kits were excluded. Remark codes on the samples (for example, Virginia Department of Environmental Quality remark code H: “Value based on field kit determination; may not be accurate”) and parameter codes (for example, USGS parameter code 99120: “Ammonia, water, filtered, field, milligrams per liter as nitrogen”) were used to identify samples that were analyzed in the field. Parameter names for parameters with no associated parameter code also were used to identify samples that were analyzed in the field (for example, “Nitrogen, Ammonia, Hach TNT” and “Ortho Phosphate, Hach method 8048”). Information on whether the samples were analyzed in the field or laboratory was not always provided, so some samples that were analyzed in the field may have been inadvertently retained.

Remark Code Harmonization

There were more than 1,400 unique remark codes in the original dataset, which included all parameters. Many of these were not defined, and many others were ambiguously defined. Data with undefined or ambiguous remark codes were excluded. Remark codes indicating poor data quality (for example, “Contamination present in the sample”), censored data, or data below a laboratory detection limit were of interest to this study. Data with one or more remark codes indicating poor data quality were excluded. Data with a remark code indicating laboratory censoring were retained. There were about 65 unique remark codes in the data indicating laboratory censoring; these were unified under a single remark code of “<.” In some cases, there was no remark code provided, but a comment field associated with a data value indicated that the value was censored; these values also were given a remark code of “<.”

Some sources reported a censored remark code for an observation but did not report an associated value. When more than 1 percent of the data for a parameter at a site had a censored remark code but no values, then all of the data (censored and uncensored) for that site and parameter were excluded. When less than 1 percent of the data for a parameter at a site had a censored remark code with a missing value and censored values were reported for other observations, then the missing censored values were set to the reported censored value that was closest in time for that site and parameter.

Other dataset sources did not have any nutrient data values with remark codes to indicate censoring. When contacted, a number of these sources confirmed that some of the data were censored, but data with remark codes were not uploaded to STORET and (or) their data with remark codes were not made publicly available through their own databases. The reasons for this practice are unknown, but it creates the potential for substantially biased analyses. When no censored remark codes were present for any data from a given source, all of the data from that source were excluded. Some sources furnished data through both the STORET database and their own database. On occasion, remark codes for some agencies were present in one of these databases but not the other. For evaluating overall censoring, data from the two databases were evaluated separately; as a result, for some sources, data from one database were excluded while data from the other database were retained. The true amount of censoring in any excluded data could not be determined; therefore, the amount of bias in the final trend results could not be determined if data from those sources were used.

In some cases, when an agency reported data with censored remark codes, units were provided for uncensored data but not for censored data. The reasons for this practice are unknown, but it creates the potential for substantially biased analyses if only the censored data were excluded because the units were missing. If all of the missing units were for censored data for a parameter from a given source, and all uncensored data were reported in the same units throughout the record (for example, all uncensored data for a parameter were reported in units of milligrams per liter or micrograms per liter), then those same units were assumed to apply to all censored data for that parameter from that source. If necessary, all retained data were converted to milligrams per liter.

Adjustments to U.S. Geological Survey Data Based on Historical Memorandums and Reports

Several adjustments to USGS data reported by the NWQL were necessary based on information contained in USGS technical memorandums and reports. Based on the NWQL technical memorandum 1997–10 (accessed April 1, 2015, at http://wwwnwql.cr.usgs.gov/dyn.shtml?techmemo), censored and uncensored values of dissolved ammonia (parameter 00608) reported with values below 0.02 were recensored to less than 0.02 mg/L for samples analyzed by method I–2522–90 (USGS method codes CL035, CL036, and CL037) between October 1, 1994, and September 30, 1997 (tables 3–1 and 3–2). Analytical methods are presented in this report using the 5-digit internal laboratory code that is embedded in the analyte identification field in the NWQL online catalog (http://nwql.cr.usgs.gov/usgs/catalog/index.cfm). Based on Patton and Truitt (2000), 0.1 mg/L was subtracted from the results for total and dissolved Kjeldahl nitrogen samples (parameters 00623 and 00625, respectively) collected between January 1, 1986, and October 1, 1991, that were analyzed by methods I–2552–85 (USGS method code CL051) and I–4552–85 (USGS method code CL083) because corrections for digestion blank concentrations (nearly equal to 0.1 mg/L) were not subtracted from the results originally reported. If the resulting value was below the detection limit at the time, the value was censored at the detection limit. Based on NWQL technical memorandum 1998–07 (accessed April 1, 2015, at http://wwwnwql.cr.usgs.gov/dyn.shtml?techmemo), results for dissolved and total phosphorus samples (parameters 00666 and 00665, respectively) collected between October 1, 1991, and September 30, 1998, that were analyzed by methods I–2610–91 and I–4610–91 (USGS method codes KJ004, KJ005, KJ009, KJ0010, KJ011) and reported below the 0.03 mg/L long-term method detection limit (LT–MDL) were censored to less than 0.05 mg/L (appendix 3).

These corrections were applied to data from the NWIS database that had the analyzing entity defined as NWQL, or when the analyzing entity was missing. In the latter case, which was true of many samples in the NWIS database, the analyzing entity was assumed to be NWQL. While this may not have always been a correct assumption and may have led to inappropriate adjustment of some data analyzed by other laboratories, it enabled appropriate adjustment to a large amount of NWQL-generated data (when analyzing agency was populated, about 91 percent of the samples were analyzed at NWQL). Because similarly detailed histories of method performance and data quality from laboratories other than the NWQL were not readily available, no adjustments were made to data from other laboratories.

Field blanks test for bias in concentrations that could result from contamination during sample collection and handling, shipping, and laboratory analysis. As part of the NAWQA Project, the quality of nutrient data collected at NAWQA sites during 1992–2001 was examined (Mueller and Titus, 2005). The study determined no contamination issues with nitrite plus nitrate values and Kjeldahl nitrogen values but did find indication of contamination with some ammonia, orthophosphate, and total phosphorus samples; however, this contamination was determined to be small: “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” (Mueller and Titus, 2005). The absolute value of the average rate of change at all of the trend sites in this report was 0.09 mg/L as N for ammonia, 0.05 mg/L as P for orthophosphate, and 0.09 mg/L as P for total phosphorus (see “Trend Results” section of this report). According to table 3 in Mueller and Titus (2005), the 80th percentile of the concentrations detected in the field blanks in streams was 0.02 mg/L as N for ammonia, less than 0.01 mg/L as P for orthophosphate, and less than 0.01 mg/L as P for total phosphorus; therefore, the magnitude of any potential contamination is likely to be much smaller than the environmental change at the trend sites in this report and would likely have a negligible effect. Although a more recent systematic evaluation of the quality of NAWQA data—including nutrients and other parameters—has not been published, it is unlikely that there have been major changes in the degree of contamination during nationwide field sampling or in the degree of contamination in the laboratory (as supported by the analysis of laboratory bias in this study; see appendix 4). In addition, when the magnitude of any contamination randomly varies over time, it is unlikely to affect trend analysis, but systematic increases or decreases over time in the magnitude of contamination could contribute to false trends. There is no evidence that field contamination has been systematically increasing or decreasing over time (Mueller and Titus, 2005). Because data were used collected by multiple USGS projects (not limited to NAWQA) and because potential contamination could have negligible effect on the trend analysis, an attempt was not made to correct the data based on the findings in Mueller and Titus (2005); however, the findings of that report may be useful in the interpretation of trend results.

Recensoring U.S. Geological Survey Data

On October 1, 1998, the NWQL began implementing new reporting conventions for nondetections (Oblinger Childress and others, 1999). Prior to October 1, 1998, nondetections were censored to a method detection level (MDL), the concentration at which the risk of a false positive detection (analyte reported as present at the MDL when not in the sample) is no more than 1 percent, or to a minimum reporting level (MRL), an inconsistently defined level of detection capability. The NWQL deemed the use of inconsistent MRL as inappropriate and also observed that the risk of a false negative occurrence at the MDL (analyte reported as not present when present at the MDL concentration) could be as much as 50 percent. To make the determination of detection limits more consistent and to reduce the risk of false negatives to no more than 1 percent, the NWQL began using a laboratory reporting level (LRL) in combination with a LT–MDL on October 1, 1998, based on Oblinger Childress and others (1999). The LRL typically was twice the LT–MDL. Between October 1, 1998, and October 1, 2010, concentrations measured between the LRL and the LT–MDL were reported as estimated concentrations with an "E" remark code and nondetections were censored to the higher LRL. Beginning on October 1, 2010, the NWQL censored inorganic nondetections to the LT–MDL (https://water.usgs.gov/admin/memo/QW/qw10.07.html).

The reporting convention in place between October 1, 1998, and October 1, 2010, creates potential bias in statistical analyses because it improperly represents the proportions of the data below the LT–MDL and the LRL (Helsel, 2005). In particular, observations that are truly below the LT–MDL will be reported as censored at the higher LRL; thus, the probability that an observation falls between the LT–MDL and the LRL may be overestimated, and the probability that it falls below the LT–MDL may be underestimated (Helsel, 2005). Depending on the true percentage of data below the LT–MDL, this can cause an upward bias that will adversely affect statistical analyses when these affected data are pooled with data subject to conventional censoring (that is data not recensored from the LT–MDL to the LRL). To avoid this problem, there are two potential solutions: (1) The LT–MDL becomes the reporting limit when values reported as censored at the higher LRL are recensored to the lower LT–MDL, and uncensored "E" values between the LT–MDL and the LRL are not recensored. (2) The LRL becomes the reporting limit when all uncensored "E" values below the LRL are censored at the LRL. To retain the extra information captured by the "E" values, the first solution was used for USGS data from the NWQL that were originally reported using the LRL approach. The "E" remark code may have been applied to other values for unrelated performance-based reasons as described in Oblinger Childress and others (1999); these values were unchanged.

Similar to the description in the "Adjustments to U.S. Geological Survey Data Based on Historical Memorandums and Reports" section, these adjustments were applied to data from the NWIS database that had the analyzing entity defined as NWQL or when the analyzing entity was missing. For data analyzed at other laboratories, metadata documenting reporting conventions for censored data were not readily available; therefore, all censored data from laboratories other than the NWQL were unaltered.

Duplicate Samples

It was not uncommon to receive the same water-quality data for the same site and collection date from more than one source during the compilation of data. Duplicates were defined as samples from the same site, date, time, and fraction. Duplicate samples were removed from the dataset. It was not possible to determine which of the duplicate samples should be removed on the basis of data source because sometimes a source provided longer or more completely populated records for some parameters (for example, there may have been duplicate data obtained from STORET and provided directly by a State agency but overall the STORET record was longer for total nitrogen whereas the State record was longer for total phosphorus). As a result, each parameter was considered separately when deciding which duplicate samples to retain and which to exclude. Note that duplicate samples did not always have identical values for a given parameter. If one of the duplicate samples was from USGS and one was from another agency, the USGS sample was retained. If both were USGS samples (for example, the duplicate samples were obtained from two different USGS NWIS servers), the sample from the primary host server (home State for each site) was retained. If information on the primary host server was not available, one of the duplicate samples was randomly retained. If both samples were from agencies other than USGS, and one sample had been provided by the CBP, then the sample from the CBP was retained. If both samples were from agencies other than USGS and the CBP, then the sample from the longest period of record was retained. Note that priority was not given to USGS samples because of an a priori expectation that USGS data were of better quality, but because typically more metadata were available on how samples were collected, analyzed, and stored.

Zero, Negative, and Missing Concentration Values

An analytical determination of zero or negative concentration is not possible, so zero or negative values in the data were considered suspect. For a given site and parameter combination, if a censored remark code was frequently associated with a zero or negative concentration value, all zero and negative values were assumed to represent censored values and were set to the reported censored value that was closest in time for that site and parameter combination. In addition, missing values that were reported with a censored remark code were assumed to be censored values without an associated laboratory detection level. These values were also set to the reported censored value that was closest in time for that site and parameter combination. If there were no reported censored values for a site and parameter combination, then zero, negative, or missing censored concentration values were set to the lowest uncensored value for that site and parameter combination and a censored remark code was assigned.

In some cases, these assumptions may have resulted in an incorrect value. For example, the zero, negative, or missing censored values may have been from samples that were not analyzed in the same laboratory as other samples for the given parameter or for samples with a laboratory error. Alternatively, the assumed detection limit assigned to these values may be wrong. The approach used in this study to populate zero, negative, or missing censored values was unlikely to affect the subsequent trend results when such values collectively constituted a small fraction of the total data for a site and parameter combination; therefore, when zero, negative, and missing censored values made up more than 1 percent of the overall dataset for a site and parameter combination, all data for that site and parameter combination were not included in the determination of trend.

Calculation of Target Parameters

In many cases, certain parameters were not directly determined in a laboratory analysis and were calculated by summing other parameters. When a parameter was not directly determined in a laboratory analysis, the following algorithms were used to calculate the value:

$$\textit{Nitrite plus nitrate} = \textit{nitrite} + \textit{nitrate}, \tag{2}$$

$$\textit{Kjeldahl nitrogen (filtered)} = \textit{organic nitrogen (filtered)} + \textit{ammonia}, \tag{3}$$

$$\textit{Kjeldahl nitrogen (unfiltered)} = \textit{organic nitrogen (unfiltered)} + \textit{ammonia}, \tag{4}$$

$$\textit{Dissolved nitrogen} = \textit{nitrite plus nitrate} + \textit{Kjeldahl nitrogen (filtered)}, \tag{5}$$

$$\textit{Total phosphorus} = \textit{particulate phosphorus} + \textit{dissolved phosphorus}, \tag{6}$$

$$\textit{Total nitrogen} = \textit{particulate nitrogen} + \textit{dissolved nitrogen}, \tag{7}$$

$$\textit{Total nitrogen} = \textit{Kjeldahl nitrogen (unfiltered)} + \textit{nitrite plus nitrate, and} \tag{8}$$

$$\textit{Total nitrogen} = \textit{Kjeldahl nitrogen (unfiltered)} + \textit{nitrate.} \tag{9}$$

The priority among the three total nitrogen calculations is as shown; that is, equation 7 was the first priority, equation 8 was the second priority, and equation 9 was the third priority.

As explained in the “Unfiltered and Filtered Fractions” subsection of the “Nutrient Data Processing” section, nitrite, nitrate, and ammonia concentrations from all fractions (including filtered, unfiltered, total, or missing) were pooled in the trend calculation for each parameter. For other trend calculations, only data specified as filtered were used for organic nitrogen (filtered) and Kjeldahl nitrogen (filtered); only data specified as unfiltered were used for organic nitrogen (unfiltered) and Kjeldahl nitrogen (unfiltered); only data specified as filtered were used for dissolved phosphorus and dissolved nitrogen; and only data specified as particulate or nonfilterable were used for particulate nitrogen and particulate phosphorus. Also, nitrate and nitrite plus nitrate were considered to be indistinguishable, so either could be used in the calculation of total nitrogen. If both were available in the same sample, nitrite plus nitrate was prioritized.

In cases when one or more of the summands was censored, interval censoring (Hirsch and De Cicco, 2015) was used. Consider the calculation of total nitrogen for the sample collected at the Potomac River at Chain Bridge, Washington, D.C. (Site 01646580), on June 7, 1999. Nitrite plus nitrate was reported as 0.596 mg/L, and Kjeldahl nitrogen was reported as less than 0.1 mg/L. Interval censoring accounts for the fact that the exact Kjeldahl nitrogen value is unknown but lies somewhere between 0 and 0.1; therefore, the lowest the total nitrogen value could be is 0.596 plus 0, or 0.596; the highest the total nitrogen value could be is 0.596 plus 0.1, or 0.696. The resulting interval censored value for total nitrogen is (0.596, 0.696).

If a value directly measured in a laboratory and a new calculated value ultimately were available for the same parameter in the same sample, the laboratory value was prioritized. For some samples, a calculated value was already present in the original data record along with laboratory values for the summands. Because the algorithm used to calculate that value was unknown and the treatment of any censored summands during the calculation was unknown (for example, one-half of the detection limit is often substituted for a censored value), calculations 2 through 9 listed above were made when summand values were available. The new calculated value was retained in place of the original calculated value. It was not always clear whether a provided value was calculated or measured in a laboratory analysis. For example, nitrite, nitrate, and nitrite plus nitrate can each be determined analytically or by calculation, depending on the laboratory methods used. Because laboratory methods were infrequently and inconsistently reported, it is possible that a measured value was inadvertently replaced with a new calculated value in an unknown number of samples. The new calculated value might differ from the original measured value; for example, in samples when one or more of the values involved in the calculation were censored, or when particulate and dissolved fractions were summed to determine total nitrogen rather than a value determined by persulfate digestion.

Multiple steps were used to determine the final trend model input values used in this study. Note that, because of the adjustments to USGS data based on historical memorandum, the recensoring of some USGS data, the summation of individual analytes for total parameters, the use of unrounded data from NWIS, and the use of interval censoring, the final USGS values used in this study will not always match the corresponding value stored in the NWIS database.

Partial and total nutrient values also were used for a more general quality screen of data from each source. All data from a given source were pooled together, and partial and total values were compared (for example, paired nitrite plus nitrate and total nitrogen values were compared). Censored values were excluded from these comparisons because of the potential for differences in the respective detection limits. If partial values exceeded total values by more than 25 percent in more than 50 percent of paired samples, there was a high likelihood of laboratory performance issues, and all data from that source were excluded for those parameters. This comparison was made for Kjeldahl nitrogen and total nitrogen, nitrite plus nitrate and total nitrogen, nitrate and total nitrogen, particulate nitrogen and total nitrogen, total dissolved nitrogen and total nitrogen, orthophosphate and total phosphorus, particulate phosphorus and total phosphorus, and total dissolved phosphorus and total phosphorus.

Identification of a Colocated or Nearby Streamgage

Following the steps described in the previous subsections of the “Nutrient Data Processing” section, all site and parameter combinations with at least 3 years of consecutive data with at least quarterly sampling (see “Coverage of Changing Seasons” section in this report for details on how quarters were defined) were paired with a streamgage from the final site-streamgage list assembled in the “Streamgage Selection” section of this report. Any site not meeting these minimum data requirements and without an acceptable streamgage match was excluded from further consideration.

Merging Colocated or Neighboring Water-Quality Sites

After assigning a streamgage to each site, the retained sites represented two classes of sites: (1) sites that had a sufficient amount of data over a long enough period of record (at least for the shortest trend period of 10 years), to be used on their own for trend analysis (single sites); (2) sites that had data for a shorter period of record (at least 3 consecutive years of quarterly data) that could be combined into a composite site with either other colocated or neighboring composite sites to form a single record for trend analysis or a colocated or neighboring stand-alone site to increase the number of trend periods that could be evaluated at the stand-alone site. Single and composite sites were only combined when additional trend periods could be evaluated after combining; they were not combined to only fill in short gaps or add to the existing data at the single site when the combination did not increase the number of trend periods that could be evaluated. Data for periods shorter than 3 years were not used to fill in gaps or extend the record at other composite or single sites because it was decided that the nonrandom variability introduced into those types of combined time series would outweigh the benefits during trend analysis.

Following methods described in the “Streamgage Matching” section, potential colocated and neighboring sites were identified, and sites with intervening influences were excluded. A maximum of two sites was combined to minimize nonrandom variability. If more than two sites were available, then the two sites with the longest period of record for the parameter of interest were selected for the composite site. If more than one of the remaining sites merged well with the primary record, then the site with the longest record was retained. A list of composite sites with the individual sites for each parameter used in trend analysis can be found in appendix 1, table 1–7. The subsequent processing steps described below were applied to the data record for merged sites and single sites.

Degree of Censoring

Following guidelines suggested for the trend method used in this study, site and parameter combinations with more than 50 percent censoring were excluded from trend analysis (Hirsch and De Cicco, 2015). Left-censored and interval-censored values, if less than 50 percent of the dataset, were included in this determination. Right-censored data were removed from the dataset.

Outliers

A small number of values in the dataset appeared to be typographical errors. For example, a total phosphorus concentration of 500 mg/L is extremely unlikely to occur in natural waters. It is far more likely that the units were reported incorrectly and that the true value is 500 µg/L or 0.5 mg/L. In contrast, a total phosphorus concentration of 50 mg/L is more ambiguous. That concentration is very high relative to concentrations found in most natural waters (Dubrovsky and others, 2010), but it could occur during locally extreme conditions such as an overflow of untreated sewage or a heavy runoff event immediately after nearby fertilizer application. In a smaller study involving far fewer sites and data sources, it might be possible to definitively establish the accuracy of ambiguous data points by checking laboratory records, field forms, or other sources of historical information, but that was not possible in a study of this size because of resource constraints, so only the most extreme outliers—those that were almost certainly typographical errors—were excluded. Under a normal distribution, 99.9999998 percent of the data should be within six standard deviations from the mean; only the most extreme values will be outside of that range. For each site and parameter pair, the data were divided into four seasons. Within each season, any data value further than six standard deviations from the mean in log space were excluded as a likely typographical error. With this approach, there was a risk that typographical errors representing less extreme values were retained. There was also a risk that values representative of true environmental conditions were excluded.

Data Thinning

The nutrient trend model used in this study weights days with multiple water-quality samples more heavily than other days during calibration. As a result, data for each site and parameter combination were thinned to have no more than one sample per day. If there was more than one sample per day, one sample was randomly selected to be retained in the dataset. The presence of data points spaced closely together in time also can introduce serial correlation into a time series. With some trend methods, serial correlation can increase the probability of incorrectly rejecting the test's null hypothesis of no trend, leading to an incorrect determination that a significant trend has occurred (Hirsch and Slack, 1984); The trend method in this study relied on bootstrapping to determine p-values (see "Nutrient, Sediment, Major Ion, Salinity, and Carbon Trend Analysis Method" section in this report), an approach that is less affected by serial correlation. For that reason, the datasets were not thinned for data points close in time on different days.

Final Screening for Data Coverage

Once data from colocated or neighboring sites had been merged, and the censoring, outlier, and data thinning steps had been completed, the resulting data series were screened again by parameter. In this final screening step, the requirements for data coverage were more stringent. All sites not meeting these criteria were excluded for trend analysis.

Coverage Throughout Target Trend Periods

- The target trend periods, in water years, were (1) 1972–2012, (2) 1982–2012, (3) 1992–2012, and (4) 2002–2012. Data were permitted to begin in the specified start year or the following year. For example, a start year of either 1972 (October 1, 1971) or 1973 (October 1, 1972) was acceptable in trend period 1, but 1974 (October 1, 1973) was not. Data were permitted to end in the specified start year or the preceding year. An end year of either 2011 (September 30, 2011) or 2012 (September 30, 2012) was acceptable in all trend periods, but 2010 was not.
- The first 2 years and the last 2 years of a trend period were required to have at least quarterly samples. See the "Coverage of Changing Seasons" section below for more detail on how quarters were defined.
- Overall, 70 percent of all years during a trend period were required to have at least quarterly samples. Any gap must have occurred in the middle of the record.
- Longer gaps in quarterly data were permitted when there was at least a 10-year span meeting the criteria beginning in any of the start years (for example, data from 1972 to1981 or 1982 to 1991) and a 10-year span meeting the criteria ending in either 2011 or 2012. Data during each of those decades must have met the other detailed criteria. Gaps of any length were permitted in the years between the beginning and ending decades.
- At least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during a given trend period in order to ensure the full range of flows were represented. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. For the shortest trend period (2002–12) of one decade, at least 14 percent of the samples were required to be high-flow samples. See the "Coverage of Changing Streamflow" section below for more detail on how high-flow samples were defined.

Coverage of Changing Seasons

Nutrient concentrations vary throughout the year in response to changes in sunlight, temperature, biological activity, and nutrient sources (Dubrovsky and others, 2010). Samples representing seasonal conditions throughout the year are needed to fully characterize changes in concentration. To ensure seasonal coverage, 70 percent of years during a trend period were required to have at least one sample per quarter; the remaining 30 percent of years could be missing some or all quarterly samples. This allowed for reasonable changes in sampling design over time.

To evaluate seasonal coverage, the water year was divided into quarters using three different divisions of the water year (table 4). Each quarter consisted of 3 calendar months (table 4). A site was accepted if the samples met any one of the quarterly definitions in table 4; the quarterly definition was permitted to vary from year to year. This method of testing for quarterly samples accommodated variations in real-world sampling schedules. The primary sampling-design artifact this accommodated was the occurrence of two samples in one quarter and none in the preceding or following quarter, which may have occurred if a sample was collected slightly earlier or slightly later than dictated by a strict 3-month schedule. As an example, consider a site sampled in January, late March, July, and October during a particular year. The source collecting these data is clearly collecting quarterly samples; however, the sample that should have been collected in April was collected in late March. Without a flexible quarter definition, this year would not have had quarterly sampling. This flexible method of defining quarters also provides a buffer for sources that sample rivers that may freeze in the winter. For various reasons, including the weather, a winter quarterly sample may need to be collected anytime between November and February.

Table 4. Definitions of quarters used to test sites for a sufficient number of samples for trend analysis.

Quarter definition based on water year division	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.
Division I	1	1	1	2	2	2	3	3	3	4	4	4
Division II	4	1	1	1	2	2	2	3	3	3	4	4
Division III	4	4	1	1	1	2	2	2	3	3	3	4

Coverage of Changing Streamflow

Nutrient concentrations also vary in response to changes in streamflow (Hem, 1984; Dubrovsky and others, 2010). Samples representing a range of streamflow conditions are needed to fully characterize changes in concentration over time. Low- or moderate-flow samples are not in short supply. To identify high-flow samples, streamflow on the date of each sample was compared to the 85th percentile of flows in that month during the corresponding decade of the trend period at that site (depending on the trend period, that would include some or all of the following decades: 1972–81, 1982–91, 1992–2001, and 2002–12). For example, for a sample collected on March 1, 1998, the 85th percentile of all daily streamflows in March during the decade 1992–2001 was determined. If the streamflow on March 1, 1998, was greater than the 85th percentile of March streamflows during that decade, the sample was considered to be a high-flow sample. This assessment was completed by decade rather than once for the full trend period to ensure that samples at high streamflows were distributed throughout the trend period, rather than potentially being concentrated in one part of it. It also allowed for the identification of high-flow samples collected during prolonged low-flow periods (for example, droughts). Such high-flow samples would have been collected during flows that were elevated relative to an unusually low baseline; the flows may not have been particularly high relative to overall average conditions at a site.

To ensure appropriate streamflow coverage, at least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during the trend period. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. The relaxation of the 14-percent criterion to 10 percent in some decades was necessary to accommodate sites with changing sampling designs over time (for example, the collection of high-flow samples increased somewhat or decreased somewhat over time). For the shortest trend period (2002–12) of only one decade, at least 14 percent of the samples were required to be high-flow samples during that period. These thresholds were based on sensitivity testing described in the “Storm Sampling” section of this report.

Data Used in Model Calibration

For site and parameter combinations that met the screening criteria, all available data before, during, and after the trend period (and if applicable, during any gaps) were retained to improve the model calibration (see the “Influence of New Data on Trend Estimates” section). Ultimately, model estimates only during the screened trend period were retained. In the case of longer records with data in two different decades separated by gaps, model estimates made in the years between the beginning and ending decades were excluded. Model input data are available in De Cicco and others (2017).

There were 850 sites evaluated for nutrient trends with 2,211 combinations of sites and parameters. Table 5 summarizes the number of sites evaluated for each parameter and trend period. A complete list of sites and parameters evaluated for nutrient trends is reported in appendix 1, table 1–6.

Sediment Parameters of Interest and Harmonization

The target sediment parameters for trend analysis were SSC and TSS. Parameter code, analytical method, parameter name, and fraction were used to harmonize the SSC and TSS records. Parameters that were not a variation on the target parameters were excluded. For example, analyses were excluded if the compiled data contained numerous analyses of fractions of the total sediment or solids, such as total dissolved solids, fixed solids, and volatile solids. Additionally, any analyses of specific size fractions of the sediment, such as less than 62.5 microns, were excluded. Numerous parameters were ambiguous because of missing metadata and were used only when the source specified their interpretation of the parameter (for example, some sources did not include the analysis fraction but did include the laboratory method, which provided unambiguous information about the analysis).

Naming conventions for SSC and TSS varied widely among data sources. For example, all of the variations on total suspended solids are shown in appendix 2, table 2–2. Numerous steps were taken to harmonize different parameter names that represented the same method of analysis for

Table 5. Number of trend sites for each parameter and trend period combination.

[--, no trend sites for this parameter and trend period]

Parameter group	Parameter	Trend period			
		1972–2012	1982–2012	1992–2012	2002–12
Carbon	Alkalinity	63	114	241	446
Carbon	Dissolved organic carbon	--	--	15	48
Carbon	Total organic carbon	5	10	61	209
Ecology	Diatoms	--	--	3	35
Ecology	Fish	--	--	4	47
Ecology	Invertebrates	--	--	9	49
Major ion and salinity	Bromide	--	--	2	32
Major ion and salinity	Calcium	27	50	98	216
Major ion and salinity	Chloride	57	95	163	367
Major ion and salinity	Magnesium	28	51	97	214
Major ion and salinity	Potassium	19	44	80	162
Major ion and salinity	Sodium	20	44	83	171
Major ion and salinity	Specific conductance	138	286	529	859
Major ion and salinity	Sulfate	44	71	143	304
Major ion and salinity	Total dissolved solids	35	50	123	280
Nutrient	Ammonia	24	103	200	344
Nutrient	Nitrate	54	161	286	498
Nutrient	Orthophosphate, filtered	5	46	146	279
Nutrient	Orthophosphate, unfiltered	--	--	--	29
Nutrient	Total nitrogen	12	85	172	331
Nutrient	Total phosphorus	41	145	304	489
Pesticide	2,6-Diethylaniline	--	--	1	--
Pesticide	Acetochlor	--	--	--	35
Pesticide	Alachlor	--	--	5	27
Pesticide	Atrazine	--	--	6	67
Pesticide	Carbaryl	--	--	5	31
Pesticide	Carbofuran	--	--	4	--
Pesticide	Chlorpyrifos	--	--	6	18
Pesticide	Cyanazine	--	--	5	3
Pesticide	Dacthal	--	--	4	18
Pesticide	Deethylatrazine	--	--	6	62
Pesticide	Desulfinylfipronil	--	--	--	37
Pesticide	Desulfinylfipronil amide	--	--	--	10
Pesticide	Diazinon	--	--	6	30
Pesticide	Dieldrin	--	--	2	2
Pesticide	S-Ethyl dipropylthioarbamate (EPTC)	--	--	4	11
Pesticide	Ethoprophos	--	--	1	3
Pesticide	Fipronil	--	--	--	32
Pesticide	Fipronil sulfide	--	--	--	23
Pesticide	Fipronil sulfone	--	--	--	14
Pesticide	Fonofos	--	--	4	--
Pesticide	Malathion	--	--	2	6
Pesticide	Metolachlor	--	--	5	58
Pesticide	Metribuzin	--	--	4	20
Pesticide	Molinate	--	--	1	1
Pesticide	Pendimethalin	--	--	6	11
Pesticide	Prometon	--	--	6	54
Pesticide	Propanil	--	--	--	1
Pesticide	Propargite	--	--	1	--
Pesticide	Propyzamide	--	--	--	3
Pesticide	Simazine	--	--	6	56
Pesticide	Tebuthiuron	--	--	4	25
Pesticide	Thiobencarb	--	--	--	1
Pesticide	Trifluralin	--	--	6	5
Sediment	Suspended sediment concentration	9	25	66	107
Sediment	Total suspended solids	5	133	252	437

sediment concentration. In many cases, the analysis represented by the parameter name was not clear. Metadata critical to precisely identifying and quantifying each data value—such as analytical method, subsampling, and units—often were incomplete, rendering the data value ambiguous. When any of these critical metadata were missing, the data were not used. In select cases, agencies were contacted to fill in the gaps in metadata. Any data values that were reported without units, such that it could not be unambiguously determined whether they were reported in milligrams per liter, micrograms per liter, or some other unit, were excluded.

All remaining concentration data with complete and unambiguous metadata were converted to milligrams per liter of SSC or TSS. Converted concentrations were rounded to the number of decimal places in the original value, with the exception of values converted only from micrograms per liter to milligrams per liter (because of the possibility of rounding a converted value to zero).

Sediment Data Processing

In order to reduce the nutrient dataset to a more reasonable size and to eliminate sites with very limited data that were not suitable for trend analysis, only sites that had at least 3 years of consecutive sediment data with at least quarterly sampling initially were retained. Details on how quarters were defined can be found in the “Coverage of Changing Seasons” section in this report.

Discrete Samples

Discrete samples were used in the study to avoid introducing sampling inconsistency. Composite samples—multiple discrete samples collected over an extended period of time (often more than 24 hours or throughout a storm event) that were composited together into a single sample—were excluded from the study.

Samples Analyzed in the Field

To maintain basic comparability in precision and accuracy of analytical determinations, data were restricted to samples analyzed in laboratories. Remark codes on samples (for example, Virginia Department of Environmental Quality remark code H: “Value based on field kit determination; may not be accurate”) and sample types (“Field Msr/Obs”) were used to identify samples that were analyzed in the field. Information on whether the samples were analyzed in the field or laboratory was not always provided, so some samples that were analyzed in the field may have been inadvertently retained.

Remark Code Harmonization

There were over 1,400 unique remark codes in the original dataset including all parameters. Many of these were not defined, and many others were ambiguously defined. Data with undefined or ambiguous remark codes were excluded. Of most interest for this study were remark codes indicating poor data quality (for example, “Contamination present in the sample”) and remark codes indicating censored data or data below a laboratory detection limit. Data with one or more remark codes indicating poor data quality were excluded. Data with a remark code indicating laboratory censoring were retained. There were about 65 unique remark codes in the data indicating laboratory censoring; these were unified under a single remark code of “<.” In some cases, there was no remark code provided, but a comment field associated with a data value indicated that the value was censored; these values also were given a remark code of “<.”

Some sources reported a censored remark code for an observation but did not report an associated value. When more than 1 percent of the data for a parameter at a site had a censored remark code but no values, then all of the data (censored and uncensored) for that site and parameter were excluded. When less than 1 percent of the data for a parameter at a site had a censored remark code with a missing value, and censored values were reported for other observations, then the missing censored values were set to the reported censored value that was closest in time for that site and parameter.

In some cases when an agency reported data with censored remark codes, units were provided for uncensored data but not for censored data. Again, the reasons for this practice are unknown, but it creates the potential for substantially biased analyses if only the censored data were excluded because the units were missing. If all of the missing units were for censored data for a parameter from a given source, and all uncensored data were reported in the same units throughout the record (for example, all uncensored data for a parameter were reported in units of milligrams per liter or micrograms per liter), then those same units were assumed to apply to all censored data for that parameter from that source. If necessary, all retained data were converted to milligrams per liter.

Duplicate Samples

It was not uncommon to have received the same water-quality data for the same site and collection date from more than one source during the data compilation effort. Duplicates were defined as samples from the same site, date, time, and fraction. Duplicate samples were removed from the dataset. It was not possible to determine which of the duplicate samples should be removed based solely on data source because sometimes a source provided longer or more completely populated records for some parameters (for example, there may have been duplicate data obtained from STORET and provided directly by a State agency, but overall the STORET record was longer for SSC and the State record was longer for TSS). As a result, each parameter was considered separately when deciding which duplicate samples to retain and which to exclude. Note that duplicate samples did not always have identical values for a given parameter. If one of the duplicate samples was from USGS and one was from another agency, the USGS sample was retained. If both were USGS samples (for example, the duplicate samples were obtained from two different USGS NWIS servers), the sample from the primary host server (home State for each site) was retained. If information on the primary host server was not available, one of the duplicate samples was randomly retained. If both samples were from agencies other than USGS, and one sample had been provided by the CBP, then the sample from the CBP was retained. If both samples were from agencies other than USGS and the CBP, then the sample from the longest period of record was retained. Note that priority was not given to USGS samples because of an a priori expectation that USGS data were of better quality, but because typically more metadata were available on how samples were collected, analyzed, and stored.

Zero, Negative, and Missing Concentration Values

An analytical determination of zero or negative concentration is not possible, so zero or negative values in the data were considered suspect. For a given site and parameter combination, if a censored remark code was frequently associated with a zero or negative concentration value, all zero and negative values were assumed to represent censored values and were set to the reported censored value that was closest in time for that site and parameter combination. In addition, missing values that were reported with a censored remark code were assumed to be censored values without an associated laboratory detection limit. These values were also set to the reported censored value that was closest in time for that site and parameter combination. If there were no reported censored values for a site and parameter combination, then zero, negative, or missing censored concentration values were set to the lowest uncensored value for that site and parameter combination, and a censored remark code was assigned.

In some cases, these assumptions may have resulted in an incorrect value. For example, the zero, negative, or missing censored values may have been from samples that were not analyzed in the same laboratory as other samples for the given parameter or samples for which there was some laboratory error. Alternatively, the assumed detection limit assigned to these values may be wrong. The approach used in this study to populate zero, negative, or missing censored values was unlikely to affect the subsequent trend results when such values collectively constituted a small fraction of the total data for a site and parameter combination; therefore, when zero, negative, and missing censored values made up more than 1 percent of the overall dataset for a site and parameter combination, all data for that site and parameter combination were not included in the determination of trend.

Identification of a Colocated or Nearby Streamgage

Following the steps described in the previous subsections of the "Sediment Data Processing" section, all site and parameter combinations with at least 3 years of consecutive data with at least quarterly sampling (see "Coverage of Changing Seasons" section in this report for details on how quarters were defined) were paired with a streamgage from the final site-streamgage list assembled in the "Streamgage Selection" section of this report. Any site not meeting these minimum data requirements and without an acceptable streamgage match was excluded from further consideration.

Merging Colocated or Neighboring Water-Quality Sites

After assigning a streamgage to each site, the retained sites represented two classes of sites: (1) sites that had a sufficient amount of data over a long enough period of record (at least for the shortest trend period of 10 years) to be used on their own for trend analysis (single sites), and (2) sites that had data for a shorter period of record (at least 3 consecutive years of quarterly data) that could be combined into a composite site with either other colocated or neighboring composite sites to form a single record for trend analysis or a colocated or neighboring stand-alone site to increase the number of trend periods that could be evaluated at the stand-alone site. Single and composite sites were only combined when additional trend periods could be evaluated after combining; they were not combined to only fill in short gaps or add to the existing data at the single site when the combination did not increase the number of trend periods that could be evaluated. Data for periods shorter than 3 years were not used to fill in gaps or extend the record at other composite or single sites because it was decided that the nonrandom variability introduced into those types of combined time series would outweigh the benefits during trend analysis.

Following methods described in the "Streamgage Matching" section, potential colocated and neighboring sites were identified, and sites with intervening influences were excluded. A maximum of two sites was combined to minimize nonrandom variability. If more than two sites were available, then the two sites with the longest period of record for the parameter of interest were selected for the composite site. If more than one of the remaining sites merged well with the primary record, then the site with the longest record was retained. A list of composite sites with the individual sites for each parameter used in trend analysis can be found in appendix 1, table 1–7. The subsequent processing steps described below were applied to the data record for merged sites and single sites.

Degree of Censoring

Following guidelines suggested for the trend method used in this study, site and parameter combinations with more than 50 percent censoring were excluded from trend analysis (Hirsch and De Cicco, 2015). Left-censored and interval-censored values, if less than 50 percent of the dataset, were included in this determination. Right-censored data were removed from the dataset.

Outliers

A small number of values in the data appeared to be typographical errors. For example, consider a case when reported SSC values in a river are between 45,000 mg/L and 51,000 mg/L for several days, and then an SSC value of 660 mg/L is reported, followed by 2 days with reported SSC values between 55,000 and 60,000 mg/L. It is extremely unlikely that SSC values would drop so low just for 1 intervening day. It is far more likely that there was a typographical error, and the true value is 66,000 mg/L. In contrast, a SSC value of 123,000 mg/L is more ambiguous. That concentration is very high relative to concentrations found in most natural waters, but it could occur during locally extreme conditions such as a large storm event, or in a river with high suspended sediment loads. For example, during an initial evaluation of SSC data for this study, a survey of over 789,000 records in NWIS between 1965 and 2014 determined that the median SSC value was 58 mg/L and the 95th percentile SSC value was 3,070 mg/L (U.S. Geological Survey, 2016). However, at USGS Site 08353000, the Rio Puerco near Bernardo, New Mexico, the median SSC value was 32,200 mg/L and the 95th percentile SSC value was 179,000 mg/L (U.S. Geological Survey, 2016). In a smaller study involving far fewer sites and data sources, it might be possible to definitively establish the accuracy of ambiguous data points by checking laboratory records, field forms, or other sources of historical information, but that was not possible in a study of this size because of resource constraints and only the most extreme outliers—those that were almost certainly typographical errors—were excluded. Under a normal distribution, 99.9999998 percent of the data should be within six standard deviations from the mean; only the most extreme values will be outside of that range. For each site and parameter pair, the data were divided into four seasons. Within each season, any data value further than six standard deviations from the mean in log space were excluded as a likely typographical error. With this approach, there was a risk that typographical errors representing less extreme values were retained. There was also a risk that values representative of true environmental conditions were excluded.

Data Thinning

The sediment trend model used in this study weights days with multiple water-quality samples more heavily than other days during calibration. As a result, data for each site and parameter combination were thinned to have no more than one sample per day. If there was more than one sample per day, one sample was randomly selected to be retained in the dataset. The presence of data points spaced closely together in time also can introduce serial correlation into a time series. With some trend methods, serial correlation can increase the probability of incorrectly rejecting the test's null hypothesis of no trend, leading to an incorrect determination that a significant trend has occurred (Hirsch and Slack, 1984). The trend method in this study relied on bootstrapping to determine p-values (see "Nutrient, Sediment, Major Ion, Salinity, and Carbon Trend Analysis Method" section in this report), an approach that is less affected by serial correlation. For that reason, the datasets were not thinned for data points close in time on different days.

Final Screening for Data Coverage

Once data from colocated or neighboring sites had been merged, and the censoring, outlier, and data thinning steps had been completed, the resulting data series were screened again by parameter. In this final screening step, the requirements for data coverage were more stringent. All sites not meeting these criteria were excluded.

Coverage Throughout Target Trend Periods

- The target trend periods, in water years, were (1) 1972–2012, (2) 1982–2012, (3) 1992–2012, and (4) 2002–12. Data were permitted to begin in the specified start year or the following year. For example, a start year of either 1972 (October 1, 1972) or 1973 (October 1, 1972) was acceptable in trend period 1, but 1974 (October 1, 1973) was not. Data were permitted to end in the specified start year or the preceding year. An end year of either 2011 (September 30, 2011) or 2012 (September 30, 2012) was acceptable in all trend periods, but 2010 was not.
- The first 2 years and the last 2 years of a trend period were required to have at least quarterly samples. See the "Coverage of Changing Seasons" section below for more detail on how quarters were defined.
- Overall, 70 percent of all years during a trend period were required to have at least quarterly samples. Any gap must have occurred in the middle of the record.
- Longer gaps in quarterly data were permitted when there was at least a 10-year span meeting the criteria beginning in any of the start years (for example, data from 1972 to 1981 or 1982 to 1991) and a 10-year span meeting the criteria ending in either 2011 or 2012. Data during each of those decades must have met the other detailed criteria. Gaps of any length were permitted in the years between the beginning and ending decades.
- At least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during a given trend period in order to ensure the full range of flows were represented. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. For the shortest trend period (2002–12) of one decade, at least 14 percent of the samples were required to be high-flow samples. See the "Coverage of Changing Streamflow" section below for more detail on how high-flow samples were defined.

Coverage of Changing Seasons

Suspended sediment concentrations vary throughout the year in response to changes in temperature and sediment sources (Walling, 2005; Garcia, 2007). Samples representing seasonal conditions throughout the year are needed to fully characterize changes in concentration. To ensure seasonal coverage, 70 percent of years during a trend period were required to have at least one sample per quarter; the remaining 30 percent of years could be missing some or all quarterly samples. This allowed for reasonable changes in sampling design over time.

To evaluate seasonal coverage, the water year was divided into quarters using three different divisions of the water year (table 4). Each quarter consisted of 3 calendar months (table 4). A site was accepted if the samples met any one of the quarterly definitions in table 4; the quarterly definition was permitted to vary from year to year. This method of testing for quarterly samples accommodated variations in real-world sampling schedules. The primary sampling-design artifact this accommodated was the occurrence of two samples in one quarter and none in the preceding or following quarter, which may have occurred if a sample was collected slightly earlier or slightly later than dictated by a strict 3-month schedule. As an example, consider a site sampled in January, late March, July, and October during a particular year. The source collecting these data is clearly collecting quarterly samples; however, the sample that should have been collected in April was collected in late March. Without a flexible quarter definition, this year would not have had quarterly sampling. This flexible method of defining quarters also provides a buffer for sources that sample rivers that may freeze in the winter. For various reasons, including the weather, a winter quarterly sample may need to be collected anytime between November and February.

Coverage of Changing Streamflow

Sediment concentrations also vary in response to changes in streamflow (Williams, 1989; Meybeck and others, 2003). Samples representing a range of streamflow conditions are needed to fully characterize changes in concentration over time. Low- or moderate-flow samples are not in short supply. To identify high-flow samples, streamflow on the date of each sample was compared to the 85th percentile of flows in that month during the corresponding decade of the trend period at that site (depending on the trend period, that would include some or all of the following decades: 1972–81, 1982–91, 1992–2001, and 2002–12). For example, for a sample collected on March 1, 1998, the 85th percentile of all daily streamflows in March during the decade 1992–2001 was determined. If the streamflow on March 1, 1998, was greater than the 85th percentile of March streamflows during that decade, the sample was considered to be a high-flow sample. This assessment was completed by decade rather than once for the full trend period to ensure that samples at high streamflows were distributed throughout the trend period, rather than potentially being concentrated in one part of it. It also allowed for the identification of high-flow samples collected during prolonged low-flow periods (for example, droughts). Such high-flow samples would have been collected during flows that were elevated relative to an unusually low baseline; the flows may not have been particularly high relative to overall average conditions at a site.

To ensure appropriate streamflow coverage, at least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during the trend period. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. The relaxation of the 14-percent criterion to 10 percent in some decades was necessary to accommodate sites with changing sampling designs over time (for example, the collection of high-flow samples increased somewhat or decreased somewhat over time). For the shortest trend period (2002–12) of only one decade, at least 14 percent of the samples were required to be high-flow samples during that period. These thresholds were based on sensitivity testing described in the "Storm Sampling" section of this report.

Data Used in Model Calibration

For site and parameter combinations that met the screening criteria, all available data before, during, and after the trend period (and if applicable, during any gaps) were retained to improve the model calibration (see the "Influence of New Data on Trend Estimates" section). Ultimately, model estimates only during the screened trend period were retained. In the case of longer records with data in two different decades separated by gaps, model estimates made in the years between the beginning and ending decades were excluded. Model input data are available in De Cicco and others (2017).

Multiple steps were used to determine the final trend model input values used in this study. Note that because of the use of unrounded data from NWIS and the use of interval censoring, the final USGS values used in this study will not always match the corresponding value stored in the NWIS database.

There were 611 sites evaluated for sediment trends with 620 combinations of sites and parameters. Table 5 summarizes the number of sites evaluated for each parameter and trend period. A complete list of sites and parameters evaluated for sediment trends is reported in appendix 1, table 1–6.

Major Ion and Salinity Parameters of Interest and Harmonization

The target parameters for trend analysis were chloride, sulfate, calcium, magnesium, potassium, sodium, specific conductance (SC), and total dissolved solids (TDS). Parameters that clearly were not a variation on the target parameters (for example, sodium plus potassium and sodium adsorption ratio) were excluded. Some major ion and salinity parameters were ambiguous and were only used when the source specified their interpretation of the parameter (for example, some sources specified that solids be used to mean TDS and others used solids but did not specify their interpretation).

Naming conventions for a single parameter varied widely among data sources. For example, all of the variations on sulfate are shown in appendix 2, table 2–3. Numerous steps were taken to harmonize different parameter names that represented the same chemical. In some cases, the chemical represented by the parameter name was not clear. The majority of chloride, sulfate, calcium, magnesium, potassium, and sodium records were unambiguous, but SC and TDS records required more effort to harmonize. Metadata critical to identifying and quantifying each data value—such as filtration status, units, and chemical form (molecular or elemental)—often were incomplete, rendering the data value ambiguous. When any of these critical metadata were missing, the data were not used. In select cases, agencies were contacted to fill in metadata gaps.

Unlike the other parameters, SC is often determined in the field. The SC should be temperature corrected (Hem, 1985), and unless otherwise specified, values were assumed to be corrected. Records that were obviously not temperature corrected were excluded. Field samples were preferred over laboratory values when both were present.

The TDS values are commonly determined using one of three methods: (1) multiprobe, (2) sum of parameters, or (3) gravimetric (residue on evaporation at 105 °C or 180 °C). The sum of parameters and gravimetric methods measure the dissolved mineral matter in the water, but when TDS is determined using a multiprobe meter, SC (electrical conductance) is actually the parameter being measured, and a conversion factor is used to calculate TDS. While electrical conductance is related to ion concentrations, the relation cannot be expected to be simple in natural waters (Hem, 1985). The conversion factor commonly used to determine TDS from SC is 0.6, but the actual ratio of SC to TDS in natural waters generally varies between 0.55 and 0.75 (and can vary between 0.54 and 0.96) (Hem, 1985). This relation ideally should be determined for each site, and the use of a generic conversion factor can produce incorrect TDS values, especially if there is a large range in SC at a site. The TDS values determined by gravimetric methods and by summation of parameter concentrations are comparable (Hem, 1985; Tillman and Anning, 2014); therefore, only TDS values determined in the laboratory using gravimetric methods or sum of parameter methods were used in this study. Parameter code, method code, parameter name, and filtration status were used to identify TDS values that were determined by sum of parameters or combustion methods. In many cases, it was unclear what the parameter name TDS was intended to represent. When the meaning of TDS could not be determined unambiguously, the associated data were not used.

Any data values that were reported without units, such that it could not be unambiguously determined whether they were reported in milligrams per liter, micrograms per liter, or some other unit were excluded. Any sulfate data values that were reported without form, such that it could not be unambiguously determined whether they were reported in elemental form (for example, sulfate as sulfur [as S]) or in molecular form (for example, sulfate as SO_4^{2-}) also were excluded. In both cases, the choice of units or form commonly could alter the value substantially. For example, a sulfate concentration reported as SO_4^{2-} would be three times larger than the same sulfate concentration reported as S.

All remaining chloride, sulfate, calcium, magnesium, potassium, sodium, and TDS concentration data with complete and unambiguous metadata were converted to milligrams per liter. The SC concentration values were converted to micromohs per centimeter (µmohs/cm). The microsiemens (µS) is numerically the same as the micromoh (Hem, 1985). Converted concentrations were rounded to the number of decimal places in the original value, with the exception of values converted only from micrograms per liter to milligrams per liter (because of the possibility of rounding a converted value to zero).

Major Ion and Salinity Data Processing

In order to reduce the nutrient dataset to a more reasonable size and to eliminate sites with very limited data that were not suitable for trend analysis, only sites that had at least 3 years of consecutive nutrient data with at least quarterly sampling initially were retained. Details on how quarters were defined can be found in the “Coverage of Changing Seasons” section in this report.

Unfiltered and Filtered Fractions

Combining fractions (unfiltered, filtered, total, and unknown) may be necessary to extend the period of record in time or fill gaps in time, but it does introduce variability, which can make trend determination more difficult. Because filtered samples composed at least 80 percent of the total samples for calcium, magnesium, potassium, and sodium, fractions were not combined and only filtered (dissolved) values were used.

The SC measurements should be made in-situ or in unfiltered samples whenever possible (Hem, 1985). Filtered or dissolved measurements constituted a very small amount (less than 10 percent) of the total SC measurements and were excluded from this study. The TDS concentrations were used only when it could be determined that the result was based on a filtered sample using filtration status, parameter name, and (or) analysis method.

Discrete Samples

Discrete samples were used in the study to avoid introducing sampling inconsistency. Composite samples—multiple discrete samples collected over an extended period of time (often more than 24 hours or throughout a storm event) that were composited together into a single sample—were excluded from the study.

Samples Analyzed in the Field

To maintain basic comparability in precision and accuracy of analytical determinations, data were restricted to samples analyzed in laboratories (except for SC where field values were preferred). Remark codes on samples (for example, Virginia Department of Environmental Quality remark code H: "Value based on field kit determination; may not be accurate") and sample type ("Field Msr/Obs") were used to identify samples that were analyzed in the field. It was particularly important to remove TDS values that had been determined in the field using a multiprobe that was actually measuring SC and using a generic formula to calculate TDS. As mentioned in the "Major Ion and Salinity Parameters of Interest and Harmonization" section, parameter code, method code, and parameter name were used to verify that TDS values were determined in the laboratory. Information on whether the samples were analyzed in the field or laboratory was not always provided, so some samples that were analyzed in the field may have been inadvertently retained.

Remark Code Harmonization

There were over 1,400 unique remark codes in the original dataset including all parameters. Many of these were not defined, and many others were ambiguously defined. Data with undefined or ambiguous remark codes were excluded. Of most interest for this study were remark codes indicating poor data quality (for example, "Contamination present in the sample") and remark codes indicating censored data or data below a laboratory detection limit. Data with one or more remark codes indicating poor data quality were excluded. Data with a remark code indicating laboratory censoring were retained. There were about 65 unique remark codes in the data indicating laboratory censoring; these were unified under a single remark code of "<." In some cases, there was no remark code provided, but a comment field associated with a data value indicated that the value was censored; these values also were given a remark code of "<."

Some sources reported a censored remark code for an observation but did not report an associated value. When more than 1 percent of the data for a parameter at a site had a censored remark code but no values, then all of the data (censored and uncensored) for that site and parameter were excluded. When less than 1 percent of the data for a parameter at a site had a censored remark code with a missing value, and censored values were reported for other observations, then the missing censored values were set to the reported censored value that was closest in time for that site and parameter.

In some cases when an agency reported data with censored remark codes, units were provided for uncensored data but not for censored data. Again, the reasons for this practice are unknown, but it creates the potential for substantially biased analyses if only the censored data were excluded because the units were missing. If all of the missing units were for censored data for a parameter from a given source, and all uncensored data were reported in the same units throughout the record (for example, all uncensored data for a parameter were reported in units of milligrams per liter or micrograms per liter), then those same units were assumed to apply to all censored data for that parameter from that source. If necessary, all retained data were converted to milligrams per liter.

Recensoring U.S. Geological Survey Data

On October 1, 1998, the NWQL began implementing new reporting conventions for nondetections (Oblinger Childress and others, 1999). Prior to October 1, 1998, nondetections were censored to a MDL, the concentration at which the risk of a false positive detection (analyte reported as present at the MDL when not in the sample) is no more than 1 percent, or to a MRL, an inconsistently defined level of detection capability. The NWQL deemed the use of inconsistent MRLs as inappropriate and also observed that the risk of a false negative occurrence at the MDL (analyte reported as not present when present at the MDL concentration) could be as much as 50 percent. To make the determination of detection limits more consistent and to reduce the risk of false negatives to no more than 1 percent, the NWQL began using a LRL in combination with a LT–MDL on October 1, 1998, based on Oblinger Childress and others (1999). The LRL typically was twice the LT–MDL. Between October 1, 1998, and October 1, 2010, concentrations measured between the LRL and the LT–MDL were reported as estimated concentrations with an "E" remark code and nondetections were censored to the higher LRL. Beginning on October 1, 2010, the NWQL censored inorganic nondetections to the LT–MDL (https://water.usgs.gov/admin/memo/QW/qw10.07.html).

The reporting convention in place between October 1, 1998, and October 1, 2010, creates potential bias in statistical analyses because it improperly represents the proportions of the data below the LT–MDL and the LRL (Helsel, 2005). In particular, any observations that are truly below the LT–MDL will be reported as censored at the higher LRL; thus, the probability that an observation falls between the LT–MDL and the LRL may be overestimated, and the probability that it falls below the LT–MDL may be underestimated (Helsel, 2005). Depending on the true percentage of data below the LT–MDL, this can cause an upward bias that will adversely affect statistical analyses when these affected data are pooled with data subject to conventional censoring (that is, data not recensored from the LT–MDL to the LRL). To avoid this problem, there are two potential solutions: (1) The LT–MDL becomes the reporting limit. Values reported as censored at the higher LRL are recensored to the lower LT–MDL. Uncensored "E" values between the LT–MDL and the LRL are not recensored. (2) The LRL becomes the reporting limit. All uncensored "E" values below the LRL are censored at the LRL. To retain the extra information captured by the "E" values, the first solution was used for USGS data from NWQL that were originally reported using the LRL approach. The "E" remark code may have been applied to other values for unrelated performance-based reasons, as described in Oblinger Childress and others (1999); these values were unchanged.

Similar to the description in the “Adjustments to U.S. Geological Survey Data Based on Historical Memorandums and Reports” section, these adjustments were applied to data from the NWIS database that had the analyzing entity defined as NWQL or when the analyzing entity was missing. For data analyzed at other laboratories, metadata documenting reporting conventions for censored data were not readily available; therefore, all censored data from laboratories other than the NWQL were unaltered.

Duplicate Samples

It was not uncommon to have received the same water-quality data for the same site and collection date from more than one source during the data compilation effort. Duplicates were defined as samples from the same site, date, time, and fraction. Duplicate samples were removed from the dataset. It was not possible to determine which of the duplicate samples should be removed based solely on data source because sometimes one source provided longer or more completely populated records for some parameters but not all (for example, there may have been duplicate data that were obtained from STORET and provided directly by a State agency, but overall the STORET record was longer for sulfate and the State record was longer for chloride). As a result, each parameter was considered separately when deciding which duplicate samples to retain and which to exclude. Note that duplicate samples did not always have identical values for a given parameter. If one of the duplicate samples was from USGS and one was from another agency, the USGS sample was retained. If both were USGS samples (for example, the duplicate samples were obtained from two different USGS NWIS servers), the sample from the primary host server (home State for each site) was retained. If information on the primary host server was not available, one of the duplicate samples was randomly retained. If both samples were from agencies other than USGS, and one sample had been provided by the CBP, then the sample from the CBP was retained. If both samples were from agencies other than USGS and the CBP, then the sample from the longest period of record was retained. Note that priority was not given to USGS samples because of an a priori expectation that USGS data were of better quality, but because typically more metadata were available on how samples were collected, analyzed, and stored.

Zero, Negative, and Missing Concentration Values

An analytical determination of zero or negative concentration is not possible, so zero or negative values in the data were considered suspect. For a given site and parameter combination, if a censored remark code was frequently associated with a zero or negative concentration value, all zero and negative values were assumed to represent censored values and were set to the reported censored value that was closest in time for that site and parameter combination. In addition, missing values that were reported with a censored remark code were assumed to be censored values without an associated laboratory detection limit. These values were also set to the reported censored value that was closest in time for that site and parameter combination. If there were no reported censored values for a site and parameter combination, then zero, negative, or missing censored concentration values were set to the lowest uncensored value for that site and parameter combination and a censored remark code was assigned.

In some cases, these assumptions may have resulted in an incorrect value. For example, the zero, negative, or missing censored values may have been from samples that were not analyzed in the same laboratory as other samples for the given parameter, or samples for which there was some laboratory error. Alternatively, the assumed detection limit assigned to these values may be wrong. The approach used in this study to populate zero, negative, or missing censored values was unlikely to affect the subsequent trend results when such values collectively constituted a small fraction of the total data for a site and parameter combination; therefore, when zero, negative, and missing censored values made up more than 1 percent of the overall dataset for a site and parameter combination, all data for that site and parameter combination were not included in the determination of trend.

Identification of a Colocated or Nearby Streamgage

Following the steps described in the previous subsections of the “Major Ion and Salinity Data Processing” section, all site and parameter combinations with at least 3 years of consecutive data with at least quarterly sampling (see “Coverage of Changing Seasons” section in this report for details on how quarters were defined) were paired with a streamgage from the final site-streamgage list assembled in the “Streamgage Selection” section of this report. Any site not meeting these minimum data requirements and without an acceptable streamgage match was excluded from further consideration.

Merging Colocated or Neighboring Water-Quality Sites

For each target parameter, sites that had at least 3 years of consecutive data with at least quarterly sampling (see “Coverage of Changing Seasons” section in this report for details on how quarters were defined) were retained initially. All other sites were excluded. The retained sites represented two classes of sites: (1) sites that had a sufficient amount of data over a long enough period of record to possibly be used on their own for trend analysis (stand-alone sites), and (2) sites that had a shorter period of record (composite sites) that could be combined with other colocated or neighboring composite sites to form a single record for trend analysis or a colocated or neighboring stand-alone site to increase the number of trend periods that could be evaluated at the stand-alone site. Stand-alone and composite sites were only combined when additional trend periods could be evaluated after combining;

they were not combined to fill in short gaps or otherwise add to the existing data at the stand-alone site. Data from a composite site could have been used to fill in short gaps or add to the existing data at the stand-alone site; likewise, data covering periods shorter than 3 years could have been used to fill in gaps or extend the record at other composite or stand-alone sites because it was decided that the nonrandom variability introduced into those types of combined time series would outweigh the benefits during trend analysis.

Following methods described in the "Streamgage Matching" section, potential colocated and neighboring sites were identified, and sites with intervening influences were excluded. A maximum of two sites was combined to minimize nonrandom variability. If more than two sites were available, then the primary site (the site with the longest record) was identified, and each of the remaining sites were prioritized based on how well the data merged with the primary record. If more than one of the remaining sites merged well with the primary record, then the site with the longest record was retained. A list of composite sites with the individual sites for each parameter used in trend analysis can be found in appendix 1, table 1–7. The subsequent processing steps described below were applied to the data record for merged sites and any remaining stand-alone sites.

Degree of Censoring

Following guidelines suggested for the trend method used in this study, site and parameter combinations with more than 50 percent censoring were excluded from trend analysis (Hirsch and De Cicco, 2015). Left-censored and interval-censored values, if less than 50 percent of the dataset, were included in this determination. Right-censored data were removed from the dataset.

Outliers

A small number of values in the dataset appeared to be typographical errors. For example, consider a case when reported SC concentrations in a river are between 260 µmohs/cm and 300 µmohs/cm for several days, and then a SC concentration of 0.29 µmohs/cm is reported, followed by a week with reported concentrations between 280 µmohs/cm and 320 µmohs/cm. It is extremely unlikely that SC concentrations would drop so low just for one intervening day. It is far more likely that there was a typographical error, and the true value is 290 µmohs/cm. In contrast, a SC concentration of 50,000 µmohs/cm is more ambiguous. That concentration is very high relative to concentrations found in most natural waters and is the approximate conductance of seawater (Hem, 1985), but it could occur in areas where subsurface brines contribute to surface water. In a smaller study involving far fewer sites and data sources, it might be possible to definitively establish the accuracy of ambiguous data points by checking laboratory records, field forms, or other sources of historical information, but that was not possible in a study of this size because of resource constraints, and only the most extreme outliers—those that were almost certainly typographical errors—were excluded. Under a normal distribution, 99.9999998 percent of the data should be within six standard deviations from the mean; only the most extreme values will be outside of that range. For each site and parameter pair, the data were divided into four seasons. Within each season, any data value further than six standard deviations from the mean in log space was excluded as a likely typographical error. With this approach, there was a risk that typographical errors representing less extreme values were retained. There was also a risk that values representative of true environmental conditions were excluded.

Data Thinning

The major ion and salinity trend model used in this study weights days with multiple water-quality samples more heavily than other days during calibration. As a result, data for each site and parameter combination were thinned to have no more than one sample per day. If there was more than one sample per day, one sample was randomly selected to be retained in the dataset. The presence of data points spaced closely together in time also can introduce serial correlation into a time series. With some trend methods, serial correlation can increase the probability of incorrectly rejecting the test's null hypothesis of no trend, leading to an incorrect determination that a significant trend has occurred (Hirsch and Slack, 1984); The trend method in this study relied on bootstrapping to determine p-values (see "Nutrient, Sediment, Major Ion, Salinity, and Carbon Trend Analysis Method" section in this report), an approach that is less affected by serial correlation. For that reason, the datasets were not thinned for data points close in time on different days.

Final Screening for Data Coverage

Once data from colocated or neighboring sites had been merged, and the censoring, outlier, and data thinning steps had been completed, the resulting data series were screened again by parameter. In this final screening step, the requirements for data coverage were more stringent. All sites not meeting these criteria were excluded.

Coverage Throughout Target Trend Periods

- The target trend periods, in water years, were (1) 1972–2012, (2) 1982–2012, (3) 1992–2012, and (4) 2002–12. Data were permitted to begin in the specified start year or the following year. For example, a start year of either 1972 (October 1, 1971) or 1973 (October 1, 1972) was acceptable in trend period 1, but 1974 (October 1, 1973) was not. Data were permitted to end in the specified start year or the preceding year. An end year of either 2011 (September 30, 2011) or 2012 (September 30, 2012) was acceptable in all trend periods, but 2010 was not.

- The first 2 years and the last 2 years of a trend period were required to have at least quarterly samples. See the “Coverage of Changing Seasons” section below for more detail on how quarters were defined.
- Overall, 70 percent of all years during a trend period were required to have at least quarterly samples. Any gap must have occurred in the middle of the record.
- Longer gaps in quarterly data were permitted when there was at least a 10-year span meeting the criteria beginning in any of the start years (for example, data from 1972 to 1981 or 1982 to1991) and a 10-year span meeting the criteria ending in either 2011 or 2012. Data during each of those decades must have met the other criteria detailed here. Gaps of any length were permitted in the years between the beginning and ending decades.
- At least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during a given trend period in order to ensure the full range of flows were represented. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. For the shortest trend period (2002–12) of one decade, at least 14 percent of the samples were required to be high-flow samples. See the “Coverage of Changing Streamflow” section below for more detail on how high-flow samples were defined.

Coverage of Changing Seasons

Major ion and salinity concentrations vary throughout the year in response to changes in temperature (evaporation and evapotranspiration) and source water (Hem, 1985; Drever, 1997; Anning, 2003; Kaushal and others, 2005; Anning and others 2006). Samples representing seasonal conditions throughout the year are needed to fully characterize changes in concentration. To ensure seasonal coverage, 70 percent of years during a trend period were required to have at least one sample per quarter; the remaining 30 percent of years could be missing some or all quarterly samples. This allowed for reasonable changes in sampling design over time.

To evaluate seasonal coverage, the water year was divided into quarters using three different divisions of the water year (table 4). Each quarter consisted of 3 calendar months (table 4). A site was accepted if the samples met any one of the quarterly definitions in table 4; the quarterly definition was permitted to vary from year to year. This method of testing for quarterly samples accommodated variations in real-world sampling schedules. The primary sampling-design artifact this accommodated was the occurrence of two samples in one quarter and none in the preceding or following quarter, which may have occurred if a sample was collected slightly earlier or slightly later than dictated by a strict 3-month schedule. As an example, consider a site sampled in January, late March, July, and October during a particular year. The source collecting these data is clearly collecting quarterly samples; however, the sample that should have been collected in April was collected in late March. Without a flexible quarter definition, this year would not have had quarterly sampling. This flexible method of defining quarters also provides a buffer for sources that sample rivers that may freeze in the winter. For various reasons, including the weather, a winter quarterly sample may need to be collected anytime between November and February.

Coverage of Changing Streamflow

Major-ion concentrations and salinity also vary in response to changes in streamflow (Hem, 1985; Langmuir, 1997). Samples representing a range of streamflow conditions are needed to fully characterize changes in concentration over time. Low- or moderate-flow samples are not in short supply. To identify high-flow samples, streamflow on the date of each sample was compared to the 85th percentile of flows in that month during the corresponding decade of the trend period at that site (depending on the trend period, that would include some or all of the following decades: 1972–81, 1982–91, 1992–2001, and 2002–12). For example, for a sample collected on March 1, 1998, the 85th percentile of all daily streamflows in March during the decade 1992–2001 was determined. If the streamflow on March 1, 1998, was greater than the 85th percentile of March streamflows during that decade, the sample was considered to be a high-flow sample. This assessment was completed by decade rather than once for the full trend period to ensure that samples at high streamflows were distributed throughout the trend period, rather than potentially being concentrated in one part of it. It also allowed for the identification of high-flow samples collected during prolonged low-flow periods (for example, droughts). Such high-flow samples would have been collected during flows that were elevated relative to an unusually low baseline; the flows may not have been particularly high relative to overall average conditions at a site.

To ensure appropriate streamflow coverage, at least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during the trend period. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. The relaxation of the 14 percent criterion to 10 percent in some decades was necessary to accommodate sites with changing sampling designs over time (for example, the collection of high-flow samples increased somewhat or decreased somewhat over time). For the shortest trend period (2002–12) of only one decade, at least 14 percent of the samples were required to be high-flow samples during that period. These thresholds were based on sensitivity testing described in the “Storm Sampling” section of this report.

Data Used in Model Calibration

For site and parameter combinations that met the screening criteria, all available data before, during, and after the trend period (and if applicable, during any gaps) were retained to improve the model calibration (see the "Influence of New Data on Trend Estimates" section). Ultimately, model estimates only during the screened trend period were retained. In the case of longer records with data in two different decades separated by gaps, model estimates made in the years between the beginning and ending decades were excluded. Model input data are available in De Cicco and others (2017).

Multiple steps were used to determine the final trend model input values used in this study. Note that because of the recensoring of some USGS data, the use of unrounded data from NWIS, and the use of interval censoring, the final USGS values used in this study will not always match the corresponding value stored in the NWIS database.

There were 1,198 sites evaluated for major ion and salinity trends with 3,024 combinations of sites and parameters. Table 5 summarizes the number of sites evaluated for each parameter and trend period. A complete list of sites and parameters evaluated for major ion and salinity trends is reported in appendix 1, table 1–6.

Carbon Parameters of Interest and Harmonization

Carbon exists in inorganic and organic forms in aquatic ecosystems. Organic carbon refers to a wide range of chemical compounds containing at least one, and usually multiple, reduced carbon atoms (Drever, 1997). Organic carbon is the intact or partially decomposed remains of organisms including plants, microbes, and animals. In contrast, inorganic carbon contains only fully oxidized carbon atoms and exists in most environmental samples as bicarbonate, carbonate, and carbon dioxide (Butler, 1982). Alkalinity was used as the measurement of inorganic carbon because in natural waters bicarbonate is the primary contributor to alkalinity and typically the largest part of the inorganic carbon pool. Although carbon dioxide is an important component of inorganic carbon, it was not included in the analysis of inorganic carbon because a carbon dioxide concentration is highly dynamic, exchanging rapidly with organic material and with the atmosphere such that an analysis for trends in concentration is problematic.

The target parameters for trend analysis were alkalinity, dissolved organic carbon (DOC), and total organic carbon (TOC). Because these analytes were sometimes reported directly and sometimes calculated from other parameters, all available data were retained for alkalinity, acid neutralizing capacity, acidity, bicarbonate, carbonate, dissolved organic carbon, particulate organic carbon, and total organic carbon. Naming conventions for a single parameter varied widely in the data. For example, some of the variations on alkalinity are shown in appendix 2, table 2–4. When a reported parameter name could be linked to a known parameter, the parameter names and units were harmonized. Numerous steps were taken to ensure that parameter names were properly represented in the harmonized dataset. In many cases, some aspect of the metadata (filtration status, units, chemical form) was unclear. When any of the critical metadata were missing, the data were considered incompatible with the larger dataset and were excluded from analysis.

The metadata required for organic carbon analysis were similar to other parameters included in this study. It was critical to know filtration status and units on all samples. Trends in TOC and DOC were analyzed separately because they express different fractions of carbon within samples. For example, a change in suspended organic material can induce a trend in TOC but is not likely to induce trends in DOC. Similarly, changing sample fractions from filtered to unfiltered could induce a trend because particles are included in the latter analysis.

The harmonization of alkalinity parameters requires particular focus on the reported units and chemical form and the method of alkalinity analysis. Alkalinity is often reported as milligrams per liter as calcium carbonate ($CaCO_3$), but it can also be reported as microequivalents per liter, milligrams per liter as bicarbonate (HCO_3^-), and others. For consistency, all alkalinity measurements were converted to milligrams per liter as $CaCO_3$. Alkalinity reporting units varied widely and were sometimes difficult to relate to milligrams per liter as $CaCO_3$. In such cases, the data were excluded from analysis. Another common ambiguity in the data was the use of the term "carbonate alkalinity," which can refer to (1) alkalinity arising from the carbonate buffering system, which should be dominant in most natural waters; or (2) alkalinity arising from the presence of carbonate ion at high pH. The latter definition is also sometimes called phenolphthalein or phenol alkalinity because of the historical use of the phenolphthalein indicator to end titration at pH 8.3 (Hem, 1985). In cases when the definition of carbonate alkalinity could not be unambiguously determined from the metadata provided, the samples were excluded.

Alkalinity was commonly determined by sequential titration of a sample to pH 8.3 and then to pH 4.5, which expresses concentrations of CO_3^{2-} and HCO_3^-, respectively. When HCO_3^- was reported, and it could be determined from the metadata that this referred to titration from 8.3 to 4.5, a total alkalinity concentration was determined using pH and temperature information. Total alkalinity based on the carbonate system is the equivalent sum of HCO_3^-, carbonate (CO_3^{2-}), and hydroxide (OH^-). OH^- was assumed to be negligible because it is a small part of total alkalinity at environmentally relevant pH (less than 1 percent when pH is less than 9.5); likewise, at field pH less than 8, CO_3^{2-} is a negligible part of total alkalinity (less than 0.4 percent) (Butler, 1982), and so total alkalinity was assumed to be equal to HCO_3^- when field pH was less than 8. At pH greater than 8, CO_3^{2-} was calculated based on carbonate equilibria, and total alkalinity was assumed to be $CO_3^{2-} + HCO_3^-$.

$$\text{Total alkalinity}\left(\text{mg CaCO}_3\ \text{L}^{-1}\right) = \text{HCO}_3^- + \left(2 \times r\text{CO}_3 \times \text{HCO}_3^-\right) \quad (10)$$

where

HCO_3^- is the given bicarbonate concentration, in milligrams $CaCO_3$ $L^{-1,}$ and

rCO_3 is the molar ratio of CO_3^{2-} to HCO^{3-}.

Calculation of *r*CO3 proceeds from temperature and pH-dependent carbonate equilibria:

$$r\text{CO}_3 = \frac{Ka2}{10^{-\text{pH}}} \quad (11)$$

$$\log Ka2 = -107.8871 - \left(0.03252849K\right) + \left(\frac{5{,}151.79}{K}\right) + \left(38.92561 \log K\right) - \left(\frac{563{,}713.9}{K^2}\right) \quad (12)$$

where

K is field temperature, in degrees Kelvin (Plummer and Busenberg, 1982).

Similar to other parameters used in this study, temperature and pH data were subjected to screening for appropriate units, metadata, and values in the database before using in the carbonate equilibria calculations. If a sample required pH and temperature data for appropriate interpretation and the data were missing or ambiguous, then the sample was excluded from the analysis.

Carbon Data Processing

In order to reduce the carbon dataset to a more reasonable size and to eliminate sites with very limited data that were not suitable for trend analysis, only sites that had at least 3 years of consecutive nutrient data with at least quarterly sampling initially were retained. Details on how quarters were defined can be found in the “Coverage of Changing Seasons” section in this report.

Unfiltered and Filtered Fractions

Carbon exists in the dissolved and particulate phases of water samples. Particulate organic carbon (POC) is composed of sediments of terrestrial or fluvial origin, phytoplankton cells, detritus from macrophytes or higher plants, and larger microbial cells or colonies. Previous studies have shown that the particulate phase represents around 15 percent of total organic material in rivers (Stets and Striegl, 2012), but this amount varies greatly among river systems and at different streamflow conditions; therefore, metadata distinguishing between filtered and unfiltered samples were necessary for the inclusion of organic carbon samples in this trend study. TOC, POC, and DOC were treated as separate parameters.

In contrast to organic carbon, alkalinity exists primarily in the dissolved fraction, with a minor contribution from the particulate phase. Particulate inorganic carbon can arise from sediments or from in-situ formation of $CaCO_3$ particles under conditions of high temperature and pH, but the contribution of this material to overall alkalinity is expected to be small. An analysis done by Stets and Striegl (2012) determined that particulate inorganic carbon was a very small percentage of total inorganic carbon, indicating that the bulk of inorganic carbon, and of alkalinity, resides in the dissolved pool. Based on the results of Stets and Striegl (2012) and in the “Paired Sample Evaluation” section of this report, filtered and unfiltered alkalinity were considered to be indistinguishable in this study and data for both fractions were combined. Alkalinity data also were accepted from agencies that did not specify a fraction in the metadata because either fraction was acceptable. When filtered, unfiltered, and (or) missing fraction results were present in a single sample, the priority for use in trend analysis was unfiltered, then filtered, and then missing.

Discrete Samples

Discrete samples were used in the study to avoid introducing sampling inconsistency. Composite samples—multiple discrete samples collected over an extended period of time (often more than 24 hours or throughout a storm event) that were composited together into a single sample—were excluded from the study.

Samples Analyzed in the Field

To maintain basic comparability in precision and accuracy of analytical determinations, the TOC, DOC, and POC data were restricted to samples analyzed in laboratories. In contrast to other parameters in this study, field alkalinity values obtained by titration are considered to be of high quality, and field titration is the preferred analytical method over laboratory titrations. Field and laboratory alkalinity was compared in paired samples from 92 sites. Most observations fell within 10 percent of each other, and the median difference between laboratory and field alkalinity was 0 percent, and the average difference was 0.1 percent. The largest differences occurred at sites with low alkalinity (less than 30 mg/L as $CaCO_3$) emphasizing that differences largely occur because of imprecision in the methods at low alkalinity or because of errors introduced when rounding. No evidence of sample bias was found between alkalinity samples run in the field compared to the laboratory, so samples analyzed by either of these methods were considered acceptable for this study. Field alkalinity measured by field kits rather than by titration were excluded from this analysis.

Remark Code Harmonization

There were over 1,400 unique remark codes in the original dataset including all parameters. Many of these were not defined, and many others were ambiguously defined. Data with undefined or ambiguous remark codes were excluded. Of most interest for this study were remark codes indicating poor data quality (for example, “Contamination present in the sample”) and remark codes indicating censored data or data below a laboratory detection limit. Data with one or more remark codes indicating poor data quality were excluded. Data with a remark code indicating laboratory censoring were retained. There were about 65 unique remark codes in the data indicating laboratory censoring; these were unified under a single remark code of “<.” In some cases, there was no remark code provided, but a comment field associated with a data value indicated that the value was censored; these values also were given a remark code of “<.”

Some sources reported a censored remark code for an observation but did not report an associated value. When more than 1 percent of the data for a parameter at a site had a censored remark code but no values, then all of the data (censored and uncensored) for that site and parameter were excluded. When less than 1 percent of the data for a parameter at a site had a censored remark code with a missing value and censored values were reported for other observations, then the missing censored values were set to the reported censored value that was closest in time for that site and parameter.

In some cases when an agency reported data with censored remark codes, units were provided for uncensored data but not for censored data. The reasons for this practice are unknown, but it creates the potential for substantially biased analyses if only the censored data were excluded because the units were missing. If all of the missing units were for censored data for a parameter from a given source, and all uncensored data were reported in the same units throughout the record (for example, all uncensored data for a parameter were reported in units of milligrams per liter or micrograms per liter), then those same units were assumed to apply to all censored data for that parameter from that source. If necessary, all retained data were converted to milligrams per liter.

Recensoring U.S. Geological Survey Data

On October 1, 1998, the NWQL began implementing new reporting conventions for nondetections (Oblinger Childress and others, 1999). Prior to October 1, 1998, nondetections were censored to a MDL, the concentration at which the risk of a false positive detection (analyte reported as present at the MDL when not in the sample) is no more than 1 percent or to a MRL, an inconsistently defined level of detection capability. The NWQL deemed the use of inconsistent MRLs as inappropriate and also observed that the risk of a false negative occurrence at the MDL (analyte reported as not present when present at the MDL concentration) could be as much as 50 percent. To make the determination of detection limits more consistent and to reduce the risk of false negatives to no more than 1 percent, the NWQL began using a LRL in combination with a LT–MDL on October 1, 1998, based on Oblinger Childress and others (1999). The LRL typically was twice the LT–MDL. Between October 1, 1998, and October 1, 2010, concentrations measured between the LRL and the LT–MDL were reported as estimated concentrations with an “E” remark code and nondetections were censored to the higher LRL. Beginning on October 1, 2010, the NWQL censored inorganic nondetections to the LT–MDL (https://water.usgs.gov/admin/memo/QW/qw10.07.html).

The reporting convention in place between October 1, 1998, and October 1, 2010, creates potential bias in statistical analyses because it improperly represents the proportions of the data below the LT–MDL and the LRL (Helsel, 2005). In particular, observations that are truly below the LT–MDL will be reported as censored at the higher LRL; thus, the probability that an observation falls between the LT–MDL and the LRL may be overestimated, and the probability that it falls below the LT–MDL may be underestimated (Helsel, 2005). Depending on the true proportion of data below the LT–MDL, this can cause an upward bias that will adversely affect statistical analyses when these affected data are pooled with data subject to conventional censoring (that is, data not recensored from the LT–MDL to the LRL). To avoid this problem, there are two potential solutions: (1) The LT–MDL becomes the reporting limit when values reported as censored at the higher LRL are recensored to the lower LT–MDL. Uncensored “E” values between the LT–MDL and the LRL are not recensored. (2) The LRL becomes the reporting limit when all uncensored “E” values below the LRL are censored at the LRL. To retain the extra information captured by the “E” values, the first solution was used for USGS data from NWQL that were originally reported using the LRL approach. The “E” remark code may have been applied to other values for unrelated performance-based reasons, as described in Oblinger Childress and others (1999); these values were unchanged.

Similar to the description in the “Adjustments to U.S. Geological Survey Data Based on Historical Memorandums and Reports” section, these adjustments were applied to data from the NWIS database that had the analyzing entity defined as NWQL or when the analyzing entity was missing. For data analyzed at other laboratories, metadata documenting reporting conventions for censored data were not readily available; therefore, all censored data from laboratories other than the NWQL were unaltered.

Duplicate Samples

It was not uncommon to receive the same water-quality data for the same site and collection date from more than one source during the compilation of data. Duplicates were defined as samples from the same site, date, time, and fraction. Duplicate samples were removed from the dataset. It was not possible to determine which of the duplicate samples should be removed on the basis of data source because sometimes a source provided longer or more completely populated records for some parameters (for example, there may have been duplicate data obtained from STORET and provided directly by a State agency, but overall the STORET record was longer for alkalinity whereas the State record was longer for DOC). As a result, each parameter was considered separately when deciding which duplicate samples to retain and which to exclude. Note that duplicate samples did not always have identical values for a given parameter. If one of the duplicate samples was from USGS and one was from another agency, the USGS sample was retained. If both were USGS samples (for example, the duplicate samples were obtained from two different USGS NWIS servers), the sample from the primary host server (home State for each site) was retained. If information on the primary host server was not available, one of the duplicate samples was randomly retained. If both samples were from agencies other than USGS, and one sample had been provided by the CBP, then the sample from the CBP was retained. If both samples were from agencies other than USGS and the CBP, then the sample from the longest period of record was retained. Note that priority was not given to USGS samples because of an a priori expectation that USGS data were of better quality, but because typically more metadata were available on how samples were collected, analyzed, and stored.

Zero, Negative, and Missing Concentration Values

An analytical determination of zero or negative concentration is not possible, so zero or negative values in the data were considered suspect. Based on associated metadata—namely, the frequent presence of a censored remark code with a zero or negative concentration value—all zero and negative values were assumed to represent censored values. As a result, all zero and negative values were set to the reported censored value that was closest in time for that site and parameter combination. In addition, missing values that were reported with a censored remark code were assumed to be censored values without an associated laboratory detection level. These values were also set to the reported censored value that was closest in time for that site and parameter combination. If there were no reported censored values for a site and parameter combination, then zero, negative, or missing censored concentration values were set to the lowest uncensored value for that site and parameter combination and a censored remark code was assigned.

In some cases, these assumptions may have resulted in an incorrect value. For example, the zero, negative, or missing censored values may have been from samples that were not analyzed in the laboratory or for samples with a laboratory error. Alternatively, the assumed detection limit assigned to these values may be wrong. The approach used in this study to populate zero, negative, or missing censored values was unlikely to affect the subsequent trend results when such values collectively constituted a small fraction of the total data for a site and parameter combination; therefore, when zero, negative, and missing censored values made up more than 1 percent of the overall dataset for a site and parameter combination, all data for that site and parameter combination were excluded.

Calculation of Target Parameters

Alkalinity concentrations used for trend analysis were uniformly expressed as milligrams per liter as $CaCO_3$, although alkalinity was commonly reported by data sources as milligrams per liter as HCO_3^- or microequivalents per liter (µeq L^{-1}). Conversion of reported HCO_3^- to total alkalinity is discussed in the "Calculation of Target Parameters" section for carbon parameters of interest. Unit conversions to milligrams per liter as $CaCO_3$ were performed as follows:

$$\begin{aligned}&\text{milligrams per liter as CaCO}_3\\&=\left(\text{milligrams per liter as HCO}_3^-\right)\\&\times\left(\frac{1\text{ mol CaCO}_3}{2\text{ mol HCO}_3^-}\right)\times\left(\frac{100.07\text{ g CaCO}_3}{61\text{ g HCO}_3^-}\right)\end{aligned} \tag{13}$$

$$\begin{aligned}&\text{milligrams per liter as CaCO}_3=\left(\text{microequivalents alkalinity L}^{-1}\right)\\&\times\left(\frac{50.04\text{ g CaCO}_3}{1\text{ equivalent alkalinity}}\right)\times\left(\frac{1\text{ mg}}{1{,}000\text{ microequivalents}}\right)\end{aligned} \tag{14}$$

Identification of a Colocated or Nearby Streamgage

Following the steps described in the previous subsections of the "Carbon Data Processing" section, all site and parameter combinations with at least 3 years of consecutive data with at least quarterly sampling (see "Coverage of Changing Seasons" section in this report for details on how quarters were defined) were paired with a streamgage from the final site-streamgage list assembled in the "Streamgage Selection" section of this report. Any site not meeting these minimum data requirements and without an acceptable streamgage match was excluded from further consideration.

Merging Colocated or Neighboring Water-Quality Sites

After assigning a streamgage to each site, the retained sites represented two classes of sites: (1) sites that had a sufficient amount of data over a long enough period of record (at least for the shortest trend period of 10 years) to be used on their own for trend analysis (single sites), and (2) sites that had data for a shorter period of record (at least 3 consecutive years of quarterly data) that could be combined into a composite site with other colocated or neighboring composite sites to form a single record for trend analysis or a colocated or neighboring stand-alone site to increase the number of trend periods that could be evaluated at the stand-alone site. Single and composite sites were only combined when additional trend periods could be evaluated after combining; they were not combined to only fill in short gaps or add to the existing data at the single site when the combination did not increase the number of trend periods that could be evaluated. Data for periods shorter than 3 years were not used to fill in gaps or extend the record at other composite or single sites because it was decided that the nonrandom variability introduced into those types of combined time series would outweigh the benefits during trend analysis.

Following methods described in the "Streamgage Matching" section, potential colocated and neighboring sites were identified, and sites with intervening influences were excluded. A maximum of two sites were combined to minimize nonrandom variability. If more than two sites were available, then the two sites with the longest period of record for the parameter of interest were selected for the composite site. If more than one of the remaining sites merged well with the primary record, then the site with the longest record was retained. A list of composite sites with the individual sites for each parameter used in trend analysis can be found in appendix 1, table 1–7. The subsequent processing steps described below were applied to the data record for merged sites and single sites.

Degree of Censoring

Following guidelines suggested for the trend method used in this study, site and parameter combinations with more than 50 percent censoring were excluded (Hirsch and De Cicco, 2015). Left-censored and interval-censored values were included in this determination. Right-censored data were removed from the dataset.

Outliers

A small number of values in the data appeared to be typographical errors. For example, a dissolved organic carbon concentration of 500 mg/L is extremely unlikely to occur in natural waters (Hanley and others, 2013). It is far more likely that the units were reported incorrectly and that the true value is 500 µg/L, or 0.5 mg/L. In contrast, a dissolved organic carbon concentration of 50 mg/L is more ambiguous. That concentration is very high relative to concentrations found in most natural waters (Clark and others, 2007), but it could occur during locally extreme conditions such as a heavy runoff event or anthropogenic inputs from wastewater or sewer overflows. In a smaller study involving far fewer sites and data sources, it might be possible to definitively establish the accuracy of ambiguous data points by checking laboratory records, field forms, or other sources of historical information, but that was not possible in a study of this size because of resource constraints, so only the most extreme outliers—those that were almost certainly typographical errors—were excluded. Under a normal distribution, 99.9999998 percent of the data should be within six standard deviations from the mean; only the most extreme values will be outside of that range. For each site and parameter pair, the data were divided into four seasons. Within each season, any data value further than six standard deviations from the mean in log space was excluded as a likely typographical error. With this approach, there was a risk that typographical errors representing less extreme values were retained. There was also a risk that values representative of true environmental conditions were excluded.

Data Thinning

The carbon trend model used in this study weights days with multiple water-quality samples more heavily than other days during calibration. As a result, data for each site and parameter combination were thinned to have no more than one sample per day. If there was more than one sample per day, one sample was randomly selected to be retained in the dataset. The presence of data points spaced closely together in time also can introduce serial correlation into a time series. With some trend methods, serial correlation can increase the probability of incorrectly rejecting the test’s null hypothesis of no trend, leading to an incorrect determination that a significant trend has occurred (Hirsch and Slack, 1984). The trend method in this study relied on bootstrapping to determine p-values (see “Nutrient, Sediment, Major Ion, Salinity, and Carbon Trend Analysis Method” section in this report), an approach that is less affected by serial correlation. For that reason, the datasets were not thinned for data points close in time on different days.

Final Screening for Data Coverage

Once data from colocated or neighboring sites had been merged, and the censoring, outlier, and data thinning steps had been completed, the resulting data series were screened again by parameter. In this final screening step, the requirements for data coverage were more stringent. All sites not meeting these criteria were excluded.

Coverage Throughout Target Trend Periods

- The target trend periods, in water years, were (1) 1972–2012, (2) 1982–2012, (3) 1992–2012, and (4) 2002–12. Data were permitted to begin in the specified start year or the following year. For example, a start year of either 1972 (October 1, 1971) or 1973 (October 1, 1972) was acceptable in trend period 1, but 1974 (October 1, 1973) was not. Data were permitted to end in the specified start year or the preceding year. An end year of either 2011 (September 30, 2011) or 2012 (September 30, 2012) was acceptable in all trend periods, but 2010 was not.
- The first 2 years and the last 2 years of a trend period were required to have at least quarterly samples. See the “Coverage of Changing Seasons” section below for more detail on how quarters were defined.
- Overall, 70 percent of all years during a trend period were required to have at least quarterly samples. Any gap must have occurred in the middle of the record.
- Longer gaps in quarterly data were permitted when there was at least a 10-year span meeting the criteria beginning in any of the start years (for example, data from 1972 to 1981 or 1982 to 1991) and a 10-year span meeting the criteria ending in either 2011 or 2012. Data during each of those decades must have met the other detailed criteria. Gaps of any length were permitted in the years between the beginning and ending decades.

At least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during a given trend period in order to ensure the full range of flows were represented. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. For the shortest trend period (2002–12) of one decade, at least 14 percent of the samples were required to be high-flow samples. See the "Coverage of Changing Streamflow" section below for more detail on how high-flow samples were defined.

Coverage of Changing Seasons

Carbon concentrations vary throughout the year in response to changes in sunlight, temperature, biological activity, and carbon sources (Boyer and others, 1998; Dalzell and others 2007; Raymond and Oh, 2007). Samples representing seasonal conditions throughout the year are needed to fully characterize changes in concentration. To ensure seasonal coverage, 70 percent of years during a trend period were required to have at least one sample per quarter; the remaining 30 percent of years could be missing some or all quarterly samples. This allowed for reasonable changes in sampling design over time.

To evaluate seasonal coverage, the water year was divided into quarters using three different divisions of the water year (table 4). Each quarter consisted of 3 calendar months (table 4). A site was accepted if the samples met any one of the quarterly definitions in table 4; the quarterly definition was permitted to vary from year to year. This method of testing for quarterly samples accommodated variations in real-world sampling schedules. The primary sampling-design artifact this accommodated was the occurrence of two samples in one quarter and none in the preceding or following quarter, which may have occurred if a sample was collected slightly earlier or slightly later than dictated by a strict 3-month schedule. As an example, consider a site sampled in January, late March, July, and October during a particular year. The source collecting these data is clearly collecting quarterly samples; however, the sample that should have been collected in April was collected in late March. Without a flexible quarter definition, this year would not have had quarterly sampling. This flexible method of defining quarters also provides a buffer for sources that sample rivers that may freeze in the winter. For various reasons, including the weather, a winter quarterly sample may need to be collected anytime between November and February.

Coverage of Changing Streamflow

Carbon concentrations also vary in response to changes in streamflow (Dalzell and others, 2007). Samples representing a range of streamflow conditions are needed to fully characterize changes in concentration over time. Low- or moderate-flow samples are not in short supply. To identify high-flow samples, streamflow on the date of each sample was compared to the 85th percentile of flows in that month during the corresponding decade of the trend period at that site (depending on the trend period, that would include some or all of the following decades: 1972–81, 1982–91, 1992–2001, and 2002–12). For example, for a sample collected on March 1, 1998, the 85th percentile of all daily streamflows in March during the decade 1992–2001 was determined. If the streamflow on March 1, 1998, was greater than the 85th percentile of March streamflows during that decade, the sample was considered to be a high-flow sample. This assessment was completed by decade rather than once for the full trend period to ensure that samples at high streamflows were distributed throughout the trend period, rather than potentially being concentrated in one portion of it. It also allowed for the identification of high-flow samples collected during prolonged low-flow periods (for example, droughts). Such high-flow samples would have been collected during flows that were elevated relative to an unusually low baseline; the flows may not have been particularly high relative to overall average conditions at a site.

To ensure appropriate streamflow coverage, at least 14 percent of the samples were required to be high-flow samples in at least half of the decades assessed during the trend period. In the other half of the decades, at least 10 percent of the samples were required to be high-flow samples. The relaxation of the 14-percent criterion to 10 percent in some decades was necessary to accommodate sites with changing sampling designs over time (for example, the collection of high-flow samples increased somewhat or decreased somewhat over time). For the shortest trend period (2002–12) of only one decade, at least 14 percent of the samples were required to be high-flow samples during that period. These thresholds were based on sensitivity testing described in the "Storm Sampling" section of this report.

Data Used in Model Calibration

For site and parameter combinations that met the screening criteria, all available data before, during, and after the trend period (and if applicable, during any gaps) were retained to improve the model calibration (see the "Influence of New Data on Trend Estimates" section). Ultimately, model estimates only during the screened trend period were retained. In the case of longer records with data in two different decades separated by gaps, model estimates made in the years between the beginning and ending decades were excluded. Model input data are available in De Cicco and others (2017).

Multiple steps were used to determine the final trend model input values used in this study. Note that because of the recensoring of some USGS data, the summation of individual analytes for total parameters, the use of unrounded data from NWIS, and the use of interval censoring, the final USGS values used in this study will not always match the corresponding value stored in the NWIS database.

There were 584 sites evaluated for alkalinity and organic carbon trends with 795 combinations of sites and parameters. Table 5 summarizes the number of sites evaluated for parameter and trend period. A complete list of sites and parameters evaluated for major ion and salinity trends is reported in appendix 1, table 1–6.

Pesticide Data Preparation and Parameters of Interest

Pesticides are fundamentally different than the other water-quality parameters included in this trend study. To accommodate these differences, a different trend method, SEAWAVE–Q, was selected for pesticides. Because of these differences, the pesticide data preparation and screening criteria differ from methods and criteria described for other water-quality parameters for a variety of reasons, some particular to the trend analysis method and others particular to the nature of pesticide concentrations in surface water. The generalized process for preparing and screening pesticide data is presented in figure 2*B*.

Recent USGS pesticide trend studies for surface water (Sullivan and others, 2009; Ryberg and others, 2010, 2014) were based on published datasets (Martin, 2009; Martin and others, 2011) of 44 commonly used pesticides and 8 pesticide degradates (new compounds formed by transformation of a pesticide by chemical, photochemical, or biological reactions [Gilliom and others, 2006]). Pesticides in these datasets were analyzed at the NWQL using gas chromatography/mass spectrometry (GCMS; Zaugg and others, 1995; Lindley and others, 1996; Madsen and others, 2003) for water-quality samples after collection from January 1992 through September 2010 at NAWQA and NASQAN sites. Concentrations were adjusted to compensate for bias resulting from temporal changes in recovery of the GCMS analytical method as determined using results from matrix spikes (Martin and others, 2009; Martin and Eberle, 2011). Martin and others (2011) stated, "Spikes are water samples where a known amount of pesticide is added to the water sample. Recovery is the measured concentration of the pesticide divided by the expected concentration and is expressed as a percentage. Both bias in recovery and variability of recovery are characteristics of method performance." Two types of spikes were considered for the recovery analysis: (1) laboratory reagent spikes and (2) matrix spikes. Laboratory reagent spikes are quality control (QC) water samples in which a spike solution has been added to a sample of reagent-grade blank water and are primarily used to demonstrate that the sample preparation and analysis processes are in control (Martin and others, 2009). Matrix spikes are QC water samples in which a spike solution has been added to an environmental water sample (usually in the field). While field and matrix spikes provide similar information on pesticide recovery, the matrix spikes are subject to the effects of pesticide degradation that might occur in the time period between sample collection and analysis (Martin and others, 2009). "Models of recovery based on matrix spikes are deemed more appropriate for adjusting concentrations of pesticides measured in groundwater and stream-water samples than models based on laboratory reagent spikes because (1) matrix spikes are expected to more closely match the matrix of environmental water samples than are reagent spikes and (2) method performance is often matrix dependent, as was shown by higher recovery in matrix spikes for most of the pesticides" (Martin and others, 2009).

Exploratory analysis for this trend study considered the inclusion of pesticide concentration data from additional sources, such as from USGS sites not designated as NAWQA or NASQAN but having concentrations determined by the GCMS method; from any USGS site and using any analytical method; and from outside agencies using a variety of methods. A major consideration for the expansion of pesticide trend analysis to additional methods, sites, and analyzing agencies is the potential need to correct for temporal changes in the pesticide recovery rate for laboratory analyses. Analytical recovery of pesticides in laboratory quality-control matrix-spike samples and in stream-water matrices varies chemically and temporally (Bexfield, 2008; Martin and others, 2009; Martin and Eberle, 2011).

Variable pesticide recovery rates can introduce artificial time trends in pesticide concentrations, even when time trends do not actually occur in the environment. The pesticide recovery rate in laboratory spikes is routinely measured by NWQL; however, pesticide concentrations reported by NWQL are not corrected for variable recovery rates (Martin and Eberle, 2011), and most environmental laboratories do not correct for recovery (Keith, 1991; J.D. Martin, U.S. Geological Survey, oral commun., 2014). Martin and others (2009) studied pesticide recovery using laboratory reagent spikes (5,132 reagent spikes for the time period September 1993 through September 2005) and groundwater and stream-water matrix spikes (2,097 matrix spikes; 49,749 estimates of pesticide recovery for the period May 1992 through August 2006). In 2011, Martin and Eberle updated and supplemented that work with 4 years of additional data. They developed models of recovery, robust, locally weighted scatterplot smooths (lowess smooths) of matrix spikes, and found that the "magnitude of temporal change in modeled recovery in stream-water matrix spikes ranged from 10.6 percent for *p,p'*-DDE to 126.3 percent for carbaryl (table 2). The median amount of temporal change was 37.2 percent" (Martin and Eberle, 2011, p. 5). This indicates that reported concentrations are often biased high or low and that the bias changes over time.

Because of the importance of recovery adjustment for trend analysis and the lack of laboratory recovery data for non-NWQL laboratories, only NWQL data using the GCMS method were used in this trend analysis. This expanded the number of sites beyond the NASQAN and NAWQA sites used in previous pesticide trend reports. Parameters selected for inclusion in this trend analysis include the same 44 pesticides and 8 pesticide degradates used in past USGS pesticide trend studies (table 6). Table 6 groups the pesticides by use category (acaricide, herbicides, and insecticides); the parent pesticide is provided for degradates.

Table 6. Pesticides included in trend analysis.

[Sorted by use category, pesticide class, and U.S. Geological Survey (USGS) parameter code; CAS, Chemical Abstracts Service; --, not defined or not applicable; DCPA, dimethyl tetrachloroterephthalate; EPTC, S-Ethyl dipropylthioarbamate; HCH, hexachlorocyclohexane; DDE, dichlorodiphenyldichloroethylene; DDT, dichlorodiphenyltrichloroethane]

Pesticide compound	USGS parameter code	CAS number[a]	Pesticide class	Parent pesticide, if degradate
Acaricide				
Propargite	82685	2312-35-8	Sulfite ester	--
Herbicides				
Propachlor	04024	1918-16-7	Acetanilide	--
Metolachlor	39415	51218-45-2	Acetanilide	--
Alachlor	46342	15972-60-8	Acetanilide	--
Acetochlor	49260	34256-82-1	Acetanilide	--
Propyzamide	82676	23950-58-5	Amide	--
Propanil	82679	709-98-8	Amide	--
Napropamide	82684	15299-99-7	Amide	--
2,6-Diethylaniline	82660	579-66-8	Aniline	Alachlor
Dacthal (DCPA)	82682	1861-32-1	Chlorobenzoic acid ester	--
Trifluralin	82661	1582-09-8	Dinitroaniline	--
Ethalfluralin	82663	55283-68-6	Dinitroaniline	--
Benfluralin	82673	1861-40-1	Dinitroaniline	--
Pendimethalin	82683	40487-42-1	Dinitroaniline	--
Butylate	04028	2008-41-5	Thiocarbamate	--
EPTC	82668	759-94-4	Thiocarbamate	--
Pebulate	82669	1114-71-2	Thiocarbamate	--
Molinate	82671	2212-67-1	Thiocarbamate	--
Triallate	82678	2303-17-5	Thiocarbamate	--
Thiobencarb	82681	28249-77-6	Thiocarbamate	--
Simazine	04035	122-34-9	Triazine	--
Prometon	04037	1610-18-0	Triazine	--
Deethylatrazine	04040	6190-65-4	Triazine	Atrazine
Cyanazine	04041	21725-46-2	Triazine	--
Atrazine	39632	1912-24-9	Triazine	--
Metribuzin	82630	21087-64-9	Triazine	--
Terbacil	82665	5902-51-2	Uracil	--
Linuron	82666	330-55-2	Urea	--
Tebuthiuron	82670	34014-18-1	Urea	--
Insecticides				
Carbofuran	82674	1563-66-2	Carbamate	--
Carbaryl	82680	63-25-2	Carbamate	--
alpha-HCH	34253	319-84-6	Organochlorine	gamma-HCH
p,p´-DDE	34653	72-55-9	Organochlorine	DDT
gamma-HCH	39341	58-89-9	Organochlorine	--
Dieldrin	39381	60-57-1	Organochlorine	--
Fonofos	04095	944-22-9	Organothiophosphate	--
Chlorpyrifos	38933	2921-88-2	Organothiophosphate	--
Malathion	39532	121-75-5	Organothiophosphate	--
Parathion	39542	56-38-2	Organothiophosphate	--
Diazinon	39572	333-41-5	Organothiophosphate	--
Phorate	82664	298-02-2	Organothiophosphate	--
Parathion-methyl	82667	298-00-0	Organothiophosphate	--
Ethoprophos	82672	13194-48-4	Organothiophosphate	--
Terbufos	82675	13071-79-9	Organothiophosphate	--
Disulfoton	82677	298-04-4	Organothiophosphate	--
Azinphos-methyl	82686	86-50-0	Organothiophosphate	--
Fipronil	62166	120068-37-3	Phenyl pyrazole	--
Fipronil sulfide	62167	120067-83-6	Phenyl pyrazole	Fipronil
Fipronil sulfone	62168	120068-36-2	Phenyl pyrazole	Fipronil
Desulfinylfipronil amide	62169	--	Phenyl pyrazole	Fipronil
Desulfinylfipronil	62170	--	Phenyl pyrazole	Fipronil
cis-Permethrin	82687	54774-45-7	Pyrethroid	--

[a] This report contains CAS Registry Numbers®, which is a Registered Trademark of the American Chemical Society. CAS recommends the verification of the CASRNs through CAS Client Services℠.

Pesticide Data Processing

Recensoring National Water Quality Laboratory Data and Preparation of Data for Trend Analysis

The principal steps in data preparation for trend analysis were to (1) adjust concentrations to compensate for bias resulting from temporal changes in recovery of the GCMS analytical method using the loess method presented in Martin and Eberle (2011), and (2) recensor routine nondetections by site and pesticide. The maximum LT–MDL (maxLT–MDL) was determined for each pesticide using historical LT–MDLs from the NWQL.

The GCMS method provides low-level concentration estimates qualified with an “E” remark that are below the LT–MDL (Martin and others, 2011, p. 3–6). “Both pesticide concentration and sample matrix are expected to affect detection capability. Pesticide concentrations less than but close to maxLT–MDL likely are detected almost as readily as concentrations at or just above maxLT–MDL, whereas concentrations two orders of magnitude less than maxLT–MDL might rarely be detected. As was true for analytical recovery (Martin and others, 2009, p. 10–12, figs. 2 and 4), detection limits likely are a function of sample matrix. Concentrations readily detected in one matrix might not be detected in a more problematic matrix” (Ryberg and others, 2010). For these reasons, a less conservative estimate of the detection limit was used for site-pesticide combinations with five or more detections below the maxLT–MDL. The routine nondetections and low-level detections were recensored to the median of the low-level detections that fell below the maxLT–MDL (qlow50; appendix 1 of the Ryberg and others [2010] report). For site-pesticide combinations with fewer than five low-level detections, low-level detections were recensored to the maxLT–MDL. Censored values with raised reporting limits (censored at a value larger than the maxLT–MDL) were left in the dataset. Zero values were removed from the dataset. All pesticide concentrations are reported in micrograms per liter and are from filtered samples.

Pesticides collected for USGS sites 06934500 Missouri River at Hermann, Missouri, and 07022000 Mississippi River at Thebes, Illinois, between October 1, 2001, and May 1, 2008, specified NWQL schedule 2010, where the extractions were done in the field (rather than in the laboratory as other schedules) and likely resulted in contamination; however, the contamination of atrazine and metolachlor is estimated, with 99 percent confidence, not to exceed 0.0093 and 0.0053 µg/L, respectively, in no more than 2 percent of all samples (Laura Medalie, U.S. Geological Survey, written commun., September 2016). Both of these sites have been used in past trend studies (Sullivan and others, 2009; Ryberg and others, 2014).

Degree of Censoring

Pesticide concentrations in rivers and streams are often below the detection level and the highly censored datasets present a challenge for statistical analysis (Sullivan and others, 2009). The trend method SEAWAVE–Q (Vecchia and others, 2008; Sullivan and others, 2009; Ryberg and Vecchia, 2013) was selected for pesticide data analysis because it is robust to high levels of censoring and incorporates a parametric survival regression model using maximum likelihood methods for censored data (R Development Core Team, 2014; Therneau, 2014). As in previous studies (Ryberg and others, 2010; Ryberg and others, 2014), each site-pesticide combination was required to have at least 10 uncensored concentrations (detections at or above the censoring level) in each trend period (both the 10-year and 20-year trend periods).

Data Thinning

The presence of data points spaced closely together in time can introduce serial correlation into a time series. Serial correlation can increase the probability of incorrectly rejecting the null hypothesis of trend test, leading to an incorrect determination that a significant trend has occurred. To limit the influence of serial correlation in the trend tests, the data for a site and parameter combination were thinned to no more than one sample per week by retaining the sample collected closest to noon on Wednesdays.

Final Screening for Data Coverage

Once the data processing steps had been completed, the resulting data series were screened for temporal coverage. All sites not meeting these criteria were excluded for trend analysis.

Coverage Throughout Target Trend Periods

- The target trend periods were (1) 1992–2012 and (2) 2002–12. Start and end years were allowed to vary by 1 year. For example, a start year of 1992 or 1993 was acceptable in trend period 1, but 1994 was not. An end year of 2011 or 2012 was acceptable in both trend periods, but 2010 was not.

- The start and end years of the series had to have three or more quarterly samples per year.

- For the 20-year trend period, 50 percent of the trend period needed to have at least three quarterly samples per year (10 years of three or more quarterly samples), and there could be no gaps longer than 5 years.

- For the 10-year trend period, 50 percent of years needed to have at least three quarterly samples per year (5 years of three or more quarterly samples), and there could be no gaps longer than 3 years.

Coverage of Changing Seasons

To ensure seasonal coverage, 50 percent of years during a trend period were required to have at least three quarterly samples per year; the remaining 50 percent of years could be missing some quarterly samples. This allowed for reasonable changes in sampling design over time. The trend method SEAWAVE–Q is robust to gaps in the samples, provided the samples meet some criteria to be representative to the trend period. In the urban pesticide trends study (Ryberg and others, 2010, p. 7), at least 10 uncensored values were required and "the minimum sampling criteria for a particular site to be considered adequately representative of a particular 9-year trend assessment period were (1) at least 2 years with four or more samples collected during the first 5 years of the assessment period and (2) at least 2 years with four or more samples collected during the last 5 years of the assessment period." Examples are shown in figures 2–5 of the Ryberg and others (2010) report for sites with varying numbers of samples per year and with high degrees of censoring. In the pesticide trends study for major river basins of the United States, Ryberg and others (2014, p. 8) state "the minimum sampling criteria for a particular site to be considered adequately representative of a particular 10-year trend period were to have (1) at least 10 uncensored values after recensoring (calculating qlow50 when applicable and recensoring at that level), (2) at least 5 years of samples, (3) 6 or more samples in at least 2 of the first 5 years of the period, and (4) 6 or more samples in at least 2 of the last 5 years of the period." For additional information on minimum sampling criteria for determining pesticide trends, see Vecchia and others (2008), Sullivan and others (2009), and Ryberg and others (2010, 2014).

Coverage of Changing Streamflow

Pesticide concentrations in surface water have a strong seasonal signal depending on the time of application. There may be one or more general periods of application depending on the pesticide, some are postemergent applied in the spring, others may be preemergent applied in the spring and (or) fall, and other patterns of application may exist depending on the climate and crop. For example, in Vecchia and others (2008), figure 1 shows that seasonal variability in atrazine is not closely related to seasonal variability in streamflow. "Thus, the seasonal hydrologic cycle has little effect on atrazine concentration. Rather, seasonal change in application amounts is the major source of variability. The same generally holds for other pesticides and sites as well. This is in sharp contrast to other more naturally occurring or chemically stable water-quality parameters, such as dissolved major ions or nutrients, whose concentrations tend to be more directly related to streamflow" (Vecchia and others, 2008). The trend method SEAWAVE–Q specifically models the seasonality in application and a decay rate (depending on the properties of the chemical, concentrations decrease at varying rates after the peak application period). Because of these important characteristics of pesticide concentrations in surface water and the features of SEAWAVE–Q, pesticide sites do not require the same screening for high-flow samples that are needed for the parameters modeled using WRTDS.

Identification of a Colocated or Nearby Streamgage

Continuous daily streamflow data were a required input to the trend model used in this study. Therefore, the presence of daily streamflow data during the full trend period was a requirement. Following methods described in an earlier section, a colocated or nearby streamgage was identified for each water-quality site to provide daily streamflow data at each site. Any site for which a suitable streamgage could not be found was excluded.

Merging Colocated or Neighboring Sites

Only USGS pesticide samples were used for this trend analysis. For this reason, there was no need to merge colocated or neighboring sites.

Zero and Missing Streamflows

On days when streamflow was reported as zero (that is, no streamflow was present in the stream), streamflow was set to the minimum recorded value for that particular stream (0.01 ft^3/s in all cases). This was done because the trend model uses logarithmic transformations that are undefined for zero.

The trend method SEAWAVE–Q does not require streamflow on the day of sampling only but also requires long-term streamflow monitoring (at least a year prior to the start of the trend period for the national trends study) in order to calculate and incorporate streamflow anomalies. Small gaps in the streamflow record can be estimated, but gages discontinued for 30 days or more during a trend analysis period cause issues, such as a gap in the model or disqualification of the site, depending on the length and timing of the gap. The trend method SEAWAVE–Q requires streamflow over the entire trend period, plus one additional year at the beginning of the period because 1-year, 30-day, and 1-day streamflow anomalies (Vecchia and others, 2008; Ryberg and Vecchia, 2012) were used in the model. Sites with sufficient pesticide concentration data were eliminated because of lack of streamflow at the beginning or end of the period of record or missing blocks of streamflow throughout the period. In one case, a site had a small number of missing streamflow values (10 missing values for site 12510500 Yakima River at Kiona, Washington, in the 2001–12 streamflow period) and the fillMiss function of the R package waterData (Ryberg and Vecchia, 2012) was used to fill in the missing days using a structural time series model. In some cases, sites had incomplete streamflow data for 2001 and 2012. These sites were retained in the analysis because their period of record varied little from the sites with a full period of record from 2001 to 2012.

Data Used in Model Calibration

There were 70 sites evaluated for pesticide trends, with 703 combinations of sites and parameters. Table 5 summarizes the number of sites evaluated for each time period for each parameter. A complete list of sites and parameters evaluated for pesticide trends is reported in appendix 1, table 1–6. Model input data are available in Ryberg and others (2017).

Ecology Data Preparation

Ecology data are fundamentally different than the other water-quality and pesticide parameters included in this trend study. As a result, a different set of screening criteria and analysis techniques were used to evaluate trends in the ecology data. The generalized analysis collectively consisted of the Pearson correlation test (Harrell and others, 2016), linear regression (R Core Team, 2014), and the Kendall-tau test (Lorenz, 2016). The generalized process for preparing and screening ecology data is presented in figure 2C.

Data and Trend Screening for Site and Sample Selection

Publically available fish, invertebrate, and algae data were screened to include only sites having greater than or equal to six annual samples of any assemblage collected between 1993 and 2012. This differed slightly from the water-quality trend period from 1992 to 2012 because NAWQA Project ecology sampling did not begin until 1993. For comparing ecology trends to the other water-quality trends, the shortened time period allowed by the screening criteria of 1993–2011 was considered equivalent to a 1992–2012 trend. All data were collected and processed by USGS personnel as part of the NAWQA Project (Gilliom and others, 1995) following established protocols (Cuffney and others, 1993; Meador and others, 1993; Porter and others, 1993; Fitzpatrick and others, 1998; Moulton and others, 2000; Charles and others, 2002; Moulton and others, 2002). Invertebrate and algae samples represented communities collected from richest-targeted habitats (Cuffney and others, 1993; Moulton and others, 2002). Fish samples consisted of combined electrofishing and seining efforts (Meador and others, 1993; Moulton and others, 2002). Sites were then screened to include only wadeable streams.

These preliminary criteria resulted in a starting dataset of 3,441 samples collected from 104 sites. Sites and samples were independently screened for each assemblage (fish, invertebrates, and algae). Only sites and samples meeting the following criteria were retained.

- The target trend periods were 1993–2012 (hereafter referred to as "20-year trend period") and 2002–12 (hereafter referred to as "10-year trend period"). Only sites with samples collected within 1 year after the start year and (or) within 1 year before the end year were retained. For example, a start year of 1993 or 1994 was acceptable in the 20-year trend period, but 1995 was not. An end year of 2011 or 2012 was acceptable in either trend period, but 2010 was not.
- Only sites with less than or equal to a 4-year continuous data gap were allowed.
- Only sites with samples in at least 50 percent of years in each half of either trend period were retained.
- Sites having available streamflow data 240 days before each ecology sample date or those where a nearby streamgage could be substituted or used to estimate values were retained (appendix 1, tables 1–3 through 1–5).
- In some instances, more than one sample was collected at a site during the same year. In such cases, samples collected at the primary reach (Fitzpatrick and others, 1998) closest in time to the mean sampling day and month within the trend period were selected to represent that year.

The above criteria resulted in various site/sample combinations for each assemblage and trend period (appendix 1, table 1–6). For example, there were 47 fish, 49 invertebrate, and 35 diatom sites that passed the criteria to be included in the 2002–12 trend period; whereas only 4 fish, 9 invertebrate, and 3 diatom sites met the 1993–2012 trend period criteria (table 5). Note that soft algae were removed from the dataset, which preserved the remaining diatoms for analysis. The total number of samples over the 2002–12 trend period was 1,096 (355 fish, 443 invertebrates, and 298 diatoms) and 208 for 1993–2012 (122 fish, 47 invertebrate, and 39 diatoms). The number of samples per site ranged between 6 and 11 for 2002–12 and between 11 and 15 for 1993–2012.

Ecology Data Processing

Taxonomic Harmonization and Resolving Ambiguous Taxa

Fish

The taxonomic target level for fish samples was species. In a few instances, species-level identifications could not be made in the field or laboratory. In these cases, identifications documented at either family (hybrids resulting from different genera) or genus (hybrids and [or] individuals that could not be identified to species with certainty) were removed from the final dataset, with the exception of genus level freshwater sculpin identifications (family Cottidae; genus Cottus). All freshwater sculpin identifications were reduced to the genus Cottus sp. because species level identification within this group is often problematic (Young and others, 2013).

Invertebrates

Invertebrates were identified to the lowest practical taxonomic resolution in the laboratory, which was mostly genus or species (mollusks, crustaceans, and insects), but also included family, phylum, or class (aquatic worms, flatworms, and nematodes), depending on the group (Moulton and others, 2000). Several steps were taken to prepare the invertebrate data for analysis. First, names were examined for cases where changes in taxonomic classification occurred during the 20-year trend period. In such cases, names were harmonized to help alleviate any bias from changes in taxonomy. This is necessary because taxonomic revision either adds (through splitting of groups) or subtracts (through lumping of groups) the number of names associated with a particular group. For example, Wang and McCafferty (2004) used genetic information to revise the mayfly family Heptageniidae. As part of their revision, all but one species originally belonging to the genus *Stenonema* (*S. femoratum*) was placed into the genus *Maccaffertium*; therefore, all samples collected after 2004 containing *Stenonema* and *Maccaffertium* were not comparable to those collected prior to 2004 when *Maccaffertium* was unrecognized. In this example, all *Stenonema* and *Maccaffertium* were given the name *Stenonema* because any *Stenonema* collected prior to 2004 other than identifiable *S. femoratum* could not be split into *Maccaffertium* without reexamining the specimens, which were no longer available. Similar issues occurred with other Heptageniidae genera (for example, *Nixe*, *Leucrocuta*, and *Ecdyonurus*) as well as with the mayfly families Baetidae (for example, *Procloeon* and *Cloeon*) and Leptohyphidae (for example, *Tricorythodes* and *Leptohyphes*). Second, species names were reduced to genus when possible, and their counts were summed by sample. Counts of larvae, pupae, and adults with the same name were also combined. Third, ambiguous taxa were resolved by keeping only those taxa (and counts) that were unique within each sample. This approach allows ambiguous taxa across the dataset, but preserves taxa richness within each sample. Preservation of taxa richness was desirable because most endpoints were based on richness (see “Ecological Endpoints” section for a description of endpoints). Exceptions were made to this approach for measures of ecological condition when it was necessary to harmonize taxonomy with the operational taxonomic units used in the development of models for the calculations. Fourth, raw data counts were standardized to 300 for each sample by randomly resampling individuals without replacement until a count of 300 was reached. Raw count adjustments were necessary because laboratory subsampling procedures resulted in a wide range of individuals extracted from each sample, which differentially affected invertebrate taxa richness across samples (Vinson and Hawkins, 1996).

Algae

Several steps were taken to prepare the algae data for analysis. First, all soft algae were removed from the dataset, which preserved the remaining diatoms for analysis. Second, the remaining diatom names were harmonized to ensure taxonomic consistency within each trend period and among sites. Finally, diatom cell counts were summed for each taxon within each sample. The target level for diatom identification was species, but some taxa were identified to lower taxonomic levels such as subspecies or variety. All taxa identified at the genus level were dropped before metrics were calculated.

Ecological Endpoints

Selected endpoints representing ecological integrity, community composition, tolerance, and similarity were calculated separately for each site to evaluate changes over time in fish, invertebrate, and diatom communities (table 7). In some cases, it was possible to calculate identical endpoints among assemblages (for example, taxa richness, Shannon-Weaver diversity, and Bray Curtis similarity). In other cases, calculations were specific to each assemblage by necessity. For example, average community tolerances of each assemblage to specific stressors were calculated; however, the specific stressor may have been represented differently for diatoms when compared to fish and (or) invertebrates. Nationally derived indicators of ecological condition were not available for diatoms at the time of this publication so only fish and invertebrates are included. Examples of indicators commonly used to assess ecological condition include the ratio of observed (O) to expected (E) taxa (O/E) and multimetric type indices (MMI). Details of each endpoint by assemblage are described below.

Fish Endpoints

Fish richness, diversity, Bray-Curtis similarity, and ecological condition, as well as tolerance to ammonia, chloride, specific conductance, dissolved oxygen, nitrate plus nitrite, pH, phosphorus, sulfate, suspended sediment, and water temperature were calculated for each sample (table 7):

- FishRich is the total number of different fish species.
- FishDiv is a measure of diversity that takes into account the number of different taxa and how evenly their counts are distributed within a sample (Shannon and Weaver, 1963).
- FishLTsim is the Bray-Curtis Similarity (Bray and Curtis, 1957) value between each year and the start year in the trend period (year 1 with 2, year 1 with 3, year 1 with 4, year 1 with 5, and so on). This measure evaluates changes in community similarity relative to the start year. Individual species counts were fourth root transformed before Bray-Curtis Similarity values were calculated. Fourth root transformation downweights the contribution of highly abundant species while boosting the influence of those with midrange and lower abundances (Clarke and Warwick, 2001).

Table 7. Ecology metric descriptions for fish, invertebrate, and diatom communities calculated for trend assessment.

Code	Metrics
	Fish
	Community composition
FishRich	The total number of different fish species
FishDiv	Fish Shannon and Weaver diversity
FishLTsim	Fish Bray-Curtis Similarity value between each year and the start year in the trend period
	Condition indicator
FishOE	Ratio of observed (O) fish taxa to the modeled number of expected (E) fish taxa (O/E) or multimetric index (MMI) value standardized to O/E units
	Tolerance
FishAmm	Average fish community tolerance to ammonia
FishCl	Average fish community tolerance to chloride
FishSC	Average fish community tolerance to specific conductance
FishDO	Average fish community tolerance to dissolved oxygen
FishNitraNitri	Average fish community tolerance to nitrite plus nitrate
FishpH	Average fish community tolerance to pH
FishPhos	Average fish community tolerance to phosphorus
FishSul	Average fish community tolerance to sulfate
FishSusSed	Average fish community tolerance to suspended sediment
FishTemp	Average fish community tolerance to water temperature
	Invertebrates
	Community composition
InvertRich	The total number of different invertebrate taxa
InvertPerEPTr	Percent invertebrate taxa richness comprised of individuals belonging to the orders Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies)
InvertPnInr	Percent of invertebrate taxa richness comprised of noninsect taxa.
InvertDiv	Invertebrate Shannon Weaver diversity
InvertLTsim	Invertebrate Bray-Curtis Similarity value between each year and the start year in the trend period
	Condition indicator
InvertOE	Ratio of observed (O) invertebrate taxa to the modeled number of expected (E) invertebrate taxa (O/E)
	Tolerance
InvertIon	Average invertebrate community tolerance to major ions
InvertNutr	Average invertebrate community tolerance to nutrients
InvertDO.Temp	Average invertebrate community tolerance to dissolved oxygen and temperature
InvertSusSed	Average invertebrate community tolerance to suspended sediment
InvertFines	Average invertebrate community tolerance to fine sediment
	Diatoms
	Community composition
DiaRich	The total number of different diatom taxa
DiaDiv	Diatom Shannon Weaver diversity
DiatomLTsim	Diatom Bray-Curtis Similarity value between each year and the start year in the trend period
	Tolerance
DiaSenBahlTol	The number of diatom taxa sensitive to general nutrient and organic enrichment
DiaxHighCa	The number of diatom taxa with a calcium optima of greater than 40 milligrams per liter
DiaLowCl	The number of diatom taxa with a chloride optima less than 15 milligrams per liter
DiaLowO2	The number of diatom taxa that are known to tolerate down to 30 percent dissolved oxygen saturation
DiaHighO2	The number of diatom taxa that require nearly 100 percent dissolved oxygen saturation
DiaLowTN	The number of diatom taxa having a total nitrogen optima of less than or equal to 0.2 milligrams per liter as nitrogen
DiaHighTN	The number of diatom taxa having a total nitrogen optima of greater than or equal to 3 milligrams per liter as nitrogen
DiaLowTP	The number of diatom taxa having a total phosphorus optima of less than or equal to 10 micrograms per liter as phosphorus
DiaHighTP	The number of diatom taxa having a total phosphorus optima of more than or equal to 100 micrograms per liter as phosphorus
DiaHighMot	The number of diatom taxa that are considered highly motile (mobile) distinguished by the presence of an extensive raphe and (or) are large in size

- FishOE is a measure of ecological condition represented by the ratio of observed (O) taxa to the modeled number of expected (E) taxa (O/E) or as a multimetric index (MMI) value standardized to O/E units. In brief, O/E models compare the taxa observed in a sample to the taxa expected to be at a site in the absence of human disturbance. Predictions of E are based on multivariate relations between taxa and the environment derived from data collected at regional reference sites. Discriminant function analysis is used to select a subset of environmental predictor variables unrelated to human disturbance that best separates groups of reference sites previously determined by the dissimilarity in taxa occurrence among sites. Each taxon is assigned a probability of capture (PC) by calculating the frequency it occurs within each reference site group. The E is then calculated for a new site as the sum of PCs weighted by the probability that the site belongs to the reference group. The O is then calculated as the sum of each taxon predicted to be at the site that actually occurred at the site. See Moss and others (1987), Hawkins and Carlisle (2001), and Clarke and others (2003) for more details. In general, MMI models are built by selecting metrics that best discriminate between predefined reference and disturbed sites within a given basin, region, ecoregion, or other spatial stratification. Typically, the final set of metrics represents various aspects of community composition, richness, tolerance, abundance, function, as well as life history or other trait type characteristics. Once a final set of metrics are selected, values are scored on a continuous scale, summed, then rescaled to range between 0 and 100. See Karr and others (1986), Barbour and others (1999), and Hill and others (2000) for details and specific examples. Herein, both O/E and standardized MMI values were used out of necessity because previously developed models to calculate either value for all fish samples were unavailable at the time of this publication. The specific models used for fish included an O/E model previously developed for the Eastern United States (east of the 100th meridian; Meador and Carlisle, 2009), and an MMI previously developed for the Western United States (west of the 100th meridian; Whittier and others, 2007; Meador and others, 2008). All MMI values were standardized to common nondimensional O/E units (Hawkins, 2006) by dividing each value of each site by the mean of values for regional reference sites used to develop the MMI (Whittier and others, 2007). Rescaling is necessary to improve the comparability of MMI and O/E values (Hawkins, 2006). Hereafter, all standardized fish MMI values are identified as FishO/E and expressed in O/E units. The construction of O/E and MMI type models are detailed elsewhere (Karr and others, 1986; Moss and others, 1987; Barbour and others, 1999; Hawkins and Carlisle, 2001; Clarke and others, 2003). Models to calculate FishOE values were unavailable at the time of this publication for two sites (01170095 Green River at Stewartville, Massachusetts, and 01621050 Cahaba Valley Creek at Cross Creek Road at Pelham, Alabama; appendix 1, table 1–6).

- FishAmm, FishCl, FishSC, FishDO, FishNitraNitri, FishpH, FishPhos, FishSul, FishSusSed, and FishTemp are richness-based measures of community tolerance to ammonia, chloride, specific conductance, dissolved oxygen, nitrate plus nitrite, pH, phosphorus, sulfate, suspended sediment, and water temperature, respectively, based on the tolerance indicator values and methods detailed by Meador and Carlisle (2007).

Invertebrate Endpoints

Invertebrate richness; percentage of mayfly, stonefly, and caddisfly richness (EPT); percentage of noninsect richness; diversity; Bray-Curtis similarity; and ecological condition; as well as tolerance to major ions, nutrients, dissolved oxygen and temperature, suspended sediment, and fine sediment were calculated for each invertebrate sample (table 7):

- InvertRich is the sum of the number of different invertebrate taxa.

- InvertPerEPTr is the percentage of taxa richness comprised of individuals belonging to the orders Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies).

- InvertPnInr is the percentage of taxa richness comprised of noninsect taxa.

- InvertDiv is a measure of diversity that takes into account the number of different taxa and how evenly their counts are distributed within a sample (Shannon and Weaver, 1963).

- InvertLTsim is the Bray-Curtis Similarity (Bray and Curtis, 1957) value between each year and the start year in the trend period (year 1 with 2, year 1 with 3, year 1 with 4, year 1 with 5, and so on). This measure evaluates changes relative to the start year. Invertebrate taxa counts were fourth root transformed before Bray-Curtis Similarity values were calculated. Fourth root transformation downweights the contribution of highly abundant taxa while boosting the influence of those with midrange and lower abundances (Clarke and Warwick, 2001).

- InvertOE is a measure of ecological condition represented by the ratio of observed (O) taxa to the modeled number of expected (E) taxa (O/E). InvertOE values herein were calculated using previously developed models detailed in Carlisle and Meador (2007) and Carlisle and Hawkins (2008) and Yuan and others (2008). Models to calculate InvertOE values were unavailable at the time of this publication for two sites (01209710 Norwalk River at Winnipauk, Connecticut, and 06713500 Cherry Creek near Denver, Colorado; table 1–6). See the previous section under FishOE for a brief description of how O/E values are derived.

- InvertIon, InvertNutr, InvertDO.Temp, InvertSusSed, and InvertFines are richness-based measures of community tolerance to major ions, nutrients, dissolved oxygen and temperature, suspended sediment, and fine substrate based on the tolerance indicator values and methods described in Carlisle and others (2007).

Diatom Endpoints

Diatom richness; Shannon-Weaver diversity; and Bray-Curtis similarity; as well as the number of taxa sensitive to general nutrient and organic enrichment; the number that prefer high calcium, low chloride, low oxygen, high oxygen, low total nitrogen, high total nitrogen, low total phosphorus, and high total phosphorus; and the number that are highly motile were calculated for each diatom sample (table 7). Measures of ecological condition such as MMI or O/E values were not calculated for diatom communities because previously developed comparable models were unavailable at the time of this publication.

- DiaRich is the total number of different diatom taxa.
- DiaDiv is a measure of diversity that takes into account the number of different taxa and how evenly their counts are distributed within a sample (Shannon and Weaver, 1963).
- DiatomLTsim is the Bray-Curtis Similarity (Bray and Curtis, 1957) value between each year and the start year in the trend period (year 1 with 2, year 1 with 3, year 1 with 4, year 1 with 5, and so on). This measure evaluates changes relative to the start year. Diatom taxa counts were fourth root transformed before Bray-Curtis Similarity values were calculated. Fourth root transformation downweights the contribution of highly abundant taxa while boosting the influence of those with midrange and lower abundances (Clarke and Warwick, 2001).
- DiaSenBahlTol is the total number of diatom taxa sensitive to general nutrient and organic enrichment (Lowe, 1974; Lange-Bertalot, 1979; Bahls, 1993).
- DiaxHighCl is the total number of diatom taxa with a calcium optima of greater than 40 mg/L (Potapova and Charles, 2003; Porter, 2008).
- DiaLowCl is the total number of diatom taxa with a chloride optima less than 15 mg/L (Potapova and Charles, 2003; Porter, 2008).
- DiaLowO2 is number of diatom taxa that are known to tolerate as low as 30 percent dissolved-oxygen saturation (van Dam and others, 1994; Porter, 2008). Values for this metric were not calculated at two sites (01170095 Green River at Stewartville, Mass., and 08195000 Frio River at Concan, Texas) because associated taxa were not present at these sites during several years.
- DiaHighO2 is the number of diatom taxa that require nearly 100 percent dissolved-oxygen saturation (van Dam and others, 1994; Porter, 2008).
- DiaLowTN is the number of species having a total nitrogen optima less than 0.2 mg/L (National values in Potapova and Charles, 2007). Values for this metric were not calculated for one site (05420680 Wapsipinicon River near Tripoli, Iowa) because associated taxa were not present during several years (table 1–6).
- DiaHighTN is the number of taxa having a total nitrogen optima greater than 3 mg/L (National values in Potapova and Charles, 2007; Porter, 2008).
- DiaLowTP is the number of taxa having a total phosphorus optima less than or equal to 10 µg/L (National values in Potapova and Charles, 2007) (table 1–6). Values for this metric were not calculated for one site (05420680 Wapsipinicon River near Tripoli, Iowa) because associated taxa were not present during several years.
- DiaHighTP is the number of taxa having a total phosphorus optima greater than or equal to 100 µg/L (National values in Potapova and Charles, 2007) (table 1–6). Values for this metric were not calculated at two sites (01170095 Green River at Stewartville, Mass., and 08195000 Frio River at Concan, Tex.) because associated taxa were not present at these sites during several years.
- DiaHighMot is the number of taxa that are considered highly motile (mobile), as described by (Lowe, 1974) and (Spaulding and others, 2010). Values for this metric were not calculated for one site (01170095 Green River at Stewartville, Mass.) because associated taxa were not present during several years (table 1–6).

Associated Environmental Variables

Selected hydrology metrics were calculated to help describe conditions antecedent to each sample date, as was the day of year that each sample was collected. These variables were used to help account for year-to-year variability in endpoint values because of differences in year-to-year conditions prior to sampling (table 8). Trend results for

Table 8. Environmental variables used to account for variability in ecological endpoints as part of trend assessment.

Variable short name	Description
Avg15	Normalized average daily flow value of the 15 days prior to sampling
Avg60	Normalized average daily flow value of the 60 days prior to sampling
Avg240	Normalized average daily flow value of the 240 days prior to sampling
CV15	Normalized daily flow coefficient of variation of the 15 days prior to sampling
CV60	Normalized daily flow coefficient of variation of the 60 days prior to sampling
CV240	Normalized daily flow coefficient of variation of the 240 days prior to sampling
DOY	The day of the year a sample was collected where day 1 is January first normalized by the site average DOY

ecology are presented for unadjusted endpoints and for endpoints adjusted by antecedent hydrology and the day of the year samples were collected. For each site, the day of the year (DOY) was determined by normalizing the DOY each sample was collected by the average DOY samples were collected over the trend period; DOY values less than 1 indicate a sample was collected earlier rather than average in the index period, likewise values greater than 1 indicate sample collection was later than average.

Antecedent Hydrology

Average daily streamflow and the coefficient of variation of daily streamflow were calculated for 15, 60, and 240 days prior (antecedent) to each sample date. These selected antecedent periods were intended to capture site-specific conditions that may influence year-to-year differences in ecological endpoints. All values were normalized by the 10-year average for the same antecedent period for each metric and years from 2002 through 2012 (or the 20-year average for trends calculated for 1993–2012). For example, if the sample date was July 15, 2003, then the 15-day average daily streamflow (Avg15) for 15 days prior to the sample was calculated for July 1 through July 15, 2003. This value was than normalized by the average of daily streamflow values from the July 1 through July 15 of each year from 2002 to 2012. Some sites had an incomplete daily streamflow record within an antecedent or normalization period. In these cases, missing streamflow values were estimated using the Maintenance of Variance Extension type 1 (MOVE.1) method (Hirsch, 1982) and software developed by Granato (2009) (Streamflow Record Extension Facilitator; SREF version 1.0). The MOVE.1 method, also known as the line of organic correlation (Helsel and Hirsh, 2002), is one of several techniques used for augmenting streamflow records when a common streamflow period is needed among many sites (Ries, 1994). The technique assumes a linear relation exists between flows at the site with missing values (short-term site) and a nearby index site (or set of index sites) that are similar in streamflow characteristics to the short-term site. The MOVE.1 equation estimates streamflow values at the short-term site:

$$\bar{Y}_i = \bar{Y} + S_y \div S_x \left(X_i - \bar{X} \right) \tag{15}$$

where

$\bar{Y}$ and $\bar{X}$ are the mean of the concurrent streamflow, in cubic feet per second,

S_y and S_x are the standard deviation of the log of the concurrent streamflow,

X_i is the log of daily streamflow values at the index site, and

$\bar{Y}_i$ is the estimated daily streamflow value at the short-term site exponentiated back into original units.

- The streamflow record was augmented using the MOVE.1 method at a total of 10 sites (table 1–4). Additionally, streamflow from a nearby upstream or downstream streamgage was utilized at four sites by necessity because the ecology site was ungaged (table 1–3).

Day of the Year

It is often assumed that year-to-year variability in endpoint values is reduced by consistently sampling within the same index period (range of dates) each year (Cuffney and others, 1993; Moulton and others, 2002). Although this is a logical way to sample for trend studies, year-to-year conditions within the same index period vary widely. In a variance components study of NAWQA invertebrate data, Gurtz and others (2013) found timing of data collection within the index can explain more than 60 percent of variance in some endpoint values; therefore, the DOY (table 8) when each sample was collected was identified to help account for any influence of sample timing on endpoint variability in addition to the influence of antecedent streamflow conditions. The DOY was determined by normalizing the DOY each sample was collected by the average DOY for each site over the trend period; DOY endpoint values less than 1 indicate that a sample was collected earlier in the year than usual.

Trend Analysis Methods

Three models were used to determine trends for this study. The WRTDS model was used to evaluate trends in major ion, salinity, nutrient, carbon, and sediment parameters. The trend method SEAWAVE–Q was used to determine trends in pesticide parameters. A combination of the Kendall-tau test, Pearson correlation test, and linear regression were used to evaluate trends in ecology. Descriptions for each trend analysis method are provided in this section.

Nutrient, Sediment, Major Ion, Salinity, and Carbon Trend Analysis Method

The WRTDS model (Hirsch and others, 2010; Hirsch and De Cicco, 2015) was used to determine trends in concentration and load for nutrients, sediment, major ions, salinity, and carbon. The trends presented here are flow normalized (explained in the "Trends in Streamflow" section). The WRTDS package is part of two R packages: (1) Exploration and Graphics for RivEr Trends (EGRET) package, implementing methods to analyze long-term changes in water quality and streamflow (Hirsch and De Cicco, 2015), and (2) EGRETci package, implementing a set of approaches to the analysis of uncertainty associated with WRTDS trend analysis as implemented in the EGRET package (Hirsch and others, 2015).

The WRTDS model makes estimates of concentration for every day of the period of record at each site as

$$\ln(c) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \quad (16)$$

where

ln is natural log,

c is concentration, in milligrams per liter,

βi are fitted coefficients,

Q is daily mean streamflow, in cubic meters per second,

t is time, in decimal years, and

ε is the unexplained variation.

Estimates of daily concentration are multiplied by the respective daily mean streamflow to estimate daily load. The WRTDS model differs from similar regression-based load and trend models in that the fitted coefficients are not the same throughout the entire domain of the data. Weighted regression is used to estimate a unique set of coefficients for every combination of Q and t in the period of record. The weights on each observation in the calibration dataset are based on the distance in time, streamflow, and season between the observation and the particular combination Q and t being used for the estimate. This process results in unbiased estimates of daily concentration and load. Because WRTDS estimates a unique set of coefficients for every combination of Q and t in the period of record, the relations among concentration, streamflow, and time are not fixed. It is possible for both the magnitude and the sign of the coefficients in equation 16 to change, which allows for inflection points in the pattern of concentration and load to accommodate changes with streamflow and (or) time, which allows for the possibility of nonmonotonic trends over time.

Any given day in the period of record can be described by a unique combination of Q and t. For each day in the period of record, equation 16 is fit using locally weighted regression, resulting in a separate regression equation for each day. To save on computational effort, rather than doing a new estimate of the weighted regression equation for each day, a matrix of regression results is created and the estimate for any given day is determined using linear interpolation of the results stored in this matrix (Hirsch and others, 2010). For each day and model calibration, the weights on each observation in the dataset are based on a combination of three separate weights describing the similarity (or distance) of each data point to the given day relative to streamflow (log[Q]), time, and season of the year. For example, if equation 16 was being estimated for July 15, 2000, from a dataset extending from water years 1992 through 2012, data points in 2000 would be given a higher weight compared to data points in 2005. Similarly, because July 15 is in the summer, data collected in June, July, and August would be given higher weights compared to data collected in December. Finally, if daily mean streamflow on July 15, 2000, was 500 ft^3/s, then data points with streamflow around 500 ft^3/s would be given higher weights than streamflow around 5,000 ft^3/s. The user sets the windows used to determine these weights, and the three separate weights are multiplied to determine a single weight for each observation in the calibration dataset. The weights change for each model calibration.

Because the estimates of daily concentration and load are strongly influenced by random variations in streamflow, flow-normalized estimates of daily concentration and load also are computed in WRTDS. The flow-normalized estimates are designed to remove the variation in concentration or load because of random streamflow variations (but not the variation because of nonrandom seasonal streamflow variations); the effects of antecedent streamflow conditions are not removed. The temporal variation in streamflow is removed in WRTDS by assuming the streamflow that occurred on any given day of the record is one sample from the probability distribution of streamflows for that particular day of the year. To compute the flow-normalized estimate of concentration for a given date (for example, assume there are n years of record and the estimation is for the flow-normalized concentration for August 10, 1999), WRTDS uses n separate weighted regressions to estimate concentration on that date with the streamflow value set to each one of the n historical streamflow values for that day of the year (in this example, every August 10 in the period of record). The flow-normalized concentration on that date (August 10, 1999, in this example) is then calculated as the mean of the estimated concentration values from each of those n weighted regressions. Similarly, each of the n estimated concentration values is multiplied by its corresponding streamflow value to get n estimated load values; the mean of the n estimated loads is the flow-normalized load for that day. In this report, flow-normalized estimates are referred to as flow-normalized concentration or flow-normalized load; nonflow-normalized estimates are referred to as estimated concentration or estimated load.

For many of the analyses herein, the daily estimates were summarized into water-year (October 1 through September 30) annual means for estimated and flow-normalized concentration and water-year annual totals for estimated and flow-normalized load. Trends in annual mean concentration were calculated in three ways:

$$\text{Net change} = c_{t_2} - c_{t_1} \tag{17}$$

where

c_{t_1} is the annual mean concentration, in year t_1;
c_{t_2} is the annual mean concentration, in year t_2;

$$\text{Net change in percentage} = \left(\frac{\left(c_{t_2} - c_{t_1}\right)}{c_{t_1}}\right) * 100 \tag{18}$$

and

$$\text{Rate of change in percentage per year} = \frac{\left(\left(\frac{\left(c_{t_2} - c_{t_1}\right)}{c_{t_1}}\right) * 100\right)}{n} \tag{19}$$

where

n is the number of years in the time period, $t_2 - t_1$.

Trends in total annual load were calculated in a similar manner.

The confidence intervals and associated significance level of the trend are determined through a block bootstrap approach based on a time interval of 100 days to avoid oversampling any of the more densely sampled periods during the overall period of record and to broadly maintain samples from individual high or low streamflow events (Hirsch and others, 2015). The bootstrap method is used to generate a 90-percent confidence interval on the magnitude of the trend (calculated using equations 17, 18, or 19) and a likelihood statistic that is the functional equivalent to the two-sided p-value. The likelihood statistic provides information on whether the null hypothesis that there is no trend over the period of record should be rejected and provides a measure of the strength of evidence that the trend is actually occurring. Trends in concentration are evaluated separately from the trends in load because the two trends may differ.

The confidence intervals and likelihood statistic were determined only for the flow-normalized values because they are more stable than the estimated values, which display a great deal of year-to-year streamflow-driven variation. This makes flow-normalized concentration and load ideal for evaluating progress toward nutrient reduction goals: however, for studies of ecological processes in the watershed or in receiving waterbodies such as Chesapeake Bay or the Gulf of Mexico, estimated concentration and load are more relevant. Future versions of WRTDS may extend the bootstrapping concept to estimated concentrations and loads, but for this report, only trends in flow-normalized concentration and load are reported for a specific site, parameter, and trend period. A trend result of "true" indicates that the two-sided p-value generated in the output file of EGRETci for the WRTDS model is less than the critical alpha value of 0.1.

There are several high-frequency datasets (daily or near-daily sampling) in the final set of trend sites for this study. High-frequency data are more likely to contain serial correlation—correlation of a time series with its own past and future values (for example, high concentrations are more likely to follow prior high concentrations and low concentrations are more likely to follow prior low concentrations). With traditional regression methods, serial correlation can result in error estimates that are smaller than the true error, which leads to a tendency to reject the null hypothesis of no trend (it will find a trend) when in fact there is no trend. The bootstrapping approach employed in this study does not have the same challenges, and the likelihood statistic will be largely unaffected by serial correlation. As a result, high frequency data were not thinned to less than one sample per day.

Early studies established conservative data criteria for use of the WRTDS model: water-quality data containing no more than 1 percent censoring, no longer than 4 years between consecutive water-quality samples, sample size greater than 200, water-quality period of record longer than 20 years, and a complete record of daily streamflow for the entire water-quality period of record. Recent advancements in the WRTDS model have accommodated greater data censoring by employing survival regression (Hirsch and others, 2015), so data with as much as 50 percent censoring were used in this study (Hirsch and others, 2015). Data gaps of longer than 4 years in the water-quality data are problematic because these periods lack relevant calibration information, and the estimates made during these times can be unreliable; therefore, these years were removed from the time series of estimated concentration and load by use of the blankTime function in WRTDS. In this study, the entire water-quality record also was used to calibrate the WRTDS model for a site, even if only part of the record passed the final data screens for trend analysis. Estimates made outside the screened trend periods, during periods when data were sparse, ultimately were removed using the blankTime function. Other recent studies have shown that WRTDS estimates may be reliable for data records with as few as 60 samples and a period of record as short as 10 years (Hirsch and De Cicco, 2015; Chanat and others, 2016). Sensitivity analyses conducted for this study (see the "Sampling Frequency" section) indicate reliable estimates are possible with sample sizes as small as 35, but only if the samples are well distributed throughout the year (at least one sample every quarter) and if the streamflow conditions captured in the water-quality dataset are comparable to the streamflow conditions across the period of record. Because the flow-normalization process required a complete record of daily flows, all sites were paired to a streamgage that had a complete record of daily streamflow during the assessed trend periods (see the "Gage Matching" section).

The R packages EGRET 2.2.0 and EGRETci 1.0.4 were used for this study. The half-window widths used in this study were 7 years for time, 1 log cycle for streamflow at sites with drainage area greater than 250,000 km^2, and 2 log cycles for streamflow at sites with drainage area less than 250,000 km^2, and 0.5 years for season. The minimum number of observations with nonzero weight required for each regression was set to 100 or to 75 percent of total observations (whichever was smaller), and the minimum number of uncensored observations required for each regression was set to one-half of the minimum number of observations. The edgeAdjust option was set to TRUE so that the window widths near the beginning and end of the record were extended to allow for more stable estimates during those periods. The variables bootBreak (the number of bootstrap iterations, which governs a stopping rule for the adaptive bootstrapping), blockLength (block length for bootstrapping, in days), and nBoot (number of bootstrap replicates) were all set to 100 for the bootstrapping part of the modeling. A value setting of 100 is suggested in Hirsch and others (2015); a setting of 100 will provide results that are much more reproducible and precise than a lower value setting. More information on these specifications in EGRET and EGRETci can be found in Hirsch and De Cicco (2015) and Hirsch and others (2015).

All WRTDS trend models were reviewed according to the procedures described in Hirsch and De Cicco (2015), with additional review steps described herein. Even though the WRTDS model is highly flexible, there is no assurance that it will provide reasonable estimates of concentration and load for all days in a record (Hirsch and De Cicco, 2015). Graphical tools were used to identify serious problems with the model fit as well as to identify suspicious patterns in the observed data that were not identified during the automated screening process. When serious problems were identified, the model results were rejected from the study. Serious problems included (1) severe outliers in the raw data that resulted in severe outliers in the residuals; (2) severe structure in the residuals when plotted against time, streamflow, or season; (3) severe structure in concentration residuals versus estimated concentration, indicating potential for biased or inaccurate concentration estimates; or (4) a poor relation between observed and estimated concentration or observed and estimated load.

Three additional quantitative metrics were used in this study to identify extreme problems with the model fit and automatically reject models from the study. The first was an extrapolation metric indicating how much greater the highest estimated concentration was compared to the highest observed concentration:

$$\text{extrapolation metric}=(\text{maximum estimated concentration})/(2*\text{maximum observed concentration}) \qquad (20)$$

When this metric was greater than two (for example, when the highest estimated concentration was at least four times greater than the highest observed concentration), the model was rejected from the study.

The second and third metrics were variations on a flux bias statistic that was created in response to a problem that several researchers previously identified with regression-based models like WRTDS (Hirsch and De Cicco, 2015). Flux bias is based on what is called load in this report. Hirsch and De Cicco (2015) define the flux bias statistic as a dimensionless representation of the difference between the sum of the estimated fluxes on all sampled days and the sum of the true fluxes on all sampled days. A value near zero indicates that the model is nearly unbiased. A positive value indicates a positive bias, and a negative value indicates a negative bias. Values that are between -0.1 and +0.1 indicate that the bias in estimates of the long-term mean flux is likely to be less than 10 percent. In this study, a truncated flux bias statistic also was calculated. The truncated flux bias statistic was calculated in the same manner as the regular flux bias statistic, but the calculation excluded the two largest absolute values of the residuals (in real space). This metric provided supplementary information on the nature of the model bias—it was used to evaluate whether the total model flux bias resulted from relatively small biases across all or most of the estimates or if it resulted from large biases represented by one or two large residual outliers. Models where the regular and truncated flux bias statistics exceeded +/- 0.4 (potentially indicating a model bias between 25 and 100 percent across all or most of the estimates) were rejected from the study. Rejected models are noted in appendix 1, table 1–6.

The Seasonal Kendall test (SKT) is a robust, nonparametric trend test that is widely used to test for trends in environmental parameters (Hirsch and Slack, 1984). Relative to WRTDS, SKT is a more conservative trend test that requires fewer assumptions about data quality; therefore, it was used to assess the overall distribution of trends in the dataset in comparison to WRTDS. All sites and parameters analyzed for trends using WRTDS were also analyzed for trends using the SKT.

Pesticide Trend Method

The seasonal wave (SEAWAVE–Q) model (Vecchia and others, 2008; Sullivan and others, 2009; Ryberg and others, 2010, 2014; Ryberg and Vecchia, 2013) was selected as the statistical tool for analyzing trends in pesticide concentrations for this study. The SEAWAVE–Q model was specifically developed to address a "number of difficulties often found in pesticide data, such as strong seasonality in response to use patterns, high numbers of concentrations below laboratory reporting levels (LRLs), complex relations between streamflow and concentration, and intermittent or changing sampling frequencies (both interannually and intraannually)" (Vecchia and others, 2008).

The SEAWAVE–Q model is a parametric regression model specifically designed for analyzing seasonal- and flow-related variability and trends in pesticide concentrations. The model is expressed as the following:

$$Log\ C(t) = \beta_0 + \beta_1 W(t) + \beta_2 LTFA(t) + \beta_3 MTFA(t) + \beta_4 STFA(t) + \beta_5 t + \eta(t) \qquad (21)$$

where

Log denotes the base-10 logarithm;
C is pesticide concentration, in micrograms per liter;
t is decimal time, in years, with respect to an arbitrary time origin;
$\beta_0, \beta_1, \ldots,$ are regression coefficients;
W is a seasonal wave representing intraannual variability in concentration;
$LTFA$, $MTFA$, and $STFA$ are dimensionless long-term, mid-term, and short-term streamflow anomalies computed from daily streamflow (described later in this section); and
$\eta(t)$ is the model error.

The seasonal wave is a dimensionless, periodic function of time with an annual cycle, similar to a mixture of sine and cosine functions often used to model seasonality in concentration data; however, the seasonal wave is better suited for modeling seasonal behavior of pesticide data than a mixture of sines and cosines. The seasonal wave is a periodic (with a period of 1 year) solution to the following differential equation (Vecchia and others, 2008):

$$\frac{d}{dt}W(t) = \lambda(t + s^*) - \varphi W(t) \qquad (22)$$

where

$\frac{d}{dt}$ is the derivative with respect to time;
W is a seasonal wave representing intra-annual variability in concentration;
t is decimal time, in years, with respect to an arbitrary time origin;
$\lambda(.)$ is a pulse input function with $\lambda(.)$ greater than 0 during specified application season(s) and $\lambda(.)$ equal to 0 otherwise;
s^* is a seasonal shift that determines the time at which W reaches its maximum; and
ϕ is a "decay rate" corresponding to an approximate half-life of $12/\phi$ months.

As in Sullivan and others (2009), the pulse input function is selected in **seawaveQ** from a menu of 14 (2 × 7) choices based on selecting either one or two distinct application seasons (when pesticides may be transported to the stream) with time lengths from 1 to 6 months. The half-life is selected from four choices (1, 2, 3, or 4 months). The half-life is referred to as a model half-life when discussing model results in order to distinguish it from the chemical half-life of pesticides. Thus, 56 (14 application seasons × 4 half-lives) choices for the seasonal wave function are available. As described in Sullivan and others (2009), the observed concentration data were used to select the best wave function for each model and to estimate the seasonal shift (s^*) through a combination of graphical and maximum likelihood techniques.

Three dimensionless streamflow anomalies were included in the SEAWAVE–Q model to help account for flow-related variability in pesticide concentrations. The anomalies were computed using the R extension package waterData (Ryberg and Vecchia, 2012) and are based on log-transformed daily streamflow aggregated over various time scales. The first anomaly represented short-term (day-to-day) streamflow variability and was defined as

$$STFA(t) = X(t) - X_{30}(t) \qquad (23)$$

where

$STFA$ is the short-term streamflow anomaly (dimensionless);
t is decimal time, in years;
$X(t)$ is log-transformed daily flow, in cubic meters per second; and
$X_{30}(t)$ is the average of log-transformed daily streamflow for 30 days up to and including time t.

Large positive values of *STFA* and associated increases in pesticide concentrations tend to occur near the beginning of a substantial rainfall-runoff or snowmelt event, whereas negative values of *STFA* and associated decreases in pesticide concentrations tend to occur after the event passes (Vecchia and others, 2008).

The second streamflow anomaly represents mid-term (30 to 365 day) streamflow variability and was defined as

$$MTFA(t) = X_{30}(t) - X_{365}(t) \qquad (24)$$

where

$MTFA$ is the mid-term streamflow anomaly (dimensionless);
t is decimal time, in years;
$X_{30}(t)$ is the average of log-transformed daily streamflow for 30 days up to and including time t; and
$X_{365}(t)$ is the average of log-transformed daily streamflow for 365 days up to and including time t.

The third streamflow anomaly added to the model represented long-term (greater than 365 days) streamflow variability and was defined as

$$LTFA(t) = X_{365}(t) - X_* \qquad (25)$$

where

$LTFA$ is the long-term streamflow anomaly (dimensionless);

t is decimal time, in years;

$X_{365}(t)$ is the average of log-transformed daily streamflow for 365 days up to and including time t; and

X_* is the average of log-transformed daily streamflow for the specified trend period (1992–2012 or 2002–12).

Unlike *STFA*, which tends to affect pesticide concentrations in a relatively consistent manner among different sites and pesticides, *MTFA* and *LTFA* can affect pesticide concentrations in different ways and to different degrees depending on the type of pesticide, the size of the basin, and the climatic and hydrologic properties of the basin. For example, for a relatively large basin with substantial nonurban or nonagricultural runoff, higher-than-normal annual streamflow conditions (as indicated by a positive value for *LTFA*) can cause decreased pesticide concentrations because of more dilution from runoff in areas where pesticides are not used (Sullivan and others, 2009); however, in a relatively small basin with primarily agricultural runoff, higher-than-normal annual streamflow conditions can cause increased pesticide concentrations as wetter conditions lead to more shallow groundwater and overland runoff from areas with high pesticide application rates.

The SEAWAVE–Q model (equation 21) was fitted to the pesticide data using maximum likelihood methods for censored data (Sullivan and others, 2009) using the statistical software R (R Development Core Team, 2014) and the R extension package **seawaveQ** (Ryberg and Vecchia, 2013).

Statistical significance of the model was determined using the t-test (Neter and others, 1996) of significance of the model coefficients (β_0–β_5, equation 21). A p-value greater than or equal to 0.10 for the trend coefficient (β_5) indicated that the trend was not statistically significant. A p–value less than 0.10 for β_5 indicated a statistically significant upward or downward trend (also referred to as uptrend or downtrend).

Unlike the concentration trends determined using WRTDS, which are expressed in terms of the flow-normalized annual mean concentration, the trend using the SEAWAVE–Q model is interpreted as a trend in the flow-adjusted annual median concentration. Also, unlike WRTDS, the SEAWAVE–Q trend in flow-normalized annual load cannot be different (computationally) from the trend in flow-normalized annual concentration when there is no trend in flow.

Daily pesticide concentration estimates provided by **seawaveQ** (Ryberg and Vecchia, 2013) are corrected for retransformation bias (the concentration model is built on the base-10 logarithm of concentration; therefore, a bias correction is required when transforming back to the original units) and then used to calculate daily loads. The bias correction used was

$$e^{\frac{\left(\log(10)*scl\right)^2}{2}} \qquad (26)$$

where

e is the exponentiation function,

$\log(10)$ is the natural logarithm of the number 10, and

scl is the scale parameter from the seawaveQ output (Ryberg and Vecchia, 2013).

This correction, based on the quasi maximum likelihood estimator (Cohn and others, 1989) that was developed for natural logarithms, is adjusted for the SEAWAVE–Q model, which models the base-10 logarithm of the concentration. To calculate loads, the bias corrected concentration was multiplied by daily streamflow and a constant, 0.892998605, which converts the load units (micrograms per liter * cubic feet per second) to kilograms per year. Daily loads were summed to annual values.

Trends in concentration were calculated in two ways:

$$\text{Net change in percentage} = 100*\left(\left(10^{\beta_5}\right)^{t_2-t_1}-1\right) \qquad (27)$$

where

β_5 is the trend coefficient from equation 21,

t_1 is the year at the start of the trend period, and

t_2 is the year at the end of the trend period.

$$\text{Net change in microgram per liter} = \left(c_{t_1}*\left(10^{\beta_5}\right)^{t_2-t_1}\right)-c_{t_1} \qquad (28)$$

where

c_{t_1} is the median concentration from the trend line at the midpoint of year t_1, and

c_{t_2} is the median concentration from the trend line at the midpoint of year t_2.

The confidence intervals on these two trend types are determined as follows:

$$\text{Upper and lower confidence intervals in percentage}$$
$$=\left[100*\left(10^{\beta_5}\pm\left(qnorm\left(1-\frac{\alpha}{2}\right)*se\left(\beta_5\right)\right)\right)^{t_2-t_1}-1\right] \qquad (29)$$

where

α is the significance level 0.10 for a two-sided 90-percent confidence interval.,

$se(\beta_5)$ is the standard error for the trend coefficient $\beta_{5,}$ and

$qnorm()$ is the $1-\frac{\alpha}{2}$ critical value for the standard normal distribution in R.

$$\text{Upper and lower confidence intervals in micrograms} \qquad (30)$$
$$\text{per liter} = c_{t_1}*\left[10^{\beta_5}\pm\left(qnorm\left(1-\frac{\alpha}{2}\right)*se\left(\beta_5\right)\right)^{t_2-t_1}\right]-c_{t_1}$$

Trends and confidence intervals for total load were calculated in a similar manner, where the load in years t_1 and t_2 were based on bias-corrected daily estimates of concentration from the model and daily streamflow.

A likelihood value that is the functional equivalent of the two-sided p-value associated with the significance level of the trend was determined as follows:

$$\text{Likelihood} = \left(1 - \left(\frac{\text{p-value}}{2}\right)\right) \quad (31)$$

where

p-value is the p-value for the trend coefficient β_5.

Ecology Trend Method

Similar to water-quality data, there are environmental factors (for example, streamflow and climate conditions) that potentially influence ecological endpoints and may, therefore, confound trend results. To address this issue, at every site, two steps were taken to help select and account for confounding factors during trend analysis. First, the Pearson correlation analysis (Harrell and others, 2016) was used to identify the environmental variable (X, table 8) that had the strongest correlation (highest absolute r value, which measures the strength of the linear relation) with each ecological endpoint (Y, table 7). For each ecological endpoint, it was assumed that the most strongly correlated environmental variable had the highest potential to confound trend analysis. Second, the residuals were extracted from a simple linear regression model (R Core Team, 2014) of each X–Y pair, site, and trend period combination. The raw residuals represented the ecological endpoint after adjusting for the presumed confounding environmental variable.

The Kendall-tau test for trends (Lorenz, 2016) was used to detect monotonic trends in the unadjusted ecological endpoints and in the residuals (adjusted ecological endpoints) at each site within each trend period. This nonparametric analysis has been widely used to detect monotonic trends in environmental data (Esterby, 1996). Trend results for both the adjusted and unadjusted data are reported, in percentage.

$$\text{Percentage change} = 100 * \text{Sen Slope} * \left(t_2 - t_1\right) / \left(M_{t_1}\right) \quad (32)$$

where

Sen Slope is the median of the slopes computed from the data pairs used in the computation of Kendall’s tau from the Kendall-tau test for trends (Lorenz, 2016),

t_1 is the year at the start of the trend period, and

t_2 is the year at the end of the trend period, and

M_{t_1} is the predicted median value, in year t_1.

Additionally, a likelihood value calculated as the functional equivalent of a two-sided p-value associated with the significance level of the trend was determined as follows:

$$\text{Likelihood} = \left(1 - \left(\frac{\text{p-value}}{2}\right)\right) \quad (33)$$

where

p-value is the p-value from the Kendall-tau test for trends (Lorenz, 2016).

Limitations on Trend Analysis and Effects on the Weighted Regressions on Time, Discharge, and Season (WRTDS) Model Estimates

Most studies that examine changes in water quality over time are challenged by common issues pertaining to data density, sample representativeness, and the stability of estimates as new data are added to the calibration record. A successful study of water-quality trends, one in which the identified trend holds up to scrutiny, requires researchers to make sound decisions to appropriately address and resolve the aforementioned issues and any others that arise because of the data or method used. This section will briefly discuss several broad issues facing water-quality trend studies and how these issues affect WRTDS model results. Specifically, this section presents several sensitivity tests to demonstrate how WRTDS estimates are affected by the frequency of sample collection throughout the year, the frequency of storm-sample collection, the precision (rounding) of water-quality data, and the incremental addition of new data to the calibration dataset. This discussion is not intended to be an exhaustive characterization of these issues or their effect on WRTDS results but provides examples of some issues faced in this study and other water-quality trend studies.

Storm Sampling

Water-quality samples collected during high-flow periods are important for an unbiased WRTDS calibration but usually require extra effort because of the random nature of hydrologic events. The relation between concentration and streamflow is one of the primary relations modeled in WRTDS and is not always well defined when sites lack a sufficient number of high-flow samples. As a result, the model will extrapolate beyond the calibration conditions to make estimates of load and concentration at high flows that might be invalid and possibly result in inaccurate trend estimates. Additionally, sites lacking a sufficient number of high-flow samples are prone to overprediction and underprediction during high-flow periods, which affects annual estimates of mean concentration, load, and long-term trends.

The issue of sufficient high-flow coverage has been acknowledged in other investigations involving the estimation of concentrations, loads, and trends using streamflow-concentration regression methods such as WRTDS (Preston and others, 1989; Chanat and others, 2016). Because the authors of this report are unaware of studies that have developed and applied consistent, objective criteria for the determination of adequate representation of flows to a large number of sites, that effort was undertaken in this study. Experiments were conducted to determine the number of high-flow samples that could be removed (by random subsampling) from the data record at a site before the estimates from WRTDS were substantially different from estimates using all data available for a site. Criteria were developed using (1) sites having a range of watershed areas and (2) multiple water-quality parameters to evaluate differences in suspended and dissolved phase behavior.

Prior to running the WRTDS experiments, it was necessary to develop an operational definition of a high-flow sample. The term "high flow" is subjective because intersite and intrasite variation, as well as project-specific goals, confound a simple, universal definition of high flow. For the purposes of this study, high streamflow was defined as the 85th percentile of the decadal monthly flow. Decades were used in order to account for prolonged droughts or wet periods. The 85th percentile of the decadal monthly streamflow was determined as follows:

1. Divide the available streamflow record for each streamgage into decades corresponding to the trend periods used in this study: 1972–81, 1982–91, 1992–2001, and 2002–12.
2. Extract all of the mean daily January streamflow values in a decade (January 1, 1972, January 2, 1972, * * * January 31, 1981) and calculate the 85th percentile (n= 31 days × 10 years = 310 streamflow values).
3. Repeat step 2 for each month in a decade.
4. Repeat steps 2 and 3 for each decade.

As table 9 shows for two streamgages, the 85th percentile high-flow threshold value varies seasonally and in response to multiyear precipitation cycles. No single threshold of streamflow can be reasonably selected to identify high-flow samples. For the sites shown in table 9, a relatively low threshold value appropriate for summer high-flow events would lead to a characterization of all (or nearly all) winter samples as high flow; similarly, a relatively high threshold value that properly characterizes winter high-flow events would miss nearly every summer storm. Similar, but smaller, variations in high streamflow occur across decades for each month.

The 85th percentile of the decadal monthly streamflow was selected as an approach that worked well to establish the high-flow threshold value for a range of sites across the country. Decadal monthly streamflow percentiles ranging from the 75th to 95th were evaluated for use as the high-flow threshold. In general, lower percentiles were found to be suitable for larger streams, like the Delaware River at Trenton, New Jersey, that respond to storms over a longer duration and therefore spend relatively more time at higher flows (fig. 3*A*). Higher percentiles were needed to characterize smaller, flashier streams, like Fanno Creek at Durham, Oregon, that spend relatively more time near their seasonal base-flow condition (fig. 3*B*). In the example 4 years shown in figures 3*A* and 3*B*, the 85th percentile of decadal monthly streamflow spans a 6-fold range of streamflow values at Delaware River and spans more than a 35-fold range of streamflow values at Fanno Creek.

Figure 4 shows the distribution of total phosphorus samples at decadal monthly streamflow quantiles ranging from 75th to 95th percentiles for the same time interval at the same sites depicted in figure 3. The seasonal and multiyear differences reflected in figure 4 are typical of those observed at other sites that were examined. Note that at both sites, samples collected at the 75th percentile were collected at flows generally above the seasonal base streamflow but often fall short of storm peaks when concentrations of many parameters are near their maximum and are not high on the rising or falling limb of storms. Samples collected at the 85th percentile of flows capture more storm peaks and are more spread out over the 4-year span than those at the 95th percentile of flows.

Once the method to identify a high-flow sample was determined, experiments were undertaken with WRTDS to evaluate the sensitivity of the results to a reduction in the number of high-flow samples. The experimental data were generated by removing high-flow samples from nitrate and total phosphorus datasets at six water-quality sites. The six sites were: 01463500, Delaware River at Trenton, New Jersey; 01654000_22340331, Accotink Creek near Annandale, Virginia; 02338000, Chattahoochee River near Whitesburg, Georgia; 06805500, Platte River at Louisville, Nebraska; 14211720, Willamette River at Portland, Oregon; and 14206950, Fanno Creek at Durham, Oregon. Descriptions of these sites are in appendix 1, table 1–6. Nitrate was selected to evaluate the effect of high-flow samples on model estimates of water-quality parameters transported primarily in the dissolved phase, and total phosphorus was selected to evaluate the effect on model estimates of parameters having a large sediment-bound fraction. The six sites were selected from sites that passed all water-quality and streamflow continuity screens (see the "Screening Streamflow Record for Streamflow Continuity" section) and were evaluated to ensure they had a relatively large percentage of high-flow samples (median = 15 percent high-flow samples among all sites and decades). Sites were selected to represent a range of watershed areas (62 to 221,107 km^2).

Using the selected sites, WRTDS models were run on the original nitrate and total phosphorus datasets and on downsampled versions of the datasets. The original datasets for each site were randomly downsampled to generate experimental datasets such that each trend decade (as defined previously in this section) contained 2, 6, 10, 14, 18, and 22 percent high-flow samples. Trend decades at some sites originally contained fewer high-flow samples than were dictated by the specified downsampling percentage; in which case, the trend decade was not downsampled. An example of a downsampled dataset for site 14211720 Willamette River at Portland, Oregon, is shown in table 10.

Table 9. The 85th percentile of decadal monthly flow for site 01463500 Delaware River at Trenton, New Jersey, and site 14206950 Fanno Creek at Durham, Oregon.

[Values reported in cubic feet per second]

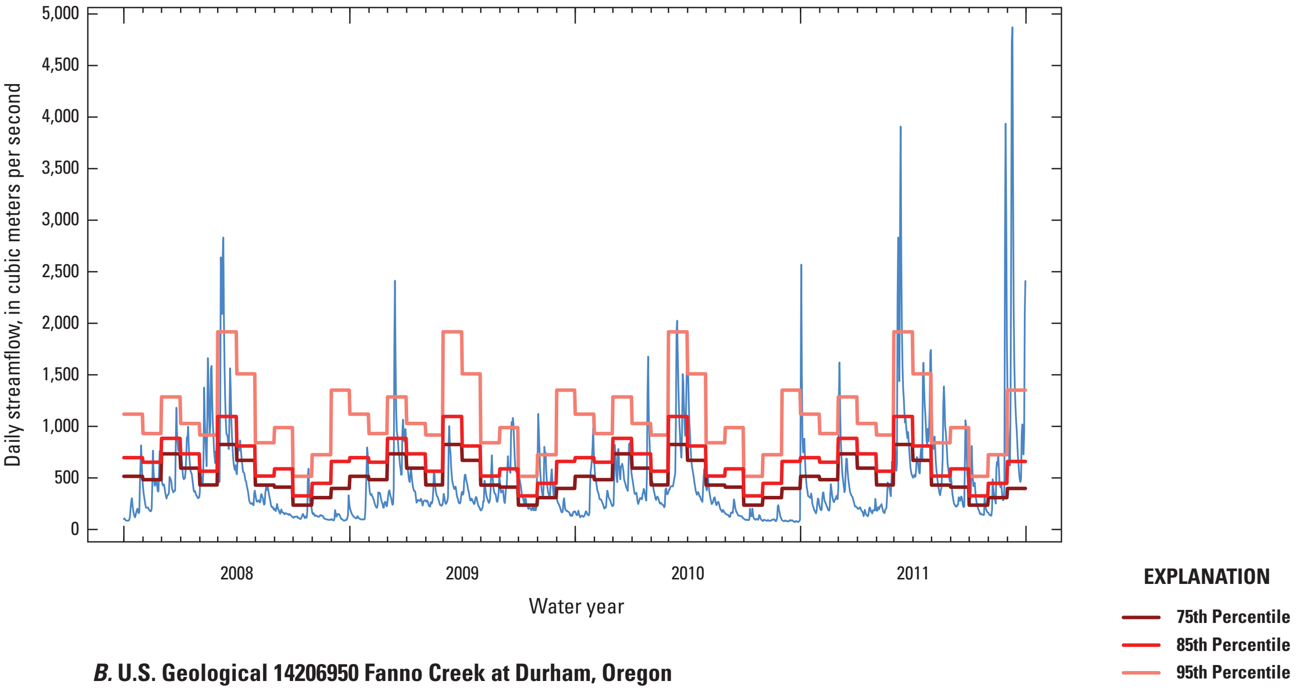
Site 01463500 Delaware River at Trenton, New Jersey

Decade	January	February	March	April	May	June	July	August	September	October	November	December	Range
1972–81	28,785	26,310	34,920	35,020	26,630	16,415	12,400	8,013	9,171	18,100	19,415	27,860	27,007
1982–91	13,930	21,655	23,100	32,560	27,400	17,300	10,800	7,360	6,765	10,165	15,015	20,160	25,795
1992–2001	19,730	17,980	28,665	39,260	18,895	14,645	9,117	7,405	6,296	9,063	16,345	22,225	32,965
2002–12	25,900	19,950	38,700	28,600	18,300	20,765	11,500	15,800	23,300	24,600	23,030	31,200	27,200
Range	*14,855*	*8,330*	*15,600*	*10,660*	*9,100*	*6,120*	*3,283*	*8,440*	*17,005*	*15,537*	*8,015*	*11,040*	
Range as percent of mean	*67*	*39*	*50*	*31*	*40*	*35*	*30*	*88*	*149*	*100*	*43*	*44*	

Site 14206950 Fanno Creek at Durham, Oregon

Decade	January	February	March	April	May	June	July	August	September	October	November	December	Range
2002–12	196	101	119	77	53	25	9	6	9	34	117	214	208
Range	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	
Range as percent of mean	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	

A. **U.S. Geological Survey 01463500 Delaware River at Trenton, New Jersey**



B. **U.S. Geological 14206950 Fanno Creek at Durham, Oregon**

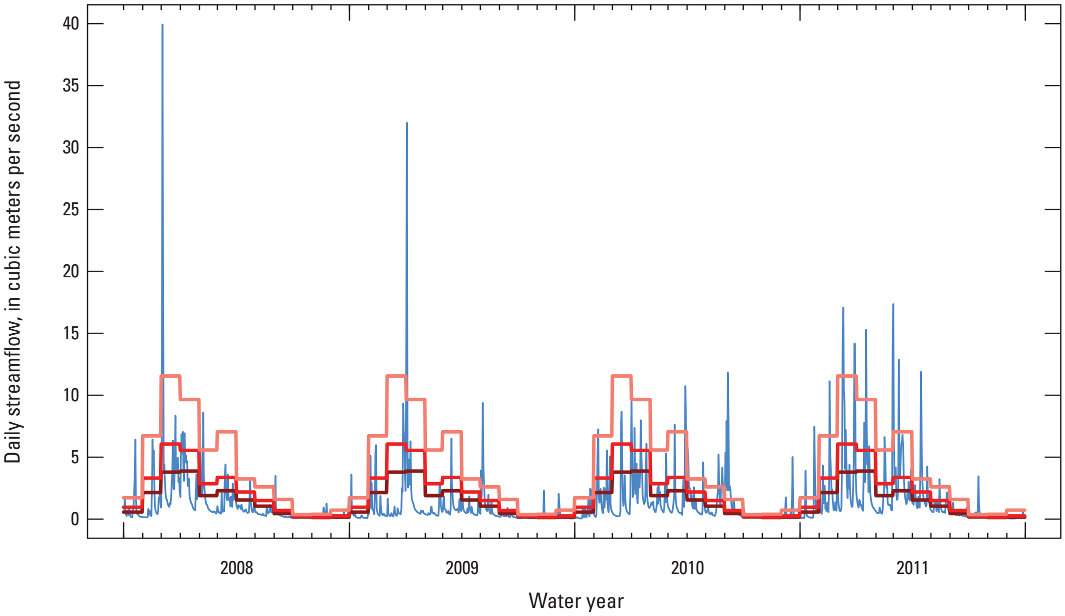
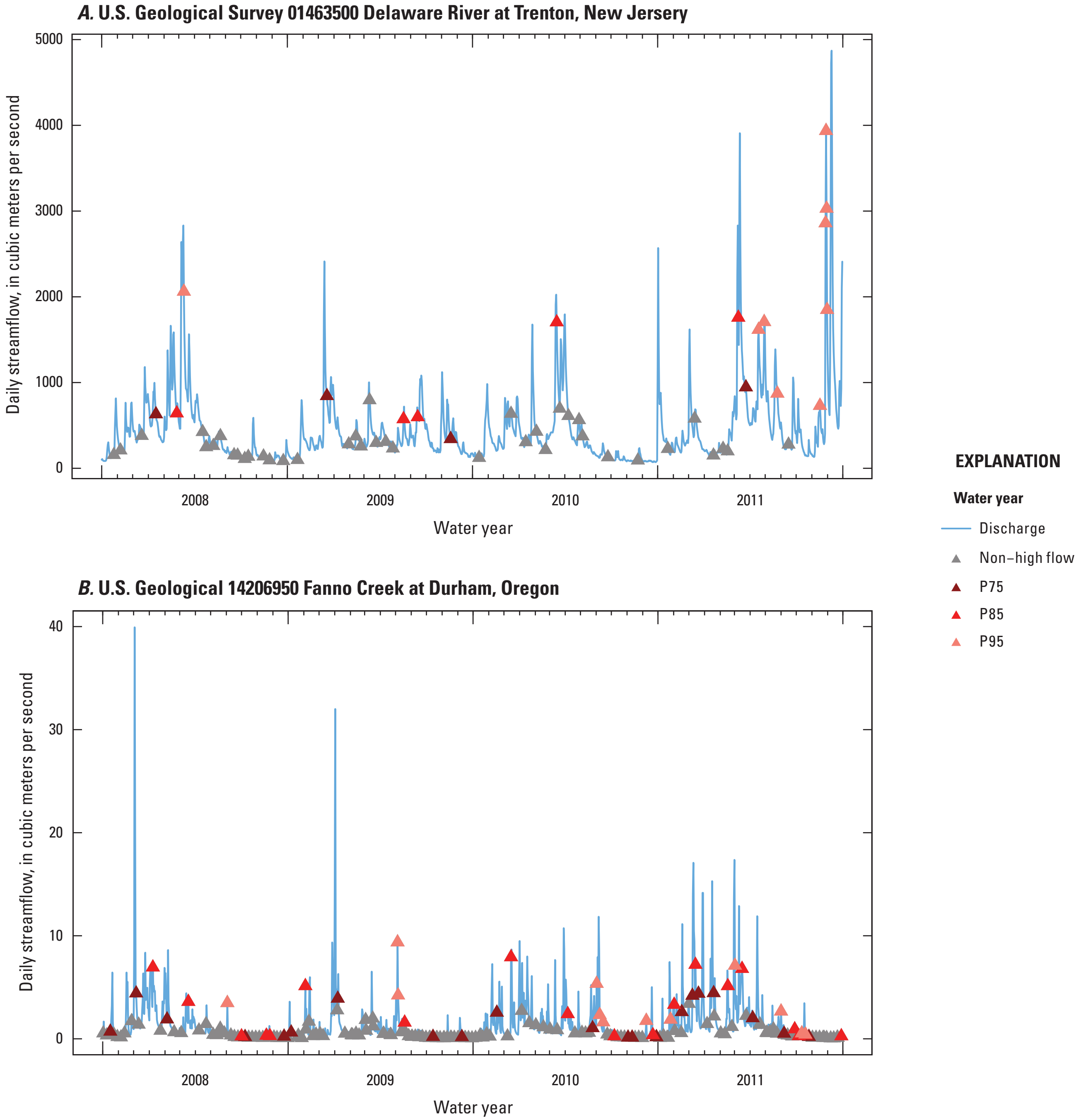


Figure 3. The 75th, 85th, and 95th decadal monthly streamflow percentiles for water years 2008–11 for sites *A*, 01463500 Delaware River at Trenton, New Jersey; and *B*, 14206950 Fanno Creek at Durham, Oregon.



A. **U.S. Geological Survey 01463500 Delaware River at Trenton, New Jersey**

B. **U.S. Geological 14206950 Fanno Creek at Durham, Oregon**

Figure 4. Collection points of total phosphorus samples exceeding the 75th, 85th, and 95th decadal monthly streamflow percentiles for water years 2008–11 for sites *A*, 01463500 Delaware River at Trenton, New Jersey; and *B*, 14206950 Fanno Creek at Durham, Oregon.

Table 10. Number of nitrate samples at site 14211720 Willamette River at Portland, Oregon, in original and downsampled dataset. Data were downsampled such that the number of high-flow samples in each trend decade was approximately 10 percent of the total number of samples in that trend decade.

Trend decade	Number of samples in original dataset	Number of high-flow samples		Percentage of high-flow samples	
		Original dataset	Downsampled dataset	Original dataset	Downsampled dataset
1972–81	82	14	8	17.1	10.5
1982–91	58	8	6	13.8	10.7
1992–2001	118	24	11	20.3	10.5
2002–12	178	24	18	13.5	10.5

Table 11. Example of high-flow sample calculations and application of high-flow screening criteria for site 06818000 Missouri River at St. Joseph, Missouri.

[Y, yes; N, no; NA, not applicable]

Site	Parameter	Percentage of high-flow samples				High-flow criteria met?			
		1972–81	1982–91	1992–2001	2002–12	1972–2012	1982–2012	1992–2012	2002–12
06818000	Nitrate	17.5	10.6	13.6	15.9	Y	N	Y	Y
06818000	Orthophosphate, filtered	NA	10.6	13.7	15.9	NA	N	Y	Y
06818000	Total nitrogen	NA	10.7	12.4	15.9	NA	N	Y	Y
06818000	Total phosphorus	17.5	10.7	12.4	15.9	Y	N	Y	Y

In general, downsampling total phosphorus data resulted in lower estimates of concentration, flow-normalized concentration, load, and flow-normalized load compared to estimates made with full datasets (fig. 5), while downsampling nitrate data resulted in higher estimates of concentration, flow-normalized concentration, load, and flow-normalized load (fig. 6). Model estimates of concentration and load in the downsampled data were generally within 10 percent of the respective model estimates using the original dataset when the downsample high-flow percentage was 10 percent or more, and nearly identical when the downsample high-flow percentage was 14 percent or more. When the percentage of high-flow samples was reduced to 6 percent or even more so, to 2 percent, model estimates were dramatically different from estimates using the original dataset containing all high-flow samples (fig. 7).

The criteria for the minimum number of high-flow samples were determined after evaluating the downsampling results. It was decided that sites would be required to have at least 10 percent high-flow samples in all trend decades to be modeled and at least 14 percent high-flow samples in half of the trend decades to be modeled. In the screening logic example shown in table 11, nitrate data from 1972 to 2012 are suitable for trend analysis because (1) all trend decades have at least 10 percent high-flow samples and (2) half of all trend decades with sufficient data have at least 14 percent high-flow samples. Nitrate data from 1982 to 2012 are not suitable for trend analysis because half the trend decades do not have at least 14 percent high-flow samples.

Sampling Frequency

The original guidance for WRTDS recommended its use at sites with long-term (at least 20 years) and dense (at least 200 samples) water-quality datasets and with continuous daily streamflow for the entire water-quality period of record (Hirsch and others, 2010). Refinements to WRTDS and more recent experience with the method have indicated that a reasonable trend estimation can be achieved for datasets as short as 10 years and as few as 60 samples (Hirsch and De Cicco, 2015; Chanat and others, 2016). Many agencies collect just four samples per year in their monitoring programs. To maximize the use of such multiagency data, this study required a sampling frequency of at least four samples per year for the first 2 years and the last 2 years of the trend period, as well as at least four samples per year in 70 percent of the years in the trend period (See “Final Screening for Data Coverage” section for more detail). Although only 28 samples were required to pass the initial WRTDS trend screening, the minimum number of samples used in any WRTDS trend model in this study was 35. To explore the effect of sample frequency on WRTDS estimates and the use of data that were collected less frequently than is typically recommended for the model, a sensitivity test was completed at two sites using water-quality data for nitrate and total phosphorus.

Water-quality records from two sites having greater than 1,000 observations (approximately 25 samples per year) and a maximum trend period of 1972–2012 or 1982–2012 were downsampled to explore the influence of sample frequency on WRTDS estimates. At each site, all available data were used to calibrate WRTDS, and then each complete dataset was downsampled iteratively, without replacement, to 12, 9, 6, 4, and 2 samples per year. Downsampled datasets were then used to calibrate WRTDS, estimate annual concentrations and loads, and calculate trends in flow-normalized concentration and flow-normalized load for each of the trend periods used in this study (that is 1972–2012, 1982–2012, 1992–2012, and 2002–12). Each iteration of downsampling began with the previously downsampled dataset. Downsampling was accomplished by dividing the water-quality record into a predefined number of seasons corresponding to the desired number of samples per year and then randomly selecting without replacement one water-quality sample from each season. The downsampling routine did not specifically consider the proportion of samples collected during high flows, though the complete datasets had been previously screened using high-flow criteria from this study. Therefore, the downsampled datasets may have fewer high-flow samples than is ideal. The results of this downsampling exercise reflect a combination of the effects of having overall fewer samples and a loss of seasonal and high-flow resolution in the WRTDS model.

Figure 8 shows the time series of annual flow-normalized concentration and flow-normalized load estimates from models calibrated using the complete dataset and using all downsampled datasets for nitrate and total phosphorus at sites mdWIL0013 and iaDMRWQN 6 (appendix 1, table 1–6). These results hereafter are referred to as all-data, ds-12, ds-9, ds-6, ds-4, and ds-2 estimates or downsampled estimates. For both sites and parameters, the overall shape of the time series of all-data estimates typically shows more inflection points and undulations over time compared to the time series of downsampled estimates. In most cases, the time series of ds-12 estimates roughly match the variations shown in the time series of all-data estimates (fig. 8). The overall agreement between the time series appears to diminish as the calibration data are downsampled more intensely. For some site-parameter combinations, the most variability between the various downsampled estimates occurs at the beginning or end of the maximum trend period. For example, at site mdWIL0013 near the beginning of the record, downsampled estimates of annual flow-normalized nitrate load range from being nearly the same as the all-data estimate to over 100,000 kilograms per year (kg/yr) higher (for ds-4), whereas estimates later in the record vary by about half this range or less (fig. 8*C*). However, for other site-parameter combinations, large differences between the all-data and downsampled flow-normalized estimates occur throughout the record. For example, at mdWIL0013, total phosphorus load varies by about 20,000 kg/yr between downsampled estimates across most of the record (fig. 8*G*).

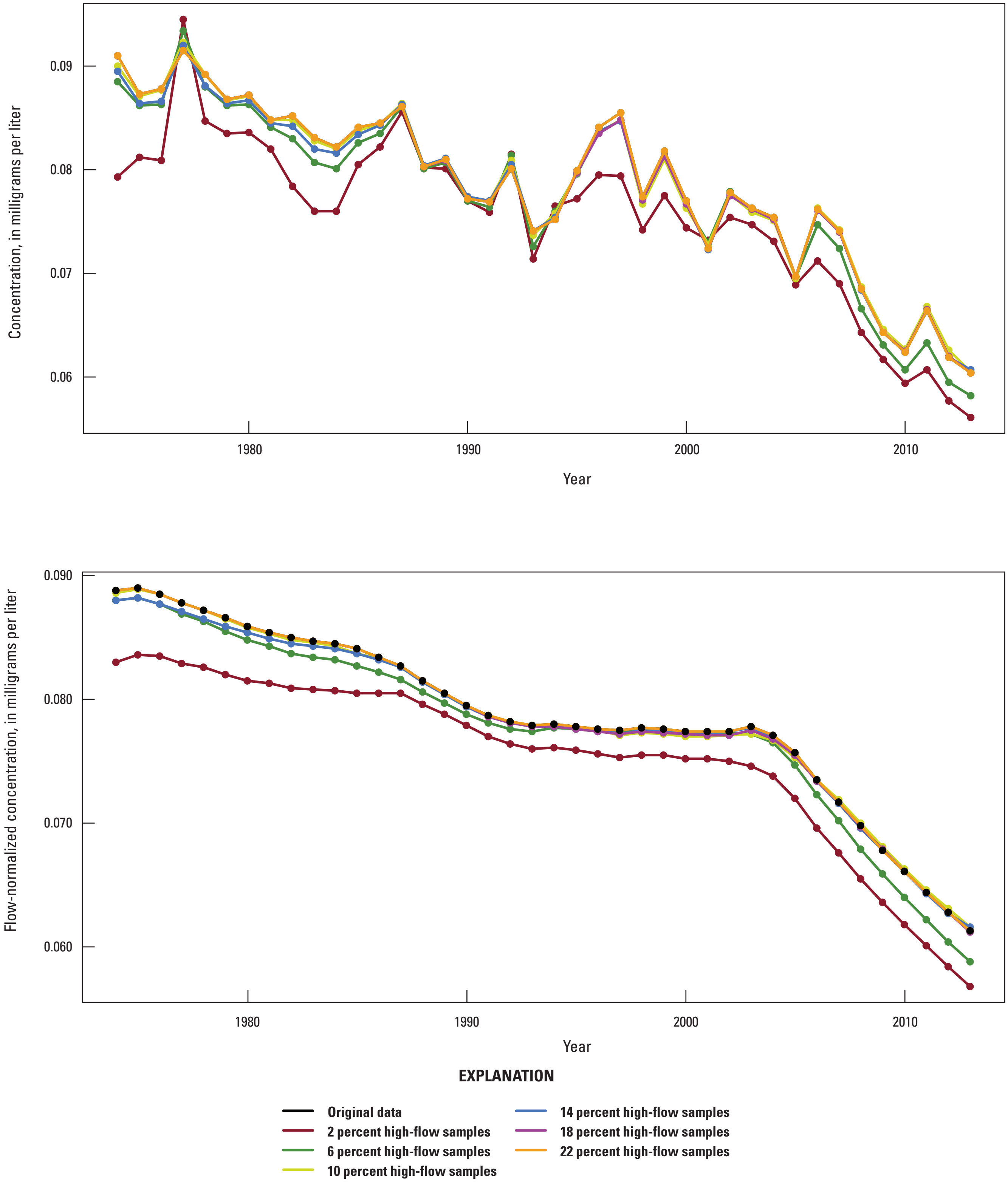


Figure 5. Annual concentration, load, flow-normalized concentration, and flow-normalized load of total phosphorus at site 14211720 Willamette River at Portland, Oregon, showing the original data and data that has been downsampled to contain 2, 6, 10, 14, 18, and 22 percent high-flow samples per trend decade.

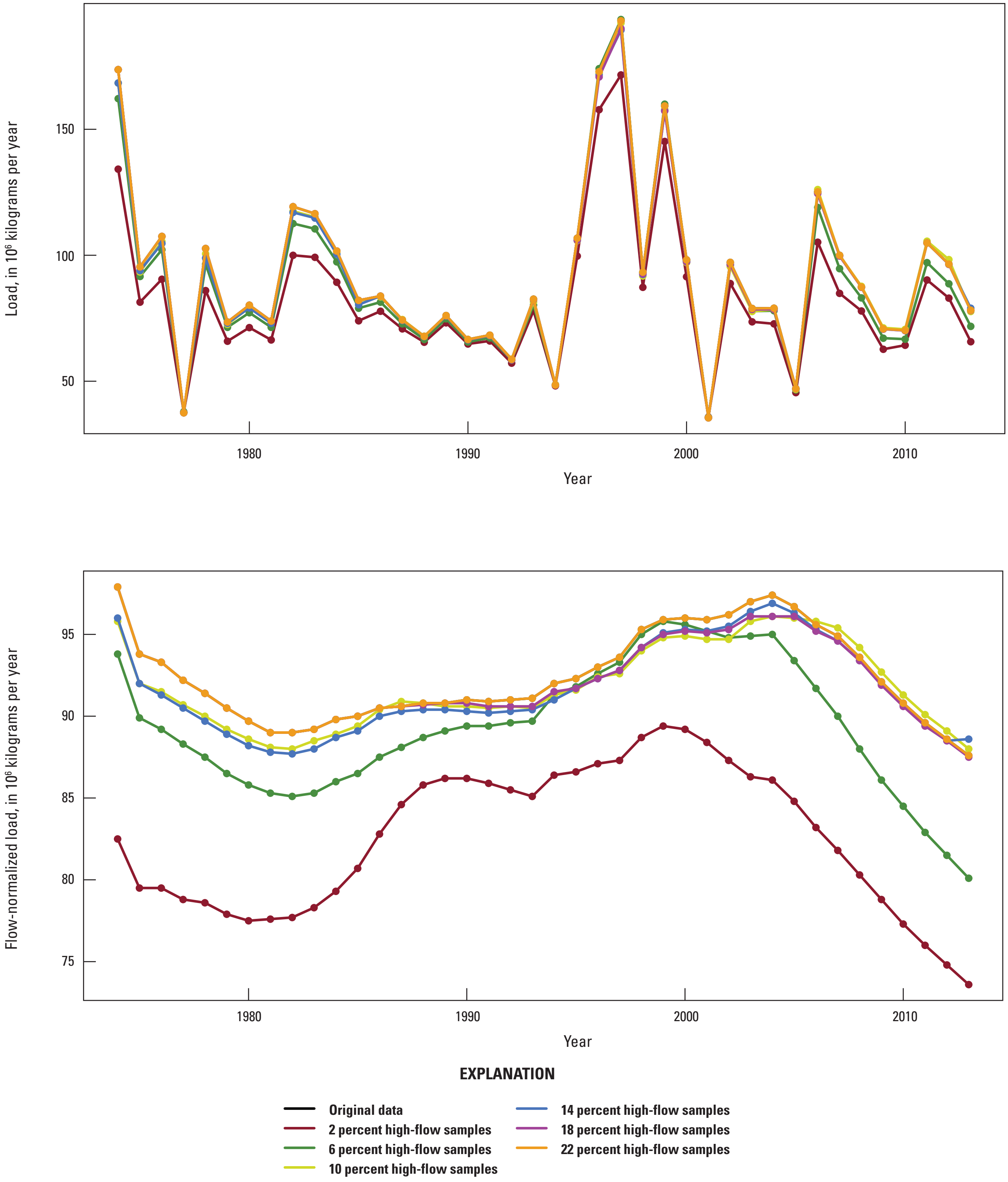


Figure 5.—Continued Annual concentration, load, flow-normalized concentration, and flow-normalized load of total phosphorus at site 14211720 Willamette River at Portland, Oregon, showing the original data and data that has been downsampled to contain 2, 6, 10, 14, 18, and 22 percent high-flow samples per trend decade.

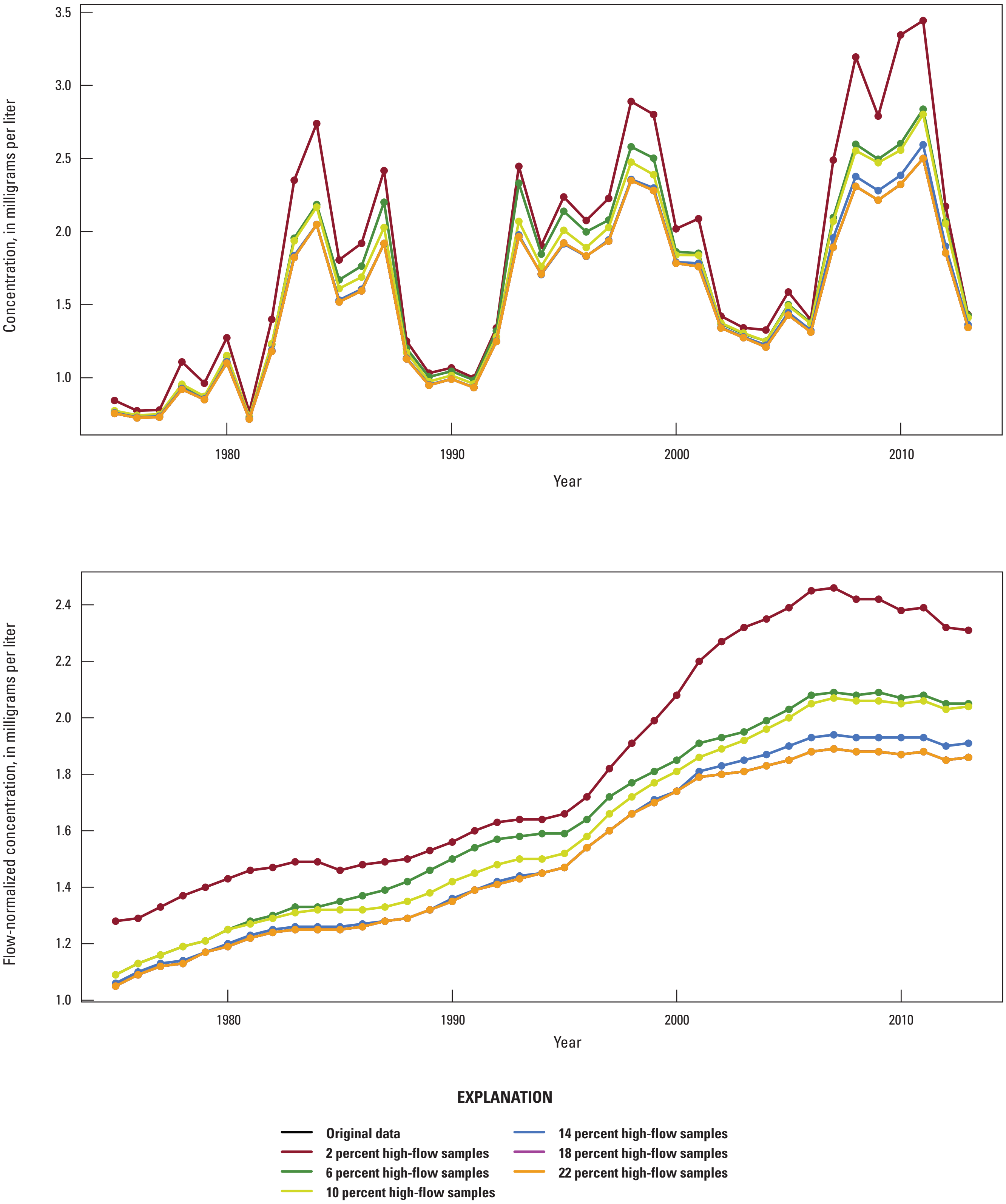


Figure 6. Annual concentration, load, flow-normalized concentration, and flow-normalized load of nitrate at site 06805500 Platte River near Louisville, Nebraska, showing the original data and data that has been downsampled to contain 2, 6, 10, 14, 18, and 22 percent high-flow samples per trend decade.

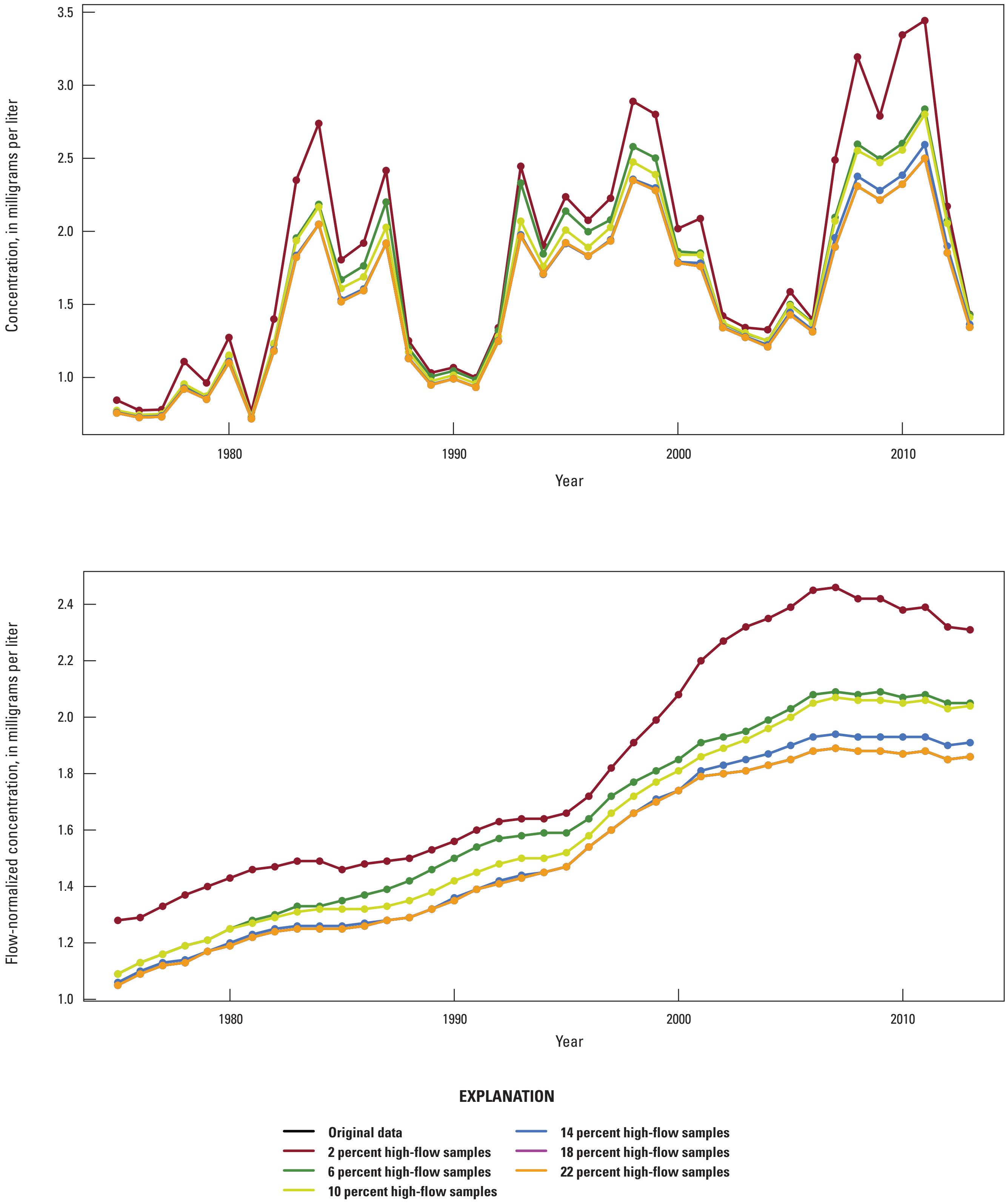


Figure 6.—Continued Annual concentration, load, flow-normalized concentration, and flow-normalized load of nitrate at site 06805500 Platte River near Louisville, Nebraska, showing the original data and data that has been downsampled to contain 2, 6, 10, 14, 18, and 22 percent high-flow samples per trend decade.

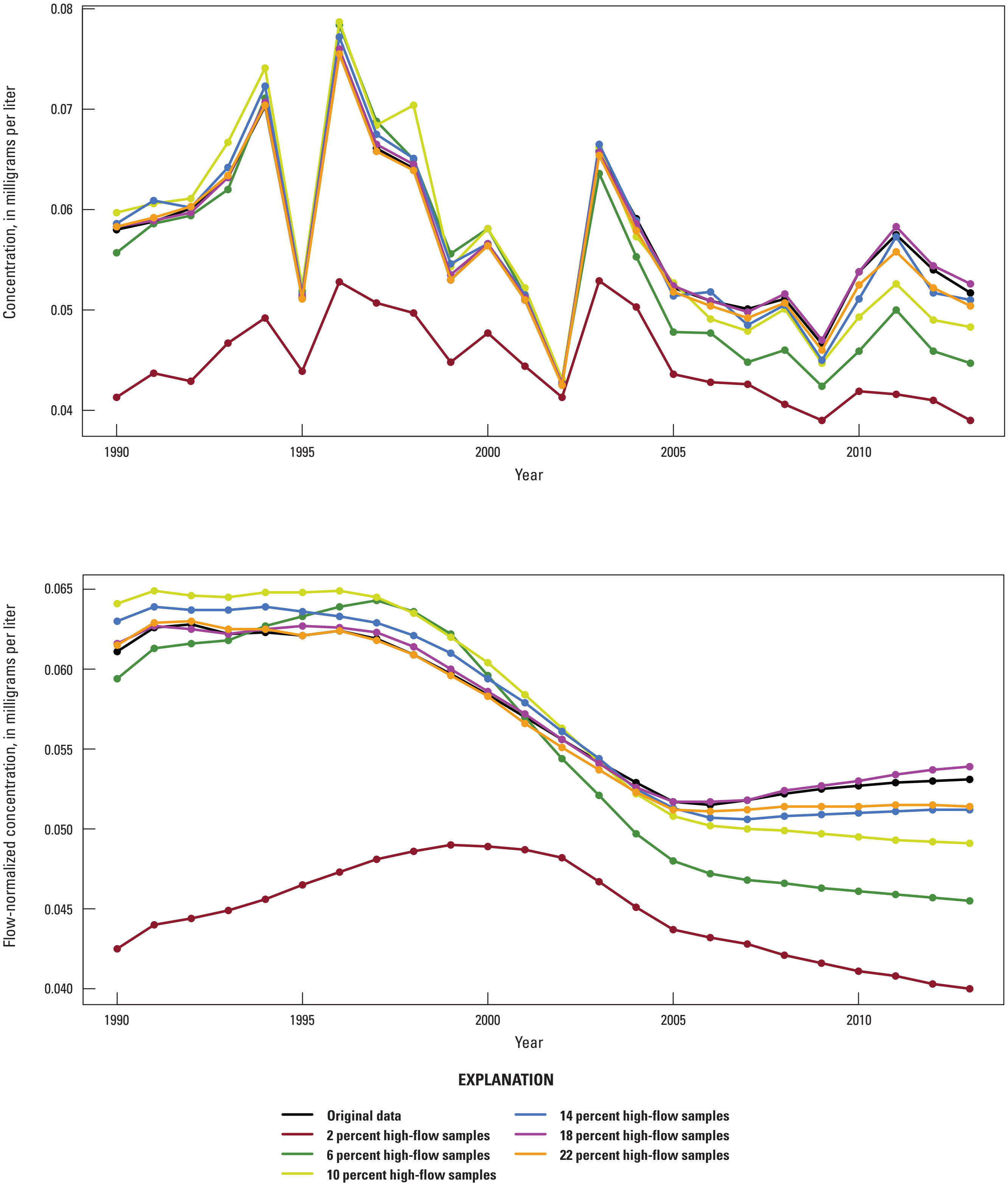


Figure 7. Annual concentration, flow-normalized concentration, load, and flow-normalized load of total phosphorus at site 01654000 Accotink Creek near Annandale, Virginia, showing the original data and data that has been downsampled to contain 2, 6, 10, 14, 18, and 22 percent high-flow samples per trend decade.

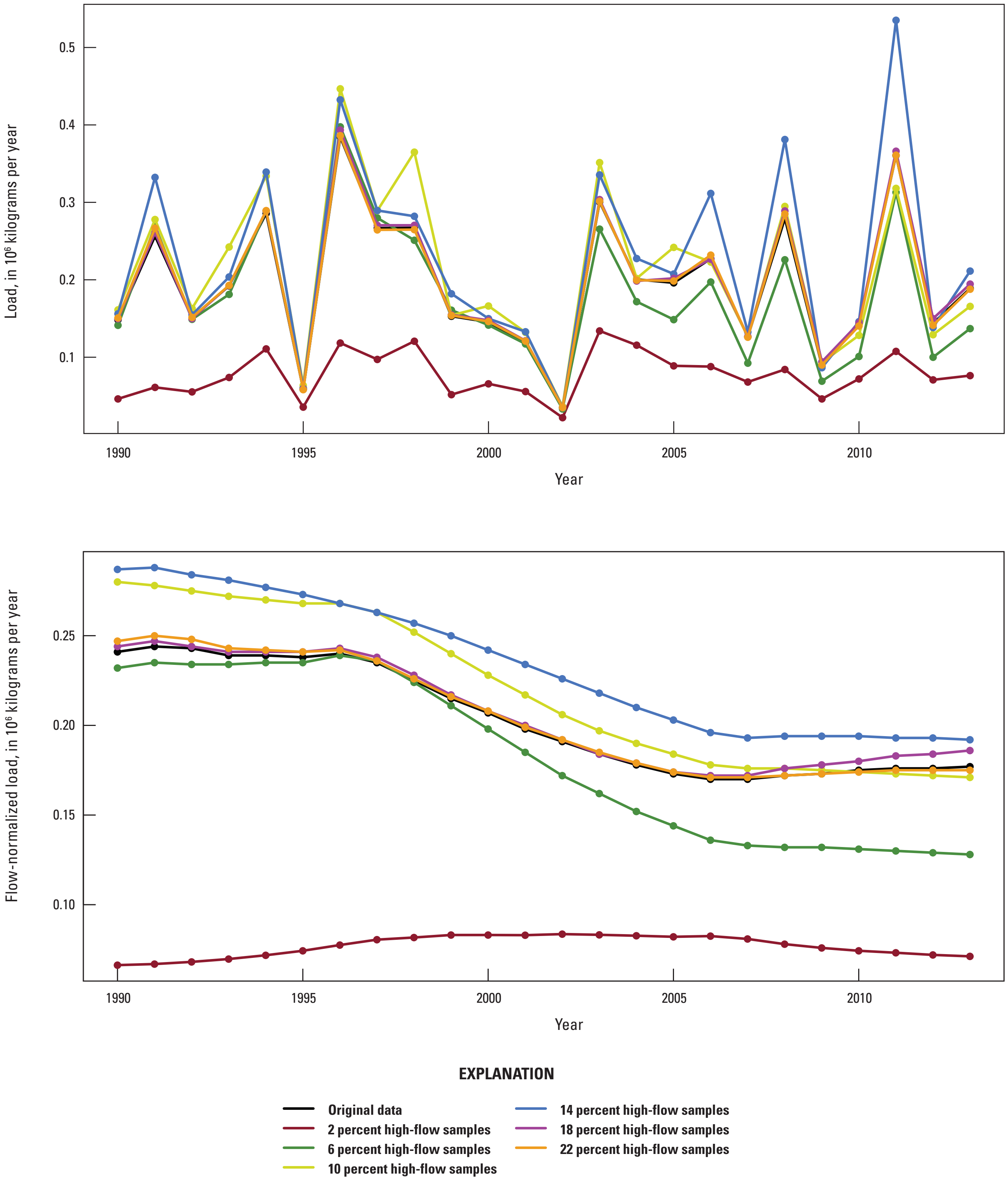


Figure 7.—Continued Annual concentration, flow-normalized concentration, load, and flow-normalized load of total phosphorus at site 01654000 Accotink Creek near Annandale, Virginia, showing the original data and data that has been downsampled to contain 2, 6, 10, 14, 18, and 22 percent high-flow samples per trend decade.

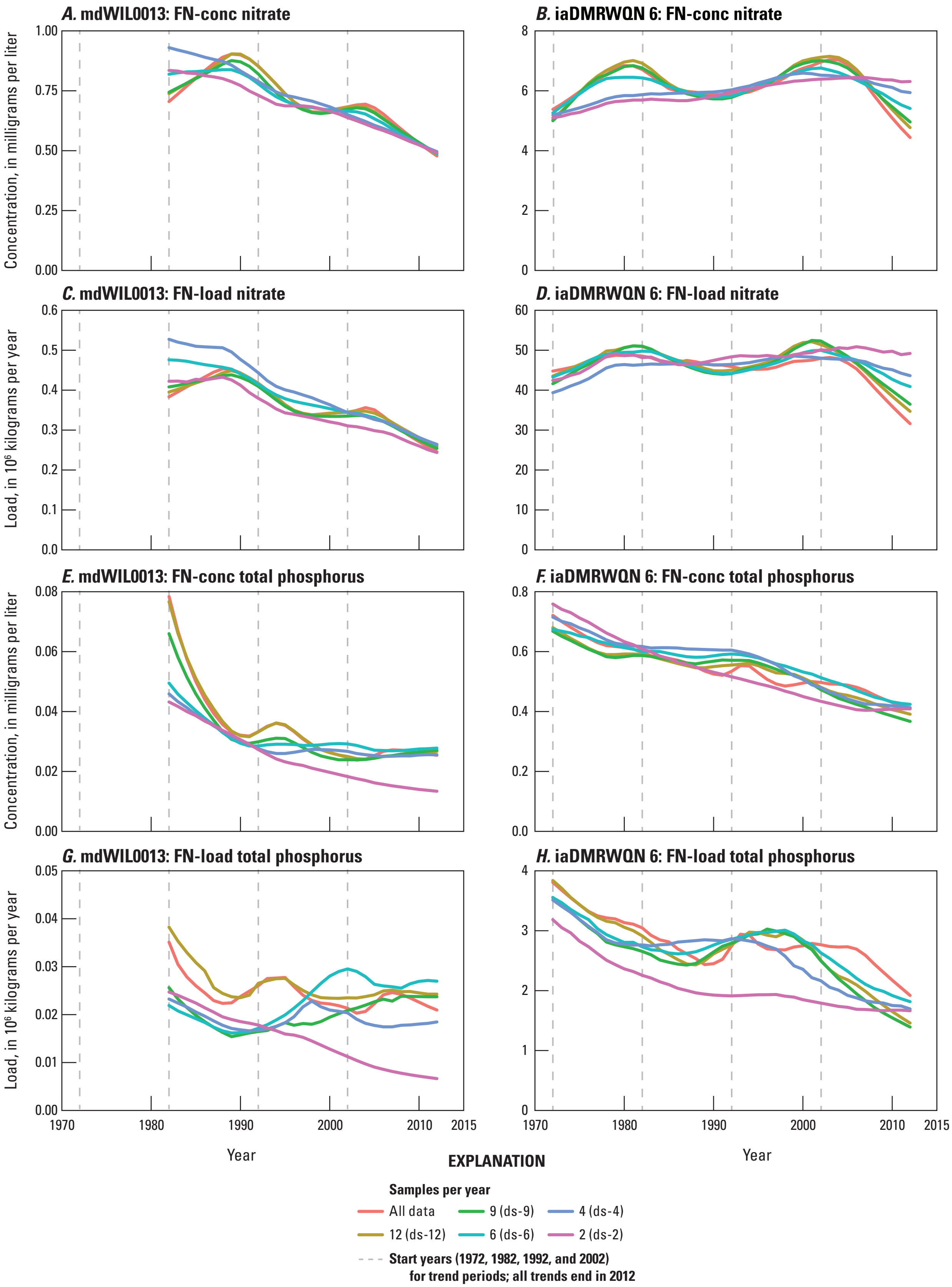


Figure 8. Time series of annual flow-normalized (FN) estimates of nitrate and total phosphorus concentration and load (FN-conc and FN-load) from models calibrated using the complete dataset ("All-data") and each downsampled (ds) dataset containing 12, 9, 6, 4, or 2 samples per year for sites iaDMRWQN 6 and mdWIL0013.

An important consideration for this study was to determine how trend estimates varied with different sample frequencies in the calibration datasets. To explore this, trends determined using the downsampled estimates were compared to trends determined using all-data estimates. Trends, described as the amount or percentage change in annual flow-normalized concentration or flow-normalized load between any 2 years, were calculated for the endpoints of all possible trend periods used in this study for these two parameters and two sites. Errors in the downsampled trends are shown in figure 9 for each site and parameter by trend period and estimate type (concentration or load). Errors were determined by calculating the percentage difference in the downsampled trend compared to the all-data trend (shown on the y-axis with a horizontal line at zero for reference). Errors were larger for downsampled load trends compared to downsampled concentration trends across all trend periods, sites, and parameters. Also, total phosphorus trends at site mdWIL0013 were prone to greater errors across most sampling frequencies compared to site iaDMRWQN 6 and to nitrate. These greater errors may be due to the much larger percentage of censored data at this site. About 30 percent of the total phosphorus concentrations were censored for site mdWIL0013, whereas less than 1 percent of the values were censored for the other sites and parameters.

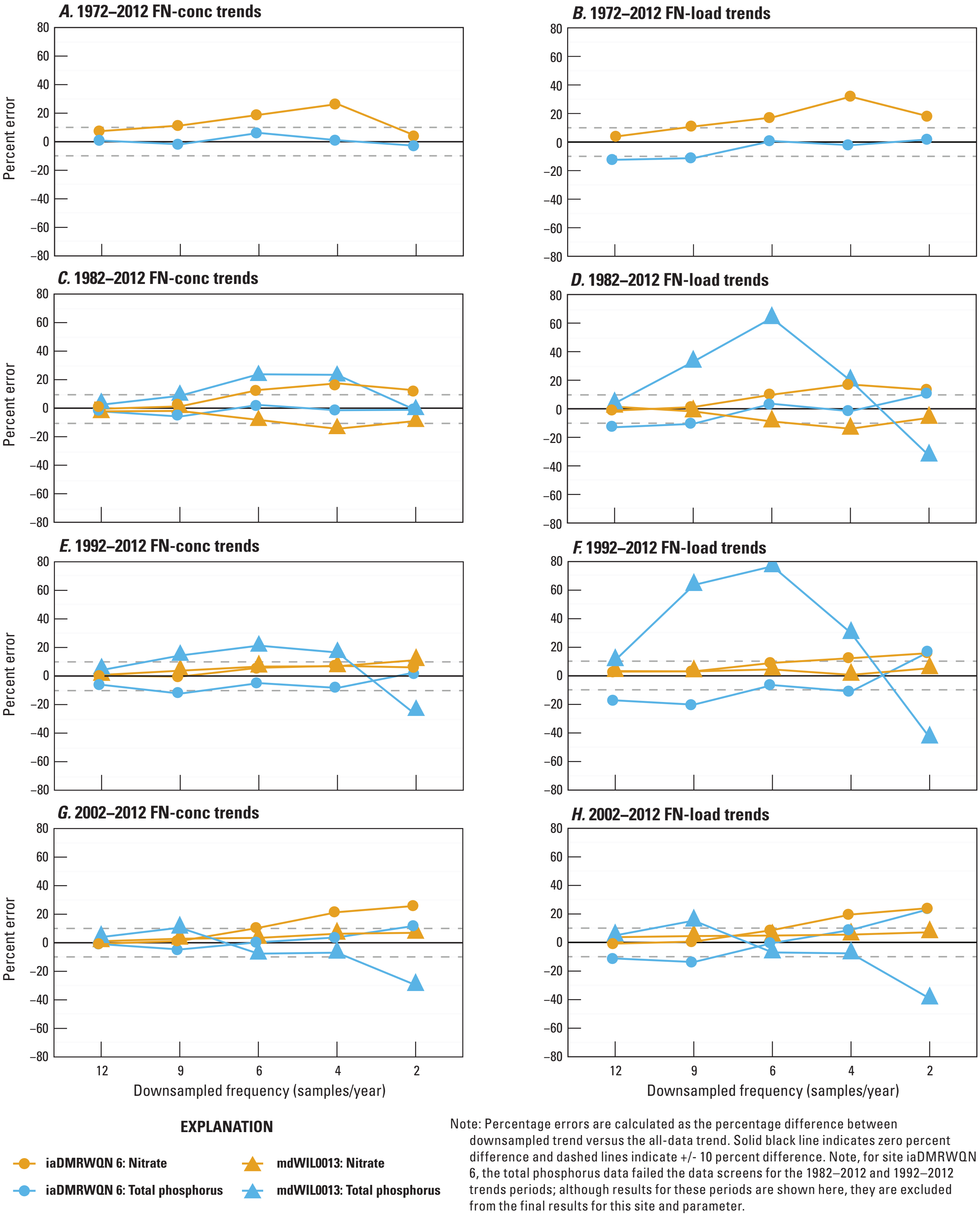
Figure 9 shows that many downsampled trends are within +/- 10 percent (dashed horizontal lines) of the all-data trend. Generally increasing errors with decreasing sample frequency is evident in the 2002–12 downsampled trends for both sites. Across the other trend periods, errors become increasingly or decreasingly large as sample frequency decreases (as opposed to remaining relatively constant), but this pattern is not necessarily monotonic. Importantly, the ds-2 trends do not always have the most error. Depending on the site, parameter, and trend period, the ds-4, ds-6, and ds-9 trends can have more error than the ds-2 trend. When the most accurate and least accurate downsampled trends were selected for each combination of site, parameter, estimate type (concentration or load), and trend period (28 trend scenarios total), as expected, the ds-12 trends were consistently the most accurate (table 12). Out of 28 trend scenarios, over half were best estimated with the ds-12 trend. The ds-2 trend was least accurate in more than 40 percent of the trend scenarios, consistently exhibiting greater error compared to the other downsampled trend estimates.

Table 12. Comparison of accuracy in trend estimates between downsampled datasets, across all 28 scenarios.

Sample frequency	Number of times downsampled trend was most accurate	Number of times downsampled trend was least accurate
ds-2	4	12
ds-4	2	7
ds-6	4	4
ds-9	3	3
ds-12	15	2

In addition to trend magnitude, the direction of the trend is an important component of trend analysis; thus, the direction of each downsampled trend was compared to the direction of the corresponding all-data trend. For this comparison, the 80 or 60 individual downsampled trends for each site were considered; this includes downsampled trends for 2 parameters, 2 estimate types (concentration or load), 5 different downsampled datasets (2, 4, 6, 9, and 12 samples), and 3 or 4 trend periods, depending on the site. For site iaDMRWQN 6, when these 80 downsampled trends were compared to their corresponding all-data trend, only 2 downsampled concentration trends and 2 downsampled load trends differed in trend direction. Specially, at site iaDMRWQN 6, the ds-2 and ds-4 nitrate concentration and load trends for 1972–2012 were positive when the all-data trend for this period was negative (fig. 8*B* and 8*D*, compare ending position of pink, blue, and red lines with their beginning position). For site mdWIL0013, the majority of the 60 downsampled concentration or load trends also had the same direction as the corresponding all-data trend, though 8 downsampled trends differed in direction. These eight conflicting downsampled trends all occur for total phosphorus concentration or load, mainly during the later two trend periods. Specifically, for the 2002–12 trend period, the ds-2, ds-4, and ds-6 trends for total phosphorus concentration were negative when the all-data concentration trend was slightly positive (fig. 8*E*), and the ds-9 and ds-12 trends for total phosphorus load were positive when the all-data load trend was slightly negative (fig. 8*G*). Also, the 1992–2012 trend period for the same site and consistuent had positive ds-4, ds-6, and ds-9 load trends when the all-data load trend was negative (fig. 8*G*). Note however, that the 1992–2012 trend is not reported for this site and parameter in the results of this study because the data failed to meet the data screens for this period. In total, about 8 percent of the downsampled trends had a different direction than the corresponding all-data trend, the majority of which occurred for total phosphorus trends at the highly censored mdWIL0013 site (accounting for 6 percent of all cases). For these two sites and parameters, the directions of the downsampled trends were generally consistent with the all-data trend, and this appears to be quite consistent across sample frequency. Differences in the trend direction also appear to be most sensitive for datasets that have considerable censoring. Bootstrap analysis was not performed for these downsampled trends, so it is possible that some or all of the eight conflicting downsampled trends were not significant.

While the errors associated with downsampled trends can be large, at times much greater than 10 percent depending on the site and parameter (fig. 9), the directions of the downsampled trends appear to typically agree with the direction of the all-data trend. This indicates that the decision to allow a minimum of four samples per year for 70 percent of the record (though four samples per year for the first 2 and last 2 years in the record will be required) may lead to trend estimates with some amount of unknown error in magnitude, but the estimates are likely to be correct in direction. Also,



Note: Percentage errors are calculated as the percentage difference between downsampled trend versus the all-data trend. Solid black line indicates zero percent difference and dashed lines indicate +/- 10 percent difference. Note, for site iaDMRWQN 6, the total phosphorus data failed the data screens for the 1982–2012 and 1992–2012 trends periods; although results for these periods are shown here, they are excluded from the final results for this site and parameter.

Figure 9. Percentage errors in flow-normalized concentration (FN-conc) and flow-normalized load (FN-load) trends by sampling frequency, site, and parameter.

the ds-2 trends exhibited the most error (table 12) across the various trend periods, indicating that the decision to limit the sample frequency to a minimum of four samples per year will help to avoid more consistent error associated with lower sample frequencies. This sensitivity analysis only explored the influence of sample frequency on WRTDS trend estimates at two sites for two parameters, thus making it impossible to draw a firm conclusion on overall errors anticipated from different sample frequencies. To more fully explore the influence of sample frequency on trend estimates and to develop a better characterization of the anticipated errors, a more robust analysis involving the use of a large number of densely sampled sites or simulated data is required.

To ensure the representativeness of the calibration datasets, this study required that the four samples per year roughly occur in each of the four seasons of the year (see “Data Processing” sections), and that a certain number of samples were collected at or above a high-flow threshold (see “Storm Sampling” section). These additional data requirements help to ensure that the calibration dataset for each site captures the range of seasonal and hydrologic variability that occurs; furthermore, additional evaluations were completed during the review of the WRTDS model residuals as described in the “Nutrient, Sediment, Major Ion, Salinity, and Carbon Trend Analysis Method” section, which included visually assessing the data for seasonal coverage and comparing the range of streamflow conditions captured by the calibration dataset to the range of streamflow conditions across the entire period of record.

Rounding of Water-Quality Data

The precision of the water-quality data used in this study varied by site, parameter, and analyzing agency because of the analytical methods used for sample determination, laboratory-specific detection limits, and reporting conventions. Additionally, the precision of a specific parameter, at a single site, analyzed by a single agency may vary over time because of upgrades in analytical equipment, updates to detection limits, and changes in reporting conventions over time.

To illustrate the influence of rounding on WRTDS estimates, sensitivity tests that applied various types of rounding were completed at two sites. At site 14211720 Willamette River at Portland, Oregon, total phosphorus data, which were originally reported to a precision greater than 0.001 (fig. 10*A*), were rounded to 3, 2, and 1 decimal place (figs. 10*B*, 10*C*, and 10*D*); each of these rounded datasets and the unrounded dataset were used to calibrate WRTDS. The annual flow-normalized concentration and flow-normalized load estimates derived from datasets rounded to 3 and 2 decimal places are indistinguishable from the estimates derived from the unrounded dataset (figs. 10*E* and 10*F*). The estimates derived from the dataset rounded to only 1 decimal place are biased by about 0.025 mg/L for flow-normalized concentration estimates and 50,000 kg/yr for flow-normalized load estimates compared with estimates derived using the unrounded dataset (figs. 10*E* and 10*F*).

As a second example, filtered orthophosphate data at site 06805500 Platte River at Louisville, Nebraska, originally contained different rounding precisions for different parts of the record. For example, from water year 1982 through mid-1996, data were reported to 2 decimal places, and from mid-1996 though the end of the record, data were reported to 3 decimal places, except for half of the year in 2007, when data were reported to 5 decimal places. The WRTDS estimates were determined using the original data at this site, and the results were compared to three scenarios: (1) all data rounded to 2 decimal places, (2) data for 1982–96 rounded to 2 decimal places and data for 1997–2012 rounded to 3 decimal places, and (3) data for 1982–90 rounded to 1 decimal place, data for 1991–2000 rounded to 2 decimal places, and data for 2001–12 rounded to 3 decimal places, simulating upgrades to analytical equipment with increased precision (figs. 11*A*, 11*B*, 11*C*, and 11*D*). The flow-normalized concentration and flow-normalized load estimates from these three scenarios are nearly identical when compared to the flow-normalized estimates derived from the unrounded data (figs. 11*E* and 11*F*). There are small departures in the estimates derived from the rounded datasets, but the overall shape of the time series remains consistent. Based on the results from these two sites, when rounding appears to minimally affect WRTDS estimates, no attempt was made to change the precision of the original data.

Influence of New Data on Trend Estimates

The WRTDS model and all regression-based water-quality models are sensitive to the incremental incorporation of new data. Estimates of concentration and load are typically updated periodically (for example, annually or semidecadally) as new water-quality data become available. Recalibration of regression-based models with updated data can lead to changes in the model coefficients, historical estimates, and model diagnostics. With each addition of new data and successive recalibration of the model, previous estimates of concentration and load for the same year can change, affecting the trend estimated between two past points in time. For example, total nitrogen data from 1995 through 2005 can be used to calibrate WRTDS, estimate annual concentrations for each year in the record, and calculate a trend between these start and end years using the annual estimates. When new data are added in 2006 and the model is recalibrated, annual total nitrogen estimates are available through 2006. Because additional calibration data have been used in the regression model, the total nitrogen estimates for 1995 and 2005 may have changed to some extent, possibly causing the 1995–2005 trend to also change. Changes in estimates between successive model calibrations are particularly evident during the last few years of the calibration period when the influence of new observations is greatest because the time weight is greatest; thus, these changes to historical estimates as new data are added to the calibration dataset are referred to as “edge effects” in this study. Estimates in the last few years of the calibration period are more vulnerable to edge effects because of the window-weighting

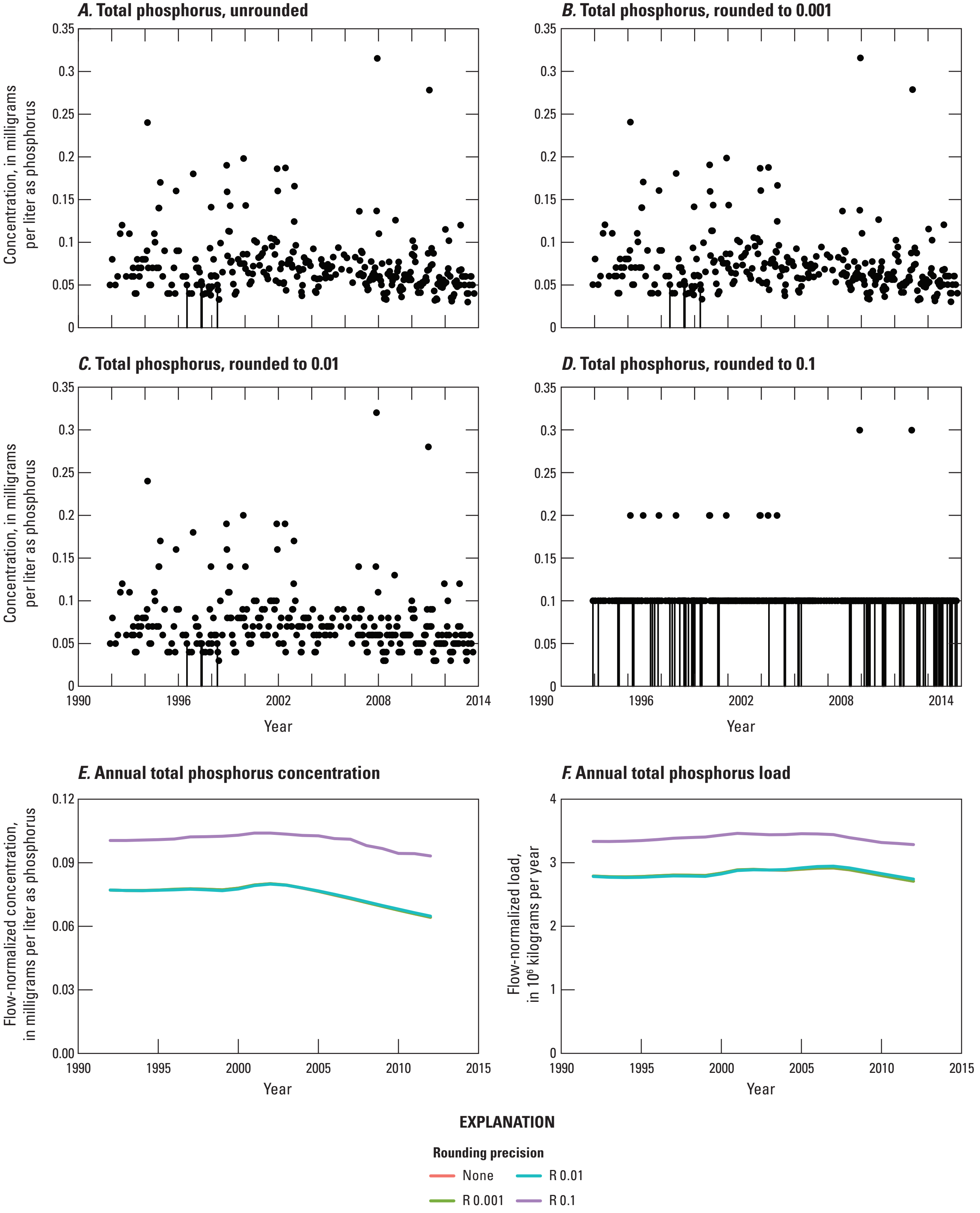


Figure 10. Total phosphorus data at site 14211720 Willamette River at Portland, Oregon, *A*, unrounded (None); and three rounding schemes including, *B*, all data rounded to 3 decimal places (R 0.001); *C*, all data rounded to 2 decimal places (R 0.01); *D*, all data rounded to 1 decimal place (R 0.1); plus *E*, annual flow-normalized concentration; and *F*, flow-normalized load estimates from WRTDS when using the unrounded and variously rounded datasets.

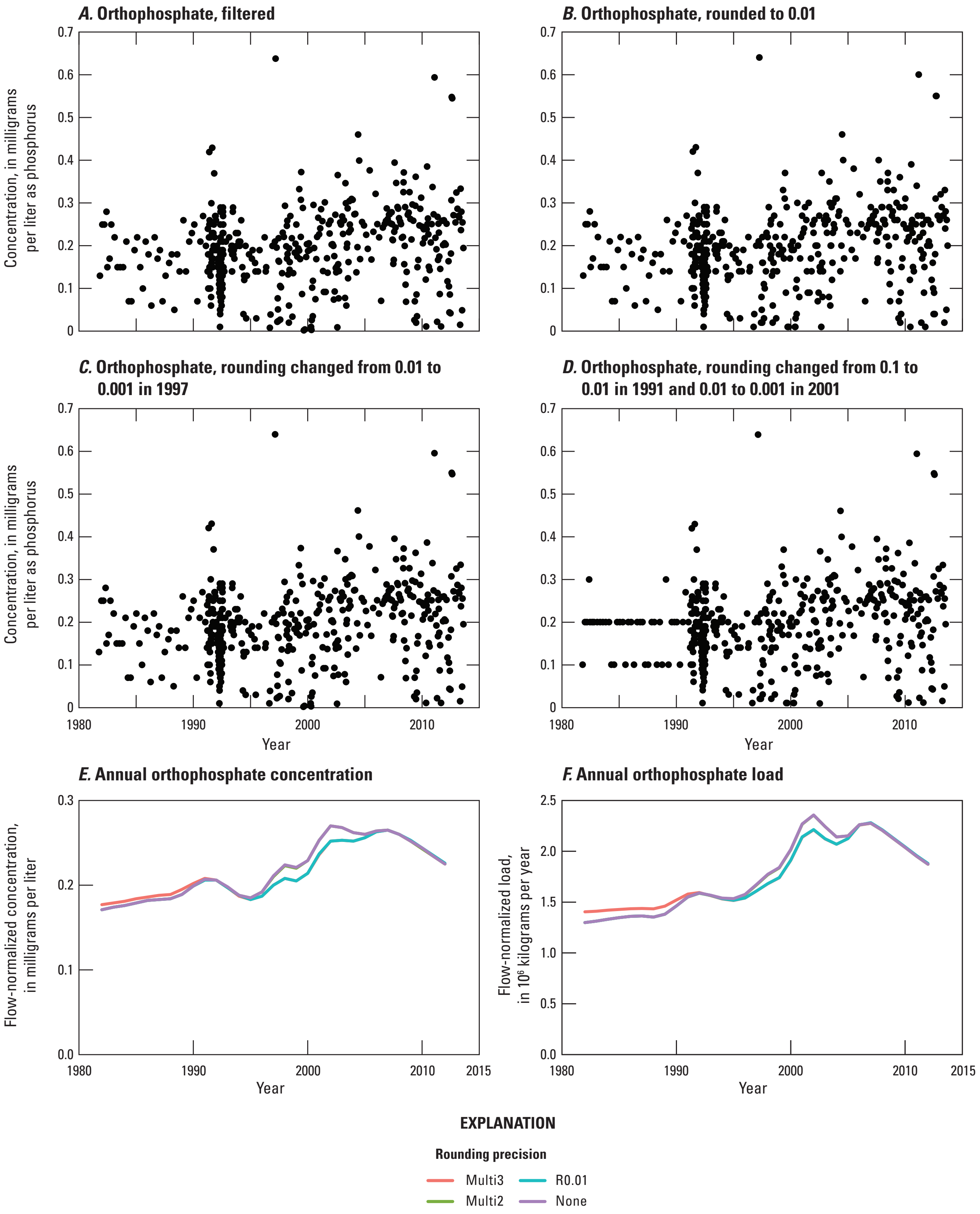


Figure 11. Filtered orthophosphate data at site 06805500 Platte River at Louisville, Nebraska, *A*, unrounded (None); *B*, all data rounded to 2 decimal places (R0.01); *C*, data for 1982–96 rounded to 2 decimal places and data for 1997–2012 rounded to 3 decimal places (Multi2); *D*, data for 1982–90 rounded to 1 decimal place and data for 1991–2000 rounded to 2 decimal places, data for 2001–12 rounded to 3 decimal places (Multi3); plus *E*, annual flow-normalized concentration; and *F*, flow-normalized load estimates for all three rounding scenarios (R0.01 from B; Multi2 from C; and Multi3 from D) and the unrounded data (None from A).

approach used for calibration in WRTDS—when the half window width for time is set to the default of 7 years, only estimates within 7 years of the end of the calibration record may change with the addition of new data; however, if the default window of 7 years is expanded to acquire more samples in the window, estimates older than 7 years also may be affected.

Edge effects can be explored using successive calibrations, which are sets of model estimates (vectors of annual estimates of concentration and load) where new data have been incrementally added to the end of the record and models are recalibrated. For example, an initial set of annual estimates can be generated for the period from 1994 to 2000. Successive calibrations are executed by incorporation of new data (1994–2001, 1994–2002, 1994–2003, and so on). Chanat and others (2016) explored the influence of incremental incorporation of new data on annual flow-normalized water-quality estimates in the Chesapeake Bay. When they completed successive calibrations for flow-normalized concentration of two parameters at several sites, they observed that in the beginning of the record there was a small amount of translation of the successive calibrations up and down the y-axis but little difference in the overall shape; these translations primarily are due to slight changes in the probability distribution of streamflow as more years of streamflow are included in the streamflow normalization process, which can influence flow-normalized estimates across the entire record. Chanat and others (2016) also observed that near the end of the record, estimates of flow-normalized concentration varied widely between successive calibrations.

Chanat and others (2016) also compared edge effects of model results from WRTDS to results from ESTIMATOR, a frequently used 7-parameter regression-based trends model equation originally described by Cohn and others (1992), using flow-normalized (for WRTDS) or flow-adjusted (for ESTIMATOR) concentrations of dissolved inorganic nitrogen and total phosphorus. For ESTIMATOR, Chanat and others (2016) found flow-adjusted estimates in the beginning of the record could change dramatically with the addition of new data at the end of the record. A result of the window-weighting approach used during model calibration in WRTDS is that historical estimates remain relatively constant as more data are added to the end of the record. Historical estimates from regression-based models like ESTIMATOR can vary substantially across the period of record (not just at the end) as new data are added because all data are weighted the same. Importantly, the sensitivity of model coefficients and estimates to the incremental incorporation of new data are not issues unique to WRTDS but are common among regression-based approaches.

Chanat and others (2016) used an earlier version of WRTDS, which did not include the edgeAdjust feature that has been incorporated into newer versions of WRTDS (version 2.2.0 of EGRET, used in this study) to rectify some of the previously described edge effects. In previous versions of WRTDS, if the half window width for time was 7 years, then the model calibrated for observations in the final year of the water-quality record would use observations only from the previous 7 years (because observations 8 years and older have zero weights for time). This small window at the end (or beginning) of the record allowed for increased flexibility and detection of inflection points but could also lead to highly variable estimates of concentration and load with each successive addition of new water-quality data. In version 2.2.0 of WRTDS, when edgeAdjust is set to TRUE, the half window width expands near the beginning (and end) of the record, so that the model will always use at least 14 years of data in each calibration (Hirsch and De Cicco, 2015). Use of this feature forces trends to be more linear near the start and end of the record and lowers the possibility of changes in slope with the addition of new data; however, it also decreases the sensitivity of the model to detect changes around an emerging inflection point. For a detailed explanation of edgeAdjust, see Hirsch and De Cicco (2015).

While the addition of the edgeAdjust feature helps minimize changes in historical estimates as new data are added to the calibration record, it does not completely rectify the issue. To illustrate the influence of the incremental incorporation of new data on WRTDS estimates with the edgeAdjust feature in place, successive calibrations were completed for total phosphorus and nitrate at sites 01646580 Potomac River at Chain Bridge, at Washington, D.C., and 02215500 Ocmulgee River at Lumber City, Georgia. The calibration periods for these sites were water years 1982–2012 and 1972–2012, respectively. To demonstrate the effects of adding new data each year, the initial calibration for each site-parameter combination was run using the period from 1982 or 1972, depending on the site, to 1992; 20 additional calibrations were then run with the start year always set at 1982 or 1972 and the end year ranging from 1992 to 2012. The longest calibration record used, either 1982–2012 or 1972–2012, depending on the site, is hereafter referred to as the "whole-record" calibration. Annual concentration, load, flow-normalized concentration, and flow-normalized load estimates for both sites and parameters and all successive calibrations are shown in figures 12–15.

For all site-parameter combinations, the concentration and load estimates for the first 10 or 20 years (depending on the site) of the record are identical or nearly identical between successive calibrations (figs. 12*A*,*B*–15*A*,*B*). This is due to the window-weighting approach used to estimate the regression models in WRTDS, which stabilizes estimates that are 7 or more years from the end of the calibration record. The slight variations in estimates between successive calibrations in the first 10 or 20 years of the record likely are due to the window widths being iteratively expanded beyond 7 years because of model requirements (WRTDS requires a minimum number of samples and will automatically relax user set window widths if needed to fulfill that minimum). Estimates in the later part of the record show more variability between calibrations. This variability primarily occurs within the last 7 years of the end of a successive calibration (for example, annual estimates between 1991 and 1997 for the 1982–97 successive calibration).

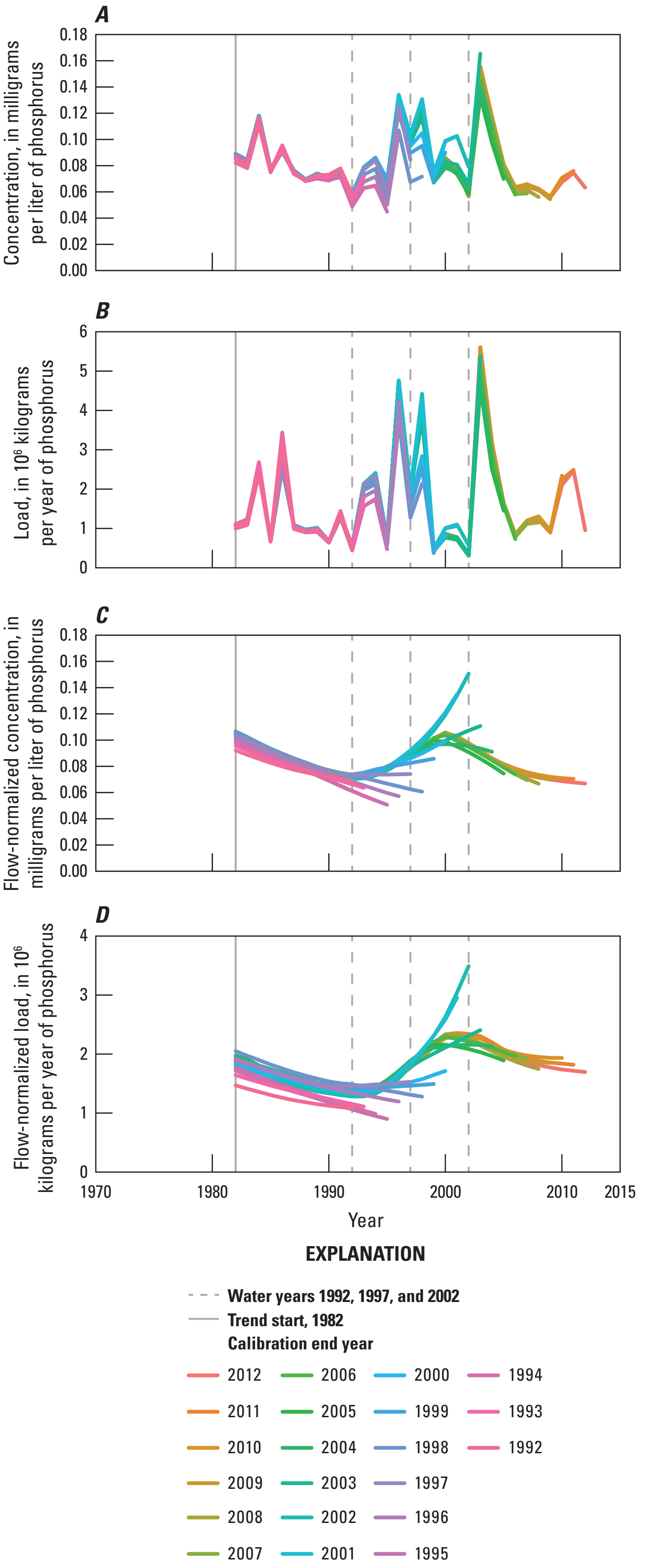


Figure 12. Annual time series of concentration, load, flow-normalized concentration, and flow-normalized load estimates for total phosphorus at site 01646580 Potomac River at Chain Bridge, at Washington, D.C. Initial calibration period from 1982 to 1992 with successive calibrations extending through 2012.

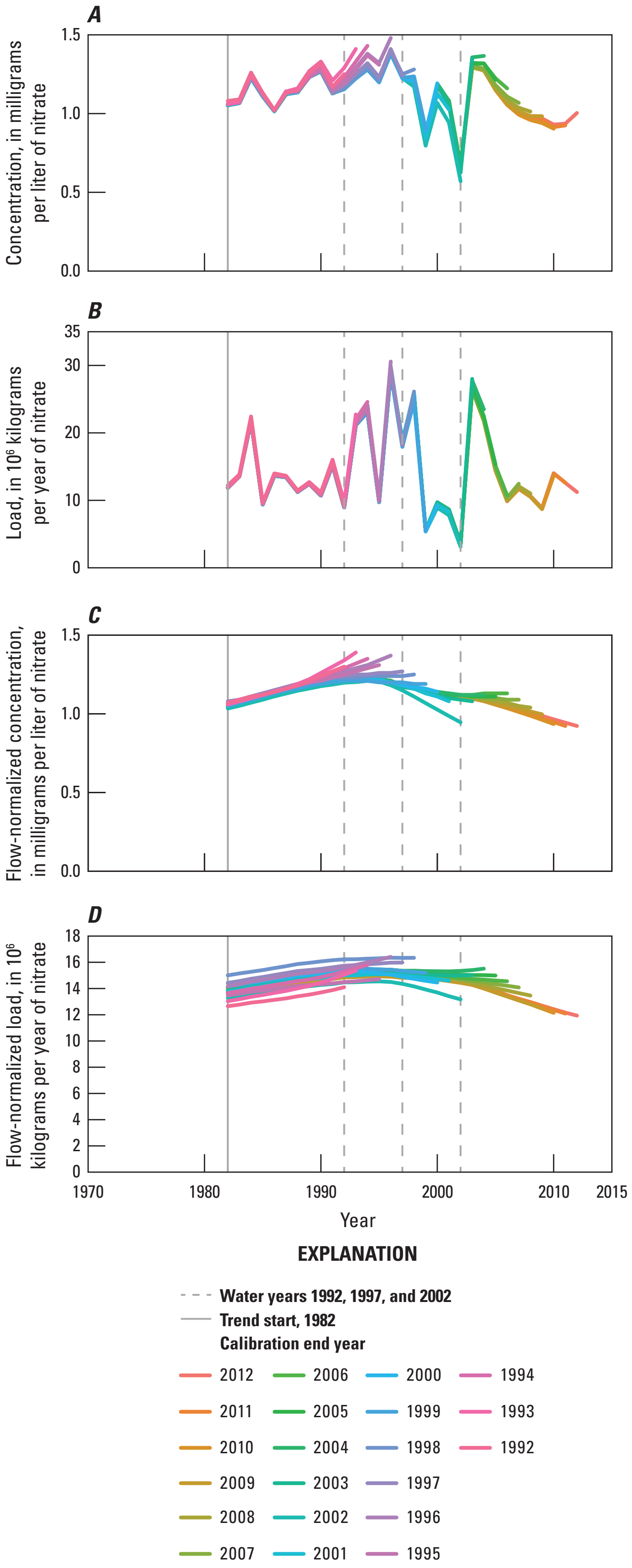


Figure 13. Annual time series of concentration, load, flow-normalized concentration and flow-normalized load estimates for nitrate at site 01646580 Potomac River at Chain Bridge, at Washington, D.C. Initial calibration period from 1982 to 1992 with successive calibrations extending through 2012.

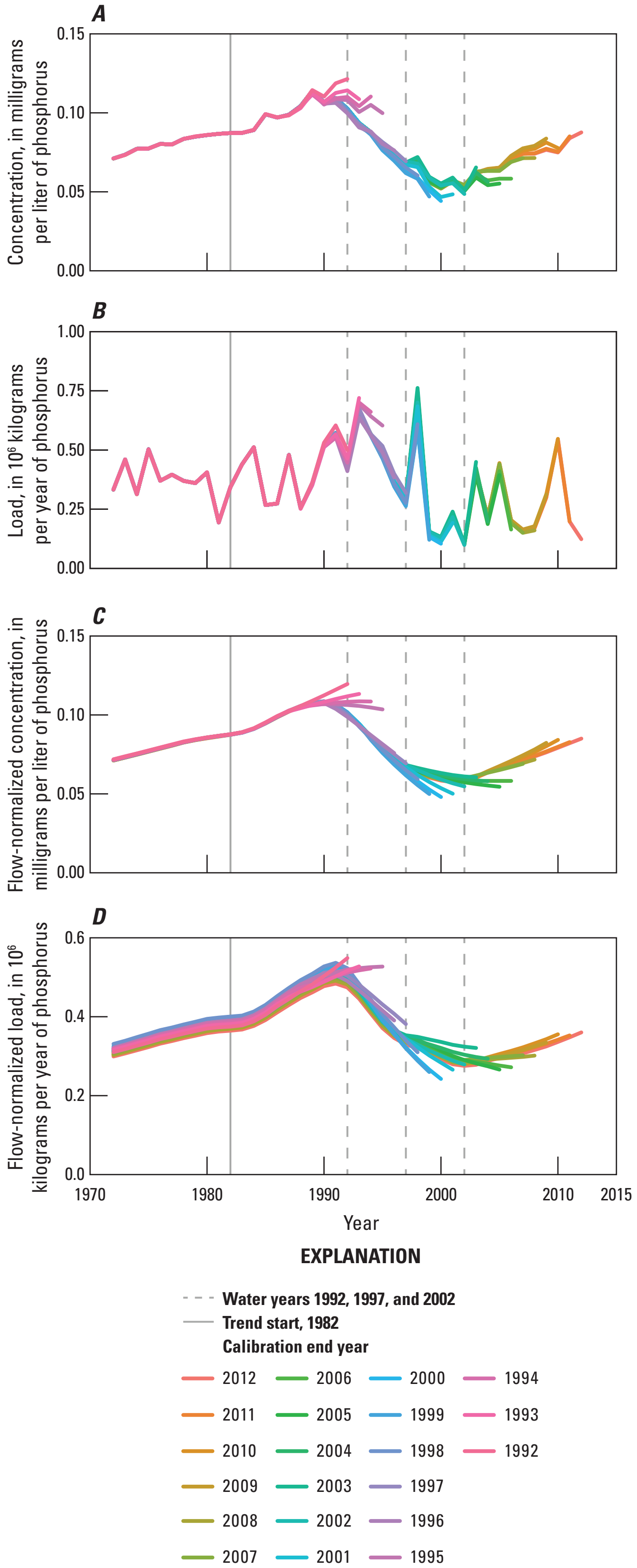


Figure 14. Annual time series of concentration, load, flow-normalized concentration and flow-normalized load estimates for total phosphorus at site 02215500 Ocmulgee River at Lumber City, Georgia. Initial calibration period from 1972 to 1992 with successive calibrations extending through 2012.

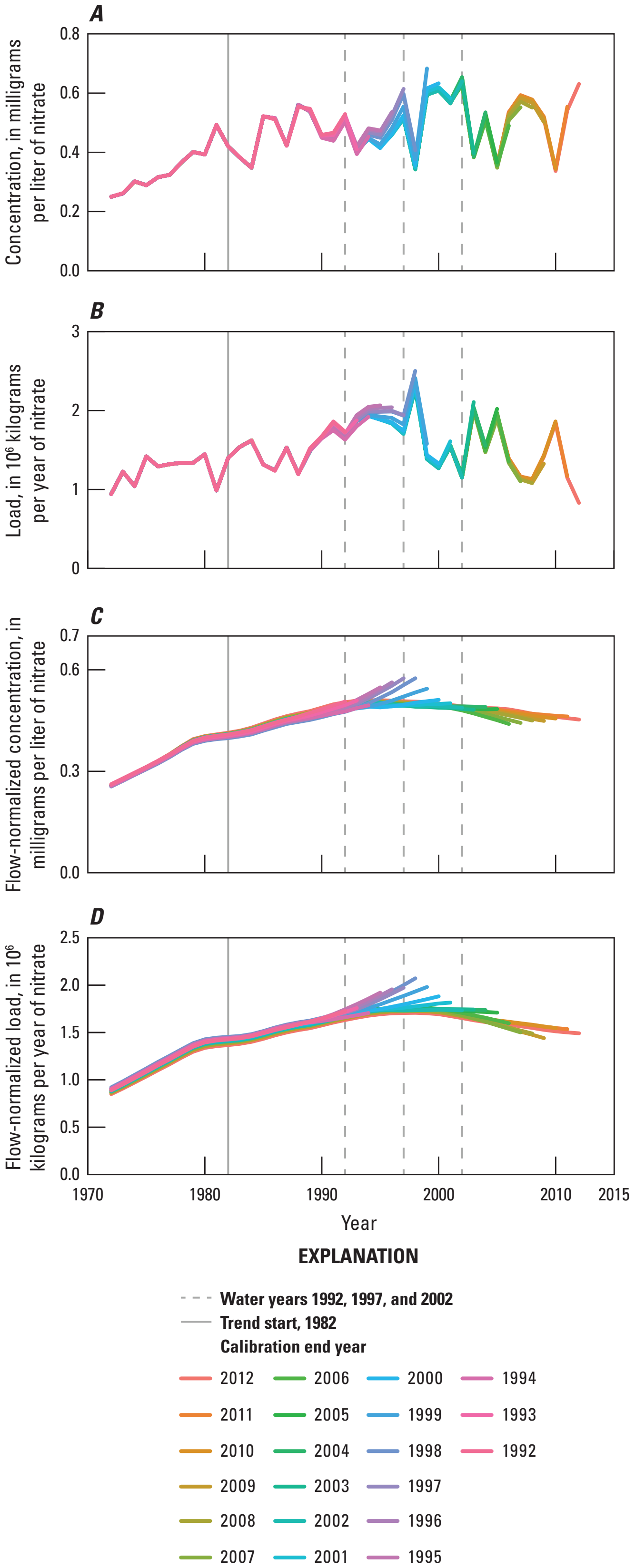


Figure 15. Annual time series of concentration, load, flow-normalized concentration and flow-normalized load estimates for nitrate at site 02215500 Ocmulgee River at Lumber City, Georgia. Initial calibration period from 1972 to 1992 with successive calibrations extending through 2012.

For both sites and parameters, the concentration estimates from successive calibrations are typically within 20 percent of the estimate from the whole-record calibration (figs. 13*A*–15*A*). The one exception is for total phosphorus concentration at site 01646580 (fig. 12*A*), where the estimates from successive calibrations can differ by more than twice that range when compared to estimates from the whole-record calibration; for example, the 2002 estimate for total phosphorus concentration from the 1982–2002 calibration (0.076 mg/L) is 40 percent larger than the 2002 estimate (0.055 mg/L) from the whole-record (1982–2012) calibration (fig. 12*A*). Compared to the concentration estimates, the variability of load estimates between successive calibrations appears to be more specific to the site and parameter. The greatest variability occurs at site 01646580, where estimates of total phosphorus load can vary by 50 percent or more between successive calibrations and the whole-record calibration (for example, the 1998 estimate shown in fig. 12*B*); however, for nitrate at the same site, the variability of annual load estimates between successive calibrations is much smaller, always less than 15 percent (fig. 13*B*). At site 02215500 for total phosphorus and nitrate, the variability of annual load estimates is typically less than +/- 20 percent (figs. 14*B* and 15*B*).

Similar to the concentration and load estimates, flow-normalized concentration and flow-normalized load estimates from successive calibrations are generally in agreement for the early part of the record. As observed by Chanat and others (2016), some of the site-parameter combinations show translations of the successive calibration curves up and down the y-axis. These translations (where the general shape of the curve is maintained) are most obvious for the flow-normalized load estimates (figs. 12*D*–15*D*). The overall shape of the flow-normalized estimates from successive calibrations is generally consistent throughout the record, although both flow-normalized concentration and flow-normalized load estimates display the most variability between successive calibrations around inflection points. Compared to flow-normalized estimates of nitrate concentration, flow-normalized estimates of total phosphorus concentration display more variability between successive calibrations and can be more than 40 percent different when compared to the estimate from the whole-record calibration (for example, the flow-normalized total phosphorus concentration estimates for 2002 at site 01646580 [fig. 12*C*] and for 1995 at site 02215500 [fig. 14*C*]). Flow-normalized estimates of nitrate concentration from all successive calibrations at both sites are generally no more than 15 percent different from the estimate derived from the whole-record calibration (figs. 13*C* and 15*C*). A similar pattern is observed for successive calibrations of flow-normalized load, where estimates of total phosphorus (figs. 12*D* and 14*D*) display more variability than estimates of nitrate (figs. 13*D* and 15*D*). Also similar to the estimates of concentration and load, estimates of flow-normalized concentration and flow-normalized load appear to stabilize when the year of the estimate is more than 7 years from the end of the calibration record. However, all flow-normalized estimates still exhibit some translational variability (shifting of the curve along the y-axis), typically within 10 to 5 percent of the estimate from the whole-record calibration, because of changes in the streamflow distribution as more data are added to the calibration dataset.

The robustness of the estimated trends between successive calibrations is an important consideration for trend analysis. For this study, trends are defined as the amount or percentage change in flow-normalized concentration and flow-normalized load between any 2 years. To explore the variability in trend estimates between successive calibrations, three trend periods were defined (1982–92, 1982–97, and 1982–2002), and the percentage change in flow-normalized concentration and flow-normalized load of nitrate and total phosphorus were calculated for each successive calibration at both sites (total phosphorus and nitrate). These trend estimates determined using successive calibrations were then compared to the trend calculated using the flow-normalized estimates from the whole-record calibration (1982–2012), hereafter referred to as the “whole-record” trend estimate. The selected trend periods all begin in 1982 (solid vertical line in figs. 12–15) and end in 1992, 1997, and 2002 (dashed vertical lines in figs. 12–15). Figure 16 shows the evolution of these three trend estimates as new data are incrementally incorporated into the calibration dataset, grouped by site and estimate type. Specifically, each panel in figure 16 shows the percentage difference of the 1982–92, 1982–97, and 1982–2002 trend estimates calculated from each successive calibration compared to the whole-record trend estimate versus the number of years the trend end year is from the end of the successive calibration. For example, the first 1982–2002 trend estimate for total phosphorus flow-normalized concentration at site 01646580 was made in 2002 when there were zero years between endpoints for trend and for the calibration record (fig. 16*A*, blue square at X=0). This first estimate of the 1982–2002 trend was around 55 percent greater than the whole-record trend estimate of the 1982–2002 trend (determined using the 1982–2012 calibration record). When the calibration period was extended through 2003, the 1982–2002 trend estimate dropped to only about 10 percent different than the whole-record estimate for the 1982–2002 trend. When the calibration period was extended again through 2004, the 1982–2002 trend estimate dropped to around -4 percent different from the whole-record trend estimate. With each incremental addition of a year’s worth of new data to the calibration dataset, the percentage difference in the 1982–2002 trend estimate continues to fluctuate within a few percentage of zero until the calibration end year is approximately 7 years from the end of the trend period. Therefore, for this example, the 1982–2002 trend estimate is less than 1 percent different from the whole-record trend estimate when the calibration period is 1982–2009 or longer (fig. 16*A*, 7 years on x-axis; see also the 1982–2002 and 1982–2009 successive calibrations in fig. 12*C*).

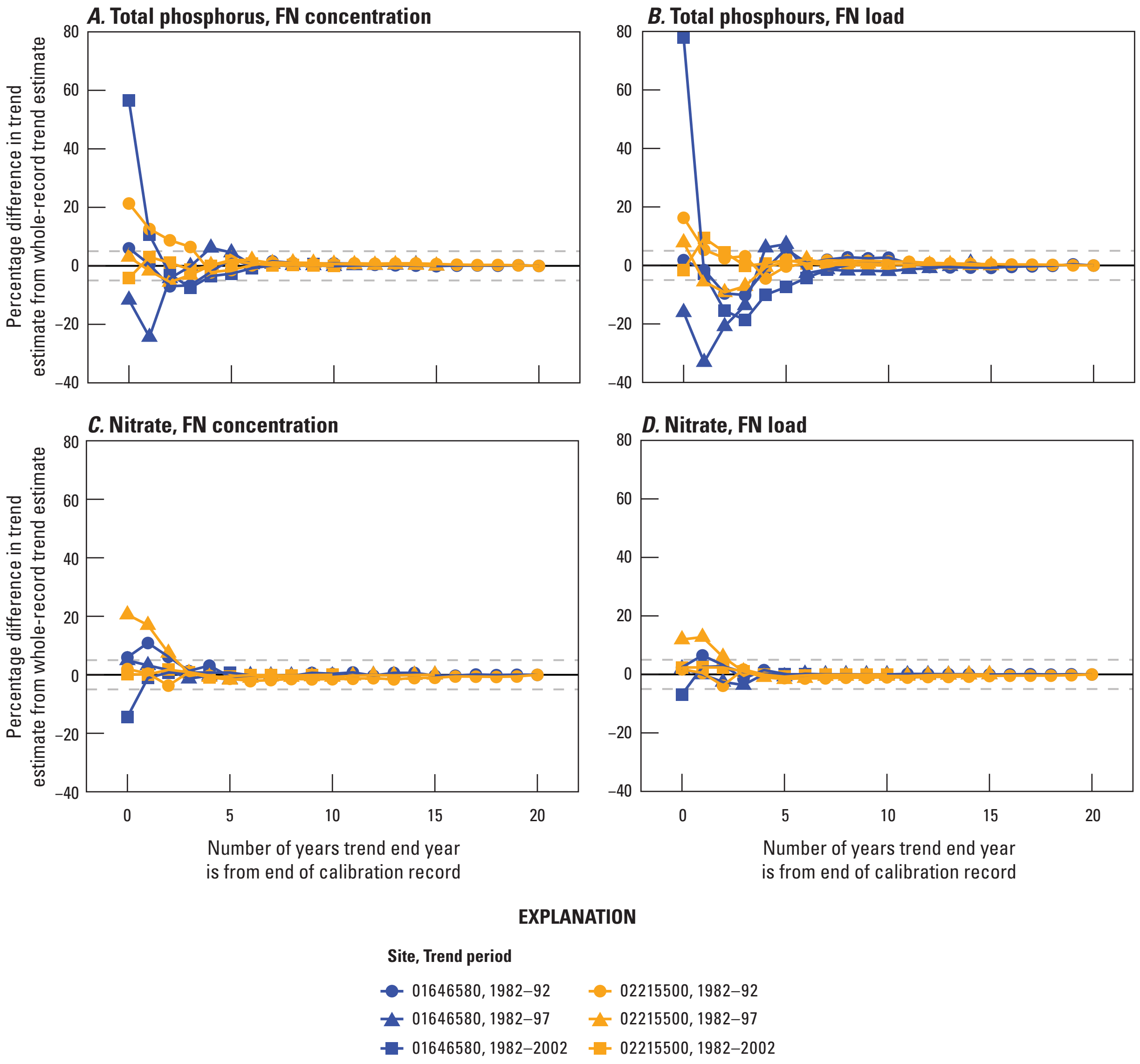


Figure 16. Evolution of flow-normalized (FN) trend estimates for 1982–92, 1982–97, and 1982–2002 with the incremental incorporation of new data to calibration dataset for nitrate and total phosphorus at sites 01646580 Potomac River at Chain Bridge, at Washington, D.C., and 02215500 Ocmulgee River at Lumber City, Georgia.

When taken together—across both sites, both parameters, and all three trend periods—the trend estimates appear to stabilize to within about 5 percent difference of the corresponding whole-record trend estimate when the calibration record extends 5 years beyond the end year of the trend period and remains within just a few percentage difference of the whole-record trend estimate after 7 years or more (fig. 16). Accordingly, when the end year of the trend period is within 5–7 years of the calibration end year, the trend estimates exhibit more variability as new data are added to the calibration record. Generally, this variability is typically within +/- 20 percent of the whole-record trend estimate. However, the total phosphorus flow-normalized concentration and load trends for 1982–97 and 1982–2002 at site 01646580 exhibit some trend estimates that vary outside of this range (figs. 16*A*–16*B*). The end years of these trend periods occur near a large inflection point in the whole-record calibration. While trend estimates for most of the trend periods generally increase in accuracy as the calibration period is extended beyond the end year of the trend period, this change is not necessarily linear, and the least accurate trend estimates do not always occur when the trend period and the calibration period have the same end year. For example, trend estimates of nitrate flow-normalized load at site 02215500 for the 1982–92 trend period are the most different from the whole-record trend estimate with the 1982–1994 successive calibration, when the calibration end year was 2 years beyond the end year of the trend period (fig. 16*D*). The patterns observed for these sites and parameters indicate the variability of trend estimates caused by the incremental incorporation of new data in the calibration dataset likely varies by site, parameter, and if the trend period ends near an emerging inflection point.

Continuing with the same examples, even though the magnitude of trend estimates varied a great deal between successive calibrations and the whole-record calibration, the direction of the trend estimates was relatively stable as new data were incrementally added to the calibration record. For the 1982–92 trend period, there are trend estimates from 20 successive calibrations that can be compared; for the 1982–97 trend period, there are 15 trend estimates for comparison; and for the 1982–2002 trend period, there are 10 trend estimates. In total, that equals 45 trend estimates for each site, parameter, and estimate type (concentration or load) combination to compare to the corresponding whole-record trend estimate. For total phosphorus trends at site 01646580, there were only 2 concentration and 2 load trend estimates derived from the successive calibrations that differ in direction from the corresponding whole-record trend estimate. For nitrate trends at the same site, there was only 1 concentration and 1 load trend estimate that differed in direction. For site 02215500, the direction of all the trend estimates calculated using the successive calibrations, for both total phosphorus and nitrate concentration and load, was the same as the corresponding whole-record trend estimate. Although trend direction reversals at these two sites between successive calibrations and the whole-record calibration were rare, other sites and parameters may have different outcomes.

As shown in the successive calibrations in figures 12–15, if a trend period is relatively short, possibly less than 5 years, changes in trend direction between successive calibrations may be more likely. For example, if a 3-year trend period between 1992 and 1995 is considered for flow-normalized total phosphorus load at site 01646580 (fig. 12*D*), the 1982–95 calibration indicates decreasing loads between 1992 and 1995. When adding 2 additional years of data to the calibration period (1982–97), the flow-normalized load estimates from that successive calibration indicate no change in load from 1992–95. After adding another 2 years of data to the calibration (1982–99 successive calibration), the load trend between 1992 and 1995 is increasing and remains so for all successive calibrations ending after 1999. This example and the successive calibrations shown in figures 12–15 indicate that the direction of trend estimates is more robust for longer trend periods.

A full exploration into the influence of the incremental incorporation of new data on WRTDS estimates has yet to be accomplished. Such an effort would help to more fully characterize edge effects at a diverse set of sites and possibly allow for greater insight into trend estimates. The successive calibrations shown in figures 12–15 indicate edge effects are most pronounced around an emerging inflection point. A better understanding of edge effects in flow-normalized estimates may help to identify emerging inflection points during the last few years of the record, arguably the years that are of greatest interest to policy makers and the general public. It is also possible that dramatically different streamflow conditions in the first or last few years of the record may also induce variable flow-normalized estimates in successive trends, though this effect has not been explored. Several recent studies (Hodgkins and Dudley, 2011; Hirsch and Archfield, 2015) have indicated increases in the frequency, magnitude, and duration of extreme hydrologic events, thus a better understanding of the influence of extreme hydrology on edge effects is likely to grow in importance in the future.

As described previously in this report, WRTDS models were calibrated using all the available data for each site-parameter combination; furthermore, all annual estimates, including those for the start and end years, whether or not these start and end years were also the first or last years of the calibration dataset, were reported (in data releases) as long as data passed all relevant screens for the trend period. This decision makes between-site comparisons more nuanced because the calibration period varies depending on the site and parameter. Some sites have trend periods that are buffered by 7 or more years of record prior to the start of the trend period or have records extending through calendar year 2014, which provides additional stability for the reported 2012 trend estimates, whereas other sites have no data beyond the trend periods. For example, annual estimates for 2011 can be reported when a site-parameter combination has a calibration period from 1965 to calendar year 2014 or from 2003 to 2011. Nevertheless, this decision more accurately characterizes changes in water quality at individual sites over time because not only are more water-quality data included in the calibration, but more streamflow data are used to develop the probability distributions from which flow-normalized estimates are generated. Streamflow can have broad periodic fluctuations in magnitude, and the capture of these fluctuations over as long a time as possible provides more robust information to the model. Because of the window-weighting approach used during the calibration of regression models in WRTDS, the decision to calibrate models using all available streamflow data has some influence on the model coefficients but could have more considerable influence on the flow-normalized estimates. Changes in the concentration-flow relation might occur during periods of missing water-quality data and could affect flow-normalized estimates by an unknown magnitude. Variability of streamflow magnitudes can be assessed through the analysis of streamflow records; however, for sites with periods of streamflow without water-quality data, determination of changes in concentration-flow relation is impossible because of the lack of data.

Limitations on Trend Analysis with the Seasonal Wave (SEAWAVE–Q) Model

The SEAWAVE–Q model is a time constant model, unlike WRTDS. The seasonal wave term in the model (equation 22; visualized in Ryberg and Vecchia, 2013, appendix 3) that represents intra-annual variability in concentration assumes a constant application period or periods that result in a constant time at which concentration reaches a maximum. This seasonal wave is generally very effective for capturing the pulsed input of pesticide and the decay, or decline in concentration after application; however, it does not adjust for any changes in cropping patterns or climate over the period of record that might change the ideal seasonal wave. The trend (equation 21) is constant over the trend period, when in actuality the period may experience a series of differing upward and downward trends as concentrations respond to climate, market, and other forces. The model does not adjust for potential serial correlation (as WRTDS does with bootstrapping); however, the effects of serial correlation were reduced by thinning the pesticide data to no more than weekly samples before trend analysis (see the "Data Thinning" section in the "Pesticide Data Processing" section).

Limitations of the Ecology Trend Method

Limitations of the ecology trend method are mostly artifacts of the inherent high interannual variability associated with stream ecology monitoring data combined with low statistical power associated with small sample sizes. Interannual variability among and within stream communities is influenced by a variety of factors, including natural fluctuations in climate, human induced disturbance, and outbreaks of disease and invasions of nonnative species. Accounting for disease and invasion would be difficult given the scale and design of this trend study and examining the influence of human disturbance was beyond the scope of this assessment. Fluctuations in climate are often strongly linked to changes in community structure (Mazor and others, 2009; Greenwood and Booker, 2015); however, these influences are likely assemblage, time, and regionally specific. Limited period of record, low sample density, and the inability to group similar sites because of the sparse distribution and large geographic coverage prevented any adjustment in the biological response by more than one exogenous variable during analysis. Additional limitations are discussed in the "Ecology Results" section.

Laboratory Method and Performance Considerations for Trend Interpretation

National Water Quality Laboratory Method and Procedure Changes

For trend studies, it is important to identify major changes in laboratory methods or procedures that may induce a bias in water-quality determinations and possibly produce an artificial step trend in the data. Major changes in laboratory methods and the resulting analytical values at the NWQL have been documented in a variety of resources. For the most part, these have been documented in USGS publications and internal memorandums—some of which are available online to USGS employees. Using these resources (appendix 3, table 3–1) timelines of potentially important changes in NWQL laboratory methods and procedures were developed for the target constituent groups in this study (appendix 3, tables 3–2 to 3–6).

Note that the method and procedure changes documented in these tables (tables 3–2 to 3–6) apply only to data analyzed at the NWQL, and the information contained within these tables is not exhaustive or definitive; other changes and errors may have occurred during the study period but were not documented in the available resources (table 3–1). The timelines were pieced together using disparate technical memorandums, reports, and Rapi-Notes (internal USGS memoranda from the NWQL) and may be missing important changes. Because of the time-consuming nature of this compilation, lack of public access to similar information for some other laboratories, and the total number of different laboratories that produced the data in this study, similar timelines for other laboratories were not compiled.

Changes in Laboratory Analysis and Sample Collection Methods

In trend analysis, the reported concentration of a particular target parameter is assumed to be a random variable that depends not only on the true riverine concentration but also on the way the sample is collected, preserved, and analyzed. Inevitable changes in analytical methods and in sample collection and preservation methods are nonrandom events that can manifest as step changes in concentrations and thus can lead to bias in some of the trend results. Such bias, if large enough in relation to environmental trends, may affect the interpretive and causal analyses. A complete description of this study's investigation into the occurrence of step trends including background, methods, and results is provided in appendix 4. Cases where potential laboratory or field method changes could have produced bias large enough to affect interpretation of trends on a regional or national scale also are provided in appendix 4.

The presence of step trends in flow-normalized concentrations of selected water-quality constituents was evaluated using 4-year moving windows beginning with 1974 (January 1, 1972 through December 31, 1975) and ending with 2011 (January 1, 2009 through December 31, 2012). Water-quality data consisted of a national dataset of concentrations compiled from samples collected during 1972–2012 by the USGS and various State agencies. These data are described in the “Water-Quality, Pesticide, and Ecology Data Compilation” section and in appendix 4. Data were harmonized and processed but were not screened for trend analysis. The purpose of the step-trend analysis was to compare step trends between the USGS and outside agency sites, and among the various outside agency sites, to determine cases where and (or) when anomalous trends that may reflect changes in laboratory or field sampling methods, rather than environmental changes, could cause difficulties in the interpretive and causal analyses of the final trends.

For each 4-year moving window, a two-stage flow-normalization process based on streamflow anomalies (orthogonal components of log-transformed daily streamflow representing annual, seasonal, and daily variability) was used to remove flow-related variability in concentrations, and the flow-normalized concentrations were analyzed for step trends. Highly variable (through time) and sparse (in space) site locations for both USGS and outside agency sources presented difficulties for making quantitative comparisons for many time windows and constituents. However, analysis of regional trend patterns for USGS and outside agency sites indicated that regional trend patterns for the vast majority of constituents and years generally were consistent among USGS and outside agencies, and there were few instances where anomalous trends, likely because of laboratory or field method changes, clearly stood out from regional trend patterns. Laboratory or field method changes probably had a limited effect on flow-normalized concentrations compared to flow-related variability and regional environmental variability or change. Although undocumented laboratory and field method changes likely did occur, the detection of the relatively small effects of those changes (compared to environmental variability or trends) would require a combination of higher (and more uniform) sampling frequencies, more colocated or nearly colocated sites from USGS and outside agencies, and (or) more detailed information on timing of changes along with field replicate samples.

Laboratory Performance Bias Evaluation

Temporal changes in the performance of laboratory analytical methods used to measure inorganic and organic concentrations have the potential to create artificial trends or to introduce noise that would make true environmental trends harder to detect. In order to track laboratory performance, the NWQL and the USGS District Water-Quality Laboratory in Ocala, Florida (until the Ocala laboratory was closed in 2004), submitted the results of double-blind QC reference samples to the USGS Branch of Quality Systems as part of the Inorganic Blind Sample Project and the Organic Blind Sample Project. About half of these double-blind reference samples were from natural-matrix waters derived from snowmelt, surface-waters, or ground-water sources; the remaining samples were low ionic-strength solutions made from deionized water instead of natural-matrix waters to preserve the low conductivity (Woodworth and Connor, 2003; Amy Ludtke, U.S. Geological Survey, Office of Water Quality, written commun., September 2015). The natural-matrix standard reference samples were used as is (undiluted), diluted with deionized water, or mixed in varying proportions with other standard reference samples in order to achieve a variety of concentrations within the range of concentrations that corresponds to the range typical of environmental water samples. For each reference sample, the NWQL or Ocala value was compared to the target value, which was the median concentration for the same reference sample measured by about 150 USGS and non-USGS laboratories.

The recovery of each NWQL and Ocala blind sample for all of the parameters used in the trend analysis for nutrients, major ions, salinity, and carbon was calculated as the measured concentration divided by the target concentration and was expressed as a percentage. Plots of recovery results are presented in graphs (appendix 5), grouped by pmcode, which is a combination of parameter code and method code. For example, pmcode 00930AA035 corresponds to parameter code 00930 and method code AA035. A list of pmcodes, grouped by parameter code and parameter name, is provided in appendix 5, table 5–1. Although not used to adjust results in this study, the recovery data can be used in conjunction with trend results to consider whether changes in laboratory performance over time could contribute to the trend results for a specific parameter.

Measured concentrations of nutrients, major ions, salinity, and carbon parameters to be analyzed for trends were not adjusted to compensate for changes in laboratory recovery over time for several reasons. Double-blind samples have been submitted to the NWQL since 1984 and the Ocala laboratory since 1985; no blind samples are available before that time. As a result, no recovery adjustments could be made to measured concentrations prior to 1984. Additionally, censored data could not be adjusted because detection limits were established earlier by the laboratory using a separate, predefined approach that incorporates the variability of a laboratory method (Oblinger Childress and others, 1999). Because the double-blind samples are intentionally low in sediment or are filtered prior to distribution, the samples are not representative of natural waters with higher sediment concentrations (and thus potential matrix interference). This approach may lead to inaccurate recovery values for parameters that typically have a particulate component, such as total phosphorus or Kjeldahl nitrogen. Finally, the recovery results apply only to data from the NWIS database that had the analyzing entity defined as NWQL or as District Water-Quality Laboratory, Ocala, Florida, for specific parameters and methods. A majority of the samples in the NWIS database were missing either the analyzing entity or the analytical method, so it was unclear whether these data should be adjusted. Similar adjustments also could not be made to water-quality data from non-NWQL laboratories. Although many of these laboratories engage in similar quality-control practices, it was beyond the scope of this study to compile and interpret those data. Because of these limitations, no additional evaluation of laboratory-related changes over time was undertaken as part of this study other than what is described in appendix 4.

Site-Specific Considerations when Modeling Trends with the Weighted Regressions on Time, Discharge, and Season (WRTDS) Model

All models are constrained by certain limitations. At the most basic level, WRTDS (and many other regression-based methods) describes changes in water quality using a linear model, allowing for relatively straightforward interpretation and solutions that can be extended in a variety of ways to account for nonlinear relations but nevertheless places limitations on the model structure and possibly on predictive power. For WRTDS in particular, the weighting approach used during model calibration does not accurately characterize abrupt step changes in water-quality data. Also, the flow-normalization process in WRTDS uses daily streamflow values across the entire period of record, which can affect the flow-normalized estimates when there is a trend in streamflow. This section will briefly discuss these issues and their influence on WRTDS model results. This discussion is not meant to be an exhaustive exploration or characterization of these issues but rather provides some important caveats on the results of this study.

Step Changes in Water-Quality Data

The WRTDS model uses weighted regression to calibrate each regression model. The individual weights assigned to observations during each calibration are determined using three windows: (1) time, (2) season, and (3) streamflow. This weighting scheme most accurately characterizes gradual changes in water quality over time, such as those related to incremental changes in the watershed but does not appropriately characterize step changes. Abrupt changes in water quality, for example, because of the construction or upgrade of a large wastewater treatment plant, would be smoothed out over a period of about 7 years because of the window-weighting approach used in WRTDS.

Other common approaches for trend estimation that presume a linear relation between water quality and time also struggle with accurately characterizing step changes in water quality. Nonparametric tests, such as the Mann-Kendall test (Mann, 1945; Kendall, 1975) or the Seasonal Kendall test, are typically accompanied with a Sen slope estimate that can imply, but is not based on, a linear trend over time. For trend analyses that use linear regression, temporal change is typically described as a linear or perhaps quadratic trend over time, though these models can be extended to accommodate step changes by incorporating an interaction term that divides the water-quality record into distinct periods on either side of a step change.

To illustrate the effect of a water-quality step change on WRTDS estimates, the example of total phosphorus at site mnVR–15.6_MNMCES located on the Vermillion River near Empire, Minnesota, was considered. The Empire Wastewater Treatment Facility shifted its streamflow from the Vermillion River to the Mississippi River in the late 2000s, causing a large decrease in total phosphorus concentrations (Minnesota Pollution Control Agency, 2012). The observed total phosphorus concentration data show this large abrupt decrease in 2007 (fig. 17). This abrupt change in total phosphorus concentrations is incorrectly characterized in the WRTDS output as a gradual decrease in estimated and flow-normalized concentrations beginning around 2002 and extending through 2012 (fig. 18). The residuals for this site clearly show a poor model fit for the years surrounding the step change in total phosphorus (fig. 19).

The WRTDS model results were rejected for any sites or constituents that had a visually recognizable step change in the water-quality data or in the model residuals. This visual evaluation was a manual and somewhat subjective process that occurred during review of WRTDS model diagnostics.

Trends in Streamflow

One of the largest challenges in estimating trends is deciphering the proportion of variability in concentration, load, or ecologic metrics that arise from variability in streamflow versus the proportion that is due to other changes and processes in the watershed or channel. Many water-quality parameters and ecological metrics are related to streamflow, and changes in streamflow can make it challenging to evaluate progress or

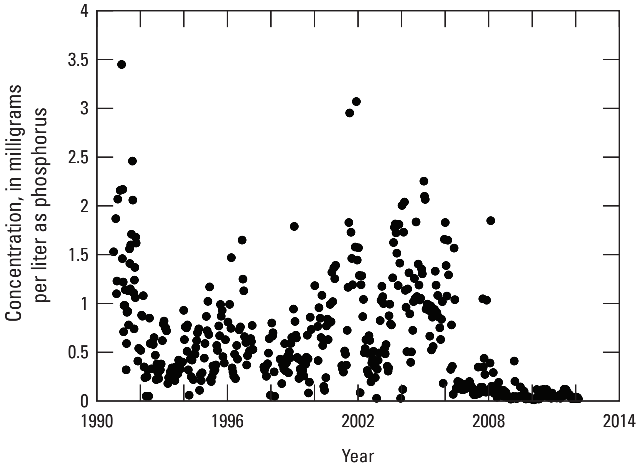
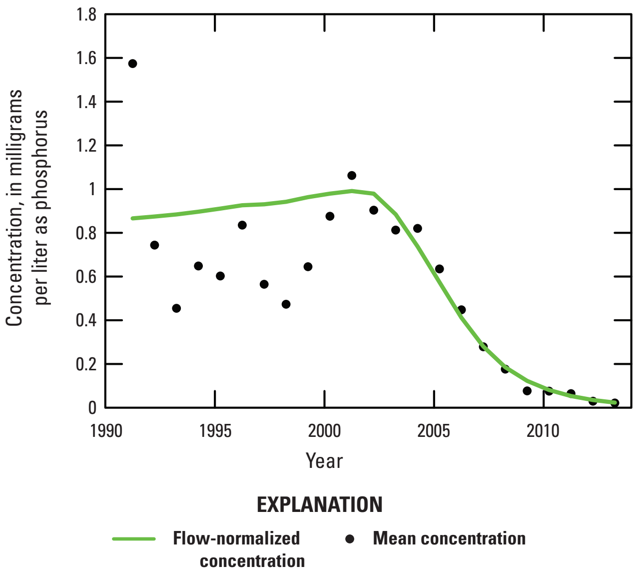


Figure 17. Observed total phosphorus concentrations at site mnVR–15.6_MNMCES Vermillion River near Empire, Minnesota, for 1990–2012.



Note: This plot includes an annual estimate for 1991 flow-normalized concentration; however, estimates outside the defined trend periods are not reported in the main compilation of results.

Figure 18. The Weighted Regressions on Time, Discharge, and Season (WRTDS) estimates of annual total phosphorus concentrations and flow-normalized total phosphorus concentrations for site mnVR–15.6_MNMCES Vermillion River near Empire, Minnesota (paired with streamgage 05345000).

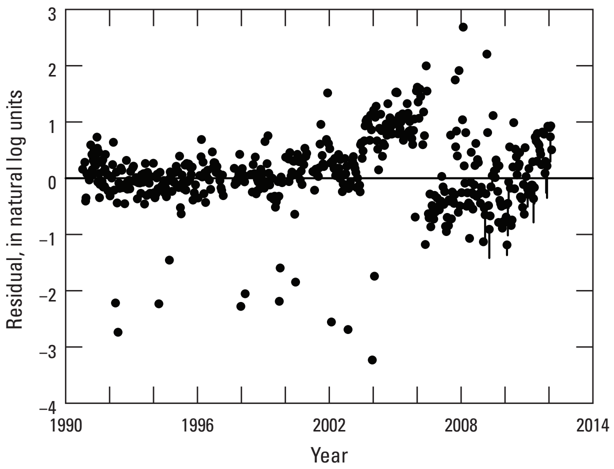


Figure 19. The Weighted Regressions on Time, Discharge, and Season (WRTDS) residuals for total phosphorus concentrations at site mnVR–15.6_MNMCES Vermillion River near Empire, Minnesota, versus year.

deterioration in response to human activities if the influence of streamflow is not considered. Trend analysis methods used in this report can filter out the effect of year-to-year variability in streamflow but do not have mechanisms to account for gradual or abrupt changes in streamflow. All three trend analysis methods used in this study to assess trends in water quality, pesticides, or ecology incorporate streamflow data as an explanatory variable in some form (for example, daily values in WRTDS and antecedent streamflow metrics for the ecology model). Because of the reliance of these models on streamflow, abrupt or gradual changes in streamflow over time may influence the results produced by these models; however, WRTDS is particularly sensitive to nonstationary streamflow, primarily because of the flow-normalization process.

Details of the analysis of gradual (monotonic) and abrupt trends in annual streamflow at all streamgages used for trend analysis in this study can be found in appendix 6, which also presents a summary of streamflow trend results for water-quality, pesticide, and ecological sites for all trend periods. Streamflow trend results for all streamgages used in this study are available in a data release (Farmer and others, 2017). The results of the annual streamflow trends were developed to assist in identifying sites where changing streamflow may be influencing the trend results. While this process was not completed for this report, these streamflow trends can serve as a basis for a future analysis. As mentioned previously, this is particularly important for the WRTDS flow-normalized estimates because the interpretation of the water-quality trends differ from standard interpretation when there is a systematic change in streamflow over time. The following subsections more fully describe the influence of nonstationary streamflow on WRTDS trend results.

Gradual Changes in Streamflow

Because the flow-normalization process is designed to remove the random variability in water quality related to the year-to-year variability in streamflow, flow-normalized load and concentration curves generally appear like a locally-weighted scatterplot smoothing (LOWESS) (Cleveland, 1979; Cleveland and Devlin, 1988) line through the annual load and concentration estimates (even though flow normalization uses a completely different process than LOWESS). LOWESS is a nonparametric regression approach that uses local weighting of nearby observations to fit a smooth line through bivariate data. When there is a gradual trend in streamflow over the calibration period, this can influence the flow-normalized estimates and cause the flow-normalized load and concentration curves to depart from the annual load and concentration estimates. For example, site 09498500 Salt River near Roosevelt, Arizona, has decreasing streamflow from the early 1980s through 2012, as demonstrated by large decreases in the annual 7-day minimum streamflow, mean daily streamflow, and 7-day maximum streamflow (fig. 20). Over the same period of time, sampled total-dissolved-solids concentrations have increased (fig. 21). The WRTDS annual concentration and load estimates from 1982–2012 are increasing and decreasing, respectively, over time, whereas the flow-normalized concentration and flow-normalized load estimates vary little over the trend period and do not follow the annual estimates (fig. 22).

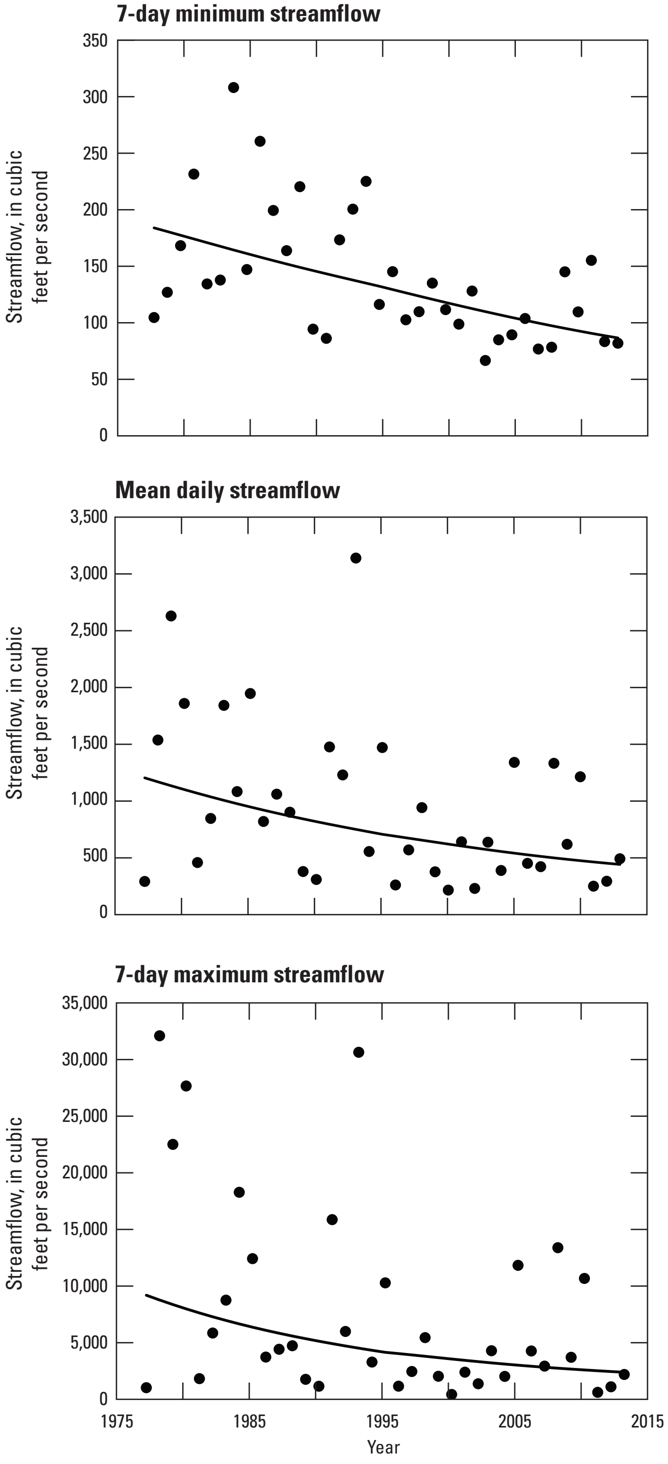


Figure 20. Annual 7-day minimum, mean daily, and 7-day maximum streamflow at site 09498500 Salt River near Roosevelt, Arizona.

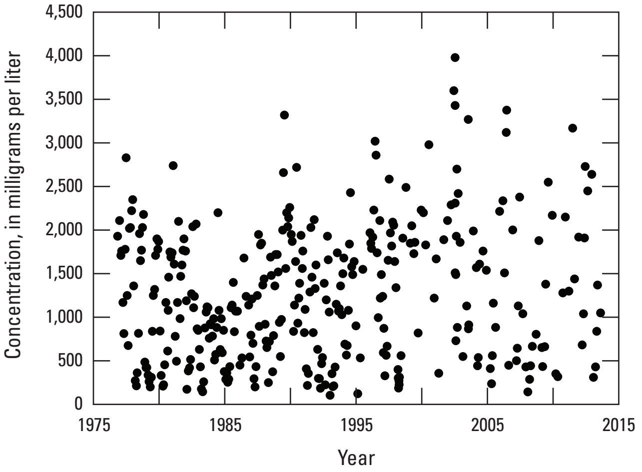


Figure 21. Total-dissolved-solids concentrations over time at site 09498500 Salt River near Roosevelt, Arizona.

Theoretically, a gradual change in streamflow equates to a gradual shift in the probability distributions of streamflow at a site over time, indicating that flows in the beginning of the record are not characteristic of those at the end of the record. Gradual changes in streamflow can be due to situations such as increasing urbanization, increasing use of artificial land drainage, gradual adoption of various land-management practices that influence streamflow (such as farm ponds or stormwater control structures), or gradual depletion of groundwater (causing decreases in base streamflow in streams). Gradual changes in streamflow can also be the result of slow shifts in regional climate. A gradual shift in the probability distributions of streamflow over time undermines the expectation of stationarity of streamflow that is a stated assumption of the flow-normalization process. In the seminal paper introducing WRTDS, Hirsch and others (2010) describe the flow-normalization process:

> Flow-normalization eliminates the influence of the temporal pattern of discharge, by viewing the discharge on any given day as a random sample of the discharges that might have taken place on that day. Thus, the method requires some means of estimating the probability distribution of discharge values for that day. Flow-normalization uses the actual historical sample of discharge values for a given day, with each historical value being assigned an equal probability of happening in any given year.

A gradual trend in streamflow clearly invalidates the assumption that streamflow on any given day is a random sample drawn from the probability distribution of streamflow values on that day across the period of record. When there is a gradual trend in streamflow, some streamflow values may be more likely to occur than others, depending on the point in time.

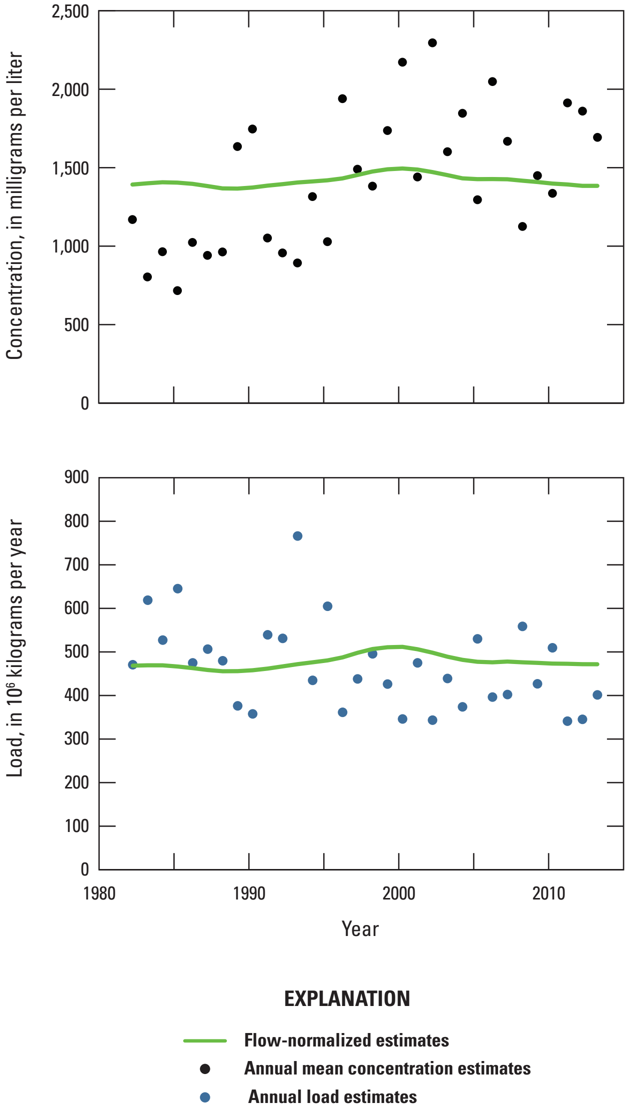


Figure 22. The Weighted Regressions on Time, Discharge, and Season (WRTDS) annual concentration and load estimates and flow-normalized curves for total dissolved solids at site 09498500 Salt River near Roosevelt, Arizona.

For sites like 09498500 (fig. 20), where there is a strong gradual trend in streamflow causing the probability distributions of daily streamflow to be nonstationary over the period of record, the theoretical description of the flow-normalized estimates differs from the definition presented in Hirsch and others (2010). Any given streamflow record can be considered the sum (or product) of seasonal, trend, and random components, and, as presented by Hirsch and others (2010), the flow-normalization process is designed based on the assumption that the trend component in the streamflow record is very small compared to the other two components. In the context of nonstationary streamflow, flow normalization removes the random year-to-year variability as a result of streamflow and uses a consistent annual set of estimated probability distributions of daily streamflow; thus, the flow normalization process is estimating concentration and load as if the probability distributions of streamflow were stationary. This hypothetical situation of stationary streamflow at a site where streamflow is not stationary leads to differences in the interpretation of flow-normalized estimates, depending on whether there is a trend in streamflow at the site.

Flow-normalized estimates show the variability of concentration or load without the direct influence of changes in streamflow, whether those changes are random, which is what occurs at sites without a trend in flow, or random and systematic, which is what occurs at sites with a trend in flow. For sites without a trend in streamflow, the flow-normalized estimates describe the general in-stream water-quality conditions, essentially the water-quality trend experienced by the biota in the stream, and these changes are typically attributed to human activities in the basin; however, for sites with trends in streamflow, the argument is made that flow-normalized estimates *do not* describe in-stream water-quality conditions and instead describe changes in the *inputs* to the stream or changes in some other in-channel processes not related to changes in flow. Changes to the input of constituents to the stream are also typically attributed to human activities in the basin or to changes in climate or other physical processes in the channel or basin (but are not correlated with the changes in flow). For example, the annual estimates of concentration at site 09498500 have increased over time, but the flow-normalized estimates of concentration have changed little, indicating that although in-stream total dissolved-solids concentrations have increased, the input of total dissolved solids to the stream has largely remained unchanged (fig. 22). Furthermore, this pattern indicates that changes in streamflow (possibly driven by humans or natural processes), not increases in the input of total dissolved solids to the stream, are affecting water quality at this site.

An improved approach that accounts for trends in the streamflow record would help to better align the interpretation of flow-normalized estimates between sites with and without streamflow trends. Accounting for trends in streamflow would allow for a better description of the general trend in in-stream concentrations and loads at a site with a trend in streamflow and would make the interpretation of flow-normalized estimates between sites with and without a trend in streamflow directly comparable. If such a method were developed, it would allow for the removal of only the random variability of streamflow and would describe the in-stream water-quality changes whether or not there is a streamflow trend at the site.

Identification of the magnitude of the streamflow trend, which results in the divergence of the flow-normalized estimates from the annual estimates and the subsequent need to qualify interpretations of the flow-normalized estimates, has yet to be studied and likely varies by site and constituent. The response of flow-normalized estimates to increasingly larger trends in streamflow is likely gradational, and identification of a single threshold for which the trend in streamflow is large enough to cause a divergence, or deciding what constitutes a divergence, of the flow-normalized curve from the annual estimates may be difficult. Current recommendations suggest professional judgement is the appropriate method for deciding if trends in streamflow are large enough to influence flow-normalized estimates (Hirsch and others, 2010; Hirsch and De Cicco, 2015). To assist with identifying sites that may have gradual trends in flow, nonparametric Mann-Kendall trend tests were run using annual streamflow statistics (7-day minimum, mean daily, and 1-day maximum) for all gages used in this study (see appendix 6 for details). Sites were flagged in Farmer and others (2017) to indicate a possible monotonic trend in streamflow if the results of these statistical tests were significant. This information can be used when evaluating the model diagnostics and interpreting the model output but did not affect any results presented in this report. Sites with significant trends in streamflow were not automatically excluded from this study because their annual and flow-normalized results still provided important information. When there was a significant trend in streamflow, the flow-normalized estimates may have diverged from the annual estimates. In these cases, the flow-normalized estimates reflect the portion of the in-stream trend that was caused by factors other than changes in streamflow, whereas the annual estimates reflect the overall in-stream trend that was caused by all factors, including random and systematic changes in streamflow. Future analyses can draw upon this information to better describe the streamflow trend that results in diverging flow-normalized and annual estimates and identify specific sites with trends in streamflow when the interpretation of the flow-normalized estimates differs from that at sites with stationary streamflow.

Abrupt Changes in Streamflow

While gradual trends in streamflow cause a nuanced interpretation of the flow-normalized estimates, abrupt changes in streamflow also are problematic for calibration and flow normalization in WRTDS. The relation of hydrology to water quality may be very different before and after a step change in streamflow. The weighting approach used to calibrate WRTDS would distribute this change across the time window (typically 14 years for this study, as the half-window width is 7 years), misrepresenting an abrupt change in the modeled estimates. Use of the entire streamflow record for flow normalization would be inappropriate because the magnitude and frequency of some flows may be very different before and after the step change. Until WRTDS is updated to divide the record into periods with similar hydrologic conditions (that is, before and after the abrupt shift in hydrology) that each meet the assumption of stationarity, the use of WRTDS with streamflow records containing a step change is inappropriate. Furthermore, it is unclear what magnitude of step change in streamflow corresponds to inappropriate use of the WRTDS model. This particular issue has yet to be fully explored.

Abrupt shifts in streamflow may be induced by engineered structures or management decisions, such as the construction or removal of a dam, large and abrupt changes in water diversions into or out of the watershed, and by regional weather patterns. For a relatively short streamflow record, an abrupt shift in hydrology that persists for a number of years, and that does not coincide with an engineered activity in the basin, can easily be construed as a step change in the probability distribution of streamflow. However, if the streamflow record was extended back in time and multiple shifts in the hydrology could be observed, it would be reasonable to conclude that there is a high degree of persistence, and the hydrology is not actually nonstationary (Cohn and Lins, 2005). In this situation, the flow-normalization process would be appropriate and would be assumed to accurately reduce the variability in water quality associated with year-to-year variations in streamflow.

To illustrate, annual mean daily streamflows since 1992 at the James River in South Dakota (site sd460761) show rather abrupt shifts in hydrologic conditions over time (fig. 23). A wet period can be observed from 1993 through 2001, a 5-year dry period follows, then another 5-year wet period, and finally perhaps a return to another dry period in the final 2 years of record. Because this record shows high autocorrelation during specific periods of time and several periods of low or high streamflows over time, the flow-normalization process is still an appropriate means of reducing the variability of the concentration and load estimates by removing the random variability as a result of streamflow. The annual WRTDS results for total dissolved-solids flow-normalized load at this site are also shown in figure 23. During the 2002–6 dry period, the annual loads plot below the flow-normalized load curve, indicating that these low values are influenced by the relatively low streamflows for those years. Overall, the flow-normalized curve smooths out the variation as a result of streamflow and reveal the long-term pattern of change of total dissolved-solids load at this site.

To assist with identifying sites that may have abrupt changes in flow, Pettitt tests (Pettitt, 1979) were run using annual streamflow statistics (7-day minimum, mean daily, and 1-day maximum) for all gages used in this study (see appendix 6 for details). Similar to the analysis of gradual trends in streamflow described above, sites were flagged in Farmer and others (2017) according to the results of the Pettitt tests to indicate if a possible step change in streamflow occurred at a site. Deciphering between periodic shifts and a step change in hydrology can be a challenge with a short streamflow record. Also, the Pettitt test only indicates if at least 1 change point occurs in the streamflow record thus sites where the Pettitt change test indicated a significant step change may actually have multiple shifts in streamflow. Although this information was not used when evaluating the model diagnostics, it can be used in future analysis to more specifically identify water-quality sites that are affected by a step change in streamflow. Sites with significant streamflow step trends were not automatically excluded.

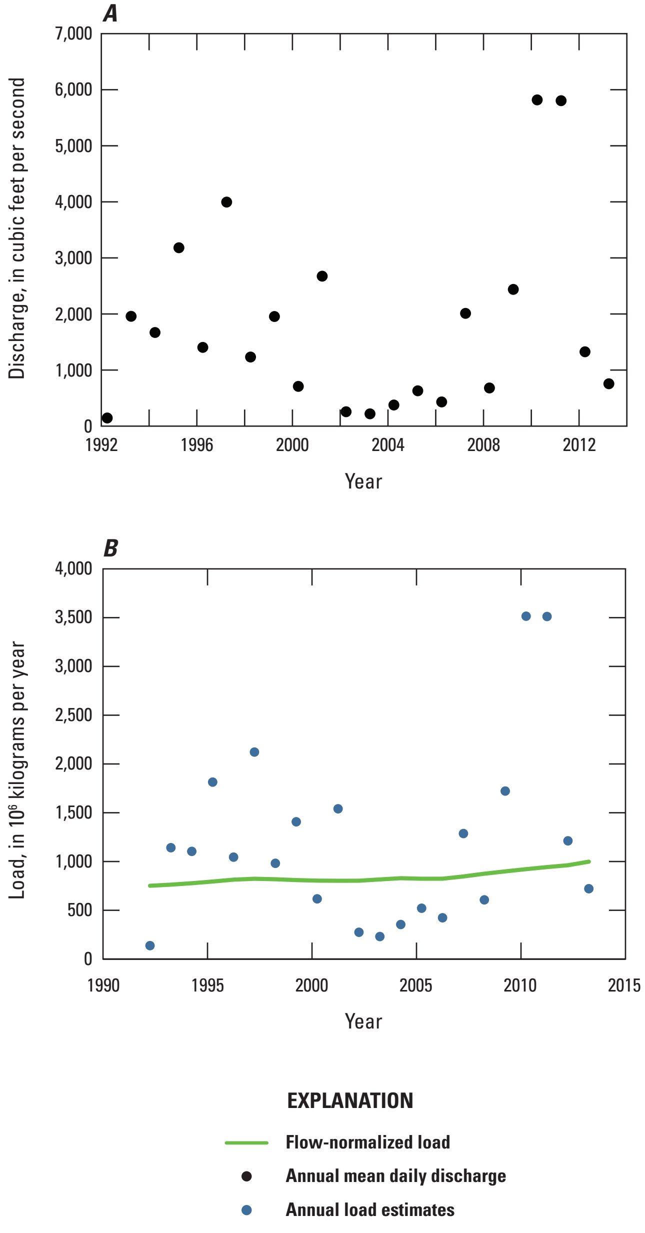


Figure 23. The *A*, mean annual daily streamflow and *B*, estimated load and flow-normalized load for total dissolved solids at site sd460761 James River in South Dakota (streamgage 06478500).

Trend Results

Trend results for each of the trend models are summarized in appendix 7 and discussed in the following sections. Table 7–1 lists sites grouped by a proposed umbrella identification that were used in the trend analysis for different parameters but were located on the same unique identifier (ComID), matched with the same streamgage, and did not have any intervening influences, so they could be used together to compare trend results for different parameters. Trend results for nutrient, sediment, major ion, salinity, and carbon parameters, pesticides, and fish, invertebrates, and diatoms are presented in tables 7–2 through 7–4, respectively. In total, nearly 12,000 trends were evaluated in this study; there were 11,729 combinations of sites, parameters, and trend periods. In some cases, trends were evaluated for different parameters at colocated or nearby sites and may need to be compared for trend interpretation or causal analysis. In many cases, the start and end years of a trend period among parameters varied by a year. For comparing trends among different water-quality, pesticide, and ecology parameters the variations in start and end years of the four trend periods allowed by the screening criteria resulted in equivalent time periods. For example, a trend determined for the period of 1993–2011 was considered equivalent to a 1992–2012 trend and a 2003–11 trend was considered equivalent to a 2002–12 trend, and so on.

Nutrient, Sediment, Major Ion, Salinity, and Carbon Results

A summary of trend results for all water-quality parameters and trend periods for both flow-normalized concentration and flow-normalized load evaluated using the WRTDS model is presented in appendix 7, table 7–2; results of rejected models are not presented. Year-to-year variability in streamflow often results in changes in water quality; therefore, flow-normalized trends in load and concentration are better suited to illustrate how water quality has changed over time. Final trend sites for each parameter evaluated for at least one trend period using the WRTDS model are shown in figures 24–42. For a specific site, parameter, and trend period, a trend result of “true” indicates that the two-sided p-value generated in the output file of the EGRETci package for the WRTDS model is less than the critical alpha value of 0.1; therefore, the null hypothesis (there is no trend in the data) should be rejected, and there is a detectable trend in the data for the specified time period. Conversely, a trend result of “false” indicates that the null hypothesis cannot be rejected, and there is no statistically significant trend in the data for the specific time period at the critical alpha value. Complete trend results, including annual estimates of concentration and load, with and without flow normalization, are reported in De Cicco and others (2017). Note that likelihood coefficients also are reported for each trend result in De Cicco and others (2017).

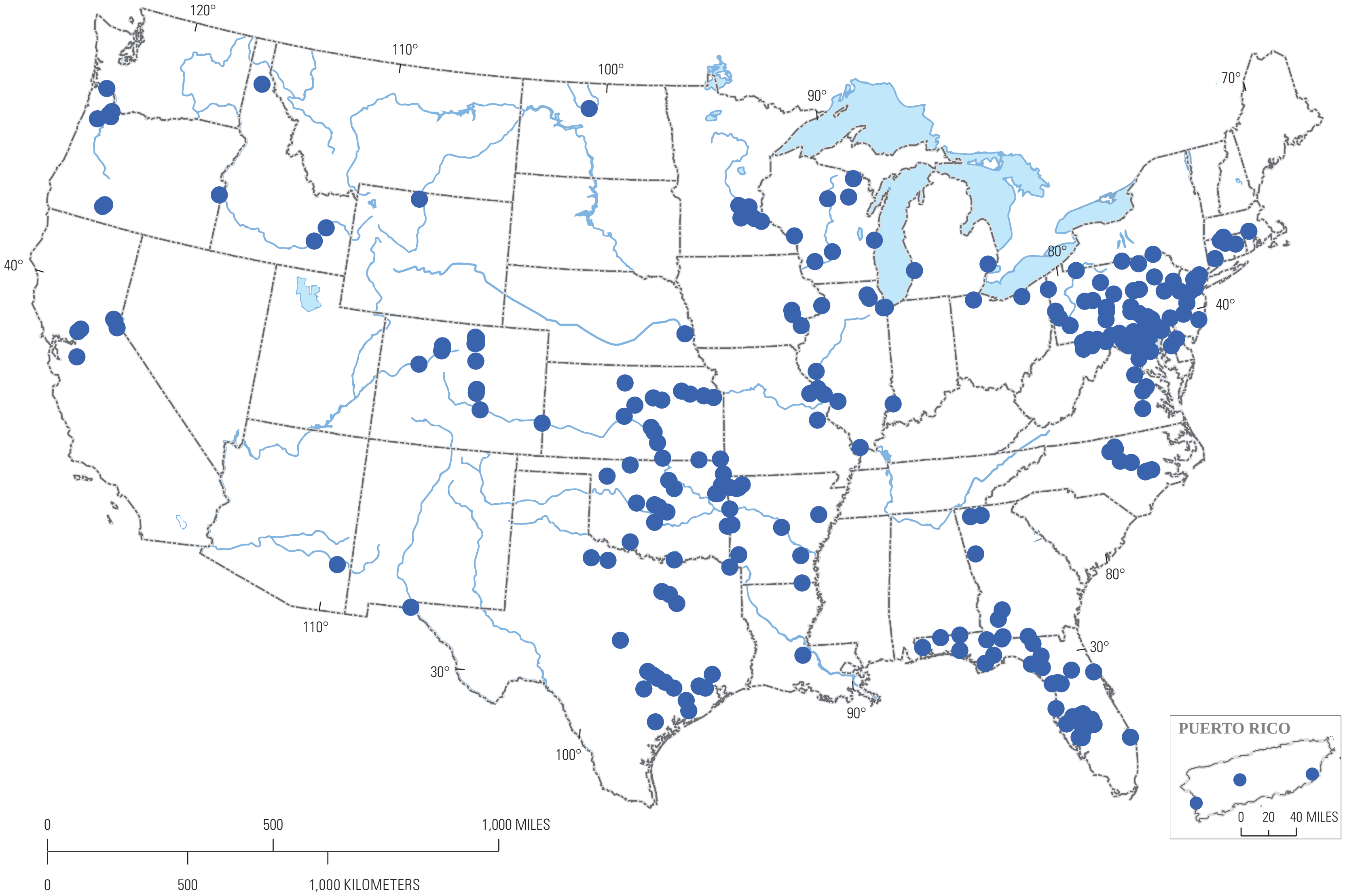


Figure 24. Final ammonia trend sites evaluated for at least one trend period.

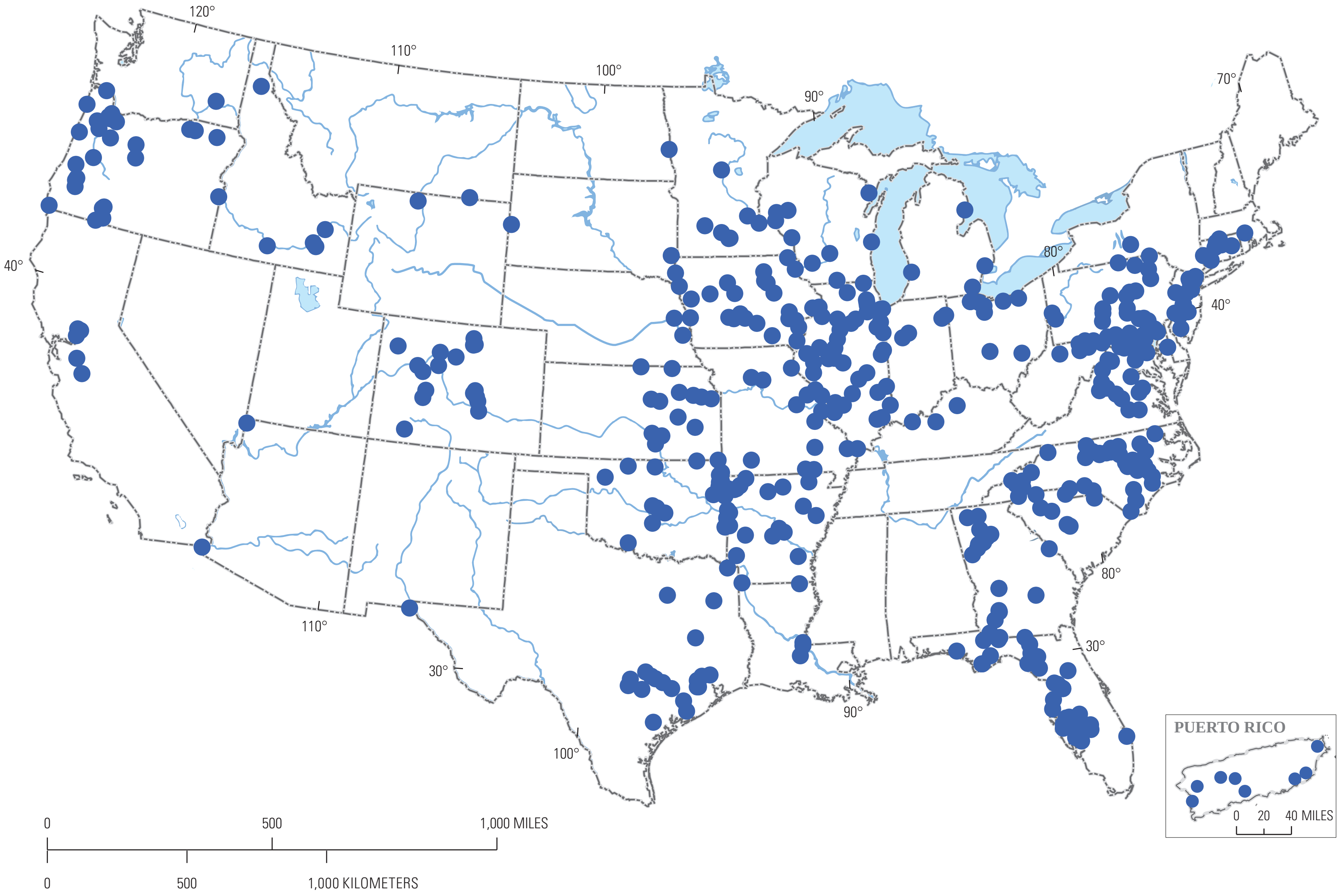


Figure 25. Final nitrate trend sites evaluated for at least one trend period.

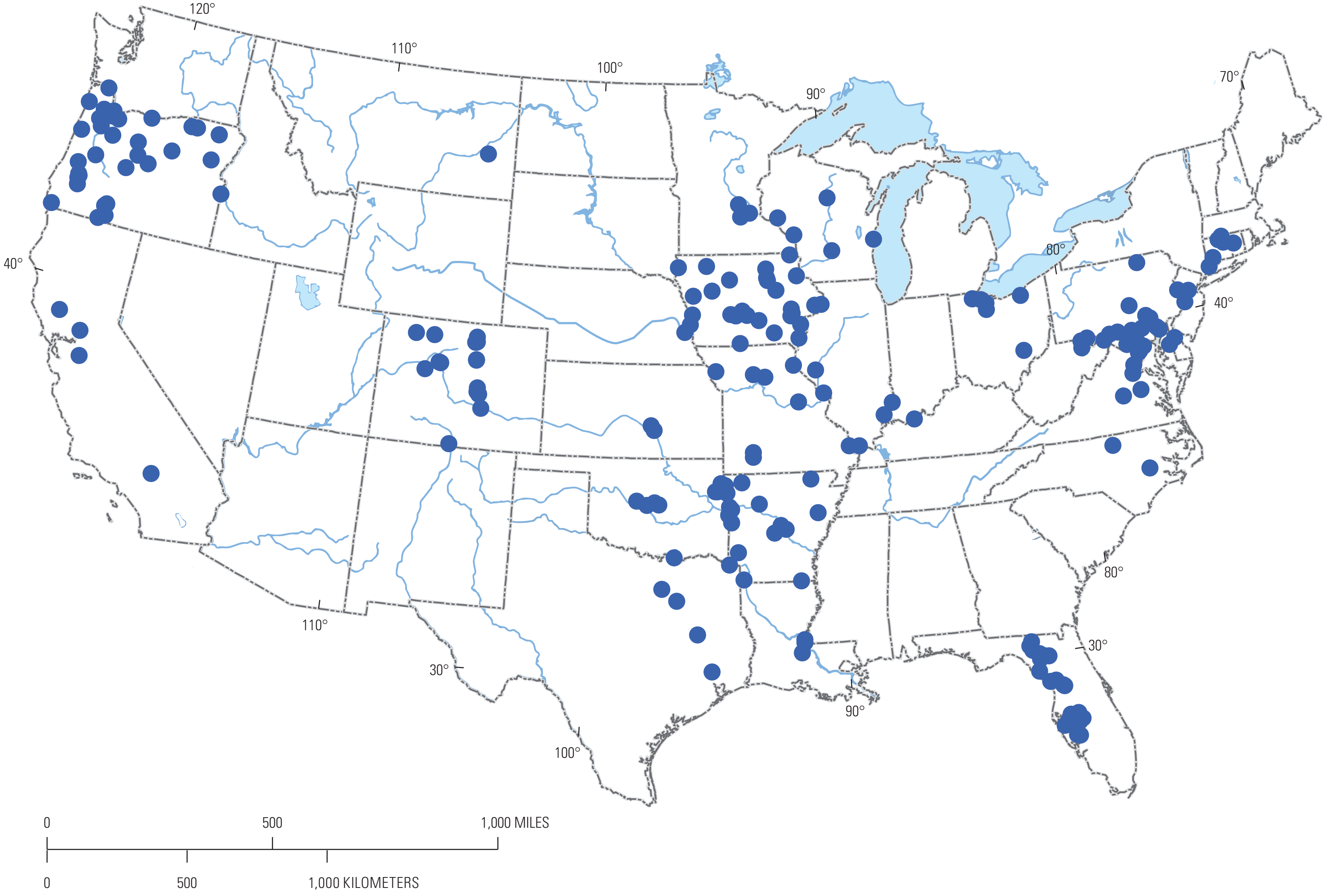


Figure 26. Final filtered orthophosphate trend sites evaluated for at least one trend period.

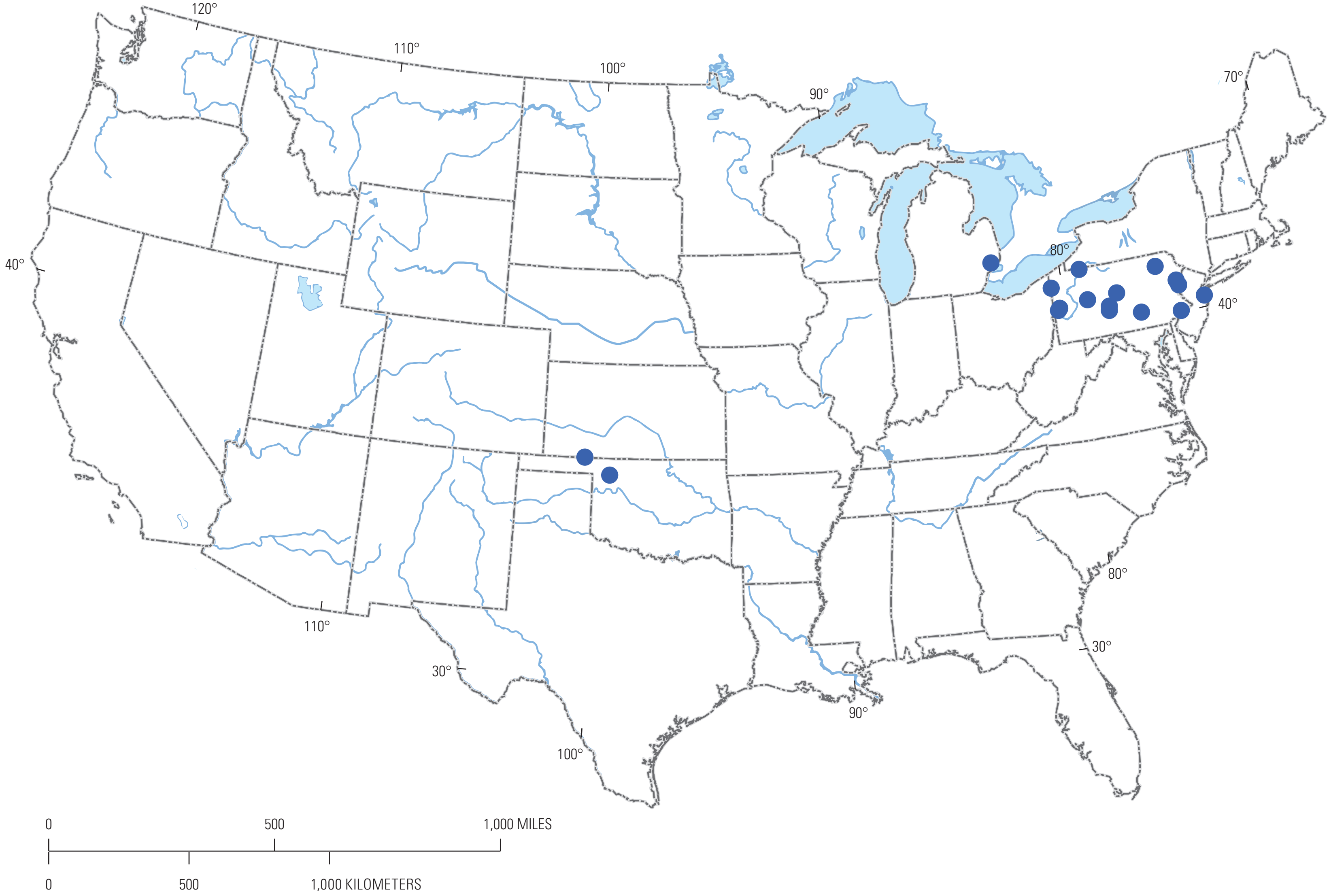


Figure 27. Final unfiltered orthophosphate trend sites evaluated for at least one trend period.

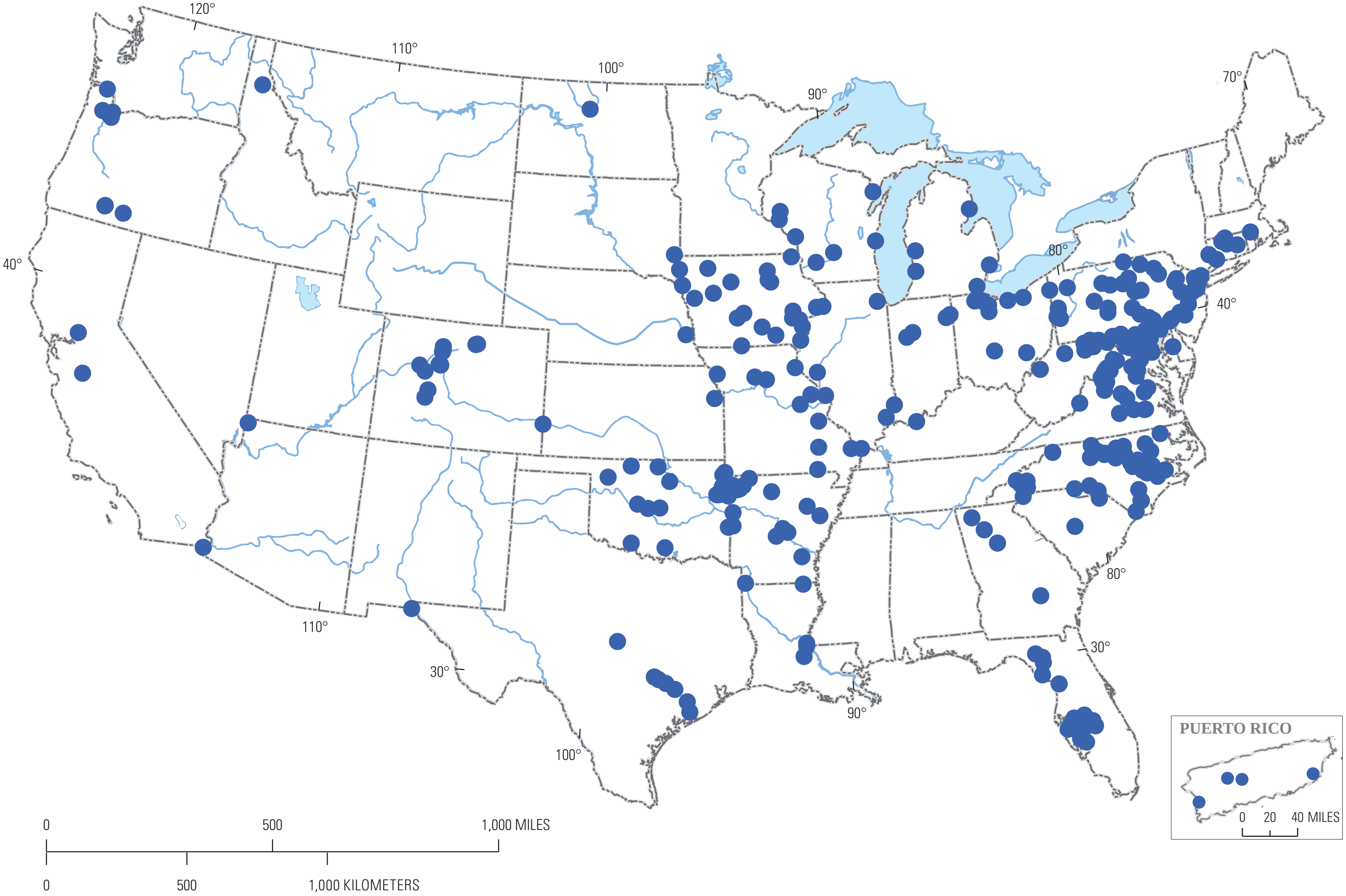


Figure 28. Final total nitrogen trend sites evaluated for at least one trend period.

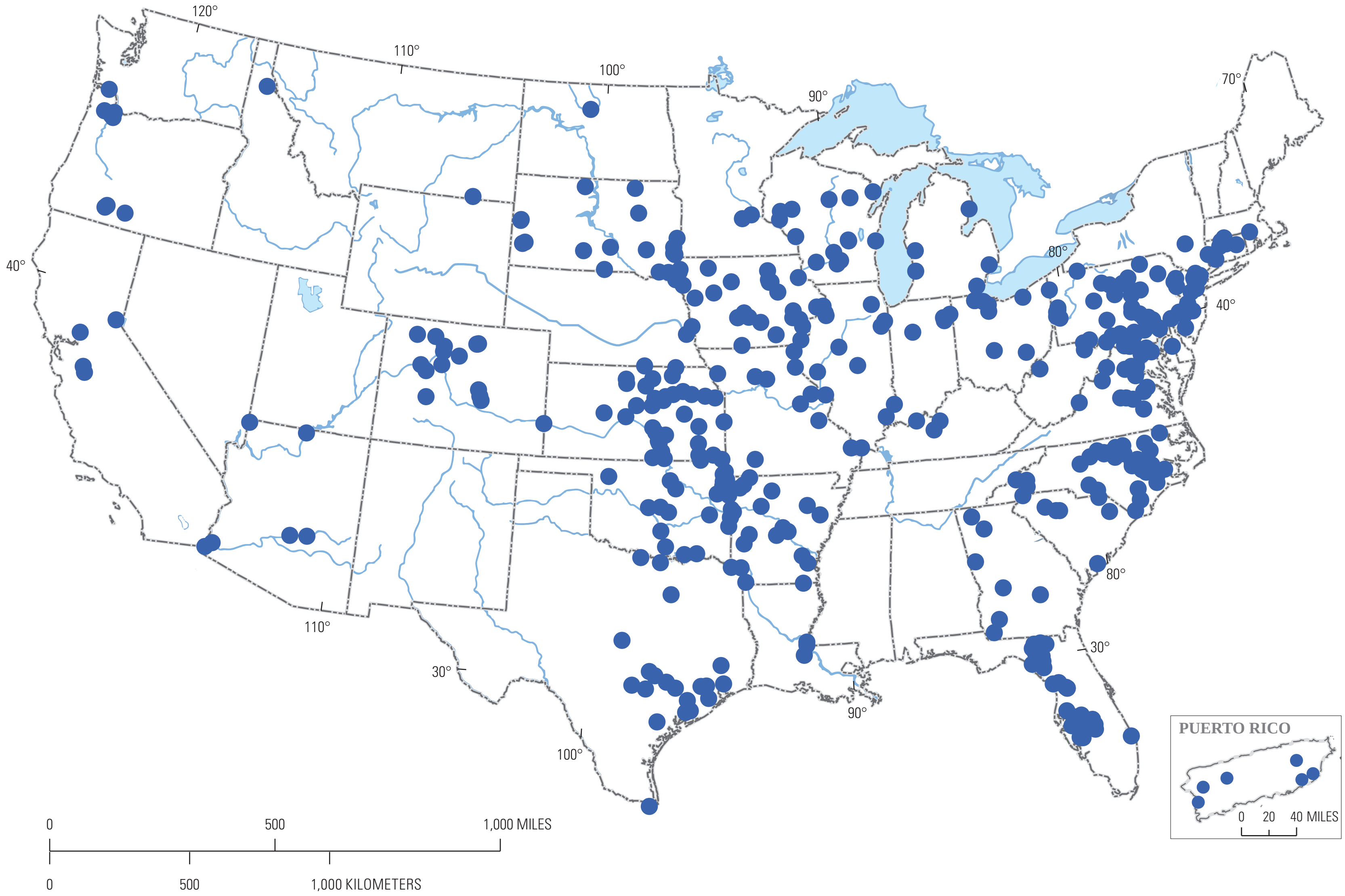


Figure 29. Final total phosphorus trend sites evaluated for at least one trend period.

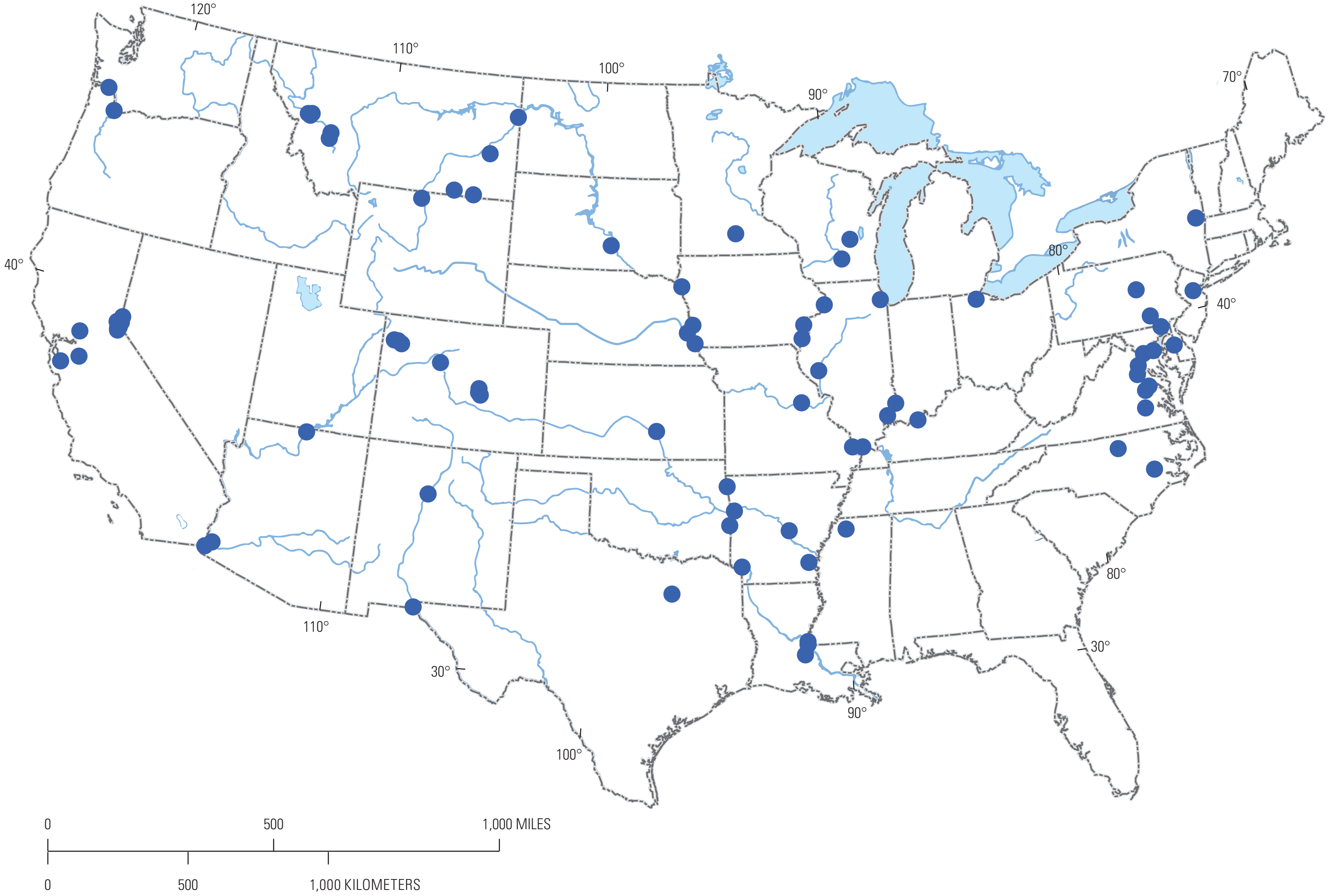


Figure 30. Final suspended sediment concentration trend sites evaluated for at least one trend period.

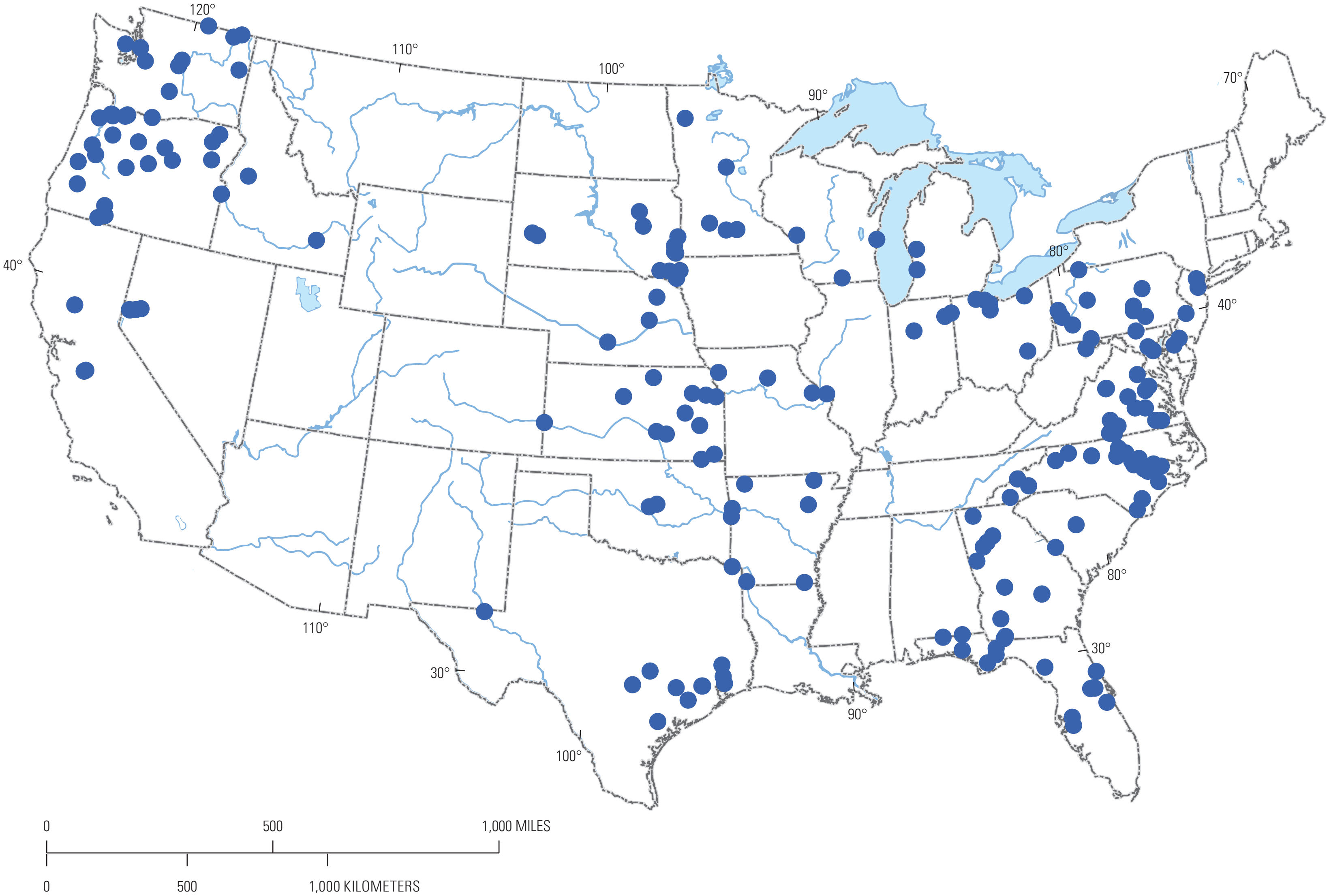


Figure 31. Final total suspended solids trend sites evaluated for at least one trend period.

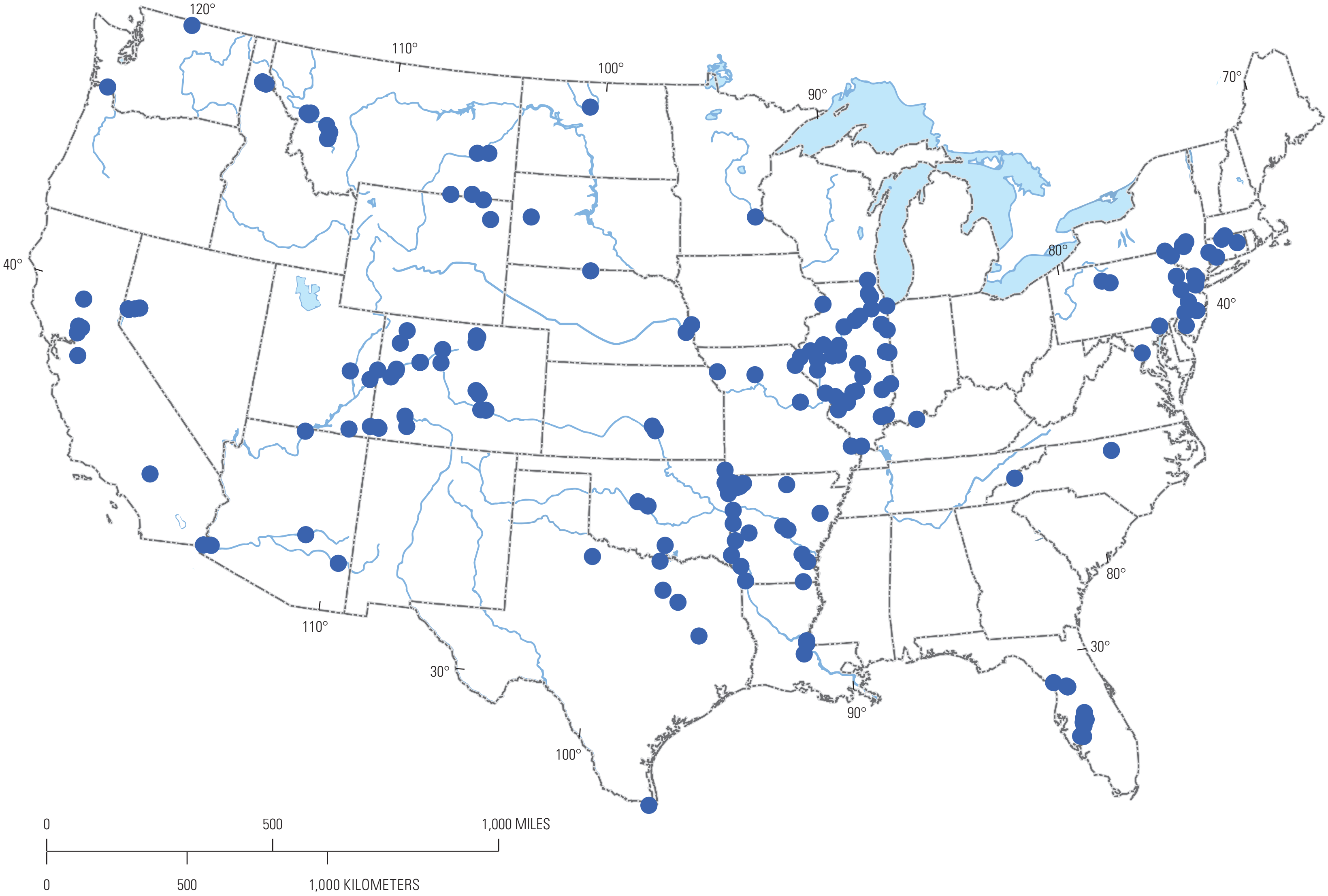


Figure 32. Final calcium trend sites evaluated for at least one trend period.

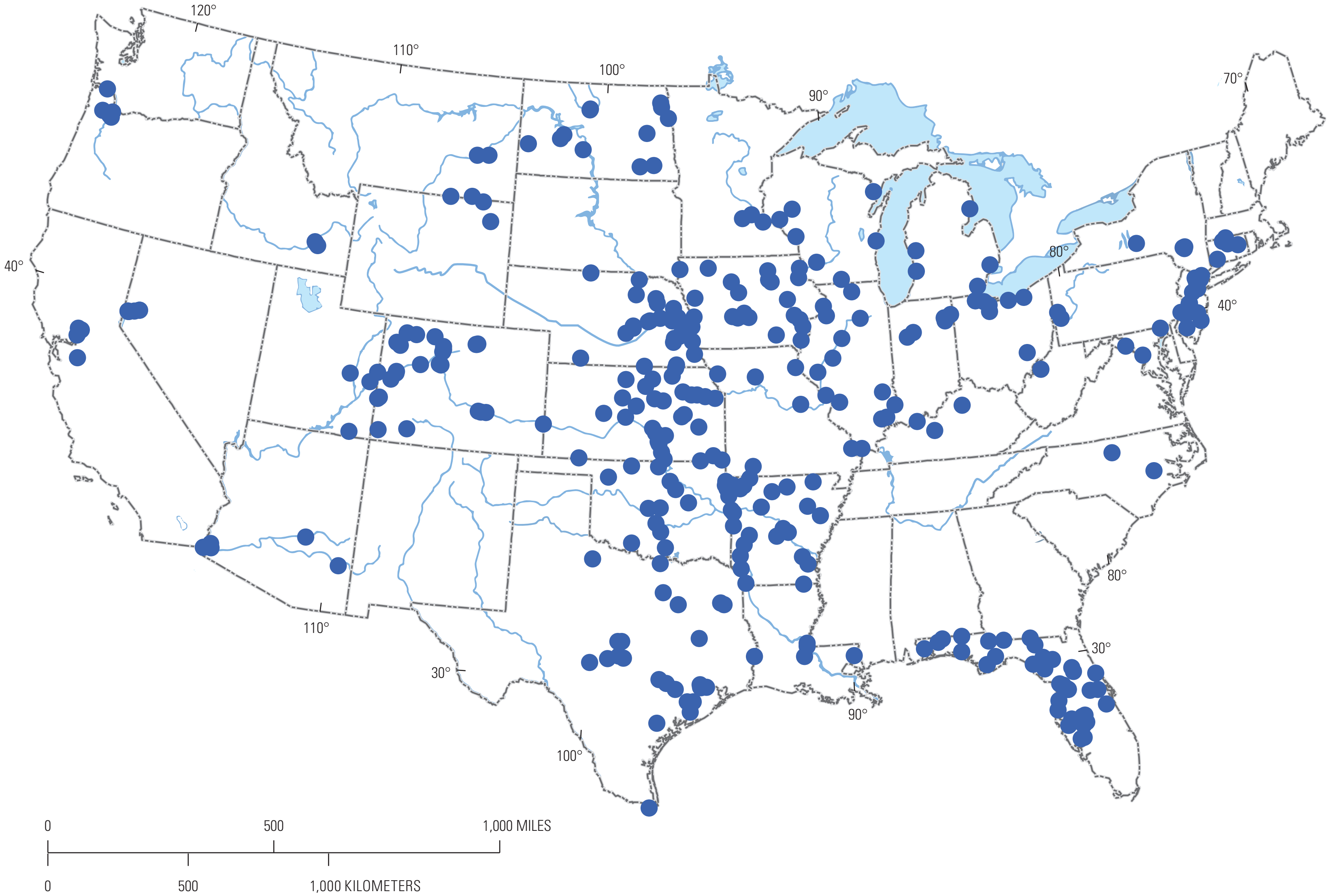


Figure 33. Final chloride trend sites evaluated for at least one trend period.

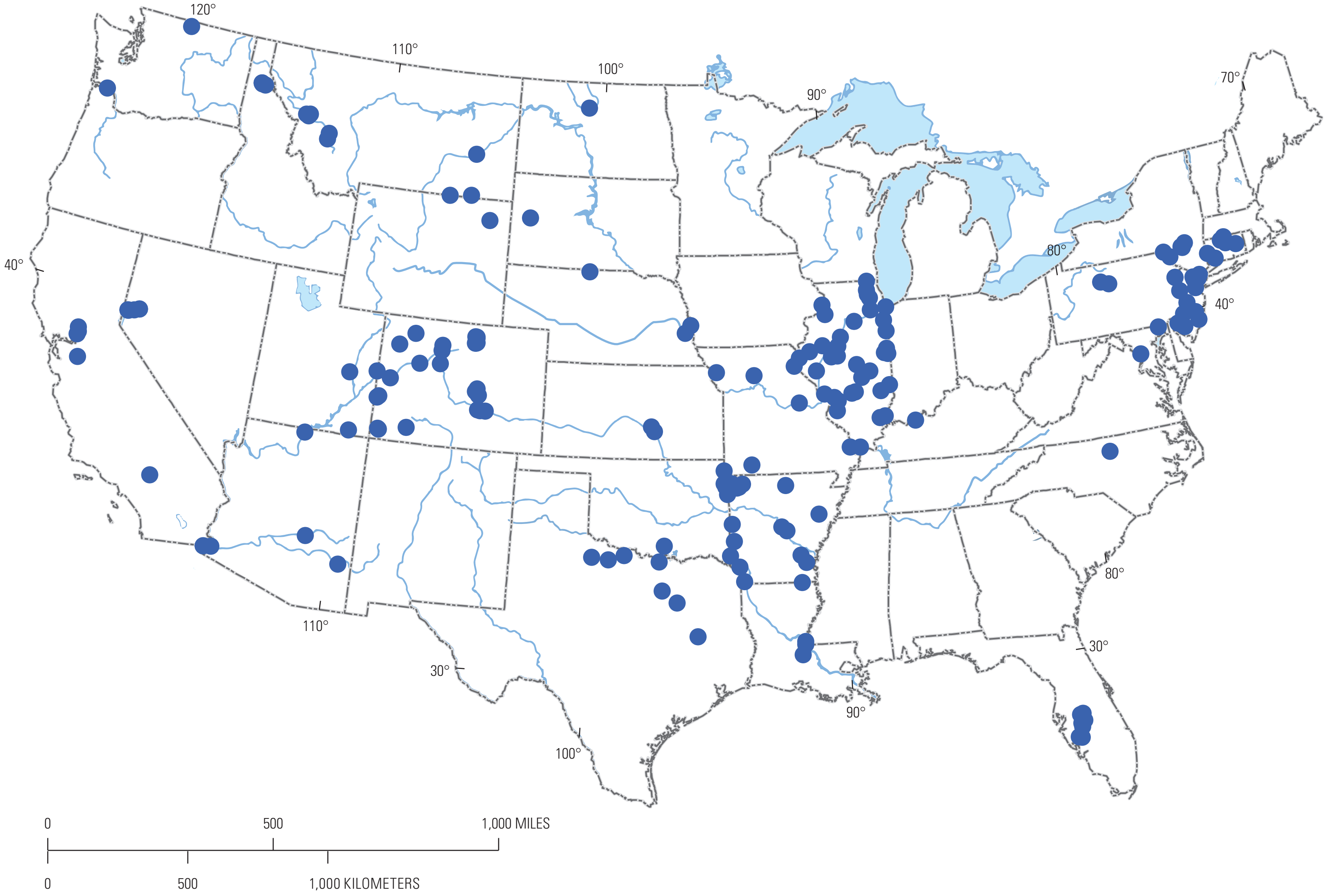


Figure 34. Final magnesium trend sites evaluated for at least one trend period.

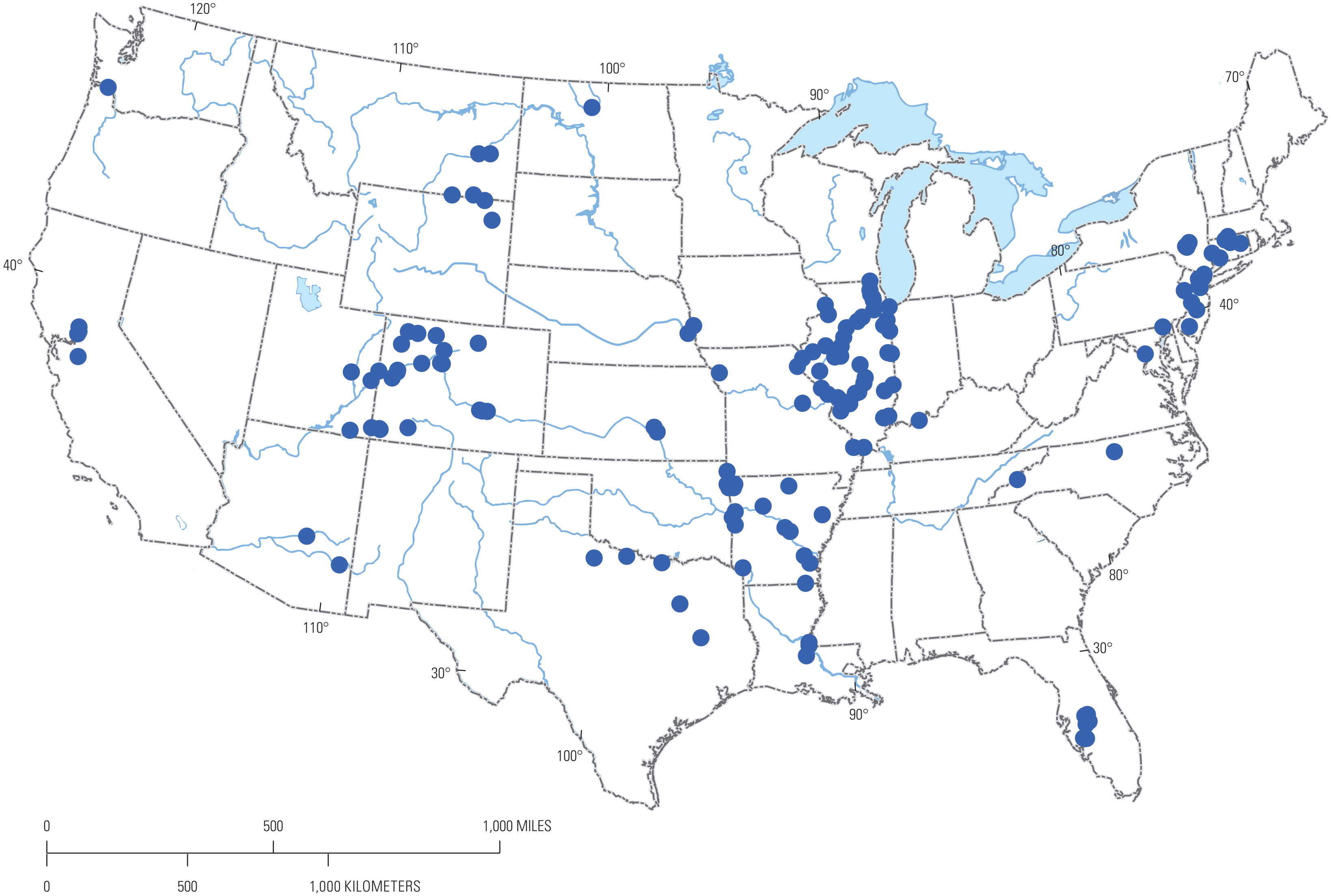


Figure 35. Final potassium trend evaluated for at least one trend period.

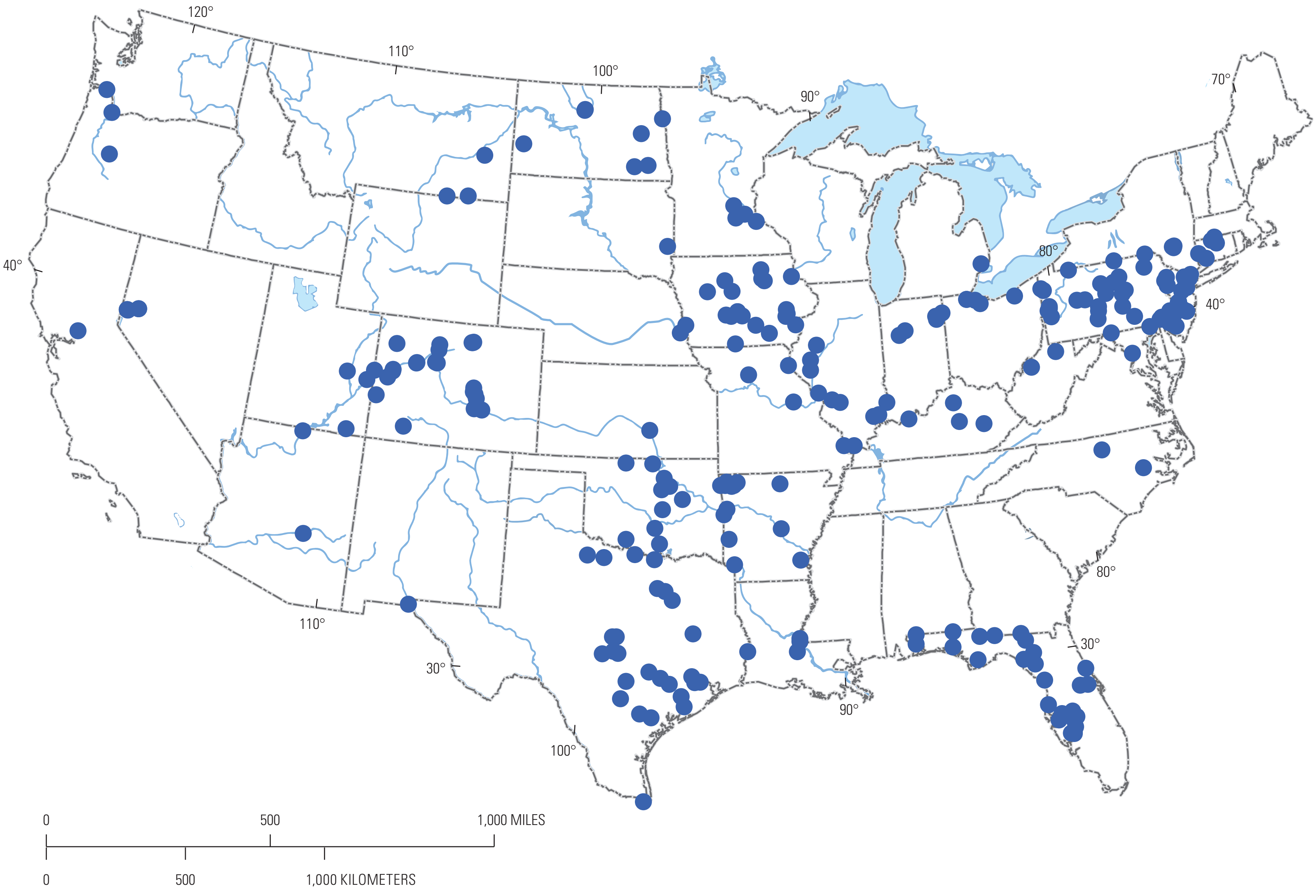


Figure 36. Final sulfate trend sites evaluated for at least one trend period.

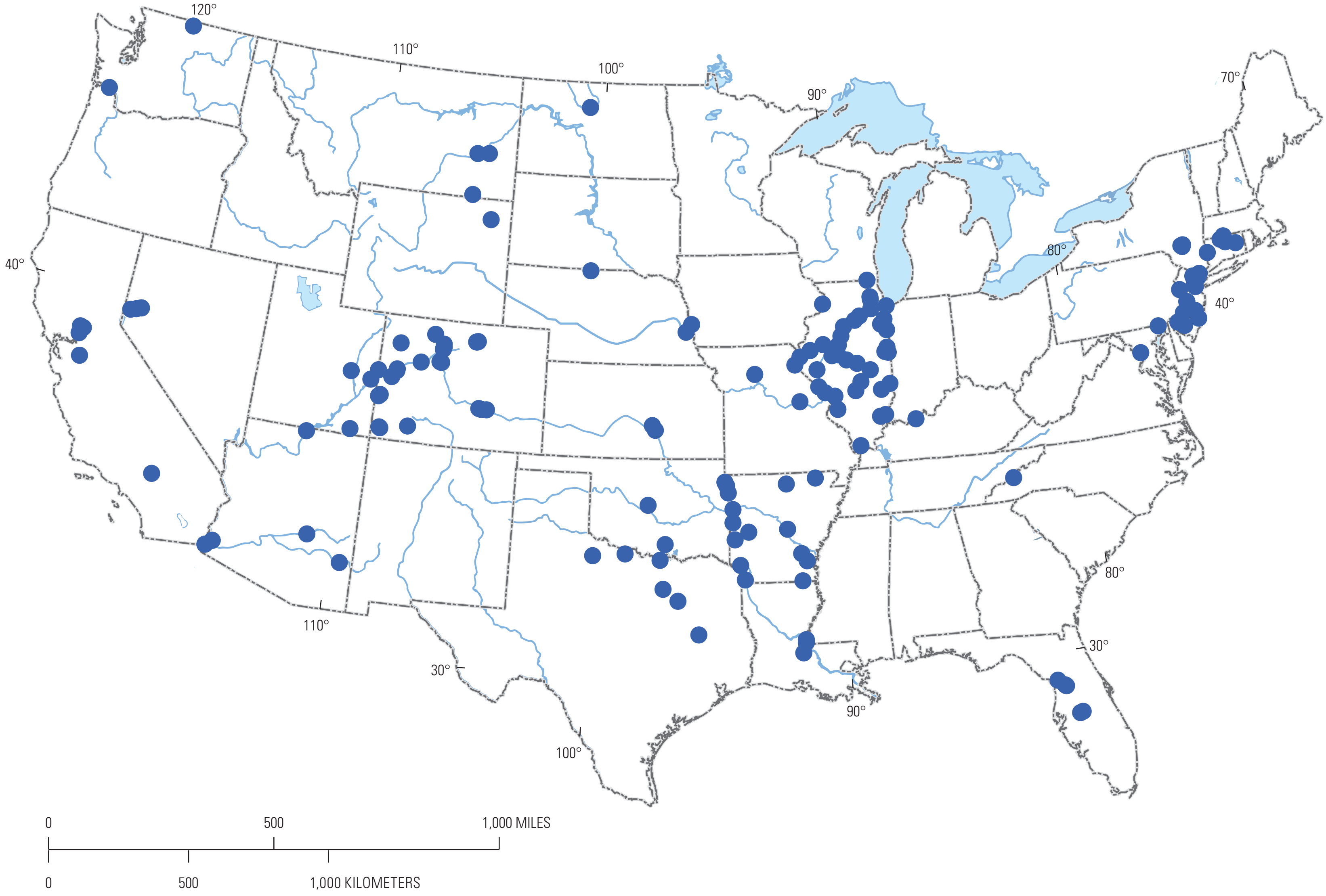


Figure 37. Final sodium trend sites evaluated for at least one trend period.

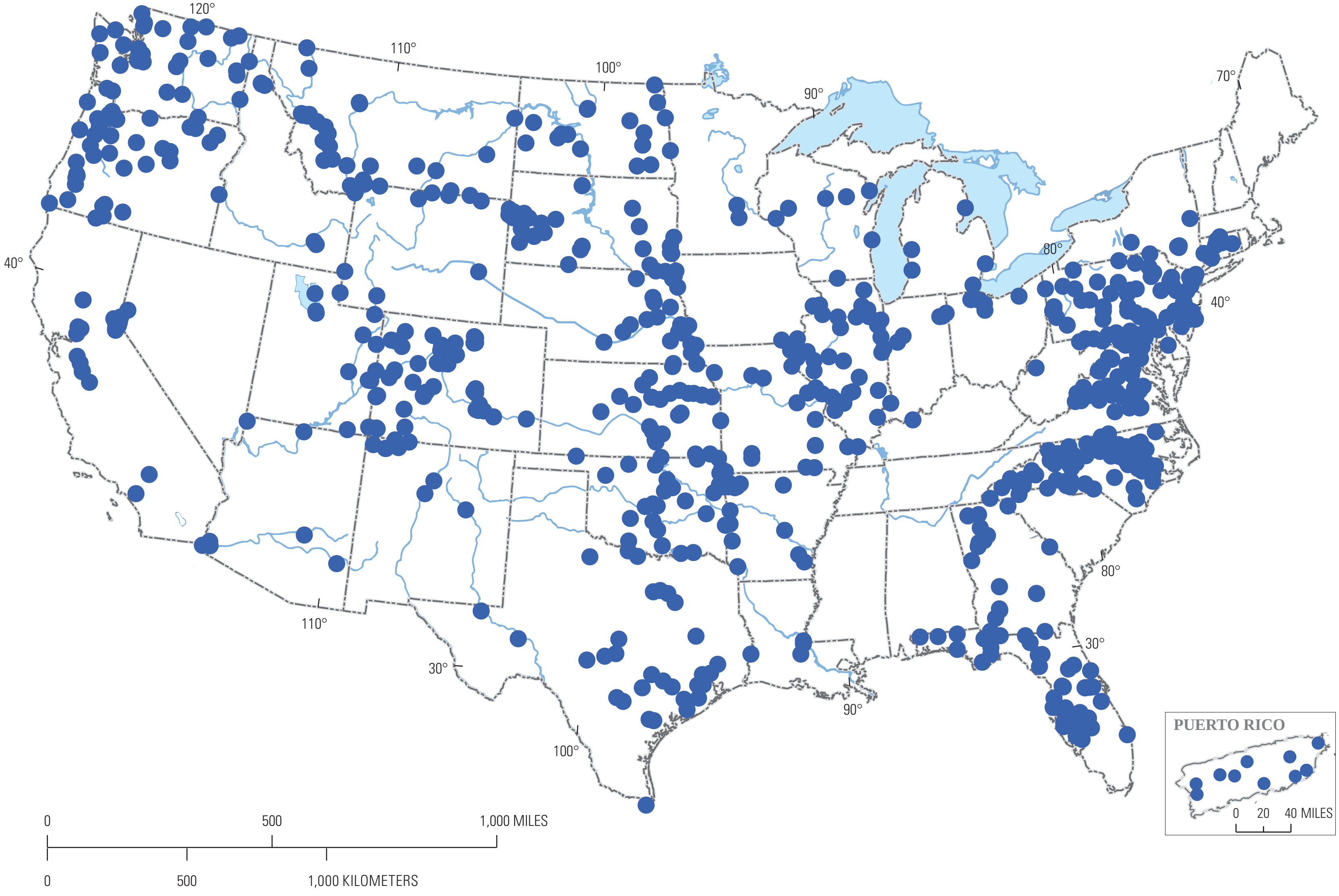


Figure 38. Final specific conductance trend sites evaluated for at least one trend period.

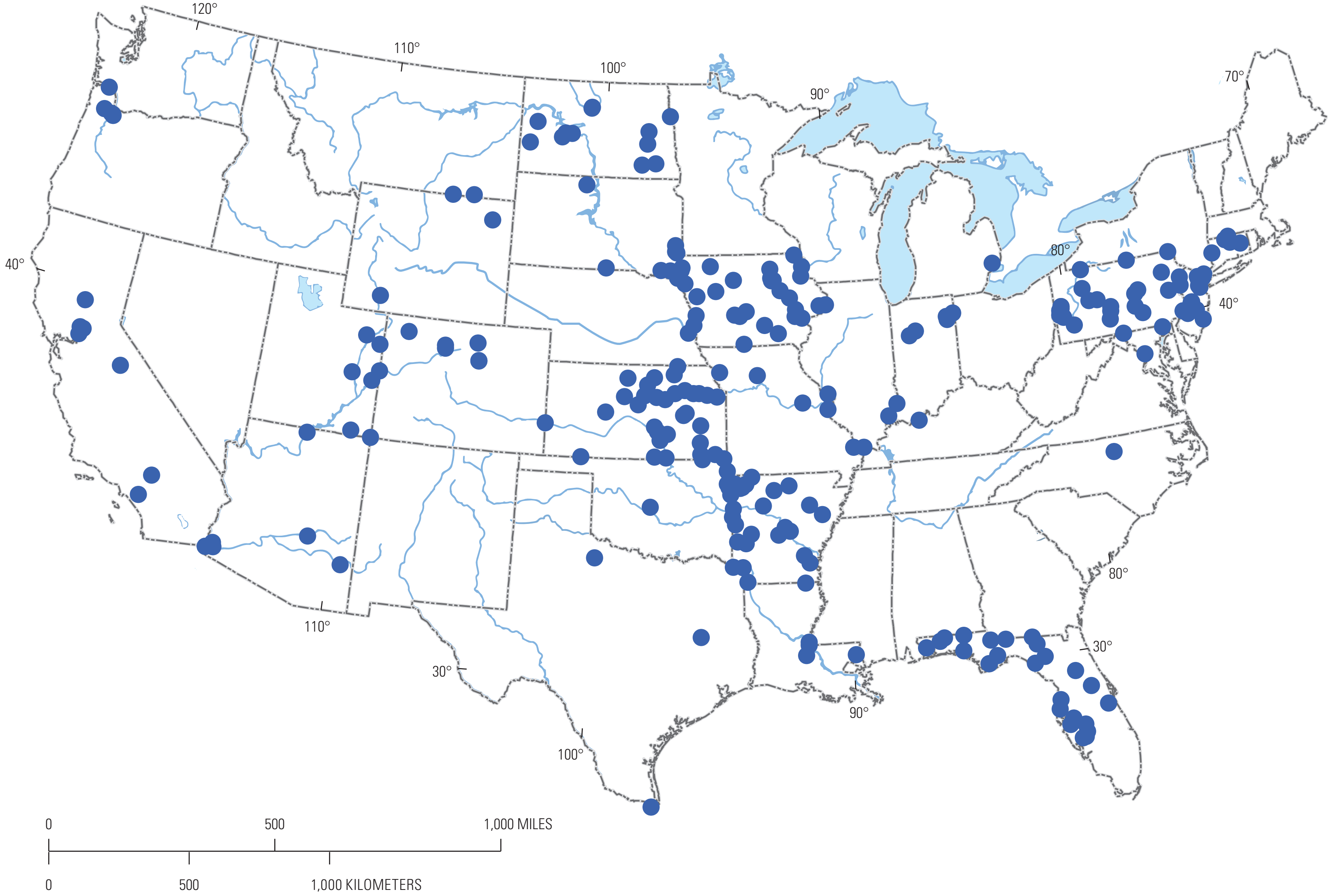


Figure 39. Final total dissolved solids trend sites evaluated for at least one trend period.

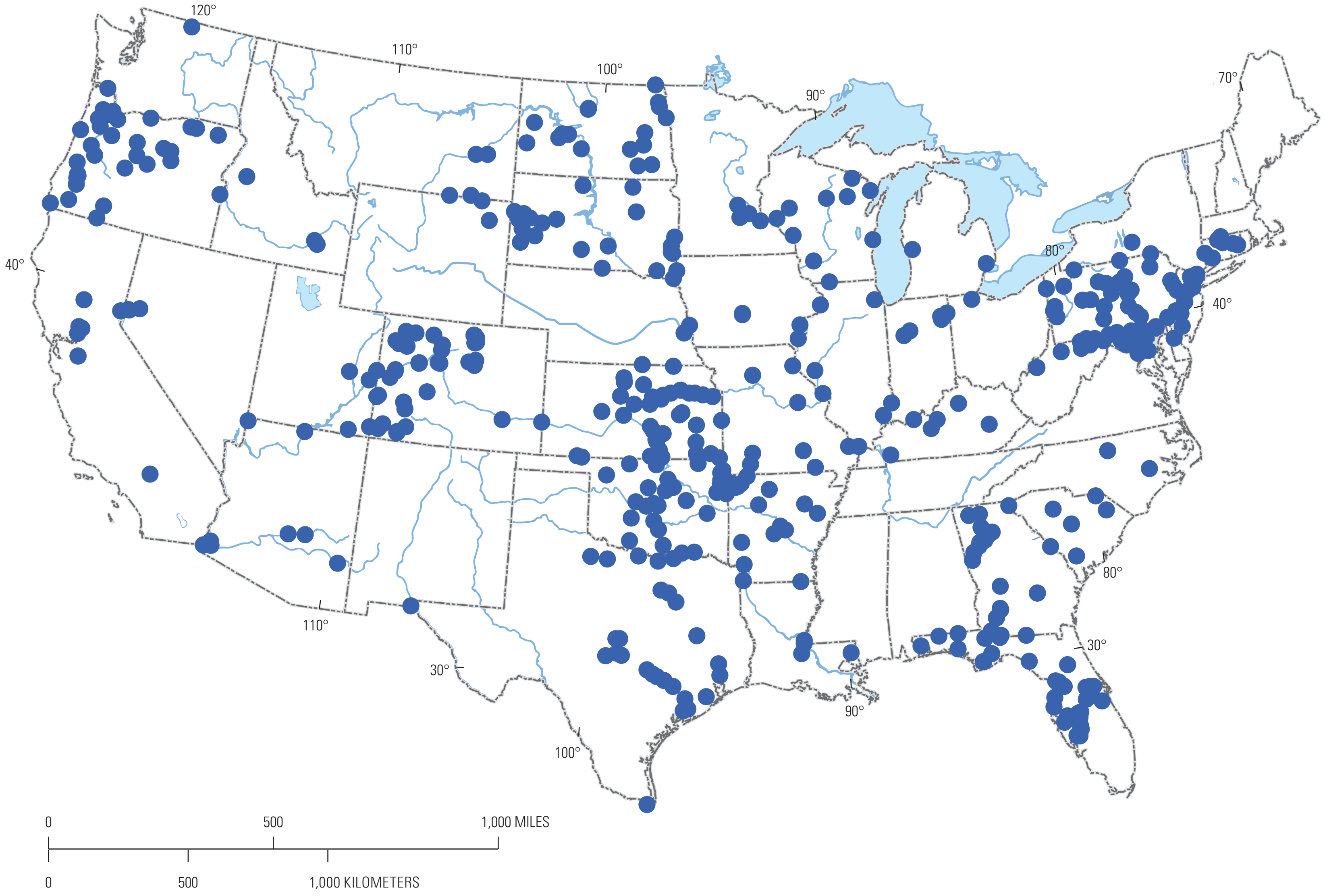


Figure 40. Final alkalinity trend sites evaluated for at least one trend period.

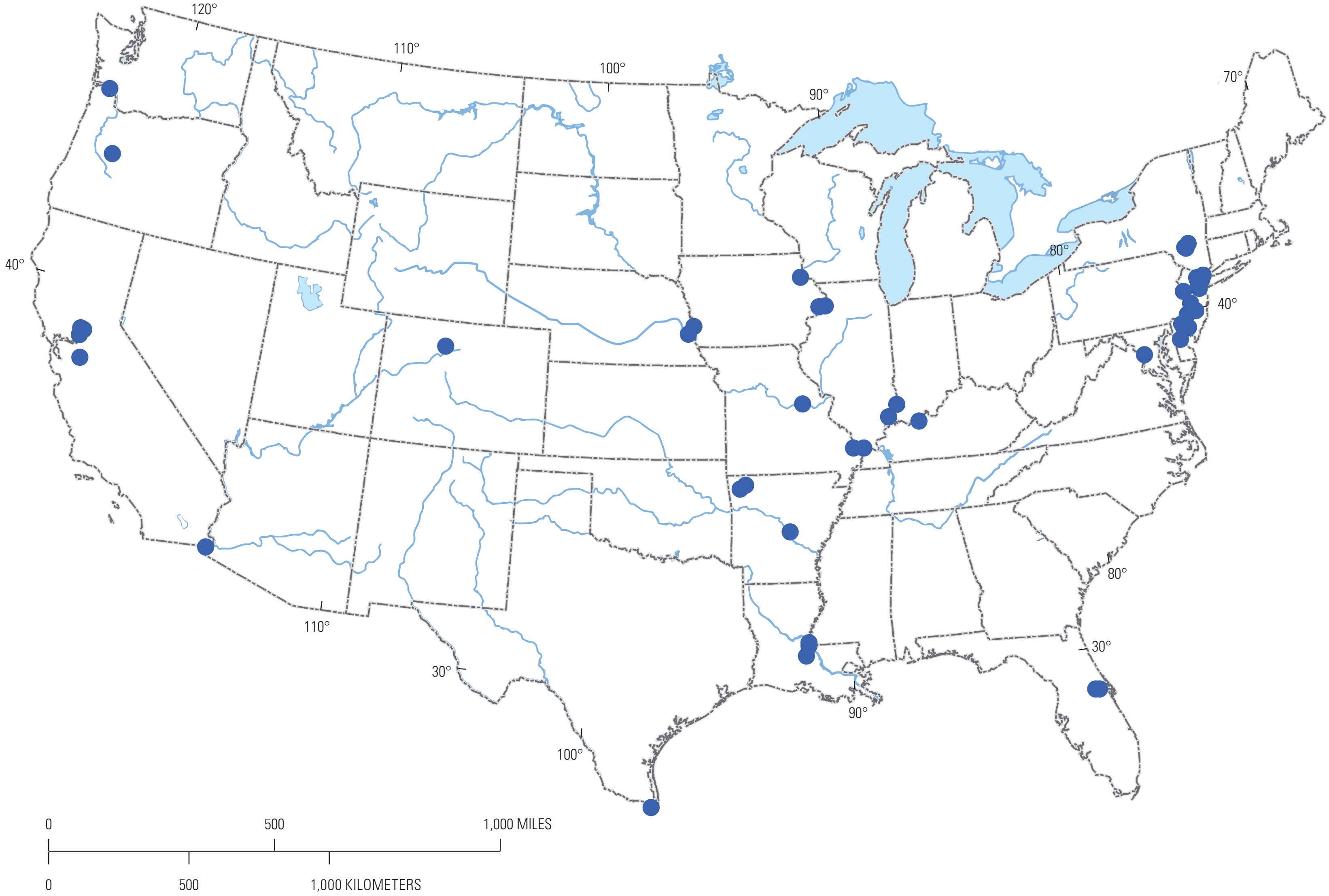


Figure 41. Final dissolved organic carbon trend sites evaluated for at least one trend period.

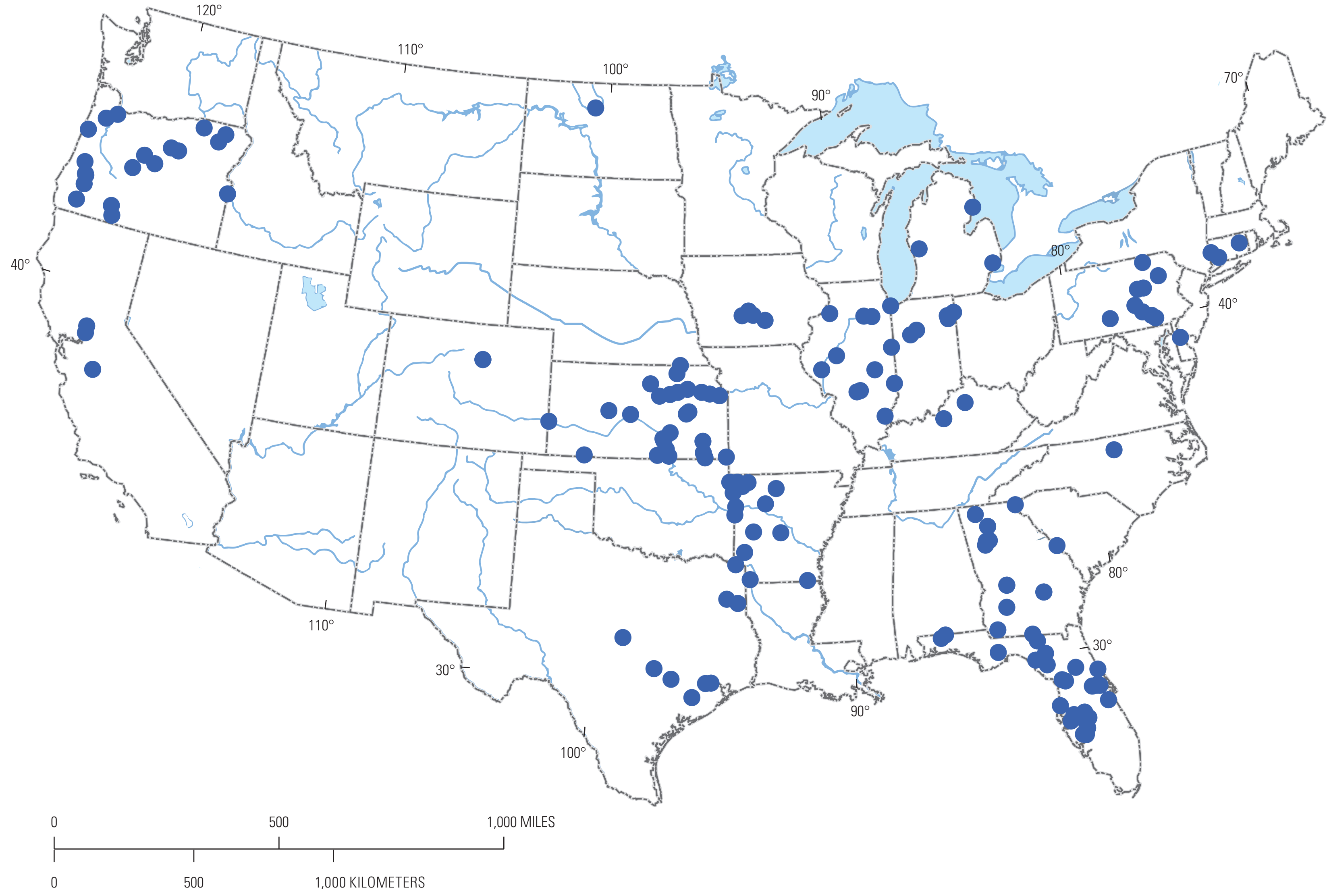


Figure 42. Final total organic carbon trend sites evaluated for at least one trend period.

The likelihood-based approach to reporting trend results is an alternative to the null-hypothesis significance testing approach. The likelihood-based approach has been used previously by the Intergovernmental Panel on Climate Change and is discussed at length in the context of water quality in Hirsch and others (2015). The likelihood-based approach gives more intuitive information on the certainty of a trend estimate and avoids the arbitrary classifications imposed by selecting an alpha value to assess significance. (For example, if the alpha value is selected to be 0.1, trend A with a p-value of 0.099 will be classified as significant, while trend B with a p-value of 0.11 will be classified as nonsignificant. In reality, there is no meaningful difference between these two p-values.) A complete discussion about the results provided by the WRTDS model and how to explore and interpret the results is included in Hirsch and De Cicco (2015) and Hirsch and others (2015). Two examples of WRTDS trend results from this study are provided below using annual results and trend period results.

At site 01377000, Hackensack River at Rivervale, New Jersey, mean annual sulfate concentrations and flow-normalized sulfate concentrations are clearly decreasing over time (fig. 43*A*). Similarly, estimates of annual sulfate load and flow-normalized sulfate loads are decreasing over time (fig. 43*B*). Sulfate trends were evaluated for the 1982–2012, 1992–2012, and 2002–12 time periods at the site. All three trend periods were determined to have significant decreasing trends in both flow-normalized concentration and flow-normalized load.

Total phosphorus at site 02089500, the Neuse River at Kinston, North Carolina, is a good example of how trend results can be different for concentrations than for loads, and how trend results can be different depending on the trend period. Total phosphorus concentrations and loads follow a similar pattern to each other but do not behave consistently over time (figs. 44*A* and 44*B*). Mean annual total phosphorus concentrations and flow-normalized total phosphorus concentrations decrease during the beginning of the 1982–2012 trend period and then increase towards the end of the 1982–2012 trend period (fig. 44*A*). The trend results indicate a significant decreasing trend in total phosphorus concentrations for 1982–2012 and a significant increasing trend in total phosphorus concentration for 2002–12; there is no significant trend in total phosphorus concentration for 1992–2012. Similar to the concentrations, estimates of annual total phosphorus load and flow-normalized total phosphorus load decrease during the beginning of the 1982–2012 trend period and then increase slightly towards the end of the 1982–2012 trend period (fig. 44*B*). The trend results indicate a significant decreasing trend in total phosphorus load for 1982–2012; there is no significant trend in total phosphorus loads for 1992–2012 or 2002–12.

A discussion of the SKT and the results of the comparison of trend results from the SKT and WRTDS are described in appendix 8. Water-quality and streamflow datasets used in the Seasonal Kendall trend tests and complete results for the SKT are in Mills and others (2017).

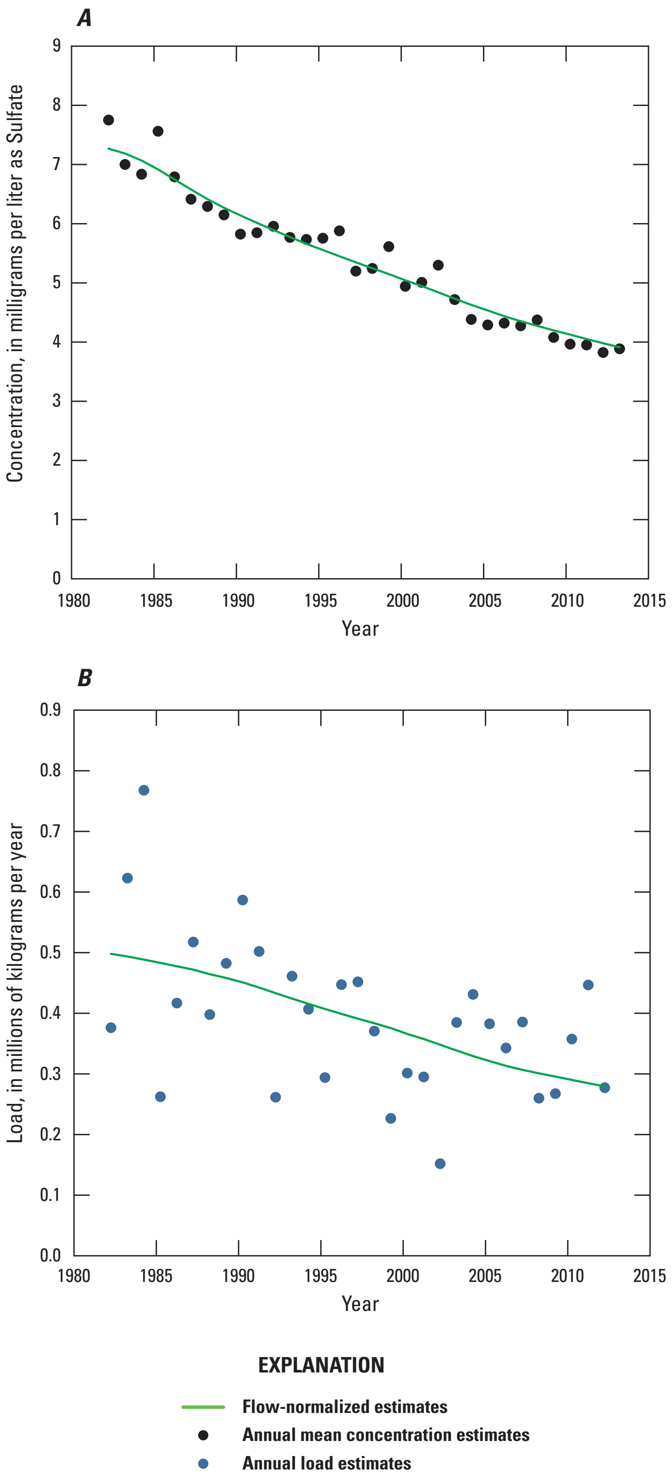


Figure 43. The Weighted Regressions on Time, Discharge, and Season (WRTDS) estimates of *A*, annual mean sulfate concentrations and flow-normalized sulfate concentrations; and *B*, annual sulfate loads and flow-normalized sulfate loads for site 01377000 Hackensack River at Rivervale, New Jersey.

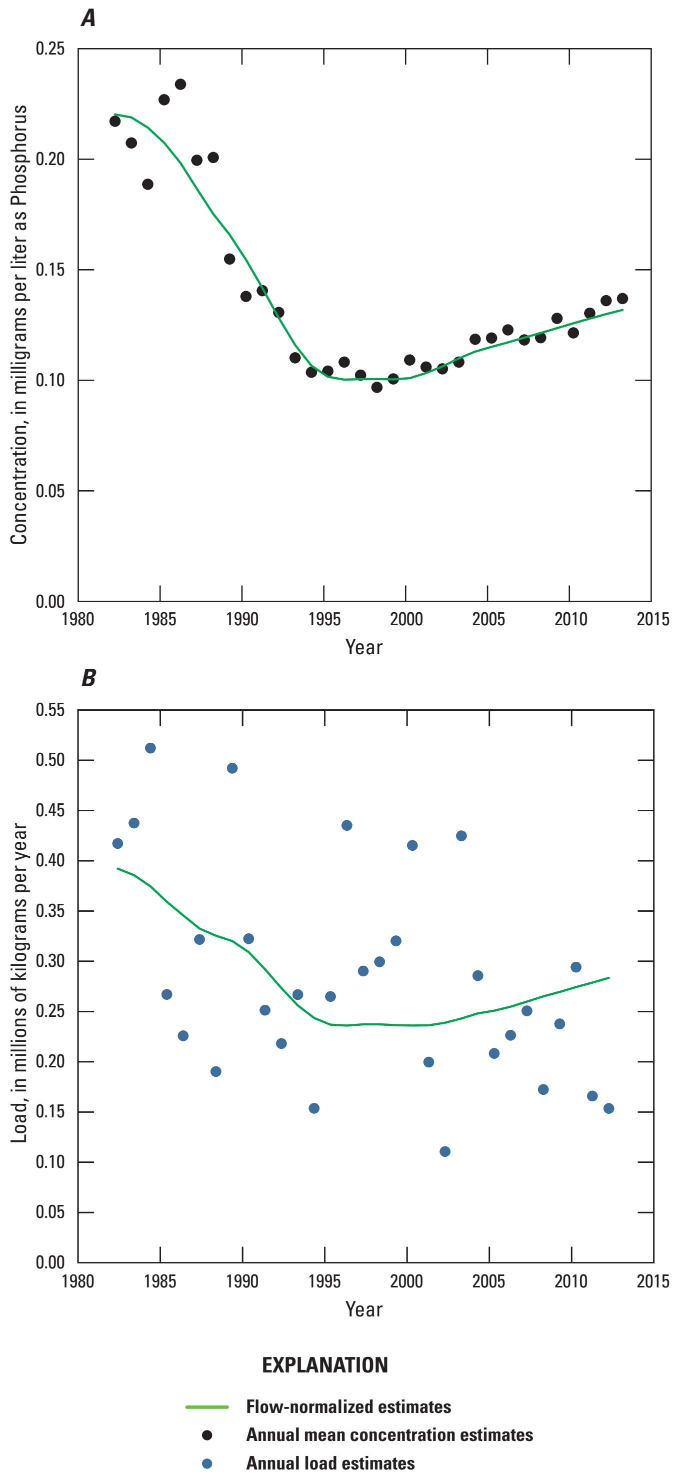


Figure 44. The Weighted Regressions on Time, Discharge, and Season (WRTDS) estimates of *A*, annual mean total phosphorus concentrations and flow-normalized total phosphorus concentrations; and *B*, annual total phosphorus loads and flow-normalized total phosphorus loads for site 02089500 Neuse River at Kinston, North Carolina.

Pesticide Results

A summary of trend results for all pesticide parameters and trend periods for both flow-adjusted concentration and load evaluated using the SEAWAVE–Q model is presented in appendix 7, table 7–3. Final pesticide trend sites evaluated using the SEAWAVE–Q model are shown in figure 45 (the pesticides for which trends are reported vary from site to site, depending on the degree of data censoring). Complete trend results for pesticides, including annual concentrations and load estimates, are reported in Ryberg and others (2017). The trends in pesticide loads are the same as the trends in pesticide concentrations, under an assumption that there are no trends in streamflow (see "Pesticide Trend Method" section).

To illustrate the SEAWAVE–Q model, trend analysis results for concentrations of one pesticide, atrazine (USGS parameter code 39632) at a single site during the 2002–12 analysis period, site 03303280 Ohio River at Cannelton Dam at Cannelton, Indiana, are shown in figure 46. Atrazine is a preemergent or postemergent herbicide used to control broadleaf weeds and some grassy weeds (U.S. Environmental Protection Agency, 2006) and is one of the most widely used agricultural pesticides in the United States, with use declining in the Ohio River Basin upstream from this site (a significant trend of -2.1 percent per year in the period 2001–10; Ryberg and others, 2014). The SEAWAVE–Q results show a statistically significant downward trend for atrazine concentration (-8.6 percent per year), which agrees well with atrazine concentration trend analysis at this same site for the period 2001–10 in a previous USGS study (a significant trend of -6.7 percent per year; Ryberg and others, 2014).

The measured and fitted (estimated) atrazine concentrations (fig. 46) show patterns of seasonality (the within-year oscillations correspond to seasonal applications of atrazine that peak in summer), trend (a general decline), and flow-related variability (the seasonality and trend terms are smooth, whereas the flow-related variability causes the smaller variations in the fitted concentrations, as well as contributing to the overall trend). More in-depth examples of how streamflow anomalies and the fitted seasonal wave interact in SEAWAVE–Q models are provided for three pesticides in Ryberg and others (2010). Visualizations of the seasonal wave term are provided in appendix 3 of Ryberg and Vecchia (2013).

Ecology Results

Trend results for all three assemblages (fish, invertebrates, and diatom) for both trend periods are reported in tabular format in a single table (Zuellig and Riskin, 2017) with site locations illustrated in figure 47. A summary of the results are reported in appendix 7, table 7–4, where trends are indicated by "true" or "false" for each site, trend-period, and ecological endpoint combination for both adjusted and unadjusted data.

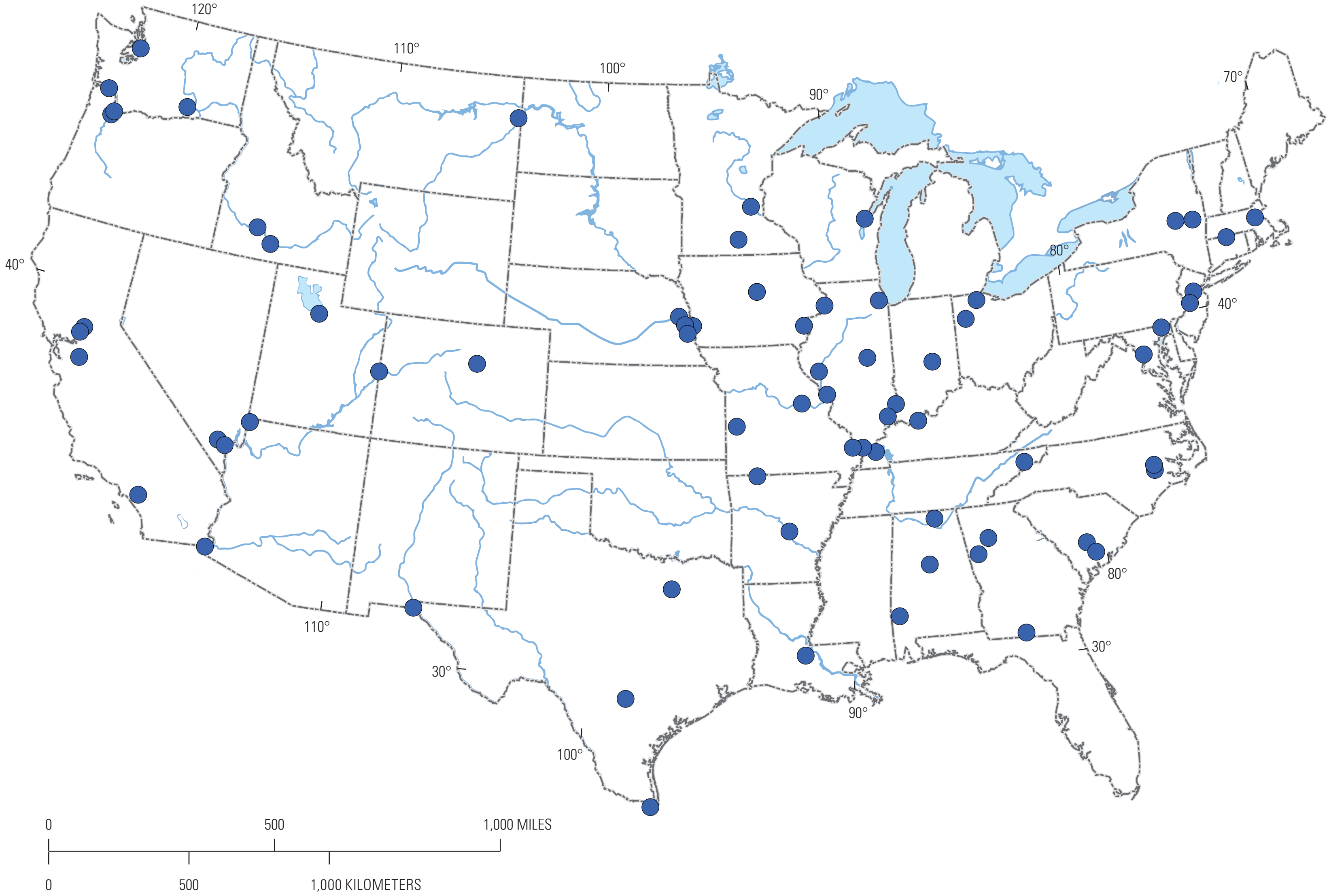


Figure 45. Final pesticide trend sites evaluated for at least one trend period.

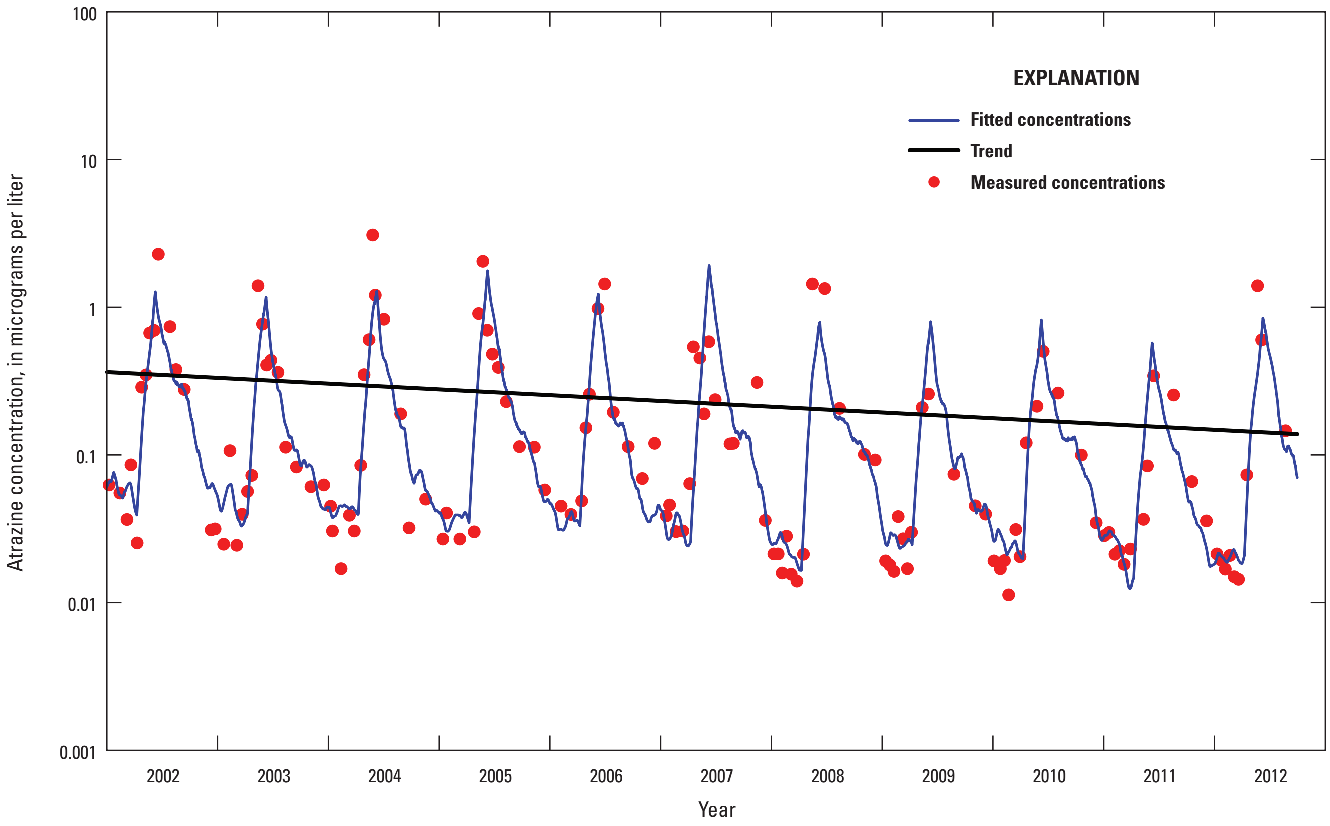


Figure 46. Measured and fitted atrazine concentration and trend line, site 03303280 Ohio River at Cannelton Dam at Cannelton, Indiana.

A trend result of true signifies the p-value was ≤ the critical value of 0.1, indicating the null-hypothesis (no trend in the data) was rejected. A false result indicates the null-hypothesis was not rejected, signifying no trend in the data (p-value > 0.1). Likelihood coefficients could not be calculated for the ecology trends determined using the Kendall-tau test (as was done for the nutrient, sediment, major ion, salinity and carbon trends determined using WRTDS) because of the limitations of the ecology data. Therefore, caution should be used when interpreting p-values near the critical significance threshold ($p \leq 0.1$) as there is likely little biological difference between trends with p-values just above and below any threshold. Complete results for diatom and fish endpoints at a single site (07053250, Yocum Creek near Oak Grove, Arkansas) are reported in appendix 7, table 7–5, with two examples (diatom endpoints 1993–2012 and fish endpoints 2002–2012) briefly described below.

Diatom Trend Result Example

At site 07053250 Yocum Creek near Oak Grove Arkansas, diatom taxa preferring chloride concentrations below 15 mg/L (DiaLowCl) decreased over the 1993–2012 trend period (pO = 0.025, Per.ChangeO ~ -30, appendix 7, table 7–5). After endpoints were adjusted for antecedent flow, the DiaLowCl result remained similar (pR = 0.004, Per. ChangeR ~ -29); however, taxa tolerant to high total nitrogen (DiaHighTN, pR = 0.049, Per.ChangeR ~ 63) and phosphorus (DiaHighTP, pR = 0.063, Per.ChangeR ~ 69) concentrations, as well as low dissolved oxygen (DiaLowDO, pR = 0.006, Per. ChangeR ~ 223) increased over the trend period. In this case, antecedent streamflow conditions may have been masking the response of taxa tolerant to high nutrients (>3 mg/L total nitrogen and > 10 µg/L total phosphorus) and low dissolved oxygen (down to 30 percent saturation).

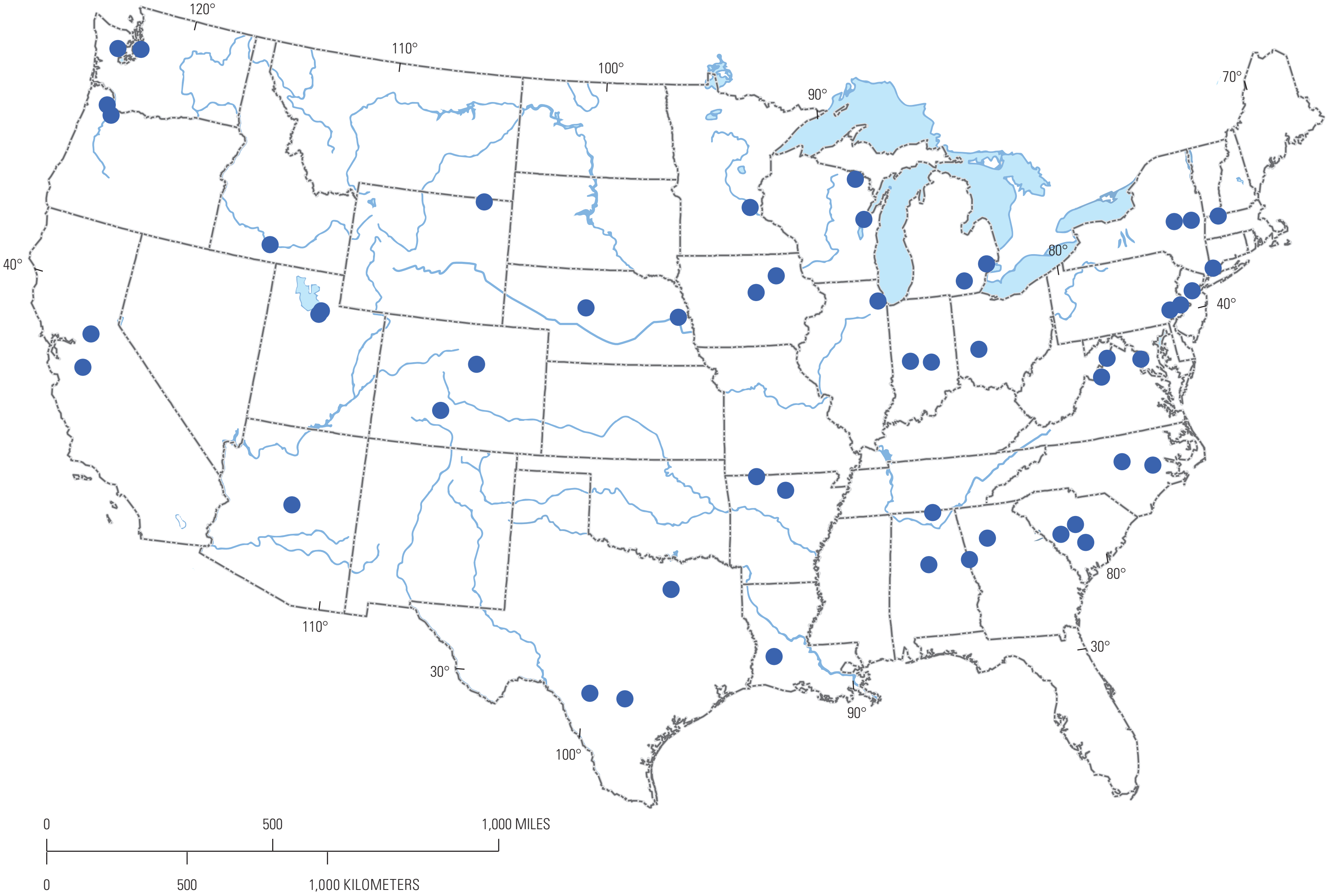


Figure 47. Final ecology trend sites evaluated for at least one trend period.

Fish Trend Result Example

At site 07053250 Yocum Creek near Oak Grove, Arkansas, fish diversity (FishDiv) decreased (pO = 0.009, Per.ChangeO ~ -25) along with community tolerance to ammonia (FishAmm, pO = 0.013, Per.ChangeO ~ -13) and low dissolved oxygen (FishDO, pO = 0.035, Per.ChangeO ~ -10) (appendix 7, table 7–5) decreased over the trend period 2002–12. After endpoints were adjusted for antecedent flow, trends in tolerance to ammonia and low dissolved oxygen were no longer detectable at the critical threshold (FishAmm, pR = 0.536; FishDO, pR = 0.266). In this case, antecedent streamflow conditions covaried with trends in community tolerance to ammonia and low dissolved oxygen. Results in decreasing fish diversity were similar, but the influence of antecedent streamflow was not as clear. For example, after adjustment, the resulting p value was very near the threshold (FishDiv, pR = 0.108). Technically, this indicates the decreasing trend in fish diversity after adjustment was undetectable at the critical value. Unfortunately, the acceptance of the null hypothesis (no trend in the data) under these circumstances is an obvious artifact of the threshold value itself and reflective of the limitations of hypothesis testing.

Summary

One of the major goals of the National Water-Quality Assessment (NAWQA) Project is to determine how water quality changes over time. To support that goal, long-term consistent and comparable monitoring has been conducted on rivers and streams throughout the Nation. Other U.S. Geological Survey (USGS) projects, as well as projects conducted by many other Federal, State, Tribal, regional, and local agencies have collected long-term water-quality data to support their own assessments of changing water-quality conditions. For the first time, in an effort meant to be as inclusive as possible, data from multiple sources have been aggregated, screened, standardized, and used to support a comprehensive assessment of surface-water-quality trends in the United States. Collectively, these trend results will be used to provide insight into how natural variation and human activities have contributed to water-quality changes over time in the Nation's rivers and streams.

This report documents (1) the data compilation and processing steps used to identify stream and river sites throughout the Nation with data suitable for trend analysis, (2) the statistical methods used to determine trends, (3) considerations for water-quality data and streamflow data when modeling trends, (4) sensitivity analyses for selecting data and interpreting trend results with the Weighted Regressions on Time, Discharge, and Season method, and (5) a brief summary of the trend results presented as the presence or absence of a significant trend at each site for each parameter and trend period for which data are available. The scope of this study includes trends in water-quality concentrations and loads (nutrients, sediment, major ions, salinity, and carbon), pesticide concentrations and loads, and metrics for biological communities (fish, invertebrates, and algae) for four time periods: 1972–2012, 1982–2012, 1992–2012, and 2002–12. This report documents the methods used to determine trends in water quality and ecology in detail because they are vital to ensuring the quality and interpretability of the results. The final trend results are presented, along with an example of how to interpret the results from each trend model; an overall or in-depth interpretation of the trend results, such as causal analysis, is not included in this report.

Water-quality monitoring and aquatic ecology data (invertebrates) were compiled from multiple sources for this study and include ambient monitoring data that were readily accessible from Federal, State, Tribal, regional, and local government agencies and nongovernmental organizations. The primary sources of data were the USGS National Water Information System (NWIS) database; the USGS Aquatic Bioassessment (BioData) database, the U.S. Environmental Protection Agency STOrage and RETrieval Data Warehouse, Water Quality Exchange (STORET) database, and additional repositories for monitoring data that are not included in STORET. More than 185 million water-quality and ecological records from approximately 480,000 sites and over 600 agencies were initially evaluated for this trend study.

For many rivers, concentration and streamflow were correlated, and properly accounting for this relation during trend analyses allowed for a better determination of general changes in water quality over time. This was important for parameters that are strongly related to streamflow and sites where streamflow is highly variable; furthermore, if trends in load were of interest, streamflow data were essential to calculate parameter load (mass per time), which is the product of water-quality concentration (a mass per volume) and an associated streamflow rate (volume per time). Streamflow data were used in this analyses as an explanatory variable in each of three trend methods and to calculate water-quality loads. The USGS streamflow gages (henceforth, referred to as "streamgages") were the primary source of streamflow data.

Because streamflow data are required for water-quality, pesticide, and ecology trend analysis, sites with water-quality, pesticide, and (or) ecology data had to be assessed to determine whether there was a suitable nearby streamgage. Streamgage matching was performed that paired each water-quality site or group of sites with a streamgage. The goal was to identify the nearest streamgage having a drainage area within 10 percent of the site drainage area having a streamflow record that encompassed as much of the water-quality period of record as possible and that did not have intervening influences between the two locations that would cause streamflow to be substantially different. In many cases, a water-quality or pesticide site could not be matched to a streamgage, and the water-quality site was dropped from further consideration.

The initial pool of sites for this study included over half a million sites. Because of the large number of sites, it was not possible to tailor data-processing choices to each individual site; instead, data-processing choices were made at a high level and applied consistently to all sites. This approach avoided introducing differences in methodology that could invalidate comparisons of the final trend results across nationwide sites and parameters. A single set of decisions appropriate for all sites was not possible because of the wide variation in data characteristics and quality across sites. In the end, some sites that could have been salvaged by making individual site-based data processing and screening choices were excluded from the study.

The high-level decisions made in this study provided a robust approach to data processing and screening when the goal was to compare trends regionally and nationally at a large number of sites and for a large number of parameters. When the goal is to compare trends at a smaller subset of sites and for a smaller number of parameters, different choices could be possible—either because consistency can be achieved in a different manner or because it is reasonable in scope to tailor data-processing choices to each site. In these situations, trend results that were slightly different in magnitude and (or) uncertainty than those from this study could be obtained for the same site, parameter, and trend period, but any differences should be minor when the underlying data described the full range of conditions at a site and when similar and appropriate trend methods were used.

Similar water-quality parameters were compared to determine whether they could be combined for trend analysis. The goals of combining parameters were to (1) increase the number of sites for which trends can be calculated, and (2) increase the length of time over which trends can be calculated. Generally, the filtered fraction and the unfiltered or unknown fraction of the same parameter were evaluated for combination, for example filtered ammonia and unfiltered ammonia; however, some comparisons were made between different parameters, for example, filtered nitrite plus nitrate and filtered nitrate. When the analysis did not indicate substantial bias among parameter groups, the data were combined.

Different models were used to determine trends in water-quality, pesticide, and ecology parameters. The Weighted Regressions on Time, Discharge, and Season (WRTDS) method was used to determine trends in flow-normalized concentration and load for nutrients, sediment, major ions, salinity, and carbon. For comparison, all sites and parameters analyzed for trends using WRTDS were also analyzed for trends using the Seasonal Kendall method. The seasonal wave (SEAWAVE–Q) model was selected as the statistical tool for analyzing trends in pesticide concentrations for this study. The Kendall-tau test for trends was used to detect monotonic trends in the unadjusted ecological endpoints before and after adjustment for the effects of climatic variability.

Most studies that examine changes in water quality over time are challenged by common issues pertaining to data density, sample representativeness, and the stability of estimates as new data are added to the calibration record. A successful study of water-quality trends, one in which the identified trend holds up to scrutiny, requires researchers to make sound decisions to appropriately address and resolve the aforementioned issues and any others that arise because of the data or method used. Several sensitivity tests are presented to demonstrate how WRTDS estimates were affected by the frequency of storm-sample collection, the frequency of sample collection throughout the year, the precision (rounding) of water-quality data, and the incremental addition of new data to the calibration dataset.

Water-quality samples collected during high-flow periods were important for an unbiased WRTDS calibration but usually required extra effort because of the random nature of hydrologic events. The relation between concentration and streamflow is one of the primary relations modeled in WRTDS and is not always well defined when sites lack a sufficient number of high-flow samples. The criteria for the minimum number of high-flow samples were determined after evaluating the results from downsampled calibration data. It was decided that sites would be required to have at least 10 percent high-flow samples in all trend decades to be modeled, and at least 14 percent high-flow samples in half of the trend decades to be modeled.

The original guidance for WRTDS recommended its use at sites with long-term (at least 20 years) and dense (at least 200 samples) water-quality datasets and with continuous daily streamflow for the entire water-quality period of record. To maximize the use of the compiled multiagency data, this study required a sampling frequency of at least four samples per year for the first 2 years and the last 2 years of the trend period, as well as at least four samples per year in 70 percent of the years in the trend period.

The precision of the water-quality data used in this study varied by site, constituent, and analyzing agency because of the analytical methods used for sample determination, laboratory-specific detection limits, and reporting conventions. Additionally, the precision of a specific constituent, at a single site, analyzed by a single agency may have varied over time because of upgrades in analytical equipment, updates to detection limits, and changes in reporting conventions. To illustrate the influence of rounding on WRTDS estimates, sensitivity tests that applied various types of rounding were completed at two sites. Based on the results from these two sites, when rounding appears to minimally affect WRTDS estimates, no attempt was made to change the precision of the original data.

The WRTDS model and all regression-based water-quality models are sensitive to the incremental incorporation of new data. Estimates of concentration and load are typically updated periodically (for example, annually or semidecadally) as new water-quality data become available. Recalibration of regression-based models with updated data can lead to changes in the model coefficients, historical estimates, and model diagnostics. Because previous estimates of concentration and load might change with updated calibration data, trends estimated between two past points in time might also change. These changes were explored using successive calibrations, sets of model estimates (vectors of annual estimates of concentration and load) where new data were incrementally added to the end of the record, and models were recalibrated. The successive calibrations indicated changes occurring with the addition of new calibration data are most pronounced around an emerging inflection point.

For trend studies, it is important to identify major changes in laboratory methods or procedures that may induce a bias in water-quality determinations and possibly produce an artificial step trend in the data. Major changes in laboratory methods and the resulting analytical values at the National Water Quality Laboratory (NWQL) have been documented in a variety of resources and are compiled herein.

To evaluate the potential for laboratory changes to affect environmental concentrations, the presence of step trends in concentrations of selected water-quality data were evaluated separately for data from the USGS and data from all other agencies using 4-year moving windows beginning in 1974. For each 4-year moving window, a flow-related variability in concentrations was removed, and the resulting flow-normalized concentrations were analyzed for step trends across groups of sites in major geographic regions. The results for USGS data were compared to results for data from all other agencies indicated that, for the vast majority of parameters and years, regional trend patterns generally were consistent among sites with USGS data and sites with data from other agencies, indicating that there were few instances where anomalous step trends were induced by method changes at USGS laboratories.

One of the largest challenges in estimating trends was deciphering the proportion of variability in concentration, load, or ecological metrics that arose from variability in streamflow versus the proportion that was due to other changes and processes in the watershed or channel. Many water-quality parameters and ecological metrics were related to streamflow, and changes in streamflow make it challenging to evaluate progress or deterioration in response to human activities if the influence of streamflow was not considered. Strong gradual trends in streamflow caused a nuanced interpretation of the flow-normalized estimates which no longer describe the in-stream water-quality conditions but instead suggest changes to the input of the parameter to the stream. Abrupt changes in streamflow also were problematic for calibration and flow normalization in WRTDS. Sites with significant gradual trends or step trends in streamflow were assessed using the Mann-Kendall test and the Pettitte test. The WRTDS trend models were not excluded based solely on the results of these streamflow trends but were listed in an associated data release.

In total, nearly 12,000 water-quality, pesticide, and ecology trends were evaluated in this study; there were 11,729 combinations of sites, parameters, and trend periods. Trend results for each of the trend models are summarized here. Complete trend results are available in the associated data releases.

Acknowledgments

Compiling data from multiple sources for this report was a huge task and the authors would like to thank Denise Argue, Karen Beaulieu, Dave Bender, Anne Choquette, Jeff Deacon, Laura Medalie, Loren Metzger, Jennifer Savoie, Josh Valder, Christine Wieben, and John Wilson for all of their work with the data compilation. Considerable effort by multiple people was also put into all aspects of data processing and trend analysis. The authors would like to thank Denise Argue, Scott Ator, Daren Carlisle, Jeff Deacon, Neil Dubrovsky, Bob Gilliom, Bob Hirsch, Candice Hopkins, Casey Lee, Dave Lorenz, Amy Ludtke, Jeff Martin, Mike Meador, Emily Read, David Saad, Greg Schwartz, Sarah Spaulding, Ted Struzeski, Elizabeth Whiddon, Christine Wieben, Dan Wise, and Dave Wolock for their contributions to this study. In addition, the authors would like to thank the reviewers for this report and for all of the associated data releases: Laura Bexfield, Jeff Deacon, Leslie DeSimone, Amy McHugh, Laura Medalie, Matt Miller, Callie Oblinger, Kristin Romanok, and Kelly Smalling.

References

Anning, D.W., 2003, Assessment of selected inorganic constituents in streams in the central Arizona basins Study area, Arizona and Northern Mexico, through 1998: U.S. Geological Survey Water-Resources Investigations Report 03–4063, 116 p. [Also available at https://pubs.usgs.gov/wri/wri034063/.]

Anning, D.W., Bauch, N.J., Gerner, S.J., Flynn, M.E., Hamlin, S.N., Moore, S.J., Schaefer, D.H., Anderholm, S.K., and Spangler, L.E., 2007, revised 2010, Dissolved solids in basin-fill aquifers and streams in the Southwestern United States: U.S. Geological Survey Scientific Investigations Report 2006–5315, v. 1.1, 168 p. [Also available at https://pubs.usgs.gov/sir/2006/5315/.]

Bahls, L.L., 1993, Periphyton bioassessment methods for Montana streams: Helena, Montana, Water Quality Bureau, Department of Health and Environmental Services, 69 p.

Barbour, M.T., Gerritsen, J., Snyder, B.D., and Stribling, J.B., 1999, Rapid bioassessment protocols for use in wadeable streams and rivers—Periphyton, benthic macroinvertebrates, and fish - second edition: U.S. Environmental Protection Agency, Office of Water, Washington, D.C., EPA 841–B–99–002, 344 p.

Bexfield, L.M., 2008, Decadal-scale changes of pesticides in ground water of the United States, 1993–2003: Journal of Environmental Quality, v. 37, supplement 5, p. S226–S239.

Bray, J.R., and Curtis, J.T., 1957, An ordination of the upland forests communities of southern Wisconsin: Ecological Monographs, v. 27, p. 325–349.

Butler, J. N., 1982, Carbon dioxide equilibria and their applications: Meno Park, California, Addison-Wesley Publishing Company, Inc.

Carlisle, D.M., and Hawkins, C.P., 2008, Land use and the structure of Western U.S. stream invertebrate assemblages—Predictive models and ecological traits: Journal of the North American Benthological Society, v. 27, p. 986–999.

Carlisle, D.M., and Meador, M.R., 2007, A biological assessment of streams in the Eastern United States using a predictive model for macroinvertebrate assemblages: Journal of the American Water Resources Association, v. 43, p. 1194–1207.

Carlisle, D.M., Meador, M.R., Moulton, S.R., II, and Ruhl, P.M., 2007, Estimation and application of indicator values for common macroinvertebrate genera and families of the United States: Ecological Indicators, v. 7, p. 22–33.

Chanat, J.G., Moyer, D.L., Blomquist, J.D., Hyer, K.E., and Langland, M.J., 2016, Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay nontidal water-quality monitoring network, results through water year 2012: U.S. Geological Survey Scientific Investigations Report 2015–5133, 88 p. [Also available at http://pubs.er.usgs.gov/publication/sir20155133.]

Charles, D.F., Knowles, C., and Davis, R., 2002, Protocols for the analysis of algal samples collected as part of the U.S. Geological Survey National Water-Quality Assessment Program: Philadelphia, Penn., The Academy of Natural Sciences, Patrick Center for Environmental Research Report No. 02–06, 124 p.

Clark, J.M., Lane, S.N., Chapman, P.J., and Adamson, J. K., 2007, Export of dissolved organic carbon from an upland peatland during storm events—Implications for flux estimates: Journal of Hydrology, v. 347, no. 3–4, p. 438–447, doi:http://dx.doi.org/10.1016/j.jhydrol.2007.09.030.

Clarke, K.R., and Warwick, R.M., 2001, Change in marine communities—An approach to statistical analysis and interpretation (1st ed.): Plymouth, United Kingdom, Plymouth Marine Laboratory, 172 p.

Clarke, R.T., Wright, J.F., and Furse, M.T., 2003, RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers: Ecological Modeling, v.160, p. 219–233.

Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots: Journal of American Statisticial Association, v. 74, p. 829–836.

Cleveland, W.S., and Devlin, S.J., 1988, Locally-weighted regression—An approach to regression analysis by local fitting: Journal of American Statistical Association, v. 83, p. 596–610.

Cohn, T.A., Caulder, D.L., Gilroy, E.J., Zynjuk, L.D., and Summers, R.M., 1992, The validity of a simple statistical model for estimating fluvial constituent loads—An empirical-study involving nutrient loads entering Chesapeake Bay: Water Resources Research, v. 28, no. 9, p. 2353–2363.

Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., and Wells, D.K., 1989, Estimating constituent loads: Water Resources Research, v. 25, no. 5, p. 937–942.

Cohn, T.A., and Lins, H.F., 2005, Nature's style—Naturally trendy: Geophysical Research Letters, v. 32, no. L23402, doi:10.1029/2005GL024476.

Corsi, S.R., De Cicco, L.A., Lutz, M.A., and Hirsch, R.M., 2015, River chloride trends in snow-affected urban watersheds—Increasing concentrations outpace urban growth rate and are common among all seasons: Science of the Total Environment, v. 508, p. 488–497, doi: 10.1016/j.scitotenv.2014.12.012.

Cuffney, T.F., Gurtz, M.E., and Meador, M.R., 1993, Methods for collecting benthic invertebrate samples as part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 93–406, 66 p.

Dalzell, B.J., Filley, T.R., and Harbor, J.M., 2007, The role of hydrology in annual organic carbon loads and terrestrial organic matter export from a midwestern agricultural watershed: Geochimica et Cosmochimica Acta, v. 71, issue 6, p. 1448–1462, doi:10.1016/j.gca.2006.12.009.

De Cicco, L.A., Sprague, L.A., Murphy, J.C., Riskin, M.L., Falcone, J.A., Stets, E.G., Oelsner, G.P., and Johnson, H.M., 2017, Water-quality and streamflow datasets used in the Weighted Regressions on Time, Discharge, and Season (WRTDS) models to determine trends in the Nation's rivers and streams, 1972–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7KW5D4H.

Drever, J.I., 1997, The geochemistry of natural waters (3d ed.): Upper Saddle River, N.J., Prentice-Hall.

Dubrovsky, N.M., Burow, K.R., Clark, G.M., Gronberg, J.M., Hamilton P.A., Hitt, K.J., Mueller, D.K., Munn, M.D., Nolan, B.T., Puckett, L.J., Rupert, M.G., Short, T.M., Spahr, N.E., Sprague, L.A., and Wilber, W.G., 2010, The quality of our Nation's waters—Nutrients in the Nation's streams and groundwater, 1992–2004: U.S. Geological Survey Circular 1350, 174 p.

Esterby, S.R., 1996, Review of methods for the detection and estimation of trends with emphasis on water quality applications: Hydrological Processes, v. 10, p. 127–149.

Falcone, J.A., Baker, N.T., and Price, C.V., 2017, Watershed boundaries for study sites of the U.S. Geological Survey Surface Water Trends project: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F78S4N29.

Farmer, W.H., Murphy, J.C., Riskin, M.L., Ryberg, K.R., and Zuellig, R.E., 2017, Daily streamflow datasets used to analyze trends in streamflow at sites also analyzed for trends in water quality and ecological condition in the Nation's rivers and streams: U.S. Geological Survey data release, http://www.dx.doi.org/10.5066/F7D798JN.

Fitzpatrick, F.A., Waite, I.R., D'Arconte, P.J., Meador, M.R., Maupin, M.A., and Gurtz, M.E., 1998, Revised methods for characterizing stream habitat in the National Water-Quality Assessment Program: U.S. Geological Survey Water-Resources Investigations Report 98–4052, 67 p.

Garcia, M.H., 2008, Sedimentation engineering—Processes, measurements, modeling, and practice: Reston, Virginia, American Society of Civil Engineers, v. 110 of ASCE Manuals and Reports on Engineering American Society of Civil Engineers Publishing. p. 1132. ISBN: 9870784408148

Gilliom, R.J., Alley, W.M., and Gurtz, M.E., 1995, Design of the National Water-Quality Assessment Program—Occurrence and distribution of water-quality conditions: U.S. Geological Survey Circular 1112, 33 p.

Gilliom and others, 2006, The quality of our Nation's waters—Pesticides in the Nation's streams and ground water, 1992–2001: U.S. Geological Survey Circular 1291,172 p.

Granato, G.E., 2009, Computer programs for obtaining and analyzing daily mean streamflow data from the U.S. Geological Survey National Water Information System web site: U.S. Geological Survey Open-File Report 2008–1362, 123 p. on CD–ROM, appendix 3 of 5.

Gray, J.R., Glysson, G.D., Turcios, L.M., and Schwarz, G.E., 2000, Comparability of suspended-sediment concentration and total suspended solids data: U.S. Geological Survey Water-Resources Investigations Report 00–4191, 14 p.

Gurtz, M.E., Van Sickle, J., Carlisle, D.M., and Paulsen, S.G., 2013, Design tradeoffs for trend assessment in aquatic biological monitoring programs: U.S. Geological Survey Open-File Report 2013–1222, 15 p., 1 appendix.

Hanley, K.W., Wollheim, W.M., Salisbury, J., Huntington, T., and Aiken, G., 2013, Controls on dissolved organic carbon quantity and chemical character in temperate rivers of North America: Global Biogeochemical Cycles, v. 27, no. 2, p. 492–504, doi:10.1002/gbc.20044.

Harrell, F.E., Jr., *with* contributions from Dupont, C., and many others, 2016, Hmisc: Harrell Miscellaneous, R package, version 3.17–4. https://cran.r-project.org/web/packages/Hmisc/index.html.

Hawkins, C.P., 2006, Taxonomic completeness as a measure of biological integrity: Ecological Applications, v. 16, p. 1277–1294.

Hawkins, C.P., and Carlisle, D.M., 2001, Use of predictive models for assessing the biological integrity of wetlands and other aquatic habitats, *in* Rader, R., Batzer, D., and Wissinger, S., eds., Bioassessment and management of North American freshwater wetlands: New York, John Wiley & Sons, p. 59–83.

Helsel, D.R., 2005, Nondetects and data analysis—Statistics for censored environmental data: Hoboken, N.J., John Wiley & Sons, 250 p.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 510 p. [Also available at http://pubs.usgs.gov/twri/twri4a3/.]

Hem, J.D., 1985, Study and interpretation of the chemical characteristics of natural water: U.S. Geological Survey Water Supply Paper 2254, 263 p.

Hill, B.H., Herlihy, A.T., Kaufmann, P.R., Stevenson, R.J., McCormick, F.H., and Johnson, C.B., 2000, Use of periphyton assemblage data as an index of biotic integrity: Journal of the North American Benthological Society, v. 19, p. 50–67.

Hirsch, R.M., 1982, A comparison of four streamflow record extension techniques: Water Resources Research, v. 18, no. 4, p. 1081–1088.

Hirsch, R.M, and Archfield, S.A., 2015, Flood trends—Not higher but more often: Nature Climate Change, v. 5, p. 198–199, doi:10.1038/nclimate2551.

Hirsch, R.M., Archfield, S.A., and De Cicco, L.A., 2015, A bootstrap method for estimating uncertainty of water quality trends: Environmental Modelling and Software, v. 73, p. 148–166.

Hirsch, R.M., and De Cicco, L.A., 2015, User guide to exploration and graphics for RivEr Trends (EGRET) and dataRetrieval—R packages for hydrologic data (version 2.0, February 2015): U.S. Geological Survey Techniques and Methods, book 4, chap. A10, 93 p., http://dx.doi.org/10.3133/tm4A10.

Hirsch, R.M., Moyer, D.L., and Archfield, S.A., 2010, Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay River inputs: JAWRA Journal of the American Water Resources Association, v. 46, p. 857–880. doi: 10.1111/j.1752-1688.2010.00482.x.

Hirsch, R.M., and Slack, J.R., 1984, A nonparametric trend test for seasonal data with serial dependence: Water Resources Research, v. 20, no. 6, p. 727–732, doi:10.1029/WR020i006p00727.

Hodgkins, G.A., and Dudley, R.W., 2011, Historical summer base flow and stormflow trends for New England rivers: Water Resources Research, v. 47, no. W07528, doi:10.1029/2010WR009109.

Hooper, R.P., Aulenbach, B.T., and Kelly, V.J., 2001, The national stream quality accounting network—A flux-based approach to monitoring the water quality of large rivers: Hydrological Processes, v. 15, no. 7, p. 1089–1106.

Jarvie, H.P., Withers, P.J.A., and Neal, Colin, 2002, Review of robust measurement of phosphorus in river water—Sampling, storage, fractionation, and sensitivity: Hydrology and Earth System Sciences, v. 6, no. 1, p. 113–132. DOI:10.5194/hess-6-113-2002.

Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., and Schlosser, I.J., 1986, Assessing biological integrity in running waters—A method and its rationale: Illinois Natural History Survey Special Publication 5, p. 28.

Kaushal, S.S., Groffman, P.M., Likens, G.E., Belt, K.T., Stack, W.P., Kelly, V.R., Band, L.E. and Fisher, G.T., 2005, Increased salinization of fresh water in the Northeastern United States: Proceedings of the National Academy of Sciences of the United States of America, v. 102, no. 38, p. 13517−13520.

Keith, L.H., 1991, Environmental sampling and analysis—A practical guide: Chelsea, Mich., Lewis Publishers, 143 p.

Kendall, M.G., 1975, Rank correlation methods (4th ed.): London, Charles Griffin, 202 p.

Kent, R., and Belitz, K., 2004, Concentrations of dissolved solids and nutrients in water sources and selected streams of the Santa Ana Basin, California, October 1998-September 2001: U.S. Geological Survey Water-Resources Investigations Report 03–4326, 61 p. [Also available at http://pubs.usgs.gov/wri/wrir034326/.]

Kornberg A., Rao, N.N., and Ault-Riché, D., 1999, Inorganic polyphosphate—A molecule of many functions: Annual Review of Biochemistry, v. 68, p. 89–125, 10.1146/annurev.biochem.68.1.89.

Lange-Bertalot, H., 1979, Pollution tolerance of diatoms as a criterion for water quality estimation: Nova Hedwigia, v. 64, p. 285–304.

Langmuir, D., 1997, Aqueous environmental geochemistry: Upper Saddle River, N.J., Prentice-Hall.

Lindley, C.E., Stewart, J.T., and Sandstrom, M.W., 1996, Determination of low concentrations of acetochlor in water by automated solid-phase extraction and gas chromatography with mass-selective detection: Journal of AOAC International, v. 79, no. 4, p. 962–966.

Lorenz, D.L., 2016, smwrStats: An R package for the analysis of hydrologic data, version 0.7.5, https://github.com/USGS-R/smwrStats/tree/v0.7.5

Lowe, R.L., 1974, Environmental requirements and pollution tolerance of freshwater diatoms: Cincinnati, Ohio, U.S. Environmental Protection Agency, Office of Research and Development, National Environmental Research Center, EPA–670/4–74–005, 334 p.

MacCoy, D.M., 2011, BioData—A national aquatic bioassessment database: U.S. Geological Survey Fact Sheet 2011–3112, 4 p.

Madsen, J.E., Sandstrom, M.W., and Zaugg, S.D., 2003, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—A method supplement for the determination of fipronil and degradates in water by gas chromatography/mass spectrometry: U.S. Geological Survey Open-File Report 02–462, 11 p. [Also available at http://nwql.usgs.gov/Public/pubs/OFR02-462/OFR02-462.html.]

Maluk, T.L., 2000, Spatial and seasonal variability of nutrients, pesticides, bacteria, and suspended sediment in the Santee River Basin and coastal drainages, North and South Carolina, 1995–97: U.S. Geological Survey Water-Resources Investigations Report 00–4076, 46 p. [Also available at https://pubs.er.usgs.gov/publication/wri004076.]

Mann, H.B., 1945, Nonparametric tests against trend: Econometrica, v. 13, p. 245–249.

Martin, J.D., 2009, Sources and preparation of data for assessing trends in concentrations of pesticides in streams of the United States, 1992–2006: U.S. Geological Survey Scientific Investigations Report 2009–5062, 41 p. [Also available at http://pubs.usgs.gov/sir/2009/5062.]

Martin, J.D., and Eberle, Michael, 2011, Adjustment of pesticide concentrations for temporal changes in analytical recovery, 1992–2010: U.S. Geological Survey Data Series 630, 11 p., 5 app. [Also available at http://pubs.usgs.gov/ds/630/.]

Martin, J.D., Eberle, Michael, and Nakagaki, Naomi, 2011, Sources and preparation of data for assessing trends in concentrations of pesticides in streams of the United States, 1992–2010: U.S. Geological Survey Data Series 655, 23 p., 5 appendixes. [Also available at http://pubs.usgs.gov/ds/655/.]

Martin, J.D., Stone, W.W., Wydoski, D.S., and Sandstrom, M.W., 2009, Adjustment of pesticide concentrations for temporal changes in analytical recovery, 1992–2006: U.S. Geological Survey Scientific Investigations Report 2009–5189, 23 p., plus appendixes. [Also available at http://pubs.usgs.gov/sir/2009/5189/.]

McKay, L., Bondelid, T., Dewald, T., Johnston, C., Moore, R., and Rea, A., 2012, NHDPlus version 2—User guide, 2012: U.S. Environmental Protection Agency. [Also available at http://www.horizon-systems.com/NHDPlus/index.php.]

Meador, M.R., and Carlisle, D.M., 2007, Quantifying tolerance indicator values for common steam fish species of the United States: Ecological Indicators, v. 7, p. 329–338.

Meador, M.R., and Carlisle, D.M., 2009, Predictive models for fish assemblages in eastern USA streams—Implications for assessing biodiversity: Transactions of the American Fisheries Society, v. 138, p. 725–740.

Meador, M.R., Cuffney, T.F., and Gurtz, M.E., 1993, Methods for sampling fish communities as a part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 93–104, 40 p.

Meador, M.R., Whittier, T.R., Goldstein, R.M., Hughes, R.M., and Peck, D.V., 2008, Assessing relations between anthropogenic disturbance and biological integrity for western USA streams and rivers: Transactions of the American Fisheries Society, v. 137, p. 13–22.

Meybeck, M., Laroche, L., Dürr, H.H., and Syvitski, J.P.M., 2003, Global variability of daily total suspended solids and their fluxes in rivers: Global and Planetary Change, v. 39, p. 65–93.

Mills, T.J., Sprague, L.A., Murphy, J.C., Riskin, M.L., Falcone, J.A., Stets, E.G., Oelsner, G.P, and Johnson, H.M., 2017, Water-quality and streamflow datasets used in Seasonal Kendall trend tests for the Nation’s rivers and streams, 1972–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7QN64VT.

Minnesota Pollution Control Agency, 2012, Vermillion River watershed monitoring and assessment report: Minnesota Pollution Control Agency document number wq-ws3-07040001c, p. 74.

Moulton, S.R., Carter, J.L., Grotheer, S.A., Cuffney, T.F., and Short, T.M., 2000, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Processing, taxonomy, and quality control of benthic macroinvertebrate samples: U.S. Geological Survey Open-File Report 00–212, 49 p.

Moulton, S.R., Kennen, J.G., Goldstein, R.M., and Hambrook, J.A., 2002, Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 02–150, 75 p.

Moss, D., Furse, M.T., Wright, J.F., and Armitage, P.D., 1987, The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data: Freshwater Biology, v. 17, p. 41–52.

Mullaney, J.R., 2016, Nutrient, organic carbon, and chloride concentrations and loads in selected Long Island Sound tributaries—Four decades of change following the passage of the Federal Clean Water Act: U.S. Geological Survey Scientific Investigations Report 2015–5189, 47 p., http://dx.doi.org/10.3133/sir20155189.

Mueller, D.K., and Titus, C.J., 2005, Quality of nutrientdata from streams and ground water sampled during water years 1992–2001: U.S. Geological Survey Scientific Investigations Report 2005–5106, 27 p.

Neter, J., Kutner, M.H., Nachtsheim, C.J., and Wasserman, William, 1996, Applied linear statistical models (4th ed.): Boston, Mass., WCB/McGraw-Hill, 1408 p.

Oblinger Childress, C.J., 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. [Also available at http://water.usgs.gov/owq/OFR_99-193/index.html.]

Patton, C.J., and Truitt, E.P., 2000, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Determination of ammonium plus organic nitrogen by a Kjeldahl digestion method and an automated photometric finish that includes digest cleanup by gas diffusion: U.S. Geological Survey Open-File Report 00–170, 31 p.

Pettitt, A.N., 1979, A non-parametric approach to the change point problem: Journal of the Royal Statistical Society Series C, Applied Statistics 28, p. 126–135.

Plummer, L.N., and Busenberg, E., 1982, The solubilities of calcite, aragonite and vaterite in CO_2-H_2O solutions between 0 and 90 °C, and an evaluation of the aqueous model for the system $CaCO_3$-CO_2-H_2O: Geochimica et Cosmochimica Acta, v. 46, p. 1011–1040.

Potapova, M., and Charles, D.F., 2003, Distribution of benthic diatoms in U.S. rivers in relation to conductivity and ionic composition: Freshwater Biology, v. 48, p. 1311–1328.

Potapova, M., and Charles, D.F., 2007, Diatom metrics for monitoring eutrophication in rivers of the United States: Ecological Indicators, v. 7, p. 48–70.

Porter, S.D., 2008, Algal attributes—An autecological classification of algal taxa collected by the National Water-Quality Assessment Program: U.S. Geological Survey Data Series 329, http://pubs.usgs.gov/ds/ds329/.

Porter, S.D., Cuffney, T.F., Gurtz, M.E., and Meador, M.R., 1993, Methods for collecting algal samples as part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 93–409, 39 p.

Preston, S.D., Bierman, V.J., and Silliman, S.E., 1989, An evaluation of methods for the estimation of tributary mass loads: Water Resources Research, v. 25, no. 6, p. 1379–1389.

R Development Core Team, 2014, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, accessed August 22, 2014, at http://www.R-project.org.

Raymond, P. A., and Oh, Neung-Hwan, 2007, An empirical study of climatic controls on riverine C export from three major U.S. watersheds: Global Biogeochemical Cycles, v. 21,GB2002, doi:10.1029/2006GB002783.

Ries, K.G., III, 1994, Estimation of low-flow duration discharges in Massachusetts: U.S. Geological Survey Water-Supply Paper 2418, 50 p.

Rinella, F.A., and Janet, M.L., 1998, Seasonal and spatial variability of nutrients and pesticides in streams of the Willamette Basin, Oregon, 1993–95: U.S. Geological Survey Water-Resources Investigations Report 97–4082–C, 59 p. [Also available at http://pubs.er.usgs.gov/usgspubs/wri/wri974082C.]

Ryberg, K.R., Murphy, J.C., Falcone, J.A., Riskin, M.L., Wieben, C.M., and Vecchia, A.V., 2017, Pesticide concentration and streamflow datasets used to evaluate pesticide trends in the Nation's rivers and streams, 1992–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7BC3WPC.

Ryberg, K.R., and Vecchia, A.V., 2012, waterData—An R package for retrieval, analysis, and anomaly calculation of daily hydrologic time series data, version 1.0: U.S. Geological Survey Open-File Report 2012–1168, 8 p. [Also available at http://pubs.usgs.gov/of/2012/1168/.]

Ryberg, K.R., and Vecchia, A.V., 2013, seawaveQ—An R package providing a model and utilities for analyzing trends in chemical concentrations in streams with a seasonal wave (seawave) and adjustment for streamflow (Q) and other ancillary variables: U.S. Geological Survey Open-File Report 2013–1255, 13 p., accessed August 21, 2016, at http://pubs.usgs.gov/of/2013/1255/.

Ryberg, K.R., Vecchia, A.V., Gilliom, R.J., and Martin, J.D., 2014, Pesticide trends in major rivers of the United States, 1992–2010: U.S. Geological Survey Scientific Investigations Report 2014–5135, 63 p., accessed August 21, 2016, at http://dx.doi.org/10.3133/sir20145135.

Ryberg, K.R., Vecchia, A.V., Martin, J.D., and Gilliom, R.J., 2010, Trends in pesticide concentrations in urban streams in the United States, 1992–2008: U.S. Geological Survey Scientific Investigations Report 2010–5139, 101 p., accessed August 21, 2016, at http://pubs.usgs.gov/sir/2010/5139/.

Shannon, C.E., and Weaver, W., 1963, The mathematical theory of communication: Urbana, Ill., University of Illinois Press, p. 144.

Spaulding, S.A., Lubinski, D.J., and Potapova, M., 2010, Diatoms of the United States: University of Colorado Boulder Web page, accessed on October 10, 2014, at http://westerndiatoms.colorado.edu.

Sprague, L.A., Oelsner, G.P., and Argue, D.M., 2017, Challenges with secondary use of multi-source water-quality data in the United States: Water Research, v. 110, p. 252–261, https://dx.doi.org/10.1016/j.watres.2016.12.024.

Stets, E.G., and Striegl, R.G., 2012, Carbon export by rivers draining the conterminous United States: Inland Waters, v. 2, no. 4, p. 177–184, doi:10.5268/IW-2.4.510.

Sullivan, D.J., Vecchia, A.V., Lorenz, D.L., Gilliom, R.J., and Martin, J.D., 2009, Trends in pesticide concentrations in corn-belt streams, 1996–2006: U.S. Geological Survey Scientific Investigations Report 2009–5132, 75 p. [Also available at http://pubs.usgs.gov/sir/2009/5132/.]

Tarapchak, S.J., Bigelow, S.M., and Rubitschum, C., 1982, Overestimation of orthophosphorus concentrations in surface waters of southern Lake Michigan—Effects of acid and ammonium molybdate: Canadian Journal of Fisheries and Aquatic Sciences, v. 39, p. 296–304.

Tillman, F.D, and Anning, D.W., 2014, Updated estimates of long-term average dissolved-solids loading in streams and rivers of the Upper Colorado River Basin: U.S. Geological Survey Open-File Report 2014–1148, 11 p., https://dx.doi.org/10.3133/ofr20141148.

Therneau, T.M., 2014, survival—A package for survival analysis in S: R package version 2.37–7, accessed August 22, 2014, at http://CRAN.R-project.org/package=survival.

U.S. Army Corps of Engineers, 2015, National inventory of dams: U.S. Army Corps of Engineers web page, accessed May 6, 2015, at http://geo.usace.army.mil/pgis/f?p=397:1:0.

U.S. Department of Agriculture, 2014, U.S. and State-level farm income and wealth statistics: U.S. Department of Agriculture Economic Research Service, accessed July 27, 2016, at http://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics.aspx.

U.S. Environmental Protection Agency, 2006, Decision documents for atrazine: U.S. Environmental Protection Agency [variously paged], accessed August 21, 2016, at https://www3.epa.gov/pesticides/chem_search/reg_actions/reregistration/red_PC-080803_1-Apr-06.pdf.

U.S. Environmental Protection Agency, 2010, Clean watersheds needs survey 2008—Report to Congress: U.S. Environmental Protection Agency Office of Wastewater Management Report EPA–832–R–10–002, 154 p.

U.S. Environmental Protection Agency, 2015, SRS names for carbon, nitrogen, and phosphorous characteristics, accessed February 24, 2015, at http://www.epa.gov/storet/wqx/wqx_downloads.html.

U.S. Environmental Protection Agency, 2016, National rivers and streams assessment 2008–2009—A collaborative survey: U.S. Environmental Protection Agency Office of Water and Office of Research and Development Report EPA/841/R–16/007, 131 p.

U.S. Geological Survey, 2013, USGS water data for the Nation: National Water Information System: Web Interface, accessed May 2013, at http://waterdata.usgs.gov/nwis/ or http://dx.doi.org/10.5066/F7P55KJN.

van Dam, H., Mertens, A., and Sinkeldam, J., 1994, A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands: Netherlands Journal of Aquatic Ecology, v. 28, no. 1, p. 117–133.

Vecchia, A.V., Martin, J.D., and Gilliom, R.J., 2008, Modeling variability and trends in pesticide concentrations in streams: Journal of the American Water Resources Association, v. 44, no. 5, p. 1308–1324.

Vinson, M.R., and Hawkins, C.P., 1996, Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams: Journal of the North American Benthological Society, v. 15, p. 392–399.

Walling, D.E., 2005, Tracing suspended sediment sources in catchments and river systems. Science of the Total Environment, v. 344, p. 159–184. [Also available at http://dx.doi.org/10.1016/j.scitotenv.2005.02.011.]

Wang, T.Q., and McCafferty, W.P., 2004, Heptageniidae (Ephemeroptera) of the world, part I—Phylogenetic higher classification: Transactions of the American Entomological Society, v. 130, p. 11–45.

Westland, A.D., and Boisclair, I., 1974, The analytical separation of phosphate from natural water by ion exchange: Water Research, v. 8, p. 467–470.

Whittier, T.R., Hughes, R.M., Stoddard, J.L., Lomnicky, G.A., Peck, D.V., and Herlihy, A.T., 2007, A structured approach for developing indices of biotic integrity—Three examples from stream and rivers in the Western USA: Transactions of the American Fisheries Society, v. 136, p. 718–735.

Williams, G.P., 1989, Sediment concentrations versus matter discharge during hydrologic events in rivers: Journal of Hydrology, v. 111, p. 89–106.

Woodworth, M.T., and Connor, B.F., 2003, Results of the U.S. Geological Survey's analytical evaluation program for standard reference samples distributed in March 2003: U.S. Geological Survey Open-File Report 03–261, 109 p.

Young, M.K., McKelvey, K.S., Pilgrim, K.L., and Schwartz, M.K., 2013, DNA barcoding at riverscape scales—Assessing biodiversity among fishes of the genus Cottus (Teleostei) in northern Rocky Mountain streams: Molecular Ecology Resources, v. 13, p. 583–595.

Yuan, L.L., Hawkins, C.P., and Van Sickle, J., 2008, Effects of regionalization decisions on an O/E index for the US national assessment: Journal of the North American Benthological Society v. 27, p. 892–905.

Zaugg, S.D., Sandstrom, M.W., Smith, S.G., and Fehlberg, K.M., 1995, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Determination of pesticides in water by C–18 solid-phase extraction and capillary-column gas chromatography/mass spectrometry with selected-ion monitoring: U.S. Geological Survey Open-File Report 95–181, 49 p.

Zuellig, R.E., and Riskin, M.L., 2017, Ecological community datasets used to evaluate the presence of trends in ecological communities in selected rivers and streams across the United States, 1992–2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7G44ND3.

Publishing support provided by
Lafayette Publishing Service Center

Oelsner and others—**Water-Quality Trends in the Nation's Rivers and Streams, 1972–2012—Data Preparation, Statistical Methods, and Trend Results**—SIR 2017–5006

ISSN 2328-0328 (online)
https://doi.org/10.3133/sir20175006